SIAM J. OPTIM.
Vol. 18, No. 3, pp. 961–979

STABILITY OF ε-APPROXIMATE SOLUTIONS TO CONVEX STOCHASTIC PROGRAMS*

W. RÖMISCH† AND R. J.-B. WETS‡

Abstract. An analysis of convex stochastic programs is provided when 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 continuously 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. We discuss the consequences 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, approximate solutions, probability metrics, two-stage models, random recourse

AMS subject classifications. 90C15, 90C31

DOI. 10.1137/060657716

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

$$\min\left\{\mathbb{E}^P f_0(x) := \int_\Xi f_0(\xi, x)\, P(d\xi) : x \in X\right\}, \tag{1.1}$$

where X is a closed convex subset of $\mathbb{R}^m$, Ξ is 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 lower semicontinuous (lsc) function*,[1] and, in particular, this means

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

*Received by the editors April 20, 2006; accepted for publication (in revised form) April 5, 2007; published electronically October 4, 2007. This work was supported by the DFG Research Center MATHEON *Mathematics for key technologies* in Berlin and by an NSF grant of the second author.
http://www.siam.org/journals/siopt/18-3/65771.html

†Humboldt-University Berlin, Institute of Mathematics, D-10099 Berlin, Germany (romisch@math.hu-berlin.de).
‡University of California at Davis, Department of Mathematics, Davis, CA 95616-8633 (rjbwets@ucdavis.edu).

[1] The concept of a random lsc function is due to Rockafellar [21], 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 [23, Chapter 14], [33].

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 infimum, the solution set, and the set of ε-approximate solutions to (1.1), i.e.,

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

Since, in practice, the underlying probability distribution P is often not known precisely, the stability behavior of the stochastic program (1.1) 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 subsets of $\mathcal{P}(\Xi)$ equipped with some probability (semi)metric.

Early work on stability of stochastic programs is reported in [11, 19, 27] and later in [1]. Quantitative stability of two-stage models was studied, e.g., in [25, 26, 29, 18]. A recent survey of stability results in stochastic programming is given in [24]. Most of the recent contributions to (quantitative) stability use the general framework and the results of [3, 14] and [23, Chapter 7J], respectively.

In the present paper, we take up an issue brought to the fore in [38, section 4]. Since solutions derived, when actually solving (1.1), 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 [23, Chapter 7J]. The corresponding estimates make use of the epi-distance between the objective functions of (1.1) 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 semimetric of the form

$$d_{\mathcal{F}}(P,Q) = \sup \left\{ \left| \int_\Xi f(\xi) P(d\xi) - \int_\Xi f(\xi) Q(d\xi) \right| \,:\, f \in \mathcal{F} \right\}, \tag{1.2}$$

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 $\mathbb{B}$ denoting the unit ball in $\mathbb{R}^m$ 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. Two-stage stochastic programs arise as deterministic equivalents of improperly posed random linear programs of the form

$$\min\{ cx : x \in X, T(\xi)x = h(\xi) \},$$

where X is polyhedral and the (technology) matrix $T(\xi)$ and the vector $h(\xi)$ depend on a random vector ξ. Given a realization of ξ, a possible deviation $h(\xi) - T(\xi)x$

is compensated for by the additional cost $q(\xi)y(\xi)$, where $y = y(\xi)$ belongs to a polyhedral set Y and satisfies $W(\xi)y = h(\xi) - T(\xi)x$. Here, the cost coefficient $q(\xi)$ and the compensation or recourse matrix $W(\xi)$ (may) depend on the realization. The modeling idea consists in adding the expected compensation cost $\mathbb{E}[q(\xi)y(\xi)]$ to cx. By minimizing the objective function $cx + \mathbb{E}[q(\xi)y(\xi)]$ first with respect to $y(\xi)$, we arrive at the function

$$f_0(\xi, x) := cx + \inf\{q(\xi)y : y \in Y, W(\xi)y = h(\xi) - T(\xi)x\},$$

whose expectation has to be minimized with respect to $x \in X$. Since the decisions x and $y(\xi)$ are made before or after the realization of ξ, they are called first- and second-stage decisions, respectively.

While Lipschitz continuity properties of the integrands f_0 with respect to ξ are well understood for fixed recourse [36], much less is known for random recourse. In section 3 we deal with the following two cases: (i) full random recourse by imposing local Lipschitz continuity of the (second-stage) dual feasibility mapping and (ii) a specific lower diagonal randomness of the recourse matrix. The latter situation occurs, for example, in the following two important cases.

Let us first consider a dynamical decision process, as in a variety of applications, where the compensation idea is repeated l times after the realization of a new random vector ξ_j, $j = 1, \ldots, l$. Then we have second-stage decisions $y_j = y_j(\xi_j)$ with corresponding cost $q_j(\xi_j)y_j$ which satisfy the constraints $y_j \in Y_j$ and $W_{jj}y_j = h_j(\xi_j) - W_{jj-1}(\xi_j)y_{j-1}$ for $j = 1, \ldots, l$, where $l \in \mathbb{N}$, y_0 is the first-stage decision and $W_{jj-1}(\xi_j)$ are (random) technology matrices. This leads to the function

$$f_0(\xi, y_0) = cy_0 + \inf\left\{\sum_{j=1}^{l} q_j(\xi)y_j : W_{jj}y_j = h_j(\xi) - W_{jj-1}(\xi)y_{j-1}, y_j \in Y_j, j = 1, \ldots, l\right\},$$

where $\xi = (\xi_1, \ldots, \xi_l)$ and $q_j(\xi) := q_j(\xi_j)$, etc. The expectation of this function is to be minimized in multiperiod two-stage stochastic programming models. If we introduce the second-stage decision vector $y = (y_1, \ldots, y_l)$, the corresponding recourse matrix $W(\xi)$ is a block lower triangular matrix containing W_{jj}, $j = 1, \ldots, l$, in the main diagonal and $W_{jj-1}(\xi)$, $j = 1, \ldots, l$, in the lower diagonal (see section 4). Hence, the recourse matrix $W(\xi)$ may be random even if the jth recourse matrix W_{jj} for the decision y_j is fixed, but (at least) one of the technology matrices $W_{jj-1}(\xi)$ is random.

Another interesting case appears, second, in risk averse two-stage stochastic programming models, if risk functionals (e.g., the conditional value-at-risk [22]) are incorporated into two-stage stochastic programs. The conditional or *average value-at-risk* (at level $\alpha \in (0, 1]$) may be defined by

$$\begin{aligned} AVaR_\alpha(z) &= \frac{1}{\alpha}\int_0^\alpha VaR_\gamma(z)d\gamma = \inf\left\{r + \frac{1}{\alpha}\mathbb{E}[\max\{0, -r - z\}] : r \in \mathbb{R}\right\} \\ &= \inf\left\{r_1 + \frac{1}{\alpha}\mathbb{E}[r_2^{(2)}] : r_1 \in \mathbb{R}, r_2 \in \mathbb{R}_+ \times \mathbb{R}_+, r_2^{(1)} - r_2^{(2)} = z + r_1\right\}, \end{aligned} \tag{1.3}$$

where z is a real random variable on some probability space. If the average value-at-risk replaces the expectation in a two-stage model with fixed recourse, the latter is of the form

$$\min\left\{cx + AVaR_\alpha(q(\xi)y) : x \in X, y \in Y, Wy = h(\xi) - T(\xi)x\right\}. \tag{1.4}$$

Using the two-stage representation (1.3) of $AVaR_\alpha$, the preceding optimization problem is equivalent to (1.1) with

$$f_0(\xi,(x,r_1)) := cx + r_1 + \inf\left\{ \frac{1}{\alpha} r_2^{(2)} : y \in Y, r_2 \geq 0, r_2^{(1)} - r_2^{(2)} = q(\xi)y + r_1, \right.$$
$$\left. Wy = h(\xi) - T(\xi)x \right\},$$

where (x,r_1) is the first-stage decision varying in $X \times \mathbb{R}$. When introducing the second-stage decision (y,r_2), the recourse cost $q_{\mathrm{avar}}(\xi)$, recourse matrix $W_{\mathrm{avar}}(\xi)$, and cone Y_{avar} take on the form

(1.5)

$$q_{\mathrm{avar}}(\xi) = \begin{pmatrix} 0 \\ 0 \\ \alpha^{-1} \end{pmatrix}, \quad W_{\mathrm{avar}}(\xi) = \begin{pmatrix} W & 0 & 0 \\ q(\xi)^\top & -1 & 1 \end{pmatrix}, \quad \text{and} \quad Y_{\mathrm{avar}} = Y \times \mathbb{R}^2_+ .$$

Hence, the recourse matrix gets random if the recourse cost of the original model is random. The same lower diagonal randomness effect appears if general polyhedral convex risk measures are used instead of $AVaR$ (see [7, section 4.1.1]).

In sections 3 and 4 we characterize the local Lipschitz continuity behavior of the functions $\mathcal{F}_\rho$. We also 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 3.6 and 4.3 for the cases (i) and (ii), respectively) 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 (1.2). Our analysis will be based on the general perturbation results for optimization models in [23, 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 (1.1) is well defined and one can proceed with the perturbation analysis. We consider

$$\mathcal{F} = \{f_0(\cdot,x) : x \in X\},$$
$$\mathcal{P}_\mathcal{F} = \left\{ Q \in \mathcal{P}(\Xi) : \int_\Xi \inf_{x \in X \cap \rho\mathbb{B}} f_0(\xi,x) Q(d\xi) > -\infty , \right.$$
$$\left. \sup_{x \in X \cap \rho\mathbb{B}} \int_\Xi f_0(\xi,x) Q(d\xi) < \infty \quad \forall\, \rho > 0 \right\},$$

where $\mathbb{B}$ is 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. [23, 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

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

Hence, $d_{\mathcal{F},\rho}$ is a distance of type (1.2), 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, and symmetric and satisfies the triangle inequality; i.e., it is a semimetric on $\mathcal{P}_{\mathcal{F}}$. In general, however, the class $\mathcal{F}_\rho$ will not be rich enough to guarantee that $d_{\mathcal{F},\rho}(P,Q)=0$ implies $P=Q$. A valuable consequence of the definition of the class $\mathcal{P}_{\mathcal{F}}$ is that the function $x\mapsto \mathbb{E}^Q f_0(x) = \int_\Xi f_0(\xi,x)\,Q(d\xi)$ is lsc at any Q belonging to $\mathcal{P}_{\mathcal{F}}$ by appealing to Fatou's lemma. Moreover, it is convex on $\mathbb{R}^m$ and finite on X for any such Q.

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 nonempty closed set. The *ρ-distance* between two nonempty closed sets is by definition

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

In fact, it is just a pseudodistance 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:

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

Indeed one always has [23, Proposition 4.37(a)]

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

for $\rho'\ge 2\rho+\max\{d_{C_1}(0),d_{C_2}(0)\}$. Our main result is stated in terms of this latter distance notion. If we let $\rho\to\infty$, we end up with the Pompeiu–Hausdorff distance

$$d\!l_\infty(C,D)=\lim_{\rho\to\infty}d\!l_\rho(C,D)=\lim_{\rho\to\infty}\hat{d\!l}_\rho(C,D)$$

between the closed nonempty sets C and D; see [23, Corollary 4.38].

The distance between (lsc) functions is measured in terms of the distance between their epi-graphs, 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 epi-graphs (in $\mathbb{R}^{m+1}$), it is convenient to rely on the "unit ball" to be $\mathbb{B}\times[-1,1]$; this brings us to an "auxiliary" distance $\hat{d\!l}^+_\rho(f_1,f_2)$ defined as the infimum of all $\eta\ge 0$ such that for all $x\in\rho\mathbb{B}$,

$$\min_{y\in\mathbb{B}(x,\eta)}f_2(y)\le\max\{f_1(x),-\rho\}+\eta,\qquad \min_{y\in\mathbb{B}(x,\eta)}f_1(y)\le\max\{f_2(x),-\rho\}+\eta.$$

For lsc $f_1,\ f_2:\mathbb{R}^n\to\mathbb{R}$, not identically ∞, one has [23, Theorem 7.61]

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

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

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

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

hold 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* (1.1)*; more precisely,*

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

with

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

Proof. For any $Q \in \mathcal{P}_{\mathcal{F}}$, the function $\mathbb{E}^Q f_0$ is lsc, 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 [23, 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 [23] yields the estimates

$$\begin{aligned} |v(P) - v(Q)| &\leq \hat{d\kern-0.4em l}_\rho^+(\mathbb{E}^P f_0, \mathbb{E}^Q f_0), \\ \emptyset \neq S(Q) &\subseteq S(P) + \Psi_P(\hat{d\kern-0.4em l}_\rho^+(\mathbb{E}^P f_0, \mathbb{E}^Q f_0))\mathbb{B} \end{aligned}$$

for any $Q \in \mathcal{P}_{\mathcal{F}}$ with $\hat{d\kern-0.4em l}_\rho^+(\mathbb{E}^P f_0, \mathbb{E}^Q f_0) < \delta$.

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

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

and, thus, $\hat{d\kern-0.4em l}_\rho^+(F_P, F_Q) \leq \eta$. Letting η pass to its lower limit leads to

$$\hat{d\kern-0.4em l}_\rho^+(F_P, F_Q) \leq \max_{x \in X \cap \rho\mathbb{B}} |\mathbb{E}^P f_0(x) - \mathbb{E}^Q f_0(x)| = d_{\mathcal{F},\rho}(P,Q). \tag{2.1}$$

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. [24, Example 26]). Furthermore, explicit descriptions of conditioning functions ψ_P of stochastic programs (like linear or quadratic growth at solution sets) are known only 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 [28].

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 nonempty. 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 $\mathbb{E}^P f_0$ and $\mathbb{E}^Q f_0$ are proper, lsc, and convex on $\mathbb{R}^m$. Let ρ_0 be chosen such that both $S(P) \cap \rho_0 \mathbb{B}$ and $S(Q) \cap \rho_0 \mathbb{B}$ are nonempty 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 [23, Theorem 7.69], the inclusion

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

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

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

The same argument works with P and Q interchanged. Finally, we appeal to the estimate (2.1) to complete the proof. $\square$

The above estimate for ε-approximate solution sets allows for the solution sets to be unbounded and, thus, extends [24, Theorem 13]. 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 nonempty and 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 nonempty and bounded. Since the functions $\mathbb{E}^P f_0$ and $\mathbb{E}^Q f_0$ are lsc 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. [20, 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 $\hat{\rho}$. $\square$

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 (1.1) 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 $\mathcal{P}_{\mathcal{F}}$ is 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 $\{\xi \in \Xi : \|\xi - \xi_0\| \leq r\}$ for some $\xi_0 \in \Xi$ and any $r > 0$. In many cases, the growth modulus $L(r)$ does not depend on x, particularly when x is varying only in a bounded subset of $\mathbb{R}^m$. Hence, function classes of the form

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

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 (1.2), but with $\mathcal{F} = \mathcal{F}_H$, i.e.,

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

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

$$\mathcal{P}_H(\Xi) := \left\{Q \in \mathcal{P}(\Xi) : \int_\Xi \max\{1, H(\|\xi - \xi_0\|)\}\|\xi - \xi_0\|Q(d\xi) < \infty\right\} \tag{2.2}$$

(cf. [8, 17]). 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

$$\mathcal{P}_r(\Xi) := \left\{Q \in \mathcal{P}(\Xi) : \int_\Xi \|\xi\|^r Q(d\xi) < \infty\right\} \tag{2.3}$$

of probability measures having finite rth order moments.

3. Stability of two-stage recourse models. We consider the linear two-stage stochastic program with recourse,

$$\min\left\{cx + \int_\Xi q(\xi)y(\xi)P(d\xi) : W(\xi)y(\xi) = h(\xi) - T(\xi)x,\ y(\xi) \in Y,\ x \in X\right\}, \tag{3.1}$$

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 (3.1) may be rewritten equivalently as a minimization problem with respect to the first stage decision x. We define the function $f_0 : \Xi \times \mathbb{R}^m \to \overline{\mathbb{R}}$ by

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

where the optimal value function Φ and the dual feasible set $D(\xi)$ are given by

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

with $W(\xi)^\top$ denoting the transpose of $W(\xi)$ and Y^* the polar cone of Y.

The (equivalent) minimization problem can thus be expressed as

$$\min\left\{\int_\Xi f_0(\xi, x)P(d\xi) :\ x \in X\right\}. \tag{3.2}$$

In order to utilize the general stability results of section 2, we first recall some well-known properties of the function Φ (cf. [34]).

LEMMA 3.1. *For any $\xi \in \Xi$, the function $\Phi(\xi, \cdot, \cdot)$ is finite and continuous on the polyhedral set $\mathcal{D}(\xi) \times W(\xi)Y$, where $\mathcal{D}(\xi) := \{u \in \mathbb{R}^{\overline{m}} : \{z \in \mathbb{R}^d : W(\xi)^\top z - u \in Y^*\} \neq \emptyset\}$. Furthermore, the function $\Phi(\xi, u, \cdot)$ is piecewise linear convex on the polyhedral set $W(\xi)Y$ for fixed $u \in \mathcal{D}(\xi)$, and $\Phi(\xi, \cdot, t)$ is piecewise linear concave on $\mathcal{D}(\xi)$ for fixed $t \in W(\xi)Y$.*

We impose the following conditions on problem (3.2).

(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 and render problem (3.2) 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

$$\mathcal{F}_\rho := \{f_0(\cdot, x) : x \in X \cap \rho\mathbb{B}\}, \tag{3.3}$$

we obtain our first stability result for model (3.1) 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 nonempty and bounded. Then there exist constants $\rho > 0$ and $\hat{\varepsilon} > 0$ such that*

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

hold 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. It extends the results in [24, section 3.1] to two-stage models with *random* recourse. 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 $W(\xi) \equiv W$ [36], and have been used to analyze quantitative stability in [18]. Our first result for random recourse matrices follows the ideas in [37]. 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 $\mathbb{R}^d$); i.e., there exists a constant $L > 0$, an element $\xi_0 \in \Xi$, and a nondecreasing function $h : \mathbb{R}_+ \to \mathbb{R}_+$ with $h(0) = 0$ such that*

$$d\!\!l_\infty(D(\xi), D(\tilde{\xi})) \leq L \max\{1, h(\|\xi - \xi_0\|), h(\|\tilde{\xi} - \xi_0\|)\}\|\xi - \tilde{\xi}\| \tag{3.4}$$

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

$$f_0(\xi, x) - f_0(\tilde{\xi}, x) \le \hat{L}(\rho) \max\{1, H(\|\xi - \xi_0\|), H(\|\tilde{\xi} - \xi_0\|)\}\|\xi - \tilde{\xi}\|, \tag{3.5}$$
$$f_0(\xi, x) - f_0(\xi, \tilde{x}) \le \hat{L} \max\{1, H(\|\xi - \xi_0\|)\|\xi - \xi_0\|\}\|x - \tilde{x}\| \tag{3.6}$$

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

$$H(t) := h(t)t \quad \forall t \in \mathbb{R}_+. \tag{3.7}$$

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 $x, \tilde{x} \in X \cap \rho\mathbb{B}$, and $\xi, \tilde{\xi} \in \Xi$, one has the estimate

$$f_0(\xi, x) - f_0(\tilde{\xi}, \tilde{x}) \le cx + (h(\xi) - T(\xi)x)z^*(\xi) - (h(\tilde{\xi}) - c\tilde{x} - T(\tilde{\xi})\tilde{x})z(\tilde{\xi}), \tag{3.8}$$

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

$$\|z^*(\xi) - \bar{z}(\tilde{\xi}; \xi)\| \le \mathbb{d}_\infty(D(\xi), D(\tilde{\xi})) \le L \max\{1, h(\|\xi - \xi_0\|), h(\|\tilde{\xi} - \xi_0\|)\}\|\xi - \tilde{\xi}\|. \tag{3.9}$$

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 (3.4), we have

$$\|\bar{z}(\tilde{\xi}; \xi)\| \le \max\{r, L\} \max\{1, h(\|\tilde{\xi} - \xi_0\|)\}\|\tilde{\xi} - \xi_0\|. \tag{3.10}$$

Now, we proceed with our estimate (3.8) when $x = \tilde{x}$, exploiting the affine linearity of $h(\cdot)$ and $T(\cdot)$, (3.9) and (3.10). 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, (3.5) is proved with $\hat{L}(\rho) = \bar{L}(1 + \rho)$. Finally, we return to (3.8) 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 $\xi \in \Xi$, $x, \tilde{x} \in X \cap \rho\mathbb{B}$. Here, we used that $\|z^*(\xi)\|$ can be bounded in the same way as $\bar{z}(\tilde{\xi};\xi)$ in (3.10). ◻

The next examples illustrate the local Lipschitz continuity property (3.4) of the dual feasibility mapping D.

Example 3.4. Let $\overline{m} = 4$, $d = 2$, $Y = \mathbb{R}^4_+$, and $\Xi = \mathbb{R}$, and consider the random recourse costs and 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 $W(\xi)Y = \mathbb{R}^2$ (complete recourse) and $D(\xi) = [0, \xi^2] \times \{\xi\}$. Hence, the conditions (A1) and (A2) and (3.4) are satisfied with $h(t) \equiv t$.

Example 3.5. We consider the second-stage program arising in the equivalent optimization problem to AVaR minimization (1.4) in section 1. Its dual feasible set is of the form

$$\begin{aligned} D_{\mathrm{avar}}(\xi) &= \left\{ z = (z_1, z_2) \in \mathbb{R}^d \times \mathbb{R} : W_{\mathrm{avar}}(\xi)^\top z - q_{\mathrm{avar}}(\xi) \in Y^*_{\mathrm{avar}} \right\} \\ &= \left\{ (z_1, z_2) \in \mathbb{R}^d \times [0, \alpha^{-1}] : W^\top z_1 + q(\xi) z_2 \in Y^* \right\} \\ &= \left\{ (z_1, u) \in \mathbb{R}^d \times \mathbb{R}^{\overline{m}} : W^\top z_1 + u \in Y^*, u \in [0, \alpha^{-1}] q(\xi) \right\} \end{aligned}$$

due to (1.5), where $u \in [0, \alpha^{-1}]q(\xi)$ means that, for every $j = 1, \dots, \overline{m}$, $0 \le u_j \le \alpha^{-1} q_j(\xi)$ holds if $q_j(\xi) \ge 0$ and $\alpha^{-1} q_j(\xi) \le u_j \le 0$ otherwise. Hence, if (A2) is satisfied, the set-valued mapping $\xi \to D_{\mathrm{avar}}(\xi)$ is Lipschitz continuous on Ξ with respect to the Pompeiu–Hausdorff distance $\mathbb{d}_\infty$ since its graph is convex polyhedral [35]. This means that Proposition 3.3 applies with $h(t) \equiv 1$.

We can reformulate the conclusions of the preceding proposition in terms of the Fortet–Mourier metrics defined on $\mathcal{P}_H(\Xi)$, the space (2.2) of probability measures.

COROLLARY 3.6. *Let the assumptions of Proposition* 3.3 *be satisfied,* $P \in \mathcal{P}_H(\Xi)$*, and* $S(P)$ *be nonempty and bounded. Then there exist constants* $\hat{L} > 0$*,* $\rho > 0$*, and* $\hat{\varepsilon} > 0$ *such that*

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

hold 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* (3.7)*, and* $\zeta_H(P, Q)$ *is the Fortet–Mourier metric on* $\mathcal{P}_H(\Xi)$*.*

Proof. The estimate (3.5) implies $d_{\mathcal{F},\rho}(P, Q) \le \hat{L}\zeta_H(P, Q)$ 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 $\mathbb{d}_\infty$ due to [35]. Hence, $H(t) \equiv t$ and $\mathcal{F}_H = \mathcal{F}_2$, and then the previous result boils down to [18, Proposition 3.2].

4. Two-stage multiperiod models. If the second stage of a stochastic program with recourse models a (stochastic) dynamical decision process (see section 1), our two-stage problem takes on the form

(4.1)
$$\min \left\{ c y_0 + \sum_{j=1}^{l} q_j(\xi) y_j : y_0 \in X, y_j \in Y_j, W_{jj} y_j = h_j(\xi) - W_{jj-1}(\xi) y_{j-1}, j = 1, \dots, l \right\},$$

where for $j = 1, \ldots, l$, $Y_j \in \mathbb{R}^{\overline{m}_j}$ 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 $y = (y_1, \ldots, y_l)$, the polyhedral set Y is the Cartesian product of polyhedral sets $Y_j \in \mathbb{R}^{\overline{m}_j}$, $j = 1, \ldots, l$, the element $T(\xi)x$ has the components $T_1(\xi)x := W_{10}(\xi)x$ and $T_j(\xi)x = 0$, $j = 2, \ldots, l$, and the random recourse matrix $W(\xi)$ is of the form

$$W(\xi) = \begin{pmatrix} W_{11} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ W_{21}(\xi) & W_{22} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & W_{32}(\xi) & W_{33} & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\vdots & \\ 0 & 0 & 0 & 0 & \cdots & W_{l-1l-2}(\xi) & W_{l-1l-1} & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & W_{ll-1}(\xi) & W_{ll} \end{pmatrix}, \tag{4.2}$$

i.e., all matrices W_{jj}, $j = 1, \ldots, l$, in the diagonal of $W(\xi)$ are nonstochastic. 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

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

where the function Ψ_1 is given by the recursion

$$\Phi_j(\xi, u_{j-1}) := \inf \left\{ q_j(\xi) y_j + \Psi_{j+1}(\xi, y_j) : W_{jj} y_j = u_{j-1}, y_j \in Y_j \right\}, \tag{4.3}$$

$$\Psi_j(\xi, y_{j-1}) := \Phi_j(\xi, h_j(\xi) - W_{jj-1}(\xi) y_{j-1}) \tag{4.4}$$

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

While the continuity and growth properties of the function $f_0(\cdot, x)$ in case $l = 1$ may be derived from Lemma 3.1, we need an extended result for establishing Lipschitz continuity properties of the inf-projection Φ_j for $j = 1, \ldots, l$. The results in [39] were developed precisely to deal with the present situation. To state the result, we denote by D^∞ the recession 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 [39, 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 $\{(u_1, y) \in \mathbb{R}^n \times Y : \|u_1\| \le r, \|y\| \le r\}$ with constant $L(r)$ for every $r > 0$. Then, $\Phi(\cdot)$ is Lipschitz continuous on $\{(u_1, u_2) \in \operatorname{dom} \Phi : \|u_1\| \le r, \|u_2\| \le r\}$ with constant $L_M L(K_M \max\{1, r\})$ for every $r > 0$, where $L_M \ge 1$ and $K_M \ge 1$ 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 $d\!\!l_\infty$ (with constant $L_M \ge 1$) since its graph is polyhedral [23, 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 $u = (u_1, u_2), \tilde{u} = (\tilde{u}_1, \tilde{u}_2) \in \operatorname{dom} \Phi \cap \{(u_1, u_2) \in \mathbb{R}^n \times \mathbb{R}^d : \|u_1\| \le r, \|u_2\| \le r\}$. 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 $\|y(u_2) - y(\tilde{u}_2)\| \le L_M \|u_2 - \tilde{u}_2\|$. In particular, there exists a constant $K_M \ge 1$ such that

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

We obtain

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

and that completes the proof. □

PROPOSITION 4.2. *Let $W(\xi)$ be as described by* (4.2). *Assume the relatively complete recourse condition* (A1) *is satisfied and that* $\ker(W_{jj}) \cap Y_j^\infty = \{0\}$ *for* $j = 1, \ldots, l-1$. *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 $x, \tilde{x} \in X \cap \rho\mathbb{B}$:*

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

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 $v_j \in \mathbb{R}^{r_j}$ and $j = 1, \ldots, l$. Furthermore, the set-valued mappings M_j from $\mathbb{R}^{r_j}$ to $\mathbb{R}^{\overline{m}_j}$ 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 $j = 2, \ldots, l$. 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 (4.3) is solvable. First we consider the functions Φ_l and Ψ_l:

$$\begin{aligned} \Phi_l(\xi, u_{l-1}) &= \inf\{q_l(\xi) y_l : W_{ll} y_l = u_{l-1}, y_l \in Y_l\}, \\ \Psi_l(\xi, y_{l-1}) &= \Phi_l(\xi, h_l(\xi) - W_{ll-1}(\xi) y_{l-1}). \end{aligned}$$

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

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

on the set $\{(\xi, x) \in \Xi \times X : \|\xi\| \le r, \|x\| \le \rho\}$.

This yields the assertions about f_0 and completes the proof. $\square$

Due to the previous result we obtain

$$\mathcal{P}_{\mathcal{F}} \supseteq \mathcal{P}_{l+1}(\Xi) = \{Q \in \mathcal{P}(\Xi) : \int_{\Xi} \|\xi\|^{l+1} Q(d\xi) < \infty\}$$
$$\text{and} \quad \frac{1}{L \max\{1, \rho\}} f_0(x, \cdot) \in \mathcal{F}_{l+1}(\Xi)$$

for each $x \in X \cap \rho\mathbb{B}$, and arrive, after specializing Theorem 3.2, at the following.

COROLLARY 4.3. *Let $W(\xi)$ be as described by (4.2). Assume the relatively complete recourse condition* (A1) *is satisfied and that* $\ker(W_{jj}) \cap Y_j^{\infty} = \{0\}$ *for* $j = 1, \ldots, l-1$.

Then there exist constants $L > 0$ and $\hat{\varepsilon} > 0$ such that for any $\varepsilon \in (0, \hat{\varepsilon})$ the estimates

$$|v(P) - v(Q)| \le L\, \zeta_{l+1}(P, Q),$$
$$d\!\!\!\!\;l_{\infty}(S_{\varepsilon}(P), S_{\varepsilon}(Q)) \le \frac{L}{\varepsilon} \zeta_{l+1}(P, Q)$$

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

The case $l = 1$ corresponds to the situation of two-stage models with fixed recourse, and that situation was already covered by [24, Theorem 24]. We note that the corollary remains valid for the slightly more general situation that $W_{jj-1}(\xi) y_{j-1}$ in (4.1) is replaced by $\sum_{i=1}^{j-1} W_{ji}(\xi) y_i$, and, hence, all lower diagonal blocks of $W(\xi)$ are random. We also note that the corollary applies to recourse matrices of the form (1.5) in risk averse two-stage models with polyhedral convex risk functionals.

If the recent stability result [10, Theorem 2.1] for linear multistage models is restricted to the two-stage model (4.1), it implies the existence of positive constants L and δ such that

$$|v(P) - v(Q)| \le L\, \ell_{l+1}(P, Q) \tag{4.5}$$

holds for every $Q \in \mathcal{P}_{l+1}(\Xi)$ with $\ell_{l+1}(P, Q) < \delta$; the distance ℓ_r denotes the L_r-*minimal* or Wasserstein metric

$$\ell_r(P, Q) := \left(\inf \left\{ \int_{\Xi \times \Xi} \|\xi - \tilde{\xi}\|^r \eta(d\xi, d\tilde{\xi}) : \eta \in \mathcal{P}(\Xi \times \Xi),\ \pi_1 \eta = P,\ \pi_2 \eta = Q \right\} \right)^{1/r} \tag{4.6}$$

on $\mathcal{P}_r(\Xi)$ for any $r \ge 1$, where π_1 and π_2 denote the projections onto the first and second components, respectively. It is known that sequences in $\mathcal{P}_r(\Xi)$ converge with respect to both metrics ζ_r and ℓ_r if they converge weakly and if their rth 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 (4.6). Such solutions exist according to [17, Theorem 8.1.1]. Then the duality theorem [17, Theorem 5.3.2] for the Fortet–Mourier metric of order r implies, via Hölder's inequality, the estimate

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

Since the convergence of probability measures with respect to ℓ_r and ζ_r implies the convergence of their rth order absolute moments, the stability result for optimal values obtained in Corollary 4.3 implies (4.5) (with some constant $L>0$). However, the convergence of $\zeta_r(P,P_n)$ to 0 may be faster than $\ell_r(P,P_n)$ for some sequence (P_n) of probability measures, as illustrated in [18, Example 3.4]. Hence, the stability result for optimal values in Corollary 4.3 strictly extends the estimate (4.5) for multiperiod two-stage stochastic programs.

5. Empirical approximations of two-stage models. Let $\xi_1,\xi_2,\dots,\xi_n,\dots$ be independent and identically distributed Ξ-valued random variables on some probability space $(\Omega,\mathcal{A},\mathbb{P})$ having the common distribution P, i.e., $P=\mathbb{P}\xi_1^{-1}$. We consider the empirical measures

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

and the *empirical approximation* of the stochastic program (1.1) with sample size n, i.e.,

$$\min\left\{\frac{1}{n}\sum_{i=1}^n f_0(\xi_i(\cdot),x) : x\in X\right\}. \tag{5.1}$$

Since the objective function of (5.1) is a random lsc function from $\mathbb{R}^m\times\Omega$ to $\overline{\mathbb{R}}$, the optimal value $v(P_n(\cdot))$ of (5.1) is measurable from Ω to $\overline{\mathbb{R}}$ and the ε-approximate solution set $S_\varepsilon(P_n(\cdot))$ is a closed-valued measurable set-valued mapping from Ω to $\mathbb{R}^m$ (see Chapter 14 and, in particular, Theorem 14.37 of [23]).

Qualitative and quantitative results on the asymptotic behavior of solutions to (5.1) are given, e.g., in [2, 6, 13] and [12, 15, 16, 18, 30], 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

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

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

$$d_{\mathcal{F}}(P_n(\cdot),P) = \sup_{f\in\mathcal{F}}|P_n(\cdot)f-Pf| \tag{5.2}$$

to 0 in terms of some stochastic convergence. Since the supremum in (5.2) is nonmeasurable in general, the outer probability $\mathbb{P}^*$ (defined by $\mathbb{P}^*(B) = \inf\{\mathbb{P}(A) : B \subset A, A \in \mathcal{A}\}$ for any subset B of Ω) is used to describe convergence in probability and almost surely, respectively (cf. [32]).

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 $\mathbb{P}^*$-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

$$\mathbb{P}^*(\{\omega : \sqrt{n}\, d_{\mathcal{F}}(P_n(\omega), P) \geq \varepsilon\}) \leq C_{\mathcal{F}}(\varepsilon) \tag{5.3}$$

holds for all $\varepsilon > 0$ and $n \in \mathbb{N}$.

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 [32]). 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 [32, Theorem 2.4.1] and [31, Theorem 1.3].

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

$$N_{[\,]}(\varepsilon, \mathcal{F}, L_1(\Xi, P)) < \infty \tag{5.4}$$

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

$$N_{[\,]}(\varepsilon, \mathcal{F}, L_2(\Xi, P)) \leq \left(\frac{R}{\varepsilon}\right)^r \tag{5.5}$$

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 (3.3) in section 3 and derive conditions implying the assumptions of Theorem 5.1, particularly the assumptions (5.4) and (5.5) 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 $H : \mathbb{R}_+ \to \mathbb{R}_+$ be defined by* (3.7). *If $P \in \mathcal{P}_H(\Xi)$, then $\mathcal{F}_\rho = \{f_0(\cdot, x) : x \in X \cap \rho\mathbb{B}\}$ is a P-Glivenko–Cantelli class for any $\rho > 0$, i.e.,*

$$\lim_{n\to\infty} \sup_{x \in X\cap\rho\mathbb{B}} \left| \int_\Xi f_0(\xi, x) P_n(\omega)(d\xi) - \int_\Xi f_0(\xi, x) P(d\xi) \right| = 0 \quad \mathbb{P}\text{- } a.s. \tag{5.6}$$

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.,

$$\mathbb{P}\left(\left\{\omega : \sqrt{n} \sup_{x\in X\cap\rho\mathbb{B}} \left| \int_\Xi f_0(\xi, x)(P_n(\omega) - P)(d\xi) \right| \geq \varepsilon \right\}\right) \leq (K(R)\varepsilon r^{-\frac{1}{2}})^r \exp(-2\varepsilon^2) \tag{5.7}$$

holds for some constant $K(R) > 0$*, any* $\varepsilon > 0$*, and* $n \in \mathbb{N}$.

Proof. According to (3.6) 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 $x, \tilde{x} \in X \cap \rho\mathbb{B}$, 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 [32, Theorem 2.7.11] that

$$N_{[\,]}(2\varepsilon\|F\|_{P,1}, \mathcal{F}_\rho, L_1(\Xi, P)) \leq N(\varepsilon, X \cap \rho\mathbb{B}, \mathbb{R}^m) \leq K\varepsilon^{-m} \tag{5.8}$$

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 (5.8) and obtain that $N_{[\,]}(\varepsilon, \mathcal{F}_\rho, L_1(\Xi, P))$ is finite for all $\varepsilon > 0$. Thus, condition (5.4) in Theorem 5.1 is satisfied.

If Ξ is bounded, the class $\mathcal{F}_\rho$ is uniformly bounded and condition (5.5) in Theorem 5.1 is also satisfied due to (5.8). It remains to note that the supremum $\sup_{x \in X \cap \rho\mathbb{B}}$ may be replaced by a supremum with respect to a countable dense subset of $X \cap \rho\mathbb{B}$. Hence, the suprema in (5.6) and (5.7) 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 the case of the two-stage model (3.2) with random recourse.

6. Conclusions. The quantitative stability results of section 3 extend earlier work for two-stage models with fixed recourse [18] and for multiperiod two-stage models [10]. Since Theorem 3.2 is stated in terms of the (uniform) semidistances $d_{\mathcal{F}_\rho}$, it allows two types of applications. First, it is possible to utilize metric entropy results and to quantify the asymptotic behavior of statistical approximations to two-stage stochastic programs with random recourse. Second, the analysis of continuity properties of the convex random lsc functions f_0 enables bounding semidistances by appropriate Fortet–Mourier metrics. Such metrics are easier to handle due to their relations to mass transportation problems and their dual representations, particularly for computational purposes (e.g., in scenario reduction algorithms developed in [5, 9]).

The general stability results for model (1.1) in section 2 provide continuity properties of infima and (approximate) solution sets relative to changes of the original probability distribution. They are simple consequences of general perturbation results for optimization problems. Presently, they are stated in terms of the uniform probability semidistance $d_{\mathcal{F}_\rho}$ on the space of probability measures, although the same results would be valid in terms of the corresponding epi-distances $\hat{d\!l}_\rho$ or $d\!l_\rho$, too. Such epi-distances would allow for richer spaces of probability measures $\mathcal{P}_\mathcal{F}$ and for extended real-valued objective functions $\mathbb{E}^P f_0(x)$ with different effective domains, respectively. But, since a theory for epi-counterparts of uniform distances of Fortet–Mourier type and of uniform large deviation results (see (5.3)) is not yet developed, the achieved generality would appear to be wasted. If, however, these gaps are filled in the future, the framework developed in section 2 forms the basis for extending the present results in sections 3, 4, and 5.

Acknowledgments. The authors wish to thank the guest editor Darinka Dentcheva and an anonymous referee for helpful comments.

REFERENCES

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

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

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

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

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

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

[7] A. Eichhorn and W. Römisch, *Polyhedral risk measures in stochastic programming*, SIAM J. Optim., 16 (2005), pp. 69–95.

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

[9] H. Heitsch and W. Römisch, *A note on scenario reduction for two-stage stochastic programs*, Oper. Res. Lett., to appear.

[10] H. Heitsch, W. Römisch, and C. Strugarek, *Stability of multistage stochastic programs*, SIAM J. Optim., 17 (2006), pp. 511–525.

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

[12] Y. M. Kaniovski, A. J. King, and R. J.-B. Wets, *Probabilistic bounds (via large deviations) for the solutions of stochastic programming problems*, Ann. Oper. Res., 56 (1995), pp. 189–208.

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

[14] D. Klatte, *On quantitative stability for non-isolated minima*, Control Cybernet., 23 (1994), pp. 183–200.

[15] G. Pflug, *Stochastic optimization and statistical inference*, in Stochastic Programming, A. Ruszczyński and A. Shapiro, eds., Handbooks Oper. Res. Management Sci. 10, Elsevier, Amsterdam, 2003, pp. 427–482.

[16] G. Pflug, A. Ruszczyński, and R. Schultz, *On the Glivenko–Cantelli problem in stochastic programming: Linear recourse and extensions*, Math. Oper. Res., 23 (1998), pp. 204–220.

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

[18] S. T. Rachev and W. Römisch, *Quantitative stability in stochastic programming: The method of probability metrics*, Math. Oper. Res., 27 (2002), pp. 792–818.

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

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

[21] 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 Math. 543, Springer-Verlag, Berlin, 1976, pp. 157–207.

[22] R. T. Rockafellar and S. Uryasev, *Conditional value-at-risk for general loss distributions*, J. Banking & Finance, 26 (2002), pp. 1443–1471.

[23] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, 2nd ed., Springer-Verlag, New York, 2004.

[24] W. Römisch, *Stability of stochastic programming problems*, in Stochastic Programming, A. Ruszczyński and A. Shapiro, eds., Handbooks Oper. Res. Management Sci. 10, Elsevier, Amsterdam, 2003, pp. 483–554.

[25] W. Römisch and R. Schultz, *Stability analysis for stochastic programs*, Ann. Oper. Res., 30 (1991), pp. 241–266.

[26] W. Römisch and R. Schultz, *Lipschitz stability for stochastic programs with complete recourse*, SIAM J. Optim., 6 (1996), pp. 531–547.

[27] 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, and F. Nožička, eds., Akademie Verlag, Berlin, 1987, pp. 327–343.

[28] R. Schultz, *Strong convexity in stochastic programs with complete recourse*, J. Comput. Appl. Math., 56 (1994), pp. 3–22.

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

[30] A. Shapiro, *Monte Carlo sampling methods*, in Stochastic Programming, A. Ruszczyński and A. Shapiro, eds., Handbooks Oper. Res. Management Sci. 10, Elsevier, Amsterdam, 2003, pp. 353–425.

[31] M. Talagrand, *Sharper bounds for Gaussian and empirical processes*, Ann. Probab., 22 (1994), pp. 28–76.

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

[33] W. Vervaat, *Random Upper Semicontinuous Functions and Extremal Processes*, Report MS-R8801, Center for Wiskunde en Informatica, Amsterdam, 1988.

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

[35] D. Walkup and R. J.-B. Wets, *A Lipschitzian characterization of convex polyhedra*, Proc. Amer. Math. Soc., 23 (1969), pp. 167–173.

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

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

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

[39] R. J.-B. Wets, *Lipschitz continuity of inf-projections*, Comput. Optim. Appl., 25 (2003), pp. 269–282.