MATHEMATICS OF OPERATIONS RESEARCH
Vol. 27, No. 4, November 2002, pp. 792–818
Printed in U.S.A.

QUANTITATIVE STABILITY IN STOCHASTIC PROGRAMMING: THE METHOD OF PROBABILITY METRICS

SVETLOZAR T. RACHEV AND WERNER RÖMISCH

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.

1. Introduction. Stochastic programming is concerned with models for optimization problems under (stochastic) uncertainty which requires 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:

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

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 $j=0,\dots,d$, 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 $\xi\mapsto \operatorname{epi} f_j(\xi,\cdot):=\{(x,r)\in\mathbb{R}^m\times\mathbb{R}: f_j(\xi,x)\le r\}$ 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., Rockafellar and Wets 1997).

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

Received December 24, 2000; revised January 10, 2002.
MSC 2000 subject classification. Primary: 90C15, 90C31, 60E05, 60B10.
OR/MS subject classification. Primary: Programming/stochastic.
Key words. Stochastic programming, quantitative stability, probability metrics, Fortet-Mourier metrics, empirical approximations, two-stage models, mixed-integer, chance-constrained models, stable portfolio models.

0364-765X/02/2704/792/$05.00
1526-5471 electronic ISSN,

$S(\mu)$, respectively, i.e.,

$$M(\mu)=\left\{x\in X\colon\ \int_{\Xi} f_j(\xi,x)\,\mu(d\xi)\le 0\,,\ j=1,\ldots,d\right\},$$

$$v(\mu)=\inf\left\{\int_{\Xi} f_0(\xi,x)\,\mu(d\xi)\,:\,x\in M(\mu)\right\},$$

$$S(\mu)=\left\{x\in M(\mu)\colon\ \int_{\Xi} f_0(\xi,x)\,\mu(d\xi)=v(\mu)\right\}.$$

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 this paper, *stability* refers to quantitative continuity properties of the optimal value function $v(\cdot)$ and the solution-set mapping $S(\cdot)$ at μ, where both $v(\cdot)$ and $S(\cdot)$ are regarded as mappings given on a certain set of probability measures.

We illustrate the abstract problem by two examples. The first is described in §5, namely, the classical problem of finding a portfolio with minimal risk in case the asset returns are modeled by a multivariate stable probability distribution. The risk depends on the (normalized) spectral measure of the underlying stable distribution and is incompletely known or even unknown in practical situations. Hence, it is desirable that the optimal portfolio does not change too much if one works with an approximation of the spectral measure instead of the true one. The second example is a variant of the classical newsboy problem (see, e.g., Artstein and Wets 1994, Dupačová 1994), and is described next.

EXAMPLE 1.1. A newsboy must place a daily order for a number x of copies of a newspaper. He has to pay r dollars for each copy and sells a copy at c dollars, where $0<r<c$. The daily demand ξ is random with (discrete) probability distribution $\mu\in\mathscr{P}(\mathbb{N})$ and the remaining copies $y(\xi)=\max\{0,x-\xi\}$ have to be removed. The newsboy wishes to minimize his expected costs, i.e., he minimizes

$$\begin{aligned}\int_{\mathbb{R}} f_0(\xi,x)\,\mu(d\xi) &:= \int_{\mathbb{R}}[(r-c)x+c\max\{0,x-\xi\}]\,\mu(d\xi)\\ &= (r-c)x+c\sum_{k\in\mathbb{N}}\pi_k\max\{0,x-k\},\\ &= rx-cx\sum_{\substack{k\in\mathbb{N}\\ k\ge x}}\pi_k-\sum_{\substack{k\in\mathbb{N}\\ k<x}}\pi_k k,\end{aligned}$$

where π_k is the probability of demand $k\in\mathbb{N}$. In order to maintain convexity of the model, we consider decisions x in $\mathbb{R}_+$. The infimum is attained at $\bar{x}$ being the largest $x\in\mathbb{R}_+$, such that $\mu(\{\xi\ge x\})=\sum_{k\in\mathbb{N},k\ge x}\pi_k\le r/c$. However, the newsboy does not know the probability distribution μ and he has to use some approximation instead. For instance, his decision might be based on n independent identically distributed observations ξ_i, $i=1,\ldots,n$, of the demand, i.e., on approximating μ by the empirical measure μ_n (cf., §4) and on solving the approximate problem,

$$\min_{x\in\mathbb{R}_+}\left\{(r-c)x+\frac{c}{n}\sum_{i=1}^{n}\max\{0,x-\xi_i\}\right\}. \tag{2}$$

Of course, this approach is only justified if some optimal solutions x_n of the approximate problems are close to $\bar{x}$ for sufficiently large n. The newsboy problem is a specific two-stage stochastic program. Its discussion will be continued in Examples 2.10 and 4.6.

In order to study continuity of $v(\cdot)$ and $S(\cdot)$ at the original distribution μ, the choice of a proper distance for probability distributions becomes important. Fortunately, there exists a diversity of probability metrics addressing different goals and based on various constructions (see, e.g., Rachev 1991). For the perturbation analysis of Model (1), a distance which compares expectations of a variety of nonlinear functions seems to be a natural choice. More precisely, distances having the form

$$d_{\mathcal{F}}(\mu,\nu) = \sup_{f\in\mathcal{F}} \left| \int_{\Xi} f(\xi)\,\mu(d\xi) - \int_{\Xi} f(\xi)\,\nu(d\xi) \right|, \tag{3}$$

where $\mathcal{F}$ denotes a class of measurable functions from Ξ to $\overline{\mathbb{R}}$ and μ, ν belongs to $\mathcal{P}_{\mathcal{F}}$, will be suitable in our stability framework. Such a distance (3) is called *Zolotarev's pseudometric* or a distance having *ζ-structure* (see, Zolotarev 1983, Rachev 1991). Clearly, $d_{\mathcal{F}}$ satisfies all properties of a pseudometric on $\mathcal{P}_{\mathcal{F}}$, 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 analyzing 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 (3) with the class $\mathcal{F}$ given by $\{f_j(\cdot,x)\colon x\in X\cap \operatorname{cl}\mathcal{U},\ j=0,\dots,d\}$, where $\mathcal{U}$ is a properly chosen open subset of $\mathbb{R}^m$, forms a *minimal information (m.i.) metric* for quantitative stability of (1). The corresponding results (Theorems 2.3 and 2.4) 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 Attouch and Wets (1993), Rockafellar and Wets (1997), and Klatte (1987, 1994).

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 controlling 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 (as, e.g., $\mathcal{F}=L_1(\Xi,\mu)$). Although it appears somewhat vague here, such properly enlarged classes $\mathcal{F}_c$ will be called *canonical classes*, and the corresponding ζ-distance d_c an *ideal* or *canonical metric*. In §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 pth order *Fortet-Mourier metric* ζ_p ($p\ge 1$) defined on $\mathcal{P}_p(\Xi)$ by

$$\zeta_p(\mu,\nu) := \sup_{f\in\mathcal{F}_p(\Xi)} \left| \int_{\Xi} f(\xi)(\mu-\nu)(d\xi) \right| \tag{4}$$

for $\mu,\nu\in\mathcal{P}_p(\Xi) := \{\nu\in\mathcal{P}(\Xi)\colon \int_{\Xi}\|\xi\|^p\,\nu(d\xi)<\infty\}$ (see, Fortet and Mourier 1953 and Rachev 1991). Here,

$$\begin{aligned}\mathcal{F}_p(\Xi) := \big\{ f:\Xi\longmapsto\mathbb{R} : |f(\xi)-f(\tilde\xi)| \\ \le \max\{1,\|\xi\|^{p-1},\|\tilde\xi\|^{p-1}\}\|\xi-\tilde\xi\|,\ \forall\,\xi,\tilde\xi\in\Xi\big\}.\end{aligned}$$

It is known that a sequence of measures in $\mathcal{P}_p(\Xi)$ converges with respect to ζ_p iff it converges weakly and the corresponding sequence of pth absolute moments converges as well (Theorem 6.2.1 in Rachev 1991). Later on, such 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 §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 §§3.2 and 3.3).

When using stability results for designing or analyzing approximation schemes or estimation procedures, further properties of metrics or function classes have to be derived sometimes. This issue is addressed in a *third step*. In §4, for example, we derive entropy numbers of certain function classes and new rates of convergence for empirical approximations of stochastic programs using m.i. metrics.

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 Kall (1987), Robinson and Wets (1987), Wets (1989), and Artstein and Wets (1994) and to the work on quantitative stability in Römisch and Wakolbinger (1987), Römisch and Schultz (1991), and Artstein (1994) (see also the overviews Dupačová 1990 and Schultz 2000). First attempts at finding suitable probability metrics for specific models were undertaken in Römisch and Schultz (1991) and Schultz (1996). 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., Dupačová and Wets 1988, King and Wets 1991, Vogel 1994, Artstein and Wets 1995, Robinson 1996, and Pflug 1998 for qualitative results, and Kaňková 1994, King and Rockafellar 1993, Kaniovski et al. 1995, Shapiro 1996, Gröwe 1997, and Shapiro and Homen-de-Mello 2000 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 Henrion and Römisch 1999 and Pflug 1999) 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 §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 §4, together with applications to the models in §3. In §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 probability distribution μ and the perturbed model

$$\min\left\{\int_\Xi f_0(\xi, x)\,\nu(d\xi)\colon\ x\in X,\ \int_\Xi f_j(\xi, x)\,\nu(d\xi)\le 0,\ j=1,\ldots,d\right\}. \tag{5}$$

We assume throughout this section that the sets X and Ξ are closed and that the functions f_j $(j = 0, \ldots, d)$ are normal integrands. For any nonempty open set $\mathcal{U} \subseteq \mathbb{R}^m$, we consider the following sets of functions, elements, and probability measures:

$$\mathcal{F}_{\mathcal{U}} := \{f_j(\cdot, x) : x \in X \cap \operatorname{cl} \mathcal{U}, j = 0, ..., d\},$$

$$M_{\mathcal{U}}(\nu) := \left\{x \in X \cap \operatorname{cl} \mathcal{U}\colon \int_{\Xi} f_j(\xi, x)\,\nu(d\xi) \le 0, j = 1, \dots, d\right\} \qquad (\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}(\Xi)),$$

$$\mathcal{P}_{\mathcal{F},\mathcal{U}}(\Xi) := \left\{\nu \in \mathcal{P}(\Xi) : -\infty < \int_{\Xi} \inf_{\substack{x\in X\\ \|x\|\le r}} f_j(\xi, x)\,\nu(d\xi) \quad \text{for each } r > 0 \quad \text{and} \right.$$

$$\left. \sup_{x\in X\cap \operatorname{cl}\mathcal{U}} \int_{\Xi} f_j(\xi, x)\,\nu(d\xi) < \infty \quad \text{for each } j = 0, \dots, d\right\},$$

and the probability pseudometric on $\mathcal{P}_{\mathcal{F},\mathcal{U}} := \mathcal{P}_{\mathcal{F},\mathcal{U}}(\Xi)$:

$$d_{\mathcal{F},\mathcal{U}}(\mu, \nu) := \sup_{f\in\mathcal{F}_{\mathcal{U}}} \left| \int_{\Xi} f(\xi)(\mu-\nu)\,(d\xi)\right|$$

$$= \sup_{x\in X\cap\operatorname{cl}\mathcal{U}} \max_{j=0,\dots,d} \left| \int_{\Xi} f_j(\xi, x)(\mu-\nu)\,(d\xi)\right|.$$

Our general assumptions and the Fatou Lemma imply that the objective function of (5) is lower semicontinuous on X and the constraint set of (5) is closed for each $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}(\Xi)$. Our first results compile some further basic properties of the model (5).

PROPOSITION 2.1. *Let $\mathcal{U}$ be a nonempty open subset of $\mathbb{R}^m$. Then the mapping $(x, \nu) \mapsto \int_{\Xi} f_j(\xi, x)\,\nu(d\xi)$ from $(X \cap \operatorname{cl}\mathcal{U}) \times (\mathcal{P}_{\mathcal{F},\mathcal{U}}, d_{\mathcal{F},\mathcal{U}})$ to $\overline{\mathbb{R}}$ is lower semicontinuous for each $j = 0, \dots, d$.*

PROOF. Let $j = 0, \dots, d$, $x \in X \cap \operatorname{cl}\mathcal{U}$, $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$, (x_n) be a sequence in $X \cap \operatorname{cl}\mathcal{U}$ such that $x_n \to x$ and let (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

$$\int_{\Xi} f_j(\xi, x)\,\nu(d\xi) \le \liminf_{n\to\infty} \int_{\Xi} f_j(\xi, x_n)\,\nu(d\xi)$$

$$\le \liminf_{n\to\infty} \left\{ d_{\mathcal{F},\mathcal{U}}(\nu, \nu_n) + \int_{\Xi} f_j(\xi, x_n)\,\nu_n(d\xi)\right\}$$

$$= \liminf_{n\to\infty} \int_{\Xi} f_j(\xi, x_n)\,\nu_n(d\xi),$$

completing the proof. □

PROPOSITION 2.2. *Let $\mathcal{U}$ be a nonempty open subset of $\mathbb{R}^m$. Then 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 $\mathbb{R}^m$ is closed.*

PROOF. Let (ν_n) be a sequence converging to ν in $(\mathcal{P}_{\mathcal{F},\mathcal{U}}, d_{\mathcal{F},\mathcal{U}})$ and let (x_n) be a sequence converging to x in $\mathbb{R}^m$ such that $x_n \in M_{\mathcal{U}}(\nu_n)$ for each $n \in \mathbb{N}$. Clearly, we have $x \in X \cap \operatorname{cl}\mathcal{U}$. For $j \in \{1, \dots, d\}$, we obtain from Proposition 2.1 that

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

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., Rockafellar and Wets 1997 and

Henrion and Römisch 1999), we consider the set-valued mapping $y \mapsto M_y(\mu)$ from $\mathbb{R}^d$ to $\mathbb{R}^m$, where

$$M_y(\mu) = \left\{ x \in X\colon \int_\Xi f_j(\xi, x)\,\mu(d\xi) \le y_j,\ j = 1, \dots, d \right\},$$

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

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

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 Robinson (1987) and Klatte (1987). For any nonempty set $\mathcal{U} \subseteq \mathbb{R}^m$ and any $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$, we set

$$v_\mathcal{U}(\nu) = \inf\left\{ \int_\Xi f_0(\xi, x)\,\nu(d\xi)\colon x \in M_\mathcal{U}(\nu) \right\},$$

$$S_\mathcal{U}(\nu) = \left\{ x \in M_\mathcal{U}(\nu)\colon \int_\Xi f_0(\xi, x)\,\nu(d\xi) = v_\mathcal{U}(\nu) \right\}.$$

A nonempty set $\mathcal{S} \subseteq \mathbb{R}^m$ is called a *complete local minimizing* (CLM) *set* of (5) 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 main stability result. Although its proof partly parallels arguments in the proof of Theorem 1 of Klatte (1987), we include it here since our model assumptions are slightly more general compared to those in Klatte (1987, 1994).

THEOREM 2.3. *Let $\mu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$ 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 $d \ge 1$, the function $x \mapsto \int_\Xi f_0(\xi, x)\,\mu(d\xi)$ is Lipschitz continuous on $X \cap \mathrm{cl}\,\mathcal{U}$;*

(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

$$|v(\mu) - v_\mathcal{U}(\nu)| \le L\, d_{\mathcal{F},\mathcal{U}}(\mu, \nu) \tag{6}$$

holds and $S_\mathcal{U}(\nu)$ is a CLM set of (5) 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 (6) is valid with $L = 1$ and for all $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$.

PROOF. We consider the (localized) parametric optimization problem

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

where the probability measure ν is regarded as a parameter varying in the pseudometric space $\mathcal{P}_{\mathcal{F},\mathcal{U}}$. Proposition 2.2 says 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 $\mathrm{cl}\,\mathcal{U}$ is compact. Furthermore, we know by Proposition 2.1 that the function g from $(X \cap \mathrm{cl}\,\mathcal{U}) \times \mathcal{P}_{\mathcal{F},\mathcal{U}}$ 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

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

holds and Berge's classical stability analysis (see Bank et al. 1982, Theorem 4.2.1(3)) 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 $(X \cap \mathrm{cl}\,\mathcal{U}) \times \mathcal{P}_{\mathcal{F},\mathcal{U}}$. Then we may conclude from Theorem 4.2.1 in Bank et al. (1982) 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 $\bar{x} \in S(\mu)$:

$$(7) \qquad M_{\mathcal{U}}(\mu) \cap B(\bar{x}, \bar{\varepsilon}) \subseteq M_{\mathcal{U}}(\nu) + a d_{\mathcal{F},\mathcal{U}}(\mu, \nu)\mathbb{B}, \quad \text{whenever } d_{\mathcal{F},\mathcal{U}}(\mu, \nu) < \bar{\varepsilon},$$

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 (7), 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 $x \in X \cap \mathrm{cl}\,\mathcal{U}$, $j \in \{1, \dots, d\}$ 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 $\mathbb{B}(\bar{x}, (a+1)\bar{\varepsilon}) \subseteq \mathcal{U}$. Hence, we have $\mathbb{B}(x, a\bar{\varepsilon}) \subseteq \mathcal{U}$ 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 $y_j = -d_{\mathcal{F},\mathcal{U}}(\mu, \nu)$, $j = 1, \dots, d$, the above estimate implies that $M_y(\mu) \cap \mathrm{cl}\,\mathcal{U} \subseteq M_{\mathcal{U}}(\nu)$. Because of the choice of $\bar{\varepsilon}$, we have $d(x, M_y(\mu) \cap \mathrm{cl}\,\mathcal{U}) = d(x, M_y(\mu))$ 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

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

which is equivalent to the property (7). 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 (5) w.r.t. $\mathcal{U}$ for each such ν.

Moreover, we obtain from (iii), for any $x \in M_{\mathcal{U}}(\nu) \cap \mathbb{B}(\bar{x}, \bar{\varepsilon})$, the estimate

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

which is equivalent to the inclusion

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

Since $S(\mu)$ is compact, we may continue as in the proof of Theorem 1 of Klatte (1987) 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 $\mathbb{B}(\bar{x}, \bar{\varepsilon})$ 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 $\|\tilde{x} - \bar{x}\| \le \hat{a} d_{\mathcal{F},\mathcal{U}}(\mu, \nu)$. For $d \ge 1$, we continue

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

where $L_g \ge 0$ denotes a Lipschitz constant of $g(\cdot, \mu)$ on $X \cap \operatorname{cl} \mathcal{U}$. 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 $\hat{a} = 0$. □

The previous result asserts that the multifunction $S_{\mathcal{U}}$ is nonempty near μ and (Berge) upper semicontinuous at μ. In order to quantify this upper semicontinuity property, a growth condition on the objective function in a neighbourhood of the solution set to the original problem (1) is needed (e.g., Attouch and Wets 1993 and Klatte 1994). Instead of imposing a specific growth condition (as, e.g., quadratic growth), we consider the *growth function*,

$$(8) \quad \psi_\mu(\tau) := \min\left\{ \int_\Xi f_0(\xi, x)\, \mu(d\xi) - v(\mu) \colon d(x, S(\mu)) \ge \tau, x \in M_{\mathcal{U}}(\mu) \right\} \quad (\tau \in \mathbb{R}_+),$$

of Problem (1) on $\operatorname{cl} \mathcal{U}$, i.e., near its solution set $S(\mu)$, and the associated function,

$$(9) \qquad \Psi_\mu(\eta) := \eta + \psi_\mu^{-1}(2\eta) \qquad (\eta \in \mathbb{R}_+),$$

where we set $\psi_\mu^{-1}(t) := \sup\{\tau \in \mathbb{R}_+ \colon \psi_\mu(\tau) \le t\}$. Both functions ψ_μ and Ψ_μ depend on the data of (1), and in particular, on μ. They are lower semicontinuous on $\mathbb{R}_+$; ψ_μ is nondecreasing, Ψ_μ increasing, and both vanish at 0 (cf., Rockafellar and Wets 1997, Theorem 7.64). Our second main stability result establishes a quantitative upper semicontinuity property of (localized) solution sets and identifies the function Ψ_μ as modulus of semicontinuity. Parts of its proof are similar to arguments in the proof of Theorem 7.64 in Rockafellar and Wets (1997).

THEOREM 2.4. *Let the assumptions of Theorem* 2.3 *be satisfied and* $\mu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$. *Then there exists a constant* $\widehat{L} \ge 1$, *such that it holds for any* $\nu \in \mathcal{P}_{\mathcal{F},\mathcal{U}}$ *that*

$$(10) \qquad \emptyset \ne S_{\mathcal{U}}(\nu) \subseteq S(\mu) + \Psi_\mu(\widehat{L} d_{\mathcal{F},\mathcal{U}}(\mu, \nu))\mathbb{B},$$

where $\mathbb{B}$ *is the closed unit ball in* $\mathbb{R}^m$ *and* Ψ_μ *is given by* (9). *In case* $d = 0$, *the estimate* (10) *is valid with* $\widehat{L} = 1$.

PROOF. Let $L > 0$ be the constant in Theorem 2.3, 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 2.3, there exists an element $\bar{x} \in M_{\mathcal{U}}(\mu)$ such that $\|\tilde{x} - \bar{x}\| \le \hat{a}\delta$, where $\delta := d_{\mathcal{F},\mathcal{U}}(\mu, \nu)$. Let $L_\mu \ge 0$ denote a Lipschitz constant of the function $x \mapsto \int_\Xi f_0(\xi, x)\, \mu(d\xi)$ on $X \cap \operatorname{cl} \mathcal{U}$. Then the definition of ψ and Theorem 2.3 imply that

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

Hence, we obtain

$$\begin{aligned} d(\tilde{x}, S(\mu)) &\le \hat{a}\delta + d(\tilde{x}, S(\mu) + \hat{a}\delta\mathbb{B}) \\ &\le \hat{a}\delta + \psi_\mu^{-1}(\delta(1 + L_\mu \hat{a} + L)) \\ &\le \widehat{L}\delta + \psi_\mu^{-1}(2\widehat{L}\delta) = \Psi_\mu(\widehat{L}\delta), \end{aligned}$$

where $\widehat{L} := \max\{\hat{a}, \frac{1}{2}(1 + L_\mu \hat{a} + L)\} \ge 1$. In case $d = 0$ we may choose $\hat{x} = \tilde{x}$, $\hat{a} = 1$, $L = 1$ 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 2.4 in case $d = 0$.

REMARK 2.5. Problem (1) is said to have kth order growth at the solution set for some $k \ge 1$ if $\psi_\mu(\tau) \ge \gamma \tau^k$ for each small $\tau \in \mathbb{R}_+$ and some $\gamma > 0$, i.e.,

$$\int_\Xi f_0(\xi, x)\,\mu(d\xi) - v(\mu) \ge \gamma d(x, S(\mu))^k \quad \text{for each feasible } x \text{ close to } S(\mu).$$

Then $\Psi_\mu(\eta) \le \eta + (2\eta/\gamma)^{1/k} \le C\eta^{1/k}$ for some constant $C > 0$ and sufficiently small $\eta \in \mathbb{R}_+$. In this case, Theorem 2.4 provides Hölder continuity of $S_{\mathcal{U}}$ at μ with rate $1/k$. Important particular cases are linear and quadratic growth for $k = 1$ and $k = 2$, respectively.

REMARK 2.6. Criteria for the metric regularity of multifunctions are given, e.g., in §9G of Rockafellar and Wets (1997) and in Mordukhovich (1994). Here, we do not intend to provide a specific sufficient condition for Assumption (iii) of Theorem 2.3, but note that the constraint functions $\int_\Xi f_j(\xi, \cdot)\,\mu(d\xi)$ $(j = 1, \ldots, d)$ are often nondifferentiable in stochastic programming and refer to the general results in Mordukhovich (1994) and to Henrion and Römisch (1999), where metric regularity in case of chance constrained stochastic programs is discussed.

REMARK 2.7. In Theorems 2.3 and 2.4, the localized optimal values $v_{\mathcal{U}}(\nu)$ and solution sets $S_{\mathcal{U}}(\nu)$ of the (perturbed) model (5) 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 (5) is convex and $S_{\mathcal{U}}(\nu)$ is a CLM set, (b) the constraint set of (5) is contained in some bounded set $\mathcal{V} \subset \mathbb{R}^m$ not depending on ν and it holds that $\mathcal{V} \subseteq \mathcal{U}$.

REMARK 2.8. Let $d = 0$, $S(\mu)$ be nonempty and assume that the objective satisfies a quadratic growth condition on $X \cap \mathcal{U}$, where $\mathcal{U}$ is a convex open neighbourhood of $S(\mu)$ (i.e., $\psi_\mu(\tau) \ge \gamma\tau^2$ for small $\tau \in \mathbb{R}_+$). Then the estimate,

$$\begin{aligned} (11) \qquad & \sup_{x \in S_{\mathcal{U}}(\nu)} d(x, S(\mu)) \\ &\quad \le \frac{1}{\gamma} \sup\left\{ \|x^*\| : x^* \in \partial\left(\int_\Xi f_0(\xi, \cdot)(\nu - \mu)\,(d\xi) \right)(x),\, x \in X \cap \operatorname{cl} \mathcal{U} \right\}, \end{aligned}$$

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., Mordukhovich 1994 and Rockafellar and Wets 1997).

Compared to the estimate in Theorem 2.4 based on function values, the above bound (11) uses subdifferentials. While Theorem 2.4 is valid in rather general situations, the assumptions implying (11) are more restrictive. However, (11) may lead to Lipschitz-type results in case of quadratic growth (see the conclusions for two-stage stochastic programs in Shapiro 1994, Römisch and Schultz 1996, and Dentcheva and Römisch 2000). For a proof of (11), the reader is referred to Bonnans and Shapiro (2000) and Shapiro (1994).

The two stability theorems illuminate the role of the distance $d_{\mathcal{F},\mathcal{U}}$ as a *minimal information* (m.i.) probability pseudometric for stability, i.e., as a pseudometric 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.

A distance d_c bounding $d_{\mathcal{F},\mathcal{U}}$ from above will be called an *ideal* or *canonical probability metric* associated with (1) if it has ζ-structure (3) generated by some class of functions $\mathcal{F} = \mathcal{F}_c$ from Ξ to $\overline{\mathbb{R}}$ such that $\mathcal{F}_c$ contains the functions $Cf_j(\cdot, x)$ for each $x \in X \cap \operatorname{cl} \mathcal{U}$, $j = 0, \ldots, d$, and some normalizing constant $C > 0$, and such that any function in $\mathcal{F}_c$ shares typical analytical properties with some function $f_j(\cdot, x)$.

In our applications, we clarify such typical analytical properties. Here, we only mention that typical integrands in stochastic programming are nondifferentiable, but piecewise Lipschitz continuous with discontinuities at boundaries of polyhedral sets.

To form an idea of how to associate a canonical metric, we consider the pth order Fortet-Mourier metric introduced in §1. Then the following result is an immediate consequence of the general theorems. It was already announced in §1.4 of Rachev and Rüschendorf (1998) in a slightly more general formulation.

COROLLARY 2.9. *Let $d = 0$ and 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 $\mathbb{R}^m$ for each $\xi \in \Xi$.*

(iii) *There exist constants $L > 0$, $p \geq 1$, such that $(1/L)f_0(\cdot, x) \in \mathcal{F}_p$ for each $x \in X \cap \operatorname{cl} \mathcal{U}$.*

Then there exists a constant $\delta > 0$ such that

$$|v(\mu) - v(\nu)| \leq L\zeta_p(\mu, \nu) \quad \text{and}$$

$$\varnothing \neq S(\nu) \subseteq S(\mu) + \Psi_\mu(L\zeta_p(\mu, \nu))\mathbb{B}$$

whenever $\nu \in \mathcal{P}_p(\Xi)$ and $\zeta_p(\mu, \nu) < \delta$. Here, the function Ψ_μ is given by (9).

PROOF. The assumptions of Theorem 2.3 are satisfied. Hence, the result is a consequence of the Theorems 2.3 and 2.4 and the fact that (iii) is equivalent to

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

for each $\xi, \tilde{\xi} \in \Xi$ and $x \in X \cap \operatorname{cl} \mathcal{U}$, and hence, it implies $d_{\mathcal{F},\mathcal{U}}(\mu, \nu) \leq L\zeta_p(\mu, \nu)$ for all $\mu, \nu \in \mathcal{P}_p(\Xi)$. Furthermore, the localized optimal values $v_{\mathcal{U}}$ and solution sets $S_{\mathcal{U}}$ may be replaced by v and S, respectively, because of the convexity assumption (ii) if ν is close to μ (see Remark 2.7). □

EXAMPLE 2.10 (EXAMPLE 1.1 CONTINUED). In this case, the set $\mathcal{F}_{\mathcal{U}}$ is a specific class of piecewise linear functions and has the form $\{(r - c)x + c \max\{0, x - \cdot\}: x \in X \cap \operatorname{cl} \mathcal{U}\}$. Furthermore, $\int_\Xi f_0(\xi, x)\, \mu(d\xi)$ is also piecewise linear and Corollary 2.9 applies with $L := c$, $p := 1$ and a linear function Ψ_μ. Hence, the solution set S behaves upper Lipschitzian at μ with respect to ζ_1.

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

$$\min\left\{cx + \int_\Xi q(\xi)y(\xi)\, \mu(d\xi) \colon Wy(\xi) = h(\xi) - T(\xi)x,\ y(\xi) \geq 0,\ x \in X\right\}, \tag{12}$$

where $c \in \mathbb{R}^m$, $X \subseteq \mathbb{R}^m$ is a polyhedron, Ξ is a polyhedron in $\mathbb{R}^s$, W is an $(r, \bar{m})$-matrix, $\mu \in \mathcal{P}(\Xi)$, and the vectors $q(\xi) \in \mathbb{R}^{\bar{m}}$, $h(\xi) \in \mathbb{R}^r$ and the (r, m)-matrix $T(\xi)$ depend affine linearly on $\xi \in \Xi$.

In (12), x is the here-and-now or first-stage decision and $y(\cdot)$ 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 (12) may be rewritten equivalently as a minimization problem with respect to the first-stage decision x. Defining the integrand $f_0\colon \Xi \times \mathbb{R}^m \to \overline{\mathbb{R}}$ by

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

where $\text{pos} W = \{Wy\colon\ y \in \mathbb{R}^{\bar{m}}_+\}$, $D = \{u \in \mathbb{R}^{\bar{m}}\colon\ \{z \in \mathbb{R}^r\colon\ W'z \le u\} \neq \varnothing\}$ and $\Phi(u, t) = \inf\{uy\colon\ Wy = t, y \ge 0\}$ $((u, t) \in \mathbb{R}^{\bar{m}} \times \mathbb{R}^r)$, the equivalent minimization problem takes the form

$$\min\left\{\int_\Xi f_0(\xi, x)\, \mu(d\xi)\colon\ x \in X\right\}. \tag{13}$$

In order to utilize the general stability results of §2, we first recall some well-known properties of the function Φ, which are derived in Walkup and Wets (1969) (see also Nožička et al. 1974).

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

$$\bigcup_{j=1}^N \mathcal{K}_j = posW \times D, \quad \text{int } \mathcal{K}_i \cap \text{int } \mathcal{K}_j = \varnothing, \qquad i \neq j,$$

$$\Phi(u, t) = C_j u \cdot t, \qquad \textit{for each } (u, t) \in \mathcal{K}_j, j = 1, \ldots, N.$$

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

In order to have Problem (13) well defined, we introduce the following assumptions:

(A1) *For each pair $(\xi, x) \in \Xi \times X$, it holds that $h(\xi) - T(\xi)x \in \text{pos}W$ and $q(\xi) \in D$.*

(A2) $\mu \in \mathcal{P}_2(\Xi)$, *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 $\Xi \times X \subseteq \text{dom} f_0$.

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

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

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

$$|f_0(\xi, x)| \le Kr \max\{1, \|\xi\|^2\}.$$

PROOF. From Lemma 3.1 and (A1), we conclude that f_0 is continuous on dom f_0, and hence, on $\Xi \times \mathbb{R}^m$. This implies that f_0 is a normal integrand (cf., Rockafellar and Wets 1997, Example 14.31). It is also a convex integrand since the properties of Φ in Lemma 3.1

imply that $f_0(\xi,\cdot)$ is convex for each $\xi\in\Xi$. In order to verify the Lipschitz property of f_0, let $x\in X$ with $\|x\|\le r$ and consider for each $j=1,\dots,N$ 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 $\Xi_j := \{\xi\in\Xi\colon (q(\xi), h(\xi)-T(\xi)x)\in\mathcal{K}_j\}$ 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:

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

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

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

where we have used for the last two estimates that $\|\xi(\lambda)\|\le\max\{\|\xi\|,\|\tilde\xi\|\}$ 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 Kall (1976) and in Theorem 7.7 of Wets (1974). The second estimate of the proposition, in particular, is a consequence of those results. Furthermore, from Lemma 3.1, we conclude the estimate

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

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

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

$$\int_\Xi \inf_{\substack{x\in X\\ \|x\|\le r}} f_0(\xi,x)\,\nu(d\xi) \ge -Kr\left(1+\int_\Xi\|\xi\|^2\,\nu(d\xi)\right) > -\infty,$$

$$\sup_{x\in X\cap\mathcal{U}}\left|\int_\Xi f_0(\xi,x)\,\nu(d\xi)\right| \le K\rho\left(1+\int_\Xi\|\xi\|^2\,\nu(d\xi)\right) < \infty,$$

if $\nu \in \mathcal{P}(\Xi)$ has a finite, second-order moment. Hence, for any nonempty bounded $\mathcal{U} \subseteq \mathbb{R}^m$, 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.,

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

The following stability result for optimal values and solution sets of the two-stage problem (13) is now a direct consequence of Corollary 2.9 and Proposition 3.2.

THEOREM 3.3. *Let* (A1) *and* (A2) *be satisfied and let* $S(\mu)$ *be nonempty and let* $\mathcal{U}$ *be an open, bounded neighbourhood of* $S(\mu)$.

Then there exist constants $L > 0$ *and* $\delta > 0$ *such that*

$$|v(\mu) - v(\nu)| \le L\zeta_2(\mu, \nu) \quad \textit{and}$$
$$\varnothing \neq S(\nu) \subseteq S(\mu) + \Psi_\mu(L\zeta_2(\mu, \nu))\mathbb{B}$$

whenever $\nu \in \mathcal{P}_2(\Xi)$ *and* $\zeta_2(\mu, \nu) < \delta$, *where* Ψ_μ *is given by* (9).

PROOF. The result is a consequence of Corollary 2.9 with $p = 2$. The assumptions (ii) and (iii) of Corollary 2.9 are verified in Proposition 3.2. $\square$

The theorem establishes quantitative stability of $v(\cdot)$ and $S(\cdot)$ with respect to ζ_2 in the two-stage case for fairly general situations. Hence, ζ_2 is the *canonical metric* for two-stage models with fixed recourse.

In Römisch and Schultz (1991), stability of two-stage models with complete recourse (i.e., $\operatorname{pos} W = \mathbb{R}^r$) was studied with respect to the L_2-minimal or L_2-Wasserstein metric W_2, defined by

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

for all μ, $\nu \in \mathcal{P}_2(\Xi)$. Because of the estimate,

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

which is essentially a consequence of Schwarz's inequality, and the fact that convergence with respect to W_2 implies convergence of second-order moments, ζ_2 may be bounded above by W_2. Hence, Theorem 3.3 remains valid if ζ_2 is replaced by W_2 (with possibly different constants) and, thus, Theorem 3.3 extends Theorem 2.4 in Römisch and Schultz (1991). This extension is strict, since the following example shows that ζ_2 and W_2 may have different asymptotic properties.

EXAMPLE 3.4. Let $\Xi := \mathbb{R}_+$, $\mu := \delta_0$ and $\mu_n := (1 - 1/n)\delta_0 + (1/n)\delta_{\xi_n}$ for each $n \in \mathbb{N}$, where (ξ_n) is an unbounded nondecreasing sequence in $\mathbb{R}_+$ such that $((1/\sqrt{n})\xi_n)$ 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 $n \in \mathbb{N}$ with $\xi_n \ge 1$:

$$\zeta_2(\mu, \mu_n) = \int_0^\infty \max\{1, \xi\} |F_\mu(\xi) - F_{\mu_n}(\xi)| \, d\xi = \frac{1}{n} + \int_1^{\xi_n} \frac{\xi}{n} \, d\xi = \frac{1}{2n}(\xi_n^2 + 1),$$
$$W_2(\mu, \mu_n) = \left(\int_0^1 |F_\mu^{-1}(t) - F_{\mu_n}^{-1}(t)|^2 dt \right)^{1/2} = \left(\int_{1-\frac{1}{n}}^1 \xi_n^2 \, dt \right)^{1/2} = \frac{1}{\sqrt{n}} \xi_n.$$

Here we have used the explicit representations of ζ_2 and W_2 in case of probability measures on (subsets of) $\mathbb{R}$ (see Rachev 1991, Chapters 5.4 and 13.1), 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 $F^{-1}(t) := \sup\{\xi \in \mathbb{R}: F(\xi) \le t\}$ ($t \in [0,1]$).

Concluding this section, we mention that in case the recourse costs $q(\cdot)$ are nonstochastic or the technology matrix $T(\cdot)$ and the right-hand side $h(\cdot)$ are nonstochastic, 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, Lipschitz stability results can be obtained in case of quadratic growth at $S(\mu)$ by exploiting the estimate in Remark 2.8 (see Shapiro 1994, Römisch and Schultz 1996, and Dentcheva and Römisch 2000). On the other hand, strong convexity properties of the objective function (cf., Schultz 1994) provide sufficient conditions for quadratic growth.

3.2. Mixed-integer two-stage models. Next, we allow for mixed-integer decisions in both stages and consider the program,

$$\min\left\{cx + \int_\Xi \Phi(h(\xi) - T(\xi)x)\,\mu(d\xi) \colon x \in X\right\}, \tag{14}$$

where

$$\Phi(t) := \min\left\{qy + \bar{q}\bar{y} \colon Wy + \overline{W}\bar{y} = t,\ y \in \mathbb{Z}_+^{\hat{m}},\ \bar{y} \in \mathbb{R}_+^{\bar{m}}\right\} \qquad (t \in \mathbb{R}^r), \tag{15}$$

$c \in \mathbb{R}^m$, X is a closed subset of $\mathbb{R}^m$, Ξ a polyhedron in $\mathbb{R}^s$, $q \in \mathbb{R}^{\hat{m}}$, $\bar{q} \in \mathbb{R}^{\bar{m}}$, W, and $\overline{W}$ are $(r, \hat{m})$- and $(r, \bar{m})$-matrices, respectively, $h(\xi) \in \mathbb{R}^r$ and the (r, m)-matrix $T(\xi)$ are affine linear functions of $\xi \in \mathbb{R}^s$, 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 (14) well defined:

(B1) The matrices W and $\overline{W}$ 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

$$\mathcal{T} := \left\{t \in \mathbb{R}^r \colon t = Wy + \overline{W}\bar{y},\ y \in \mathbb{Z}_+^{\hat{m}},\ \bar{y} \in \mathbb{R}_+^{\bar{m}}\right\}.$$

(B3) There exists an element $u \in \mathbb{R}^r$ such that $W'u \le q$ and $\overline{W}'u \le \bar{q}$.

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 (15) 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., $\mathcal{T} = \bigcup_{i\in\mathbb{N}} \mathcal{B}_i$ such that*

(i) *Each of the sets has a representation $\mathcal{B}_i = \{b_i + pos\overline{W}\} \setminus \bigcup_{j=1}^{N_0}\{b_{ij} + pos\overline{W}\}$, where $b_i, b_{ij} \in \mathbb{R}^r$ for $i \in \mathbb{N}$ and $j = 1, \ldots, N_0$. Moreover, there exists an $N_1 \in \mathbb{N}$, such that for any $t \in \mathcal{T}$ the ball $\mathbb{B}(t, 1)$ in $\mathbb{R}^r$ 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, $\tilde{t} \in \mathcal{T}$:

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

Part (i) of the lemma is proved in §5.6 of Bank et al. (1982) and in Lemma 2.5 of Schultz (1996), (ii) is derived as Lemma 2.3 in Schultz (1996) and the remaining properties of Φ are established in Blair and Jeroslow (1977). 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 to take into account the tail behaviour of μ.

In order to state the stability results for Model (14), we consider the following probability metric with ζ-structure on $\mathcal{P}_1(\Xi)$ for any $k \in \mathbb{N}$:

$$\begin{aligned}(16)\qquad \zeta_{1,ph_k}(\mu,\nu) &:= \sup\left\{\left|\int_P f(\xi)(\mu-\nu)(d\xi)\right| : f \in \mathcal{F}_1(\Xi), P \in \mathcal{B}_{ph_k}(\Xi)\right\}\\ &= \sup\left\{\left|\int_\Xi f(\xi)\chi_P(\xi)(\mu-\nu)(d\xi)\right| : f \in \mathcal{F}_1(\Xi), P \in \mathcal{B}_{ph_k}(\Xi)\right\}.\end{aligned}$$

Here, $\mathcal{B}_{ph_k}(\Xi)$ denotes the set of all polyhedra which are contained in Ξ and have at most k faces, χ_P the characteristic function of P and $\mathcal{F}_1(\Xi)$ is defined in §1.

THEOREM 3.6. *Let the conditions* (B1)–(B3) *be satisfied, let* $\mu \in \mathcal{P}_p(\Xi)$ *for some* $p > 1$, $S(\mu)$ *be nonempty, and* $\mathcal{U} \subseteq \mathbb{R}^m$ *be an open bounded neighbourhood of* $S(\mu)$.

Then there exist constants $L > 0$, $\delta > 0$, *and* $k \in \mathbb{N}$ *such that*

$$(17)\qquad \begin{aligned}|v(\mu) - v_{\mathcal{U}}(\nu)| &\le L\zeta_{1,ph_k}(\mu,\nu)^{(1/(1+(r/(p-1))))}\\ \varnothing \ne S_{\mathcal{U}}(\nu) &\subseteq S(\mu) + \Psi_\mu(L\zeta_{1,ph_k}(\mu,\nu)^{(1/(1+(r/(p-1))))})\mathbb{B}\end{aligned}$$

and $S_{\mathcal{U}}(\nu)$ *is a CLM set for* (14) *w.r.t.* $\mathcal{U}$ *whenever* $\nu \in \mathcal{P}_1(\Xi)$ *and* $\zeta_{1,ph_k}(\mu,\nu) < \delta$. *Here, the function* Ψ_μ *is given by* (9).

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 (see Rockafellar and Wets 1997, Example 14.31). Using Lemma 3.5 we obtain the estimate

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

for each pair $(\xi,x) \in \Xi \times X$. Since $h(\xi)$ and $T(\xi)$ depend affine linearly on ξ, there exists a constant $C_1 > 0$ such that $|f_0(\xi,x)| \le C_1 \max\{1, \|\xi\|\}$ holds for each pair $(\xi,x) \in \Xi \times (X \cap \operatorname{cl}\mathcal{U})$. Hence, $\mathcal{P}_{\mathcal{F},\mathcal{U}}(\Xi) \supseteq \mathcal{P}_1(\Xi)$ and Theorem 2.3 applies with $d = 0$ and the distance $d_{\mathcal{F},\mathcal{U}}$ on $\mathcal{P}_1(\Xi)$. It remains to show that the estimate

$$(18)\qquad d_{\mathcal{F},\mathcal{U}}(\mu,\nu) = \sup_{x \in X \cap \operatorname{cl}\mathcal{U}} \left|\int_\Xi f_0(\xi,x)(\mu-\nu)(d\xi)\right| \le C\zeta_{1,ph_k}(\mu,\nu)^{(1/(1+(r/(p-1))))}$$

is valid for some constant $C > 0$ and sufficiently small $\zeta_{1,ph_k}(\mu,\nu)$. 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 Schultz (1996) 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., Schultz 1996, page 1143). 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 $x \in X \cap \mathrm{cl}\,\mathcal{U}$ and consider the following disjoint Borel subsets of Ξ:

$$\Xi_{j,x}^R := \{\xi \in \Xi\colon\ h(\xi) - T(\xi)x \in B_j\} \qquad (j = 1, \dots, N),$$
$$\Xi_{0,x}^R := \Xi \setminus \bigcup_{j=1}^N \Xi_{j,x}^R \subseteq \{\xi \in \Xi\colon\ \|h(\xi) - T(\xi)x\|_\infty > R\}.$$

From Lemma 3.5, we conclude that there exists a constant $L_1 > 0$ (which does not depend on $x \in X \cap \mathrm{cl}\,\mathcal{U}$, $j = 1, \dots, N$ and $R > 0$) such that each function $f_{j,x}^R(\cdot) := cx + \Phi|_{B_j}(h(\cdot) - T(\cdot)x)$ is Lipschitz continuous on $\Xi_{j,x}^R$ with constant L_1. We extend each function $f_{j,x}^R(\cdot)$ to the whole of Ξ by preserving the Lipschitz constant L_1.

For each $\nu \in \mathcal{P}_1(\Xi)$, we may continue:

$$\begin{aligned}(19)\qquad \left|\int_\Xi f_0(\xi, x)(\mu - \nu)\,(d\xi)\right| &= \left|\sum_{j=0}^N \int_{\Xi_{j,x}^R} f_0(\xi, x)(\mu - \nu)\,(d\xi)\right| \\ &\le \sum_{j=1}^N \left|\int_{\Xi_{j,x}^R} f_{j,x}^R(\xi)(\mu - \nu)\,(d\xi)\right| + I_x^R(\mu, \nu) \\ &\le NL_1 \sup_{f \in \mathcal{F}_1(\Xi),\, j=1,\dots,N} \left|\int_\Xi f(\xi)\chi_{\Xi_{j,x}^R}(\mu - \nu)\,(d\xi)\right| \\ &\quad + I_x^R(\mu, \nu),\end{aligned}$$

where $I_x^R(\mu, \nu) := |\int_{\Xi_{0,x}^R} f_0(\xi, x)(\mu - \nu)\,(d\xi)|$.

For each $j = 1, \dots, N$ the closures of the sets B_j are polyhedra with a number of faces which is bounded above by some number not depending on j, N, and R. Hence, the same is true for the closures of the sets $\Xi_{j,x}^R$, i.e., for $\mathrm{cl}\,\Xi_{j,x}^R = \{\xi \in \Xi\colon\ h(\xi) - T(\xi)x \in \mathrm{cl}\,B_j\}$, where the corresponding number $k \in \mathbb{N}$ does not, in addition, depend on $x \in X \cap \mathrm{cl}\,\mathcal{U}$. For each such $\Xi_{j,x}^R$, we now consider a sequence of closed polyhedra $P_{j,x}^R$, which are contained in $\Xi_{j,x}^R$ and have at most k faces, such that their characteristic functions $\chi_{P_{j,x}^R}$ converge pointwise to the characteristic function $\chi_{\Xi_{j,x}^R}$. Then the sequence consisting of the elements $|\int_\Xi f(\xi)\chi_{P_{j,x}^R}(\xi)(\mu - \nu)\,(d\xi)|$ converges to $|\int_\Xi f(\xi)\chi_{\Xi_{j,x}^R}(\xi)(\mu - \nu)\,(d\xi)|$ while each element is bounded by $\zeta_{1,ph_k}(\mu, \nu)$. Hence, the estimate (19) may be continued to

$$(20)\qquad \left|\int_\Xi f_0(\xi, x)(\mu - \nu)\,(d\xi)\right| \le \kappa L_1 R^r \zeta_{1,ph_k}(\mu, \nu) + I_x^R(\mu, \nu).$$

Since there exists a constant $C_2 > 0$ such that $\|h(\xi) - T(\xi)x\|_\infty \le C_2 \max\{1, \|\xi\|_\infty\}$ holds for each pair $(\xi, x) \in \Xi \times (X \cap \mathrm{cl}\,\mathcal{U})$, the following upper bound for $I_x^R(\mu, \nu)$ holds for sufficiently large $R > 0$:

$$I_x^R(\mu, \nu) \le C_1 \int_{\{\xi \in \Xi: \|\xi\|_\infty \ge R/C_2\}} \|\xi\|(\mu + \nu)\,(d\xi).$$

Clearly, the set $\{\xi \in \Xi\colon\ \|\xi\|_\infty \ge R/C_2\}$ is contained in the union of a finite number, say M, of Borel sets whose closures are polyhedra. Using the same arguments as for deriving the estimate (20), we obtain for sufficiently large k,

$$\int_{\{\xi \in \Xi: \|\xi\|_\infty \ge R/C_2\}} \|\xi\|\nu\,(d\xi) \le M\zeta_{1,ph_k}(\mu, \nu) + \int_{\{\xi \in \Xi: \|\xi\|_\infty \ge R/C_2\}} \|\xi\|\mu\,(d\xi).$$

Hence, we get from (20) and the previous estimate that

$$d_{\mathcal{F},\mathcal{U}}(\mu,\nu) \le \kappa(L_1 R^r + MC_1)\zeta_{1,ph_k}(\mu,\nu) + 2C_1 \int_{\{\xi\in\Xi:\|\xi\|_\infty \ge R/C_2\}} \|\xi\|\,\mu(d\xi) \tag{21}$$

$$\le \kappa(L_1 R^r + MC_1)\zeta_{1,ph_k}(\mu,\nu) + CR^{1-p} \int_{\Xi} \|\xi\|^p\,\mu(d\xi) \tag{22}$$

for some constant $C > 0$. Inserting $R := \zeta_{1,ph_k}(\mu,\nu)^{-1/(p-1+r)}$ for sufficiently small $\zeta_{1,ph_k}(\mu,\nu)$ into (21), (22) implies the desired estimate (18). □

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 improves slightly.

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*

$$|v(\mu) - v_{\mathcal{U}}(\nu)| \le L\zeta_{1,ph_k}(\mu,\nu),$$
$$\varnothing \ne S_{\mathcal{U}}(\nu) \subseteq S(\mu) + \Psi_\mu(L\zeta_{1,ph,k}(\mu,\nu))\mathbb{B},$$

and $S_{\mathcal{U}}(\nu)$ *is a CLM set of* (14) *w.r.t.* $\mathcal{U}$ *whenever* $\nu \in \mathcal{P}(\Xi)$ *and* $\zeta_{1,ph_k}(\mu,\nu) < \delta$.

PROOF. Since Ξ is bounded, it holds that $\mathcal{P}_1(\Xi) = \mathcal{P}(\Xi)$. Moreover, the term $I_x^R(\mu,\nu)$ in the previous proof vanishes for each $x \in X \cap \mathrm{cl}\,\mathcal{U}$, $\nu \in \mathcal{P}(\Xi)$ and sufficiently large $R > 0$. Hence, (20) and Theorem 3.6 imply the assertion. □

Hence, the probability metric ζ_{1,ph_k} is a *canonical metric* for (general) linear mixed-integer two-stage stochastic programs.

REMARK 3.8. Since $\Xi \in \mathcal{B}_{ph_k}(\Xi)$ for sufficiently large $k \in \mathbb{N}$, we obtain from (16) by choosing $P := \Xi$ and $f \equiv 1$, respectively,

$$\max\{\zeta_1(\mu,\nu), \alpha_{ph_k}(\mu,\nu)\} \le \zeta_{1,ph_k}(\mu,\nu) \tag{23}$$

for large k and all $\mu,\nu \in \mathcal{P}_1(\Xi)$. Here, α_{ph_k} denotes the polyhedral discrepancy,

$$\alpha_{ph_k}(\mu,\nu) := \sup\{|\mu(P) - \nu(P)| : P \in \mathcal{B}_{ph_k}(\Xi)\}. \tag{24}$$

Hence, convergence with respect to ζ_{1,ph_k} implies weak convergence, convergence of first-order absolute moments, and convergence with respect to the polyhedral discrepancy α_{ph_k}. Using the technology of the proof of Proposition 3.1 in Schultz (1996), it can be shown that

$$\zeta_{1,ph_k}(\mu,\nu) \le \alpha_{ph_k}(\mu,\nu)^{1/(s+1)} \tag{25}$$

holds for all $\mu,\nu \in \mathcal{P}(\Xi)$ if $\Xi \subset \mathbb{R}^s$ is bounded. This illuminates the relation between the stability results stated in this section and those in Schultz (1996) for $T(\cdot) \equiv T$, $\Xi = \mathbb{R}^s$ and in terms of some polyhedral discrepancy with exponents bounded by $1/(s+1)$. In view of (23) and (25), the metric ζ_{1,ph_k} is stronger than α_{ph_k} in general, but in case of bounded Ξ they metrize the same topology on $\mathcal{P}(\Xi)$. However, our analysis may lead to improved rates of convergence (e.g., for empirical approximations in Example 4.5).

3.3. Linear chance constrained models. In this section, we study consequences of the general stability analysis of §2 to linear chance-constrained stochastic programs of the form

$$\min\{cx\colon x \in X,\ \mu(\{\xi \in \Xi : T(\xi)x \ge h(\xi)\}) \ge p\}, \tag{26}$$

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 $H(x) = \{\xi \in \Xi : T(\xi)x \geq h(\xi)\}$, we observe that the program (26) is a particular case of the general stochastic program (1). In order to see that the general assumptions of §2 are 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 (Rockafellar and Wets 1997, Example 14.31). Moreover, $p - 1 \leq f_1(x, \xi) \leq p$ holds and, for any nonempty open and 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 §2:

$$\mathcal{P}_{\mathcal{F},\mathcal{U}}(\Xi) = \left\{ \nu \in \mathcal{P}(\Xi) \colon \sup_{x \in X \cap \mathrm{cl}\, \mathcal{U}} \max_{j=0,1} \left| \int_{\Xi} f_j(\xi, x)\, \nu(d\xi) \right| < \infty \right\} = \mathcal{P}(\Xi),$$

$$\begin{aligned} d_{\mathcal{F},\mathcal{U}}(\mu, \nu) &= \sup_{x \in X \cap \mathrm{cl}\, \mathcal{U}} \max_{j=0,1} \left| \int_{\Xi} f_j(\xi, x)(\mu - \nu)(d\xi) \right| \\ &= \sup_{x \in X \cap \mathrm{cl}\, \mathcal{U}} |\mu(H(x)) - \nu(H(x))| \qquad (\mu,\ \nu \in \mathcal{P}(\Xi)). \end{aligned}$$

Such pseudometrics were already used in the stability analysis of Römisch and Schultz (1991a, b). Since the sets $H(x)$ are polyhedra with a uniformly bounded number of faces, the polyhedral discrepancy α_{ph_k} on $\mathcal{P}(\Xi)$ (see (24)) for some $k \in \mathbb{N}$ is a natural candidate for a *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 $\mathcal{U} \subseteq \mathbb{R}^m$ is an open-bounded neighbourhood of $S(\mu)$.*

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

Then there exist constants $L > 0$, $\delta > 0$ and $k \in \mathbb{N}$ such that

$$|v(\mu) - v_{\mathcal{U}}(\nu)| \leq L\alpha_{ph_k}(\mu, \nu),$$

$$\varnothing \neq S_{\mathcal{U}}(\nu) \subseteq S(\mu) + \Psi_\mu(L\alpha_{ph_k}(\mu, \nu))\mathbb{B}$$

and $S_{\mathcal{U}}(\nu)$ is a CLM set for (26) *w.r.t. $\mathcal{U}$ whenever $\nu \in \mathcal{P}(\Xi)$ and $\alpha_{ph_k}(\mu, \nu) < \delta$. Here, the function Ψ_μ is given by* (9).

PROOF. Clearly, all assumptions of Theorem 2.3 are satisfied for the special situation considered in this section. Hence, the result follows from the Theorems 2.3 and 2.4 by taking into account the estimate $d_{\mathcal{F},\mathcal{U}}(\mu, \nu) \leq \alpha_{ph_k}(\mu, \nu)$. □

REMARK 3.10. Let the function $g(x) := \mu(\{\xi \in \Xi \colon T(\xi)x \geq h(\xi)\})$ be locally Lipschitz continuous on X. Then Condition (ii) of Proposition 3.9 is satisfied if the constraint qualification $\partial(-g)(\bar{x}) \cap (-N_X(\bar{x})) = \varnothing$ 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., Mordukhovich 1994). Similar constraint qualifications were used in Römisch and Schultz (1991b) for a specific model of the form (26) (with $r = 1$, $s = m + 1$ and a multivariate normal distribution μ) and in Henrion and Römisch (1999) for models with nonstochastic technology matrix $T(\cdot)$. Henrion and Römisch (1999) also provide conditions implying quadratic growth of the function ψ_μ (see (8)).

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 $(\Omega, \mathcal{A}, \mathbb{P})$ having the joint distribution μ, i.e., $\mu = \mathbb{P}\xi_1^{-1}$. We consider the empirical measures

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

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

$$\text{(27)} \qquad \min\left\{\frac{1}{n}\sum_{i=1}^{n} f_0(\xi_i(\cdot), x)\colon x \in X,\ \frac{1}{n}\sum_{i=1}^{n} f_j(\xi_i(\cdot), x) \le 0,\ j = 1, \ldots, d\right\}.$$

Since the objective and constraint functions of (27) are normal integrands from $\Omega \times \mathbb{R}^m$ to $\overline{\mathbb{R}}$, the constraint set is closed valued and measurable from Ω to $\mathbb{R}^m$, and hence, the optimal value $v(\mu_n(\cdot))$ of (27) is measurable from Ω to $\overline{\mathbb{R}}$ and the solution set $S(\mu_n(\cdot))$ is a closed-valued measurable multifunction from Ω to $\mathbb{R}^m$ (see Rockafellar and Wets 1997, Chapter 14, and in particular, Theorem 14.37). 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 $\mathbb{R}^m$.

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

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

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 van der Vaart and Wellner (1996), concepts were described that allow for overcoming the possible nonmeasurability of the supremum. To simplify matters here, we call a class $\mathcal{F}$ of measurable functions from Ξ to $\overline{\mathbb{R}}$ *permissible* for $\mu \in \mathcal{P}(\Xi)$ 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 (cf., van der Vaart and Wellner 1996, Example 2.3.4). 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 $n \in \mathbb{N}$ 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 $\overline{\mathbb{R}}$. Instances of permissible classes $\mathcal{F}$ are given in Examples 4.3–4.5.

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 $\mathbb{P}$-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 Dudley 1984 and van der Vaart and Wellner 1996). To introduce these concepts, let $\mathcal{F}$ be a subset of the normed space $L_p(\Xi, \mu)$ (for some $p \ge 1$) 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 $\{g \in L_p(\Xi, \mu)\colon \|g - f\|_p < \varepsilon\}$ needed to cover $\mathcal{F}$. Given two functions f_1 and f_2 from $L_p(\Xi, \mu)$, the set $[f_1, f_2] := \{f \in L_p(\Xi, \mu)\colon f_1(\xi) \le f(\xi) \le f_2(\xi)$ 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 $N(\varepsilon, \mathcal{F}, L_p(\Xi, \mu)) \le N_{[\,]}(2\varepsilon, \mathcal{F}, L_p(\Xi, \mu))$, 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 Dudley 1984, Theorem 6.1.5). We also refer to Talagrand (1996) for further criteria of Glivenko-Cantelli classes and to Pflug et al. (1998) for applications to stochastic programming.

To state our next results, we denote the set of all real-valued random variables on $(\Omega, \mathcal{A}, \mathbb{P})$ by $\mathcal{X}(\mathbb{R})$, where equality is understood as equality $\mathbb{P}$-almost surely, and introduce the *Ky Fan distance* κ of two real random variables $\mathcal{X}, \mathcal{Y} \in \mathcal{X}(\mathbb{R})$ by

$$\kappa(\mathcal{X}, \mathcal{Y}) := \inf\{\eta \geq 0\colon \mathbb{P}(|\mathcal{X} - \mathcal{Y}| > \eta) \leq \eta\}. \tag{28}$$

It is known that the infimum in (28) is attained and that κ metrizes convergence in probability in $\mathcal{X}(\mathbb{R})$ (see, e.g., Dudley 1989, §9.2). By means of the Ky Fan metric, the quantitative stability results of §2 directly translate into estimates for the empirical optimal values and solution sets.

PROPOSITION 4.1. *Assume that the Conditions* (i)–(iii) *of Theorem* 2.3 *are satisfied, that $\mathcal{F}_{\mathcal{U}}$ is permissible for μ and a μ-Glivenko-Cantelli class. Then it holds for sufficiently large $n \in \mathbb{N}$ that*

$$\kappa(v(\mu), v_{\mathcal{U}}(\mu_n(\cdot))) \leq \max\{1, L\}\kappa(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu), 0),$$
$$\kappa\left(\sup_{x \in S_{\mathcal{U}}(\mu_n(\cdot))} d(x, S(\mu)), 0\right) \leq \Psi_\mu(\kappa(d_{\mathcal{F},\mathcal{U}}(\mu_n(\cdot), \mu), 0)),$$

where $L > 0$ is the constant in Theorem 2.3 *and Ψ_μ the associated function* (9).

Moreover, for $\mathbb{P}$-almost all $\omega \in \Omega$ the set $S_{\mathcal{U}}(\mu_n(\omega))$ is a CLM set of (27) *with respect to $\mathcal{U}$ for sufficiently large $n \in \mathbb{N}$.*

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 2.3. Then Theorem 2.3 implies

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

for sufficiently large $n \in \mathbb{N}$, since $\mathcal{F}_{\mathcal{U}}$ is a Glivenko-Cantelli class, and thus, the sequence (ε_n) tends to 0. Hence, we obtain from (28) 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_μ be the function (9). Then we conclude from Theorem 2.4 that

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

for sufficiently large $n \in \mathbb{N}$, since it holds that $\Psi_\mu(\varepsilon_n) \geq \varepsilon_n$ (see Theorem 2.3) and since $(\Psi_\mu(\varepsilon_n))$ 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 because of Proposition 2.1. Since $\mathcal{F}_{\mathcal{U}}$ is a μ-Glivenko-Cantelli class, there exists a set $A \in \mathcal{A}$ with $\mathbb{P}(A) = 0$ 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 $\omega \in \Omega \backslash A$ and for sufficiently large $n \in \mathbb{N}$. This completes the proof. $\square$

Note that in case of a fixed constraint set (i.e., $d = 0$) both estimates in Proposition 4.1 are valid for each $n \in \mathbb{N}$ (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* 2.3 *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) $N(\varepsilon, \mathcal{F}_{\mathcal{U}}, L_2(\Xi, \nu)) \leq (R/\varepsilon)^r$ *for any discrete* $\nu \in \mathcal{P}(\Xi)$ *with finite support.*

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

Then, with Ψ_μ *given by* (9), *the following rates of convergence are valid*:

$$\kappa(v(\mu), v_{\mathcal{U}}(\mu_n(\cdot))) = O((\log n)^{1/2} n^{-1/2})$$

$$\kappa\left(\sup_{x \in S_{\mathcal{U}}(\mu_n(\cdot))} d(x, S(\mu)), 0\right) = O(\Psi_\mu((\log n)^{1/2} n^{-1/2})).$$

PROOF. In both cases, (i) and (ii), we obtain from Theorem 1.3 in Talagrand (1994) that

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

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

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

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

Both estimates and convergence rates in Propositions 4.1 and 4.2 could be formulated alternatively in terms of certain confidence bounds as in Pflug (1999) and in §5 of Henrion and Römisch (1999). However, such bounds typically contain unknown constants (like the constant δ appearing in Theorem 2.3 or the entropy constant R). Finally, we discuss applications of the results to linear two-stage and chance-constrained stochastic programming models, thereby extending earlier work in Pflug (1999, Example 4.4) and Schultz (1996, Example 4.5).

EXAMPLE 4.3. (LINEAR CHANCE-CONSTRAINED 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 $\{\xi_1, \ldots, \xi_l\}$ 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 $N(\varepsilon, \{\chi_B\colon B \in \mathcal{B}\}, L_2(\Xi, \nu)) \leq K\varepsilon^{-2r}$ for each $\varepsilon > 0$ and $\nu \in \mathcal{P}(\Xi)$, and some constant $K > 0$ depending only on r (van der Vaart and Wellner 1996, Theorem 2.6.4).

For any polyhedral set, $\Xi \subseteq \mathbb{R}^s$ and $k \in \mathbb{N}$, the class $\mathcal{B}_{ph_k}(\Xi)$ 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, since the set of all polyhedra in $\mathcal{B}_{ph_k}(\Xi)$ having vertices at rational points in $\mathbb{R}^s$ plays the role of the countable subset in the definition of permissibility. Since Condition (i) in Proposition 4.2 is satisfied, it follows as in the proof of 4.2 from Talagrand (1994) that

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

holds for each $k \in \mathbb{N}$. This rate of convergence applies directly to empirical approximations of the linear chance-constrained model in §3.3 (see Proposition 3.9).

EXAMPLE 4.4. (TWO-STAGE MODELS WITHOUT INTEGRALITY). Let f_0 be defined as in §3.1 and let (A1) and (A2) be satisfied. Then, for each nonempty open and bounded subset $\mathcal{U}$ of $\mathbb{R}^m$, the class $\mathcal{F}_{\mathcal{U}} = \{f_0(\cdot, x)\colon x \in X \cap \operatorname{cl}\mathcal{U}\}$ is a subset of $L_1(\Xi, \mu)$. $\mathcal{F}_{\mathcal{U}}$ is also permissible for μ, since any class $\{f_0(\cdot, x)\colon x \in X_c\}$ with X_c being a countable and dense subset of $X \cap \operatorname{cl}\mathcal{U}$ may be used as the countable subset of $\mathcal{F}_{\mathcal{U}}$ in the definition of permissibility. Because of the Lipschitz continuity property of $f_0(\xi, \cdot)$ with Lipschitz constant $\widehat{L}\max\{1, \|\xi\|^2\}$ (see Proposition 3.2), the bracketing numbers of $\mathcal{F}_{\mathcal{U}}$ are bounded by the covering numbers of $X \cap \operatorname{cl}\mathcal{U}$ (see van der Vaart and Wellner 1996, Theorem 2.7.11). In particular, if μ has a finite $2p$th order moment for some $p \geq 1$, it holds with $F(\xi) := \widehat{L}\max\{1, \|\xi\|^2\}$ $(\xi \in \Xi)$ that

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

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

$$\kappa(v(\mu), v(\mu_n(\cdot))) = O((\log n)^{1/2} n^{-1/2})$$
$$\kappa\left(\sup_{x \in S(\mu_n(\cdot))} d(x, S(\mu)), 0\right) = O(\Psi_\mu((\log n)^{1/2} n^{-1/2})).$$

If Ξ is unbounded, $\mathcal{F}_{\mathcal{U}}$ is not uniformly bounded and Proposition 4.2 does not apply.

EXAMPLE 4.5. (MIXED-INTEGER TWO-STAGE MODELS). Let f_0 be defined as in §3.2, and let (B1)–(B3) be satisfied and Ξ be bounded. Then, for each nonempty open and bounded subset $\mathcal{U}$ of $\mathbb{R}^m$, the class

$$\mathcal{F}_{\mathcal{U}} = \left\{ f_0(\cdot, x) = \sum_{j=1}^{N} (cx + \Phi(h(\cdot) - T(\cdot)x)\chi_{\Xi^R_{j,x}}(\cdot)\colon x \in X \cap \operatorname{cl}\mathcal{U} \right\}$$

is a subset of $L_p(\Xi, \mu)$ for each $p \geq 1$. Here, the sets $\Xi^R_{j,x}$ $(j = 1, \ldots, N)$ are constructed in the proof of Theorem 3.6 such that the function $\Phi(h(\cdot) - T(\cdot)x)$ is Lipschitz continuous (with a uniform constant $L_1 > 0$) on each of them. Furthermore, for each $X \cap \operatorname{cl}\mathcal{U}$, they form disjoint Borel sets, their closures are in $\mathcal{B}_{ph_k}(\Xi)$ for some $k \in \mathbb{N}$, and it holds that $\bigcup_{j=1}^N \Xi^R_{j,x} = \Xi$ if R is chosen sufficiently large such that $\{\xi \in \Xi\colon \|h(\xi) - T(\xi)x\|_\infty > R\} = \varnothing$ for each $x \in X \cap \operatorname{cl}\mathcal{U}$.

Let $f_0^j(\cdot, x)$ denote a Lipschitz extension of the function $cx + \Phi(h(\cdot) - T(\cdot)x)$ from $\Xi^R_{j,x}$ to Ξ by preserving the Lipschitz constant L_1 $(j = 1, \ldots, N)$. Furthermore, let $\mathcal{F}^j_{\mathcal{U}} := \{f_0^j(\cdot, x)\colon x \in X \cap \operatorname{cl}\mathcal{U}\}$ and $\mathcal{G}^j_{\mathcal{U}} := \{\chi_{\Xi^R_{j,x}}\colon x \in X \cap \operatorname{cl}\mathcal{U}\}$ $(j = 1, \ldots, N)$.

Now, we use a permanence property of the uniform covering numbers (cf., van der Vaart and Wellner 1996, §2.10.3). Let $\nu \in \mathcal{P}(\Xi)$ be discrete with finite support. Then it holds that

$$(29)\qquad N(\varepsilon C_0, \mathcal{F}_{\mathcal{U}}, L_2(\Xi, \nu)) \leq \prod_{j=1}^{N} N(\varepsilon C_j, \mathcal{F}^j_{\mathcal{U}}, L_2(\Xi, \nu_j)) N(\varepsilon \widehat{C}_j, \mathcal{G}^j_{\mathcal{U}}, L_2(\Xi, \hat{\nu}_j)),$$

where C_0, $C_j > 1$, $\widehat{C}_j$, $j = 1, \ldots, N$, are certain constants and ν_j, $\hat{\nu}_j$, $j = 1, \ldots, N$, certain discrete measures having finite support. The constants depend on the bounds of the uniformly bounded classes $\mathcal{F}^j_{\mathcal{U}}$ and $\mathcal{G}^j_{\mathcal{U}}$, $j = 1, \ldots, N$. Since the latter classes satisfy Condition (i) of Proposition 4.2 (see Examples 4.3 and 4.4), (29) implies that $\mathcal{F}_{\mathcal{U}}$ satisfies (i), too. Hence, we obtain the same rates of convergence for mixed-integer two-stage models as in Example 4.4 for two-stage models without integrality requirements.

EXAMPLE 4.6. (EXAMPLE 1.1 AND 2.10 CONTINUED). According to Example 2.10, the class $\mathcal{F}_{\mathcal{U}}$ has the form $\mathcal{F}_{\mathcal{U}} = \{f_0(\cdot, x) = (r-c)x + c\max\{0, x-\cdot\}\colon x \in X \cap \operatorname{cl}\mathcal{U}\}$. Hence, it is a subset of $L_p(\Xi, \mu)$ if $\int_\Xi \|\xi\|^p \mu(d\xi) = \sum_{k\in\mathbb{N}} \pi_k k^p < \infty$ $(p \geq 1)$. As in Example 4.4, we obtain

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

and hence, Proposition 4.2 provides the same rate of convergence of the solution sets $S(\mu_n(\cdot))$ of (2) as in Example 4.4 with linear Ψ_μ.

5. Stability of stable portfolios with minimal risk. Stable probability distributions for modelling asset returns were proposed and discussed in the fundamental work of Mandelbrot (1963) and in Ziemba (1974) and Mittnik and Rachev (1993), for instance. We also refer to the recent monograph (Rachev and Mittnik 2000), 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 $\mathbb{R}^s$ is of the form

$$\begin{aligned}\varphi(t) &= \mathbb{E}\exp(i\langle t, \varrho\rangle)\\ &= \exp\left\{-\int_{\Sigma^s} |\langle \xi, t\rangle|^\alpha \left(1 - i\operatorname{sign}\langle \xi, t\rangle \tan\frac{\pi\alpha}{2}\right)\Gamma(d\xi) + i\langle m_0, t\rangle\right\},\end{aligned}$$

where $\Sigma^s = \{\xi \in \mathbb{R}^s\colon \langle \xi, \xi\rangle = 1\}$ is the unit sphere in $\mathbb{R}^s$, Γ is a finite Borel measure on Σ^s (*spectral measure*), $\alpha \in (1, 2)$ is the *stability index*, $m_0 \in \mathbb{R}^s$ is the *shift* of the stable law and $\langle\cdot,\cdot\rangle$ denotes the (Euclidean) scalar product in $\mathbb{R}^s$.

Denoting by $x \in \mathbb{R}^s_+$ 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., $r_{\alpha,\Gamma}(x) := \int_{\Sigma^s} |\langle x, \xi\rangle|^\alpha \Gamma(d\xi)$. For a discussion of the risk of a stable portfolio and related aspects, the reader is referred to Cheng and Rachev (1995) and to §8.4 of Rachev and Mittnik (2000). The classical problem of the choice of the *efficient portfolio* corresponds to the optimization problem,

$$\min\left\{\int_{\Sigma^s} |\langle x, \xi\rangle|^\alpha \Gamma(d\xi)\colon x \in \mathbb{R}^s_+, \sum_{i=1}^s x_i = 1\right\}, \tag{30}$$

which fits into the form of the stochastic programming model (1) by putting $m = s$, $d = 0$, $f_0(\xi, x) = |\langle x, \xi\rangle|^\alpha$, X to be the standard simplex $\{x \in \mathbb{R}^s_+ : \sum_{i=1}^s x_i = 1\}$, $\Xi = \Sigma^s$, 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 (30), 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 $\mathbb{R}^s$ for each $\xi \in \Xi$, and for any $x, \tilde{x} \in X$ and $\xi, \tilde{\xi} \in \Xi$ it holds that*

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

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 $|\langle x, \xi\rangle| \leq \|x\|\,\|\xi\| = \|x\| \leq (\sum_{i=1}^s x_i)^{1/2} = 1$ 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 $\int_{\Sigma^s}|\langle x_* - x^*, \xi\rangle|^2\Gamma(d\xi) = 0$ holds for any two elements $x_*, x^* \in S(\alpha, \Gamma)$. This leads to calling a spectral measure $\Gamma \in \mathscr{P}(\Sigma^s)$ *nonsingular* if the relation $\int_{\Sigma^s}|\langle x, \xi\rangle|^2\Gamma(d\xi) = 0$ implies $x = 0$. Example 5.3 illustrates that relevant spectral measures are nonsingular, indeed.

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

$$\sup_{x \in S(\tilde{\alpha}, \widetilde{\Gamma})} d(x, S(\alpha, \Gamma)) \leq \Psi_\Gamma(\alpha\zeta_1(\Gamma, \widetilde{\Gamma}) + \exp(-1)|\alpha - \tilde{\alpha}|)$$

whenever $(\tilde{\alpha}, \widetilde{\Gamma}) \in (1, 2) \times \mathscr{P}(\Sigma^s)$ *and* $\zeta_1(\Gamma, \widetilde{\Gamma}) + |\alpha - \tilde{\alpha}| < \delta$.

Moreover, if Γ *is nonsingular,* (30) *has a unique solution and it holds for the associated function* Ψ_Γ (*see* (9)) *that* $\Psi_\Gamma(\eta) \leq C\eta^{1/2}$ *for some* $C > 0$ *and sufficiently small* $\eta \in \mathbb{R}_+$.

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 $\Gamma, \widetilde{\Gamma} \in \mathscr{P}(\Sigma^s)$, we obtain

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

Then the first part of the result is an immediate consequence of Lemma 5.1 and Theorem 2.4. The additional assumption implies that the mapping $x \mapsto (\int_{\Sigma^s}|\langle x, \xi\rangle|^2\Gamma(d\xi))^{1/2}$ from $\mathbb{R}^s$ to $\mathbb{R}$ is a norm on $\mathbb{R}^s$. 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 $x \in \mathbb{R}^s$. We conclude that the risk function $r_{\alpha,\Gamma}$ is strongly convex on X and that (30) 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, the function $\psi_\Gamma(\tau) = \frac{1}{4}c\alpha(\alpha-1)\tau^2$ $(\tau \in \mathbb{R}_+)$ is a growth function of (30). Referring to Remark 2.5 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 (Samorodnitsky and Taqqu 1994, Proposition 2.3.7).

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

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

where $\beta_k \in \mathbb{R}$, $\gamma_k \geq 0$ depend on the parameters of $\varrho^{(k)}$ and on A, and the vectors $s_k \in \Sigma^s$ are normalized columns of A (see Samorodnitsky and Taqqu 1994, Example 2.3.6). Hence, it holds that

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

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

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

$$\begin{aligned} \kappa\left(\sup_{x \in S(\alpha_n, \Gamma_n)} d(x, S(\alpha, \Gamma)), 0 \right) &= O(\kappa(\zeta_1(\Gamma_n, \Gamma), 0)^{1/2} + \kappa(\alpha_n, \alpha)^{1/2}) \\ &= O((\log n)^{1/4} n^{-1/4}), \end{aligned}$$

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 $\mathbb{R}$. Finally, we note that Theorem 5.2 also applies to the study of convergence properties of the estimation procedures described in Chapter 8 of Rachev and Mittnik (2000).

Acknowledgments. 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 thank the mentioned institutions for their hospitality and support. Moreover, suggestions by two anonymous referees are gratefully acknowledged.

References

Artstein, Z. 1994. Sensitivity with respect to the underlying information in stochastic programs. *J. Comput. Appl. Math.* **56** 127–136.

———, R. J.-B. Wets. 1994. Stability results for stochastic programs and sensors, allowing for discontinuous objective functions. *SIAM J. Optim.* **4** 537–550.

———, ———. 1995. Consistency of minimizers and the SLLN for stochastic programs. *J. Convex Anal.* **2** 1–17.

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

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

Blair, C. E., R. G. Jeroslow. 1977. The value function of a mixed integer program. *Discrete Math.* **19** 121–138.

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

Cheng, B. N., S. T. Rachev. 1995. Multivariate stable commodities in the futures market. *Math. Finance* **5** 133–153.

Dentcheva, D., W. Römisch. 2000. Differential stability of two-stage stochastic programs. *SIAM J. Optim.* **11** 87–112.

Dudley, R. M. 1984. A course on empirical processes. P. L. Hennequin, ed. *École d'Été de Probabilities de Saint-Flour XII—1982, Lecture Notes in Mathematics* 1097. Springer-Verlag, Berlin. 2–142.

———. 1989. *Real Analysis and Probability*. Wadsworth & Brooks/Cole, Pacific Grove.

Dupačová, J. 1990. Stability and sensitivity analysis for stochastic programming. *Ann. Oper. Res.* **27** 115–142.

———. 1994. Applications of stochastic programming under incomplete information. *J. Comput. Appl. Math.* **56** 113–125.

———, R. J.-B. Wets. 1988. Asymptotic behaviour of statistical estimators and of optimal solutions of stochastic optimization problems. *Ann. Statist.* **16** 1517–1549.

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

Gröwe, N. 1997. Estimated stochastic programs with chance constraints. *Eur. J. Oper. Res.* **101** 285–305.

Henrion, R., W. Römisch. 1999. Metric regularity and quantitative stability in stochastic programs with probabilistic constraints. *Math. Programming* **84** 55–88.

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

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

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

Kaňková, V. 1994. A note on estimates in stochastic programming. *J. Comput. Appl. Math.* **56** 97–112.

King, A. J., R. T. Rockafellar. 1993. Asymptotic theory for solutions in statistical estimation and stochastic programming. *Math. Oper. Res.* **18** 148–162.

———, R. J.-B. Wets. 1991. Epi-consistency of convex stochastic programs. *Stochastics and Stochastics Rep.* **34** 83–92.

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

———. 1994. On quantitative stability for non-isolated minima. *Control Cybernetics* **23** 183–200.

Mandelbrot, B. B. 1963. New methods in statistical economics. *J. Political Econom.* **71** 421–440.

Mittnik, S., S. T. Rachev. 1993. Modeling asset returns with alternative stable distributions. *Econometric Rev.* **12** 261–330.

Mordukhovich, B. S. 1994. Lipschitzian stability of constraint systems and generalized equations. *Nonlinear Anal., Theory, Methods Appl.* **22** 173–206.

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

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

———. 1999. Stochastic programs and statistical data. *Ann. Oper. Res.* **85** 59–78.

Rachev, S. T. 1991. *Probability Metrics and the Stability of Stochastic Models*. Wiley, Chichester, U.K.

———, S. Mittnik. 2000. *Stable Paretian Models in Finance*. Wiley, Chichester, U.K.

———, L. Rüschendorf. 1998. *Mass Transportation Problems, Vol. I: Theory*. Springer, New York.

Robinson, S. M. 1987. Local epi-continuity and local optimization. *Math. Programming* **37** 208–223.

———. 1996. Analysis of sample-path optimization. *Math. Oper. Res.* **21** 513–528.

———, R. J.-B. Wets. 1987. Stability in two-stage stochastic programming. *SIAM J. Control Optim.* **25** 1409–1416.

Rockafellar, R. T., R. J.-B. Wets. 1997. *Variational Analysis*. Springer, Berlin.

Römisch, W., R. Schultz. 1991a. Distribution sensitivity in stochastic programming. *Math. Programming* **50** 197–226.

———, ———. 1991b. Stability analysis for stochastic programs. *Ann. Oper. Res.* **30** 241–266.

———, ———. 1996. Lipschitz stability for stochastic programs with complete recourse. *SIAM J. Optim.* **6** 531–547.

———, A. Wakolbinger. 1987. Obtaining convergence rates for approximations in stochastic programming. J. Guddat, H. T. Jongen, B. Kummer, F. Nožička, eds. *Parametric Optimization and Related Topics*. Akademie-Verlag, Berlin. 327–343.

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

Schultz, R. 1994. Strong convexity in stochastic programs with complete recourse. *J. Comput. Appl. Math.* **56** 3–22.

———. 1995. On structure and stability in stochastic programs with random technology matrix and complete integer recourse. *Math. Programming* **70** 73–89.

———. 1996. Rates of convergence in stochastic programs with complete integer recourse. *SIAM J. Optim.* **6** 1138–1152.

———. 2000. Some aspects of stability in stochastic programming. *Ann. Oper. Res.* **100** 55–84.

Shapiro, A. 1994. Quantitative stability in stochastic programming. *Math. Programming* **67** 99–108.

———. 1996. Simulation-based optimization—Concergence analysis and statistical inference. *Comm. Statist.—Stochastic Models* **12** 425–454.

———, T. Homem-de-Mello. 2000. On rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs. *SIAM J. Optim.* **11** 70–86.

Talagrand, M. 1994. Sharper bounds for Gaussian and empirical processes. *Ann. Probab.* **22** 28–76.

———. 1996. The Glivenko-Cantelli problem, ten years later. *J. Theoret. Probab.* **9** 371–384.

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

Vogel, S. 1994. A stochastic approach to stability in stochastic programming. *J. Comput. Appl. Math.* **56** 65–96.

Walkup, D. W., R. J.-B. Wets. 1969. Lifting projections of convex polyhedra. *Pacific J. Math.* **28** 465–475.

Wets, R. J.-B. 1994. Stochastic programs with fixed recourse: The equivalent deterministic program. *SIAM Rev.* **16** 309–339.

———. 1989. Stochastic programming. G. L. Nemhauser, A. H. G. Rinnooy Kan, M. J. Todd, eds. *Handbooks in Operations Research and Management Science, Vol. 1, Optimization.* North-Holland, Amsterdam. 573–629.

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

Zolotarev, V. M. 1983. Probability metrics. *Theory Probab. Appl.* **28** 278–302.

S. T. Rachev: Department of Statistics and Applied Probability, University of California, Santa Barbara, California 93106, and Institute of Statistics and Mathematical Economics, University of Karlsruhe, Postfach 6980, D-76128 Karlsruhe, Germany; e-mail: rachev@pstat.ucsb.edu

W. Römisch: Institute of Mathematics, Humboldt University Berlin, D-10099, Berlin, Germany; e-mail: romisch@mathematik.hu-berlin.de