Stability of ε-approximate solutions to convex stochastic programs

Werner Römisch
Humboldt-University Berlin
Institute of Mathematics
10099 Berlin, Germany

and

Roger J-B Wets
University of California at Davis
Department of Mathematics
Davis, CA 95616-8633, U.S.A.

Abstract

An analysis of convex stochastic programs is provided if the underlying probability distribution is subjected to (small) perturbations. It is shown, in particular, that ε-approximate solution sets of convex stochastic programs behave Lipschitz continuous with respect to certain distances of probability distributions that are generated by the relevant integrands. It is shown that these results apply to linear two-stage stochastic programs with random recourse. Consequences are discussed on associating Fortet-Mourier metrics to two-stage models and on the asymptotic behavior of empirical estimates of such models, respectively.

Key Words: Stochastic programming, quantitative stability, ε-approximate solutions, probability metrics, two-stage models, random recourse

1991 Mathematics Subject Classification: 90C15, 90C31

1 Introduction

Stochastic programming deals with models for optimization problems under (stochastic) uncertainty that require a decision on the basis of probabilistic information about random data. Typically, deterministic equivalents of such models are, finite- or infinite dimensional, nonlinear programs depending on the properties of the distribution of the random components of the problems. Their solutions depend on the probability distribution of the random data via certain expectation functionals. Many deterministic equivalents of stochastic programming models take the form

$$\min \ \mathbb{E}^P f_0(x) := \int_\Xi f_0(\xi, x)\, P(d\xi) : x \in X \tag{1}$$

where X a closed convex subset of $\mathbb{R}^m$, Ξ a closed subset of $\mathbb{R}^s$, P is a Borel probability measure on Ξ and $\mathbb{E}^P$ denotes expectation with respect to P. The function f_0 from $\mathbb{R}^m \times \Xi$ to $\overline{\mathbb{R}} = [-\infty, \infty]$ is a *convex random lsc (lower semicontinuous) function* [1] and, in particular, this means

[1]The concept of a random lsc function is due to Rockafellar [19] who introduced it in the context of the calculus of variations under the name of 'normal integrand.' Further properties of random lsc functions are set forth in [20, Chapter 14], [30].

- $(\xi, x) \mapsto f_0(\xi, x)$ is Borel measurable ,
- for all $\xi \in \Xi$, $f_0(\cdot, \xi)$ is lsc and convex.

It is part of the stochastic programming folklore, repeatedly observed in practice, that the solutions, or at least the approximating solutions, are quite robust with respect to reasonable perturbations of the probability distribution of the random components of the problem. In this paper, we substantiate this belief by focusing our analysis on the approximating solutions for which we are able to derive Lipschitz continuity without even requiring fixed (deterministic) recourse.
In the following, we denote by $\mathcal{P}(\Xi)$ the set of all Borel probability measures on Ξ and by $v(P)$, $S(P)$ and $S_\varepsilon(P)$ ($\varepsilon \geq 0$) the optimal value, the solution set and the set of ε-approximate solutions to (1), i.e.,

$$\begin{aligned} v(P) &:= \inf \mathbb{E}^P f_0 := \inf \{\mathbb{E}^P f_0(x) : x \in X\} \\ S(P) &:= \operatorname{argmin} \mathbb{E}^P f_0 := S_0(P), \\ S_\varepsilon(P) &:= \varepsilon\text{-argmin}\, \mathbb{E}^P f_0 := \{x \in X \,:\, \mathbb{E}^P f_0(x) \leq v(P) + \varepsilon\}. \end{aligned}$$

Since, in practice, the underlying probability distribution P is often not known precisely, the stability behavior of the stochastic program (1) when changing (perturbing, estimating, approximating) P is important. Here, stability refers to continuity properties of the optimal value function $v(.)$ and of the set-valued mapping $S_\varepsilon(.)$ at P, where both $v(.)$ and $S_\varepsilon(.)$ are regarded as mappings given on certain subset of $\mathcal{P}(\Xi)$ equipped with some probability (semi) metric.

Early work on stability of stochastic programs is reported in [9, 17, 24] and later in [1]. Quantitative stability of two-stage models was studied, e.g., in [22, 23, 26, 16]. A recent survey of stability results in stochastic programming is given in [21]. Most of the recent contributions to (quantitative) stability use the general framework and the results of [3, 12] and [20, Chapter 7J], respectively.

In the present paper, we take up an issue brought to the fore in [34, Section 4]. Since solutions derived, when actually solving (1), are usually ε-approximate solutions of an approximating problem where P has been replaced by an approximating measure Q, it is crucial to investigate the (quantitative) continuity properties of the (set-valued) mapping ε-argmin as a function of P, i.e., $P \mapsto S_\varepsilon(P)$, from $\mathcal{P}$ of probability measures to the space of closed convex subsets of $\mathbb{R}^m$.

Quantitative perturbation results for ε-approximate solutions in optimization are given in [4] and [20, Chapter 7J]. The corresponding estimates make use of the epi-distance between the objective functions of (1) and its perturbations. In our analysis, the corresponding subset $\mathcal{P}$ of probability measures is determined by satisfying certain moment conditions that are related to growth properties of the integrand f_0 with respect to ξ. The epi-distances of the objective functions can be bounded by some probability semi-metric of the form

$$d_{\mathcal{F}}(P, Q) = \sup \left\{ \left| \int_\Xi \psi(\xi) P(d\xi) - \int_\Xi \psi(\xi) Q(d\xi) \right| \,:\, \psi \in \mathcal{F} \right\} \tag{2}$$

where $\mathcal{F}$ is an appropriate class of measurable functions from Ξ to $\overline{\mathbb{R}}$ and P, Q are probability measures in $\mathcal{P}$. First, we show in Section 2 that classes of the form $\mathcal{F}_\rho = \{f_0(\cdot, x) : x \in X \cap \rho\mathbb{B}\}$, for some $\rho > 0$, and the corresponding distance $d_{\mathcal{F}_\rho}$ are suitable to derive the desired stability results.

In Section 3 we then provide characterizations of the function classes $\mathcal{F}_\rho$ for two-stage models with random recourse. While the continuity of the integrands f_0 with respect to ξ is well understood for fixed recourse [32], much less is known for random recourse. We deal with the following two cases: (i) full random recourse by imposing local Lipschitz continuity of the dual feasibility mapping and (ii) lower diagonal randomness of the recourse matrix. The latter situation occurs for multi-period two-stage models with random technology matrices. Based on these characterizations, we show that the distances $d_{\mathcal{F}_\rho}$ are bounded by Fortet-Mourier (type) metrics and that the metric entropy of $\mathcal{F}_\rho$ in terms of bracketing numbers is reasonably "small". In this way, we obtain new results on stability (Corollaries 3.5 and 4.3) and on the asymptotic behavior of nonparametric statistical estimates (Theorem 5.2) of random recourse models.

2 Quantitative Stability

Given the original probability measure P and a perturbation Q of P we will give quantitative estimates of the distance between $(v(Q), S_\varepsilon(Q))$ and $(v(P), S_\varepsilon(P))$ in terms of a probability metric of the type (2). Our analysis will be based on the general perturbation results for optimization models in [20, Section 7J].

Let us now introduce functions spaces and probability measures that are useful for characterizing classes of probability distributions such that the stochastic program (1) is well-defined and one can proceed with the perturbation analysis. We consider

$$\begin{aligned} \mathcal{F} &= \{f_0(\cdot, x) \,:\, x \in X\}, \\ \mathcal{P}_{\mathcal{F}} &= \{Q \in \mathcal{P}(\Xi) : \int_\Xi \inf_{x \in X \cap \rho\mathbb{B}} f_0(\xi, x) Q(d\xi) > -\infty \text{ , and} \\ &\qquad \sup_{x \in X \cap \rho\mathbb{B}} \int_\Xi f_0(\xi, x) Q(d\xi) < \infty \text{ , for all } \rho > 0\}, \end{aligned}$$

where $\mathbb{B}$ denotes the closed unit ball in $\mathbb{R}^m$. We note that the infimum function $\xi \mapsto \inf_{x \in X \cap \rho\mathbb{B}} f_0(\xi, x)$ is measurable for each $\rho > 0$ as f_0 is a random lsc function, cf. [20, Theorem 14.37].

For any $\rho > 0$ and probability measures P, $Q \in \mathcal{P}_{\mathcal{F}}$ we consider their $d_{\mathcal{F},\rho}$ distance defined by

$$d_{\mathcal{F},\rho}(P, Q) = \sup_{x \in X \cap \rho\mathbb{B}} |\mathbb{E}^P f_0(x) - \mathbb{E}^Q f_0(x)| \, .$$

Hence, $d_{\mathcal{F},\rho}$ is a distance of type (2) where the relevant class of functions is $\mathcal{F}_\rho = \{f_0(\cdot, x) : x \in X \cap \rho\mathbb{B}\}$. It is nonnegative, finite, symmetric and satisfies the triangle inequality, i.e., it is a semi-metric on $\mathcal{P}_{\mathcal{F}}$. In general, however, the class $\mathcal{F}_\rho$ will not be rich enough to guarantee $d_{\mathcal{F},\rho}(P,Q) = 0$ implies $P = Q$. A valuable consequence of the definition of the class $\mathcal{P}_{\mathcal{F}}$ is that, for any Q belonging to $\mathcal{P}_{\mathcal{F}}$, the function $x \mapsto I\!E^Q f_0(x) = \int_{\Xi} f_0(\xi, x)\, Q(d\xi)$ is lsc, by appealing to Fatou's lemma, and convex on $I\!R^m$.

Since our statements and proofs rely extensively on estimates for the epi-distance between (lsc) functions, we include a brief review of the relevant definitions and implications. Let $d_C(x) = d(x, C)$ denote the distance of a point to a non-empty closed set. The $\rho-$*distance* between two non-empty closed sets is by definition

$$d\!l_\rho(C, D) = \sup_{||x|| \leq \rho} |d_C(x) - d_D(x)|.$$

In fact, it is just a pseudo-distance from which one can build a metric on the hyperspace of closed sets, for example, by setting $d\!l(C,D) = \int_0^\infty d\!l_\rho(C,D) e^{-\rho}\, d\rho$. Estimates for the ρ-distance can be obtained by relying on a 'truncated' Pompeiu-Hausdorff type distance:

$$\hat{d\!l}_\rho(C,D) = \inf\Big\{\eta \geq 0 :\ C \cap \rho I\!B \subset D + \eta I\!B,\ \ D \cap \rho I\!B \subset C + \eta I\!B\Big\}.$$

Indeed one always has [20, Proposition 4.37(a)],

$$\hat{d\!l}_\rho(C_1, C_2) \leq d\!l_\rho(C_1, C_2) \leq \hat{d\!l}_{\rho'}(C_1, C_2)$$

for $\rho' \geq 2\rho + \max\{d_{C_1}(0), d_{C_2}(0)\}$. Our main result is stated in terms of this latter distance notion. If we let $\rho \to \infty$, we end up with $d\!l_\rho(C,D)$ and $\hat{d\!l}_\rho(C,D)$ tending to $d\!l_\infty(C,D)$, the *Pompeiu-Hausdorff* distance between the closed non-empty sets C and D, see [20, Corollary 4.38].

The distance between (lsc) functions is measured in terms of the distance between their epigraphs, so for $\rho > 0$,

$$d\!l_\rho(f,g) = d\!l_\rho(\operatorname{epi} f, \operatorname{epi} g), \qquad \hat{d\!l}_\rho(f,g) = \hat{d\!l}_\rho(\operatorname{epi} f, \operatorname{epi} g).$$

and $d\!l(f,g) = d\!l(\operatorname{epi} f, \operatorname{epi} g)$. However, since our sets are epigraphs (in $I\!R^{m+1}$), it is convenient to rely on the 'unit ball' to be $I\!B \times [-1,1]$, this brings us to an 'auxiliary' distance $\hat{d\!l}^+_\rho(f_1, f_2)$ defined as the infimum of all $\eta \geq 0$ such that for all $x \in \rho I\!B$,

$$\min_{B(x,\eta)} f_2 \leq \max\{\, f_1(x), -\rho\,\} + \eta \quad \min_{B(x,\eta)} f_1 \leq \max\{\, f_2(x), -\rho\,\} + \eta.$$

For lsc f_1, $f_2 : I\!R^n \to \overline{I\!R}$, not identically ∞, one has ([20, Theorem 7.61]),

$$\hat{d\!l}^+_{\rho/\sqrt{2}}(f_1, f_2) \leq \hat{d\!l}_\rho(f_1, f_2) \leq \sqrt{2}\, \hat{d\!l}^+_\rho(f_1, f_2).$$

Our first stability result, already announced in [5], is concerned with the solution set $S(P)$, rather than $S_\varepsilon(P)$ that will be dealt with later on.

Theorem 2.1 *Let $P \in \mathcal{P}_{\mathcal{F}}$ and suppose $S(P)$ is non-empty and bounded. Then there exist constants $\rho > 0$ and $\delta > 0$ such that*

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

holds for all $Q \in \mathcal{P}_{\mathcal{F}}$ with $d_{\mathcal{F},\rho}(P,Q) < \delta$, where Ψ_P is a conditioning function *associated with our given problem (1), more precisely,*

$$\Psi_P(\eta) := \eta + \psi_P^{-1}(2\eta), \quad \eta \geq 0$$

with

$$\psi_P(\tau) := \min\{\mathbb{E}^P f_0(x) - v(P) : d(x, S(P)) \geq \tau\}, \quad \tau \geq 0.$$

Proof: For any $Q \in \mathcal{P}_{\mathcal{F}}$, the function $\mathbb{E}^Q f_0$ is lower semicontinuous, proper and convex. Define

$$F_Q(x) := \begin{cases} \mathbb{E}^Q f_0(x), & x \in X, \\ +\infty, & \text{else}, \end{cases}$$

for each $Q \in \mathcal{P}_{\mathcal{F}}$ and rely on [20, Theorem 7.64] to derive the result. Let $\bar{\rho} > 0$ be chosen such that $S(P) \subset \bar{\rho}\mathbb{B}$ and $v(P) \geq -\bar{\rho}$. For $\rho > \bar{\rho}$ and δ such that $0 < \delta < \min\{\frac{1}{2}(\rho - \bar{\rho}), \frac{1}{2}\psi_P(\frac{1}{2}(\rho - \bar{\rho}))\}$, since F_Q and F_P are convex, Theorem 7.64 of [20] yields the estimates

$$\begin{aligned} |v(P) - v(Q)| &\leq \hat{d\!l}_\rho^+(\mathbb{E}^P f_0, \mathbb{E}^Q f_0) \\ \emptyset \neq S(Q) &\subseteq S(P) + \Psi_P(\hat{d\!l}_\rho^+(\mathbb{E}^P f_0, \mathbb{E}^Q f_0))\mathbb{B} \end{aligned}$$

for any $Q \in \mathcal{P}_{\mathcal{F}}$ with $\hat{d\!l}_\rho^+(\mathbb{E}^P f_0, \mathbb{E}^Q f_0) < \delta$. Now, let η be chosen such that $\eta \geq \max_{x \in X \cap \rho\mathbb{B}} |\mathbb{E}^P f_0(x) - \mathbb{E}^Q f_0(x)|$. Clearly, the inequalities

$$\begin{aligned} \min_{y \in x + \eta\mathbb{B}} F_Q(y) &\leq \max\{F_P(x), -\rho\} + \eta \\ \min_{y \in x + \eta\mathbb{B}} F_P(y) &\leq \max\{F_Q(x), -\rho\} + \eta \end{aligned}$$

are trivially satisfied when $x \notin X$. When $x \in X \cap \rho\mathbb{B}$, we have

$$\begin{aligned} \min_{y \in x + \eta\mathbb{B}} F_Q(y) &\leq F_Q(x) \leq F_P(x) + \eta = \max\{F_P(x), -\rho\} + \eta \\ \min_{y \in x + \eta\mathbb{B}} F_P(y) &\leq F_P(x) \leq F_Q(x) + \eta \leq \max\{F_P(x), -\rho\} + \eta. \end{aligned}$$

and, thus, $\hat{d\!l}_\rho^+(F_P, F_Q) \leq \eta$. Letting η pass to its lower limit leads to

$$\hat{d\!l}_\rho^+(F_P, F_Q) \leq \max_{x \in X \cap \rho\mathbb{B}} |\mathbb{E}^P f_0(x) - \mathbb{E}^Q f_0(x)| = d_{\mathcal{F},\rho}(P,Q). \tag{3}$$

Since the function Ψ_P is increasing, the proof is complete. ■

Simple examples of two-stage stochastic programs show that, in general, the set-valued mapping $S(.)$ is not inner semicontinuous at P (cf. [21, Example 26]). Furthermore, explicit descriptions of conditioning functions ψ_P of stochastic programs (like linear or quadratic growth at solution sets) are only known in some specific cases, for example, for linear two-stage stochastic programs with finite discrete distribution or with strictly positive densities of random right-hand sides [25].

As we shall see, we are in much better shape, when we consider the stability properties of the sets $S_\varepsilon(\cdot)$ of ε-approximate solutions. Indeed, $S_\varepsilon(\cdot)$ even satisfies a Lipschitz property under rather mild assumptions.

Theorem 2.2 *Let $P, Q \in \mathcal{P}_\mathcal{F}$ and such that the corresponding solution sets $S(P)$ and $S(Q)$ are non-empty. Then there exist constants $\rho > 0$ and $\bar{\varepsilon} > 0$ such that*

$$\hat{d\!l}_\rho(S_\varepsilon(P), S_\varepsilon(Q)) \leq \frac{4\rho}{\varepsilon} d_{\mathcal{F},\rho+\varepsilon}(P, Q)$$

holds for any $\varepsilon \in (0, \bar{\varepsilon})$ where $d_{\mathcal{F},\rho+\varepsilon}(P, Q) < \varepsilon$.

Proof: The assumptions imply that both $I\!E^P f_0$ and $I\!E^Q f_0$ are proper, lsc and convex on $I\!R^m$. Let ρ_0 be chosen such that both $S(P) \cap \rho_0 I\!B$ and $S(Q) \cap \rho_0 I\!B$ are non-empty and $\min\{v(P), v(Q)\} \geq -\rho_0$. For $\rho > \rho_0$ and $0 < \varepsilon < \bar{\varepsilon} = \rho - \rho_0$, one obtains, from the proof of [20, Theorem 7.69], the inclusion

$$S_\varepsilon(P) \cap \rho I\!B \subseteq S_\varepsilon(Q) + \frac{2\eta}{\varepsilon + 2\eta} 2\rho I\!B \subseteq S_\varepsilon(Q) + \frac{4\rho}{\varepsilon} \eta I\!B,$$

for all $\eta > \hat{d\!l}^+_{\rho+\varepsilon}(I\!E^P f_0, I\!E^Q f_0)$. This implies

$$S_\varepsilon(P) \cap \rho I\!B \subseteq S_\varepsilon(Q) + \frac{4\rho}{\varepsilon} \hat{d\!l}^+_{\rho+\varepsilon}(I\!E^P f_0, I\!E^Q f_0) I\!B.$$

The same argument works with P and Q interchanged. Finally, we appeal to the estimate (3) to complete the proof. ■

The above estimate for ε-approximate solution sets allows for the solution sets to be unbounded. The result becomes somewhat more tangible if the original solution set $S(P)$ is assumed to be bounded.

Corollary 2.3 *Let $P \in \mathcal{P}_\mathcal{F}$ and $S(P)$ be non-empty, bounded. Then there exist constants $\hat{\rho} > 0$ and $\hat{\varepsilon} > 0$ such that*

$$d\!l_\infty(S_\varepsilon(P), S_\varepsilon(Q)) \leq \frac{4\hat{\rho}}{\varepsilon} d_{\mathcal{F},\hat{\rho}+\varepsilon}(P, Q)$$

holds for any $\varepsilon \in (0, \hat{\varepsilon})$ and $Q \in \mathcal{P}_\mathcal{F}$ such that $d_{\mathcal{F},\hat{\rho}+\varepsilon}(P, Q) < \varepsilon$.

Proof: Let δ and ρ be the constants from Theorem 2.1 and put $\hat{\varepsilon} = \delta$. Let $\varepsilon \in (0, \hat{\varepsilon})$ and $Q \in \mathcal{P}_{\mathcal{F}}$ such that $d_{\mathcal{F},\rho+\varepsilon}(P,Q) < \varepsilon$. Then $S(Q)$ is also non-empty and bounded. Since the functions $\mathbb{E}^P f_0$ and $\mathbb{E}^Q f_0$ are lower semicontinuous and convex, the level sets $S_{\hat{\varepsilon}}(P)$ and $S_{\hat{\varepsilon}}(Q)$ are bounded since the sets $S_0(P)$ and $S_0(Q)$ are bounded (cf. [18, Corollary 8.7.1]). Next we choose ρ_0 as in Theorem 2.2 and $\hat{\rho}$ such that $\hat{\rho} > \max\{\rho, \rho_0 + \hat{\varepsilon}\}$ and both level sets $S_{\hat{\varepsilon}}(P)$ and $S_{\hat{\varepsilon}}(Q)$ are contained in $\hat{\rho}\mathbb{B}$. Then the result follows from Theorem 2.2 by taking into account that

$$\hat{d\!l}_{\hat{\rho}}(S_\varepsilon(P), S_\varepsilon(Q)) = d\!l_\infty(S_\varepsilon(P), S_\varepsilon(Q))$$

holds because of the choice of $\hat{\rho}$. ■

The results illuminate the role of the probability distances $d_{\mathcal{F},\rho}$ given that the parameter $\rho > 0$ is properly chosen. These probability metrics process the minimal information about problem (1) and allow us to derive, remarkable stability properties for the optimal values and (approximate) solutions. Clearly, the preceding stability results remain valid if the set $\mathcal{F}_\rho$ is enlarged to a set $\hat{\mathcal{F}}$ and the set $\mathcal{P}_{\mathcal{F}}$ reduced to a subset on which the new distance $d_{\hat{\mathcal{F}}}$ is finite and well-defined.

Hence, it is important to identify classes $\hat{\mathcal{F}}$ of functions that contain $\{f_0(\cdot, x) : x \in X \cap \rho\mathbb{B}\}$ for any $\rho > 0$. For many convex stochastic programming problems the functions $f_0(\cdot, x)$, $x \in X$, are locally Lipschitz continuous on Ξ with certain Lipschitz constants $L(r)$ on the sets $\{\xi \in \Xi : \|\xi - \xi_0\| \le r\}$ for some $\xi_0 \in \Xi$ and any $r > 0$. In many cases, the growth modulus $L(r)$ does not depend on x, in particular when x is only varying in a bounded subset of $\mathbb{R}^m$. Hence, function classes of the form

$$\mathcal{F}_H := \{\psi : \Xi \to \mathbb{R} : \psi(\xi) - \psi(\tilde{\xi}) \le \max\{1, H(\|\xi - \xi_0\|), H(\|\tilde{\xi} - \xi_0\|)\}\|\xi - \tilde{\xi}\|, \forall \xi, \tilde{\xi} \in \Xi\}$$

are of particular interest, where $H : \mathbb{R}_+ \to \mathbb{R}_+$ is nondecreasing, $H(0) = 0$ and $\xi_0 \in \Xi$. The distances introduced in (2), but with $\mathcal{F} = \mathcal{F}_H$, i.e.,

$$d_{\mathcal{F}_H}(P,Q) = \sup\left\{ \left| \int_\Xi \psi(\xi)P(d\xi) - \int_\Xi \psi(\xi)Q(d\xi) \right| : \psi \in \mathcal{F}_H \right\} := \zeta_H(P,Q)$$

are so-called *Fortet-Mourier metrics*, denoted by ζ_H and defined on

$$\mathcal{P}_H(\Xi) := \{Q \in \mathcal{P}(\Xi) : \int_\Xi \max\{1, H(\|\xi - \xi_0\|)\}\|\xi - \xi_0\|Q(d\xi) < \infty\}$$

(cf. [7, 15]). Important special cases come to light when the function H has the polynomial form $H(t) := t^{r-1}$ for $r \ge 1$. The corresponding function classes and distances are denoted by $\mathcal{F}_r$ and ζ_r, respectively. The distances ζ_r are well defined on the set

$$\mathcal{P}_r(\Xi) := \{Q \in \mathcal{P}(\Xi) : \int_\Xi \|\xi\|^r Q(d\xi) < \infty\} \tag{4}$$

of probability measures having finite r-th order moments.

3 Stability of Two-Stage Recourse Models

We consider the linear two-stage stochastic program with recourse,

$$\min\Big\{cx + \int_{\Xi} q(\xi)y(\xi)P(d\xi) : W(\xi)y(\xi) = h(\xi) - T(\xi)x,\, y(\xi) \in Y,\, x \in X\Big\}, \qquad (5)$$

where $c \in \mathbb{R}^m$, $X \subseteq \mathbb{R}^m$ and $\Xi \subseteq \mathbb{R}^s$ are polyhedral, $Y \subseteq \mathbb{R}^{\overline{m}}$ is a polyhedral cone and $P \in \mathcal{P}(\Xi)$. We assume that $q(\xi) \in \mathbb{R}^{\overline{m}}$, $h(\xi) \in \mathbb{R}^d$, the recourse matrix $W(\xi) \in \mathbb{R}^{d\times\overline{m}}$ and the technology matrix $T(\xi) \in \mathbb{R}^{d\times n}$ may depend affinely on $\xi \in \Xi$.
Denoting by $\Phi(\xi, q(\xi), h(\xi) - T(\xi)x)$, the value of the optimal second stage decision, problem (5) may be rewritten equivalently as a minimization problem with respect to the first stage decision x. Defining the integrand $f_0 : \Xi \times \mathbb{R}^m \to \overline{\mathbb{R}}$ by

$$f_0(\xi, x) = \begin{cases} cx + \Phi(\xi, q(\xi), h(\xi) - T(\xi)x) & \text{if } h(\xi) - T(\xi)x \in W(\xi)Y,\ D(\xi) \neq \emptyset, \\ +\infty & \text{otherwise,} \end{cases}$$

where, the optimal value function Φ and the dual feasibility set $D(\xi)$, are defined by

$$\begin{aligned} \Phi(\xi, u, t) &:= \inf\{uy : W(\xi)y = t, y \in Y\}, \qquad (\xi, u, t) \in \Xi \times \mathbb{R}^{\overline{m}} \times \mathbb{R}^d \\ D(\xi) &:= \{z \in \mathbb{R}^r : W(\xi)^\top z - q(\xi) \in Y^*\}, \quad \xi \in \Xi, \end{aligned}$$

where $W(\xi)^\top$ is the transposed of $W(\xi)$ and Y^* the polar cone of Y.
The (equivalent) minimization problem can thus be expressed as

$$\min\Big\{\int_{\Xi} f_0(\xi, x)P(d\xi) :\ x \in X\Big\}. \qquad (6)$$

In order to utilize the general stability results of Section 2, we first recall some well-known properties of the function Φ (cf. [31]).

Lemma 3.1 *For any $\xi \in \Xi$, the function $\Phi(\xi, \cdot, \cdot)$ is finite and continuous on the polyhedral set $D(\xi) \times W(\xi)Y$. Furthermore, the function $\Phi(\xi, u, \cdot)$ is piecewise linear convex on the polyhedral set $W(\xi)Y$ for fixed $u \in D(\xi)$, and $\Phi(\xi, \cdot, t)$ is piecewise linear concave on $D(\xi)$ for fixed $t \in W(\xi)Y$.*

We impose the following conditions on problem (6):

(A1) *relatively complete recourse*: for any $(\xi, x) \in \Xi \times X$, $h(\xi) - T(\xi)x \in W(\xi)Y$;

(A2) *dual feasibility*: $D(\xi) \neq \emptyset$ holds for all $\xi \in \Xi$.

Conditions (A1) and (A2) are standard ones and render problem (6) well-defined. Due to Lemma 3.1 they imply that f_0 is a convex random lsc function with $\Xi \times X \subseteq \operatorname{dom} f_0$. As earlier, with the notation

$$\mathcal{F}_\rho := \{f_0(\cdot, x) : x \in X \cap \rho\mathbb{B}\}, \qquad (7)$$

we obtain our first stability result for model (5) as immediate consequences of Theorem 2.1 and Corollary 2.3.

Theorem 3.2 *Suppose the stochastic program satisfies the relatively complete recourse (A1) and the dual feasibility (A2) conditions, $P \in \mathcal{P}_{\mathcal{F}}$ and $S(P)$ is non-empty and bounded. Then there exist constants $\rho > 0$ and $\hat{\varepsilon} > 0$ such that*

$$\begin{aligned} |v(P) - v(Q)| &\leq d_{\mathcal{F},\rho}(P,Q) \\ d\!\!l_\infty(S_\varepsilon(P), S_\varepsilon(Q)) &\leq \frac{4\rho}{\varepsilon} d_{\mathcal{F},\rho+\varepsilon}(P,Q) \end{aligned}$$

holds for any $\varepsilon \in (0, \hat{\varepsilon})$ and each $Q \in \mathcal{P}_{\mathcal{F}}$ such that $d_{\mathcal{F},\rho+\varepsilon}(P,Q) < \varepsilon$.

The theorem establishes Lipschitz stability of $v(.)$ and S_ε in the two-stage case for fairly general situations. However, the set of (perturbed) probability measures $\mathcal{P}_{\mathcal{F}}$ and, in particular, the metrics $d_{\mathcal{F},\rho}$ are rather sophisticated and could be difficult to use in applications.

To overcome this difficulty, we need to explore quantitative continuity properties of the integrand f_0. Such properties are well known in case of *fixed recourse*, i.e., in case $W(\xi) \equiv W$ [32], and have been used to analyze quantitative stability in [16]. Our first result for random recourse matrices follows the ideas in [33]. There, it is shown that (semi)continuity properties of parametric optimal value functions are consequences of the (semi)continuity of the primal and dual feasibility mapping with respect to the relevant parameters. Next, we verify that a local Lipschitz property of the dual feasible set-valued mapping $\xi \mapsto D(\xi)$ in addition to (A1) implies local Lipschitz continuity of $f_0(\cdot, x)$ with the modulus not depending on having x vary only in a bounded set.

Proposition 3.3 *Suppose the stochastic program satisfies the relatively complete recourse (A1) and the dual feasibility (A2) conditions. Assume also that the mapping $\xi \mapsto D(\xi)$ is bounded-valued and locally Lipschitz continuous on Ξ with respect to the Pompeiu-Hausdorff distance (on the subsets of $I\!R^d$), i.e., there exists a constant $L > 0$, an element $\xi_0 \in \Xi$ and a nondecreasing function $h : I\!R_+ \to I\!R_+$ with $h(0) = 0$ such that*

$$d\!\!l_\infty(D(\xi), D(\tilde{\xi})) \leq L \max\{1, h(\|\xi - \xi_0\|), h(\|\tilde{\xi} - \xi_0\|)\} \|\xi - \tilde{\xi}\| \tag{8}$$

holds for all $\xi, \tilde{\xi} \in \Xi$.
Then, for any $\rho > 0$, there exist constants $\hat{L} > 0$ and $\hat{L}(\rho) > 0$ such that

$$f_0(\xi, x) - f_0(\tilde{\xi}, x) \leq \hat{L}(\rho) \max\{1, H(\|\xi - \xi_0\|), H(\|\tilde{\xi} - \xi_0\|)\} \|\xi - \tilde{\xi}\| \tag{9}$$
$$f_0(\xi, x) - f_0(\xi, \tilde{x}) \leq \hat{L} \max\{1, H(\|\xi - \xi_0\|) \|\xi - \xi_0\|\} \|x - \tilde{x}\| \tag{10}$$

for all $\xi, \tilde{\xi} \in \Xi$, $x, \tilde{x} \in X \cap \rho I\!B$, where H is defined by

$$H(t) := h(t)t, \; \forall t \in I\!R_+. \tag{11}$$

Proof: Let $\rho > 0$. Due to (A1) and (A2), the function $f_0(\cdot, x)$ is real-valued for every $x \in X$. For any $x, \tilde{x} \in X \cap \rho I\!B$ and $\xi, \tilde{\xi} \in \Xi$, one has the estimate

$$f_0(\xi, x) - f_0(\tilde{\xi}, \tilde{x}) \leq cx + (h(\xi) - T(\xi)x)z^*(\xi) - (h(\tilde{\xi}) - c\tilde{x} - T(\tilde{\xi})\tilde{x})z(\tilde{\xi}), \tag{12}$$

where $z^*(\xi) \in D(\xi)$ is a dual solution of the second-stage problem and $z(\tilde{\xi})$ is some element in $D(\tilde{\xi})$. We denote by $\bar{z}(\tilde{\xi};\xi)$ the projection of $z^*(\xi)$ onto $D(\tilde{\xi})$, i.e.,

$$d(z^*(\xi), D(\tilde{\xi})) = \|z^*(\xi) - \bar{z}(\tilde{\xi};\xi)\|,$$

yielding

$$\|z^*(\xi) - \bar{z}(\tilde{\xi};\xi)\| \le d\!\!l_\infty(D(\xi), D(\tilde{\xi})) \le L\max\{1, h(\|\xi-\xi_0\|), h(\|\tilde{\xi}-\xi_0\|)\}\|\xi-\tilde{\xi}\|. \quad (13)$$

As $D(\xi_0)$ is bounded, there exists $r > 0$ such that $\|z\| \le r$ for each $z \in D(\xi_0)$. As the estimate

$$d(\bar{z}(\tilde{\xi};\xi), D(\xi_0)) \le L\max\{1, h(\|\tilde{\xi}-\xi_0\|)\}\|\tilde{\xi}-\xi_0\|$$

holds for all $\xi, \tilde{\xi} \in \Xi$ according to (8), we have

$$\|\bar{z}(\tilde{\xi};\xi)\| \le \max\{r, L\}\max\{1, h(\|\tilde{\xi}-\xi_0\|)\}\|\tilde{\xi}-\xi_0\|. \quad (14)$$

Now, we proceed with our estimate (12) when $x = \tilde{x}$, exploiting the affine linearity of $h(\cdot)$ and $T(\cdot)$, (13) and (14). Setting $z(\tilde{\xi}) := \bar{z}(\tilde{\xi};\xi)$ we obtain

$$\begin{aligned}
f_0(\xi,x) - f_0(\tilde{\xi},x) &\le (h(\xi) - T(\xi)x)(z^*(\xi) - \bar{z}(\tilde{\xi};\xi)) \\
&\quad -((h(\tilde{\xi}) - h(\xi)) - (T(\tilde{\xi}) - T(\xi))x)\bar{z}(\tilde{\xi};\xi) \\
&\le \|h(\xi) - T(\xi)x\|\|z^*(\xi) - \bar{z}(\tilde{\xi};\xi)\| \\
&\quad +(\|h(\tilde{\xi}) - h(\xi)\| + \|T(\tilde{\xi}) - T(\xi)\|\|x\|)\|\bar{z}(\tilde{\xi};\xi)\| \\
&\le \Big(KL(1+\rho)\max\{1, \|\xi-\xi_0\|\}\max\{1, h(\|\xi-\xi_0\|), h(\|\tilde{\xi}-\xi_0\|)\} \\
&\quad +\tilde{K}\max\{r, L\}(1+\rho)\max\{1, h(\|\tilde{\xi}-\xi_0\|)\}\|\tilde{\xi}-\xi_0\|\Big)\|\xi-\tilde{\xi}\| \\
&\le \bar{L}(1+\rho)\max\{1, H(\|\xi-\xi_0\|), H(\|\tilde{\xi}-\xi_0\|)\}\|\xi-\tilde{\xi}\|
\end{aligned}$$

for each $\xi, \tilde{\xi} \in \Xi$ and some positive constants K, $\tilde{K}$ and $\bar{L}$. Thus, (9) is proved with $\hat{L}(\rho) = \bar{L}(1+\rho)$. Finally, we return to (12) in case $\xi = \tilde{\xi}$, choosing $\bar{z}(\xi) = z^*(\xi)$, we arrive at the estimate

$$\begin{aligned}
f_0(\xi,x) - f_0(\xi,\tilde{x}) &\le c(x-\tilde{x}) + T(\xi)(\tilde{x}-x)z^*(\xi) \le (\|c\| + \|T(\xi)\|\|z^*(\xi)\|)\|x-\tilde{x}\| \\
&\le \hat{L}\max\{1, H(\|\xi-\xi_0\|)\|\xi-\xi_0\|\}\|x-\tilde{x}\|
\end{aligned}$$

for some constant $\hat{L} > 0$ and all $\xi \in \Xi$, $x, \tilde{x} \in X \cap \rho I\!B$. Here, we used that $\|z^*(\xi)\|$ can be bounded in the same way as $\bar{z}(\tilde{\xi};\xi)$ in (14). ∎

Our next example illustrates the local Lipschitz continuity property (8) of the dual feasibility mapping D.

Example 3.4 *Let* $\overline{m} = 4$, $d = 2$, $Y = I\!R^4_+$, $\Xi = I\!R$ *and consider the random (second-stage) costs and recourse matrix*

$$W(\xi) = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -\xi & 0 & 1 & -1 \end{pmatrix} \qquad q(\xi) = \begin{pmatrix} 0 \\ 0 \\ \xi \\ -\xi \end{pmatrix}$$

Then $W(\xi)Y = I\!R^2$ *(complete recourse) and* $D(\xi) = [0, \xi^2] \times \{\xi\}$. *Hence, the conditions (A1), (A2) and (8) are satisfied with* $h(t) = t$ *for each* $t \in I\!R_+$.

We can reformulate the conclusions of the preceding proposition in terms of the Fortet-Mourier metrics defined on $\mathcal{P}_H(\Xi)$, the space of probability measures.

Corollary 3.5 *Let the assumptions of Proposition 3.3 be satisfied, $P \in \mathcal{P}_H(\Xi)$ and $S(P)$ be non-empty and bounded.*
Then there exist constants $\hat{L} > 0$, $\rho > 0$ and $\hat{\varepsilon} > 0$ such that

$$\begin{aligned} |v(P) - v(Q)| &\leq \hat{L}\zeta_H(P, Q) \\ d\!\!l_\infty(S_\varepsilon(P), S_\varepsilon(Q)) &\leq \frac{4\rho\hat{L}}{\varepsilon}\zeta_H(P, Q) \end{aligned}$$

holds for any $\varepsilon \in (0, \hat{\varepsilon})$ and each $Q \in \mathcal{P}_H(\Xi)$ such that $\zeta_H(P, Q) < \varepsilon$ where H is defined by (11), $\zeta_H(P, Q)$ is Fortet-Mourier metric on $\mathcal{P}_H(\Xi)$.

Proof: The estimate (9) implies $d_{\mathcal{F},\rho}(P, Q) \leq \hat{L}\zeta_H(P, Q)$ with $\hat{L} = \hat{L}(\rho)$ and, hence, the result follows from Theorem 3.2. ■

When $W(\xi) \equiv W$, the mapping $\xi \mapsto D(\xi)$ is even Lipschitz continuous with respect to the Pompeiu-Hausdorff distance $d\!\!l_\infty$. Hence, $H(\xi) \equiv \xi$ and $\mathcal{F}_H = \mathcal{F}_2$, and then the previous result boils down to [16, Proposition 3.2].

4 Two-Stage Multi-period Models

If the second stage of a stochastic program with recourse models a (stochastic) dynamical decision process, as is the case in a variety of applications, our two-stage problem takes on the form:

$$\min\Big\{cy_0 + \sum_{j=1}^{\ell} q_j(\xi)y_j : y_0 \in X, y_j \in Y_j, W_{jj}y_j = h_j(\xi) - W_{jj-1}(\xi)y_{j-1}, j = 1, \ldots, \ell\Big\} \quad (15)$$

where for $j = 1, \ldots, \ell$, $Y_j \in I\!R^{\overline{m}_j}$ are polyhedral sets for some finite ℓ and first-stage decision $x := y_0$; the matrices $W_{j,j-1}(\xi)$ are (potentially) stochastic. Then the second stage program has separable block structure, i.e., the recourse variable y has the form $y = (y_1, \ldots, y_\ell)$, the polyhedral set Y is the Cartesian product of polyhedral sets $Y_j \in I\!R^{\overline{m}_j}$, $j = 1, \ldots, \ell$, the element $T(\xi)x$ has the components $T_1(\xi)x := W_{10}(\xi)x$ and $T_j(\xi)x = 0$, $j = 2, \ldots, \ell$, and the random recourse matrix $W(\xi)$ is of the form

$$W(\xi) = \begin{pmatrix} W_{11} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ W_{21}(\xi) & W_{22} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & W_{32}(\xi) & W_{33} & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\vdots & \\ 0 & 0 & 0 & 0 & \cdots & W_{\ell-1\ell-2}(\xi) & W_{\ell-1\ell-1} & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & W_{\ell\ell-1}(\xi) & W_{\ell\ell} \end{pmatrix}, \quad (16)$$

i.e., all matrices W_{jj}, $j = 1, \ldots, \ell$, in the diagonal of $W(\xi)$ are non-stochastic. Denoting by $q_j(\xi)$ and $h_j(\xi)$ the components of $q(\xi)$ and $h(\xi)$, respectively, the integrand f_0 is of the form

$$\begin{aligned} f_0(\xi, x) &= cx + \inf\Big\{\sum_{j=1}^{\ell} q_j(\xi)y_j : W_{jj}y_j = h_j(\xi) - W_{jj-1}(\xi)y_{j-1}, y_j \in Y_j, j = 1, \ldots, \ell\Big\} \\ &=: cx + \Psi_1(\xi, x), \end{aligned}$$

where the function Ψ_1 is given by the recursion

$$\Phi_j(\xi, u_{j-1}) := \inf\Big\{q_j(\xi)y_j + \Psi_{j+1}(\xi, y_j) : W_{jj}y_j = u_{j-1}, y_j \in Y_j\Big\} \tag{17}$$

$$\Psi_j(\xi, y_{j-1}) := \Phi_j(\xi, h_j(\xi) - W_{jj-1}(\xi)y_{j-1}) \tag{18}$$

for $j = \ell, \ldots, 1$, where $y_0 = x$ and $\Psi_{\ell+1}(\xi, y_\ell) \equiv 0$.

While the continuity and growth properties of the function $f_0(\cdot, x)$ in case $\ell = 1$ may be derived from Lemma 3.1, we need an extended result for establishing Lipschitz continuity properties of the inf-projection Φ_j for $j = 1, \ldots, \ell$. The results in [35] were developed precisely to deal with the present situation. To state the result, we denote by D^∞ the horizon cone of a convex set $D \subseteq \mathbb{R}^m$. It consists of all elements $x_d \in \mathbb{R}^m$ such that $x + \lambda x_d \in D$ for all $x \in D$ and $\lambda \in \mathbb{R}_+$. Clearly, we have $D^\infty = \{0\}$ if D is bounded. Furthermore, D^∞ is polyhedral if D is polyhedral. Next we record [35, Proposition 4.4] and provide a self-contained proof for the convenience of the reader.

Lemma 4.1 *Let $h \in \mathbb{R}^d$, $W \in \mathbb{R}^{d\times n}$ and $Y \subseteq \mathbb{R}^n$ be polyhedral. Let $u = (u_1, u_2) \in \mathbb{R}^n \times \mathbb{R}^d$ and*

$$\Phi(u) := \inf\{f(u_1, y) : Wy = h - u_2, y \in Y\}$$

Assume that $\ker(W) \cap Y^\infty = \{0\}$ *and that f is Lipschitz continuous on $\{(u_1, y) \in \mathbb{R}^n \times Y : \|u_1\| \leq r, \|y\| \leq r\}$ with constant $L(r)$ for every $r > 0$. Then, $\Phi(\cdot)$ is Lipschitz continuous on $\{(u_1, u_2) \in \operatorname{dom}\Phi : \|u_1\| \leq r, \|u_2\| \leq r\}$ with constant $L_M L(K_M \max\{1, r\})$ for every $r > 0$, where $L_M \geq 1$ and $K_M \geq 1$ are constants depending only on the set-valued mapping $M(u_2) := \{y \in Y : Wy = h - u_2\}$ from $\mathbb{R}^d$ to $\mathbb{R}^n$.*

Proof: The condition $\ker(W) \cap Y^\infty = \{0\}$ is equivalent to the local boundedness of the mapping M. M is Lipschitz continuous with respect to the Pompeiu-Hausdorff distance $d\!\!I_\infty$ (with constant $L_M \geq 1$) since its graph is polyhedral [20, Example 9.35]. Since the set $M(u_2)$ is compact, Φ is finite for all pairs (u_1, u_2) such that $u_2 \in \operatorname{dom} M$. Now, let $r > 0$ and $u = (u_1, u_2), \tilde{u} = (\tilde{u}_1, \tilde{u}_2) \in \operatorname{dom}\Phi \cap \{(u_1, u_2) \in \mathbb{R}^n \times \mathbb{R}^d : \|u_1\| \leq r, \|u_2\| \leq r\}$. Then there exist $y(u_2) \in M(u_2)$ and $y(\tilde{u}_2) \in M(\tilde{u}_2)$ such that $\Phi(u) = f(u_1, y(u_2))$ and $\|y(u_2) - y(\tilde{u}_2)\| \leq L_M\|u_2 - \tilde{u}_2\|$. In particular, there exists a constant $K_M \geq 1$ such that

$$\max\{\|y(u_2)\|, \|y(\tilde{u}_2)\|\} \leq K_M \max\{1, \|u_2\|, \|\tilde{u}_2\|\} \leq K_M \max\{1, r\}.$$

We obtain

$$\begin{aligned}\Phi(\tilde{u}) - \Phi(u) &\leq f(\tilde{u}_1, y(\tilde{u}_2)) - f(u_1, y(u_2)) \\ &\leq L(K_M \max\{1, r\})(\|\tilde{u}_1 - u_1\| + \|y(\tilde{u}_2) - y(u_2)\|) \\ &\leq L_M L(K_M \max\{1, r\})(\|\tilde{u}_1 - u_1\| + \|\tilde{u}_2 - u_2\|)\end{aligned}$$

and that completes the proof. ■

Proposition 4.2 *Let $W(\xi)$ be as described by (16). Assume the relatively complete recourse condition (A1) is satisfied and that* $\ker(W_{jj}) \cap Y_j^\infty = \{0\}$ *for* $j = 1, \ldots, \ell - 1$. *Then, there exist constants $L > 0$, $\hat{L} > 0$ and $K > 0$ such that the following holds for all $\xi, \tilde{\xi} \in \Xi$ and $x, \tilde{x} \in X \cap \rho I\!B$:*

$$\begin{aligned}|f_0(\xi, x) - f_0(\tilde{\xi}, x)| &\leq L \max\{1, \rho, \|\xi\|^\ell, \|\tilde{\xi}\|^\ell\}\|\xi - \tilde{\xi}\|, \\ |f_0(\xi, x) - f_0(\xi, \tilde{x})| &\leq \hat{L} \max\{1, \|\xi\|^{\ell+1}\}\|x - \tilde{x}\|, \\ |f_0(\xi, x)| &\leq K \max\{1, \rho, \|\xi\|^{\ell+1}\}.\end{aligned}$$

Proof: Due to the assumptions, all sets of the form $M_j(v_j) := \{y_j \in Y_j : W_{jj}y_j = v_j\}$ are bounded polyhedra for all $v_j \in I\!R^{r_j}$ and $j = 1, \ldots, \ell$. Furthermore, the set-valued mappings M_j from $I\!R^{r_j}$ to $I\!R^{\overline{m}_j}$ are Lipschitz continuous on $\operatorname{dom} M_j$ with constant L_j. Due to (A1), we have recursively $h_j(\xi) - W_{jj-1}(\xi)y_{j-1} \in \operatorname{dom} M_j$ for all $y_{j-1} \in Y_{j-1}$, $y_0 = x \in X$, $\xi \in \Xi$ and $j = 2, \ldots, \ell$. Hence, if Lemma 4.1 is used recursively by setting $\Phi = \Phi_j$, $f_j(u_1, y_j) := q_j(\xi)y_j + \Psi_{j+1}(\xi, y_j)$ with $u_1 = \xi$ and $u_2 = u_{j-1}$, each subproblem (17) is solvable. First we consider the functions Φ_ℓ and Ψ_ℓ.

$$\begin{aligned}\Phi_\ell(\xi, u_{\ell-1}) &= \inf\{q_\ell(\xi)y_\ell : W_{\ell\ell}y_\ell = u_{\ell-1}, y_\ell \in Y_\ell\} \\ \Psi_\ell(\xi, y_{\ell-1}) &= \Phi_\ell(\xi, h_\ell(\xi) - W_{\ell\ell-1}(\xi)y_{\ell-1}).\end{aligned}$$

Then the Lipschitz constant of f_j on $\{(\xi, y_\ell) \in \Xi \times Y_\ell : \|\xi\| \leq r, \|y_\ell\| \leq r\}$ has the form $L_\ell \max\{1, r\}$ and Lemma 4.1 implies that Φ_ℓ has the Lipschitz constant $\hat{L}_\ell \max\{1, r\}$ on $\{(\xi, u_{\ell-1}) \in \Xi \times \operatorname{dom} M_\ell : \|\xi\| \leq r, \|u_{\ell-1}\| \leq r\}$. Due to the term $W_{\ell\ell-1}(\xi)y_{\ell-1}$ in the definition of Ψ_ℓ, however, the function Ψ_ℓ has the Lipschitz constant $\tilde{L}_\ell \max\{1, r^2\}$ on $\{(\xi, y_{\ell-1}) \in \Xi \times Y_{\ell-1} : \|\xi\| \leq r, \|y_{\ell-1})\| \leq r\}$. Since Ψ_ℓ enters the definition of $f_{\ell-1}$ and the infimum, $\Phi_{\ell-1}$ is Lipschitz continuous with constant $\hat{L}_{\ell-1} \max\{1, r^2\}$ on $\{(\xi, u_{\ell-2}) \in \Xi \times \operatorname{dom} M_{\ell-1} : \|\xi\| \leq r, \|u_{\ell-2}\| \leq r\}$ according to Lemma 4.1. Due to the term $W_{\ell-1\ell-2}(\xi)y_{\ell-2}$, the function $\Psi_{\ell-1}$ is Lipschitz continuous with constant $\tilde{L}_{\ell-1} \max\{1, r^3\}$ on $\{(\xi, y_{\ell-2}) \in \Xi \times Y_{\ell-2} : \|\xi\| \leq r, \|y_{\ell-2})\| \leq r\}$ etc. This process may be continued until one concludes that Φ_1 is Lipschitz continuous with constant $\hat{L}_1 \max\{1, r^\ell\}$ on $\{(\xi, u_0) \in \Xi \times \operatorname{dom} M_1 : \|\xi\| \leq r, \|u_0\| \leq r\}$. Hence, the function Ψ_1 depending on (ξ, x) satisfies the following Lipschitz continuity property

$$|\Psi_1(\xi, x) - \Psi_1(\tilde{\xi}, \tilde{x})| \leq \tilde{L}_1 \max\{1, \rho, r^\ell\}(\max\{1, \rho\}\|\xi - \tilde{\xi}\| + \max\{1, r\}\|x - \tilde{x}\|)$$

on the set $\{(\xi, x) \in \Xi \times X : \|\xi\| \leq r, \|x\| \leq \rho\}$.
Thus, yields the assertions about f_0 and completes the proof. ■

Due to the previous result we obtain

$$\mathcal{P}_{\mathcal{F}} \supseteq \mathcal{P}_{\ell+1}(\Xi) = \{Q \in \mathcal{P}(\Xi) : \int_{\Xi} \|\xi\|^{\ell+1} Q(d\xi) < \infty\}$$

and

$$\frac{1}{L \max\{1,\rho\}} f_0(x,\cdot) \in \mathcal{F}_{\ell+1}(\Xi)$$

for each $x \in X \cap \rho I\!B$, and arrive, after specializing Theorem 3.2, to the following:

Corollary 4.3 *Let $W(\xi)$ be as described by (16). Assume the relatively complete recourse condition (A1) is satisfied and that $\ker(W_{jj}) \cap Y_j^\infty = \{0\}$ for $j = 1,\ldots,\ell-1$. Then there exist constants $L > 0$ and $\hat{\varepsilon} > 0$ such that for any $\varepsilon \in (0,\hat{\varepsilon})$ the estimates*

$$\begin{aligned} |v(P) - v(Q)| &\leq L\,\zeta_{\ell+1}(P,Q) \\ d\!l_\infty(S_\varepsilon(P), S_\varepsilon(Q)) &\leq \frac{L}{\varepsilon}\zeta_{\ell+1}(P,Q) \end{aligned}$$

hold whenever $Q \in \mathcal{P}_{\ell+1}(\Xi)$ and $\zeta_{\ell+1}(P,Q) < \varepsilon$.

The case $\ell = 1$ corresponds to the situation of two-stage models with fixed recourse, and that situation was already covered by [21, Theorem 24]. Note that the corollary remains valid for the slightly more general situation that all lower diagonal blocks of $W(\xi)$ are random. If the recent stability result [8, Theorem 2.1] for linear multistage models is restricted to the two-stage model (15), it implies the existence of positive constants L and δ such that

$$|v(P) - v(Q)| \leq L\,\mathcal{W}_{\ell+1}(P,Q) \tag{19}$$

holds for each $Q \in \mathcal{P}_{\ell+1}(\Xi)$ with $\mathcal{W}_{\ell+1}(P,Q) < \delta$; the distance $\mathcal{W}_r$ denotes the r-th order *Wasserstein metric*

$$\mathcal{W}_r(P,Q) := \Big(\inf\Big\{\int_{\Xi\times\Xi} \|\xi - \tilde{\xi}\|^r \eta(d\xi, d\tilde{\xi})\,|\eta \in \mathcal{P}(\Xi\times\Xi),\, \pi_1\eta = P,\ \pi_2\eta = Q\Big\}\Big)^{1/r} \tag{20}$$

on $\mathcal{P}_r(\Xi)$ for any $r \geq 1$, where π_1 and π_2 denote the projections onto the first and second component, respectively. It is known that sequences in $\mathcal{P}_r(\Xi)$ converge with respect to both metrics ζ_r and $\mathcal{W}_r$ if they converge weakly and if their r-th order absolute moments converge. To derive a quantitative estimate, let $\eta^* \in \mathcal{P}(\Xi\times\Xi)$ be a solution of the minimization problem on the right-hand side of (20). Such solutions exist according to [15, Theorem 8.1.1]. Then the duality theorem [15, Theorem 5.3.2] for the Fortet-Mourier metric of order r implies, via Hölder's inequality, the estimate

$$\begin{aligned} \zeta_r(P,Q) &\leq \int_{\Xi\times\Xi} \max\{1,\|\xi\|,\|\tilde{\xi}\|\}^{r-1}\|\xi-\tilde{\xi}\|\eta^*(d\xi,d\tilde{\xi}) \\ &\leq \Big(\int_{\Xi\times\Xi} \max\{1,\|\xi\|,\|\tilde{\xi}\|\}^{r}\eta^*(d\xi,d\tilde{\xi})\Big)^{\frac{r-1}{r}} \Big(\int_{\Xi\times\Xi} \|\xi-\tilde{\xi}\|^r\eta^*(d\xi,d\tilde{\xi})\Big)^{\frac{1}{r}} \\ &= \Big(\int_{\Xi\times\Xi} \max\{1,\|\xi\|,\|\tilde{\xi}\|\}^{r}\eta^*(d\xi,d\tilde{\xi})\Big)^{\frac{r-1}{r}} \mathcal{W}_r(P,Q) \\ &\leq \Big(1 + \int_{\Xi} \|\xi\|^r (P+Q)(d\xi)\Big)^{\frac{r-1}{r}} \mathcal{W}_r(P,Q). \end{aligned}$$

Hence, the stability result for optimal values obtained in Corollary 4.3 extends (19); this extension is 'strict,' as illustrated in [16, Example 3.4].

5 Empirical Approximations of Two-Stage Models

Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be independent identically distributed Ξ-valued random variables on some probability space $(\Omega, \mathcal{A}, I\!P)$ having the common distribution P, i.e., $P = I\!P\xi_1^{-1}$. We consider the empirical measures

$$P_n(\omega) := \frac{1}{n}\sum_{i=1}^{n} \delta_{\xi_i(\omega)} \qquad (\omega \in \Omega;\ n \in I\!N)$$

and the *empirical approximation* of the stochastic program (1) with sample size n, i.e.,

$$\min\Big\{\frac{1}{n}\sum_{i=1}^{n} f_0(\xi_i(\cdot), x) : x \in X\Big\}. \tag{21}$$

Since the objective function of (21) is a random lsc function from $I\!R^m \times \Omega$ to $\overline{I\!R}$, the optimal value $v(P_n(\cdot))$ of (21) is measurable from Ω to $\overline{I\!R}$ and the ε-approximate solution set $S_\varepsilon(P_n(\cdot))$ is a closed-valued measurable set-valued mapping from Ω to $I\!R^m$ (see Chapter 14 and, in particular, Theorem 14.37 of [20]).
Qualitative and quantitative results on the asymptotic behavior of solutions to (21) are given, e.g., in [2, 6, 11] and [10, 13, 14, 16, 27], respectively.
Due to the results in the previous sections, the asymptotic behavior of $v(P_n(\cdot))$ and $S_\varepsilon(P_n(\cdot))$ is closely related to uniform convergence properties of the empirical process

$$\{\sqrt{n}(P_n(\cdot) - P)f = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(f(\xi_i(\cdot)) - Pf)\}_{f\in\mathcal{F}}$$

indexed by the class $\mathcal{F} = \{f_0(x, \cdot) : x \in X\}$. Here, we set $Qf := \int_\Xi f(\xi)Q(d\xi)$ for any $Q \in \mathcal{P}(\Xi)$ and $f \in \mathcal{F}$. Uniform convergence properties refer to the convergence, or to the convergence rate, of

$$d_{\mathcal{F}}(P_n(\cdot), P) = \sup_{f\in\mathcal{F}} |P_n(\cdot)f - Pf| \tag{22}$$

to 0 in terms of some stochastic convergence. Since the supremum in (22) is non-measurable in general, the outer probability $I\!P^*$ is used to describe convergence in probability and almost surely, respectively (cf. [29]).

The class $\mathcal{F}$ is called a *P-Glivenko-Cantelli class* if the sequence $(d_{\mathcal{F}}(P_n(\cdot), P))$ of random variables converges to 0 $I\!P^*$-almost surely or, equivalently, in outer probability. The empirical process is called *uniformly bounded in outer probability with tail $C_{\mathcal{F}}(\cdot)$* if the function $C_{\mathcal{F}}(\cdot)$ is defined on $(0, \infty)$ and decreasing to 0, and the estimate

$$I\!P^*(\{\omega : \sqrt{n}\, d_{\mathcal{F}}(P_n(\omega), P) \geq \varepsilon\}) \leq C_{\mathcal{F}}(\varepsilon)$$

holds for all $\varepsilon > 0$ and $n \in I\!N$.

Whether a given class $\mathcal{F}$ is a P-Glivenko-Cantelli class or the empirical process is uniformly bounded in outer probability, depends on the size of the class $\mathcal{F}$ measured in terms of *bracketing numbers*, or of the corresponding *metric entropy numbers* defined as their logarithms (see [29]). To introduce this concept, let $\mathcal{F}$ be a subset of the normed linear space $L_p(\Xi, P)$ (for some $p \geq 1$) equipped with the usual norm $\|f\|_{P,p} = (P|f|^p)^{\frac{1}{p}}$. The bracketing number $N_{[]}(\varepsilon, \mathcal{F}, L_p(\Xi, P))$ is the minimal number of *brackets* $[l, u] = \{f \in L_p(\Xi, P) : l \leq f \leq u\}$ with $\|l - u\|_{P,p} < \varepsilon$ needed to cover $\mathcal{F}$. The following result provides criteria for the desired properties in terms of bracketing numbers. For its proof we refer to [29, Theorem 2.4.1] and [28, Theorem 1.3].

Theorem 5.1 *Let $\mathcal{F}$ be a class of real-valued functions on Ξ. If*

$$N_{[]}(\varepsilon, \mathcal{F}, L_1(\Xi, P)) < \infty, \tag{23}$$

holds for every $\varepsilon > 0$, 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

$$N_{[]}(\varepsilon, \mathcal{F}, L_2(\Xi, P)) \leq \left(\frac{R}{\varepsilon}\right)^r \tag{24}$$

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

Next we consider the class $\mathcal{F} := \mathcal{F}_\rho$ of integrands defined by (7) in Section 3 and derive conditions implying the assumptions of Theorem 5.1, in particular, the assumptions (23) and (24) for the bracketing numbers $N_{[]}(\varepsilon, \mathcal{F}_\rho, L_p(\Xi, P))$ with $p \in \{1, 2\}$.

Theorem 5.2 *Let the assumptions of Proposition 3.3 be satisfied and $H : I\!R_+ \to I\!R_+$ defined by (11). If $P \in \mathcal{P}_H(\Xi)$, then $\mathcal{F}_\rho = \{f_0(\cdot, x) : x \in X \cap \rho I\!B\}$ is a P-Glivenko-Cantelli class for any $\rho > 0$, i.e.,*

$$\lim_{n\to\infty} \sup_{x \in X \cap \rho I\!B} \left| \int_\Xi f_0(\xi, x) P_n(\omega)(d\xi) - \int_\Xi f_0(\xi, x) P(d\xi) \right| = 0 \quad I\!P - a.s.. \tag{25}$$

If, in addition, Ξ is bounded, then the empirical process indexed by $\mathcal{F}_\rho$ is uniformly bounded in probability with exponential tail, i.e.,

$$I\!P(\{\omega : \sqrt{n} \sup_{x \in X \cap \rho I\!B} \left| \int_\Xi f_0(\xi, x)(P_n(\omega) - P)(d\xi) \right| \geq \varepsilon\}) \leq (K(R)\varepsilon r^{-\frac{1}{2}})^r \exp(-2\varepsilon^2) \tag{26}$$

holds for some constant $K(R) > 0$, any $\varepsilon > 0$ and $n \in I\!N$.

Proof: According to (10) in Proposition 3.3, the functions $f_0(\xi, \cdot)$ satisfy the Lipschitz property

$$f_0(\xi, x) - f_0(\xi, \tilde{x},) \leq \hat{L} \max\{1, H(\|\xi - \xi_0\|)\|\xi - \xi_0\|\}\|x - \tilde{x}\|$$

for all $x, \tilde{x} \in X \cap \rho I\!B$ and $\xi \in \Xi$. Setting $F(\xi) := \hat{L} \max\{1, H(\|\xi - \xi_0\|)\|\xi - \xi_0\|\}$ for all $\xi \in \Xi$, we conclude from [29, Theorem 2.7.11] that

$$N_{[]}(2\varepsilon\|F\|_{P,1}, \mathcal{F}_\rho, L_1(\Xi, P))) \leq N(\varepsilon, X \cap \rho I\!B, I\!R^m) \leq K\varepsilon^{-m} \tag{27}$$

holds for some $K > 0$ and all $\varepsilon > 0$. Since $\|F\|_{P,1}$ is finite, we may replace ε by $\varepsilon/2\|F\|_{P,1}$ in (27) and obtain that $N_{[]}(\varepsilon, \mathcal{F}_\rho, L_1(\Xi, P))$ is finite for all $\varepsilon > 0$. Thus, condition (23) in Theorem 5.1 is satisfied.

If Ξ is bounded, the class $\mathcal{F}_\rho$ is uniformly bounded and condition (24) in Theorem 5.1 is also satisfied due to (27). It remains to note that the supremum $\sup_{x \in X \cap \rho \mathbb{B}}$ may replaced by a supremum with respect to a countable dense subset of $X \cap \rho \mathbb{B}$. Hence, the suprema in (25) and (26) are measurable with respect to $\mathcal{A}$ and, thus, the outer probability $\mathbb{P}^*$ can be replaced by $\mathbb{P}$. ■

When combining the previous result with Theorem 3.2, we arrive at conditions implying a Glivenko-Cantelli result and a large deviation result for the distances of empirical ε-approximate solution sets $S_\varepsilon(P_n(\cdot))$ to $S_\varepsilon(P)$ in case of the two-stage model (6) with random recourse.

6 Conclusions

The quantitative stability results of Section 3 extend earlier work for two-stage models with fixed recourse [16] and for multi-period two-stage models [8]. Theorem 3.2 allows two types of applications. The general version in terms of the semi-distances $d_{\mathcal{F}_\rho}$ makes it possible to utilize metric entropy results and to quantify the asymptotic behavior of statistical approximations to two-stage stochastic programs. The analysis of continuity properties of the integrands f_0 enables to bound the semi-distances by Fortet-Mourier metrics, which are easier to handle due to their dual representations, in particular, for computational purposes (e.g., for scenario reduction [5]).

Acknowledgments

This work was supported by the DFG Research Center Matheon *Mathematics for key technologies* in Berlin and by a NSF grant of the second author.

References

[1] Z. Artstein and R.J-B Wets, Stability results for stochastic programs and sensors, allowing for discontiuous objective functions, *SIAM Journal on Optimization* 4 (1994), 537–550.

[2] Z. Artstein and R.J-B Wets, Consistency of minimizers and the SLLN for stochastic programs, *Journal of Convex Analysis* 2 (1995), 1–17.

[3] H. Attouch and R.J-B Wets, Quantitative stability of variational systems II. A framework for nonlinear conditioning, *SIAM Journal on Optimization* 3 (1993), 359–381.

[4] H. Attouch and R.J-B Wets, Quantitative stability of variational systems III. ε-approximate solutions, *Mathematical Programming* 61 (1993), 197–214.

[5] J. Dupačová, N. Gröwe-Kuska and W. Römisch, Scenario reduction in stochastic programming: An approach using probability metrics, *Mathematical Programming*, Ser. A 95 (2003), 493–511.

[6] J. Dupačová and R.J-B Wets, Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems, *The Annals of Statistics* 16 (1988), 1517–1549.

[7] R. Fortet and E. Mourier, Convergence de la répartition empirique vers la répartition théorique, *Ann. Sci. Ecole Norm. Sup.* 70 (1953), 266–285.

[8] H. Heitsch, W. Römisch and C. Strugarek: Stability of multistage stochastic programs, Preprint 255, DFG Research Center MATHEON Mathematics for key technologies, 2005 and *SIAM Journal on Optimization* (to appear).

[9] P. Kall, On approximations and stability in stochastic programming, in: *Parametric Optimization and Related Topics* (J. Guddat, H.Th. Jongen, B. Kummer, F. Nožička, Eds.), Akademie-Verlag, Berlin 1987, 387–407.

[10] Y.M. Kaniovski, A.J. King and R.J-B Wets, Probabilistic bounds (via large deviations) for the solutions of stochastic programming problems, *Annals of Operations Research* 56 (1995), 189–208.

[11] A.J. King and R.J.-B. Wets, Epi-consistency of convex stochastic programs, *Stochastics and Stochastics Reports* 34 (1991), 83–92.

[12] D. Klatte, On quantitative stability for non-isolated minima, *Control and Cybernetics* 23 (1994), 183–200.

[13] G. Pflug, Stochastic programs and statistical data, *Annals of Operations Research* 85 (1999), 59–78.

[14] G. Pflug, A. Ruszczyński and R. Schultz: On the Glivenko-Cantelli problem in stochastic programming: Linear recourse and extensions, *Mathematics of Operations Research* 23 (1998), 204–220.

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

[16] S.T. Rachev and W. Römisch, Quantitative stability in stochastic programming: The method of probability metrics, *Mathematics of Operations Research* 27 (2002), 792–818.

[17] S.M. Robinson and R.J-B Wets, Stability in two-stage stochastic programming, *SIAM Journal on Control and Optimization* 25 (1987), 1409–1416.

[18] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton 1970.

[19] R.T. Rockafellar, Integral functionals, normal integrands and measurable selections, in: *Nonlinear Operators and the Calculus of Variations* (J. Gossez and L. Waelbroeck Eds.), Lecture Notes in Mathematics 543, Springer, Berlin 1976, 157–207.

[20] R.T. Rockafellar and R.J-B Wets, *Variational Analysis*, Springer, 2004 (2nd edition).

[21] W. Römisch, Stability of stochastic programming problems, in: *Stochastic Programming* (A. Ruszczyński and A. Shapiro Eds.), Handbooks of Operations Research and Management Science, Volume 10, Elsevier, Amsterdam 2003, 483–554.

[22] W. Römisch and R. Schultz, Stability analysis for stochastic programs, *Annals of Operations Research* 30 (1991), 241–266.

[23] W. Römisch and R. Schultz, Lipschitz stability for stochastic programs with complete recourse, *SIAM Journal on Optimization*, 6 (1996), 531–547.

[24] W. Römisch and A. Wakolbinger, Obtaining convergence rates for approximations in stochastic programming, in: *Parametric Optimization and Related Topics* (J. Guddat, H.Th. Jongen, B. Kummer, F. Nožička, Eds.), Akademie Verlag, Berlin 1987, 327–343.

[25] R. Schultz, Strong convexity in stochastic programs with complete recourse, *Journal of Computational and Applied Mathematics* 56 (1994), 3–22.

[26] A. Shapiro, Quantitative stability in stochastic programming, *Mathematical Programming* 67 (1994), 99–108.

[27] A. Shapiro: Monte Carlo sampling methods, in: *Stochastic Programming* (A. Ruszczyński and A. Shapiro Eds.), Handbooks of Operations Research and Management Science, Volume 10, Elsevier, Amsterdam 2003, 353–425.

[28] M. Talagrand, Sharper bounds for Gaussian and empirical processes, *Annals of Probability* 22 (1994), 28–76.

[29] A.W. van der Vaart and J.A. Wellner, *Weak Convergence and Empirical Processes*, Springer, New York 1996.

[30] W. Vervaat. Random upper semicontinuous functions and extremal processes. Report MS-R8801, Center for Wiskunde en Informatica, Amsterdam, 1988.

[31] D.W. Walkup and R.J-B Wets, Lifting projections of convex polyhedra, *Pacific Journal of Mathematics* 28 (1969), 465–475.

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

[33] R.J-B Wets, On the continuity of the value of a linear program and of related polyhedral-valued multifunctions, *Mathematical Programming Study* 24 (1985), 14–29.

[34] R.J-B Wets, Challenges in stochastic programming, *Mathematical Programming* 75 (1996), 115–135.

[35] R.J-B Wets, Lipschitz continuity of inf-projections, *Computational Optimization and Applications* 25 (2003), 269–282.