Noname manuscript No.
(will be inserted by the editor)

Approximation and contamination bounds for probabilistic programs

Martin Branda · Jitka Dupačová

Received: date / Accepted: date

Abstract Development of applicable robustness results for stochastic programs with probabilistic constraints is a demanding task. In this paper we follow the relatively simple ideas of output analysis based on the contamination technique and focus on construction of computable global bounds for the optimal value function. Dependence of the set of feasible solutions on the probability distribution rules out the straightforward construction of these concavity-based global bounds for the perturbed optimal value function whereas local results can still be obtained. Therefore we explore approximations and reformulations of stochastic programs with probabilistic constraints by stochastic programs with suitably chosen recourse or penalty-type objectives and fixed constraints. Contamination bounds constructed for these substitute problems may be then implemented within the output analysis for the original probabilistic program.

Keywords Stochastic programs with probabilistic constraints · output analysis · contamination technique

1 Modeling issues

Classical stochastic programming (SP) models aim at hedging against consequences of possible realizations of random parameters — *scenarios* — so that the expected final outcome or position is the best possible.

Modeling part of realistic applications consists of a clear declaration of random factors to be taken into account, of distinguishing between hard and soft

Martin Branda
Charles University in Prague, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics,
E-mail: branda@karlin.mff.cuni.cz

Jitka Dupačová
Charles University in Prague, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics,
Tel.: +420-221913280
Fax: +420-222323316
E-mail: dupacova@karlin.mff.cuni.cz

constraints and of a choice of a sensible optimality criterion. The starting point may be formulation of a deterministic problem which would be solved if no randomness is considered, e.g.

$$\min \{f(x) \,:\, x \in \mathcal{X}, g_k(x) \leq 0,\, k = 1, \ldots, m\}$$

with $\mathcal{X} \subset \mathbb{R}^n$ and with real functions f, $g_k \, \forall k$.

Taking into account the presence of a random factor ω and the fact that a decision x has to be chosen before ω occurs, a reformulation of the minimization problem is needed. Two prevailing approaches have been used to this purpose:

- static expected penalty models,
- probabilistic programs.

For penalty type models, $\mathcal{X}$ is defined by hard constraints plus some other conditions that guarantee plausible properties of the model, whereas soft constraints, such as $g_k(x, \omega) \leq 0$, are reflected by penalties included into the random objective function. For probabilistic programs, probabilistic reliability-type constraints are introduced.

1.1 Stochastic programming models with penalties

The basic SP model with penalties is of the form

$$\min_{x \in \mathcal{X}} E_P f(x, \omega). \tag{1}$$

It is identified by

- a known probability distribution P of random parameter ω whose support Ω is a closed subset of $\mathbb{R}^s$; E_P denotes the corresponding expectation. In the sequel, the same character ω will be used both for the random vector and its realization.
- a given, nonempty, closed set $\mathcal{X} \subset \mathbb{R}^n$ of decisions x which is independent of P, that is $\mathcal{X}$ remains fixed even if several probability distributions P are considered.
- a preselected random objective f from $\mathcal{X} \times \Omega$ to the extended reals — a loss or a cost caused by the decision x when scenario ω occurs. As a function of ω, f is measurable for each fixed $x \in \mathcal{X}$ and such that its expectation $E_P f(x, \omega)$ is well defined. The structure of f may be quite complicated e.g. for multistage problems. For convex $\mathcal{X}$, a frequent assumption is that f is lower semicontinuous and convex with respect to x, i.e. f is *convex normal integrand.*

An example of (1) is the two-stage stochastic linear program with fixed recourse where $\mathcal{X}$ is convex polyhedral and the random objective function $f(x, \omega) = c^\top x + q(x, \omega)$ involves the second-stage (recourse) function q defined as

$$q(x, \omega) = \min_y \{q^\top y \,:\, Wy = b(\omega) - T(\omega)x,\, y \in \mathbb{R}^r_+\}. \tag{2}$$

Vector $q \in \mathbb{R}^r$ and recourse matrix $W(m, r)$ are fixed, $b(\omega)$, $T(\omega)$ are of consistent dimensions with components affine linear in ω.

1.2 Probabilistic constraints

Instead of (1) one may consider stochastic programs

$$\min_{x \in \mathcal{X}(P)} F(x,P) := E_P f(x,\omega) \tag{3}$$

in which the set of feasible solutions $\mathcal{X}(P) \subset \mathbb{R}^n$ depends on the probability distribution P.

A special type of (3) is *probabilistic programming* obtained when $\mathcal{X}(P) = \mathcal{X} \cap \mathcal{X}_\varepsilon(P)$ with $\mathcal{X}_\varepsilon(P)$ defined e.g. by the *joint probabilistic constraint*

$$\mathcal{X}_\varepsilon(P) := \{x \in \mathbb{R}^n : P(g(x,\omega) \leq 0) \geq 1 - \varepsilon\} \tag{4}$$

with $g : \mathbb{R}^n \times \Omega \to \mathbb{R}^m$ and $\varepsilon \in (0,1)$ fixed, chosen by the decision maker. It is a reliability type constraint which can be written as

$$H(x,P) := P(\max_{k=1,\dots,m} g_k(x,\omega) \leq 0) \geq 1 - \varepsilon. \tag{5}$$

We make use of the following convention: If $V(\omega)$ is a predicate on ω, we write $P(V(\omega))$ instead of $P(\{\omega \in \Omega : V(\omega)\})$.

Individual probability constraints are a special type of probabilistic constraints which treat conditions $g_k(x,\omega) \leq 0$ separately: Given probability thresholds $\varepsilon_1, \dots, \varepsilon_m$ the feasible solutions are $x \in \mathcal{X}$ that fulfill m individual probabilistic constraints

$$P(g_k(x,\omega) \leq 0) \geq 1 - \varepsilon_k,\ k = 1,\dots,m. \tag{6}$$

This is a relatively easy structure of problem, namely, if ω_k are separated being the right-hand sides of constraints, i.e. $g_k(x,\omega) = \omega_k - g_k(x)\,\forall k$. The constraints of (6) become

$$g_k(x) \geq u_{1-\varepsilon_k}(P_k),\ k = 1,\dots,m \tag{7}$$

whose right-hand sides $u_{1-\varepsilon_k}(P_k)$ are quantiles of marginal probability distributions P_k of ω_k. For concave $g_k(x)\,\forall k$ the set of feasible decisions is convex and for linear objective function and linear $g_k(x)\,\forall k$ the resulting problem is a linear program.

Such results are no more valid for joint probability constraints. Even for random right-hand sides only, special requirements on the probability distribution P are needed; cf. log-concave or quasiconcave probability distributions [29]. These seminal results on convexity properties of the set $\mathcal{X}_\varepsilon(P)$ and of the related function $H(x,P)$ are due to Prékopa; see e.g. [27], [28]. They have been reported and further extended in various monographs and collections devoted to stochastic programming, e.g. [1], [31].

Formally, the independence of the set of feasible solutions of P can be achieved by means of the extended real indicator functions. Problem (3) can be e.g. written as

$$\min_x \left[F(x,P) + \mathrm{ind}_{\mathcal{X}(P)}(x)\right] \tag{8}$$

with $\mathcal{X}(P) \subset \mathcal{X} \subseteq \mathbb{R}^n$, $\mathcal{X}$ a fixed closed set independent of P, and the indicator function $\mathrm{ind}_{\mathcal{X}(P)}(x) := 0$ for $x \in \mathcal{X}(P)$ and $+\infty$ otherwise. However, the resulting extended real objective function in (8) is then very likely to lose the convenient properties of the original objective function $F(x, P)$ in (3). Semicontinuity properties of the corresponding indicator functions, see e.g. [19], can be obtained under various sets of assumptions about the function g and/or its components g_k and about the probability distribution P. These play an important role in qualitative and quantitative stability analysis with respect to changes of the probability measure; see [30].

Klein Haneveld [24] suggested to replace probability constraints (4) and (6) by *Integrated Chance Constraints,* ICC

$$E_P(\max_k [g_k(x, \omega)]^+) \leq \beta \text{ and } E_P([g_k(x, \omega)]^+) \leq \beta_k \, \forall k, \tag{9}$$

respectively, with fixed nonnegative values β, β_k. Mathematical properties of ICC are much nicer and from the modeling point of view, it is convenient that integrated chance constraints *quantify the size* of infeasibilities.

The two prevailing types of static stochastic programs — with penalties and with probabilistic constraints — are not competitive but rather complementary. Contrary to penalty models, probabilistic programs capture the reliability requirements or risk restrictions even in cases which do not allow for reasonably accurate evaluation of penalties, e.g. [14]. A suggestion of [28] is to apply probabilistic constraints (4), (5) or (6) and at the same time, to extend the objective function for an expected penalty term which is active whenever the original constraints $g_k(x, \omega) \leq 0$ are not fulfilled:

> Prékopa [28] "...we are convinced that the best way of operating a stochastic system is to operate it with a prescribed (high) reliability and at the same time use penalties to punish discrepancies."

Another suggestion of [28] was to assign a probabilistic constraint on the second-stage variables y in the two-stage stochastic program as a way how to restrict a possible unfeasibility of the second-stage constraints in incomplete recourse problems. Integrated chance constraints (9) of [24] aim at a similar goal.

Such extensions may be useful when modeling real problems. An example of a "mixed" model and its properties was recently presented in [3]; see Example 3. In addition, ideas of multiobjective programming can be used if the choice of the penalty function is not clear and multiple penalty functions are therefore considered.

For various reasons, probability distribution P may not be precisely specified and for applications of models (1) or (3) it is important to know how sensitive are the obtained results on changes in P. There exist various results in this direction, see e.g. [30] and references therein. The first issue of interest is then the sensitivity of the obtained optimal solutions with respect to perturbances of the probability distribution which may be quantified, inter alia, by bounds on the "error" in the perturbed optimal value. In this paper we shall follow the relatively simple ideas of output analysis based on the contamination technique and focus on construction of computable global lower and upper bounds for the optimal value function. The brief presentation of the contamination technique in Section 2 reveals that, due to the dependence of the set of feasible solutions on the probability distribution, the straightforward construction of global contamination bounds for probabilistic programs is hardly possible. Therefore reformulations of probabilistic programs to expected penalty-type problems are offered in Section 3. Such approach can be helpful in numerical solution of probabilistic programs, see e.g. [4], [34]. Moreover, as illustrated in Section 4, it opens a possibility to construct approximate contamination bounds for the probabilistic programs in question.

2 Contamination technique

For derivation of contamination bounds one mostly assumes that the stochastic program is reformulated as

$$\min_{x \in \mathcal{X}} F(x,P) := E_P f(x,\omega) = \int_{\Omega} f(x,\omega) P(d\omega) \tag{10}$$

where P is a fixed fully specified probability distribution of the random parameter $\omega \in \Omega \subset \mathbb{R}^s$, $\mathcal{X} \subset \mathbb{R}^n$ is nonempty, closed, *independent* of P and $f : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}$ such that the expectation E_P is well defined.

Denote $\varphi(P)$ the optimal value and $\mathcal{X}^*(P)$ the set of optimal solutions of (10). Possible changes or perturbations of probability distribution P are modeled using contaminated distributions P_λ,

$$P_\lambda := (1-\lambda)P + \lambda Q,\ \lambda \in [0,1] \tag{11}$$

with Q another *fixed* probability distribution. Limiting thus the analysis to a selected direction $Q - P$ only, the results are directly applicable but they are less general than quantitative stability results with respect to arbitrary (but small) changes in P summarized e.g. in [30]. On the other hand, contaminated probability distributions may also capture substantial changes in P : for example with P, Q carried by different beds of scenarios the contaminated distribution is carried by the pooled sample and resistance of results obtained for P with respect to additional scenarios – atoms of Q – can be analyzed; see e.g. [11], [13], [15].

Via contamination, robustness analysis with respect to changes in probability distribution P gets reduced to a much simpler analysis with respect to a scalar parameter λ : The objective function in (10) is *linear* in P, hence the perturbed objective

$$F(x,\lambda) := \int_{\Omega} f(x,\omega) P_\lambda(d\omega) = (1-\lambda)F(x,P) + \lambda F(x,Q)$$

is linear in λ. Suppose for simplicity that stochastic program (10) has an optimal solution for all considered distributions P_λ, $0 \le \lambda \le 1$ of the form (11). Then the optimal value function

$$\varphi_{PQ}(\lambda) := \min_{x \in \mathcal{X}} F(x, P_\lambda)$$

is concave on $[0,1]$ which implies its continuity and existence of directional derivatives on $(0,1)$. Its continuity and existence of the directional derivative at the point $\lambda = 0$ is a property related with stability results for the stochastic program in question. In general, one needs a nonempty, bounded set of optimal solutions

$\mathcal{X}^*(P)$ of the initial stochastic program (10). This assumption together with stationarity of derivatives $\frac{dF(x,\lambda)}{d\lambda} = F(x,Q) - F(x,P)$ is used to derive the form of the directional derivative

$$\varphi'_{PQ}(0^+) = \min_{x \in \mathcal{X}^*(P)} F(x,Q) - \varphi(P) \tag{12}$$

which enters the upper bound for the optimal value function $\varphi_{PQ}(\lambda) = \varphi(P_\lambda)$:

$$\varphi(P) + \lambda\varphi'_{PQ}(0^+) \geq \varphi_{PQ}(\lambda) \geq (1-\lambda)\varphi(P) + \lambda\varphi(Q),\ \lambda \in [0,1]; \tag{13}$$

for details see [8], [11] and references therein. Formula (12) follows e.g. by application of Theorem I. in Chapter III. of [6] provided that $\mathcal{X}$ is compact and the both objectives $F(x,P)$, $F(x,Q)$ are continuous in x.

If $x^*(P)$ is the *unique* optimal solution of (10), $\varphi'_{PQ}(0^+) = F(x^*(P),Q) - \varphi(0)$, i.e. the *local change of the optimal value function caused by a small change of* P *in direction* $Q - P$ *is the same as that of the objective function at* $x^*(P)$. If there are multiple optimal solutions, each of them leads to an upper bound $\varphi'_{PQ}(0^+) \leq F(x(P),Q) - \varphi(P)$, $x(P) \in \mathcal{X}^*(P)$. Relaxed contamination bounds can be then written as

$$(1-\lambda)\varphi(P) + \lambda F(x(P),Q) \geq \varphi_{PQ}(\lambda) \geq (1-\lambda)\varphi(P) + \lambda\varphi(Q) \tag{14}$$

valid for an arbitrary optimal solution $x(P) \in \mathcal{X}^*(P)$ and for all $\lambda \in [0,1]$.

To construct contamination bounds (13) or (14) one exploits concavity property of the optimal value function $\varphi_{PQ}(\lambda)$ and the existence and the problem specific form of the directional derivative $\varphi'_{PQ}(0^+)$. For problems with $F(x,P)$ *concave* in P and $\mathcal{X}$ *independent* of P concavity of the optimal value function $\varphi(\lambda)$ is preserved. Under additional assumptions, e.g. convexity of the stochastic program (10), one may then apply general results by [2], [6], [17] and others to get the existence and the form of the directional derivative

$$\varphi'_{PQ}(0^+) = \min_{x \in \mathcal{X}^*(P)} \frac{d}{d\lambda} F(x,P_\lambda)\Big|_{\lambda=0+} \tag{15}$$

which enters contamination bounds (13); see [11], [12].

Also convexity with respect to x can be relaxed, e.g. Theorem 8 of [10] and the general result in Theorem 4.26 of [2]. It means that contamination bounds can be derived also for mixed integer SLP with recourse [7].

In the present paper we shall discuss the role of contamination bounds in output analysis for stochastic programs with probabilistic constraints and related SP problem formulations. As the set of feasible solutions depends on P, the optimal value function $\varphi_{PQ}(\lambda)$ is not concave in general so that a direct application of the contamination technique will be successful only exceptionally.

EXAMPLE 1

As the first example consider the stochastic linear program with individual probabilistic constraints and random right-hand sides ω_k

$$\min_{x \in \mathcal{X}} \{c^\top x \,:\, P(\omega_k - T_k x \leq 0) \geq 1 - \varepsilon_k,\ k = 1,\dots,m\}.$$

(We assume for simplicity that $\mathcal{X} = \mathrm{I\!R}^n$.) It reduces to a linear program whose right-hand sides are the corresponding quantiles $u_{1-\varepsilon_k}(P_k)$ of the marginal probability distributions P_k :

$$\varphi(P) := \min\{c^\top x \,:\, T_k x \geq u_{1-\varepsilon_k}(P_k),\, k = 1, \ldots, m\}. \tag{16}$$

Assume that the optimal value $\varphi(P)$ of (16) is finite. Using duality theory for linear programming, it can be expressed as

$$\varphi(P) = \max_z\{\sum_k z_k u_{1-\varepsilon_k}(P_k) \,:\, T^\top z = c,\, z \in \mathrm{I\!R}^m_+\} \tag{17}$$

where T is the (m, n) matrix composed of rows $T_k,\ k = 1, \ldots, m$. This is a problem whose set of feasible solutions is fixed and only the objective function depends, in a nonlinear way, on probability distribution P; hence, it seems to be in the form suitable for construction of contamination bounds for optimal value $\varphi_{PQ}(\lambda) := \varphi(P_\lambda)$. Denote $z^*(P)$ an optimal solution of (17).

For one-dimensional probability distribution P and under assumptions about existence and continuity of its positive density p on a neighborhood of the quantile $u_\alpha(P)$ we get derivatives of quantiles $u_\alpha(P_\lambda)$ of contaminated probability distribution P_λ at $\lambda = 0^+$, cf. [32]: Let Γ denote the distribution function of the contaminating probability distribution Q. Then

$$\frac{d}{d\lambda} u_\alpha((1-\lambda)P + \lambda Q)\Big|_{\lambda=0^+} = \frac{\alpha - \Gamma(u_\alpha(P))}{p(u_\alpha(P))}. \tag{18}$$

Hence, if the contaminated objective function $\sum_k z_k u_{1-\varepsilon_k}((1-\lambda)P_k + \lambda Q_k)$ of (17) is convex in λ, we get (maximization type) contamination bounds with

$$\varphi'_{PQ}(0^+) = \sum_k z^*_k(P) \frac{1 - \varepsilon_k - \Gamma_k(u_{1-\varepsilon_k}(P_k))}{p_k(u_{1-\varepsilon_k}(P_k))}$$

where Γ_k denote marginal distribution functions of probability distribution Q.

The main obstacle is that convexity with respect to λ of quantiles $u_\alpha(P_\lambda)$ cannot be guaranteed (cf. [23]). It means that $\sum_k z_k u_{1-\varepsilon_k}((1-\lambda)P_k + \lambda Q_k)$, the objective function of the contaminated dual (maximization) linear program, need not be convex in λ either. To overcome the difficulties, let us follow the suggestion of [9]. Assume that the optimal solution $x^*(P)$ of (16) is unique and nondegenerated, the marginal densities p_k are for all k continuous and positive at the points $T_k x^*(P)$ and the marginal distribution functions of the contaminating probability distribution Q have continuous derivatives on the neighborhoods of the points $T_k x^*(P)$. Using (18), approximate the right-hand sides of (16) linearly

$$u_{1-\varepsilon_k}((1-\lambda)P_k + \lambda Q_k) \approx u_{1-\varepsilon_k}(P_k) + \lambda \frac{du_{1-\varepsilon_k}((1-\lambda)P_k + \lambda Q_k)}{d\lambda}\Bigg|_{\lambda=0^+}.$$

Approximate the optimal solution $x^*(P_\lambda)$ of the contaminated program

$$\min_{x \in \mathcal{X}}\{c^\top x \,:\, T_k x \geq u_{1-\varepsilon_k}((1-\lambda)P_k + \lambda Q_k),\, k = 1, \ldots, m\} \tag{19}$$

by an optimal solution $\hat{x}(P_\lambda)$ of parametric linear program

$$\min_{x\in\mathcal{X}} \left\{ c^\top x \, : T_k x \geq u_{1-\varepsilon_k}(P_k) + \lambda \frac{1-\varepsilon_k - \Gamma_k(u_{1-\varepsilon_k}(P_k))}{p_k(u_{1-\varepsilon_k}(P_k))}, \, k=1,\ldots,m \right\} \quad (20)$$

whose properties are well known; namely, the optimal value function $\hat{\varphi}(\lambda)$ of (20) is convex piecewise linear in λ and $\hat{\varphi}(0) = \varphi(P)$. This allows us to construct the convexity based contamination bounds for (20). Moreover, if the noncontaminated problem (with $\lambda = 0$) has a unique nondegenerated optimal solution, there is a $\lambda_0 > 0$ such that $\hat{\varphi}(\lambda)$ is linear on $[0, \lambda_0]$ and the optimal basis B of the linear program dual to (20) stays fixed. Namely, $\hat{x}(P) = (B^\top)^{-1} \left[u_{1-\varepsilon_i}(P_i)\right]_{i\in I}$ for I denoting the set of active constraints of (16). In virtue of our assumptions, there is $\lambda_0 > 0$ such that for $0 \leq \lambda \leq \lambda_0$, the set of active constraints remains fixed, B is optimal basis of (20) and the optimal solution is

$$\begin{aligned}\hat{x}(P_\lambda) &= x^*(P) + \lambda (B^\top)^{-1} \left[\frac{1-\varepsilon_k - \Gamma_k(u_{1-\varepsilon_k}(P_k))}{p_k(u_{1-\varepsilon_k}(P_k))} \right]_{k\in I} \\ &= x^*(P) + \lambda \mathrm{d}x^*(P; Q-P).\end{aligned}$$

For $0 \leq \lambda \leq \lambda_0$ the optimal value equals

$$\hat{\varphi}(\lambda) = \varphi(P) + \lambda c^\top (B^\top)^{-1} \left[\frac{1-\varepsilon_k - \Gamma_k(u_{1-\varepsilon_k}(P_k))}{p_k(u_{1-\varepsilon_k}(P_k))} \right]_{k\in I}$$

where the second term determines the slope of the lower contamination bound. The upper contamination bound for $\hat{\varphi}(\lambda)$ follows by convexity:

$$\hat{\varphi}(\lambda) \leq (1-\lambda)\varphi(P) + \lambda \hat{\varphi}(1).$$

This approximation was applied in [9] to real data for a water resources planning problem and similar ideas were used also for contamination of empirical VaR; cf. [15]. Notice that the obtained bounds are based on a *local* approximation (20) of the contaminated problem (19).

In the last Section we shall explore construction of approximate global contamination bounds based on ideas of multiobjective programming and on a relationship between optimal solutions of stochastic programs with probabilistic constraints and optimal solutions of stochastic programs with suitably chosen recourse or penalty type objectives and fixed constraints.

3 Probabilistic programs and their recourse reformulations

Stochastic linear programs with convex polyhedral set $\mathcal{X}$, with nonrandom matrix T composed of rows T_k, $k = 1, \ldots, m$, and with individual probabilistic constraints on random right-hand sides ω_k can be qualified as an easy case. They reduce to linear programs whose right-hand sides are the corresponding quantiles $u_{1-\varepsilon_k}(P_k)$ of the marginal probability distributions P_k, cf. (16). Moreover, under additional assumptions, local contamination bounds can be constructed, see Example 1.

In general, even individual linear probabilistic programs (6) with $g_k(x, \omega) := b_k(\omega) - T_k(\omega)x$ and right-hand sides $b_k(\omega)$ and rows $T_k(\omega)$ affine linear in ω, need not be easy to solve. Besides of application of Monte Carlo sampling methods (cf. [25], [26], [33]) the existing algorithms for probabilistic programs have been developed for special structures of the model and/or probability distributions, cf. [20], [21], [29]. Hence, approximations by easier problems are of interest.

The idea proposed already in [22] is to construct another problem with a fixed set $\mathcal{X}$ of feasible solutions related to the probabilistic program

$$\min\{E_P f(x,\omega)\ :\ x \in \mathcal{X} \cap \mathcal{X}_\varepsilon(P)\}, \tag{21}$$

with $\mathcal{X}_\varepsilon(P)$ defined by (4): One assigns penalties $N[g_k(x,\omega)]^+$ with positive penalty coefficients and solves

$$\min_{x \in \mathcal{X}} \left[E_P f(x,\omega) + N \sum_{k=1}^{m} E_P[g_k(x,\omega)]^+ \right] \tag{22}$$

instead of (21).

ILLUSTRATIVE NUMERICAL EXAMPLE

Consider the example of probabilistic program with one linear joint probabilistic constraint taken from [28]:

$$\begin{aligned} &\min 3x_1 + 2x_2 \\ &\text{subject to} \\ &x_1 + 4x_2 \geq 4,\ 5x_1 + x_2 \geq 5,\ x_1 \geq 0,\ x_2 \geq 0, \\ &P(x_1 + x_2 - 3 \geq \omega_1, 2x_1 + x_2 - 4 \geq \omega_2) \geq 1 - \varepsilon. \end{aligned} \tag{23}$$

The random components (ω_1, ω_2) have bivariate normal distribution with $E[\omega_1] = E[\omega_2] = 0$, $E[\omega_1^2] = E[\omega_2^2] = 1$, and $E[\omega_1\omega_2] = 0.2$. The corresponding simple recourse model may be formulated as follows.

$$\begin{aligned} \min\ &3x_1 + 2x_2 + N \cdot E\Big[(\omega_1 - x_1 - x_2 + 3)^+ + (\omega_2 - 2x_1 - x_2 + 4)^+\Big] \\ &\text{subject to} \\ &x_1 + 4x_2 \geq 4,\ 5x_1 + x_2 \geq 5,\ x_1 \geq 0,\ x_2 \geq 0. \end{aligned} \tag{24}$$

We used SLP-IOR, see [21], and the solver PROCON for solving the problem (23) with joint probabilistic constraint for decreasing levels ε; the solver SRAPPROX was used for simple recourse models (24) with increasing N.

When comparing results from tables 1 and 2, we observe that optimal solutions and optimal values of (23) and (24) behave similarly with increasing N and decreasing ε, respectively. A question is if there is a quantitative relation between the outputs of the two models in dependence on the choice of parameters N and ε.

The two problems (21) and (22) are not equivalent. However, by intuition one may expect that for a given $\varepsilon > 0$ there exists N large enough such that the obtained optimal solution, say $x_N(P)$, of (22) satisfies the probabilistic constraint (4). Then the corresponding value of $E_P f(x_N(P),\omega)$ may serve as an upper bound for the optimal value of (21). This conjecture is supported by the analysis of optimality conditions for the simple recourse problem with random right-hand sides; they provide a link between the simple-recourse penalty coefficients and the

Table 1 Optimal values and solutions for simple recourse model.

N	First-stage objective value	Recourse objective value	Optimal solution $\hat{x}_1^N$	$\hat{x}_2^N$
1	4.1053	2.9596	0.8421	0.7895
10	9.5596	0.9285	1.0000	3.2798
100	11.6185	0.5887	1.0000	4.3092
1000	12.9373	0.2680	1.0000	4.9686
10000	13.4786	0.0494	1.0000	5.2393
100000	13.5703	0.0052	1.0000	5.2851
1000000	13.5800	0.0004	1.0000	5.2900

Table 2 Optimal values and solutions for probabilistic program.

ε	Objective value	Optimal solution $\hat{x}_1^\varepsilon$	$\hat{x}_2^\varepsilon$
$0,2$	9.4438	1.0000	3.2219
$0,1$	10.2349	1.0000	3.6174
$0,05$	10.8899	1.0000	3.9450
$0,01$	12.1382	1.0000	4.5691
$0,001$	13.5754	1.0000	5.2877

values of the probability thresholds for the corresponding problem with individual probabilistic constraints and random right-hand sides, cf. Section 3.2 of [1]. A more complicated form of penalty may be designed as the optimal value of the recourse function $q(x,\omega)$ of the second-stage linear program $\min_y\{q^\top y : Wy \geq b(\omega) - T(\omega)x, y \geq 0\}$ with some $q \in \mathbb{R}^m_+$ and a fixed recourse matrix W; problem (22) corresponds to $q_k = N\,\forall k$ and the simple recourse matrix $W = I$.

For a numerical evidence see also [14] where a piecewise linear nonseparable penalty function $N[\max_j(x_j - \omega_j)]^+$ was applied and the results compared with those obtained for a joint probabilistic constraint of the type $P\{x_j \leq \omega_j,\, j = 1,\ldots,n\} \geq 1-\varepsilon$. The recent numerical experiments of [4] show that the penalty-based reformulations are able to generate feasible solutions of the original probabilistic program with a high reliability.

Also convexity preserving penalty functions, say $\vartheta : \mathbb{R}^m \to \mathbb{R}_+$ which are continuous, nondecreasing in their components and are equal to 0 on $\mathbb{R}^m_-$ and positive otherwise seem to be a suitable choice. This idea stems from the connection between (21) and (22) that can be recognized within the framework of *convex* multiobjective stochastic programming. For example optimal solutions of (21) with $\mathcal{X}_\varepsilon(P)$ given by (5) can be viewed as efficient solutions of the *bi-criterial* optimization problem

$$\text{“min”}_{x\in\mathcal{X}}[F(x,P), -H(x,P)]$$

which are obtained by the ϵ-constrained approach. Hence, for convex $\mathcal{X}$ and for $F(\cdot,P) = E_P f(\cdot,\omega)$ convex and $H(\cdot,P) := P(\omega : g(\cdot,\omega) \leq 0)$ concave there exists $N \geq 0$ such that these efficient solutions can be found as optimal solutions of the *parametric program*

$$\min_{x\in\mathcal{X}}[F(x,P) - N\cdot H(x,P)].$$

Similarly, such relationship can be established for stochastic programs of the form

$$\min_{x \in \mathcal{X}} \{E_P f(x,\omega) : E_P h_k(x,\omega) \le \epsilon_k,\ k = 1,\dots,m\} \tag{25}$$

with $h_k(x,\omega)$ suitable transforms of $g_k(x,\omega)$ and with a distribution dependent set of feasible solutions. Evidently, problems with probabilistic constraints and those with integrated chance constraints are special instances of (25). Again, for fixed thresholds ϵ_k, $k = 1,\dots,m$, the optimal solutions of (25) may be viewed as efficient solutions of the multiobjective problem

$$\text{“min”}_{x \in \mathcal{X}} \{E_P f(x,\omega),\ E_P h_k(x,\omega),\ k = 1,\dots,m\} \tag{26}$$

obtained by the ϵ-constrained approach.

If $\mathcal{X}$ is nonempty, convex, compact and the functions $E_P f(x,\omega), E_P h_k(x,\omega)$, $k = 1,\dots,m$, are *convex* in x on $\mathbb{R}^n$ then there exists a nonnegative parameter vector $t \in \mathbb{R}^m$, $t \ne 0$ such that the efficient points can be obtained by solving a *scalar convex optimization problem*

$$\min_{x \in \mathcal{X}} [E_P f(x,\omega) + \sum_{k=1}^{m} t_k E_P h_k(x,\omega)]. \tag{27}$$

This scalarization is a special form of scalarization by a penalty function $\vartheta : \mathbb{R}^m \to \mathbb{R}$ which must be continuous and nondecreasing in its arguments to provide efficient solutions of (26). (For relevant results on multiobjective optimization see e.g. [18].)

Problem (27) corresponds to (22) with $h_k(x,\omega) = [g_k(x,\omega)]^+$. In general, $h_k(x,\omega)$ can be a suitable penalty function applied to $g_k(x,\omega) \le 0$ and $t_k = N$, $k = 1,\dots,m$.

A rigorous proof of the relationship between optimal values and solutions of (21) and those of (22) for the penalty function $N \sum_{k=1}^{m} [g_k(x,\omega)]^+$ is due to Ermoliev et al. [16]. It is valid under modest assumptions on the nonlinear functions g_k, on continuity of the probability distribution P and on the structure of problem (21).

The approach by [16] can be further extended to a whole class of penalty functions ϑ. For functions $\vartheta : \mathbb{R}^m \to \mathbb{R}_+$ which are continuous nondecreasing in their components, equal to 0 on $\mathbb{R}^m_-$ and positive otherwise, it holds that

$$P\big(g_k(x,\omega) \le 0,\ 1 \le k \le m\big) \ge 1 - \varepsilon \iff P\big(\vartheta(g(x,\omega)) > 0\big) \le \varepsilon.$$

The considered penalty function problem can be formulated as follows

$$\varphi_N(P) = \min_{x \in \mathcal{X}} \Big[E_P f(x,\xi) + N \cdot E_P \vartheta(g(x,\omega))\Big] \tag{28}$$

with N a positive parameter. We denote $x_N(P)$ an optimal solution of (28) and $x_\varepsilon(P)$ an optimal solution of (21) with a level $\varepsilon \in (0,1)$.

Theorem 1 *For a fixed probability distribution P consider the two problems (21) and (28) and assume:*

$\mathcal{X} \neq \emptyset$ compact, $F(x,P) = E_P f(x,\omega)$ a continuous function of x,

$\vartheta : \mathbb{R}^m \to \mathbb{R}_+$ a continuous function, nondecreasing in its components, which is equal to 0 on $\mathbb{R}^m_-$ and positive otherwise,

(i) $g_k(\cdot,\omega)\,\forall k$ are almost surely continuous;
(ii) there exists a nonnegative random variable $C(\omega)$ with $E_P C^{1+\kappa}(\omega) < \infty$ for some $\kappa > 0$, such that $|\vartheta(g(x,\omega))| \leq C(\omega)$;
(iii) $E_P \vartheta(g(x^{'},\omega)) = 0$, for some $x^{'} \in \mathcal{X}$;
(iv) $P(g_k(x,\omega) = 0) = 0\,\forall k$, for all $x \in \mathcal{X}$.

Denote $\gamma = \kappa/2(1+\kappa)$, and for arbitrary $N > 0$ and $\varepsilon \in (0,1)$ put

$$\begin{aligned}\varepsilon(N) &= P\big(\vartheta(g(x_N(P),\omega)) > 0\big),\\ \alpha(N) &= N \cdot E_P \vartheta(g(x_N(P),\omega)),\\ \beta(\varepsilon) &= \varepsilon^{-\gamma} E_P \vartheta(g(x_\varepsilon(P),\omega)).\end{aligned}$$

THEN for any prescribed $\varepsilon \in (0,1)$ there always exists N large enough so that minimization (28) generates optimal solutions $x_N(P)$ which also satisfy the probabilistic constraints (4) with the given ε.

Moreover, bounds on the optimal value $\psi_\varepsilon(P)$ of (21) based on the optimal value $\varphi_N(P)$ of (28) and vice versa can be constructed:

$$\begin{aligned}\varphi_{1/\varepsilon^\gamma(N)}(P) - \beta(\varepsilon(N)) \leq \psi_{\varepsilon(N)}(P) &\leq \varphi_N(P) - \alpha(N),\\ \psi_{\varepsilon(N)}(P) + \alpha(N) \leq \quad \varphi_N(P) \quad &\leq \psi_{1/N^{1/\gamma}}(P) + \beta(1/N^{1/\gamma}),\end{aligned} \tag{29}$$

with

$$\lim_{N\to+\infty} \alpha(N) = \lim_{N\to+\infty} \varepsilon(N) = \lim_{\varepsilon\to 0_+} \beta(\varepsilon) = 0.$$

For the proof see [5].

It means that under assumptions of the Theorem, the two problems (21), (28) are *asymptotically equivalent* which is a useful theoretical result. However, the theorem does not make any statement on the convergence of optimal solutions but it relates optimal values for certain values of the levels and the penalty parameter.

Remark 1 The assumption (iii) can be very strong because it requires existence of a permanently feasible point for which the constraints are fulfilled for almost all realizations of the random vector ω. The problem is that the overall feasible set may shrink with increasing levels to the empty set, hence, the approach may fail for probability measures with an unbounded support.

Remark 2 The assumption (iv) ensures that the probability function

$$H(x,P) = P\big(g(x,\omega) \leq 0\big)$$

is continuous in the decision vector, which can be easy seen if we realize that the only point of discontinuity of the function is $g_k(x,\omega) = 0$ for any k and x.

Notice, however, that when we want to evaluate one of the bounds in (29), we must be prepared to face some problems. We solve the penalty function problem (28) taking a sufficiently large $N > 0$ to get its optimal solution $x_N(P)$ and optimal value $\varphi_N(P)$. Then we are able to compute $\alpha(N)$, $\varepsilon(N)$, hence the upper bound for the optimal value $\psi_{\varepsilon(N)}(P)$ of the probabilistic program (4) with probability level $\varepsilon(N)$. But we are not able to compute $\beta(\varepsilon(N))$ without having the solution $x_{\varepsilon(N)}(P)$ which we do not want to find or even may not able to find. We can only solve the penalty function problem with $N = 1/\varepsilon^\gamma(N)$ getting its optimal solution $x_{1/\varepsilon^\gamma(N)}(P)$ and optimal value $\varphi_{1/\varepsilon^\gamma(N)}(P)$ which is only a part of the lower bound for the optimal value $\psi_{\varepsilon(N)}(P)$. A question for future research is how to choose the parameter N so that the probability level ε is ensured. A recent numerical study in this direction can be found in [34] in the context of a beam design optimization.

The bounds (29) and the terms $\alpha(N)$, $\varepsilon(N)$ and $\beta(\varepsilon)$ depend also on the choice of the penalty function ϑ. Two special penalty functions are readily available: $\vartheta_1(u) = \sum_{k=1}^m [u_k]^+$ applied in [16] and $\vartheta_2(u) = \max_{1\le k\le m}[u_k]^+$ applied in [14].

The obtained problems with penalties and with a fixed set of feasible solutions are much simpler to solve and analyze then the probabilistic programs; namely, contamination technique can be used for output analysis of the penalty function problems (28). Exploration of this possibility and of its meaning for output analysis of the original probabilistic program is the subject of the next Section.

4 Approximate contamination bounds

As discussed in Section 2, to construct global bounds for the optimal value of stochastic programs under contamination of the probability distribution, i.e. with respect to the contamination parameter λ, we need that the optimal value function $\varphi_{PQ}(\lambda)$ is *concave* to get (13). With the set of feasible solutions dependent on P, concavity of the optimal value function $\varphi_{PQ}(\lambda)$ cannot be guaranteed. Our suggestion is to approximate probabilistic programs by penalty-type problems which possess a fixed set of feasible first-stage solutions. This in turn opens the possibility to construct approximate contamination bounds.

EXAMPLE 2.

For the general probabilistic program (21), let us view the problem

$$\varphi(P;t) = \min_{x\in\mathcal{X}} \left[E_P f(x,\omega) + \sum_{k=1}^m t_k E_P [g_k(x,\omega)]^+ \right] \tag{30}$$

with convex compact $\mathcal{X} \neq \emptyset$, convex functions $f(\cdot,\omega)$, $g_k(\cdot,\omega)$, $k = 1,\dots,m$ and a fixed nonnegative parameter vector $t \in \mathbb{R}^m$ as an acceptable substitute for the probabilistic program (21)

$$\min\{E_P f(x,\omega) \,:\, x \in \mathcal{X} \cap \mathcal{X}_\varepsilon(P)\}$$

with $\mathcal{X}_\varepsilon(P)$ defined by (4). Using the additivity of the penalty term in (30) we can rewrite (30) as

$$\min_{x\in\mathcal{X}} \Theta(x,P;t) := E_P \theta(x,\omega;t)$$

with the random objective

$$\theta(x,\omega;t) = f(x,\omega) + \sum_{k=1}^{m} t_k [g_k(x,\omega)]^+$$

which is convex in x. The set of feasible solutions $\mathcal{X}$ does not depend on P and for fixed vectors t, the contamination bounds for the optimal value $\varphi(P_\lambda;t)$ of (30) for contaminated probability distribution (11) follow the usual pattern (14). They may serve as an indicator of robustness of the optimal value of (21) and may support decisions about the choice of weights t_k.

EXAMPLE 3.

In the model of [3], the second-stage variables y are supposed to satisfy with a prescribed probability $1-\delta$ certain constraints driven by another random factor, η, whose probability distribution S is independent of P. The problem is

$$\min_{x\in\mathcal{X}} F(x,P,S) := E_P f(x,\omega,S) = c^\top x + \int_\Omega R(x,\omega,S)P(d\omega) \qquad (31)$$

where

$$R(x,\omega,S) = \min_y \{q^\top y \, : Wy = h(\omega) - T(\omega)x,\ y \in \mathrm{IR}^r_+,\ S(Ay \le \eta) \ge 1-\delta\}, \quad (32)$$

S denotes the probability distribution of s-dimensional random vector η and $A(s,r)$ is a deterministic matrix composed of rows A_j, $j = 1,\dots,s$. With respect to P, (31) is an expectation type of stochastic program, with a fixed set $\mathcal{X}$ of feasible first-stage decisions and with an incomplete recourse; hence, there is a good chance to construct contamination bounds on the optimal value $\varphi(P_\lambda, S)$ of (31) for a *fixed* probability distribution S and for the contaminated probability distribution P_λ.

Assume that T is nonrandom with a full row rank, S is fixed and such that the set $\mathcal{Y}_S := \{y \in \mathrm{IR}^r_+ :\ S(Ay \le \eta) \ge 1-\delta\}$ is convex and bounded. Let the set $\mathcal{X}^*(P,S)$ of optimal solutions of (31) be nonempty and bounded. Under further assumptions which guarantee finiteness and convexity of the second stage value function $R(\cdot,\omega,S)$ for a fixed probability distribution S (hence finiteness and convexity of the random objective value function $f(\cdot,\omega,S)$), assumptions on existence of its expectation with respect to P and Q and assumptions on joint continuity of $F(\cdot,\cdot,S)$ in the decision vector x and the probability measure P, cf. [3], the directional derivative of the optimal value function $\varphi(\cdot,S)$ at P in the direction $Q-P$ exists and is of the form (12), i.e.

$$\min_{x\in\mathcal{X}^*(P,S)} F(x,Q,S) - \varphi(P,S). \qquad (33)$$

To model perturbations with respect to S and to construct the corresponding contamination bounds is more demanding. A possibility is to keep a fixed probability distribution P and to replace the probabilistic constraint in (32) by an expected (with respect to S) penalty term added to $q^\top y$, as done e.g. in (22), for $N>0$ sufficiently large:
$R(x,\omega,S) \approx$

$$\min_y \{q^\top y + N\sum_{j=1}^{s} E_S[A_j y - \eta_j]^+ \ :\ Wy = h(\omega) - T(\omega)x,\ y \ge 0\} := \tilde{R}(x,\omega,S). \quad (34)$$

As we have shown in Theorem 1, such approximation works well under particular assumptions, e.g. for compact set of feasible solutions and under existence of a permanently feasible solution. We obtain then the approximate problem

$$\min_{x\in\mathcal{X}} \tilde{F}(x,P,S) := E_P\tilde{f}(x,\omega,S) = c^\top x + \int_\Omega \tilde{R}(x,\omega,S)P(d\omega), \tag{35}$$

where the objective function is linear in P and concave in S, $\mathcal{X}$ is independent of P, S. This allows us to construct contamination bounds for (35) with respect to the contaminated probability distribution S. In this case, however, existence of the directional derivative of the optimal value function must be examined in detail. The first step can be to get the directional derivative of $\tilde{R}(x,\omega,S)$ at S in the direction $U - S$ according to (12):

$$\tilde{R}'_{SU}(x,\omega,S) = \min_{y\in\mathcal{Y}^*(x,\omega,S)} [q^\top y + N\sum_{j=1}^{s} E_U[A_j y - \eta_j]^+] - \tilde{R}(x,\omega,S) \tag{36}$$

where $\mathcal{Y}^*(x,\omega,S)$ is the set of optimal solutions of the approximate second-stage problem (34) for fixed x, ω.

If it is possible to interchange the expectation with respect to P and the directional derivative, we can obtain the directional derivative of the objective function $\tilde{F}(x,P,S)$ of (35) which is concave in S. This allows us to use Theorem 4.26 of [2] to obtain the directional derivative of the optimal value function $\hat{\varphi}(P,S)$ of (35) at S for fixed P in the direction $U - S$:

$$\min_{x\in\mathcal{X}^*(P,S)} \int_\Omega \tilde{R}'_{SU}(x,\omega,S)P(d\omega), \tag{37}$$

where $\tilde{R}'_{SU}(x,\omega,S)$ is the directional derivative (36) of the approximated recourse function $\tilde{R}(x,\omega,S)$.

The last example illustrates the merits of the suggested use of contamination bounds constructed for the penalty-type reformulation in output analysis for the original probabilistic program, here for the stress testing with respect to an extremal scenario.

EXAMPLE 4.
In the jointly constrained probabilistic program

$$\begin{aligned} &\min x_1 + x_2 \\ &\text{subject to} \\ &P(\omega_1 x_1 + x_2 \geq 7, \omega_2 x_1 + x_2 \geq 4) \geq 1 - \varepsilon, \\ &x_1 \geq 0,\ x_2 \geq 0. \end{aligned} \tag{38}$$

the random components (ω_1, ω_2) are independent and have uniform distributions on the intervals $[1, 4]$ and $[1/3, 1]$. The explicit solution can be obtained, cf. [26]. The problem can be solved by the sample approximation method, cf. [25], [26], [33]. We verified that the sample based penalty-type reformulation provides good solutions of the original probabilistic problem. This enables us to exploit the approximate contamination bounds analysis of the initial problem (38).

The corresponding model with the penalty objective may be formulated as follows:

$$\begin{aligned} \min\ & x_1 + x_2 + N \cdot E\Big[(7 - \omega_1 x_1 - x_2)^+ + (4 - \omega_2 x_1 - x_2)^+\Big] \\ & \text{subject to} \\ & x_1 \geq 0,\ x_2 \geq 0. \end{aligned} \tag{39}$$

We chose the level $\varepsilon = 0.01$. By numerical experiments we found that the sample size which is necessary to obtain a good approximation of the probabilistic program (PS) is 500. We also found that by setting the penalty parameter N to 140, the penalty term becomes very small so that the optimal values of the both sample based problems (38) and (39) are almost the same, hence the penalty reformulation (PR) can serve as a good approximation of the probabilistic program (PP).

To stress the sample distribution we choose the extremal scenario $(\tilde{\omega}_1, \tilde{\omega}_2) = (4, 1)$. The contamination bounds are shown in Figure 1, where also optimal values of the sample-based problems PP and PR are plotted.

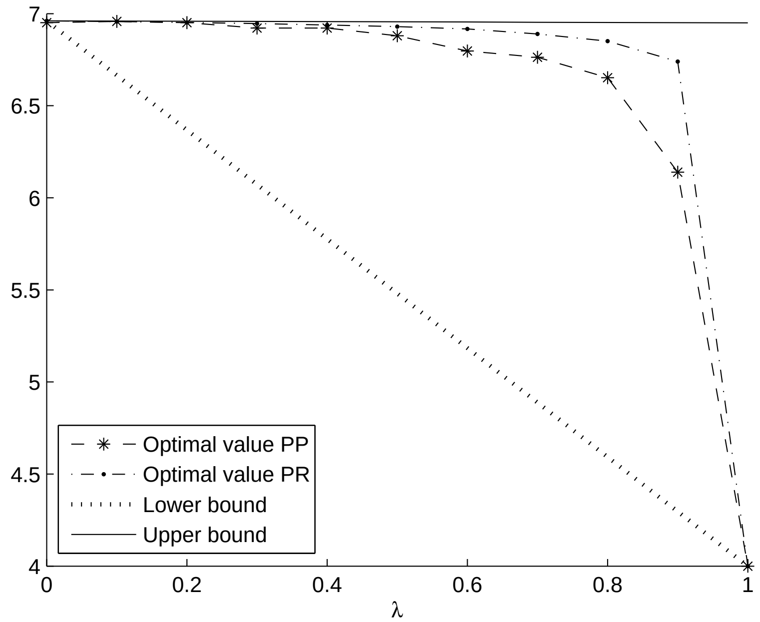

Fig. 1 Contamination bounds.

Notice, that contamination may influence the choice of N for which the penalty-type approximation is precise enough. In our example the required precision is evidently not acheived if $\lambda > 0.4$.

5 Conclusions

Reformulation of probabilistic programs by incorporating a suitably chosen penalty function into the objective helps to arrive at problems with a fixed set of feasible solutions whose optimal solutions are linked with the optimal solutions of the original probabilistic program. The recommended form of the penalty function agrees with the basic ideas of multiobjective programming and suitable properties of the approximation approach follow by generalization of results by [16].

Provided that the approximation error is low the next idea is to exploit the contamination bounds constructed for the penalty-type reformulation to test stability of the optimal value function of the original probabilistic program with respect to perturbations of the probability distribution P or its resistance with respect to additional out-of-sample scenarios. As the set of feasible solutions of probabilistic programs depends on the probability distribution, the possibility of obtaining contamination bounds for probabilistic programs directly is limited to problems of a very special form and/or under special distributional and structural assumptions.

Acknowledgements The research of the first author has been partly supported by the grant SVV 2010/261315. The second author acknowledges the support by the project “Methods of modern mathematics and their applications” – MSM 0021620839 and by the grant 402/08/0107 of the Czech Science Foundation.

References

1. Birge, J. R., Louveaux, F. (1997), *Introduction to Stochastic Programming*, Springer Series on Oper. Res., New York.
2. Bonnans, J. S., Shapiro, A. (2000), *Perturbation Analysis of Optimization Problems.* Springer Series in OR, New York.
3. Bosch, P., Jofré, A., Schultz, R. (2007), Two-stage stochastic programs with mixed probabilities, *SIAM. J. Optim.* **18**, 778–788.
4. Branda, M. (2010), Reformulation of general chance constrained problems using the penalty functions, *SPEPS* 2010-2.
5. Branda, M., Dupačová, J. (2008), Approximation and contamination bounds for probabilistic programs, *SPEPS* 2008-13 (available at www.stoprog.org).
6. Danskin, J. M. (1967), *Theory of Max-Min*, Springer, Berlin (*Econometrics and Operations Research* **5**).
7. Dobiáš, P. (2003), Contamination for stochastic integer programs, *Bull. Czech Econometric Soc.* **18**, 65–80.
8. Dupačová, J. (1986), Stability in stochastic programming with recourse – contaminated distributions, *Math. Programming Study* **27**, 133–144.
9. Dupačová, J. (1987), Stochastic programming with incomplete information: A survey of results on postoptimization and sensitivity analysis, *Optimization* **18**, 507–532.
10. Dupačová, J. (1990), Stability and sensitivity analysis in stochastic programming, *Ann. Oper. Res.* **27**, 115-142.
11. Dupačová, J. (1996), Scenario based stochastic programs: Resistance with respect to sample, *Ann. Oper. Res.* **64**, 21–38.
12. Dupačová, J. (2008) Risk objectives in two-stage stochastic programming problems, *Kybernetika* **44**, 227–242.
13. Dupačová, J., Bertocchi, M. and Moriggia, V. (1998), Postoptimality for scenario based financial models with an application to bond portfolio management. In: Ziemba, W. T. and Mulvey, J. (eds.) *World Wide Asset and Liability Modeling.* Cambridge Univ. Press, 1998, pp. 263–285.
14. Dupačová, J., Gaivoronski, A. A., Kos, Z., Szántai, T. (1991), Stochastic programming in water management: A case study and a comparison of solution techniques, *European J. Oper. Res.* **52**, 28–44.

15. Dupačová, J., Polívka, J. (2007), Stress testing for VaR and CVaR, *Quantitiative Finance* **7**, 411–421.
16. Ermoliev, T., Ermolieva, Y. M., MacDonald, G. J., Norkin, V. I. (2000), Stochastic optimization of insurance portfolios for managing expose to catastrophic risks, *Ann. Oper. Res.* **99**, 207–225.
17. Gol'štejn, E. G. (1972), *Theory of Convex Programming,* Translations of Mathematical Monographs 36, American Mathematical Society, Providence, RI.
18. Guddat, J., Guerra Vasquez, F., Tammer, K., Wendler, K. (1985), *Multiobjective and Stochastic Optimization Based on Parametric Optimization*, Akademie-Verlag, Berlin.
19. Kall, P. (1987), On approximations and stability in stochastic programming. In: Guddat, J. et al. (eds.) *Parametric Optimization and Related Topics*. Akademie-Verlag, Berlin, pp. 387–400.
20. Kall, P., Mayer, J. (2005), *Stochastic Linear Programming: Models, Theory and Computation.* Springer International Series, New York.
21. Kall, P., Mayer, J. (2005), Building and solving stochastic linear programming models with SLP-IOR. In: Wallace, S.W., Ziemba, W.T. (eds.), *Applications of stochastic programming*, MPS-SIAM Book Series on Optimization 5.
22. Kall, P., Ruszczyński, A., Frauendorfer, K. (1988), Approximation techniques in stochastic programming. In: Ermoliev, Yu., Wets, R. J-B. (eds.), *Numerical Techniques for Stochastic Optimization*, Springer, Berlin, pp.34–64.
23. Kibzun, A. I., Kan, Y. S. (1996), *Stochastic Programming Problems with Probability and Quantile Functions*, Wiley, Chichester.
24. Klein Haneveld, W. K. (1986), *Duality in Stochastic Linear and Dynamic Programming*, LNEMS **274**, Springer, Berlin.
25. Luedtke J., Ahmed, S. (2008), A sample approximation approach for optimization with probabilistic constraints, *SIAM J. Optim.* **19**, 674–633.
26. Pagnoncelli, B. K., Ahmed, S., Shapiro, A. (2009), Sample average approximation method for chance constrained programming: Theory and applications, *J Optim Theory Appl* **142**, 399–416.
27. Prékopa, A. (1971), Logarithmic concave measures with application to stochastic programming, *Acta Sci. Math. (Szeged)* **32**, 301–316.
28. Prékopa, A. (1973), Contributions to the theory of stochastic programming, *Math. Program.* **4**, 202–221.
29. Prékopa, A. (2003), Probabilistic programming, Chapter 5 in [31], pp. 267–351.
30. Römisch, W. (2003), Stability of stochastic programming problems. Chapter 8 in: [31], pp. 483–554.
31. Ruszczynski, A., Shapiro A. (eds.) (2003), *Handbook on Stochastic Programming*, Elsevier, Amsterdam.
32. Serfling, R. J. (1980), *Approximation Theorems of Mathematical Statistics*, Wiley, New York.
33. Shapiro, A. (2003), Monte Carlo sampling methods, Chapter 6 in [31], pp. 353–425.
34. Žampachová, E. (2010), *Approximations in Stochastic Optimization and Their Applications*, PhD Thesis, Brno University of Technology.