

Quantitative Stability of Two-stage Stochastic Linear Programs with Full Random Recourse

Jie Jiang^a and Zhiping Chen^a

^a School of Mathematics and Statistics, Xi'an Jiaotong University, Shaanxi, China

ARTICLE HISTORY

Compiled February 9, 2018

ABSTRACT

In this paper, we use the parametric programming technique and pseudo metrics to study the quantitative stability of the two-stage stochastic linear programming problem with full random recourse. Under the simultaneous perturbation of the cost vector, coefficient matrix and right-hand side vector, we first establish the locally Lipschitz continuity of the optimal value function and the boundedness of optimal solutions of parametric linear programs. On the basis of these results, we deduce the locally Lipschitz continuity and the upper bound estimation of the objective function of the two-stage stochastic linear programming problem with full random recourse. Then under moderate assumptions and different pseudo metrics, we obtain the quantitative stability results of two-stage stochastic linear programs with full random recourse, which improve the current results in terms of tractability and the necessary assumptions. Finally, we consider the empirical approximation of the two-stage stochastic programming model and derive the rate of convergence of the corresponding SAA method.

KEYWORDS

stochastic programming; full random recourse; parametric programming; pseudo metrics; quantitative stability; empirical approximation

1. Introduction

Stochastic programming with recourse is a well-known mathematical tool for solution of multi-period decision-making problems under uncertainty. This approach assumes that each uncertain parameter follows a given probability distribution and decisions are taken over time based on the revealed information. It has been widely applied in practical problems such as portfolio selection [1,2], traffic and transportation management [3], equilibrium and complementarity problems [4,5] and similar fields. In view of this, different issues about stochastic programmes with recourse have been extensively studied in recent years, which leads to plenty of literature in this field. One refer to excellent monographs [6,7] and references therein.

Theoretical solution of stochastic programs can be found explicitly only in very exceptional cases due to the complexity of the functional form of the stochastic models. Therefore, the necessity of numerical solution arises. By now, many approximation schemes have been proposed, for example the sample average approximation (SAA)

method. Then an important issue is: when the original continuous distribution is approximated a series of discrete probability distributions, will the optimal value and the optimal solutions of the discretization problem converge to those of the original stochastic programming problem? This kind of analysis is the so-called stability analysis. In this aspect, there are a number of researches, see for example [8–12] and the references therein. According to the obtained conclusions, stability analyses can be divided into two groups: the qualitative stability analysis and the quantitative stability analysis. The former mainly examines the convergence or continuity properties. The latter aims at deriving the quantitative error bound estimation. Of particular interest of this paper, we concentrate on the quantitative stability analysis of general two-stage stochastic linear programs with full random recourse.

The general form of the two-stage stochastic linear programming problem with full random recourse can be formulated as

$$\min_{x \in X} \mathbb{E}_P[f(x, \xi)], \quad (1)$$

where $\mathbb{E}_P[\cdot]$ denotes the expectation with respect to the probability measure P and

$$f(x, \xi) := \langle c, x \rangle + \Phi(x, \xi), \quad (2)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product of vectors in the finite dimensional Hilbert space and $\Phi(x, \xi)$ is the optimal value of the following full random linear recourse problem:

$$\inf \{ \langle q(\xi), y \rangle : W(\xi)y + T(\xi)x = h(\xi), y \in \mathbb{R}_+^m \}. \quad (3)$$

Here, $\xi : \Omega \rightarrow \mathbb{R}^s$ is a random vector defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $\Xi \subset \mathbb{R}^s$ is the support set of ξ , that is, the smallest closed subset in \mathbb{R}^s such that $P(\Xi) = 1$ where $P = \mathbb{P} \circ \xi^{-1}$; $c \in \mathbb{R}^n$; $X \subset \mathbb{R}^n$ is the feasible solution set for the first stage here-and-now decision vector $x \in \mathbb{R}^n$; $y \in \mathbb{R}^m$ is the second stage wait-and-see solution, i.e., we determine a $y(\xi)$ when we have the realization value of ξ . $W(\cdot) : \Xi \rightarrow \mathbb{R}^{r \times m}$, $T(\cdot) : \Xi \rightarrow \mathbb{R}^{r \times n}$, $h(\cdot) : \Xi \rightarrow \mathbb{R}^r$ and $q(\cdot) : \Xi \rightarrow \mathbb{R}^m$ are all matrix or vector valued mappings. They are called the recourse matrix, the technology matrix, the right-hand side vector and the cost vector, respectively. We say that the two-stage stochastic programming problem (1)-(3) has relatively complete recourse if problem (3) is solvable whenever $(x, \xi) \in X \times \Xi$, while the complete recourse implies that problem (3) is solvable whenever $(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^s$. We adopt the term ‘full random’ because all the coefficients in the second stage problem (3) are random.

Early works on stability analysis mainly focus on the continuity property of the optimal value function and the convergence of the optimal solution set with respect to some probability metrics. Kall considered in [13] the two-stage stochastic programming with complete fixed recourse, and derived the lower and upper bounds of $\mathbb{E}_P[\Phi(x, \xi)]$ on the condition that the support set Ξ is a hypercube. Under the fixed recourse assumption, Robinson and Wets established in [14] the continuity of the optimal value function and the upper semi-continuity of the optimal solution set with respect to probability metrics. In [15], Römisch and Schultz derived quantitative stability results for the two-stage stochastic program by viewing it as a parametric programming with respect to probability measures. It is worthy pointing out that they simply view the technology matrix in the second stage problem as a random parameter. The same authors employed the Wasserstein metric in [16] to obtain the quantitative stability,

and, under the complete fixed recourse assumption, established in [17] the Lipschitz continuity of the optimal solution set with respect to the Kolmogorov-Smirnov metric. Dentcheva and Römisch investigated in [18] the differentiable properties of the optimal value function and solution set mapping when only the right-hand side vector is random. Rachev and Römisch in [8] considered the general form of stochastic programming and studied the quantitative stability of the optimal value and optimal solution set by utilizing minimal information probability metrics. Based on these results, we can deduce the corresponding stability results for specific models such as the two-stage stochastic linear programming with fixed recourse. On the basis of the results in [8] and other literature, Römisch provided an excellent review in [9] about the stability analysis of stochastic programs. Recently, Liu, Römisch and Xu studied the quantitative stability of the stochastic generalized equation in [19], whose results can be applied to two-stage stochastic linear programs with fixed recourse.

All the above researches only consider the randomness of some coefficients in the second stage recourse problem (3). To the best of our knowledge, there are only two papers that investigate the quantitative stability of two-stage stochastic linear programs with full random recourse. The first attempt was conducted by Römisch and Wets in [10]. They established the Lipschitz continuity assertions of the optimal value and ϵ -approximation solution sets with respect to the adopted pseudo metric, but only obtained the quantitative stability results for the ϵ -approximation solution set. Meanwhile, Römisch and Wets mainly utilized generic perturbation results for optimization problems in [20] and employed an abstract probability metric, the so-called minimal information metric. Thus it is difficult to apply their method and results due to the abstract and complexity. To derive tractable conclusions with simple techniques, Han and Chen in [21] did another try. They first established the locally Lipschitz continuity of the feasible solution set of parametric linear programs by using Hoffman constant, which was used to deduce the locally Lipschitz continuity of and an upper bound on the objective function of two-stage stochastic linear programs. On the basis of these results, they obtained the quantitative stability assertions of the optimal value and the optimal solution set of problem (1)-(3) under the Fortet-Mourier probability metric. Although Han and Chen give a relatively simple demonstration and employ a specific probability metric to derive the upper bound estimation, they require that the feasible solution set of the second stage recourse problem (3) is bounded and satisfies some regularity conditions, which are rather restrictive and difficult to verify.

The limitations in [10] and [21] inspire us to establish general quantitative stability conclusions under weak and natural assumptions by using techniques easy to follow. To this end, we conduct the perturbation analysis of parametric linear programs with respect to all the coefficients based on the well-known Farkas' lemma. The boundedness of optimal solutions and the locally Lipschitz continuity of the optimal value function are derived. With these quantitative perturbation results for parametric linear programs, we obtain different quantitative stability results for the two-stage stochastic linear program with full random recourse under different conditions. Our results avoid the shortcomings in [10,21] and other relevant literature.

On the other hand, the empirical approximation has been an important tool for the solution of stochastic optimizations. There are a number of studies in this field. Rachev and Römisch used the quantitative stability results to carry out the convergence analysis in [8] and derived the rate of convergence. This kind of methods is discussed in detail in the comprehensive review [9]. Zhao and Guan investigated in [22] the discrete approximation of several probability metrics for risk-averse two-stage stochastic programmes. Shapiro and Xu studied the exponential convergence by using the Cramér's

large deviation theorem in [23]. Based on this result, Liu and Xu in [12] estimated the rate of convergence of the SAA method for the stochastic optimization with second order dominance constraints. As an application of our quantitative stability results, we consider the SAA counterpart of problem (1)-(3) by employing a similar method as that in [8]. We will establish the rate of convergence of the empirical approximation problem to the original continuous problem as the sample size goes to infinity, which extends the discrete approximation conclusion in [8] to the full random recourse case.

Throughout this paper, we will adopt the following notations. $\mathbb{E}_P[\cdot]$ stands for the expectation operator with respect to probability measure P . $a \vee b := \max\{a, b\}$, $a \wedge b := \min\{a, b\}$ for any $a, b \in \mathbb{R}$. \mathbb{B} denotes the closed unit ball centered at zero. $\|\cdot\|$ denotes the Euclidean norm of vectors or the corresponding matrix norm induced by Euclidean vector norm, which depends on the context. For $\tilde{x} \in \mathbb{R}^n$ and $A, B \subset \mathbb{R}^n$, $d(\tilde{x}, B) = \inf_{y \in B} \|\tilde{x} - y\|$ denotes the distance between \tilde{x} and B ; the deviation distance between A and B is denoted by $d(A, B) = \sup_{x \in A} \inf_{y \in B} \|x - y\|$ and the Hausdorff distance between A and B is denoted by $d_H(A, B) = \max\{d(A, B), d(B, A)\}$. We use $\mathcal{P}(\Xi)$ to denote the collection of all probability measures over the support set Ξ ; $\mathcal{P}_q(\Xi) = \{P \in \mathcal{P}(\Xi) : \mathbb{E}_P[\|\xi\|^q] < +\infty\}$ for $q > 0$.

The rest of the paper is organized as follows. In section 2, we discuss the continuity and boundedness properties of parametric linear programs based on Farkas' lemma. With these results, we carry out the quantitative stability analysis of two-stage stochastic linear programs with full random recourse in section 3. As an application of these quantitative results, we consider the rate of convergence of the SAA method for solving problem (1)-(3) in section 4. Finally, we have a conclusion section 5.

2. Prerequisites from parametric linear programs

In this section, in view of problem (3), we consider the following parametric linear programming problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax = b, \quad x \geq 0 \end{aligned} \tag{4}$$

and its dual problem

$$\begin{aligned} \max_{y \in \mathbb{R}^m} \quad & b^T y \\ \text{s.t.} \quad & A^T y \leq c, \end{aligned} \tag{5}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$. Here, A, b and c are all treated as model parameters. For the brevity of notations, we denote $\Upsilon = (A, b, c)$. Throughout this section, we use the following notations

$$\begin{aligned} F_{pri}(\Upsilon) &= \{x \in \mathbb{R}^n : Ax = b, \quad x \geq 0\}, \\ v_{pri}(\Upsilon) &= \min \{c^T x : x \in F_{pri}(\Upsilon)\}, \\ S_{pri}(\Upsilon) &= \{x \in F_{Pri}(\Upsilon) : c^T x = v_{pri}(\Upsilon)\} \end{aligned}$$

and

$$\begin{aligned} F_{dual}(\Upsilon) &= \{y \in \mathbb{R}^m : A^T y \leq c\}, \\ v_{dual}(\Upsilon) &= \max \{b^T x : x \in F_{dual}(\Upsilon)\}, \\ S_{dual}(\Upsilon) &= \{y \in F_{dual}(\Upsilon) : b^T y = v_{dual}(\Upsilon)\} \end{aligned}$$

to represent the feasible solution sets, optimal values and optimal solution sets of problem (4) and problem (5), respectively.

By the strong duality theorem of linear programs, we know that if problem (4) has finite optimal value, then its dual problem (5) has finite optimal value and their optimal values are the same. From the notations above, we know $\Upsilon \in \mathbb{R}^{m \times n} \times \mathbb{R}^m \times \mathbb{R}^n$. For further discussion, we use $\text{Pri}\emptyset \subseteq \mathbb{R}^{m \times n} \times \mathbb{R}^m \times \mathbb{R}^n$ to denote the set of parameters such that for any $\Upsilon \in \text{Pri}\emptyset$, problem (4) is infeasible. Similarly, let $\text{Dual}\emptyset$ be the collection of the parameter Υ for which problem (5) is infeasible.

To continue the discussion, we give the following two equivalent statements of the well-known Farkas' lemma. They will be useful for our later discussion.

Lemma 2.1 (Farkas' Lemma). *For given $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, exactly one of each group is true:*

- (i) (a) *There exists an x such that $Ax = b, x \geq 0$,*
 (b) *There exists a y satisfying $A^T y \geq 0, b^T y < 0$;*
- (ii) (a) *There exists an x that satisfies $Ax \leq b$,*
 (b) *There exists a y that satisfies $y \geq 0, A^T y = 0, b^T y < 0$.*

Based on this lemma, we have the following proposition.

Proposition 2.2. *The following two assertions hold:*

- (i) *If $d(\Upsilon, \text{Dual}\emptyset) > 0$, then for any $x \in F_{pri}(\Upsilon)$, we have*

$$\|x\| \leq \frac{\max\{\|b\|, c^T x\}}{d(\Upsilon, \text{Dual}\emptyset)};$$

- (ii) *If $d(\Upsilon, \text{Pri}\emptyset) > 0$, then for any $y \in F_{dual}(\Upsilon)$, we have*

$$\|y\| \leq \frac{\max\{\|c\|, -b^T y\}}{d(\Upsilon, \text{Pri}\emptyset)}.$$

Proof. (i) Let $x \in F_{pri}(\Upsilon)$. If $x = 0$, the conclusion obviously holds. Otherwise, we have $Ax = b, x \geq 0$ and $x \neq 0$. We consider the perturbation $\Delta\Upsilon := (\Delta A, 0, \Delta c)$ to Υ . Concretely, we define

$$\begin{aligned} \Delta A &= -\frac{bx^T}{\|x\|^2}, \\ \Delta c &= -\frac{\max\{0, c^T x + \varepsilon\}}{\|x\|^2} \cdot x \end{aligned}$$

for any scalar $\varepsilon > 0$.

From the definition of $\Delta\Upsilon$, we have

$$\begin{aligned}\Delta Ax &= -\frac{bx^T x}{\|x\|^2} = -b, \\ \Delta c^T x &= -\frac{\max\{0, c^T x + \varepsilon\}}{\|x\|^2} \cdot x^T x = -\max\{0, c^T x + \varepsilon\}.\end{aligned}$$

Hence, we have

$$\begin{aligned}x &\geq 0, \\ (A + \Delta A)x &= Ax - b = 0, \\ (c + \Delta c)^T x &= c^T x - \max\{0, c^T x + \varepsilon\} \leq -\varepsilon < 0.\end{aligned}$$

According to the Farkas' lemma, we have that the system $(A + \Delta A)^T y \leq (c + \Delta c)$ is inconsistent, which means $d(\Upsilon, \text{Dual}\emptyset) \leq \|\Delta\Upsilon\|$. Note that

$$\|\Delta\Upsilon\| = \max\{\|\Delta A\|, \|\Delta c\|\} \leq \frac{\max\{\|b\|, c^T x + \varepsilon\}}{\|x\|},$$

so

$$\|x\| \leq \frac{\max\{\|b\|, c^T x + \varepsilon\}}{\|\Delta\Upsilon\|} \leq \frac{\max\{\|b\|, c^T x + \varepsilon\}}{d(\Upsilon, \text{Dual}\emptyset)}.$$

This holds for arbitrary $\varepsilon > 0$, we obtain that

$$\|x\| \leq \frac{\max\{\|b\|, c^T x\}}{d(\Upsilon, \text{Dual}\emptyset)}.$$

(ii) Let $y \in F_{\text{dual}}(\Upsilon)$. The conclusion obviously holds when $y = 0$, so we assume $y \neq 0$ in what follows. $y \in F_{\text{dual}}(\Upsilon)$ means $A^T(-y) \geq -c$. We consider the following perturbation $\Delta\Upsilon := (\Delta A, \Delta b, 0)$, where

$$\begin{aligned}\Delta A &= -\frac{yc^T}{\|y\|^2}, \\ \Delta b &= \frac{\max\{0, -b^T y + \varepsilon\}}{\|y\|^2} \cdot y\end{aligned}$$

for any $\varepsilon > 0$. Then we have

$$\begin{aligned}\Delta A^T(-y) &= -\frac{cy^T(-y)}{\|y\|^2} = c, \\ \Delta b^T(-y) &= \frac{\max\{0, -b^T y + \varepsilon\}}{\|y\|^2} \cdot y^T(-y) = -\max\{0, -b^T y + \varepsilon\}.\end{aligned}$$

This implies

$$\begin{aligned}(A + \Delta A)^T(-y) &\geq -c + c = 0, \\ (b + \Delta b)^T(-y) &= b^T(-y) - \max\{0, -b^T y + \varepsilon\} \leq -\varepsilon < 0.\end{aligned}$$

According to the Farkas' lemma, there does not exist an x such that $(A + \Delta A)x = (b + \Delta b), x \geq 0$. So we must have $\|\Delta \Upsilon\| \geq d(\Upsilon, \text{Pri}\emptyset)$. Note that

$$\|\Delta \Upsilon\| = \max\{\|\Delta A\|, \|\Delta b\|\} \leq \frac{\max\{\|c\|, -b^T y\}}{\|y\|}.$$

Thus, we obtain

$$\|y\| \leq \frac{\max\{\|c\|, -b^T y\}}{\|\Delta \Upsilon\|} \leq \frac{\max\{\|c\|, -b^T y\}}{d(\Upsilon, \text{Pri}\emptyset)}.$$

□

Remark 1. Proposition 2.2 requires that $d(\Upsilon, \text{Dual}\emptyset) > 0$ and $d(\Upsilon, \text{Pri}\emptyset) > 0$. They imply that both the dual problem and the primal problem are feasible. Moreover, they ensure the feasibility of problem (4) or (5) under sufficiently small perturbation. Otherwise, if $d(\Upsilon, \text{Dual}\emptyset) = 0$ (or $d(\Upsilon, \text{Pri}\emptyset) = 0$), the dual problem (or the primal problem) would probably be infeasible under any small perturbation.

According to Proposition 2.2, if $x \in S_{\text{pri}}(\Upsilon)$ and $y \in S_{\text{dual}}(\Upsilon)$, it is easy to derive the following conclusions.

Corollary 2.3. *The following two assertions hold:*

(i) *If $d(\Upsilon, \text{Dual}\emptyset) > 0$, then for any $x^* \in S_{\text{pri}}(\Upsilon)$, we have*

$$\|x^*\| \leq \frac{\max\{\|b\|, v_{\text{pri}}(\Upsilon)\}}{d(\Upsilon, \text{Dual}\emptyset)};$$

(ii) *If $d(\Upsilon, \text{Pri}\emptyset) > 0$, then for any $y^* \in S_{\text{dual}}(\Upsilon)$, we have*

$$\|y^*\| \leq \frac{\max\{\|c\|, -v_{\text{dual}}(\Upsilon)\}}{d(\Upsilon, \text{Pri}\emptyset)}.$$

Proposition 2.4. *Suppose $d(\Upsilon, \text{Dual}\emptyset) > 0$ and $d(\Upsilon, \text{Pri}\emptyset) > 0$, then we have*

$$-\frac{\|c\| \|b\|}{d(\Upsilon, \text{Dual}\emptyset)} \leq v_{\text{pri}}(\Upsilon) = v_{\text{dual}}(\Upsilon) \leq \frac{\|c\| \|b\|}{d(\Upsilon, \text{Pri}\emptyset)}.$$

Proof. When $v_{\text{pri}}(\Upsilon) > 0$, the lower bound on the left-hand side obviously holds. So without loss of generality, we assume that $v_{\text{pri}}(\Upsilon) \leq 0$. Then for any $x^* \in S_{\text{pri}}(\Upsilon)$, we have

$$v_{\text{pri}}(\Upsilon) = c^T x^* \geq -|c^T x^*| \geq -\|c\| \|x^*\| \geq -\frac{\|c\| \|b\|}{d(\Upsilon, \text{Dual}\emptyset)}.$$

The last inequality comes from Corollary 2.3 and $v_{pri}(\Upsilon) \leq 0$.

Analogously, the right-hand side upper bound obviously holds if $v_{dual}(\Upsilon) \leq 0$. For $v_{dual}(\Upsilon) > 0$ and $y^* \in S_{dual}(\Upsilon)$, we have

$$v_{dual}(\Upsilon) = b^T y^* \leq \|b\| \|y^*\| \leq \frac{\|b\| \|c\|}{d(\Upsilon, \text{Pri}\emptyset)}.$$

Finally, $d(\Upsilon, \text{Dual}\emptyset) > 0$ and $d(\Upsilon, \text{Pri}\emptyset) > 0$ guarantee that $v_{pri}(\Upsilon) = v_{dual}(\Upsilon)$. Thus, we complete the proof. \square

As for the model parameters, we have the following description.

Lemma 2.5. *The following assertions hold:*

(i) *For any Υ such that $d(\Upsilon, \text{Pri}\emptyset) > 0$, we have*

$$\frac{\|\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset)} \geq 1;$$

(ii) *For any Υ such that $d(\Upsilon, \text{Dual}\emptyset) > 0$, we have*

$$\frac{\|\Upsilon\|}{d(\Upsilon, \text{Dual}\emptyset)} \geq 1.$$

Proof. Note that $\text{Pri}\emptyset$ is a cone, so is $\text{cl}(\text{Pri}\emptyset)$. Let $d(\Upsilon, \text{Pri}\emptyset) = \|\Upsilon - \tilde{\Upsilon}\|$, where $\tilde{\Upsilon} \in \text{cl}(\text{Pri}\emptyset)$. Then we have $\lambda\tilde{\Upsilon} \in \text{cl}(\text{Pri}\emptyset)$ for any $\lambda > 0$ and

$$d(\Upsilon, \text{Pri}\emptyset) = \|\Upsilon - \tilde{\Upsilon}\| \leq \|\Upsilon - \lambda\tilde{\Upsilon}\| \leq \|\Upsilon\| + \lambda \|\tilde{\Upsilon}\|.$$

Letting $\lambda \downarrow 0$ proves (i). (ii) can be proved similarly. \square

With Lemma 2.5, we can deduce the boundedness of the optimal solutions of the primal problem and those of the dual problem.

Theorem 2.6. *Assume $d(\Upsilon, \text{Pri}\emptyset) > 0$ and $d(\Upsilon, \text{Dual}\emptyset) > 0$. Then for any $x^* \in S_{pri}(\Upsilon)$ and $y^* \in S_{dual}(\Upsilon)$, we have*

$$\|x^*\| \leq \frac{\|b\|}{d(\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset)}$$

and

$$\|y^*\| \leq \frac{\|c\|}{d(\Upsilon, \text{Pri}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Dual}\emptyset)}.$$

Proof. We know from Corollary 2.3 and Proposition 2.4 that

$$\begin{aligned}
\|x^*\| &\leq \frac{\max\{\|b\|, v_{pri}(\Upsilon)\}}{d(\Upsilon, \text{Dual}\emptyset)} \\
&\leq \frac{\|b\|}{d(\Upsilon, \text{Dual}\emptyset)} \max\left\{1, \frac{\|c\|}{d(\Upsilon, \text{Pri}\emptyset)}\right\} \\
&\leq \frac{\|b\|}{d(\Upsilon, \text{Dual}\emptyset)} \max\left\{1, \frac{\|\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset)}\right\} \\
&= \frac{\|b\|}{d(\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset)}.
\end{aligned}$$

The last equality comes from Lemma 2.5. Similarly,

$$\begin{aligned}
\|y^*\| &\leq \frac{\max\{\|c\|, -v_{dual}(\Upsilon)\}}{d(\Upsilon, \text{Pri}\emptyset)} \\
&\leq \frac{\|c\|}{d(\Upsilon, \text{Pri}\emptyset)} \max\left\{1, \frac{\|b\|}{d(\Upsilon, \text{Dual}\emptyset)}\right\} \\
&\leq \frac{\|c\|}{d(\Upsilon, \text{Pri}\emptyset)} \max\left\{1, \frac{\|\Upsilon\|}{d(\Upsilon, \text{Dual}\emptyset)}\right\} \\
&= \frac{\|c\|}{d(\Upsilon, \text{Pri}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Dual}\emptyset)}.
\end{aligned}$$

□

It can be deduced from

$$\frac{\|\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset)} \geq 1, \quad \frac{\|\Upsilon\|}{d(\Upsilon, \text{Dual}\emptyset)} \geq 1$$

and $d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset) \leq \min\{d(\Upsilon, \text{Pri}\emptyset), d(\Upsilon, \text{Dual}\emptyset)\}$ that

$$\frac{\|\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset)} \geq \max\left\{\frac{\|\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset)}, \frac{\|\Upsilon\|}{d(\Upsilon, \text{Dual}\emptyset)}\right\} \geq 1.$$

With the above preparation, we can now consider the general perturbation to Υ and derive an upper bound estimation about the difference between the optimal value of the original problem and that of the perturbed one.

Theorem 2.7. *Suppose that $d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset) > \|\Delta\Upsilon\| \geq 0$. Then we have*

$$\begin{aligned}
&|v_{pri}(\Upsilon + \Delta\Upsilon) - v_{pri}(\Upsilon)| \\
\leq &\|\Delta A\| \frac{\|c + \Delta c\| \vee \|c\|}{d(\Upsilon, \text{Pri}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|b + \Delta b\| \vee \|b\|}{d(\Upsilon + \Delta\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon + \Delta\Upsilon\|}{d(\Upsilon + \Delta\Upsilon, \text{Pri}\emptyset)} \\
&+ \|\Delta b\| \frac{\|c + \Delta c\| \vee \|c\|}{d(\Upsilon, \text{Pri}\emptyset) \wedge d(\Upsilon + \Delta\Upsilon, \text{Pri}\emptyset)} \cdot \frac{\|\Upsilon\| \vee \|\Upsilon + \Delta\Upsilon\|}{d(\Upsilon, \text{Dual}\emptyset) \wedge d(\Upsilon + \Delta\Upsilon, \text{Dual}\emptyset)} \\
&+ \|\Delta c\| \frac{\|b + \Delta b\| \vee \|b\|}{d(\Upsilon, \text{Dual}\emptyset) \wedge d(\Upsilon + \Delta\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon\| \vee \|\Upsilon + \Delta\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset) \wedge d(\Upsilon + \Delta\Upsilon, \text{Pri}\emptyset)}.
\end{aligned}$$

Proof. Let $\Delta\Upsilon = (\Delta A, \Delta b, \Delta c)$ and $y^* \in S_{dual}(\Upsilon + \Delta\Upsilon)$. Denote $\Delta\Upsilon' = (0, 0, \Delta c - \Delta A^T y^*)$. We first consider the case that

$$|v_{pri}(\Upsilon + \Delta\Upsilon) - v_{pri}(\Upsilon)| = v_{pri}(\Upsilon + \Delta\Upsilon) - v_{pri}(\Upsilon).$$

Take $x^* \in S_{pri}(\Upsilon)$, we have $x^* \in F_{pri}(\Upsilon + \Delta\Upsilon')$ due to the special perturbation $\Delta\Upsilon'$. Then we have the following estimation

$$\begin{aligned} v_{pri}(\Upsilon + \Delta\Upsilon') &\leq (c + \Delta c - \Delta A^T y^*)^T x^* = c^T x^* + \Delta c^T x^* - (y^*)^T \Delta A^T x^* \\ &= v_{pri}(\Upsilon) + \Delta c^T x^* - (y^*)^T \Delta A^T x^*. \end{aligned}$$

Therefore,

$$\begin{aligned} &v_{pri}(\Upsilon + \Delta\Upsilon') - v_{pri}(\Upsilon) \\ &\leq \Delta c^T x^* - (y^*)^T \Delta A^T x^* \\ &\leq \|\Delta c\| \|x^*\| + \|y^*\| \|\Delta A\| \|x^*\| \\ &\leq \|\Delta c\| \frac{\|b\|}{d(\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset)} \\ &\quad + \|\Delta A\| \frac{\|b\|}{d(\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset)} \frac{\|c\|}{d(\Upsilon + \Delta\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon + \Delta\Upsilon\|}{d(\Upsilon + \Delta\Upsilon, \text{Pri}\emptyset)}. \end{aligned}$$

On the other hand, we estimate $v_{dual}(\Upsilon + \Delta\Upsilon) - v_{dual}(\Upsilon + \Delta\Upsilon')$. Since $y^* \in S_{dual}(\Upsilon + \Delta\Upsilon)$, $y^* \in F_{dual}(\Upsilon + \Delta\Upsilon')$, we have

$$\begin{aligned} v_{dual}(\Upsilon + \Delta\Upsilon) - v_{dual}(\Upsilon + \Delta\Upsilon') &\leq \Delta b^T y^* \\ &\leq \|\Delta b\| \frac{\|c\|}{d(\Upsilon + \Delta\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon + \Delta\Upsilon\|}{d(\Upsilon + \Delta\Upsilon, \text{Pri}\emptyset)}. \end{aligned}$$

The assumption $d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset) > \|\Delta\Upsilon\| \geq 0$ ensures that both the primal problem and the dual problem are feasible. Then the strong duality theorem holds. Therefore, we have

$$\begin{aligned} &v_{pri}(\Upsilon + \Delta\Upsilon) - v_{pri}(\Upsilon) \\ &= v_{dual}(\Upsilon + \Delta\Upsilon) - v_{dual}(\Upsilon + \Delta\Upsilon') + v_{pri}(\Upsilon + \Delta\Upsilon') - v_{pri}(\Upsilon) \\ &\leq \|\Delta A\| \frac{\|c\|}{d(\Upsilon, \text{Pri}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|b\|}{d(\Upsilon + \Delta\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon + \Delta\Upsilon\|}{d(\Upsilon + \Delta\Upsilon, \text{Pri}\emptyset)} \\ &\quad + \|\Delta b\| \frac{\|c\|}{d(\Upsilon + \Delta\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon + \Delta\Upsilon\|}{d(\Upsilon + \Delta\Upsilon, \text{Pri}\emptyset)} \\ &\quad + \|\Delta c\| \frac{\|b\|}{d(\Upsilon, \text{Pri}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Dual}\emptyset)}. \end{aligned}$$

For the other case that $|v_{pri}(\Upsilon + \Delta\Upsilon) - v_{pri}(\Upsilon)| = v_{pri}(\Upsilon) - v_{pri}(\Upsilon + \Delta\Upsilon)$, we can set $\tilde{\Upsilon} = \Upsilon + \Delta\Upsilon$ and $\Delta\tilde{\Upsilon} = -\Delta\Upsilon$. Then $v_{pri}(\Upsilon) - v_{pri}(\Upsilon + \Delta\Upsilon) = v_{pri}(\Upsilon + \Delta\tilde{\Upsilon}) -$

$v_{pri}(\tilde{\Upsilon})$. According to the above deduction method, we obtain that

$$\begin{aligned}
& v_{pri}(\Upsilon + \Delta\tilde{\Upsilon}) - v_{pri}(\tilde{\Upsilon}) \\
\leq & \|\Delta A\| \frac{\|c + \Delta c\|}{d(\Upsilon + \Delta\Upsilon, \text{Pri}\emptyset)} \cdot \frac{\|\Upsilon + \Delta\Upsilon\|}{d(\Upsilon + \Delta\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|b + \Delta b\|}{d(\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset)} \\
& + \|\Delta b\| \frac{\|c + \Delta c\|}{d(\Upsilon, \text{Pri}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Dual}\emptyset)} \\
& + \|\Delta c\| \frac{\|b + \Delta b\|}{d(\Upsilon + \Delta\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon + \Delta\Upsilon\|}{d(\Upsilon + \Delta\Upsilon, \text{Pri}\emptyset)}.
\end{aligned}$$

Then, we complete the proof. \square

It is straightforward to derive from Theorem 2.7 the following corollary.

Corollary 2.8. *Suppose that $d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset) > 0$. Then there exists a positive constant $\delta(\Upsilon)$, such that for any $\Upsilon_1 = (A_1, b_1, c_1)$ and $\Upsilon_2 = (A_2, b_2, c_2)$ with $\|\Upsilon_1 - \Upsilon\| \leq \delta(\Upsilon)$ and $\|\Upsilon_2 - \Upsilon\| \leq \delta(\Upsilon)$, we have*

$$\begin{aligned}
d(\Upsilon_1, \text{Pri}\emptyset \cup \text{Dual}\emptyset) & \geq \frac{1}{2}d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset), \\
d(\Upsilon_2, \text{Pri}\emptyset \cup \text{Dual}\emptyset) & \geq \frac{1}{2}d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset)
\end{aligned} \tag{6}$$

and

$$\begin{aligned}
|v_{pri}(\Upsilon_1) - v_{pri}(\Upsilon_2)| & \leq \frac{8\|A_1 - A_2\| \cdot \|c_1\| \vee \|c_2\| \cdot \|b_1\| \vee \|b_2\| \cdot \|\Upsilon_1\| \cdot \|\Upsilon_2\|}{d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset)^4} \\
& + \frac{4\|b_1 - b_2\| \cdot \|c_1\| \vee \|c_2\| \cdot \|\Upsilon_1\| \vee \|\Upsilon_2\|}{d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset)^2} \\
& + \frac{4\|c_1 - c_2\| \cdot \|b_1\| \vee \|b_2\| \cdot \|\Upsilon_1\| \vee \|\Upsilon_2\|}{d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset)^2}.
\end{aligned} \tag{7}$$

It is noteworthy that for the special case $\delta(\Upsilon) = \frac{1}{2}d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset)$, (6) holds. In detail, we have

$$\begin{aligned}
d(\Upsilon_1, \text{Pri}\emptyset \cup \text{Dual}\emptyset) & \geq d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset) - \|\Upsilon_1 - \Upsilon\| \\
& \geq d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset) - \frac{1}{2}d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset) \\
& = \frac{1}{2}d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset).
\end{aligned}$$

We can derive the analogous conclusion for $d(\Upsilon_2, \text{Pri}\emptyset \cup \text{Dual}\emptyset)$. Therefore, for the above case, we can obtain (7) from (6).

As an illustration to our conclusions above, we give a simple example.

Example 2.9. Consider the following linear programming problem:

$$\begin{aligned} \min \quad & 5x_1 + 4x_2 \\ \text{s.t.} \quad & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \\ & x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

Its dual problem is

$$\begin{aligned} \max \quad & 3y_1 + 4y_2 \\ \text{s.t.} \quad & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \leq \begin{pmatrix} 5 \\ 4 \end{pmatrix}. \end{aligned}$$

Let

$$\Upsilon := \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix} \right).$$

Obviously, we have $1/2 < d(\Upsilon, \text{Pri}\emptyset) < 1$, $1/2 < d(\Upsilon, \text{Dual}\emptyset) < 1$ and $1/2 < d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset) < 1$. The optimal solution of the primal problem is $x_1 = 3$, $x_2 = 4$ and the optimal solution of the dual problem is $y_1 = 5$, $y_2 = 4$. Their common optimal value is 31. Note that $\|\Upsilon\| = \sqrt{41}$ and

$$\begin{aligned} \|(3, 4)^T\| &= 5 \leq 41 \leq \frac{\|(5, 4)^T\|}{d(\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset)}, \\ \|(5, 4)^T\| &= \sqrt{41} \leq 41 \leq \frac{\|(3, 4)^T\|}{d(\Upsilon, \text{Pri}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Dual}\emptyset)}, \end{aligned}$$

which verify Theorem 2.6.

We consider the following two small perturbations to Υ :

$$\begin{aligned} \Upsilon_1 &:= \left(\begin{pmatrix} 7/8 & 0 \\ 0 & 7/8 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix} \right), \\ \Upsilon_2 &:= \left(\begin{pmatrix} 3/4 & 0 \\ 0 & 3/4 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix} \right). \end{aligned}$$

It is easy to find that the optimal solutions and values of the primal linear programming problem with parameters Υ_1 and Υ_2 are $x_1 = 24/7, x_2 = 32/7, 248/7$ and $x_1 = 4, x_2 = 16/3, 124/3$, respectively. Note that we have

$$\begin{aligned} \|\Upsilon - \Upsilon_1\| &= \frac{1}{8} \leq \frac{d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset)}{2}, \\ \|\Upsilon - \Upsilon_2\| &= \frac{1}{4} \leq \frac{d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset)}{2}. \end{aligned}$$

Moreover, we have

$$\begin{aligned} \left| \frac{248}{7} - \frac{124}{3} \right| &= \frac{124}{21}, \\ \left\| \begin{pmatrix} 7/8 & 0 \\ 0 & 7/8 \end{pmatrix} - \begin{pmatrix} 3/4 & 0 \\ 0 & 3/4 \end{pmatrix} \right\| &= \frac{1}{8}, \\ \|(5, 4)^T\| &= \sqrt{41}, \\ \|(3, 4)^T\| &= 5, \\ \|\Upsilon_1\| &= \sqrt{41}, \\ \|\Upsilon_2\| &= \sqrt{41}. \end{aligned}$$

Then we obtain that

$$\begin{aligned} & \frac{8 \|A_1 - A_2\| \cdot \|c_1\| \vee \|c_2\| \cdot \|b_1\| \vee \|b_2\| \cdot \|\Upsilon_1\| \cdot \|\Upsilon_2\|}{d(\Upsilon, \text{Pri}\emptyset \cup \text{Dual}\emptyset)^4} \\ & \geq \frac{5 \times 41 \sqrt{41}}{1^4} = 205 \sqrt{41} > \frac{124}{21} = \left| \frac{248}{7} - \frac{124}{3} \right|, \end{aligned}$$

which verifies (7) in Corollary 2.8.

It is worth mentioning that the perturbation analysis results about parametric linear programs in [21] can not provide the error bound estimation for both the optimal solution and optimal value of above linear programs because of its restrictive assumptions. Specifically, it is assumed in [21] that the convex hull of the row vectors of the coefficient matrix contains the zero vector as an inner point. That is, for the above example, the convex hull of $\{(1, 0)^T, (0, 1)^T\}$ should contain $(0, 0)^T$ as an inner point, which doesn't hold obviously. In this regard, we extend the results in [21]. The reason why the authors in [21] impose such a strict condition is that their arguments need the boundedness of the feasible solution set. This in return is due to that the analysis in [21] relies on the Hoffman constant, while we simply use the Farkas' lemma.

With the above results, we can carry out the quantitative stability analysis of two-stage stochastic linear programs with full random recourse in what follows.

3. Stability of two-stage stochastic linear programs with full random recourse

In order to state the quantitative stability results, we need to introduce the so-called pseudo metrics. They have been extensively used in the quantitative stability study of stochastic programming problems ([8,9,24]) and distributionally robust optimization problems (see for example [25,26]). Here we focus on a class of pseudo metrics known as the ζ -structure metric.

Definition 3.1 (ζ -structure metrics). Let \mathcal{G} be a set of real-valued measurable functions on Ξ . For any two probability measures $P, Q \in \mathcal{P}(\Xi)$,

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

is called the ζ -structure metric between P and Q induced by \mathcal{G} . Also, we call \mathcal{G} the generator of $\mathbb{D}_{\mathcal{G}}(\cdot, \cdot)$.

$\mathbb{D}_{\mathcal{G}}(\cdot, \cdot)$ is a pseudo metric because $\mathbb{D}_{\mathcal{G}}(P, Q) = 0$ usually fails to imply $P = Q$ unless \mathcal{G} is rich enough. It is known from Definition 3.1 that, if we choose different \mathcal{G} s, different ζ -structure metrics can be obtained. Therefore, ζ -structure metrics include many commonly used probability metrics. For example, for $p \geq 1$, if we let

$$\mathcal{G}_{FM_p} := \left\{ g : \Xi \rightarrow \mathbb{R} : |g(\xi^1) - g(\xi^2)| \leq \max \{1, \|\xi^1\|, \|\xi^2\|\}^{p-1} \|\xi^1 - \xi^2\| \right\},$$

the corresponding ζ -structure metric is called the p -th order Fortet-Mourier metric, which is often used in the stability analysis of two-stage stochastic programs [9]. We will denote the p -th Fortet-Mourier metric by $\zeta_p(\cdot, \cdot)$. Another commonly used ζ -structure metric based on \mathcal{G}_{FM_p} , denoted by $\zeta_{p,ph_k}(\cdot, \cdot)$ with some positive integer k , is

$$\zeta_{p,ph_k}(P_{\xi}, Q_{\xi}) := \sup \left| \int_B g(\xi)(P_{\xi} - Q_{\xi})(d\xi) : g(\xi) \in \mathcal{G}_{FM_p}, B \in \mathcal{B}_{ph_k} \right|,$$

here \mathcal{B}_{ph_k} denotes the set of all the polyhedrons in Ξ which have at most k faces.

If we denote

$$\mathcal{G}_{TV} = \{g : \Xi \rightarrow \mathbb{R} : g \text{ is measurable and } \sup_{\xi \in \Xi} |g(\xi)| \leq 1\},$$

the resulting metric is the total variation metric. We use $\mathbb{D}_{TV}(P, Q)$ to denote the total variation metric between P and Q . One can refer to [22] for more details about the relationships among these pseudo metrics.

In what follows, we denote the optimal value and optimal solution set of problem (1)-(3) by $v(P)$ and $S(P)$, respectively. We call $\psi_P(\cdot) : \mathbb{R}_+ \rightarrow \mathbb{R}$ the growth function of problem (1)-(3), if

$$\psi_P(\pi) := \min\{\mathbb{E}_P[f(x, \xi)] - v(P) : d(x, S(P)) \geq \pi, x \in X\}.$$

It is not difficult to verify from its definition that $\psi_P(\cdot)$ is nondecreasing and lower semicontinuous. Its inverse function is defined as

$$\psi_P^{-1}(t) = \sup\{\pi \in \mathbb{R}_+ : \psi(\pi) \leq t\}.$$

Then, we call

$$\Psi_P(t) = t + \psi_P^{-1}(2t), \quad t \in \mathbb{R}_+,$$

the conditioning function of problem (1)-(3). Obviously, Ψ_P is lower semicontinuous and increasing. One can refer to [8,9] for details about the above functions.

Moreover, we need the following assumption.

Assumption 3.2. For the full random two-stage stochastic linear programming problem (1)-(3), we assume that:

- (A1) $\kappa(x, \xi) := d(\Upsilon(x, \xi), \bar{\text{Pri}}\emptyset \cup \bar{\text{Dual}}\emptyset) > 0$ for any $\xi \in \Xi$ and $x \in X$, where $\Upsilon(x, \xi) := (W(\xi), h(\xi) - T(\xi)x, q(\xi))$, $\bar{\text{Pri}}\emptyset(\bar{\text{Dual}}\emptyset)$ denotes the set of parameters such that

(the dual problem of) the problem (3) is infeasible;
(A2) $q(\xi), W(\xi), T(\xi)$ and $h(\xi)$ all depend affine linearly on ξ .

Remark 2. Assumption 3.2 is necessary and reasonable. (A1) requires that $\kappa(x, \xi) > 0$ for any $x \in X$ and $\xi \in \Xi$, which avoids that these coefficients locate on the boundary of $\text{Pri}\emptyset \cup \text{Dual}\emptyset$. It indicates the feasibility of problem (3) for each $(x, \xi) \in X \times \Xi$, which corresponds to the relatively complete recourse. Moreover, the rationality of (A1) lies in that it ensures the feasibility of problem (3) for any sufficiently small perturbation on coefficients. (A2) is a commonly used assumption in the quantitative stability analysis (see, for example, [9]). Mathematically, (A2) means that there exists a positive constant C , such that

$$\begin{aligned} \|\Lambda(\xi_1) - \Lambda(\xi_2)\| &\leq C \|\xi_1 - \xi_2\|, \\ \|\Lambda(\xi)\| &\leq C \max\{1, \|\xi\|\}, \end{aligned}$$

for any $\xi, \xi_1, \xi_2 \in \Xi$ and $\Lambda(\cdot) = q(\cdot), W(\cdot), T(\cdot)$, or $h(\cdot)$.

Denote $\kappa(\xi) = \min_{x \in X} \kappa(x, \xi)$. We have the following conclusion.

Lemma 3.3. *Let Assumption 3.2 hold and X be a compact set. Then $\kappa(\cdot)$ is Lipschitz continuous over Ξ .*

Proof. We know from its definition that, $\kappa(\cdot, \cdot)$ is continuous over $X \times \Xi$ under Assumption 3.2. Since X is a compact set, we obtain $\kappa(\xi) > 0$ for any $\xi \in \Xi$, and there exist $x, x' \in X$ such that $\kappa(\xi') = \kappa(x', \xi')$ and $\kappa(\xi) = \kappa(x, \xi)$ for any $\xi, \xi' \in \Xi$. Then we have

$$|\kappa(\xi') - \kappa(\xi)| = |\kappa(x', \xi') - \kappa(x, \xi)|.$$

If $|\kappa(x', \xi') - \kappa(x, \xi)| = \kappa(x', \xi') - \kappa(x, \xi)$, then $|\kappa(x', \xi') - \kappa(x, \xi)| \leq \kappa(x, \xi') - \kappa(x, \xi)$ because $\kappa(x', \xi') \leq \kappa(x, \xi')$; If $|\kappa(x', \xi') - \kappa(x, \xi)| = \kappa(x, \xi) - \kappa(x', \xi')$, then $|\kappa(x', \xi') - \kappa(x, \xi)| \leq \kappa(x', \xi) - \kappa(x', \xi')$ because $\kappa(x, \xi) \leq \kappa(x', \xi)$. To summarize, we obtain

$$|\kappa(\xi') - \kappa(\xi)| \leq \max\{|\kappa(x', \xi') - \kappa(x', \xi)|, |\kappa(x, \xi') - \kappa(x, \xi)|\}.$$

Note that

$$\begin{aligned} |\kappa(x', \xi') - \kappa(x', \xi)| &= |d(\Upsilon(x', \xi'), \text{Pri}\emptyset \cup \text{Dual}\emptyset) - d(\Upsilon(x', \xi), \text{Pri}\emptyset \cup \text{Dual}\emptyset)| \\ &\leq \|\Upsilon(x', \xi') - \Upsilon(x', \xi)\| \\ &\leq C(1 + \|x'\|) \|\xi' - \xi\|. \end{aligned}$$

The last inequality comes from (A2). Similarly, we can obtain

$$|\kappa(x, \xi') - \kappa(x, \xi)| \leq C(1 + \|x\|) \|\xi' - \xi\|.$$

Finally, we derive

$$\begin{aligned} |\kappa(\xi') - \kappa(\xi)| &\leq C \max\{1 + \|x'\|, 1 + \|x\|\} \|\xi' - \xi\| \\ &\leq \max_{x \in X} \{C(1 + \|x\|)\} \|\xi' - \xi\|. \end{aligned} \tag{8}$$

Due to the compactness of X , we have that

$$\max_{x \in X} \{C(1 + \|x\|)\} < +\infty.$$

Thus, we conclude that $\kappa(\cdot)$ is Lipschitz continuous over Ξ . \square

In what follows, we always assume that the feasible solution set of the first stage problem (1) is compact, i.e., X is a compact set. This is a conventional assumption in the quantitative stability analysis, see for example [8,9,12,19].

Lemma 3.4. *Suppose that Assumption 3.2 holds. Then, for any $\xi \in \Xi$ and $x, x_1, x_2 \in X$, there exist positive constants $\delta(\xi), L_1, L_2, L_3$, such that*

$$|f(x, \xi_1) - f(x, \xi_2)| \leq \frac{L_1}{\min\{1, \kappa^4(\xi)\}} \max\{1, \|\xi_1\|, \|\xi_2\|\}^4 \|\xi_1 - \xi_2\|, \quad (9a)$$

$$|f(x_1, \xi) - f(x_2, \xi)| \leq \frac{L_2}{\kappa^2(\xi)} \max\{1, \|\xi\|\}^3 \|x_1 - x_2\|, \quad (9b)$$

$$|f(x, \xi)| \leq \frac{L_3}{\kappa^2(\xi)} \max\{1, \|\xi\|\}^3, \quad (9c)$$

when $\|\xi_1 - \xi\| \leq \delta(\xi)$ and $\|\xi_2 - \xi\| \leq \delta(\xi)$.

Proof. Let $\Upsilon_1 = (W(\xi_1), h(\xi_1) - T(\xi_1)x, q(\xi_1))$ and $\Upsilon_2 = (W(\xi_2), h(\xi_2) - T(\xi_2)x, q(\xi_2))$, (9a) can be directly derived from Corollary 2.8.

Similarly, let $\Upsilon_1 = (W(\xi), h(\xi) - T(\xi)x_1, q(\xi))$ and $\Upsilon_2 = (W(\xi), h(\xi) - T(\xi)x_2, q(\xi))$, (9b) can also be derived from Corollary 2.8.

For (9c), notice that $f(x, \xi) = \langle c, x \rangle + \Phi(x, \xi)$, which implies $|f(x, \xi)| = \|c\| \|x\| + |\Phi(x, \xi)|$. Let $\Upsilon = (W(\xi), h(\xi) - T(\xi)x, q(\xi))$ and y^* be any dual optimal solution of problem (3) under Υ . We can deduce from Theorem 2.6 and (A2) that

$$\begin{aligned} \|y^*\| &\leq \frac{\|h(\xi) - T(\xi)x\|}{d(\Upsilon, \text{Dual}\emptyset)} \cdot \frac{\|\Upsilon\|}{d(\Upsilon, \text{Pri}\emptyset)} \\ &\leq \frac{L \max\{1, \|\xi\|\}^2}{\kappa^2(\xi)} \end{aligned}$$

for some constant $L > 0$. Then we have

$$\begin{aligned} |f(x, \xi)| &\leq \|c\| \|x\| + \|q(\xi)\| \frac{L \max\{1, \|\xi\|\}}{\kappa^2(\xi)} \\ &\leq \frac{L_3}{\kappa^2(\xi)} \max\{1, \|\xi\|\}^3, \end{aligned}$$

for some positive scalar $L_3 > 0$. \square

We can see from (8) that $\delta(\xi)$ can be set to $\kappa(\xi)/2\theta$, here $\theta := C \max_{x \in X} (1 + \|x\|)$. This observation is important for the next assertion.

Let $\Xi_R := [-R, R] \times [-R, R] \times \dots \times [-R, R] \cap \Xi$ for some scalar $R \geq 1$, and its complementary set is denoted by $\Xi_R^c = \Xi \setminus \Xi_R$. In addition, we need the following assumption to control the distance between $\Upsilon(x, \xi)$ and $\text{Pri}\emptyset \cup \text{Dual}\emptyset$ for all $x \in X$.

This is achieved by imposing a lower bound on $\kappa(\xi)$. Concretely, we make the following assumption.

Assumption 3.5. There exists a positive scalar C_0 such that $\kappa(\xi) \geq 1/\|\xi\|$ for any $\|\xi\| \geq C_0$.

Note that we can also use a more generic assumption $\kappa(\xi) \geq 1/\|\xi\|^\gamma$ here for any positive scalar γ , without affecting the following discussion. We simply let $\gamma = 1$ for the simplicity of notations. Obviously, Assumption 2 holds when $\kappa(\xi) \geq \kappa$ for some positive scalar κ . In this case, we can let

$$\|\xi\| \geq C_0 := \frac{1}{\kappa}$$

and then we have $\kappa(\xi) \geq 1/\|\xi\|$.

Due to the continuity of $\kappa(\xi)$ from (8), we have that $\min_{\{\xi \in \Xi: \|\xi\| \leq C_0\}} \kappa(\xi)$ is a finite positive number. This implies that for any positive number $C_1 (\geq C_0)$ satisfying

$$\min_{\{\xi \in \Xi: \|\xi\| \leq C_0\}} \kappa(\xi) \geq \frac{1}{C_1},$$

we have $\min_{\{\xi \in \Xi: \|\xi\| \leq C_1\}} \kappa(\xi) \geq 1/C_1$.

Lemma 3.6. *Suppose that Assumptions 3.2 and 3.5 hold. Then there exists a positive scalar L_R , which depends on R , such that*

$$|f(x, \xi^1) - f(x, \xi^2)| \leq L_R \max\{1, \|\xi^1\|, \|\xi^2\|\}^4 \|\xi^1 - \xi^2\| \quad (10)$$

for any $\xi^1, \xi^2 \in \Xi_R$.

Moreover, if Ξ is convex, L_R can be chosen as $L_1 s^2 R^4$ for sufficiently large R . Here L_1 is the constant in Lemma 3.4;

Proof. We prove the first claim by contradiction. If (10) fails, then there exist two convergent sequences $\{\xi_n^1\}_{n=1}^\infty$ and $\{\xi_n^2\}_{n=1}^\infty$, such that

$$|f(x, \xi_n^1) - f(x, \xi_n^2)| \geq n \max\{1, \|\xi_n^1\|, \|\xi_n^2\|\} \|\xi_n^1 - \xi_n^2\|. \quad (11)$$

Since $f(x, \cdot)$ is continuous, we have that $f(x, \cdot)$ is bounded on the compact set Ξ_R , which implies that $|f(x, \xi_n^1) - f(x, \xi_n^2)|$ is bounded too. To ensure that (11) holds for all n , $\{\xi_n^1\}_{n=1}^\infty$ and $\{\xi_n^2\}_{n=1}^\infty$ must converge to a common point, denoted by ξ^* . Then, for sufficiently large n , we have $\|\xi_n^1 - \xi^*\| \leq \delta(\xi^*)$ and $\|\xi_n^2 - \xi^*\| \leq \delta(\xi^*)$. This and (11) contradict with (9a) in Lemma 3.4 with respect to $\xi^* \in \Xi$. Therefore, (10) holds.

For the second claim, since $\xi^1, \xi^2 \in \Xi_R$, $\|\xi^1 - \xi^2\| \leq 2\sqrt{s}R$. We know from (9a) in Lemma 3.4 that $\delta(\xi)$ can be chosen as $\kappa(\xi)/2\theta$. From Assumption 3.5, we know that for sufficiently large R , $\kappa(\xi) \geq 1/(\sqrt{s}R)$ for any $\xi \in \Xi_R$, and $\delta(\xi)$ can be chosen as $1/(2\theta\sqrt{s}R)$ for any $\xi \in \Xi_R$. Then, we can select a sequence of points $\xi^1 + \lambda_1(\xi^2 - \xi^1)$, $\xi^1 + \lambda_2(\xi^2 - \xi^1)$, \dots , $\xi^1 + \lambda_N(\xi^2 - \xi^1)$ contained in Ξ_R with $0 < \lambda_1 < \lambda_2 < \dots < \lambda_N < 1$

and $N \leq 4\theta sR^2$, such that

$$\begin{aligned}\|\xi^1 - (\xi^1 + \lambda_1(\xi^2 - \xi^1))\| &= \lambda_1 \|\xi^2 - \xi^1\| \leq \frac{1}{2\theta\sqrt{s}R}, \\ \|\xi^1 + \lambda_{i+1}(\xi^2 - \xi^1) - (\xi^1 + \lambda_i(\xi^2 - \xi^1))\| &= (\lambda_{i+1} - \lambda_i) \|\xi^2 - \xi^1\| \leq \frac{1}{2\theta\sqrt{s}R}, \\ \|\xi^2 - (\xi^1 + \lambda_N(\xi^2 - \xi^1))\| &= (1 - \lambda_N) \|\xi^2 - \xi^1\| \leq \frac{1}{2\theta\sqrt{s}R},\end{aligned}$$

for $i = 1 \dots, N - 1$. Therefore, according to Lemma 3.4, we can deduce that

$$\begin{aligned}|f(x, \xi^1) - f(x, \xi^2)| &\leq L_1 s^2 R^4 \|\xi^1 - (\xi^1 + \lambda_1(\xi^2 - \xi^1))\| \\ &\quad + \sum_{i=1}^{N-1} L_1 s^2 R^4 \|\xi^1 + \lambda_{i+1}(\xi^2 - \xi^1) - (\xi^1 + \lambda_i(\xi^2 - \xi^1))\| \\ &\quad + L_1 s^2 R^4 \|\xi^2 - (\xi^1 + \lambda_N(\xi^2 - \xi^1))\| \\ &= L_1 s^2 R^4 \|\xi^1 - \xi^2\|,\end{aligned}$$

which completes the proof. \square

On the other hand, the following conclusion can be straightforward derived from Lemma 3.4 and Assumption 3.5:

Lemma 3.7. *Suppose that Assumptions 3.2 and 3.5 hold. Then for any $\xi \in \Xi_R$ and sufficiently large R , we have*

$$\begin{aligned}|f(x_1, \xi) - f(x_2, \xi)| &\leq L_2 s R^2 \max\{1, \|\xi\|\}^3 \|x_1 - x_2\|, \\ |f(x, \xi)| &\leq L_3 s R^2 \max\{1, \|\xi\|\}^3,\end{aligned}$$

where L_2 and L_3 are defined in Lemma 3.4.

With the above conclusions, we can obtain the following quantitative stability results.

Theorem 3.8. *Suppose that Assumptions 3.2, 3.5 hold, $P, Q \in \mathcal{P}_6(\Xi)$ and Ξ is a polyhedral set. Then, there exist constants $\delta > 0$, positive integer k and $L_4 > 0$ which only depends on P , such that*

$$\begin{aligned}|v(P) - v(Q)| &\leq L_4 \zeta_{5,ph_k}(P, Q)^{1/5} (1 + \zeta_6(P, Q)), \\ \emptyset \neq S(Q) &\subseteq S(P) + \Psi_P(L_4 \zeta_{5,ph_k}(P, Q)^{1/5} (1 + \zeta_6(P, Q)))\mathbb{B},\end{aligned}$$

when $\zeta_{5,ph_k}(P, Q) < \delta$.

Proof. Since $\Xi = \Xi_R \cup \Xi_R^c$, we have

$$\begin{aligned}
|v(P) - v(Q)| &= \left| \inf_{x \in X} \mathbb{E}_P[f(x, \xi)] - \inf_{x \in X} \mathbb{E}_Q[f(x, \xi)] \right| \\
&\leq \sup_{x \in X} |\mathbb{E}_P[f(x, \xi)] - \mathbb{E}_Q[f(x, \xi)]| \\
&\leq \sup_{x \in X} \left(\left| \int_{\Xi_R} f(x, \xi)(P - Q)(d\xi) \right| + \left| \int_{\Xi_R^c} f(x, \xi)(P - Q)(d\xi) \right| \right) \\
&\leq \sup_{x \in X} \left| \int_{\Xi_R} f(x, \xi)(P - Q)(d\xi) \right| + \sup_{x \in X} \left| \int_{\Xi_R^c} f(x, \xi)(P - Q)(d\xi) \right|. \quad (12)
\end{aligned}$$

As Ξ is polyhedral, Ξ_R is also a polyhedral set. For sufficiently large R , we know from Lemma 3.6 that

$$\left| \frac{f(x, \xi_1)}{L_1 s^2 R^4} - \frac{f(x, \xi_2)}{L_1 s^2 R^4} \right| \leq \max\{1, \|\xi_1\|, \|\xi_2\|\}^4 \|\xi_1 - \xi_2\|.$$

Thus, the first term at the right-hand side of (12) can be bounded as follows

$$\begin{aligned}
\sup_{x \in X} \left| \int_{\Xi_R} f(x, \xi)(P - Q)(d\xi) \right| &= L_1 s^2 R^4 \sup_{x \in X} \left| \int_{\Xi_R} \frac{f(x, \xi)}{L_1 s^2 R^4} (P - Q)(d\xi) \right| \\
&\leq L_1 s^2 R^4 \zeta_{5, ph_k}(P, Q),
\end{aligned}$$

where k is the number of faces of the polyhedral set Ξ_R .

For the second term at the right-hand side of (12), we have

$$\begin{aligned}
\sup_{x \in X} \left| \int_{\Xi_R^c} f(x, \xi)(P - Q)(d\xi) \right| &\leq \int_{\Xi_R^c} |f(x, \xi)| (P + Q)(d\xi) \\
&\leq \int_{\Xi_R^c} L_3 \|\xi\|^2 \max\{1, \|\xi\|\}^3 (P + Q)(d\xi) \\
&= \int_{\Xi_R^c} L_3 \|\xi\|^5 (P + Q)(d\xi).
\end{aligned}$$

The last inequality is due to Assumption 3.5. For $\|\cdot\|^p$ ($p \geq 1$), we have the following fact:

$$\begin{aligned}
|\|\xi_1\|^p - \|\xi_2\|^p| &\leq p \max\{\|\xi_1\|^{p-1}, \|\xi_2\|^{p-1}\} \|\xi_1 - \xi_2\| \\
&\leq p \max\{1, \|\xi_1\|^{p-1}, \|\xi_2\|^{p-1}\} \|\xi_1 - \xi_2\|
\end{aligned} \quad (13)$$

for any $\xi_1, \xi_2 \in \Xi$, which implies

$$\frac{\|\cdot\|^p}{p} \in \mathcal{G}_{FM_p}. \quad (14)$$

Therefore, we have

$$\begin{aligned}
\int_{\Xi_R^c} \|\xi\|^5 (P + Q)(d\xi) &\leq \frac{1}{\sqrt{s}R} \int_{\Xi_R^c} \|\xi\|^6 (P + Q)(d\xi) \\
&\leq \frac{1}{\sqrt{s}R} (\mathbb{E}_P[\|\xi\|^6] + \mathbb{E}_Q[\|\xi\|^6]) \\
&\leq \frac{1}{\sqrt{s}R} (2\mathbb{E}_P[\|\xi\|^6] + 6\zeta_6(P, Q)).
\end{aligned}$$

The last inequality is due to

$$\left| \mathbb{E}_Q[\|\xi\|^6] - \mathbb{E}_P[\|\xi\|^6] \right| \leq 6\zeta_6(P, Q),$$

which can be derived from (13) and (14). Thus we obtain

$$\sup_{x \in X} \left| \int_{\Xi_R^c} f(x, \xi)(P - Q)(d\xi) \right| \leq \frac{1}{\sqrt{s}R} (2\mathbb{E}_P[\|\xi\|^6] + 6\zeta_6(P, Q)).$$

Notice that $P \in \mathcal{P}_6(\Xi)$, we have $\mathbb{E}_P[\|\xi\|^6] < +\infty$. Let $R = \zeta_{5,ph_k}(P, Q)^{-1/5} > (1/\delta)^{1/5}$, which can be sufficiently large when δ is selected sufficiently small.

Combining the above two parts of estimations together, we finally deduce that

$$\begin{aligned}
|v(P) - v(Q)| &\leq L_1 s^2 \zeta_{5,ph_k}(P, Q)^{-4/5} \zeta_{5,ph_k}(P, Q) + \sqrt{s} \zeta_{5,ph_k}(P, Q)^{1/5} (2\mathbb{E}_P[\|\xi\|^6] + 6\zeta_6(P, Q)) \\
&\leq \left(L_1 s^2 + 2\sqrt{s} \mathbb{E}_P[\|\xi\|^6] + 6\sqrt{s} \right) \zeta_{5,ph_k}(P, Q)^{1/5} (1 + \zeta_6(P, Q)).
\end{aligned}$$

Letting $L_4 = L_1 s^2 + 2\sqrt{s} \mathbb{E}_P[\|\xi\|^6] + 6\sqrt{s}$ completes the proof of the first claim.

For the inclusion of optimal solution sets, we can show it based on the first assertion, by using a method similar to that in [9], which is thus omitted here. \square

In the next theorem, we establish the quantitative stability conclusion by adopting the total variation metric. To this end, we denote $M_R = \{\xi \in \Xi : \|\xi\| \leq R\}$ and its complementary set by $M_R^c = \Xi \setminus M_R$. It is easy to verify that Lemma 3.7 holds for any $\xi \in M_R$ when R is sufficiently large.

Theorem 3.9. *Let Assumptions 3.2, 3.5 hold and $P, Q \in \mathcal{P}_6(\Xi)$. Then, there exist constants $\bar{\delta} > 0$ and $\bar{L}_4 > 0$, such that*

$$\begin{aligned}
|v(P) - v(Q)| &\leq \bar{L}_4 \mathbb{D}_{TV}(P, Q)^{1/6} (1 + \zeta_6(P, Q)), \\
\emptyset \neq S(Q) &\subseteq S(P) + \Psi_P(\bar{L}_4 \mathbb{D}_{TV}(P, Q)^{1/6} (1 + \zeta_6(P, Q)))\mathbb{B},
\end{aligned}$$

when $\mathbb{D}_{TV}(P, Q) < \bar{\delta}$.

Proof. Analogous to the proof of Lemma 3.7, we can show under Assumption 3.5 that

$$|f(x, \xi)| \leq L_3 R^2 \max\{1, \|\xi\|\}^3 \leq L_3 R^5 \tag{15}$$

for sufficiently large R and any $\xi \in M_R$. Then we have

$$|v(P) - v(Q)| \leq \sup_{x \in X} \left| \int_{M_R} f(x, \xi)(P - Q)(d\xi) \right| + \sup_{x \in X} \left| \int_{M_R^c} f(x, \xi)(P - Q)(d\xi) \right|. \quad (16)$$

On one hand, we know from (15) that

$$\sup_{x \in X} \left| \int_{M_R} f(x, \xi)(P - Q)(d\xi) \right| \leq L_3 R^5 \mathbb{D}_{TV}(P, Q).$$

On the other hand, according to Assumption 3.5 and (9c), we have

$$\begin{aligned} \sup_{x \in X} \left| \int_{M_R^c} f(x, \xi)(P - Q)(d\xi) \right| &\leq \int_{M_R^c} L_3 \|\xi\|^2 \max\{1, \|\xi\|^3\} (P + Q)(d\xi) \\ &= L_3 \int_{M_R^c} \|\xi\|^5 (P + Q)(d\xi) \\ &\leq \frac{L_3}{R} (2\mathbb{E}_P[\|\xi\|^6] + 6\zeta_6(P, Q)). \end{aligned}$$

The above two inequalities mean that

$$|v(P) - v(Q)| \leq L_3 R^5 \mathbb{D}_{TV}(P, Q) + \frac{L_3}{R} (2\mathbb{E}_P[\|\xi\|^6] + 6\zeta_6(P, Q)).$$

Let $R = \mathbb{D}_{TV}(P, Q)^{-1/6} > (1/\bar{\delta})^{1/6}$. Then we have

$$\begin{aligned} |v(P) - v(Q)| &\leq L_3 \mathbb{D}_{TV}(P, Q)^{1/6} + \mathbb{D}_{TV}(P, Q)^{1/6} (2\mathbb{E}_P[\|\xi\|^6] + 6\zeta_6(P, Q)) \\ &\leq L_3 (1 + 2\mathbb{E}_P[\|\xi\|^6] + 6) \mathbb{D}_{TV}(P, Q)^{1/6} (1 + \zeta_6(P, Q)). \end{aligned}$$

We complete the proof by letting $\bar{L}_4 = L_3 (1 + 2\mathbb{E}_P[\|\xi\|^6] + 6)$. \square

Theorem 3.8 and Theorem 3.9 present the quantitative stability results under different pseudo metrics. The necessity for both theorems is that, as far as we know, there are no definite conclusions about the relationship between $\mathbb{D}_{TV}(P, Q)$ and $\zeta_{5,ph_k}(P, Q)$. For Theorem 3.8, it requires that the support set is polyhedral, which may be a little restrictive. Naturally, this condition fulfils when $\Xi = \mathbb{R}^s$. Despite of this, Theorem 3.8 is important and meaningful because the rate of convergence for the discrete approximation can be easily derived under this circumstance, see [8].

At the end of this section, we look at some special cases and then derive stronger conclusions. We first consider $\kappa(\xi) \geq \kappa > 0$.

Theorem 3.10. *Let Assumption 3.2 hold and $\kappa(\xi) \geq \kappa > 0$.*

(i) *If Ξ is convex and $P, Q \in \mathcal{P}_5(\Xi)$, there exists a positive constant $L_5 > 0$, such that*

$$\begin{aligned} |v(P) - v(Q)| &\leq L_5 \zeta_5(P, Q), \\ \emptyset \neq S(Q) &\subseteq S(P) + \Psi_P(L_5 \zeta_5(P, Q)) \mathbb{B}; \end{aligned}$$

(ii) If $P, Q \in \mathcal{P}_4(\Xi)$, there exists a positive constant $\bar{L}_5 > 0$ which only depends on P , such that

$$|v(P) - v(Q)| \leq \bar{L}_5 \left(\mathbb{D}_{TV}(P, Q)^{1/4} (1 + \zeta_4(P, Q)) \right),$$

$$\emptyset \neq S(Q) \subseteq S(P) + \Psi_P \left(\bar{L}_5 \left(\mathbb{D}_{TV}(P, Q)^{1/4} (1 + \zeta_4(P, Q)) \right) \right) \mathbb{B}.$$

Proof. (i) Without loss of generality, we assume that $\kappa \leq 1$ in the following proof. Note that

$$|v(P) - v(Q)| \leq \sup_{x \in X} \left| \int_{\Xi} f(x, \xi)(P - Q)(d\xi) \right|.$$

According to Lemma 3.4 and Lemma 3.6, we have

$$|f(x, \xi_1) - f(x, \xi_2)| \leq \frac{L_1}{\kappa^4} \max\{1, \|\xi_1\|, \|\xi_2\|\}^4 \|\xi_1 - \xi_2\|$$

for any $\xi_1, \xi_2 \in \Xi$. Then, we obtain

$$|v(P) - v(Q)| \leq \frac{L_1}{\kappa^4} \zeta_5(P, Q).$$

Letting $L_5 := \frac{L_1}{\kappa^4}$ establishes the assertion.

(ii) Define the same M_R as that in Theorem 3.8, we have

$$|v(P) - v(Q)| \leq \sup_{x \in X} \left| \int_{M_R} f(x, \xi)(P - Q)(d\xi) \right| + \sup_{x \in X} \left| \int_{M_R^c} f(x, \xi)(P - Q)(d\xi) \right|.$$

Similarly, we have

$$\sup_{x \in X} \left| \int_{M_R} f(x, \xi)(P - Q)(d\xi) \right| \leq \frac{L_3 R^3}{\kappa^2} \mathbb{D}_{TV}(P, Q)$$

and

$$\begin{aligned} \sup_{x \in X} \left| \int_{M_R^c} f(x, \xi)(P - Q)(d\xi) \right| &\leq \int_{M_R^c} \frac{L_3}{\kappa^2} \max\{1, \|\xi\|^3\} (P + Q)(d\xi) \\ &= \frac{L_3}{\kappa^2} \int_{M_R^c} \|\xi\|^3 (P + Q)(d\xi) \\ &\leq \frac{L_3}{\kappa^2 R} (2\mathbb{E}_P[\|\xi\|^4] + 4\zeta_4(P, Q)). \end{aligned}$$

Hence, let $R = \mathbb{D}_{TV}(P, Q)^{-1/4}$, then we obtain

$$\begin{aligned} |v(P) - v(Q)| &\leq \frac{L_3 R^3}{\kappa^2} \mathbb{D}_{TV}(P, Q) + \frac{L_3}{\kappa^2 R} (2\mathbb{E}_P[\|\xi\|^4] + 4\zeta_4(P, Q)) \\ &\leq \frac{L_3}{\kappa^2} \mathbb{D}_{TV}(P, Q)^{1/4} (2\mathbb{E}_P[\|\xi\|^4] + 4\zeta_4(P, Q)) \\ &\leq \frac{L_3}{\kappa^2} (2\mathbb{E}_P[\|\xi\|^4] + 4) \mathbb{D}_{TV}(P, Q)^{1/4} (1 + \zeta_4(P, Q)). \end{aligned}$$

This completes the proof by letting $\bar{L}_5 = \frac{L_3}{\kappa^2} (2\mathbb{E}_P[\|\xi\|^4] + 4)$. \square

Moreover, if the support set Ξ is a compact set, we can derive the following conclusion.

Corollary 3.11. *Let Assumption 3.2 hold and Ξ be a compact set.*

(i) *If $P, Q \in \mathcal{P}_5(\Xi)$, then there exists a positive constant $L_6 > 0$ such that*

$$\begin{aligned} |v(P) - v(Q)| &\leq L_6 \zeta_5(P, Q), \\ \emptyset \neq S(Q) &\subseteq S(P) + \Psi_P(L_6 \zeta_5(P, Q)) \mathbb{B}; \end{aligned}$$

(ii) *If $P, Q \in \mathcal{P}(\Xi)$, then there exists a positive constant $\bar{L}_6 > 0$ such that*

$$\begin{aligned} |v(P) - v(Q)| &\leq \bar{L}_6 \mathbb{D}_{TV}(P, Q), \\ \emptyset \neq S(Q) &\subseteq S(P) + \Psi_P(\bar{L}_6 \mathbb{D}_{TV}(P, Q)) \mathbb{B}. \end{aligned}$$

Proof. (i) Since Ξ is a compact set, we can prove by using a similar proof as that in Lemma 3.6 that, there exists a positive scalar L_6 such that

$$|f(x, \xi^1) - f(x, \xi^2)| \leq L_6 \max \{1, \|\xi^1\|, \|\xi^2\|\} \|\xi^1 - \xi^2\|,$$

for any $\xi^1, \xi^2 \in \Xi$. Thus, we can directly obtain the assertion by the definition of $\zeta_5(P, Q)$.

(ii) It can be derived directly from Lemma 3.3, (9c) in Lemma 3.4 and the definition of total variation metric. \square

Remark 3. From different perspectives, Theorem 3.8, Theorem 3.9, Theorem 3.10 and Corollary 3.11 establish the quantitative stability about two-stage stochastic programs with full random recourse. Most existing results are only applicable to the case with the Lipschitz continuous objective function (see comprehensive review [9]). Römisch and Wets discussed the full random recourse model in [10]. However, they adopt an abstract probability matrix by specifying the pseudo metric generator straightforward. Han and Chen [21] considered two-stage stochastic programs with full random recourse and obtained the locally Lipschitz continuity of the objective function under some uniform boundedness assumptions, which are rather restrictive. Our above stability results overcome these limitations under moderate conditions and can be applied to the convergence or approximation analysis, as we will show in the next section.

4. Empirical approximation

In this section, we adopt the method in the pioneering work [8] to study the empirical approximation of problem (1)-(3). The discrete approximation of stochastic optimizations is important and useful for the practical solution and estimation. Specifically, we will mainly estimate the rate of convergence of the SAA method based on the quantitative stability results established in section 3.

Assume that we have a sequence of independent identically distributed (i.i.d.) samples, denoted by $\xi^1, \xi^2, \dots, \xi^N, \dots$, having the common probability distribution P . For any positive integer N , we can construct the discrete approximation distribution P_N as follows:

$$P_N = \frac{1}{N} \sum_{i=1}^N \delta_{\xi^i}.$$

Then, the SAA counterpart of the two-stage stochastic linear program with full random recourse (1)-(3) can be described as

$$\min_{x \in X} \mathbb{E}_{P_N}[f(x, \xi)] = \frac{1}{N} \sum_{i=1}^N f(x, \xi^i).$$

To ensure the measurability of $\mathbb{D}_{\mathcal{G}}(P_N, P)$, the so-called permissibility is put forward.

Definition 4.1 (permissibility, [8]). \mathcal{G} is called permissible for some $P \in \mathcal{P}(\Xi)$ if there exists a countable subset $\mathcal{G}_0 \subseteq \mathcal{G}$, such that for each $g \in \mathcal{G}$, there exists a sequence of measurable functions $\{g_n\}_{n=1}^{\infty} \subseteq \mathcal{G}_0$ converging pointwise to g and such that $\mathbb{E}_P[g_n(\xi)] \rightarrow \mathbb{E}_P[g(\xi)]$ as n tends to infinity.

The concept of permissibility leads to

$$\begin{aligned} \mathbb{D}_{\mathcal{G}}(P_N, P) &= \sup_{g \in \mathcal{G}} |\mathbb{E}_{P_N}[g(\xi)] - \mathbb{E}_P[g(\xi)]| \\ &= \sup_{g \in \mathcal{G}_0} |\mathbb{E}_{P_N}[g(\xi)] - \mathbb{E}_P[g(\xi)]|, \end{aligned}$$

that is, we can reduce the generator of the pseudo metric to a countable subset. This will guarantee the measurability of $\mathbb{D}_{\mathcal{G}}(P_N, P)$ when we take the supremum in the pseudo metric. We know from [8, Example 4.4] that the class $\mathcal{G} = \{f(x, \xi) : x \in X, f(x, \xi) \text{ is defined in (2)}\}$ is permissible when $P \in \mathcal{P}_3(\Xi)$ and Ξ is a compact set. In practice, without loss of generality, we can select $\mathcal{G}_0 := \{f(x, \xi) : x \in X_c, f(x, \xi) \text{ is defined in (2)}\}$ where X_c is a countable and dense subset contained in X .

Moreover, we need to define the so-called covering number and bracketing number. Let $L_p(\Xi, P)$ be the collection of all measurable functions that have finite p -th order absolute moments with respect to P , and it equips with the usual norm $\|F(\xi)\|_p = \mathbb{E}_P[|F(\xi)|^p]^{1/p}$ for any $F(\xi) \in L_p(\Xi, P)$. Then the covering number and bracketing number can be defined as follows.

Definition 4.2 (covering number and bracketing number [8]). Let $\mathcal{F}(\Xi) \subseteq L_p(\Xi, P)$ denote a class of measurable functions.

- (i) The covering number $N(\epsilon, \mathcal{F}(\Xi), L_p(\Xi, P))$ is the minimal number of open balls $\{g \in L_p(\Xi, P) : \|g - h\|_p < \epsilon\}$ needed to cover $\mathcal{F}(\Xi)$;
- (ii) The bracketing number $N_{[\cdot]}(\epsilon, \mathcal{F}(\Xi), L_p(\Xi, P))$ is the minimal number of ϵ -brackets needed to cover $\mathcal{F}(\Xi)$, here the ϵ -bracket is defined as $[h_1, h_2] := \{g \in L_p(\Xi, P) : h_1(\xi) \leq g(\xi) \leq h_2(\xi) \text{ for } P\text{-a.e.}\}$ for $h_1, h_2 \in L_p(\Xi, P)$ and $\|h_1 - h_2\|_p < \epsilon$.

We can directly see from the above definitions that $N(\epsilon, \mathcal{F}(\Xi), L_p(\Xi, P)) \leq N_{[\cdot]}(2\epsilon, \mathcal{F}(\Xi), L_p(\Xi, P))$. To describe the rate of convergence of $v(P_N)$ tending to $v(P)$, we introduce the so-called Ky Fan distance.

Definition 4.3 (Ky Fan distance). The Ky Fan distance between two real-valued random variables \mathcal{X} and \mathcal{Y} on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is

$$\mathbb{K}(\mathcal{X}, \mathcal{Y}) := \inf\{\eta \geq 0 : \mathbb{P}(|\mathcal{X} - \mathcal{Y}| > \eta) \leq \eta\}.$$

For notational convenience and without loss of generality, we always assume that Ξ is a compact set in the following demonstration. For the noncompact case, we know from Theorem 3.10 that the similar conclusion can be derived under the assumption $\kappa(\xi) \geq \kappa$. With the Ky Fan distance and the quantitative stability results in Corollary 3.11, we can directly obtain the following conclusions:

Proposition 4.4. *Let Assumption 3.2 hold, Ξ be a compact set and $P \in \mathcal{P}_5(\Xi)$. Then we have*

$$\begin{aligned} \mathbb{K}(v(P_N), v(P)) &\leq \max\{1, L_6\} \mathbb{K}(\zeta_5(P_N, P), 0), \\ \mathbb{K}(d(S(P_N), S(P)), 0) &\leq \Psi_P(\max\{1, L_6\} \mathbb{K}(\zeta_5(P_N, P), 0)), \end{aligned}$$

where L_6 is defined in Corollary 3.11.

Proof. A similar proof can be found in [8, Proposition 4.1]. We give a brief proof here for completeness. Denote $\epsilon_N = \mathbb{K}(\zeta_5(P_N, P), 0)$, we have from Corollary 3.11 that

$$\mathbb{P}(|v(P_N) - v(P)| > L_6 \epsilon_N) \leq \mathbb{P}(\zeta_5(P_N, P) > \epsilon_N) \leq \epsilon_N.$$

The last inequality is due to the definition of the Ky Fan distance. Then, we obtain

$$\mathbb{P}(|v(P_N) - v(P)| > \max\{1, L_6\} \epsilon_N) \leq \max\{1, L_6\} \epsilon_N,$$

which implies $\mathbb{K}(v(P_N), v(P)) \leq \max\{1, L_6\} \epsilon_N$.

Similarly, for the second claim, it is known from Corollary 3.11 that

$$\begin{aligned} \mathbb{P}(d(S(P_N), S(P)) > \Psi_P(L_6 \epsilon_N)) &\leq \mathbb{P}(\Psi_P(L_6 \zeta_5(P_N, P)) > \Psi_P(L_6 \epsilon_N)) \\ &= \mathbb{P}(\zeta_5(P_N, P) > \epsilon_N) \\ &\leq \epsilon_N \leq \Psi_P(\epsilon_N). \end{aligned}$$

Therefore, we have

$$\mathbb{K}(d(S(P_N), S(P)), 0) \leq \max\{\Psi_P(\epsilon_N), \Psi_P(L_6 \epsilon_N)\} = \Psi_P(\max\{1, L_6\} \epsilon_N),$$

which completes the proof. \square

Proposition 4.5. *Let Assumption 3.2 hold, Ξ be a compact set and $P \in \mathcal{P}_5(\Xi)$. Then, the following rates of convergence hold:*

$$\begin{aligned}\mathbb{K}(v(P_N), v(P)) &= O((\log N)^{1/2} \cdot N^{-1/2}), \\ \mathbb{K}(d(S(P_N), S(P)), 0) &= O(\Psi_P((\log N)^{1/2} \cdot N^{-1/2})).\end{aligned}$$

Proof. Denote $\mathcal{F}(\Xi) = \{f(x, \xi) : x \in X, f(x, \xi) \text{ is defined in (2)}\}$. Thanks to [8, Example 4.4], we know that

$$N_{[\cdot]}(2\epsilon \|F(\xi)\|_1, \mathcal{F}(\Xi), L_1(\Xi, P)) \leq N(\epsilon, X, \mathbb{R}^n) \leq C\epsilon^{-n},$$

where $F(\xi) = \frac{L_3}{\kappa^2} \{1, \|\xi\|^3\}$ and κ is the minimum of $\kappa(\xi)$ over Ξ . Then according to [8, Proposition 4.2], we straightforward obtain the conclusion. \square

Remark 4. Interested readers can refer to [9] for more comprehensive review of the techniques used here. It is noteworthy that we can also discuss the qualitative convergence by adopting some classical theories, such as the uniformly large number theorem and the large deviation theorem. Specifically, under assumptions that the objective function $f(x, \xi)$ is Lipschitz continuous with respect to x and the Lipschitz modulus of the moment generating function is uniformly bounded near zero, the exponential rate of convergence can be established, see for example [23]. However, the relevant coefficients in these results usually depend on the accuracy parameter ϵ . Therefore, we can not deduce from these studies the explicit expressions like those in Proposition 4.5.

5. Concluding remarks

Under mild assumptions, we established the continuity and boundedness of optimal values and optimal solutions, respectively, of parametric linear programmes. On the basis of these results, we have obtained the locally Lipschitz continuity and upper bound estimation of the objective function of two-stage stochastic linear programs with full random recourse, which leads to different types of quantitative stability results under different conditions. Finally, we considered the empirical approximation of the two-stage stochastic linear programming with full random recourse and derived the rate of convergence of the SAA method for solving it.

Our results improve the existing stability results for two-stage stochastic linear programs with full random recourse. We rely on moderate assumptions to derive the quantitative stability results with commonly used pseudo metrics. Our assumptions are reasonable and easy to understand. What's more important, our quantitative results can be used to obtain the rate of convergence of the corresponding SAA method.

There are still several issues to settle. For example, how to extend our results to the multistage case, how to generalize our results to the risk-averse situation, how to design efficient algorithms for these models based on the obtained quantitative stability results. They are all left for future study.

Funding

This work was supported by the National Natural Science Foundation of China (Grant Nos. 11571270 and 71371152).

References

- [1] Gülten S, Ruszczyński A. Two-stage portfolio optimization with higher-order conditional measures of risk. *Annals of Operations Research*. 2015;229(1):409–427.
- [2] Dentcheva D, Ruszczyński A. Risk-averse portfolio optimization via stochastic dominance constraints. In: *Handbook of financial econometrics and statistics*. Springer; 2015. p. 2281–2302.
- [3] Chen X, Wets RJB, Zhang Y. Stochastic variational inequalities: residual minimization smoothing sample average approximations. *SIAM Journal on Optimization*. 2012; 22(2):649–673.
- [4] Liu Y, Xu H, Yang SJS, et al. Distributionally robust equilibrium for continuous games: Nash and stackelberg models. *European Journal of Operational Research*. 2018; 265(2):631–643.
- [5] Chen X, Sun H, Xu H. Discrete approximation of two-stage stochastic and distributionally robust linear complementarity problems [<https://arxiv.org/pdf/1706.06847.pdf>]; 2017.
- [6] Ruszczyński A, Shapiro A. *Stochastic programming*. Vol. 10. Elsevier Amsterdam; 2003.
- [7] Shapiro A, Dentcheva D, Ruszczyński A. *Lectures on stochastic programming: modeling and theory*. SIAM; 2009.
- [8] Rachev ST, Römisch W. Quantitative stability in stochastic programming: The method of probability metrics. *Mathematics of Operations Research*. 2002;27(4):792–818.
- [9] Römisch W. Stability of stochastic programming problems. *Handbooks in operations research and management science*. 2003;10:483–554.
- [10] Römisch W, Wets RB. Stability of ε -approximate solutions to convex stochastic programs. *SIAM Journal on Optimization*. 2007;18(3):961–979.
- [11] Dentcheva D, Römisch W. Stability and sensitivity of stochastic dominance constrained optimization models. *SIAM Journal on Optimization*. 2013;23(3):1672–1688.
- [12] Liu Y, Xu H. Stability analysis of stochastic programs with second order dominance constraints. *Mathematical Programming*. 2013;142(1-2):435–460.
- [13] Kall P. On approximations and stability in stochastic programming. *Parametric optimization and related topics*. 1987;:387–407.
- [14] Robinson SM, Wets RJB. Stability in two-stage stochastic programming. *SIAM Journal on Control and Optimization*. 1987;25(6):1409–1416.
- [15] Römisch W, Schultz R. Distribution sensitivity in stochastic programming. *Mathematical Programming*. 1991;50(1):197–226.
- [16] Römisch W, Schultz R. Stability analysis for stochastic programs. *Annals of Operations Research*. 1991;30(1):241–266.
- [17] Römisch W, Schultz R. Lipschitz stability for stochastic programs with complete recourse. *SIAM Journal on Optimization*. 1996;6(2):531–547.
- [18] Dentcheva D, Römisch W. Differential stability of two-stage stochastic programs. *SIAM Journal on Optimization*. 2000;11(1):87–112.
- [19] Liu Y, Römisch W, Xu H. Quantitative stability analysis of stochastic generalized equations. *SIAM Journal on Optimization*. 2014;24(1):467–497.
- [20] Rockafellar RT, Wets RJB. *Variational analysis*. Springer Science & Business Media; 2009.
- [21] Han Y, Chen Z. Quantitative stability of full random two-stage stochastic programs with recourse. *Optimization Letters*. 2015;9(6):1075–1090.
- [22] Zhao C, Guan Y. Data-driven risk-averse two-stage stochastic program with ζ -structure probability metrics [http://www.optimization-online.org/DB_FILE/2015/07/5014.pdf];

- 2015.
- [23] Shapiro A, Xu H. Stochastic mathematical programs with equilibrium constraints, modelling and sample average approximation. *Optimization*. 2008;57(3):395–418.
 - [24] Rachev ST, Stoyanov SV, Fabozzi FJ. A probability metrics approach to financial risk measures. John Wiley & Sons; 2011.
 - [25] Zhang J, Xu H, Zhang L. Quantitative stability analysis for distributionally robust optimization with moment constraints. *SIAM Journal on Optimization*. 2016;26(3):1855–1882.
 - [26] Zhang J, Xu H, Zhang L. Quantitative stability analysis of stochastic quasi-variational inequality problems and applications. *Mathematical Programming*. 2017;165(1):433–470.