Quantitative stability in stochastic programming: The method of probability metrics

Svetlozar T. Rachev[1] and Werner Römisch[2]

[1] Department of Statistics and Applied Probability, University of California, Santa Barbara, CA 93106-3110, U.S.A., and

Institute of Statistics and Mathematical Economics, University of Karlsruhe, Postfach 6980, D-76128 Karlsruhe, Germany

[2] Institute of Mathematics, Humboldt University Berlin, D-10099 Berlin, Germany

Abstract

Quantitative stability of optimal values and solution sets to stochastic programming problems is studied when the underlying probability distribution varies in some metric space of probability measures. We give conditions that imply that a stochastic program behaves stable with respect to a minimal information (m.i.) probability metric that is naturally associated with the data of the program. Canonical metrics bounding the m.i. metric are derived for specific models, namely for linear two-stage, mixed-integer two-stage and chance constrained models. The corresponding quantitative stability results as well as some consequences for asymptotic properties of empirical approximations extend earlier results in this direction. In particular, rates of convergence in probability are derived under metric entropy conditions. Finally, we study stability properties of stable investment portfolios having minimal risk with respect to the spectral measure and stability index of the underlying stable probability distribution.

Key Words: Stochastic programming, quantitative stability, probability metrics, Fortet-Mourier metrics, empirical approximations, two-stage models, chance constrained models, stable portfolio models
1991 MSC: 90C15, 90C31, 60E05, 60B10

1 Introduction

Stochastic programming is concerned with models for optimization problems under (stochastic) uncertainty that require a decision on the basis of given probabilistic information on random data. Typically, deterministic equivalents of such models represent finite-dimensional nonlinear programs. These programs and their solutions depend on the probability distribution of the random data via certain expectation functions in the objective and/or in the constraints. Such deterministic equivalents as well as many statistical decision problems take the following form

$$\min\{\int_{\Xi} f_0(\xi,x)\mu(d\xi) : x \in X, \int_{\Xi} f_j(\xi,x)\mu(d\xi) \le 0\,,\, j=1,...,d\} \tag{1}$$

where the (nonempty) set $X \subseteq \mathbb{R}^m$ of deterministic constraints is closed, Ξ is a closed subset of $\mathbb{R}^s$, the functions f_j from $\Xi \times \mathbb{R}^m$ to the extended reals $\overline{\mathbb{R}}$ are normal integrands for $j = 0,...,d$, and μ is a Borel probability measure on Ξ. Here, the set X is used to describe all constraints not depending on μ, and the set Ξ contains the supports of the relevant measures and provides some flexibility for formulating the models and the corresponding assumptions. Recall that f_j is a normal integrand if its epigraphical mapping $\xi \mapsto \text{ epi } f_j = \{(x,r) \in \mathbb{R}^m \times \mathbb{R} : f_j(\xi,x) \le r\}$ is closed-valued and measurable, which implies, in particular, that $f_j(\xi,\cdot)$ is lower semicontinuous for each $\xi \in \Xi$ and $f_j(\cdot,x)$ is measurable for each $x \in \mathbb{R}^m$ (see e.g. [38]).
In what follows we denote the set of all Borel probability measures on Ξ by $\mathcal{P}(\Xi)$, the feasible set of (1) by $M(\mu)$, and the optimal value and the solution set of (1) by $v(\mu)$ and $S(\mu)$, respectively, i.e.,

$$\begin{aligned} M(\mu) &= \{x \in X : \int_{\Xi} f_j(\xi,x)\mu(d\xi) \le 0\,,\, j=1,...,d\}, \\ v(\mu) &= \inf\{\int_{\Xi} f_0(\xi,x)\mu(d\xi) : x \in M(\mu)\}, \\ S(\mu) &= \{x \in M(\mu) : \int_{\Xi} f_0(\xi,x)\mu(d\xi) = v(\mu)\}. \end{aligned}$$

Since the underlying probability distribution μ is often incompletely known in applied models, the stability behaviour of the stochastic program (1) when changing (perturbing, estimating, approximating) μ is important. In the present paper, stability refers to continuity properties of the optimal value function $v(.)$ and the solution-set mapping $S(.)$ at μ, where both $v(.)$ and $S(.)$ are regarded as mappings given on a certain set of probability measures equipped with some metric distance. It will turn out that probability distances having the form

$$d_{\mathcal{F}}(\mu,\nu) = \sup_{f \in \mathcal{F}} |\int_{\Xi} f(\xi)\mu(d\xi) - \int_{\Xi} f(\xi)\nu(d\xi)|\,, \tag{2}$$

where $\mathcal{F}$ denotes a class of measurable functions from Ξ to $\overline{\mathbb{R}}$ and μ, ν belongs to (a suitable subset of) $\mathcal{P}_{\mathcal{F}}$, appear as natural and suitable distances of probability measures for the perturbation analysis of model (1). Such a probability distance is called *Zolotarev's pseudometric* or a distance having *ζ-structure* (see [59, 32]). Clearly, $d_{\mathcal{F}}$ satisfies all properties of a pseudometric on a subset of $\mathcal{P}_{\mathcal{F}}$, where $d_{\mathcal{F}}$ is finite, and it is a metric if the class $\mathcal{F}$ is rich enough to preserve that $d_{\mathcal{F}}(\mu,\nu) = 0$ implies $\mu = \nu$.
Our approach to analyze the stability behaviour of stochastic programming models consists of a number of steps that are intrinsically connected with and inspired by the *method of probability metrics*, i.e., by adapting a suitable probability metric to the optimization model and/or to its perturbations.
In a *first step* we show that a probability metric of the form (2) with the class $\mathcal{F}$ given by $\{f_j(\cdot,x) : x \in X \cap cl\,\mathcal{U}, j = 0,\ldots,d\}$, where $\mathcal{U}$ is a properly chosen subset of $\mathbb{R}^m$, forms a *minimal information (m.i.) metric* that implies the quantitative stability of

(1). The corresponding results (Theorems 2.2 and 2.3) work under quite weak assumptions on the underlying data of (1). In particular, differentiability or even continuity assumptions on the functions $x \mapsto \int_{\Xi} f_j(\xi,x)\mu(d\xi)$ are avoided if possible for the sake of generality. Our approach is inspired by the perturbation analysis in [4, 38] and [23, 24]. Since the m.i. metrics are often rather involved and difficult to handle, we look, in a *second step*, for another metric with ζ-structure by enlarging the class $\mathcal{F}$ and, hence, bounding the m.i. metric from above. We propose to control this enlargement procedure in such a way that each function in the enlarged class $\mathcal{F}_c$ shares the essential analytical properties with some function $f_j(\cdot,x)$ appearing in (1). More precisely, we consider normalized enlarged classes such that the functions $f_j(\cdot,x)$ are proportional to some function in $\mathcal{F}_c$ and that the corresponding probability metric $d_c = d_{\mathcal{F}_c}$ enjoys pleasant properties (e.g. a duality and convergence theory). In particular, it has to be avoided that the classes $\mathcal{F}_c$ become too rich and, hence, are not specifically adjusted to the model (like the class of all measurable functions satisfying some integrability condition). Properly enlarged classes $\mathcal{F}_c$ are called *canonical classes* and the corresponding ζ-distance d_c a *canonical metric*. In Section 3 we show for three types of stochastic programs how such canonical metrics come to light in a natural way by revealing the analytical properties of the relevant integrands $f_j(\cdot,x)$. At the same time, we obtain quantitative stability results for all models thereby extending earlier work. An important example of a probability metric with the potential of serving as a canonical metric for stochastic programs with locally Lipschitz continuous integrands is the *Fortet-Mourier metric* of order $p \geq 1$ defined on $\mathcal{P}_p(\Xi) := \{\nu \in \mathcal{P}(\Xi) : \int_{\Xi} \|\xi\|^p \nu(d\xi) < \infty\}$ by

$$\zeta_p(\mu,\nu) := \sup_{f \in \mathcal{F}_p(\Xi)} \left| \int_{\Xi} f(\xi)(\mu-\nu)(d\xi) \right|, \qquad \text{where}$$

$$\mathcal{F}_p(\Xi) := \{f : \Xi \mapsto I\!R \,:\, |f(\xi) - f(\tilde{\xi})| \leq \max\{1, \|\xi\|^{p-1}, \|\tilde{\xi}\|^{p-1}\}\|\xi - \tilde{\xi}\|\,, \forall \xi, \tilde{\xi} \in \Xi\}$$

(see [14], [32]). Fortet-Mourier metrics appear as canonical metrics in case of integrands $f_j(\cdot,x)$ with local Lipschitz constants growing polynomially. In particular, they are canonical for two-stage models without integrality requirements (see Section 3.1). For two-stage models containing integer variables and for chance constrained models, the relevant integrands are discontinuous and their canonical classes contain products of (locally) Lipschitzian functions and of characteristic functions of sets describing regions of continuity (see Sections 3.2 and 3.3).

When using stability results for designing or analyzing approximation schemes or estimation procedures, further properties of metrics and function classes have to be derived sometimes. This issue is addressed in a *third step*. In Section 4, for example, we derive entropy numbers of certain (canonical) function classes and discuss consequences for empirical approximations of stochastic programs.

Earlier work on stability analysis of stochastic programs was mostly directed to conditions implying continuity properties of optimal values and solution sets with respect to the topology of weak convergence of probability measures and distances metrizing the weak topology, respectively. We refer to the corresponding qualitative studies in [18], [37], [57] and [2], and to the work on quantitative stability in [42], [39], [1] (see

also the overviews [12] and [47]). First attempts of finding suitable probability metrics for specific models were undertaken in [40] and [46]. In the present paper, we take up this question and come to different conclusions for two-stage models.
In parallel to this development, much work was directed to the study of convergence properties of random approximations or statistical estimation procedures in stochastic programming (see e.g. [13], [22], [54], [3], [36] and [30] for qualitative results, and [20], [21], [19], [49], [15] and [50] for rates of convergence, large deviation results and central limit theorems). In the present paper, we take up the idea of using bounds for empirical processes (raised in [16] and [29]) and extend some of the earlier work.
Our paper is organized as follows. Section 2 contains the general perturbation results for (1) together with a discussion of how to associate canonical metrics with more specific stochastic programs. In Section 3 we discuss linear two-stage, mixed-integer two-stage and linear chance constrained stochastic programs and present perturbation results for these models with respect to the corresponding canonical probability metric. The application of the general perturbation analysis to empirical approximations as specific "perturbations" is discussed in Section 4 together with applications to the models in Section 3. In Section 5, we finally apply our analysis to a specific stochastic optimization problem, namely, the choice of stable portfolios with minimal risk.

2 Quantitative Stability

Together with the original stochastic programming problem (1) we consider a perturbation $\nu \in \mathcal{P}(\Xi)$ of the underlying probability distribution μ and the perturbed model

$$\min\{\int_{\Xi} f_0(\xi,x)\nu(d\xi) : \; x \in X, \; \int_{\Xi} f_j(\xi,x)\nu(d\xi) \leq 0\,,\; j=1,...,d\}\,. \tag{3}$$

We assume throughout this section that the general assumptions on X, Ξ and the functions f_j in Section 1 are satisfied and consider for any nonempty set $\mathcal{U} \subseteq I\!R^m$ the following sets of functions, elements and probability measures

$$\begin{aligned}
\mathcal{F}_{\mathcal{U}} &:= \{f_j(x,.) \,:\, x \in X \cap cl\,\mathcal{U}\,, j=0,...,d\}\,,\\
M_{\mathcal{U}}(\nu) &:= \{x \in X \cap cl\,\mathcal{U} : \int_{\Xi} f_j(\xi,x)\nu(d\xi) \leq 0\,,\; j=1,...,d\} \;\; (\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}(\Xi))\,,\\
\mathcal{P}_{\mathcal{F},\mathcal{U}}(\Xi) &:= \{\nu \in \mathcal{P}(\Xi) : -\infty < \int_{\Xi} \inf_{\substack{x\in X\\ \|x\|\leq r}} f_j(\xi,x)\nu(d\xi) \quad \text{for each } r>0 \text{ and}\\
&\qquad \sup_{x\in X\cap cl\,\mathcal{U}} \int_{\Xi} f_j(\xi,x)\nu(d\xi) < \infty \quad \text{for each } j=0,...,d\},
\end{aligned}$$

and the probability distance of μ, $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}} := \mathcal{P}_{\mathcal{F},\mathcal{U}}(\Xi)$:

$$d_{\mathcal{F},\mathcal{U}}(\mu,\nu) := \sup_{f\in\mathcal{F}_{\mathcal{U}}} |\int_{\Xi} f_j(\xi,x)(\mu-\nu)(d\xi)| = \sup_{x\in X\cap cl\,\mathcal{U}} \max_{j=0,\dots,d} |\int_{\Xi} f_j(\xi,x)(\mu-\nu)(d\xi)|\,.$$

Our general assumptions and the Fatou Lemma imply that the objective function of (3) is lower semicontinuous on X and the constraint set of (3) is closed for each $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}(\Xi)$. Our first result compiles some further basic properties of the model (3).

Proposition 2.1 *Let the general assumptions be satisfied and $\mathcal{U}$ be a nonempty subset of $I\!R^m$. Then the mapping $(x,\nu) \mapsto \int_{\Xi} f_j(\xi,x)\nu(d\xi)$ from $(X \cap cl\,\mathcal{U}) \times (\mathcal{P}_{\mathcal{F},\mathcal{U}}, d_{\mathcal{F},\mathcal{U}})$ to $\overline{I\!R}$ is lower semicontinuous for each $j = 0,...,d$. In particular, the graph of the set-valued mapping $\nu \mapsto M_{\mathcal{U}}(\nu)$ from $(\mathcal{P}_{\mathcal{F},\mathcal{U}}, d_{\mathcal{F},\mathcal{U}})$ into $I\!R^m$ is closed.*

Proof: Let $j = 0,...,d$, $x \in X \cap \mathcal{U}$, $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$, (x_n) be a sequence in $X \cap cl\,\mathcal{U}$ such that $x_n \to x$ and (ν_n) be a sequence converging to ν in $(\mathcal{P}_{\mathcal{F},\mathcal{U}}, d_{\mathcal{F},\mathcal{U}})$. Then the lower semicontinuity of $f_j(\xi,\cdot)$ for each $\xi \in \Xi$ and the Fatou Lemma imply that

$$\begin{aligned}\int_{\Xi} f_j(\xi,x)\nu(d\xi) &\leq \liminf_{n\to\infty} \int_{\Xi} f_j(\xi,x_n)\nu(d\xi)\\ &\leq \liminf_{n\to\infty} \{d_{\mathcal{F},\mathcal{U}}(\nu,\nu_n) + \int_{\Xi} f_j(\xi,x_n)\nu_n(d\xi)\}\\ &= \liminf_{n\to\infty} \int_{\Xi} f_j(\xi,x_n)\nu_n(d\xi)\,.\end{aligned}$$

Hence, the first assertion has been shown. In order to show that the graph of $\nu \mapsto M_{\mathcal{U}}(\nu)$ is closed, let $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$, (ν_n) be a sequence converging to ν in $(\mathcal{P}_{\mathcal{F},\mathcal{U}}, d_{\mathcal{F},\mathcal{U}})$ and (x_n) be a sequence such that $x_n \in M_{\mathcal{U}}(\nu_n)$ for each $n \in I\!N$ and $x_n \to x$. Clearly, we have $x \in X \cap cl\,\mathcal{U}$. For $j \in \{1,...,d\}$ we obtain from the above estimate

$$\int_{\Xi} f_j(\xi,x)\nu(d\xi) \leq \liminf_{n\to\infty} \int_{\Xi} f_j(\xi,x_n)\nu_n(d\xi) \leq 0\,.$$

and, hence, $x \in M_{\mathcal{U}}(\nu)$. ∎

To obtain quantitative stability results for (1), a stability property of the constraint set $M(\mu)$ when perturbing the *probabilistic constraints* is needed. Consistently with the general definition of metric regularity for multifunctions (see e.g. [38, 16]), we consider the set-valued mapping $y \mapsto M_y(\mu)$ from $I\!R^d$ to $I\!R^m$, where

$$M_y(\mu) = \{x \in X : \int_{\Xi} f_j(\xi,x)\mu(d\xi) \leq y_j\,,\ j = 1,...,d\},$$

and say that its inverse $x \mapsto M_x^{-1}(\mu) = \{y \in I\!R^d : x \in M_y(\mu)\}$ is *metrically regular* at some pair $(\bar{x},0) \in I\!R^m \times I\!R^d$ with $\bar{x} \in M(\mu) = M_0(\mu)$ if there are constants $a \geq 0$ and $\varepsilon > 0$ such that it holds for all $x \in X \cap I\!B(\bar{x},\varepsilon)$ and $y \in I\!R^d$ with $\max_{j=1,...,d}|y_j| \leq \varepsilon$:

$$d(x, M_y(\mu)) \leq a \max_{j=1,...,d} \max\{0, \int_{\Xi} f_j(\xi,x)\mu(d\xi) - y_j\}\,.$$

In order to state our general stability results we will also need localized versions of optimal values and solution sets, and we follow the concept of local stability analysis proposed in [35], [23]. For any nonempty set $\mathcal{U} \subseteq I\!R^m$ and any $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$ we set

$$\begin{aligned}v_{\mathcal{U}}(\nu) &= \inf\{\int_{\Xi} f_0(\xi,x)\nu(d\xi) : x \in M_{\mathcal{U}}(\nu)\},\\ S_{\mathcal{U}}(\nu) &= \{x \in M_{\mathcal{U}}(\nu) : \int_{\Xi} f_0(\xi,x)\nu(d\xi) = v_{\mathcal{U}}(\nu)\}.\end{aligned}$$

A nonempty set $\mathcal{S} \subseteq \mathbb{R}^m$ is called a *complete local minimizing (CLM) set* of (3) with respect to $\mathcal{U}$ if $\mathcal{U} \subseteq \mathbb{R}^m$ is open and $\mathcal{S} = S_{\mathcal{U}}(\nu) \subset \mathcal{U}$. Clearly, CLM sets are local minimizing sets, and the global minimizing set $S(\nu)$ is a CLM set with $S(\nu) = S_{\mathcal{U}}(\nu)$ if $S(\nu) \subset \mathcal{U}$. Now, we are ready to state our first stability result. Although our proof partly parallels arguments in the proof of Theorem 1 of [23], we include it here since our model assumptions are slightly more general compared to those in [23, 24].

Theorem 2.2 *Let the general assumptions be satisfied, $\mu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$, and assume that*

(i) $S(\mu)$ is nonempty and $\mathcal{U} \subseteq \mathbb{R}^m$ is an open bounded neighbourhood of $S(\mu)$;

(ii) if $d \geq 1$, the function $x \mapsto \int_\Xi f_0(\xi, x)\mu(d\xi)$ is Lipschitz continuous on $X \cap cl\,\mathcal{U}$;

(iii) the mapping $x \mapsto M_x^{-1}(\mu)$ is metrically regular at each pair $(\bar{x}, 0)$ with $\bar{x} \in S(\mu)$.

Then the multifunction $S_{\mathcal{U}}$ from $(\mathcal{P}_{\mathcal{F},\mathcal{U}}, d_{\mathcal{F},\mathcal{U}})$ to $\mathbb{R}^m$ is (Berge) upper semicontinuous at μ and there exist constants $L > 0$ and $\delta > 0$ such that

$$|v(\mu) - v_{\mathcal{U}}(\nu)| \leq L\, d_{\mathcal{F},\mathcal{U}}(\mu, \nu) \tag{4}$$

holds and $S_{\mathcal{U}}(\nu)$ is a CLM set of (3) w. r. t. $\mathcal{U}$ whenever $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$ and $d_{\mathcal{F},\mathcal{U}}(\mu, \nu) < \delta$. In case $d = 0$, the estimate (4) is valid with $L = 1$ and for all $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$.

Proof: We consider the (localized) parametric optimization problem

$$\min\{g(x, \nu) = \int_\Xi f_0(\xi, x)\nu(d\xi) : \ x \in M_{\mathcal{U}}(\nu)\},$$

where the probability measure ν is regarded as a parameter varying in the (pseudo-) metric space $\mathcal{P}_{\mathcal{F},\mathcal{U}}$. Proposition 2.1 implies that the graph of the multifunction $M_{\mathcal{U}}$ from $\mathcal{P}_{\mathcal{F},\mathcal{U}}$ to $\mathbb{R}^m$ is closed. Hence, $M_{\mathcal{U}}$ is (Berge) upper semicontinuous on $\mathcal{P}_{\mathcal{F},\mathcal{U}}$, since $cl\,\mathcal{U}$ is compact. Furthermore, we know by Proposition 2.1 that the function g from $(X \cap cl\,\mathcal{U}) \times \mathcal{P}_{\mathcal{F},\mathcal{U}}$ to $\overline{\mathbb{R}}$ is lower semicontinuous and finite. Let us first consider the case of $d = 0$. Since $g(\cdot, \nu)$ is lower semicontinuous, $S_{\mathcal{U}}(\nu)$ is nonempty for each $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$. Let $x_* \in S(\mu)$, $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$ and $\tilde{x} \in S_{\mathcal{U}}(\nu)$. Then the estimate

$$|v(\mu) - v_{\mathcal{U}}(\nu)| \leq \max\{\int_\Xi f_0(\xi, x_*)(\nu - \mu)(d\xi), \int_\Xi f_0(\xi, \tilde{x})(\mu - \nu)(d\xi)\} \leq d_{\mathcal{F},\mathcal{U}}(\mu, \nu)$$

holds and Berge's classical stability analysis (see Theorem 4.2.1(3) in [5]) implies that the multifunction $S_{\mathcal{U}}$ from $(\mathcal{P}_{\mathcal{F},\mathcal{U}}, d_{\mathcal{F},\mathcal{U}})$ to $\mathbb{R}^m$ is (Berge) upper semicontinuous at μ. In case $d \geq 1$, condition (ii) implies that the function g is even continuous on $(X \cap cl\,\mathcal{U}) \times \mathcal{P}_{\mathcal{F},\mathcal{U}}$. Then we may conclude from Theorem 4.2.1 in [5] that $S_{\mathcal{U}}$ is (Berge) upper semicontinuous at μ if $M_{\mathcal{U}}$ satisfies the following (lower semicontinuity) property at some pair $(\bar{x}, \mu)$ with $\bar{x} \in S(\mu)$:

$$M_{\mathcal{U}}(\mu) \cap B(\bar{x}, \bar{\varepsilon}) \subseteq M_{\mathcal{U}}(\nu) + a\, d_{\mathcal{F},\mathcal{U}}(\mu, \nu)\mathbb{B}\ , \text{ whenever } d_{\mathcal{F},\mathcal{U}}(\mu, \nu) < \bar{\varepsilon}\ , \tag{5}$$

where $\mathbb{B}$ denotes the closed unit ball in $\mathbb{R}^m$, $a \geq 0$ is the corresponding constant in condition (iii) and $\bar{\varepsilon} > 0$ is sufficiently small. To establish property (5), let $\bar{x} \in$

$S(\mu)$, and $a = a(\bar{x}) \geq 0$, $\varepsilon = \varepsilon(\bar{x}) > 0$ be the metric regularity constants from (iii). First we observe that the estimate $\int_\Xi f_j(\xi, x)(\nu - \mu)(d\xi) \leq d_{\mathcal{F},\mathcal{U}}(\mu, \nu)$ holds for any $x \in X \cap cl\,\mathcal{U}$, $j \in \{1, ..., d\}$ and $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$. Next we choose $\bar{\varepsilon} = \bar{\varepsilon}(\bar{x})$ such that $0 < \bar{\varepsilon} < \varepsilon$ and $I\!B(\bar{x}, (a+1)\,\bar{\varepsilon}) \subseteq \mathcal{U}$. Hence, we have $I\!B(x, a\bar{\varepsilon}) \subseteq \mathcal{U}$ for any $x \in B(\bar{x}, \bar{\varepsilon})$. Let $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$ be such that $d_{\mathcal{F},\mathcal{U}}(\mu, \nu) < \bar{\varepsilon}$. Putting $y_j = -d_{\mathcal{F},\mathcal{U}}(\mu, \nu)$, $j = 1, ..., d$, the above estimate implies that $M_y(\mu) \cap cl\,\mathcal{U} \subseteq M_{\mathcal{U}}(\nu)$. Due to the choice of $\bar{\varepsilon}$ we have $d(x, M_y(\mu) \cap cl\,\mathcal{U}) = d(x, M_y(\mu))$, for any $x \in M_{\mathcal{U}}(\mu) \cap B(\bar{x}, \bar{\varepsilon})$, and, hence, we obtain from the metric regularity condition (iii) the estimate

$$\begin{aligned} d(x, M_{\mathcal{U}}(\nu)) &\leq d(x, M_y(\mu) \cap cl\,\mathcal{U}) = d(x, M_y(\mu)) \\ &\leq a \max_{j=1,\ldots,d} \max\{0, \int_\Xi f_j(\xi, x)\mu(d\xi) + d_{\mathcal{F},\mathcal{U}}(\mu, \nu)\} \\ &\leq a\, d_{\mathcal{F},\mathcal{U}}(\mu, \nu)\,, \end{aligned}$$

which is equivalent to the property (5). Hence, $S_{\mathcal{U}}$ is (Berge) upper semicontinuous at μ and there exists a constant $\hat{\delta} > 0$ such that $S_{\mathcal{U}}(\nu) \subset \mathcal{U}$ for any $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$ with $d_{\mathcal{F},\mathcal{U}}(\mu, \nu) < \hat{\delta}$. Thus $S_{\mathcal{U}}(\nu)$ is a CLM set of (3) w.r.t. $\mathcal{U}$ for each such ν.
Moreover, we obtain from (iii) for any $x \in M_{\mathcal{U}}(\nu) \cap I\!B(\bar{x}, \bar{\varepsilon})$ the estimate

$$\begin{aligned} d(x, M_{\mathcal{U}}(\mu)) &= d(x, M_0(\mu) \cap cl\,\mathcal{U}) = d(x, M_0(\mu)) \\ &\leq a \max_{j=1,\ldots,d} \max\{0, \int_\Xi f_j(\xi, x)\mu(d\xi)\} \\ &\leq a \max_{j=1,\ldots,d} \max\{0, \int_\Xi f_j(\xi, x)\mu(d\xi) - \int_\Xi f_j(\xi, x)\nu(d\xi)\} \\ &\leq a\, d_{\mathcal{F},\mathcal{U}}(\mu, \nu)\,, \end{aligned}$$

which is equivalent to the inclusion

$$M_{\mathcal{U}}(\nu) \cap B(\bar{x}, \bar{\varepsilon}) \subseteq M_{\mathcal{U}}(\mu) + a\, d_{\mathcal{F},\mathcal{U}}(\mu, \nu) I\!B\,.$$

Since $S(\mu)$ is compact, we may continue as in the proof of Theorem 1 of [23] by exploiting a finite covering argument and arriving at two analogues of both inclusions, where a neighbourhood $\mathcal{N}$ of $S(\mu)$ appears in their left-hand sides instead of the balls $I\!B(\bar{x}, \bar{\varepsilon})$ and a uniform constant $\hat{a}$ appears instead of a in their right-hand sides. Moreover, there exists a uniform constant $\hat{\varepsilon} > 0$ such that the (new) inclusions are valid whenever $d_{\mathcal{F},\mathcal{U}}(\mu, \nu) < \hat{\varepsilon}$. Now, we choose $\delta > 0$ such that $\delta \leq \min\{\hat{\delta}, \hat{\varepsilon}\}$ and $S_{\mathcal{U}}(\nu) \subset \mathcal{N}$ whenever $d_{\mathcal{F},\mathcal{U}}(\mu, \nu) < \delta$.
Let $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$ and $\tilde{x} \in S_{\mathcal{U}}(\nu) \subseteq M_{\mathcal{U}}(\nu) \cap \mathcal{N}$. Then there exists an element $\bar{x} \in M_{\mathcal{U}}(\nu)$ satisfying $\|\tilde{x} - \bar{x}\| \leq \hat{a}\, d_{\mathcal{F},\mathcal{U}}(\mu, \nu)$. For $d \geq 1$ we continue

$$\begin{aligned} v(\mu) \leq g(\bar{x}, \mu) &\leq g(\tilde{x}, \nu) + |g(\bar{x}, \mu) - g(\tilde{x}, \nu)| \\ &\leq v_{\mathcal{U}}(\nu) + |g(\bar{x}, \mu) - g(\tilde{x}, \mu)| + |g(\tilde{x}, \mu) - g(\tilde{x}, \nu)| \\ &\leq v_{\mathcal{U}}(\nu) + L_g\|\bar{x} - \tilde{x}\| + d_{\mathcal{F},\mathcal{U}}(\mu, \nu) \\ &\leq v_{\mathcal{U}}(\nu) + (L_g\hat{a} + 1) d_{\mathcal{F},\mathcal{U}}(\mu, \nu)\,, \end{aligned}$$

where $L_g \geq 0$ denotes a Lipschitz constant of $g(., \mu)$ on $X \cap cl\,\mathcal{U}$. Exchanging the role of μ and ν, we arrive at the desired continuity property of $v_{\mathcal{U}}$ by putting $L = L_g\hat{a} + 1$. To complete the proof, it remains to note that in case of $d = 0$ we may choose $\bar{x} = \tilde{x} \in X$ and $\hat{a} = 0$. ■

In order to derive quantitative stability of solution sets, a growth condition on the objective function in a neighbourhood of the solution set to the original problem (1) is important (e.g. [4], [24]). The next result establishes a quantitative upper semicontinuity property of (localized) solution sets and shows how the modulus of semicontinuity relates to the growth behaviour of the objective. Parts of its proof are similar to some arguments in the proof of Theorem 7.64 in [38].

Theorem 2.3 *Let the assumptions of Theorem 2.2 be satisfied and $\mu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$. Then there exists a constant $\hat{L} \geq 1$ such that it holds for each $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$ that*

$$\emptyset \neq S_{\mathcal{U}}(\nu) \subseteq S(\mu) + \Psi(\hat{L} d_{\mathcal{F},\mathcal{U}}(\mu,\nu))\mathbb{B}\,. \tag{6}$$

Here $\Psi(\eta) := \eta + \psi^{-1}(\eta)$ and $\psi(\tau) := \min\{\int_{\Xi} f_0(\xi,x)\mu(d\xi) - v(\mu) \,:\, d(x,S(\mu)) \geq \tau\,,\, x \in M_{\mathcal{U}}(\mu)\}$ $(\eta,\tau \in \mathbb{R}_+)$, i.e., ψ is the growth function of problem (1) on $cl\,\mathcal{U}$. (The functions ψ and Ψ are lower semicontinuous on $\mathbb{R}_+$; ψ is nondecreasing and Ψ is increasing, both vanish at 0 and we set $\psi^{-1}(t) := \sup\{\tau \in \mathbb{R}_+ \,:\, \psi(\tau) \leq t\}$.)

Proof: The properties of ψ and Ψ mentioned above are verified in the proof of Theorem 7.64 in [38]. Let $L > 0$ be the constant in Theorem 2.2, and let $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$ and $\tilde{x} \in S_{\mathcal{U}}(\nu)$. As argued in the proof of Theorem 2.2, there exists an element $\bar{x} \in M_{\mathcal{U}}(\mu)$ such that $\|\tilde{x} - \bar{x}\| \leq \hat{a}\delta$, where $\delta := d_{\mathcal{F},\mathcal{U}}(\mu,\nu)$. Let $L_\mu \geq 0$ denote a Lipschitz constant of the function $x \mapsto \int_{\Xi} f_0(\xi,x)\mu(d\xi)$ on $X \cap cl\,\mathcal{U}$. Then the definition of ψ and Theorem 2.2 imply that

$$\begin{aligned}
\delta(1 + L_\mu\hat{a} + L) &\geq \delta(1 + L_\mu\hat{a}) + v_{\mathcal{U}}(\nu) - v(\mu) \\
&= \delta(1 + L_\mu\hat{a}) + \int_{\Xi} f_0(\xi,\tilde{x})\nu(d\xi) - v(\mu) \\
&\geq \delta L_\mu\hat{a} + \int_{\Xi} f_0(\xi,\tilde{x})\mu(d\xi) - v(\mu) \\
&\geq \int_{\Xi} f_0(\xi,\bar{x})\mu(d\xi) - v(\mu) \geq \psi(d(\bar{x},S(\mu))) \\
&\geq \inf_{y \in \mathbb{B}(\tilde{x},\hat{a}\delta)} \psi(d(y,S(\mu))) = \psi(d(\tilde{x},S(\mu) + \hat{a}\delta\mathbb{B}))\,.
\end{aligned}$$

Hence, we obtain

$$\begin{aligned}
d(\tilde{x},S(\mu)) &\leq \hat{a}\delta + d(\tilde{x},S(\mu) + \hat{a}\delta\mathbb{B}) \\
&\leq \hat{a}\delta + \psi^{-1}(\delta(1 + L_\mu\hat{a} + L)) \leq \hat{L}\delta + \psi^{-1}(\hat{L}\delta) = \Psi(\hat{L}\delta)\,,
\end{aligned}$$

where $\hat{L} := \max\{\hat{a}, 1 + L_\mu\hat{a} + L\} \geq 1$. In case that $d = 0$ we may choose $\hat{x} = \tilde{x}$, $\hat{a} = 1$ and $L_\mu = 0$. This completes the proof. ■

Next, we briefly comment on some aspects of the general stability theorems, namely, specific growth conditions, verification of condition (iii), localization issues and an extension of Theorem 2.3 in case $d = 0$.

Remark 2.4 *If* $\psi(\tau) = \gamma\tau^k$ *for some* $k \geq 1$ *and* $\gamma > 0$ *(k-th order growth at the solution set), we have* $\Psi(\eta) = \hat{L}\eta + (\frac{\hat{L}}{\gamma}\eta)^{\frac{1}{k}} \leq C\eta^{\frac{1}{k}}$ *for some constant* $C > 0$ *and sufficiently small* $\eta \geq 0$*. In this case, Theorem 2.3 provides the Hölder continuity of* $S_{\mathcal{U}}$ *at* μ *with rate* $\frac{1}{k}$*. Note that for* $d = 0$ *the function* Ψ *is of the form* $\Psi(\eta) = \eta + \psi^{-1}(2\eta)$ $(\eta \in I\!R_+)$*, since then we have* $L = 1$ *in Theorem 2.2.*

Remark 2.5 *Criteria for the metric regularity of multifunctions are given e.g. in Section 9G of [38] and in [27]. Here, we do not provide a specific sufficient condition for assumption (iii) of Theorem 2.2, but note that the constraint functions* $\int_{\Xi} f_j(\xi,\cdot)\mu(d\xi)$ *are often nondifferentiable in stochastic programming and refer to the general results in [27] and to [16], where metric regularity in case of chance constrained stochastic programs is discussed.*

Remark 2.6 *In the Theorems 2.2 and 2.3 the localized optimal values* $v_{\mathcal{U}}(\nu)$ *and solution sets* $S_{\mathcal{U}}(\nu)$ *of the (perturbed) model (3) may be replaced by their global versions* v *and* S *if there exists a constant* $\delta_0 > 0$ *such that for each* $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$ *with* $d_{\mathcal{F},\mathcal{U}}(\mu,\nu) < \delta_0$ *either of the following conditions is satisfied: (a) The model (3) is convex and* $S_{\mathcal{U}}(\nu)$ *is a CLM set, (b) the constraint set of (3) is contained in some bounded set* $\mathcal{V} \subset I\!R^m$ *not depending on* ν *and it holds that* $\mathcal{V} \subseteq \mathcal{U}$*.*

Remark 2.7 *Let* $d = 0$*,* $S(\mu)$ *be nonempty and assume that the objective satisfies a second order growth condition on* $X \cap \mathcal{U}$*, where* $\mathcal{U}$ *is a convex open neighbourhood of* $S(\mu)$ *(i.e.,* $\psi(\tau) = \gamma\tau^2$ *for small* $\tau \in I\!R_+$*). Then the estimate*

$$\sup_{x \in \mathcal{S}_{\mathcal{U}}(\nu)} d(x, S(\mu)) \leq \frac{1}{\gamma} \sup\{\|x^*\| \,:\, x^* \in \partial(\int_{\Xi} f_0(\xi,\cdot)(\nu-\mu)(d\xi))(x)\,,\, x \in X \cap cl\,\mathcal{U}\}\,,$$

holds provided that the function $\int_{\Xi} f_0(\xi,\cdot)(\nu-\mu)(d\xi)$ *is locally Lipschitz continuous on* X*. Here, "*∂*" denotes the Mordukhovich subdifferential (cf. [27, 38]).*
For the proof of such an estimate and for some conclusions on Lipschitz stability results in two-stage stochastic programming, respectively, the reader is referred to [7, 48] and [48, 41, 9].

The two stability theorems illuminate the role of the distance $d_{\mathcal{F},\mathcal{U}}$ as a *minimal information (m.i.)* probability (pseudo-) metric for stability, i.e., as a (pseudo-) metric processing the minimal information of problem (1), implying quantitative stability of its optimal values and solution sets. Furthermore, notice that both results remain valid when bounding $d_{\mathcal{F},\mathcal{U}}$ from above by another distance and when reducing the set $\mathcal{P}_{\mathcal{F},\mathcal{U}}$ to a subset on which this distance is defined and finite. Such a distance will be called a *canonical probability metric* d_c associated with (1) if it has ζ-structure (2) generated by some class of functions $\mathcal{F} = \mathcal{F}_c$ from Ξ to $\overline{I\!R}$ such that $\mathcal{F}_c$ contains the functions $Cf_j(\cdot,x)$ for each $x \in X \cap cl\,\mathcal{U}$, $j = 0,\ldots,d$, and some normalizing constant $C > 0$, and such that the functions in $\mathcal{F}_c$ have the same analytical properties as $f_j(\cdot,x)$, $j = 1,\ldots,d$. Typical analytical properties defining canonical function classes $\mathcal{F}_c$ relevant in stochastic programming are piecewise Lipschitz continuity properties.
To give an idea of how to associate a canonical metric, we show that for a generalized

Fortet-Mourier metric (cf. Section 1). Let $g : I\!R_+ \to I\!R_+$ be a nondecreasing function with $g(0) = 0$ and ξ_0 be an element in Ξ, and consider the following class of continuous functions from Ξ to $I\!R$

$$\mathcal{F}_g(\Xi) := \{f : |f(\xi) - f(\tilde{\xi})| \le \max\{1, g(\|\xi - \xi_0\|), g(\|\tilde{\xi} - \xi_0\|)\}\|\xi - \tilde{\xi}\|, \forall \xi, \tilde{\xi} \in \Xi\},$$

and the corresponding probability metric with ζ-structure generated by $\mathcal{F}_g$

$$\zeta_g(\mu, \nu) := d_{\mathcal{F}_g(\Xi)}(\mu, \nu) = \sup_{f \in \mathcal{F}_g(\Xi)} |\int_\Xi f(\xi)(\mu - \nu)(d\xi)|$$

defined on $\mathcal{P}_g(\Xi) := \{\nu \in \mathcal{P}(\Xi) : \int_\Xi \max\{1, g(\|\xi\|)\}\|\xi\|\nu(d\xi) < \infty\}$.

Then the following result is an immediate consequence of the general results. It was already announced in Section 1.4 of [34].

Corollary 2.8 *In addition to the general assumptions, let $d = 0$, $\xi_0 \in \Xi$, $g : I\!R_+ \to I\!R_+$ be nondecreasing with $g(0) = 0$ and $\mu \in \mathcal{P}_g(\Xi)$. Furthermore, assume that*
(i) $S(\mu)$ is nonempty and $\mathcal{U}$ is an open, bounded neighbourhood of $S(\mu)$;
(ii) X is convex and $f_0(\xi, \cdot)$ is convex on $I\!R^m$ for each $\xi \in \Xi$;
(iii) there exists a constant $L > 0$ such that $\frac{1}{L} f_0(\cdot, x) \in \mathcal{F}_g$ for each $x \in X \cap cl\,\mathcal{U}$.
Then there exists a constant $\delta > 0$ such that

$$\begin{aligned} |v(\mu) - v(\nu)| &\le L\zeta_g(\mu, \nu) \quad \text{and} \\ \emptyset \neq S(\nu) &\subseteq S(\mu) + \Psi(L\zeta_g(\mu, \nu))I\!B \end{aligned}$$

whenever $\nu \in \mathcal{P}_g(\Xi)$ and $\zeta_g(\mu, \nu) < \delta$.

Here $\Psi(\eta) := \eta + \psi^{-1}(2\eta)$ $(\eta \in I\!R_+)$ and $\psi(\tau) := \min\{\int_\Xi f_0(\xi, x)\mu(d\xi) - v(\mu) : d(x, S(\mu)) \ge \tau, x \in X \cap cl\,\mathcal{U}\}$ $(\tau \in I\!R_+)$.

Proof: The assumptions of Theorem 2.2 are satisfied. Hence, the result is a consequence of the Theorems 2.2 and 2.3 and the fact that (iii) is equivalent to

$$|f_0(\xi, x) - f_0(\tilde{\xi}, x)| \le L \max\{1, g(\|\xi - \xi_0\|), g(\|\tilde{\xi} - \xi_0\|)\}\|\xi - \tilde{\xi}\|$$

for each $\xi, \tilde{\xi} \in \Xi$ and $x \in X \cap cl\,\mathcal{U}$, and, hence, it implies $d_{\mathcal{F},\mathcal{U}}(\mu, \nu) \le L\zeta_g(\mu, \nu)$ for all $\mu, \nu \in \mathcal{P}_g(\Xi)$. Furthermore, the localized optimal values $v_\mathcal{U}$ and solution sets $S_\mathcal{U}$ may be replaced by v and S, respectively, due to the convexity assumption (ii) if ν is close to μ (see Remark 2.6). ■

3 Stability of Linear Two-Stage and Chance Constrained Models

3.1 Linear Two-Stage Models

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

$$(7) \qquad \min\{cx + \int_\Xi q(\xi)y(\xi)\mu(d\xi) : Wy(\xi) = h(\xi) - T(\xi)x, y(\xi) \ge 0, x \in X\}$$

where $c \in I\!R^m$, $X \subseteq I\!R^m$ is a polyhedron, Ξ is a polyhedron in $I\!R^s$, W is an $(r, \overline{m})$-matrix, $\mu \in \mathcal{P}(\Xi)$, and the vectors $q(\xi) \in I\!R^{\overline{m}}$, $h(\xi) \in I\!R^r$ and the (r, m)-matrix $T(\xi)$ depend affine linearly on $\xi \in \Xi$.
In (7), x is the here-and-now or first stage decision and $y(.)$ the (stochastic) recourse or second stage action, which is needed to compensate a violation of the constraint $T(\xi)x = h(\xi)$. For given x the recourse action $y(\xi)$ is chosen such that it meets the constraints and is optimal for the (stochastic) recourse costs $q(\xi)$. Denoting by $\Phi(q(\xi), h(\xi) - T(\xi)x)$ the value of the optimal second stage decision, problem (7) may be rewritten equivalently as a minimization problem with respect to the first stage decision x. Defining the integrand $f_0 : \Xi \times I\!R^m \to \overline{I\!R}$ by

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

where $\operatorname{pos} W = \{Wy : y \in I\!R^{\overline{m}}_+\}$, $D = \{u \in I\!R^{\overline{m}} : \{z \in I\!R^r : W'z \leq u\} \neq \emptyset\}$ and $\Phi(u, t) = \inf\{uy : Wy = t, y \geq 0\}$ $((u, t) \in I\!R^{\overline{m}} \times I\!R^r)$, the equivalent minimization problem takes the form

$$\min\{\int_{\Xi} f_0(\xi, x)\mu(d\xi) : \; x \in X\}. \tag{8}$$

In order to utilize the general stability results of Section 2, we first recall some well-known properties of the function Φ, which are derived in [55] (see also [28]).

Lemma 3.1 *The function Φ is finite and continuous on the $(\overline{m}+r)$-dimensional polyhedral cone $pos W \times D$ and there exist $(r, \overline{m})$-matrices C_j and $(\overline{m} + r)$-dimensional polyhedral cones $\mathcal{K}_j$, j=1,...,N, such that*

$$\begin{aligned} \bigcup_{j=1}^{N} \mathcal{K}_j &= pos W \times D \, , \; int \mathcal{K}_i \cap int \mathcal{K}_j = \emptyset \, , \, i \neq j, \\ \Phi(u, t) &= C_j u \cdot t \, , \; \text{ for each } \; (u, t) \in \mathcal{K}_j, \; j = 1, ..., N. \end{aligned}$$

Moreover, $\Phi(u, .)$ is convex on $pos W$, for each $u \in D$, and $\Phi(., t)$ is concave on D for each $t \in pos W$.

In order to have problem (8) well defined we introduce the following assumptions:
(A1) For each pair $(\xi, x) \in \Xi \times X$ it holds that $h(\xi) - T(\xi)x \in \operatorname{pos} W$ and $q(\xi) \in D$.
(A2) $\mu \in \mathcal{P}(\Xi)$ has a finite second order moment, i.e., $\int_{\Xi} \|\xi\|^2 \mu(d\xi) < \infty$.
Condition (A1) combines the two usual conditions: *relatively complete recourse* and *dual feasibility*. It implies that $\Xi \times X \subseteq \operatorname{dom} f_0$.

Proposition 3.2 *Let (A1) be satisfied. Then f_0 is a normal convex integrand. Furthermore, there exist constants $L > 0$, $\hat{L} > 0$ and $K > 0$ such that the following holds for all $\xi, \tilde{\xi} \in \Xi$ and $x, \tilde{x} \in X$ with $\|x\| \leq r$:*

$$\begin{aligned} |f_0(\xi, x) - f_0(\tilde{\xi}, x)| &\leq L r \max\{1, \|\xi\|, \|\tilde{\xi}\|\} \|\xi - \tilde{\xi}\| \, , \\ |f_0(\xi, x) - f_0(\xi, \tilde{x})| &\leq \hat{L} \max\{1, \|\xi\|^2\} \|x - \tilde{x}\| \, , \\ |f_0(\xi, x)| &\leq K r \max\{1, \|\xi\|^2\} \, . \end{aligned}$$

Proof: From Lemma 3.1 and (A1) we conclude that f_0 is continuous on $\mathrm{dom}\, f_0$ and, hence, on $\Xi \times I\!R^m$. This implies that f_0 is a normal integrand (cf. Example 14.31 in [38]). It is also a convex integrand since the properties of Φ in Lemma 3.1 imply that $f_0(\xi, .)$ is convex for each $\xi \in \Xi$. In order to verify the Lipschitz property of f_0, let $x \in X$ with $\|x\| \leq r$ and consider for each $j = 1, \ldots, N$ and $\xi \in \Xi_j$ the function

$$g_j(\xi) := f_0(\xi, x) = \Phi(q(\xi), h(\xi) - T(\xi)x) = C_j q(\xi) \cdot (h(\xi) - T(\xi)x)\,,$$

where $\Xi_j := \{\xi \in \Xi \,:\, (q(\xi), h(\xi) - T(\xi)x) \in \mathcal{K}_j\}$ are polyhedral subsets of Ξ, and C_j and $\mathcal{K}_j$ are the matrices and the polyhedral cones from Lemma 3.1, respectively. Since $q(\cdot)$, $h(\cdot)$ and $T(\cdot)$ depend affine linearly on ξ, the function g_j depends quadratically on ξ and linearly on x. Hence, there exists a constant $L_j > 0$ such that g_j satisfies the following Lipschitz property:

$$|g_j(\xi) - g_j(\tilde{\xi})| \leq L_j r \max\{1, \|\xi\|, \|\tilde{\xi}\|\} \|\xi - \tilde{\xi}\| \quad \text{for all } \xi, \tilde{\xi} \in \Xi_j.$$

Now, let $\xi\,,\, \tilde{\xi} \in \Xi$, assume that $\xi \in \Xi_i$ and $\tilde{\xi} \in \Xi_k$ for some $i\,,\, k \in \{1, \ldots, N\}$ and consider the line segment $[\xi, \tilde{\xi}] = \{\xi(\lambda) = (1-\lambda)\xi + \lambda\tilde{\xi} \,:\, \lambda \in [0,1]\}$. Since $[\xi, \tilde{\xi}] \subseteq \Xi$, there exist indices $i_j\,,\, j = 1, \ldots, l$ such that $i_1 = i$, $i_l = k$, $[\xi, \tilde{\xi}] \cap \Xi_{i_j} \neq \emptyset$ for each $j = 1, \ldots, l$ and $[\xi, \tilde{\xi}] \subseteq \bigcup_{j=1}^{l} \Xi_{i_j}$. Furthermore, there exist increasing numbers $\lambda_{i_j} \in [0,1]$ for $j = 0, \ldots, l-1$ such that $\xi(\lambda_{i_0}) = \xi(0) = \xi$, $\xi(\lambda_{i_j}) \in \Xi_{i_j} \cap \Xi_{i_{j+1}}$ and $\xi(\lambda) \notin \Xi_{i_j}$ if $\lambda_{i_j} < \lambda \leq 1$. Then we obtain

$$\begin{aligned}
|f_0(\xi, x) - f_0(\tilde{\xi}, x)| &= |g_{i_1}(\xi) - g_{i_l}(\tilde{\xi})| \\
&\leq \sum_{j=0}^{l-1} |g_{i_{j+1}}(\xi(\lambda_{i_j})) - g_{i_{j+1}}(\xi(\lambda_{i_{j+1}}))| \\
&\leq \sum_{j=0}^{l-1} L_{i_{j+1}} r \max\{1, \|\xi(\lambda_{i_j})\|, \|\xi(\lambda_{i_{j+1}})\|\} \|\xi(\lambda_{i_j}) - \xi(\lambda_{i_{j+1}})\| \\
&\leq \max_{j=1,\ldots,N} L_j r \max\{1, \|\xi\|, \|\tilde{\xi}\|\} \sum_{j=0}^{l-1} \|\xi(\lambda_{i_j}) - \xi(\lambda_{i_{j+1}})\| \\
&\leq \max_{j=1,\ldots,N} L_j r \max\{1, \|\xi\|, \|\tilde{\xi}\|\} \|\xi - \tilde{\xi}\|\,,
\end{aligned}$$

where we have used for the last two estimates that $\|\xi(\lambda)\| \leq \max\{\|\xi\|, \|\tilde{\xi}\|\}$ for each $\lambda \in [0,1]$ and that $|\lambda - \tilde{\lambda}| \|\xi - \tilde{\xi}\| = \|\xi(\lambda) - \xi(\tilde{\lambda})\|$ holds for all $\lambda, \tilde{\lambda} \in [0,1]$.
Lipschitz continuity of f_0 with respect to x is shown in Theorem 10 of [17] and in Theorem 7.7 of [56]. The second estimate of the proposition, in particular, is a consequence of those results. Furthermore, from Lemma 3.1 we conclude the estimate

$$\begin{aligned}
|f_0(\xi, x)| &\leq \sup_{\|x\| \leq r} \{|cx| + \max_{j=1,\ldots,N} |C_j q(\xi) \cdot (h(\xi) - T(\xi)x)|\} \\
&\leq \|c\| r + (\max_{j=1,\ldots,N} \|C_j\|) \|q(\xi)\| (\|h(\xi)\| + \|T(\xi)\| r)
\end{aligned}$$

for any pair $(\xi, x) \in \Xi \times X$ with $\|x\| \leq r$. Then the third estimate follows again from the fact that $q(.)$, $h(.)$ and $T(.)$ depend affine linearly on ξ. ■

The estimate in Proposition 3.2 implies that, for any $r > 0$, any nonempty bounded $\mathcal{U} \subseteq I\!R^m$ and some $\rho > 0$, it holds

$$\begin{aligned} \int_\Xi \inf_{\substack{x\in X \\ \|x\|\le r}} f_0(\xi,x)\nu(d\xi) &\ge -Kr(1+\int_\Xi \|\xi\|^2\nu(d\xi)) > -\infty\,, \\ \sup_{x\in X\cap\mathcal{U}} |\int_\Xi f_0(\xi,x)\nu(d\xi)| &\le K\rho(1+\int_\Xi \|\xi\|^2\nu(d\xi)) < \infty\,. \end{aligned}$$

if $\nu \in \mathcal{P}(\Xi)$ has a finite second order moment. Hence, for any nonempty bounded $\mathcal{U} \subseteq I\!R^m$ the set of probability measures $\mathcal{P}_{\mathcal{F},\mathcal{U}}$ contains the set of measures on Ξ having finite second order moments, i.e.,

$$\mathcal{P}_{\mathcal{F},\mathcal{U}} \supseteq \{\nu\in\mathcal{P}(\Xi) : \int_\Xi \|\xi\|^2\nu(d\xi) < \infty\} = \mathcal{P}_2(\Xi)\,.$$

The following stability result for optimal values and solution sets of the two-stage problem (8) is now a direct consequence of Corollary 2.8 and Proposition 3.2.

Theorem 3.3 *Let (A1) and (A2) be satisfied and let $S(\mu)$ be nonempty and $\mathcal{U}$ be an open, bounded neighbourhood of $S(\mu)$.*
Then there exist constants $L > 0$ and $\delta > 0$ such that

$$\begin{aligned} |v(\mu)-v(\nu)| &\le L\zeta_2(\mu,\nu) \quad \textit{and} \\ \emptyset \ne S(\nu) &\subseteq S(\mu)+\Psi(L\zeta_2(\mu,\nu))I\!B \end{aligned}$$

whenever $\nu \in \mathcal{P}_2(\Xi)$ and $\zeta_2(\mu,\nu) < \delta$, where Ψ is defined as in Corollary 2.8.

Proof: The result is a consequence of Corollary 2.8 with $g(t) := t$ for $t \in I\!R_+$. The assumptions (ii) and (iii) of Corollary 2.8 are verified in Proposition 3.2. ∎

The theorem establishes quantitative stability of $v(\cdot)$ and $S(\cdot)$ in the two-stage case for fairly general situations. It extends the stability result with respect to the Wasserstein metric W_2 in [40] for the situation of complete recourse, where W_2 is defined by

$$W_2(\mu,\nu) := (\inf\{\int_{\Xi\times\Xi} \|\xi-\tilde{\xi}\|^2\eta(d\xi,d\tilde{\xi}) \;:\; \eta\in\mathcal{P}(\Xi\times\Xi) \text{ with marginals } \mu \text{ and } \nu\})^{\frac{1}{2}}$$

for all $\mu,\ \nu \in \mathcal{P}_2(\Xi)$. This follows from the estimate

$$\zeta_2(\mu,\nu) \le (1+\int_\Xi \|\xi\|^2\mu(d\xi)+\int_\Xi \|\xi\|^2\nu(d\xi))^{\frac{1}{2}}W_2(\mu,\nu)\,,$$

which is a consequence of Schwarz' inequality, and from the fact that convergence with respect to W_2 implies convergence of second order moments. Hence, Theorem 3.3 remains valid if ζ_2 is replaced by W_2 (with possibly different constants). However, the following example shows that ζ_2 and W_2 may have different asymptotic properties.

Example 3.4 *Let $\Xi := I\!R_+$, $\mu := \delta_0$ and $\mu_n := (1-\frac{1}{n})\delta_0 + \frac{1}{n}\delta_{\xi_n}$ for each $n \in I\!N$, where (ξ_n) is an unbounded nondecreasing sequence in $I\!R_+$ such that $(\frac{1}{\sqrt{n}}\xi_n)$ tends to zero. Here, δ_ξ denotes the measure on Ξ placing unit mass at $\xi \in \Xi$. Then (μ_n) converges* *weakly to μ and the second order moments converge, too. Hence, (μ_n) converges to μ with respect to ζ_2 and W_2. However, the speed of convergence of (μ_n) is different since it holds for each $n \in I\!N$ with $\xi_n \geq 1$:*

$$\zeta_2(\mu, \mu_n) = \int_0^\infty \max\{1, \xi\}|F_\mu(\xi) - F_{\mu_n}(\xi)|d\xi = \frac{1}{n} + \int_1^{\xi_n} \frac{\xi}{n} d\xi = \frac{1}{2n}(\xi_n^2 + 1)\,,$$

$$W_2(\mu, \mu_n) = (\int_0^1 |F_\mu^{-1}(t) - F_{\mu_n}^{-1}(t)|^2 dt)^{\frac{1}{2}} = (\int_{1-\frac{1}{n}}^1 \xi_n^2 dt)^{\frac{1}{2}} = \frac{1}{\sqrt{n}} \xi_n\,.$$

Here we have used the explicit representations of ζ_2 and W_2 in case of probability measures on (subsets of) $I\!R$ (see the Chapters 5.4 and 13.1 in [32]), where F_μ and F_{μ_n} are the probability distribution functions of μ and μ_n, respectively, and the (generalized) inverse function of a distribution function F is defined by $F^{-1}(t) := \sup\{\xi \in I\!R : F(\xi) \leq t\}$ $(t \in [0, 1])$.

Concluding this section, we mention that in case the recourse costs $q(\cdot)$ are non-stochastic or the technology matrix $T(\cdot)$ and the right-hand side $h(\cdot)$ are non-stochastic, then assumption (A2) may be weakened to $(A2)^*$ $\mu \in \mathcal{P}_1(\Xi)$, and Theorem 3.3 holds true with the metric ζ_1 instead of ζ_2. In case of random right-hand sides only stronger stability results are obtained in [40, 41, 9] by exploiting strong convexity properties of the objective function (cf. [44]).

3.2 Mixed-Integer Two-Stage Models

Next we allow for mixed-integer decisions in both stages and consider the program

$$\min\{cx + \int_\Xi \Phi(h(\xi) - T(\xi)x)\mu(d\xi) \,:\, x \in X\}\,, \tag{9}$$

where

$$\Phi(t) := \min\{qy + \bar{q}\bar{y} \,:\, Wy + \bar{W}\bar{y} = t\,,\ y \in Z\!\!\!Z_+^{\hat{m}}\,,\ \bar{y} \in I\!R_+^{\bar{m}}\}\ (t \in I\!R^r)\,, \tag{10}$$

$c \in I\!R^m$, X is a closed subset of $I\!R^m$, Ξ a polyhedron in $I\!R^s$, $q \in I\!R^{\hat{m}}$, $\bar{q} \in I\!R^{\bar{m}}$, W and $\bar{W}$ are $(r, \hat{m})$- and $(r, \bar{m})$-matrices, respectively, $h(\xi) \in I\!R^r$ and the (r, m)-matrix $T(\xi)$ are affine linear functions of $\xi \in I\!R^s$, and $\mu \in \mathcal{P}(\Xi)$.
Similarly as for the two-stage models without integrality requirements in the previous section, we need some conditions to have the model (9) well-defined:
(B1) The matrices W and $\bar{W}$ have only rational elements.
(B2) For each pair $(\xi, x) \in \Xi \times X$ it holds that $h(\xi) - T(\xi)x \in \mathcal{T}$, where
$\mathcal{T} := \{t \in I\!R^r \,:\, t = Wy + \bar{W}\bar{y},\ y \in Z\!\!\!Z_+^{\hat{m}},\ \bar{y} \in I\!R_+^{\bar{m}}\}$.
(B3) There exists an element $u \in I\!R^r$ such that $W'u \leq q$ and $\bar{W}'u \leq \bar{q}$.
The conditions (B2) and (B3) mean *relatively complete recourse* and *dual feasibility*, respectively. We note that condition (B3) is equivalent to $\Phi(0) = 0$, and that (B2) and (B3) imply that $\Phi(t)$ is finite for all $t \in \mathcal{T}$. In the context of this section, the following properties of the value function Φ of (10) on $\mathcal{T}$ are important.

Lemma 3.5 *Assume (B1)–(B3). Then there exists a countable partition of $\mathcal{T}$ into Borel subsets $\mathcal{B}_i$, i.e., $\mathcal{T} = \bigcup_{i \in I\!N} \mathcal{B}_i$ such that*

(i) each of the sets has a representation $\mathcal{B}_i = \{b_i + \text{ pos } \bar{W}\} \setminus \bigcup_{j=1}^{N_0} \{b_{ij} + \text{ pos } \bar{W}\}$, where $b_i, b_{ij} \in I\!R^r$ for $i \in I\!N$ and $j = 1, \ldots, N_0$. Moreover, there exists an $N_1 \in I\!N$ such that for any $t \in \mathcal{T}$ the ball $I\!B(t,1)$ in $I\!R^r$ is intersected by at most N_1 different subsets $\mathcal{B}_i$.

(ii) the restriction $\Phi|_{\mathcal{B}_i}$ of Φ to $\mathcal{B}_i$ is Lipschitz continuous with a constant $L_\Phi > 0$ that does not depend on i.

Furthermore, the function Φ is lower semicontinuous and piecewise polyhedral on $\mathcal{T}$ and there exist constants $\alpha > 0$ and $\beta > 0$ such that it holds for all $t, \tilde{t} \in \mathcal{T}$:

$$|\Phi(t) - \Phi(\tilde{t})| \leq \alpha \|t - \tilde{t}\| + \beta \,.$$

Part (i) of the lemma is proved in Section 5.6 of [5] and in Lemma 2.5 of [46], (ii) is derived as Lemma 2.3 in [46] and the remaining properties of Φ are established in [6]. Compared to Lemma 3.1 for optimal value functions of linear programs without integrality requirements, the representation of Φ is now given on countably many (unbounded) Borel sets. This requires an assumption on the tail behaviour of μ and its perturbations ν that is stronger than (A2).

Theorem 3.6 *Let the conditions (B1)–(B3) be satisfied and let $\mu \in \mathcal{P}_{p,K}(\Xi) := \{\nu \in \mathcal{P}_p(\Xi) : \int_\Xi \|\xi\|^p \nu(d\xi) \leq K\}$ for some constants $p > 1$ and $K > 0$. Assume that $S(\mu)$ is nonempty and $\mathcal{U} \subseteq I\!R^m$ is an open bounded neighbourhood of $S(\mu)$.*
Then there exist constants $L > 0$, $\delta > 0$ and $k \in I\!N$ such that

$$
\begin{aligned}
|v(\mu) - v_{\mathcal{U}}(\nu)| &\leq L\, d_{1,phk}(\mu,\nu)^{\frac{1}{1+\frac{r}{p-1}}} \qquad (11)\\
\emptyset \neq S_{\mathcal{U}}(\nu) &\subseteq S(\mu) + \Psi(L d_{1,phk}(\mu,\nu)^{\frac{1}{1+\frac{r}{p-1}}}) I\!B
\end{aligned}
$$

and $S_{\mathcal{U}}(\nu)$ is a CLM set for (9) w.r.t. $\mathcal{U}$ whenever $\nu \in \mathcal{P}_{p,K}(\Xi)$ and $d_{1,phk}(\mu,\nu) < \delta$.

Here, $d_{1,phk}(\mu,\nu) := \sup\{|\int_P f(\xi)(\mu-\nu)(d\xi)| : f \in \mathcal{F}_{1,lg}(P)\,,\ P$ is a polyhedron with at most k faces$\}$, $\mathcal{F}_{1,lg}(P) := \{f \in \mathcal{F}_1(P) : |f(\xi)| \leq \max\{1, \|\xi\|\}$ for each $\xi \in P\}$, $\mathcal{F}_1(P)$ and $\mathcal{P}_p(\Xi)$ are introduced in Section 1 and Ψ is defined as in Corollary 2.8.

Proof: For each pair $(\xi, x) \in \Xi \times X$ we set $f_0(\xi,x) := cx + \Phi(h(\xi) - T(\xi)x)$. Then f_0 is lower semicontinuous on $\Xi \times X$ and, hence, a normal integrand (Example 14.31 in [38]). Using Lemma 3.5 we obtain the estimate

$$|f_0(\xi,x)| \leq \|c\|\|x\| + \alpha(\|h(\xi)\| + \|T(\xi)\|\|x\|) + \beta$$

for each pair $(\xi,x) \in \Xi \times X$. Since $h(\xi)$ and $T(\xi)$ depend affine linearly on ξ, it holds that $\mathcal{P}_{\mathcal{F},\mathcal{U}}(\Xi) \subseteq \mathcal{P}_1(\Xi)$. Hence, Theorem 2.2 applies with $d = 0$ and the distance $d_{\mathcal{F},\mathcal{U}}$ on $\mathcal{P}_1(\Xi)$. It remains to show that the estimate

$$d_{\mathcal{F},\mathcal{U}}(\mu,\nu) = \sup_{x \in X \cap cl\,\mathcal{U}} \left| \int_\Xi f_0(\xi,x)(\mu-\nu)(d\xi) \right| \leq C d_{1,phk}(\mu,\nu)^{\frac{1}{1+\frac{r}{p-1}}} \qquad (12)$$

is valid for some constant $C > 0$ and sufficiently small $d_{1,phk}(\mu,\nu)$. To this end, let $\mathcal{T}_R := \mathcal{T} \cap \mathbb{B}_\infty(0,R)$ for any $R > 0$, where $\mathbb{B}_\infty$ refers to a closed ball in $\mathbb{R}^r$ equipped with the norm $\|\cdot\|_\infty$. Now, we proceed similarly as in the proof of Proposition 3.1 in [46] and partition the ball $\mathbb{B}_\infty(0,R)$ into disjoint Borel sets whose closures are $\mathbb{B}_\infty$-balls with radius 1 where possible gaps are filled with maximal balls of radius less than 1. Then the number of elements in this partition of $\mathbb{B}_\infty(0,R)$ is bounded above by $(2R)^r$. From Lemma 3.5 (i) we know that each element of this partition is intersected by at most N_1 subsets $\mathcal{B}_i$ (for some $N_1 \in \mathbb{N}$). Another consequence of Lemma 3.5 (i) is that each $\mathcal{B}_i$ splits into disjoint Borel subsets whose closures are polyhedra. Moreover, the number of such subsets can be bounded from above by a constant not depending on i (cf. also page 1143 in [46]). Hence, there exist a number $N \in \mathbb{N}$ and disjoint Borel subsets $\{B_j : j = 1,\ldots,N\}$ such that their closures are polyhedra, their union contains $\mathcal{T}_R$, and N is bounded above by κR^r, where the constant $\kappa > 0$ is independent on R. Now, let $x \in X \cap cl\,\mathcal{U}$ and consider the following disjoint Borel subsets of Ξ:

$$\begin{aligned}
\Xi^R_{j,x} &:= \{\xi \in \Xi : h(\xi) - T(\xi)x \in B_j\} \ (j = 1,\ldots,N),\\
\Xi^R_{0,x} &:= \Xi \setminus \bigcup_{j=1}^{N} \Xi^R_{j,x} \subseteq \{\xi \in \Xi : \|h(\xi) - T(\xi)x\|_\infty > R\}.
\end{aligned}$$

From Lemma 3.5 we conclude that there exists a constant $L_1 > 0$ (which does not depend on $x \in X \cap cl\,\mathcal{U}$, $j = 1,\ldots,N$ and $R > 0$) such that the function $f^R_{j,x}(\cdot) := cx + \Phi|_{B_j}(h(\cdot) - T(\cdot)x)$ is Lipschitz continuous on $\Xi^R_{j,x}$ with constant L_1 and that $|f^R_{j,x}(\xi)| \le L_1 \max\{1, \|\xi\|\}$ holds for each $\xi \in \Xi^R_{j,x}$, i.e, it has uniform linear growth on $\Xi^R_{j,x}$.
For each $\nu \in \mathcal{P}_{p,K}(\Xi)$ we may continue:

$$\begin{aligned}
|\int_\Xi f_0(\xi,x)(\mu-\nu)(d\xi)| &= |\sum_{j=0}^{N} \int_{\Xi^R_{j,x}} f_0(\xi,x)(\mu-\nu)(d\xi)|\\
&\le \sum_{j=1}^{N} |\int_{\Xi^R_{j,x}} f^R_{j,x}(\xi)(\mu-\nu)(d\xi)| + I^R_x(\mu,\nu)\\
&\le N L_1 \sup_{f \in \mathcal{F}_{1,lg}(\Xi^R_{j,x}),\, j=1,\ldots,N} |\int_{\Xi^R_{j,x}} f(\xi)(\mu-\nu)(d\xi)| + I^R_x(\mu,\nu) \qquad (13)
\end{aligned}$$

where $I^R_x(\mu,\nu) := |\int_{\Xi^R_{0,x}} f_0(\xi,x)(\mu-\nu)(d\xi)|$.
For each $j = 1,\ldots,N$ the closures of the sets B_j are polyhedra with a number of faces that is bounded above by some number not depending on j, N and R. Hence, the same is true for the closures of the sets $\Xi^R_{j,x}$, i.e., for $cl\ \Xi^R_{j,x} = \{\xi \in \Xi : h(\xi) - T(\xi)x \in cl\ B_j\}$, where the corresponding number $k \in \mathbb{N}$ does, in addition, not depend on $x \in X \cap cl\,\mathcal{U}$. For each such $\Xi^R_{j,x}$ we now consider a sequence of closed polyhedra $P^R_{j,x}$, which are contained in $\Xi^R_{j,x}$ and have at most k faces, such that their characteristic functions $\chi_{P^R_{j,x}}$ converge pointwise to the characteristic function $\chi_{\Xi^R_{j,x}}$. Then the sequence consisting of the elements $|\int_\Xi f(\xi)\chi_{P^R_{j,x}}(\xi)(\mu-\nu)(d\xi)|$ converges to $|\int_\Xi f(\xi)\chi_{\Xi^R_{j,x}}(\xi)(\mu-\nu)(d\xi)|$ while

each element is bounded by $d_{1,phk}(\mu,\nu)$. Hence, the estimate (13) may be continued to

$$|\int_{\Xi} f_0(\xi,x)(\mu-\nu)(d\xi)| \le NL_1 d_{1,phk}(\mu,\nu) + I_x^R(\mu,\nu)\,. \tag{14}$$

Since there exist constants $C_1 > 0$ and $C_2 > 0$ such that

$$|f_0(\xi,x)| \le C_1 \max\{1,\|\xi\|\} \quad \text{and} \quad \|h(\xi)-T(\xi)x\| \le C_2 \max\{1,\|\xi\|\}$$

hold for each pair $(\xi,x) \in \Xi \times (X \cap cl\,\mathcal{U})$, the following upper bound for $I_x^R(\mu,\nu)$ can be derived for sufficiently large $R > 0$:

$$\begin{aligned} I_x^R(\mu,\nu) &\le \int_{\{\xi\in\Xi\,:\,\|\xi\|>\frac{R}{C_2}\}} C_1\|\xi\|(\mu+\nu)(d\xi) \\ &\le C_1(\frac{R}{C_2})^{1-p} \int_{\{\xi\in\Xi\,:\,\|\xi\|>\frac{R}{C_2}\}} \|\xi\|^p(\mu+\nu)(d\xi) \le C_1(\frac{R}{C_2})^{1-p}2K\,. \end{aligned}$$

Altogether, we obtain from (14) and the previous estimate that

$$d_{\mathcal{F},\mathcal{U}}(\mu,\nu) \le \kappa L_1 R^r d_{1,phk}(\mu,\nu) + C_1(\frac{R}{C_2})^{1-p}2K\,. \tag{15}$$

Inserting $R := d_{1,phk}(\mu,\nu)^{-\frac{1}{p-1+r}}$ for sufficiently small $d_{1,phk}(\mu,\nu)$ into (15) implies the desired estimate (12). ■

Hence, the probability metric $d_{1,phk}$ serves as a *canonical* metric for (general) linear mixed-integer two-stage stochastic programs. In case that the underlying distribution μ and their perturbations ν have their supports in some bounded subset of $\mathbb{R}^s$, the stability result is slightly improved.

Corollary 3.7 *Let the conditions (B1)–(B3) be satisfied, Ξ be bounded and $\mu \in \mathcal{P}(\Xi)$. Assume that $S(\mu)$ is nonempty and $\mathcal{U} \subseteq \mathbb{R}^m$ is an open bounded neighbourhood of $S(\mu)$. Then there exist constants $L > 0$, $\delta > 0$ and $k \in \mathbb{N}$ such that*

$$\begin{aligned} |v(\mu) - v_{\mathcal{U}}(\nu)| &\le L\,\alpha_{phk}(\mu,\nu) \\ \emptyset \ne S_{\mathcal{U}}(\nu) &\subseteq S(\mu) + \Psi(L\alpha_{phk}(\mu,\nu))\mathbb{B} \end{aligned}$$

and $S_{\mathcal{U}}(\nu)$ is a CLM set of (9) w.r.t. $\mathcal{U}$ whenever $\nu \in \mathcal{P}(\Xi)$ and $\alpha_{phk}(\mu,\nu) < \delta$.

Here, $\alpha_{phk}(\mu,\nu) := \sup\{|\mu(P)-\nu(P)| : P$ is a polyhedron with at most k faces $\}$ and Ψ is defined as in Corollary 2.8.

Proof: Since Ξ is bounded, $\mathcal{P}_{p,K}(\Xi) = \mathcal{P}(\Xi)$ holds for each $p > 1$ and some $K = K(p) > 0$, and $\alpha_{phk}(\mu,\nu) \le d_{1,phk}(\mu,\nu) \le C\alpha_{phk}(\mu,\nu)$ holds for all $\nu \in \mathcal{P}(\Xi)$ and some $C > 0$. Moreover, the term $I_x^R(\mu,\nu)$ in the previous proof vanishes for each $x \in X \cap cl\,\mathcal{U}$, $\nu \in \mathcal{P}(\Xi)$ and sufficiently large $R > 0$. Hence, (14) and Theorem 3.6 imply the assertion. ■

Remark 3.8 *Since we may assume without loss of generality that Ξ has at most k faces, convergence with respect to $d_{1,phk}$ implies weak convergence of probability measures and the convergence of first order absolute moments. Furthermore, it holds for all $\mu, \nu \in \mathcal{P}_{p,K}(\Xi)$ $(p > 1)$ that*

$$\text{(16)}\quad \alpha_{phk}(\mu,\nu) \le d_{1,phk}(\mu,\nu) \;\le\; \Big(\int_{\Xi} \max\{1, \|\xi\|\}^p(\mu+\nu)(d\xi)\Big)^{\frac{1}{p}} \alpha_{phk}(\mu,\nu)^{1-\frac{1}{p}} \;\le\; 2(1+K)^{\frac{1}{p}} \alpha_{phk}(\mu,\nu)^{1-\frac{1}{p}}\,.$$

Hence, the convergence of sequences in $\mathcal{P}_{p,K}(\Xi)$ with respect to $d_{1,phk}$ and α_{phk}, respectively, is equivalent. Moreover, by combining the estimates (11) and (16) we arrive at the continuity property

$$\text{(17)}\qquad |v(\mu) - v_{\mathcal{U}}(\nu)| \le \bar{L}\,\alpha_{phk}(\mu,\nu)^{\frac{1-\frac{1}{p}}{1+\frac{r}{p-1}}}$$

of $v_{\mathcal{U}}$ at μ with some constant $\bar{L} > 0$. For the case that T is non-stochastic and $h(\xi) = \xi$ (i.e. $r = s$), a similar result was obtained in Theorem 4.1 of [46]. However, we note that the exponent in (17) tends to 1 as $p \to \infty$, whereas the corresponding exponent in [46] is bounded by $\frac{1}{r+1}$.

3.3 Linear Chance Constrained Models

In this section, we study consequences of the general stability analysis of Section 2 to linear chance constrained stochastic programs of the form

$$\text{(18)}\qquad \min\{cx : x \in X,\, \mu(\{\xi \in \Xi : T(\xi)x \ge h(\xi)\}) \ge p\}\,,$$

where $c \in \mathbb{R}^m$, X is a polyhedron in $\mathbb{R}^m$, Ξ a polyhedron in $\mathbb{R}^s$, $p \in (0,1)$, $\mu \in \mathcal{P}(\Xi)$, and $h(\xi) \in \mathbb{R}^r$ and the (r,m)-matrix $T(\xi)$ depend affine linearly on $\xi \in \Xi$.
Setting $d = 1$, $f_0(\xi,x) = cx$, $f_1(x,\xi) = p - \chi_{H(x)}(\xi)$, where $H(x) = \{\xi \in \Xi : T(\xi)x \ge h(\xi)\}$ and χ denotes the characteristic function of subsets of Ξ, we observe that the program (18) is a particular case of the general stochastic program (1). In order to see that condition (A) is satisfied, it remains to note that the mapping $(x,\xi) \mapsto \chi_{H(x)}(\xi)$ from $\Xi \times \mathbb{R}^m$ to $\mathbb{R}$ is upper semicontinuous since the graph of H is closed. This implies that f_1 is lower semicontinuous on $\Xi \times \mathbb{R}^m$ and, hence, a normal integrand (Example 14.31 in [38]). Moreover, $p - 1 \le f_1(x,\xi) \le p$ holds and for any nonempty bounded subset $\mathcal{U}$ of $\mathbb{R}^m$ we obtain by specifying the general class of probability measures and the m.i. probability metric in Section 2:

$$\begin{aligned}
\mathcal{P}_{\mathcal{F},\mathcal{U}}(\Xi) &= \{\nu \in \mathcal{P}(\Xi) : \sup_{x \in X \cap cl\,\mathcal{U}} \max_{j=0,1} \Big|\int_{\Xi} f_j(\xi,x)\nu(d\xi)\Big| < \infty\} = \mathcal{P}(\Xi)\\
d_{\mathcal{F},\mathcal{U}}(\mu,\nu) &= \sup_{x \in X \cap cl\,\mathcal{U}} \max_{j=0,1} \Big|\int_{\Xi} f_j(\xi,x)(\mu-\nu)(d\xi)\Big|\\
&= \sup_{x \in X \cap cl\,\mathcal{U}} |\mu(H(x)) - \nu(H(x))| \quad (\mu,\, \nu \in \mathcal{P}(\Xi))
\end{aligned}$$

Such (pseudo-) metrics were already used in the stability analysis of [39, 40]. Since the sets $H(x)$ are polyhedra with a uniformly bounded number of faces, the metric α_{phk} on

$\mathcal{P}(\Xi)$ (for some $k \in I\!N$) introduced in the previous section is a natural candidate for the *canonical* metric of linear chance constrained stochastic programs. Furthermore, the following is an immediate conclusion of our general results.

Proposition 3.9 *Let $\mu \in \mathcal{P}(\Xi)$ and assume that*

(i) $S(\mu)$ is nonempty and $\mathcal{U} \subseteq I\!R^m$ is an open bounded neighbourhood of $S(\mu)$;

(ii) the mapping $x \mapsto \{y \in I\!R : \mu(\{\xi \in \Xi : T(\xi)x \geq h(\xi)\}) \geq p - y\}$ is metrically regular at each pair $(\bar{x}, 0)$ with $\bar{x} \in S(\mu)$.

Then there exist constants $L > 0$, $\delta > 0$ and $k \in I\!N$ such that

$$\begin{aligned} |v(\mu) - v_{\mathcal{U}}(\nu)| &\leq L\,\alpha_{phk}(\mu, \nu) \\ \emptyset \neq S_{\mathcal{U}}(\nu) &\subseteq S(\mu) + \Psi(\alpha_{phk}(\mu, \nu))I\!B \end{aligned}$$

and $S_{\mathcal{U}}(\nu)$ is a CLM set for (18) w.r.t. $\mathcal{U}$ whenever $\nu \in \mathcal{P}(\Xi)$ and $\alpha_{phk}(\mu, \nu) < \delta$.

Here $\Psi(\eta) := L\eta + \psi^{-1}(L\eta)\,(\eta \in I\!R_+)$ and $\psi(\tau) := \min\{cx - v(\mu) : d(x, S(\mu)) \geq \tau\,,\, x \in X \cap cl\,\mathcal{U}\,,\, \mu(H(x)) \geq p\}\,(\tau \in I\!R_+)$.

Proof: Clearly, all assumptions of Theorem 2.2 are satisfied for the special situation considered in this section. Hence, the result follows from the Theorems 2.2 and 2.3 by taking into account the estimate $d_{\mathcal{F},\mathcal{U}}(\mu, \nu) \leq \alpha_{phk}(\mu, \nu)$. ■

Remark 3.10 *Let the function $g(x) := \mu(\{\xi \in \Xi : T(\xi)x \geq h(\xi)\})$ be locally Lipschitz continuous on X. Then condition (ii) of Proposition 3.9 is satisfied if the constraint qualification $\partial(-g)(\bar{x}) \cap (-N_X(\bar{x})) = \emptyset$ holds in case of $g(\bar{x}) = p$, where "∂" denotes the Mordukhovich subdifferential and $N_X(\bar{x})$ the normal cone to X at $\bar{x} \in X$ (cf. [27]). Similar constraint qualifications were used in [40] for a specific model of the form (18) (with $r = 1$, $s = m + 1$ and a multivariate normal distribution μ) and in [16] for models with non-stochastic technology matrix $T(\cdot)$. The paper [16] also provides conditions implying second order growth of ψ.*

4 Empirical Approximations

In this section, we analyze the approximation of the stochastic programming model (1) when estimating the underlying probability distribution $\mu \in \mathcal{P}(\Xi)$ by empirical measures. Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be independent identically distributed Ξ-valued random variables on a probability space $(\Omega, \mathcal{A}, I\!P)$ having the joint distribution μ, i.e., $\mu = I\!P\xi_1^{-1}$. We consider the empirical measures

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

and the empirical approximations of the stochastic program (1), i.e.,

$$\text{(19)} \qquad \min\{\frac{1}{n}\sum_{i=1}^{n} f_0(\xi_i(\cdot), x) \,:\, x \in X\,,\, \frac{1}{n}\sum_{i=1}^{n} f_j(\xi_i(\cdot), x) \leq 0\,,\, j = 1, \ldots, d\}\,.$$

Since the objective and constraint functions of (19) are normal integrands from $\Omega \times I\!R^m$ to $\overline{I\!R}$, the constraint set is closed-valued and measurable from Ω to $I\!R^m$ and, hence, the optimal value $v(\mu_n(\cdot))$ of (19) is measurable from Ω to $\overline{I\!R}$ and the solution set $S(\mu_n(\cdot))$ is a closed-valued measurable multifunction from Ω to $I\!R^m$ (see Chapter 14 and, in particular, Theorem 14.37 of [38]). Of course, the same conclusion is valid for the corresponding localized concepts $v_{\mathcal{U}}$ and $S_{\mathcal{U}}$ for any nonempty subset $\mathcal{U}$ of $I\!R^m$.
Another measurability question arises when studying uniform convergence properties of the empirical process

$$\{(\mu_n(\cdot) - \mu)f = \frac{1}{n}\sum_{i=1}^{n}(f(\xi_i(\cdot)) - \mu f)\}_{f\in\mathcal{F}}$$

indexed by some class $\mathcal{F}$ of functions that are integrable with respect to μ. Here, we set $\nu f := \int_{\Xi} f(\xi)\nu(d\xi)$ for any $\nu \in \mathcal{P}(\Xi)$ and $f \in \mathcal{F}$. Uniform convergence properties refer to the convergence or to rates of convergence of $\sup_{f\in\mathcal{F}} |\mu_n(\cdot)f - \mu f|$ to 0 in terms of some stochastic convergence. In [53], concepts were described that allow to overcome the possible non-measurability of the supremum. To simplify matters here, we call a class $\mathcal{F}$ of measurable functions from Ξ to $\overline{I\!R}$ *permissible* for $\mu \in \mathcal{P}(\Xi)$ if $\sup_{f\in\mathcal{F}} |\mu_n(\cdot)f - \mu f|$ is a measurable mapping from Ω to $\overline{I\!R}$. To give an example, we mention that a class $\mathcal{F}$ is permissible (cf. Example 2.3.4 in [53]) if there exists a countable subset $\mathcal{F}_0$ of $\mathcal{F}$ such that for each function $f \in \mathcal{F}$ there exists a sequence (f_n) in $\mathcal{F}_0$ converging pointwise to f and such that the sequence (μf_n) also converges to μf. Then it holds that

$$d_{\mathcal{F}}(\mu_n(\omega), \mu) = \sup_{f\in\mathcal{F}} |(\mu_n(\omega) - \mu)f| = d_{\mathcal{F}_0}(\mu_n(\omega), \mu)$$

for each $n \in I\!N$ and $\omega \in \Omega$, i.e., the analysis is reduced to a countable class and, in particular, $d_{\mathcal{F}}(\mu_n(\cdot), \mu)$ is a measurable function from Ω to $\overline{I\!R}$.
Let $\mathcal{F}$ be permissible for $\mu \in \mathcal{P}(\Xi)$. Then $\mathcal{F}$ is called a *μ-Glivenko-Cantelli class* if the sequence $(d_{\mathcal{F}}(\mu_n(\cdot), \mu))$ of random variables converges to 0 $I\!P$-almost surely or, equivalently, in probability. Whether a given class $\mathcal{F}$ is a μ-Glivenko-Cantelli class or whether even a rate of convergence of $(d_{\mathcal{F}}(\mu_n(\cdot), \mu))$ is valid, depends on the size of the class $\mathcal{F}$ measured in terms of *covering* or *bracketing numbers* or the corresponding *metric entropy numbers* defined as their logarithms (see [10, 31, 53]). To introduce these concepts, let $\mathcal{F}$ be a subset of the normed space $L_p(\Xi, \mu)$ (for some $p \geq 1$) equipped with the usual norm $\|\cdot\|_p$. The covering number $N(\varepsilon, \mathcal{F}, L_p(\Xi, \mu))$ is the minimal number of open balls $\{g \in L_p(\Xi, \mu) : \|g - f\|_p < \varepsilon\}$ needed to cover $\mathcal{F}$. Given two functions f_1 and f_2 from $L_p(\Xi, \mu)$, the set $[f_1, f_2] := \{f \in L_p(\Xi, \mu) : f_1(\xi) \leq f(\xi) \leq f_2(\xi)$ for μ-almost all $\xi \in \Xi\}$ is called an ε-bracket if $\|f_1 - f_2\|_p < \varepsilon$. Then the bracketing number $N_{[\,]}(\varepsilon, \mathcal{F}, L_p(\Xi, \mu))$ is the minimal number of ε-brackets needed to cover $\mathcal{F}$. Both numbers are related by the estimate $N(\varepsilon, \mathcal{F}, L_p(\Xi, \mu)) \leq N_{[\,]}(2\varepsilon, \mathcal{F}, L_p(\Xi, \mu))$, but, in general, there is no converse inequality. It is known that $\mathcal{F} \subset L_1(\Xi, \mu)$ is a μ-Glivenko-Cantelli class if $N_{[\,]}(\varepsilon, \mathcal{F}, L_1(\Xi, \mu)) < \infty$ for each $\varepsilon > 0$ (see Theorem 6.1.5 in [10]). We also refer to [52] for further criteria of Glivenko-Cantelli classes and to [30] for applications to stochastic programming.
To state our next results, we denote the set of all real-valued random variables on

$(\Omega, \mathcal{A}, I\!\!P)$ by $\mathcal{X}(I\!\!R)$, where equality is understood as equality $I\!\!P$-almost surely, and introduce the *Ky Fan distance* κ of two real random variables $\mathcal{X}, \mathcal{Y} \in \mathcal{X}(I\!\!R)$ by

$$\kappa(\mathcal{X}, \mathcal{Y}) := \inf\{\eta \geq 0 \,:\, I\!\!P(|\mathcal{X} - \mathcal{Y}| > \eta) \leq \eta\}\,. \tag{20}$$

It is known that the infimum in (20) is attained and that κ metrizes convergence in probability in $\mathcal{X}(I\!\!R)$ (see e.g. Sect. 9.2 in [11]). By means of the Ky Fan metric the quantitative stability results of Section 2 directly translate into estimates for the empirical optimal values and solution sets.

Proposition 4.1 *Assume that the conditions (i)–(iii) of Theorem 2.2 are satisfied, that $\mathcal{F}_{\mathcal{U}}$ is permissible for μ and a μ-Glivenko-Cantelli class. Then it holds for sufficiently large $n \in I\!\!N$ that*

$$\begin{aligned} \kappa(v(\mu), v_{\mathcal{U}}(\mu_n(\cdot))) &\leq \max\{1, L\}\kappa(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu), 0) \\ \kappa(\sup_{x \in S_{\mathcal{U}}(\mu_n(\cdot))} d(x, S(\mu)), 0) &\leq \Psi(\kappa(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu), 0))\,, \end{aligned}$$

where $L > 0$ is the constant in Theorem 2.2 and Ψ the function in Theorem 2.3. Moreover, for $I\!\!P$-almost all $\omega \in \Omega$ the set $S_{\mathcal{U}}(\mu_n(\omega))$ is a CLM set of (19) with respect to $\mathcal{U}$ for sufficiently large $n \in I\!\!N$.

Proof: Let $\varepsilon_n := \kappa(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu), 0)$ and let $L > 0$, $\delta > 0$ be the constants from Theorem 2.2. Then Theorem 2.2 implies

$$\begin{aligned} I\!\!P(|v(\mu) - v_{\mathcal{U}}(\mu_n(\cdot))| > L\varepsilon_n) &\leq I\!\!P(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu) > \min\{\delta, \varepsilon_n\}) \\ &\leq I\!\!P(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu) > \varepsilon_n) \leq \varepsilon_n \end{aligned}$$

for sufficiently large $n \in I\!\!N$, since $\mathcal{F}_{\mathcal{U}}$ is a Glivenko-Cantelli class and, thus, the sequence (ε_n) tends to 0. Hence, we obtain from (20) that

$$\kappa(v(\mu), v_{\mathcal{U}}(\mu_n(\cdot))) \leq \max\{\varepsilon_n, L\varepsilon_n\}\,.$$

Now, let $\hat{\delta} > 0$ be the corresponding constant and Ψ be the function in Theorem 2.3. Then we conclude from Theorem 2.3 that

$$\begin{aligned} I\!\!P(\sup_{x \in S_{\mathcal{U}}(\mu_n(\cdot))} d(x, S(\mu)) > \Psi(\varepsilon_n)) &\leq I\!\!P(\Psi(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu)) > \min\{\hat{\delta}, \Psi(\varepsilon_n)\}) \\ &\leq I\!\!P(\Psi(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu)) > \Psi(\varepsilon_n)\}) \\ &= I\!\!P(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu)) > \varepsilon_n\}) \leq \varepsilon_n \leq \Psi(\varepsilon_n) \end{aligned}$$

for sufficiently large $n \in I\!\!N$, since it holds that $\Psi(\varepsilon_n) \geq \varepsilon_n$ (see Theorem 2.3) and since $(\Psi(\varepsilon_n))$ tends to 0.
Finally, let $\omega \in \Omega$. Then $S_{\mathcal{U}}(\mu_n(\omega))$ is nonempty, since the objective $\int_{\Xi} f_0(\xi, \cdot)\mu(d\xi)$ is lower semicontinuous on X and the constraint set $M_{\mathcal{U}}(\mu_n(\omega))$ is compact due to Proposition 2.1. Since $\mathcal{F}_{\mathcal{U}}$ is a μ-Glivenko-Cantelli class, there exists a set $A \in \mathcal{A}$ with $I\!\!P(A) = 0$ such that $(d_{\mathcal{F},\mathcal{U}}(\mu_n(\omega), \mu))$ converges to 0 and, hence, $S_{\mathcal{U}}(\mu_n(\omega)) \subseteq \mathcal{U}$ for all $\omega \in \Omega \setminus A$ and for sufficiently large $n \in I\!\!N$. This completes the proof. ■

Note that in case of a fixed constraint set (i.e. $d = 0$) both estimates in Proposition 4.1 are valid for each $n \in I\!\!N$ (without assuming that $\mathcal{F}_{\mathcal{U}}$ forms a μ-Glivenko-Cantelli class). For the specific situation of a uniformly bounded class $\mathcal{F}_{\mathcal{U}}$ we show next how these estimates may be used to derive rates of convergence.

Proposition 4.2 *Let the assumptions of Theorem 2.2 be satisfied and assume that $\mathcal{F}_{\mathcal{U}}$ is uniformly bounded and permissible for μ, and that either of the following conditions holds for some constants $r \geq 1$, $R \geq 1$ and all $\varepsilon > 0$:*

(i) $N(\varepsilon, \mathcal{F}_{\mathcal{U}}, L_2(\Xi, \nu)) \leq (\frac{R}{\varepsilon})^r$ *for any discrete $\nu \in \mathcal{P}(\Xi)$ with finite support,*

(ii) $N_{[\,]}(\varepsilon, \mathcal{F}_{\mathcal{U}}, L_2(\Xi, \mu)) \leq (\frac{R}{\varepsilon})^r$.

Then the following rates of convergence

$$\begin{aligned} \kappa(v(\mu), v_{\mathcal{U}}(\mu_n(\cdot))) &= O((\log n)^{\frac{1}{2}} n^{-\frac{1}{2}}) \\ \kappa(\sup_{x \in S_{\mathcal{U}}(\mu_n(\cdot))} d(x, S(\mu)), 0) &= O(\Psi((\log n)^{\frac{1}{2}} n^{-\frac{1}{2}})) \end{aligned}$$

are valid, where Ψ denotes the function in Theorem 2.3.

Proof: In both cases (i) and (ii) we obtain from Theorem 1.3 in [51] that

$$\mathbb{P}(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu) > \varepsilon) \leq (K(R)\varepsilon\sqrt{\frac{n}{r}})^r \exp(-2n\varepsilon^2)$$

holds for all $\varepsilon > 0$ and $n \in \mathbb{N}$. Replacing ε by $(\log n)^{\frac{1}{2}} n^{-\frac{1}{2}}$ leads to the estimate

$$\mathbb{P}(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu) > (\log n)^{\frac{1}{2}} n^{-\frac{1}{2}}) = O((\log n)^{\frac{r}{2}} n^{-2})$$

and, hence, to $\kappa(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu), 0) = O((\log n)^{\frac{1}{2}} n^{-\frac{1}{2}})$. Now, the result follows by appealing to Proposition 4.1. ∎

Both estimates and convergence rates in Propositions 4.1 and 4.2 could be formulated alternatively in terms of certain confidence bounds as in [29] and in Section 5 of [16]. However, such bounds typically contain unknown constants (like the constant δ appearing in Theorem 2.2 or the entropy constant R). Finally, we discuss applications of the results to linear two-stage and chance constrained stochastic programming models.

Example 4.3 *(linear chance constrained and mixed-integer two-stage models)*
A class $\mathcal{B}$ of Borel sets of $\mathbb{R}^s$ is called a Vapnik-Červonenkis (VC) class of index $r = r(\mathcal{B})$ if r is finite and equal to the smallest $n \in \mathbb{N}$ for which no set of cardinality $n+1$ is shattered by $\mathcal{B}$. $\mathcal{B}$ is said to shatter a subset $\{\xi_1, \ldots, \xi_k\}$ of cardinality l in $\mathbb{R}^s$ if each of its 2^l subsets is of the form $B \cap \{\xi_1, \ldots, \xi_l\}$ for some $B \in \mathcal{B}$. For VC classes $\mathcal{B}$ it holds that $N(\varepsilon, \{\chi_B : B \in \mathcal{B}\}, L_2(\Xi, \nu)) \leq K\varepsilon^{-2r}$ for each $\varepsilon > 0$ and $\nu \in \mathcal{P}(\Xi)$, and some constant $K > 0$ depending only on r (Theorem 2.6.4 in [53]).
The class of all polyhedra with a uniformly bounded number of faces is a VC class, since the class of all closed half spaces is VC and finite intersections of VC classes are again VC. The corresponding class of characteristic functions is permissible for μ, condition (i) in Proposition 4.2 is satisfied and it follows as in the proof of 4.2 from [51] that

$$\kappa(\alpha_{phk}(\mu_n(\cdot), \mu), 0) = O((\log n)^{\frac{1}{2}} n^{-\frac{1}{2}})$$

holds for each $k \in \mathbb{N}$. This rate of convergence applies directly to empirical approximations of the linear chance constrained model in Section 3.3 (see Proposition 3.9). It *also applies directly to mixed-integer two-stage models if Ξ is bounded (Corollary 3.7). In case the assumptions of Theorem 3.6 are satisfied, we obtain from (17) the rate of convergence*

$$\kappa(v(\mu), v_{\mathcal{U}}(\mu_n(\cdot))) = O((\frac{\log n}{n})^{\frac{1-\frac{1}{p}}{2(1+\frac{r}{p-1})}}) . \tag{21}$$

Example 4.4 *(two-stage models without integrality)*
Let f_0 be defined as in Section 3.1 and let (A1) and (A2) be satisfied. Then, for each nonempty bounded subset $\mathcal{U}$ of $I\!R^m$, the class $\mathcal{F}_{\mathcal{U}} = \{f_0(\cdot, x) : x \in X \cap cl\,\mathcal{U}\}$ is a subset of $L_1(\Xi, \mu)$. Furthermore, $\mathcal{F}_{\mathcal{U}}$ is permissible for μ. Due to the Lipschitz continuity property of f_0 (see Proposition 3.2)

$$|f_0(\xi, x) - f_0(\xi, \tilde{x})| \leq \hat{L} \max\{1, \|\xi\|^2\}\|x - \tilde{x}\| ,$$

for all $\xi \in \Xi$ and $x, \tilde{x} \in X$, the bracketing numbers of $\mathcal{F}_{\mathcal{U}}$ are bounded by the covering numbers of $X \cap cl\,\mathcal{U}$ (see Theorem 2.7.11 of [53]). In particular, if μ has a finite $2p$-th order moment for some $p \geq 1$, it holds with $F(\xi) := \hat{L}\max\{1, \|\xi\|^2\}$ ($\xi \in \Xi$) that

$$N_{[]}(2\varepsilon\|F\|_p, \mathcal{F}_{\mathcal{U}}, L_p(\Xi, \mu)) \leq N(\varepsilon, X \cap cl\,\mathcal{U}, I\!R^m) \leq C\varepsilon^{-m} ,$$

for each $0 < \varepsilon < 1$ and some constant $C > 0$ depending only on m and the diameter of $X \cap cl\,\mathcal{U}$. If, in particular, Ξ is a bounded subset of $I\!R^s$, the class $\mathcal{F}_{\mathcal{U}}$ is uniformly bounded and Proposition 4.2 applies, leading to the empirical rates of convergence

$$\begin{aligned} \kappa(v(\mu), v_{\mathcal{U}}(\mu_n(\cdot))) &= O((\log n)^{\frac{1}{2}} n^{-\frac{1}{2}}) \\ \kappa(\sup_{x \in S_{\mathcal{U}}(\mu_n(\cdot))} d(x, S(\mu)), 0) &= O(\Psi((\log n)^{\frac{1}{2}} n^{-\frac{1}{2}})) . \end{aligned}$$

If Ξ is unbounded, each function in $\mathcal{F}_{\mathcal{U}}$ is unbounded and Prop. 4.2 does not apply.

5 Stability of Stable Portfolios with Minimal Risk

Stable probability distributions for modelling asset returns were proposed and discussed in the fundamental work [25] and in [58], [26] for instance. We also refer to the recent monograph [33], in which many aspects of non-Gaussian stable distributions in finance are illuminated. In the following the s-dimensional random vector ϱ represents the per share returns on all assets in a given investment portfolio. We assume that ϱ follows an α-stable law with $\alpha \in (1, 2)$, i.e., its characteristic function φ on $I\!R^s$ is of the form

$$\varphi(t) = I\!E\exp(i\langle t, \varrho\rangle) = \exp\{-\int_{\Sigma^s} |\langle \xi, t\rangle|^{\alpha}(1 - i\,\mathrm{sign}\langle \xi, t\rangle \tan\frac{\pi\alpha}{2})\Gamma(d\xi) + i\langle m_0, t\rangle\},$$

where $\Sigma^s = \{\xi \in I\!R^s : \langle \xi, \xi\rangle = 1\}$ is the unit sphere in $I\!R^s$, Γ is a finite Borel measure on Σ^s (*spectral measure*), $\alpha \in (1, 2)$ is the *stability index*, $m_0 \in I\!R^s$ is the *shift* of the stable law and $\langle\cdot, \cdot\rangle$ denotes the (Euclidean) scalar product in $I\!R^s$.
Denoting by $x \in I\!R^s_+$ the proportions of the number of shares in the portfolio, the risk of the stable portfolio with vector of returns ϱ is defined as the scaled dispersion

parameter of the probability distribution of ϱ, i.e., $r_{\alpha,\Gamma}(x) := \int\limits_{\Sigma^s} |\langle x, \xi\rangle|^\alpha \Gamma(d\xi)$. For a discussion of the risk of a stable portfolio and related aspects the reader is referred to [8] and to Section 8.4 of [33]. The classical problem of the choice of the *efficient portfolio* corresponds to the optimization problem

$$\text{(22)} \qquad \min\{\int\limits_{\Sigma^s} |\langle x, \xi\rangle|^\alpha \Gamma(d\xi) \;:\; x \in I\!R^s_+,\; \sum_{i=1}^s x_i = 1\},$$

which fits into the form of the the stochastic programming model (1) by putting $m = s$, $d = 0$, $f_0(\xi, x) = |\langle x, \xi\rangle|^\alpha$, $X = \{x \in I\!R^s_+ : \sum\limits_{i=1}^s x_i = 1\}$, $\Xi = \Sigma^s$ and $\mu = \Gamma$. Here, we assume w.l.o.g. that Γ is normalized, i.e., $\Gamma(\Sigma^s) = 1$.
Now, our aim is to study the stability of portfolios with minimal risk, i.e., of solution sets $S(\alpha, \Gamma)$ to (22), when changing or estimating the stability index α and the spectral measure Γ. We start with some useful properties of the functions f_0 and $r_{\alpha,\Gamma}$.

Lemma 5.1 *The integrand $f_0(\xi, \cdot)$ is convex on $I\!R^s$ for each $\xi \in \Xi$, and for any $x, \tilde{x} \in X$ and $\xi, \tilde{\xi} \in \Xi$ it holds*

$$|f_0(\xi, x) - f_0(\tilde{\xi}, \tilde{x})| \leq \alpha(\|\xi - \tilde{\xi}\| + \|x - \tilde{x}\|)\,.$$

The solution set $S(\alpha, \Gamma)$ is nonempty and the risk satisfies the growth condition

$$\frac{1}{4}\alpha(\alpha - 1) \int_{\Sigma^s} |\langle x - x_*, z\rangle|^2 \Gamma(dz) \leq r_{\alpha,\Gamma}(x) - v(\alpha, \Gamma)$$

for all $x \in X$ and some $x_* \in S(\alpha, \Gamma)$.

Proof: First we observe that the function $g(t) = |t|^\alpha$ is strongly convex on $[-1, 1]$ with constant $\frac{1}{2}\alpha(\alpha - 1)$. Since $|\langle x, \xi\rangle| \leq \|x\|\|\xi\| = \|x\| \leq (\sum_{i=1}^s x_i)^{\frac{1}{2}} = 1$ holds for all $x \in X$ and $\xi \in \Sigma^s$, the integrand $f_0(\xi, x) = g(\langle x, \xi\rangle)$ has the property
$f_0(\xi, \frac{1}{2}x + \frac{1}{2}\tilde{x}) \leq \frac{1}{2}f_0(\xi, x) + \frac{1}{2}f_0(\xi, \tilde{x}) - \frac{1}{8}\alpha(1 - \alpha)|\langle x - \tilde{x}, \xi\rangle|^2$ for all $\xi \in \Sigma^s$ and $x, \tilde{x} \in X$.
Hence, the risk $r_{\alpha,\Gamma}$ satisfies the convexity property
$r_{\alpha,\Gamma}(\frac{1}{2}x + \frac{1}{2}\tilde{x}) \leq \frac{1}{2}r_{\alpha,\Gamma}(x) + \frac{1}{2}r_{\alpha,\Gamma}(\tilde{x}) - \frac{1}{8}\alpha(1 - \alpha) \int_{\Sigma^s} |\langle x - \tilde{x}, \xi\rangle|^2 \Gamma(d\xi)$ for all $x, \tilde{x} \in X$.
Since $r_{\alpha,\Gamma}$ is convex, $S(\alpha, \Gamma)$ is nonempty and the desired growth condition follows from the previous estimate by choosing $\tilde{x} = x_* \in S(\alpha, \Gamma)$.
Completing the proof, we obtain for all $x \in X$ and $\xi, \tilde{\xi} \in \Xi$ that

$$\begin{aligned} |f_0(\xi, x) - f_0(\tilde{\xi}, x)| &\leq \alpha \max\{|\langle x, \xi\rangle|^{\alpha-1}, |\langle x, \tilde{\xi}\rangle|^{\alpha-1}\}|\langle x, \xi\rangle - \langle x, \tilde{\xi}\rangle| \\ &\leq \alpha\|x\|^\alpha\|\xi - \tilde{\xi}\| \leq \alpha\|\xi - \tilde{\xi}\| \end{aligned}$$

and note that the roles of ξ and x may be exchanged. ■

The estimate in Lemma 5.1 shows that $\int\limits_{\Sigma^s} |\langle x_* - x^*, \xi\rangle|^2 \Gamma(d\xi) = 0$ holds for any two elements $x_*, x^* \in S(\alpha, \Gamma)$. This motivates to call a spectral measure $\Gamma \in \mathcal{P}(\Sigma^s)$ *nonsingular* if the the relation $\int\limits_{\Sigma^s} |\langle x, \xi\rangle|^2 \Gamma(d\xi) = 0$ implies $x = 0$. Example 5.3 illustrates that relevant spectral measures are nonsingular, indeed.

Theorem 5.2 *For each* $(\alpha, \Gamma) \in (1,2) \times \mathcal{P}(\Sigma^s)$ *there exists a constant* $\delta > 0$ *such that*

$$\sup_{x \in S(\tilde{\alpha}, \tilde{\Gamma})} d(x, S(\alpha, \Gamma)) \leq \Psi(\alpha \zeta_1(\Gamma, \tilde{\Gamma}) + \mathrm{e}^{-1}|\alpha - \tilde{\alpha}|)$$

whenever $(\tilde{\alpha}, \tilde{\Gamma}) \in (1,2) \times \mathcal{P}(\Sigma^s)$ *and* $\zeta_1(\Gamma, \tilde{\Gamma}) + |\alpha - \tilde{\alpha}| < \delta$.
Moreover, if Γ *is nonsingular, then (22) has a unique solution and* $\Psi(\eta) \leq C\eta^{\frac{1}{2}}$ *holds for some* $C > 0$ *and sufficiently small* $\eta \in I\!R_+$.

Here, $\Psi(\eta) := \eta + \psi^{-1}(2\eta)\,(\eta \in I\!R_+)$, $\psi(\tau) := \min\{\int_{\Sigma^s} |\langle x, \xi\rangle|^\alpha \Gamma(d\xi) - v(\alpha, \Gamma) : d(x, S(\alpha, \Gamma)) \geq \tau\,,\, x \in X\}\,(\tau \in I\!R_+)$ *and* $\mathrm{e} := \exp(1)$.

Proof: First we show that the mapping $(x; \alpha, \Gamma) \mapsto \int_{\Sigma^s} |\langle x, \xi\rangle|^\alpha \Gamma(d\xi)$ satisfies a Lipschitz property. For all $x \in X$, $\alpha, \tilde{\alpha} \in (1,2)$ and $\Gamma, \tilde{\Gamma} \in \mathcal{P}(\Sigma^s)$ we obtain

$$\begin{aligned} |r_{\alpha,\Gamma}(x) - r_{\tilde{\alpha},\tilde{\Gamma}}(x)| &= |\int_{\Sigma^s} |\langle x, \xi\rangle|^\alpha \Gamma(d\xi) - \int_{\Sigma^s} |\langle x, \xi\rangle|^{\tilde{\alpha}} \tilde{\Gamma}(d\xi)| \\ &\leq \alpha\zeta_1(\Gamma, \tilde{\Gamma}) + \int_{\Sigma^s} ||\langle x, \xi\rangle|^\alpha - |\langle x, \xi\rangle|^{\tilde{\alpha}}| \tilde{\Gamma}(d\xi) \\ &\leq \alpha\zeta_1(\Gamma, \tilde{\Gamma}) + \sup_{t \in [-1,1]} ||t|^\alpha - |t|^{\tilde{\alpha}}| \leq \alpha\zeta_1(\Gamma, \tilde{\Gamma}) + \mathrm{e}^{-1}|\alpha - \tilde{\alpha}|\,. \end{aligned}$$

Then the first part of the result is an immediate consequence of Lemma 5.1 and Theorem 2.3. The additional assumption implies that the mapping $x \mapsto (\int_{\Sigma^s} |\langle x, \xi\rangle|^2 \Gamma(d\xi))^{\frac{1}{2}}$ from $I\!R^s$ to $I\!R$ is a norm on $I\!R^s$. Hence, there exists a constant $c = c(s, \Gamma) > 0$ such that $c\|x\|^2 \leq \int_{\Sigma^s} |\langle x, \xi\rangle|^2 \Gamma(d\xi)$ holds for all $x \in I\!R^s$. We conclude that the risk function $r_{\alpha,\Gamma}$ is strongly convex on X and that (22) has a unique solution $x_* \in X$. Furthermore, Lemma 5.1 implies the estimate

$$\frac{1}{4} c\, \alpha(\alpha - 1)\|x - x_*\|^2 \leq r_{\alpha,\Gamma}(x) - v(\alpha, \Gamma)$$

for each $x \in X$. Hence, $\psi(\tau) = \frac{1}{4}c\alpha(\alpha - 1)\tau^2$ for each $\tau \in I\!R_+$. Referring to Remark 2.4 completes the proof. ■

Example 5.3 *(Nonsingularity of discrete spectral measures)*
A spectral measure Γ *of an* α*-stable random vector* ϱ *is concentrated on a finite number of points on the unit sphere* Σ^s *iff* ϱ *can be expressed as a linear transformation of independent* α*-stable real random variables (Proposition 2.3.7 of [43]).*
Now, let $\varrho^{(k)}$ *be* d *independent* α*-stable real random variables and* $A = (a_{jk})$ *be a real* (s, d)*-matrix with rank equal to* s*. Then the random vector* ϱ *with components* $\varrho_j = \sum_{k=1}^{d} a_{jk}\varrho^{(k)}$ $(j = 1, \ldots, s)$ *is* α*-stable with (discrete) spectral measure*

$$\Gamma = \sum_{k=1}^{d} \{\frac{1}{2}(1 + \beta_k)\gamma_k\delta_{s_k} + \frac{1}{2}(1 - \beta_k)\gamma_k\delta_{-s_k}\}\,,$$

where $\beta_k \in I\!R$, $\gamma_k \geq 0$ depend on the parameters of $\varrho^{(k)}$ and on A, and the vectors $s_k \in \Sigma^s$ are normalized columns of A (see Example 2.3.6 of [43]). Hence, it holds

$$\int_{\Sigma^s} |\langle x, \xi \rangle|^2 \Gamma(d\xi) = \sum_{k=1}^{d} \gamma_k |\langle x, s_k \rangle|^2 \,,$$

and Γ is nonsingular since $span\{s_1, \ldots, s_d\} = I\!R^s$.

Let us finally consider empirical estimates α_n and Γ_n of some unknown pair $(\alpha, \Gamma) \in (1,2) \times \mathcal{P}(\Sigma^s)$ of parameters of a stable random vector ϱ. If the spectral measure Γ is nonsingular, Proposition 4.1 and Theorem 5.2 imply that

$$\kappa(\sup_{x \in S(\alpha_n, \Gamma_n)} d(x, S(\alpha, \Gamma)), 0) = O(\kappa(\zeta_1(\Gamma_n, \Gamma), 0)^{\frac{1}{2}} + \kappa(\alpha_n, \alpha)^{\frac{1}{2}}) = O((\log n)^{\frac{1}{4}} n^{-\frac{1}{4}}) \,,$$

where, as in the proof of Proposition 4.2, the specific rate of convergence follows from an estimate of the bracketing number of the set of Lipschitz continuous functions on Σ^s (see also Example 4.4) and from the classical limit theory of empirical estimates in $I\!R$. Finally, we note that Theorem 5.2 also applies for studying convergence properties of the estimation procedures desribed in Chapter 8 of [33].

Acknowledgement:
This research was supported by the Schwerpunktprogramm *Echtzeit-Optimierung großer Systeme* of the Deutsche Forschungsgemeinschaft. It was carried out during mutual visits of both authors to the University of California at Santa Barbara and the Humboldt-University Berlin. The authors should like to thank the mentioned institutions for their hospitality and support.

References

[1] Z. Artstein, Sensitivity with respect to the underlying information in stochastic programs, *Journal of Computational and Applied Mathematics* 56 (1994), 127–136.

[2] 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.

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

[4] 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.

[5] B. Bank, J. Guddat, D. Klatte, B. Kummer and K. Tammer, *Non-Linear Parametric Optimization*, Akademie-Verlag, Berlin 1982.

[6] C.E. Blair and R.G. Jeroslow, The value function of a mixed integer program, *Discrete Mathematics* 19 (1977), 121–138.

[7] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York 2000.

[8] B.N. Cheng and S.T. Rachev, Multivariate stable commodities in the futures market, *Mathematical Finance* 5 (1995), 133–153.

[9] D. Dentcheva and W. Römisch, Differential stability of two-stage stochastic programs, *SIAM Journal on Optimization* (to appear).

[10] R.M. Dudley, A course on empirical processes, in: École d'Été de Probabilities de Saint-Flour XII – 1982, *Lecture Notes in Mathematics* 1097, Springer-Verlag, Berlin 1984, 2–142.

[11] R.M. Dudley, *Real Analysis and Probability*, Wadsworth & Brooks/Cole, Pacific Grove 1989.

[12] J. Dupačová, Stability and sensitivity analysis for stochastic programming, *Annals of Operations Research* 27 (1990), 115-142.

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

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

[15] N. Gröwe, Estimated stochastic programs with chance constraints, *European Journal of Operational Research* 101 (1997), 285–305.

[16] R. Henrion and W. Römisch, Metric regularity and quantitative stability in stochastic programs with probabilistic constraints, *Mathematical Programming* 84 (1999), 55–88.

[17] P. Kall, *Stochastic Linear Programming*, Springer-Verlag, Berlin 1976.

[18] 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.

[19] 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.

[20] V. Kaňková, A note on estimates in stochastic programming, *Journal of Computational and Applied Mathematics* 56 (1994), 97–112.

[21] A.J. King and R.T. Rockafellar, Asymptotic theory for solutions in statistical estimation and stochastic programming, *Mathematics of Operations Research* 18 (1993), 148–162.

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

[23] D. Klatte, A note on quantitative stability results in nonlinear optimization, in: *Proceedings 19. Jahrestagung Mathematische Optimierung* (K. Lommatzsch Ed.), Humboldt-Universität Berlin, Sektion Mathematik, Seminarbericht Nr. 90, 1987, 77–86.

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

[25] B.B. Mandelbrot, New methods in statistical economics, *Journal of Political Economy* 71 (1963), 421–440.

[26] S. Mittnik and S.T. Rachev, Modeling asset returns with alternative stable distributions, Econometric Reviews 12 (1993), 261–330.

[27] B.S. Mordukhovich, Lipschitzian stability of constraint systems and generalized equations, *Nonlinear Analysis, Theory, Mehods & Applications* 22 (1994), 173–206.

[28] F. Nožička, J. Guddat, H. Hollatz and B. Bank, *Theorie der linearen parametrischen Optimierung*, Akademie-Verlag, Berlin 1974.

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

[30] 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.

[31] D. Pollard, Empirical Processes: Theory and Applications, *NSF-CBMS Regional Conference Series in Probability and Statistics Vol. 2*, Institute of Mathematical Statistics, 1990.

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

[33] S.T. Rachev and S. Mittnik, *Stable Paretian Models in Finance*, Wiley, Chichester 2000.

[34] S.T. Rachev and L. Rüschendorf, *Mass Transportation Problems, Vol. I: Theory*, Springer, New York 1998.

[35] S.M. Robinson, Local epi-continuity and local optimization, *Mathematical Programming* 37 (1987), 208–223.

[36] S.M. Robinson, Analysis of sample-path optimization, *Mathematics of Operations Research* 21 (1996), 513–528.

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

[38] R.T. Rockafellar and R. J-B Wets, *Variational Analysis*, Springer, Berlin 1997.

[39] W. Römisch and R. Schultz, Distribution sensitivity in stochastic programming, *Mathematical Programming* 50 (1991), 197–226.

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

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

[42] 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.

[43] G. Samorodnitsky and M.S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman&Hall, New York, 1994.

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

[45] R. Schultz, On structure and stability in stochastic programs with random technology matrix and complete integer recourse, *Mathematical Programming* 70 (1995), 73–89.

[46] R. Schultz, Rates of convergence in stochastic programs with complete integer recourse *SIAM Journal on Optimization* 6 (1996), 1138–1152.

[47] R. Schultz, Some aspects of stability in stochastic programming, *Annals of Operations Research* (to appear).

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

[49] A. Shapiro, Simulation-based optimization – Concergence analysis and statistical inference, *Communications in Statistics – Stochastic Models* 12 (1996), 425–454.

[50] A. Shapiro and T. Homem-de-Mello, On rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs, *SIAM Journal on Optimization* 11 (2000), 70–86.

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

[52] M. Talagrand, The Glivenko-Cantelli problem, ten years later, *Journal of Theoretical Probability* 9 (1996), 371–384.

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

[54] S. Vogel, A stochastic approach to stability in stochastic programming, *Journal of Computational and Applied Mathematics* 56 (1994), 65–96.

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

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

[57] R. J-B Wets, Stochastic programming, in: *Handbooks in Operations Research and Management Science, Vol. 1, Optimization* (G.L. Nemhauser, A.H.G. Rinnooy Kan, M.J. Todd Eds.), North-Holland, Amsterdam 1989, 573–629.

[58] W.T. Ziemba, Choosing investment portfolios when the returns have stable distributions, in: *Mathematical Programming in Theory and Practice* (P.L. Hammer, G. Zoutendijk Eds.), North-Holland, 1974, 443–482.

[59] V.M. Zolotarev, Probability metrics, *Theory of Probability and its Applications* 28 (1983), 278–302.