

# Probability maximization by inner approximation

**Csaba I. Fábián<sup>1</sup>, Edit Csizmás<sup>1</sup>, Rajmund Drenyovszki<sup>1</sup>, Wim van Ackooij<sup>2</sup>, Tibor Vajnai<sup>1</sup>, Lóránt Kovács<sup>1</sup>, Tamás Szántai<sup>3</sup>**

<sup>1</sup> Department of Informatics, GAMF: Faculty of Engineering and Computer Science, John von Neumann University. Izsáki út 10, 6000 Kecskemét, Hungary.

<sup>2</sup> EDF Research and Development, Department OSIRIS. 1, avenue du Général de Gaulle, F-92141 Clamart Cedex France.

<sup>3</sup> Department of Differential Equations, Institute of Mathematics, Budapest University of Technology and Economics. Műgyetem rakpart 3-9, 1111 Budapest, Hungary.

E-mails: fabian.csaba@gamf.uni-neumann.hu, csizmas.edit@gamf.uni-neumann.hu, drenyovszki.rajmund@gamf.uni-neumann.hu, wim.van-ackooij@edf.fr, vajnai.tibor@gamf.uni-neumann.hu, kovacs.lorant@gamf.uni-neumann.hu, szantai@math.bme.hu.

---

*Abstract: We solve probability maximization problems using an approximation scheme that is analogous to the classic approach of  $p$ -efficient points, proposed by Prékopa to handle chance constraints. But while  $p$ -efficient points yield an approximation of a level set of the probabilistic function, we approximate the epigraph. The present scheme is easy to implement and is immune to noise in gradient computation.*

*Keywords: stochastic programming; probabilistic constraints; applications.*

---

## 1 Introduction

A probabilistic constraint is of the following type:

$$P(g(x, \xi) \leq 0) \geq p, \quad (1)$$

where  $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^k$  is a mapping,  $\xi \in \mathbb{R}^m$  a multivariate random vector with associated probability measure  $P$  and  $p \in [0, 1]$  a user defined safety level. When  $k \geq 1$ , the terminology *joint probabilistic constraint* is also frequently employed, since we would like the random inequality system  $g(x, \xi) \leq 0$  to hold with high-enough probability.

We are interested in two general optimization problems associated with (1), namely that of maximizing the probability function and a classic problem of optimizing

under constraint (1). These appear under the following form:

$$\max P(g(x, \xi) \leq 0) \quad \text{subject to } x \in X, \quad \text{and} \quad (2)$$

$$\min c^T x \quad \text{subject to } P(g(x, \xi) \leq 0) \geq p, x \in X, \quad (3)$$

where  $X$  is a convex compact set. In many applications  $X$  is a polyhedral set  $X = \{x \in \mathbb{R}^n : Ax \leq b\}$ . We will make the assumption that  $g$  is jointly-quasi concave and that  $\xi$  admits a density (with respect to the Lebesgue-measure) disposing of generalized concavity properties as well. Under these assumptions the mapping  $x \mapsto \phi(x) := P(g(x, \xi) \leq 0)$  also disposes of generalized concavity properties. In particular problems (2) and (3) are convex optimization problems under these assumptions.

In the present paper we will deal with the special case when  $g(x, \xi) = \xi - Tx$ . Then the problems (2) and (3) become the following:

$$\max P(Tx \geq \xi) \quad \text{subject to } Ax \leq b, \quad (4)$$

and the probabilistic constrained problem

$$\min c^T x \quad \text{subject to } P(Tx \geq \xi) \geq p, Ax \leq b, \quad (5)$$

where the decision vector is  $x$ . Given are the matrices  $A, T$  and the vectors  $b, c$ , of corresponding sizes. The probability  $1 > p > 0$  is set, and the distribution of the random vector  $\xi$  is known. We assume that the feasible domains are not empty and are bounded. We assume that  $\xi$  has a continuous, logconcave distribution. It follows that the cumulative distribution function  $F(z) = P(z \geq \xi)$  is logconcave.

Probabilistic constraints arise in many applications such as water management, telecommunications, electricity network expansion, mineral blending, chemical engineering etc. (e.g., [21, 41, 52, 53, 55, 69, 76, 78]). With the advance of information technologies, new areas of application are emerging, e.g., smart grids and transportation systems.

For an overview of recent theory and algorithmic treatment of probabilistic constraints we refer to [9, 49, 50]. Other monographs dealing (partially) with probabilistic constraints are [8, 26, 38] and [37], where the latter focussed more on algorithms.

### **A brief history of methods for solving probabilistically constrained problems**

Programming under probabilistic constraints as a decision model under uncertainty, has been introduced by [7]. In this paper the authors use the term *chance constrained programming* for this model and its variants as well as extensions presented, among others, in the paper [6]. However these early chance constrained models were based on *individual chance constraints*, i.e., instead of a constraint of the type in problem (3), the following type constraints were used:  $P(g_i(x, \xi) \leq 0) \geq p_i, i = 1, \dots, k$ . Programming under probabilistic constraint with a random right hand side vector  $\xi$  (as it stands in problem (5)), having stochastically independent components, was first considered by [39]. The more general problem (3), where  $\xi$  is allowed to have stochastically dependent components, was introduced by Prékopa [44, 46] and

further investigated by him and his followers. A significant step for the numerical treatment of probabilistic constraints was laid out when convexity statements based on the theory of logconcave measures were developed by Prékopa [45, 47] and later generalized by [3, 4, 63]. Recent advances in convexity statements for probabilistic constraints are based on eventual convexity and can be found in [23, 24, 70]

In [52], Prékopa and co-authors developed a model (STABIL) for a planning problem in the Hungarian electrical energy sector, which is of the form (5). The resulting stochastic programming problem is solved by a feasible direction method of Zoutendijk [81]. It should be noted however that Zoutendijk's method lacks the global convergence property as shown in [64]. We refer to the discussion in [40] for further information.

Cutting-plane methods were also developed for the probabilistic constrained problem, approximating the level set  $M(p) := \{x \in \mathbb{R}^n : P(g(x, \xi) \leq 0) \geq p\}$ . The method of Prékopa and Szántai [53] applies a Slater point to determine where to construct the next cut. (Namely, the intersection of the boundary of  $M(p)$  on the one hand, and the interval connecting the Slater point with the current iterate on the other hand.) The method is related to that of Veinott [79]. In his solver built for the STABIL problem, Szántai [61] developed a careful interval bisection algorithm for safely computing the intersection point on the boundary of  $M(p)$  when the probability values defining the probability constraints cannot be calculated with arbitrary high precision. He also applied Veinott's technique of moving the Slater point in course of the solution process, which results in faster convergence and makes the supporting hyperplane method equivalent to a method of Zoutendijk [81]. Mayer [37] proposed a central cutting plane method, an adaptation of Elzinga and Moore [13]. Cutting-plane methods converge in less iterations than feasible direction methods do, since former gradient information is retained. These methods obviously require that one is able to compute the gradient of  $\phi(x) := P(g(x, \xi) \leq 0)$  efficiently. Identifying conditions under which  $\phi$  is differentiable has led to the development of two main research directions. The first direction exploits no specific knowledge of  $\xi$  or its underlying distribution, but only differentiability properties of its density and differentiability of  $g$ . Then under several additional assumptions, including the assumption that  $B(x) := \{z \in \mathbb{R}^m : g(x, z) \leq 0\}$  is bounded in a neighbourhood of  $x$ , one can represent the gradient of  $\phi$  as an integral over  $B(x)$  and/or its boundary  $\partial B(x)$ . We refer to [35, 36, 65–67] and the references contained therein for more on this research direction. We note here, that the condition that  $B(x)$  remains bounded around a point  $x$  rules out the study of distribution functions. The second research direction exploits specific knowledge of the underlying distribution of  $\xi$  and tries to build a link between any component of the gradient of  $\phi$  and the evaluation of a quantity akin to  $\phi$ . This direction was explored in [22, 44, 54, 60, 73–75]. When combined with sophisticated software such as for instance Genz' code [17, 19] for multivariate normal distributions, high dimensional problems can be solved with significant efficiency (e.g., a case with  $k = 168$  is examined in [72]).

In the supporting hyperplane method, the inaccuracy of evaluating  $\phi$  needs to be taken into account when computing the intersection point on the boundary of  $M(p)$ . We refer to [1] for such an approach. Still inaccuracy of  $\nabla\phi$  may result in a cut

cutting into the level set  $M(p)$ . This leads to the development of the notion of *upper-oracle* in [77] and specialized proximal ([77]) and level ([72]) bundle methods for probabilistically constrained problems with underlying convexity structure.

A non-standard dual formulation for problems of type (5) was proposed by Komáromi [27, 28]. This is a max-min formulation, the inner problem being minimization of a linear function over the level set  $M(p)$ . For the solution of the dual problem, a special feasible direction method is developed in [27].

We are going to focus on p-efficient point approaches. Other recent algorithmic approaches for probabilistically constrained programming are the penalty approach [14], scenario approximation [5], convex approximation [42], sample average approximation and integer programming [31–33, 43], binarization approaches [29, 30].

### On p-efficient point approaches

When the mapping  $g$  is of the form  $g(x, z) := z - h(x)$ , the probabilistic constraint is said to be separable and properties of  $\phi(x) = \mathbb{P}(g(x, \xi) \leq 0) := F_\xi(h(x))$  relate directly to that of the multivariate distribution function  $F_\xi$ . In this setting, Prékopa [48] initiated a new solution approach by introducing the concept of p-efficient points. A point  $z$  is p-efficient if and only if  $F_\xi(z) \geq p$  and there exists no  $z'$  such that  $z' \leq z$ ,  $z' \neq z$ ,  $F_\xi(z') \geq p$ . Prékopa, Vizvári, and Badics [56] employ this concept in the solution of problems of the type (5), where the random parameters have a discrete finite distribution. They first enumerate all the p-efficient points, and based on these, propose a convex relaxation of the problem. The relaxed probabilistic constraint prescribes the existence of a point  $z$  in the convex hull of the p-efficient points such that  $h(x) \geq z$  holds. The relaxed problem is then solved by a cutting-plane method. In essence, the cuts generated correspond to facets of the convex hull of the p-efficient points.

Prékopa [51] considers a problem equivalent to (5), where the random vector has a continuous logconcave distribution. He combines the cutting-plane method of [56] with the supporting hyperplane method of Szántai [61]. The resulting hybrid method simultaneously constructs inner and outer approximations of the level set  $M(p)$ . The supporting hyperplane method is used to generate p-efficient points in the course of the solution process. (More general stochastic programming models are also proposed in [51], but in the present paper we restrict ourselves to simpler formulations.)

Dentcheva, Prékopa, and Ruszczyński [12] consider problems of type (5), where the random parameters are integer valued. They prove that the probabilistic constraint is essentially convex, in case the random parameters have an r-concave distribution. The probabilistic constraint is formulated in a split form:  $h(x) \geq z$ , where  $z$  belongs to (a discrete version of) the level set  $M(p)$ . These authors construct a Lagrangian dual by relaxing the constraint  $h(x) \geq z$ , and observe that the dual functional splits into the sum of two functionals. The addend functionals are the respective optimal objective value functions of two simpler problems. The first auxiliary problem is a linear programming problem, and the second auxiliary problem is about minimizing a linear function over (a discrete version of) the level set  $M(p)$ . Once the

dual problem is solved, a primal optimal solution can be constructed, though technical problems may occur and need to be overcome. These authors also develop a new specialized method which separates the generation of  $p$ -efficient points and the solution of the approximate problem based on known  $p$ -efficient points. The new method, called cone generation, employs the time-honoured concept of column generation. The inherent link with integer programming is given in [80].

Dentcheva, Lai, and Ruszczyński [10] extend these results to general convex problems, and general ( $r$ -concave) distributions. The probabilistic constraint is formulated in a split form, and the Lagrangian dual is constructed by relaxing the constraint  $h(x) \geq z$ . The dual functional splits into the sum of two functionals, like in the special case discussed in [12]. The first auxiliary problem, however, is a well-structured convex programming problem, instead of the linear programming problem of [12]. The difficult part is still the second auxiliary problem, minimizing a linear function over  $M(p)$ . These authors develop a dual method, and propose a way of recovering a primal solution. Moreover, they extend the cone generation method to a general primal-dual method.

Dentcheva and Martinez [11] developed a regularized version of the dual method of [10]. Moreover they developed a progressive augmented Lagrangian method that is a primal-dual-type method. The latter method turns out to be more efficient as it requires the solution of fewer minimization problems over the level set  $M(p)$ .

A solution framework that includes and extends various existing formulations was developed by Van Ackooij, Berge, de Oliveira and Sagastizábal [71].

## Contribution

In the present paper, we construct polyhedral approximations of the epigraphs of the probabilistic functions in problems (4) and (5). This is analogous to the use of  $p$ -efficient points. But while  $p$ -efficient points yield an approximation of a level set, we approximate the epigraph. We formulate dual problems that are analogous to those of [12,27], and [10]. The present scheme yields very convenient duals, simple formulations using conjugate functions.

The solution approaches proposed in [12] and [10] can be adapted to the present approximation scheme and dual formulations. Finding a new approximation point in the present scheme is easier than finding a  $p$ -efficient point in the schemes of [12] or [10]. – In the latter schemes, finding a  $p$ -efficient point amounts to minimization over the level set  $M(p)$ . In the present scheme, an approximation point is found by unconstrained minimization.

The present simple models and methods expose an important contrast between column generation methods and direct cutting-plane methods. Direct cutting-plane methods for probabilistic functions are difficult to implement due to noisy gradient computation. A practicable implementation requires sophisticated tolerance handling. In contrast, the column generation approach is immune to noise in gradient computation.

## 2 Problem and model formulation

Using the distribution function  $F(z)$ , let  $\phi(z) = -\log F(z)$ . Of course it is a convex function, due to the logconcavity of  $F(z)$ . Taking into account the monotonicity of the distribution function, Problem (4) can be written as

$$\min \phi(z) \quad \text{subject to} \quad Ax - b \leq 0, \quad z - Tx \leq 0. \quad (6)$$

This problem has an optimal solution, due to our assumption that the feasible domain of (4) is not empty and is bounded. Introducing non-positive multiplier vectors  $y, u$  to the respective constraints, we formulate the Lagrangian relaxation of (6):

$$\begin{aligned} & \inf_{x,z} \{ \phi(z) - y^T(Ax - b) - u^T(z - Tx) \} \\ &= \inf_z \{ \phi(z) - u^T z \} + \inf_x \{ (-y^T A + u^T T)x \} + y^T b. \end{aligned}$$

The first addend is by definition  $-\phi^*(u)$ , where  $\phi^*$  is the convex conjugate of  $\phi$ . The second addend is finite iff  $-y^T A + u^T T = 0^T$ . Hence the Lagrangian dual of (6) can be written as

$$\max_{y,u \leq 0} \{ y^T b - \phi^*(u) \} \quad \text{subject to} \quad -y^T A + u^T T = 0^T. \quad (7)$$

According to the theory of convex duality, this problem has an optimal solution, since the primal problem (6) has an optimal solution.

Concerning the probabilistic constraint, let  $\pi = -\log p$ . We formulate (5) as

$$\min c^T x \quad \text{subject to} \quad Ax - b \leq 0, \quad z - Tx \leq 0, \quad \phi(z) - \pi \leq 0. \quad (8)$$

This problem has an optimal solution, due to our assumption that the feasible domain of (5) is not empty and is bounded. Introducing the multiplier vectors  $-y \geq 0$ ,  $-u \geq 0$ ,  $v \geq 0$  to the respective constraints, we formulate the Lagrangian relaxation of (8):

$$\begin{aligned} & \inf_{x,z} \{ c^T x - y^T(Ax - b) - u^T(z - Tx) + v(\phi(z) - \pi) \} \\ &= \inf_z \{ v\phi(z) - u^T z \} + \inf_x \{ (c^T - y^T A + u^T T)x \} + y^T b - v\pi. \end{aligned}$$

The first addend is by definition  $-(v\phi)^*(u)$ . The second addend is finite iff  $c^T = y^T A - u^T T$ . Hence the Lagrangian dual of (8) can be written as

$$\max \{ y^T b - v\pi - (v\phi)^*(u) \} \quad \text{subject to} \quad y, u \leq 0, \quad v \geq 0, \quad c^T = y^T A - u^T T. \quad (9)$$

*Remark.* The function  $(v, u) \mapsto (v\phi)^*(u) = \sup_z \{ u^T z - v\phi(z) \}$  is convex by definition, and given  $(\hat{v}, \hat{u})$  in the effective domain, a gradient can be computed by finding the optimal  $z$ .

In this paper we focus on unconstrained problems. The proposed algorithms can be extended to the constrained case.

### Polyhedral models

Suppose we evaluated the function  $\phi(z)$  in the points  $z_i$  ( $i = 0, 1, \dots, k$ ). These result the function  $\phi_k(z)$ , an inner approximation (polyhedral convex upper approximation) of  $\phi(z)$ , in the usual way: given  $z$ , let

$$\phi_k(z) = \min \sum_{i=0}^k \lambda_i \phi(z_i) \quad \text{such that } \lambda_i \geq 0, \sum_{i=0}^k \lambda_i = 1, \sum_{i=0}^k \lambda_i z_i = z. \quad (10)$$

If  $z \notin \text{Conv}(z_0, \dots, z_k)$ , then we have  $\phi_k(z) = +\infty$  by definition.

The following problem is the current polyhedral model of (6):

$$\min \phi_k(z) \quad \text{subject to } Ax - b \leq 0, z - Tx \leq 0. \quad (11)$$

We assume that (11) is feasible, i.e., its optimum is  $< +\infty$ . This can be ensured by the selection of the vectors  $z_0, \dots, z_k$ . The convex conjugate of  $\phi_k$  can be computed by taking into account a finite set only, hence

$$\phi_k^*(u) = \max_{0 \leq i \leq k} \{u^T z_i - \phi(z_i)\}. \quad (12)$$

– The above observation is in accordance with Chapter X Section 3.4 of Hiriart-Urruty and Lemaréchal [25]. – Of course  $-\phi_k^*$  is a cutting-plane approximation (polyhedral concave upper approximation) of  $-\phi^*$ . Hence the following problem is a cutting-plane model of (7):

$$\max_{y, u \leq 0} \{y^T b - \phi_k^*(u)\} \quad \text{subject to } -y^T A + u^T T = 0^T. \quad (13)$$

It is easy to check that (11) and (13), considered as linear programming problems, form a primal-dual pair. We are going to examine the primal problem.

### Linear programming formulation

Introducing the notation  $\phi_i = \phi(z_i)$  ( $i = 0, \dots, k$ ), the primal model problem (11) can be formulated as follows. – Dual variables corresponding to the different constraints are indicated in the right-hand column.

$$\begin{array}{llll} \min & \sum_{i=0}^k \phi_i \lambda_i & & \\ \text{such that} & \lambda_i \geq 0 & (i = 0, \dots, k), & \\ & \sum_{i=0}^k \lambda_i & = 1, & \perp \quad \vartheta \in \mathbb{R} \\ & \sum_{i=0}^k \lambda_i z_i - Tx & \leq 0, & \perp \quad u \leq 0 \\ & Ax & \leq b. & \perp \quad y \leq 0 \end{array} \quad (14)$$

Let us assume that the primal model problem (14) has a feasible solution. Let  $(\bar{\lambda}_0, \dots, \bar{\lambda}_k, \bar{x})$  and  $(\bar{\vartheta}, \bar{u}, \bar{y})$  denote an optimal solution and an optimal dual solution, respectively – both existing due to our assumption. Let moreover

$$\bar{z} = \sum_{i=0}^k \bar{\lambda}_i z_i. \quad (15)$$

**Observation 1.** We have  $\phi_k(\bar{z}) = \sum_{i=0}^k \phi_i \bar{\lambda}_i = \bar{\vartheta} + \bar{u}^T \bar{z}$ .

**Proof.** The first equality follows from the equivalence of (14) on the one hand, and (10)-(11) on the other hand.

The second inequality is a consequence of complementarity.  $\bar{\lambda}_i > 0$  implies that the reduced cost of the  $i$ th column is 0 in (14), hence  $\bar{\vartheta} + \bar{u}^T z_i = \phi_i$ . It follows that

$$\sum_{i=0}^k \phi_i \bar{\lambda}_i = \sum_{i=0}^k (\bar{\vartheta} + \bar{u}^T z_i) \bar{\lambda}_i = \bar{\vartheta} \sum_{i=0}^k \bar{\lambda}_i + \bar{u}^T \sum_{i=0}^k \bar{\lambda}_i z_i.$$

### 3 Column generation

We solve (6) by iteratively adding improving columns to the primal model (14). An optimal dual solution (i.e., shadow price vector) of the current model problem is  $(\bar{\vartheta}, \bar{u}, \bar{y})$ .

Given a vector  $z$ , we can add the corresponding column  $(1, z, 0)$  with objective component  $\phi(z)$ . This is an improving column if its reduced cost is positive; formally, if  $\bar{\rho}(z) > 0$  holds for

$$\bar{\rho}(z) := (\bar{\vartheta}, \bar{u})^T (1, z) - \phi(z) = \bar{\vartheta} + \bar{u}^T z - \phi(z). \quad (16)$$

The vector yielding the best reduced cost can be found by maximizing  $\bar{\rho}(z)$ . Let  $\bar{\mathcal{R}}$  denote the optimal objective value.

If  $\bar{\mathcal{R}}$  is small, then  $(\bar{x}, \bar{z})$  is a near-optimal solution to (6). Otherwise an improving column can be constructed to (14).

#### A practical way of finding an improving column

In order to maximize the reduced cost, we can apply a steepest descent method to  $-\bar{\rho}(z)$ , a natural starting point being  $\bar{z}$ . However, we found the computational effort prohibitive. Hence we propose to perform just a single line search. As theoretical motivation, we put forward the following well-known theorem. (It can be found in [34] or [57].)

**Theorem 1.** Let the convex function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be twice continuously differentiable. Assume that

$$\alpha I \preceq \nabla^2 f(z) \preceq \omega I \quad (z \in \mathbb{R}^n), \quad (17)$$

where  $0 < \alpha \leq \omega$ ,  $I$  is the identity matrix, and the relation  $U \preceq V$  between matrices means that  $V - U$  is positive semidefinite. We minimize  $f$  using a steepest descent method, starting from a point  $z^0$ . Let  $z^1, \dots, z^j, \dots$  denote the iterates obtained by applying exact line search at each step. Denoting  $\mathcal{F} = \min_z f(z)$ , we have

$$f(z^j) - \mathcal{F} \leq \left(1 - \frac{\alpha}{\omega}\right)^j [f(z^0) - \mathcal{F}]. \quad (18)$$

*Remark.* Similar results can be proven for the case when approximate minimizers are found in the line search procedures. See a discussion on Armijo's rule in [34].

**Corollary.** *Provided Theorem 1 is applicable to  $f(z) = -\bar{\rho}(z)$ , we can construct a fairly good improving vector in the column generation scheme. Namely, let  $\beta$  ( $0 < \beta \ll 1$ ) be given. Taking a finite (and moderate) number of steps with the steepest descent method, we find a vector  $\hat{z}$  satisfying*

$$\bar{\rho}(\hat{z}) \geq (1 - \beta) \bar{\mathcal{R}}.$$

**Proof.** Substituting  $f(z) = -\bar{\rho}(z)$  and  $z^0 = \bar{z}$  in (18), and introducing the notation  $\rho = 1 - \alpha/\omega$ , we get

$$\bar{\mathcal{R}} - \bar{\rho}(z^j) \leq \rho^j [\bar{\mathcal{R}} - \bar{\rho}(\bar{z})]. \quad (19)$$

(We have  $\mathcal{F} = -\bar{\mathcal{R}}$  by definition.) From  $\phi_k(\cdot) \geq \phi(\cdot)$  and Observation 1, we get

$$\bar{\rho}(\bar{z}) = \bar{\vartheta} + \bar{u}^T \bar{z} - \phi(\bar{z}) \geq \bar{\vartheta} + \bar{u}^T \bar{z} - \phi_k(\bar{z}) = 0$$

Due to non-negativity,  $\bar{\rho}(\bar{z})$  can be discarded in (19), and we get

$$\bar{\rho}(z^j) \geq (1 - \rho^j) \bar{\mathcal{R}}.$$

Selecting  $j$  such that  $\rho^j \leq \beta$  yields an appropriate  $\hat{z} = z^j$ .  $\square$

Setting  $j = 1$  always resulted in a good improving vector in our computational experiments. The above discussion is only meant as motivation for performing a single line search, showing that the procedure works in an ideal case. The condition (17) obviously does not hold for every  $z$  with  $f(z) = -\bar{\rho}(z)$ . However, in the case of normal distribution, there exists a bounded box  $\mathcal{Z}$  such that the probability weight outside  $\mathcal{Z}$  can be ignored. For the sake of simplicity let us assume that the polyhedron  $\mathcal{T} = \{Tx | Ax \leq b\}$  is bounded, and that  $\mathcal{T} \subset \mathcal{Z}$ . Then we'll always have  $\bar{z} \in \mathcal{Z}$ , provided the primal model (14) has been properly initialized. Starting from  $\bar{z} \in \mathcal{Z}$ , we perform a single line search. Due to special characteristics of the function  $\phi(z)$  and due to  $\bar{u} \leq 0$  being boundable, this line search can be restricted to a bounded neighborhood of  $\mathcal{Z}$ . Such restriction would justify assumption (17). However, we implemented a simple approximate line search without restriction, and still found that iterates fell into a relatively small box.

## 4 Implementation issues

For the implementation of our method and computational study we used MATLAB with the IBM ILOG CPLEX (Version 12.6.3) optimization toolbox.

### The master problem

We assume that the distribution is standard normal. Let  $r$  denote the number of the components of the random vector (equal to the number of the rows of the matrix  $T$ ).

First we look for an appropriate  $z_0 \in \mathbb{R}^r$  vector whose inclusion makes the primal model problem (14) feasible. This is done by solving the problem

$$\begin{aligned} \max \quad & t \\ \text{such that} \quad & 1t - Tx \leq 0, \\ & Ax \leq b, \end{aligned} \tag{20}$$

where  $t \in \mathbb{R}$ , and  $1 \in \mathbb{R}^r$  denotes a vector consisting of ones. If (20) has no feasible solution then the original problem is also infeasible. On the other hand, if the objective value is not bounded then probability 1 can be achieved in the original problem. Let  $z_0 = 1t^*$ , with  $t^*$  denoting an optimal solution of (20).

Let  $\mathcal{Z} \subset \mathbb{R}^r$  denote a bounded box such that the probability weight outside  $\mathcal{Z}$  can be ignored. In our case the distribution is standard normal, hence we consider an  $r$ -dimensional box  $\mathcal{Z}$  that it is symmetrical with respect to the origo. In our experiments we worked with a box such that  $P(\mathcal{Z}) \approx 0.99$ .

Let  $z^{max} = (z_1^{max}, \dots, z_r^{max})$  denote maximal vertex of  $\mathcal{Z}$ . To ease the solution of the primal model problem (14), we initialize it by adding the following vectors (besides  $z_0$ , above)

$$\begin{aligned} z_\ell &= (z_1^{max}, \dots, z_{\ell-1}^{max}, 0, z_{\ell+1}^{max}, \dots, z_r^{max}) & (\ell = 1, \dots, r), \\ z_{r+1} &= 0, \\ z_{r+2} &= z^{max}. \end{aligned} \tag{21}$$

Consequently we have  $k = r + 2$  in (14).

We solved the master problem with the CPLEX simplex solver, applying the optimality tolerance  $1E - 4$ .

### The oracle

In accordance with Section 3, our aim is to maximize the reduced cost (16). Since  $\bar{\vartheta}$  is constant in a given iteration the oracle has to find an approximate solution to the problem  $\max_z \{\bar{u}^T z - \phi(z)\}$ . This problem can be reformulated as minimizing the function  $\phi(z) - \bar{u}^T z$ . Here  $\phi(z) = -\log F(z)$  and  $F(z)$  is the multidimensional normal distribution function.  $\phi(z)$  is a convex function, due to the logconcavity of  $F(z)$ . We implemented the approximate form of the steepest descent method described in Section 3. We perform a single line search ( $j = 1$ ) and even in this single line search, we stop with an approximate minimum. Namely, we apply the golden section ratio, see, e.g. [34]. We perform only 1 or 2 golden section ratio steps.

The steepest descent direction can be found by calculating the gradient vector of the function:

$$\nabla (\phi(z) - \bar{u}^T z) = \nabla \phi(z) - \bar{u} = -\nabla \log(F(z)) - \bar{u} = -\frac{\nabla F(z)}{F(z)} - \bar{u}. \quad (22)$$

Consequently we need to calculate the function value and gradient vector of the multidimensional normal distribution function  $F(z)$ . For this computation we use the formulas in section 6.6.4 of Prékopa's book [49]. By using these formulas the calculation of the gradient of a multidimensional probability distribution function can be reduced to computing conditional distribution function values.

The numerical computation of multivariate normal distribution values was performed with the QSIMVNV Matlab function implemented by Genz [18].

## 5 Computational study

Before describing test problems and discussing computational results, let us illustrate condition (17) with a small example.

### Preliminary examinations

We illustrate the well-conditioned nature of the objective in case of a two-dimensional standard normal distribution with moderately dependent marginals (covariance 0.5).

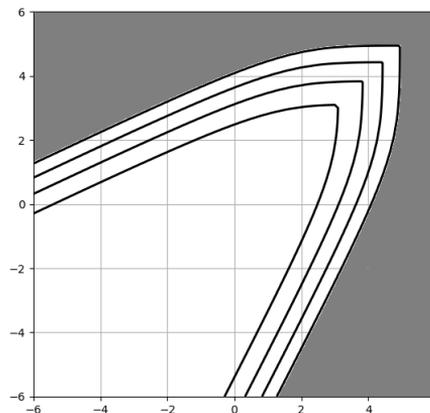


Figure 1  
Smaller eigenvalue of the Hessian  $\nabla^2 \phi(z)$   
( $-6 \leq z_1, z_2 \leq +6$ )

We depict the eigenvalues of the Hessian matrix of  $\phi(z) = -\log F(z)$ , where  $F(z)$  is the distribution function. We calculated the smaller and the larger of the two eigenvalues of the Hessian, while both components of  $z$  fall into the interval  $[-6, +6]$ .

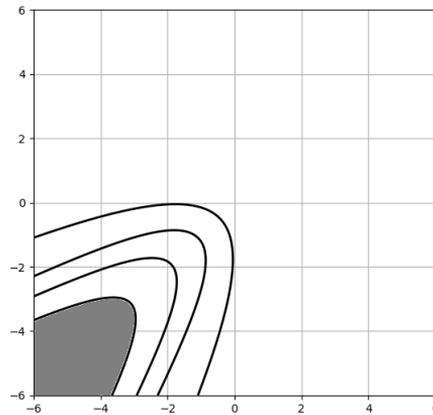


Figure 2  
Larger eigenvalue of the Hessian  $\nabla^2\phi(z)$   
( $-6 \leq z_1, z_2 \leq +6$ )

Figure 1 depicts the smaller eigenvalue. Contour lines from top right are  $1e-5, 1e-4, 1e-3, 1e-2$ . In the area not filled with gray, the smaller eigenvalue is above  $1e-5$ .

Figure 2 depicts the larger eigenvalue. Contour lines from top right are 1, 1.2, 1.4, 1.6. In the area not filled with gray, the larger eigenvalue is below 1.6.

These experiments illustrate that there is a fairly large safe domain over which  $\phi(z)$  is well-conditioned.

### Test problems

First we considered eight test problems published in [62] by T. Szántai. These problems occur at a coffee company. The company is marketing three different blends of coffee. There is a rigid set of requirements for each of the blends according their acidity, caffeine content, liquoring value, hardness and aroma. On the first day of a particular month the company found that its available supply of green coffees was limited to 8 different types. These green coffees vary according to price, quantity available and the above mentioned five taste characteristics. The demands for the company's 3 blends during the coming month are random variables with given expected values, standard deviations and correlation coefficients. The company is confronted with the problem of determining an optimum combination of available green coffees for next month's roasting operation. So they have to formulate a stochastic programming problem to satisfy all of the random demands with a prescribed (high) probability and pay the smallest possible price for the green coffees. All data and numerical results according to probability level 0.9 can be found in the paper [62]. In this paper we will call these problems 'Coffee1', ..., 'Coffee8'.

Secondly we considered an extended version of the coffee blending problem. In this extension the company is marketing five different blends of coffees and so the

multivariate normal probability distribution is five dimensional. This problem will be called 'Coffee9' in this paper.

Finally we considered a cash matching problem with fifteen dimensional normal probability distribution. In this problem we are interested in investing a certain amount of cash on behalf of a pension fund that needs to make certain payments over the coming 15 years of time. Details of this problem can be found in [10] and [20]. This problem will be called 'CashMatching' in this paper.

### Numerical results

We solved each test problem with different right-hand sides of the cost constraint. Our computational results are reported in Figures 3 - 5.

Our test problems had originally been formulated as cost minimization under a probabilistic constraint. We converted the problems to probability maximization. The right hand-sides of the cost constraints had been set in such a way that the corresponding optimal probability levels would be those listed in the column 'prescribed probability level' of our tables. For these computations we used Szántai's computer code [61].

Problem	prescribed probability level	1 GSR steps per iter			2 GSR steps per iter		
		Genz	itNum	p	Genz	itNum	p
Coffee 1	0.8	103	7	0.7998	105	6	0.7994
	0.85	78	5	0.8501	90	5	0.8504
	0.9	93	6	0.9002	186	11	0.9005
	0.95	70	5	0.9499	116	7	0.9504
	0.98	80	6	0.9798	144	11	0.9803
	0.99	70	6	0.9896	102	8	0.9900
Coffee 2	0.8	132	9	0.7998	208	12	0.8000
	0.85	107	7	0.8499	158	9	0.8499
	0.9	134	9	0.9000	166	9	0.9000
	0.95	120	8	0.9500	119	7	0.9500
	0.98	93	7	0.9800	126	8	0.9800
	0.99	84	8	0.9897	69	6	0.9897
Coffee 3	0.8	167	10	0.8000	148	8	0.8000
	0.85	129	8	0.8500	198	11	0.8500
	0.9	120	8	0.9000	152	8	0.9000
	0.95	167	11	0.9500	159	9	0.9500
	0.98	105	8	0.9800	149	9	0.9800
	0.99	71	7	0.9897	57	5	0.9897
Coffee 4	0.8	158	9	0.8000	207	11	0.8000
	0.85	172	10	0.8500	174	9	0.8500
	0.9	150	9	0.9000	155	8	0.9000
	0.95	139	9	0.9500	153	8	0.9500
	0.98	117	9	0.9800	115	7	0.9800
	0.99	69	7	0.9897	55	5	0.9897

Figure 3  
Computational results for problems 'Coffee1', ..., 'Coffee4'

Problem	prescribed probability level	1 GSR steps per iter			2 GSR steps per iter		
		Genz	itNum	p	Genz	itNum	p
Coffee 5	0.8	112	7	0.7999	131	7	0.8000
	0.85	114	7	0.8500	125	7	0.8500
	0.9	110	7	0.9000	135	7	0.9000
	0.95	112	7	0.9500	135	8	0.9499
	0.98	109	8	0.9800	113	7	0.9800
	0.99	71	6	0.9897	68	5	0.9897
Coffee 6	0.8	75	5	0.8000	104	6	0.7999
	0.85	77	5	0.8502	119	7	0.8497
	0.9	74	5	0.9004	109	6	0.9003
	0.95	81	6	0.9504	99	6	0.9506
	0.98	145	11	0.9806	84	5	0.9797
	0.99	82	6	0.9900	39	2	0.9894
Coffee 7	0.8	110	7	0.7999	127	7	0.7999
	0.85	97	6	0.8499	145	8	0.8500
	0.9	65	4	0.8997	110	6	0.8999
	0.95	70	4	0.9500	75	4	0.9500
	0.98	89	6	0.9799	108	6	0.9800
	0.99	53	4	0.9899	48	3	0.9900
Coffee 8	0.8	147	8	0.8001	105	5	0.8000
	0.85	71	4	0.8493	106	5	0.8501
	0.9	75	4	0.8999	106	5	0.8999
	0.95	80	4	0.9501	108	5	0.9500
	0.98	114	7	0.9801	114	5	0.9801
	0.99	48	3	0.9901	28	1	0.9894

Figure 4

Computational results for problems 'Coffee5', ..., 'Coffee8'

We solved each problem with two settings of the oracle, performing either 1 or 2 Golden Section Ratio (GSR) steps in course of each line search. – The corresponding data are shown under the headers '1 GSR step per iter' and '2 GSR steps per iter'. In each case, we list the number of calls to the Genz subroutine (under the header 'Genz'), the number of oracle calls (under the header 'itNum'), and the optimum found (under the header 'p').

In each case, most of the computation time was spent in the Genz subroutines. In case of the 'Coffee' problems, performing 2 GSR steps per iteration resulted in slightly less calls to the Genz subroutine than 1 GSR step did. Interestingly, the 'CashMatching' problem was solved significantly faster when performing a single GSR step per iteration, instead of two steps. All these results indicate that approximate solution of the column generation problems is sufficient.

The  $\hat{z}$  vectors returned by the oracle always fell into a relatively small box, thereby remaining in the safe domains where the respective objective functions are well-conditioned.

Problem	prescribed probability level	1 GSR steps per iter			2 GSR steps per iter		
		Genz	itNum	p	Genz	itNum	p
Coffee 9	0.8	104	5	0.7997	136	6	0.7998
	0.85	98	5	0.8502	110	5	0.8502
	0.9	80	4	0.9000	109	5	0.9001
	0.95	115	6	0.9506	155	7	0.9506
	0.96	94	5	0.9604	132	6	0.9604
	0.97	115	7	0.9705	101	5	0.9706
CashMatching	0.8	634	24	0.7957	783	29	0.7982
	0.85	873	35	0.8483	1078	40	0.8480
	0.9	581	24	0.8981	725	28	0.8982
	0.95	330	13	0.9462	441	17	0.9470
	0.98	159	6	0.9755	324	13	0.9767
	0.99	213	8	0.9863	353	14	0.9865

Figure 5

Computational results for problems 'Coffee9' and 'CashMatching'

## 6 Conclusions

The proposed probability-maximization approach is based on a polyhedral approximation of the epigraph of the probabilistic function. Finding a new approximation point in the present scheme is easier than finding a  $p$ -efficient point in the classic scheme of Dentcheva, Prékopa and Ruszczyński [12]. In the present scheme, an approximation point is found by unconstrained optimization. In LP terms, this is a column generation scheme where new columns are found by maximizing reduced cost.

The inner approximating model of the epigraph is immune to noise in gradient computation, in the following sense. Suppose that at iteration  $k$ , the next iterate  $z_{k+1}$  is just a rough approximate solution of the relevant subproblem (reduced cost-maximization). As long as  $\phi(z_{k+1})$  is computed with reasonable accuracy, the model remains a true inner approximation.

Our computational experiments indicate that rough approximate solution of the subproblems is sufficient for convergence. We also provide theoretical explanation of this observation. – A randomized version of the present algorithm is proposed with convergence proof in [15].

### Acknowledgement

This research is supported by EFOP-3.6.1-16-2016-00006 "The development and enhancement of the research potential at Pallas Athena University" project. The Project is supported by the Hungarian Government and co-financed by the European Social Fund.

## References

- [1] T. Arnold, R. Henrion, A. Möller, and S. Vigerske. A mixed-integer stochastic nonlinear optimization problem with joint probabilistic constraints. *Pacific Journal of Optimization*, 10:5–20, 2014.
- [2] M. Bertocchi, G. Consigli, and M.A.H. Dempster (eds). *Stochastic Optimization Methods in Finance and Energy: New Financial Products and Energy Market Strategies*. International Series in Operations Research and Management Science. Springer, 2012.
- [3] C. Borell. Convex set functions in  $d$ -space. *Periodica Mathematica Hungarica*, 6:111–136, 1975.
- [4] H.J. Brascamp and E.H. Lieb. On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log-concave functions and with an application to the diffusion equations. *Journal of Functional Analysis*, 22:366–389, 1976.
- [5] G. C. Calafiore and M. C. Campi. The scenario approach to robust control design. *IEEE Trans. Automat. Control*, 51:742–753, 2006.
- [6] A. Charnes and W. Cooper. Deterministic equivalents for optimizing and satisficing under chance constraints. *Operations Research*, 11:18–39, 1963.
- [7] A. Charnes, W.W. Cooper, and G.H. Symonds. Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4:235–263, 1958.
- [8] D. Dentcheva. Optimization models with probabilistic constraints. In G. Calafiore and F. Dabbene, editors, *Probabilistic and Randomized Methods for Design under Uncertainty*, pages 49–97. Springer, 1st edition, 2006.
- [9] D. Dentcheva. *Optimisation Models with Probabilistic Constraints. Chapter 4 in [59]*. MPS-SIAM series on optimization. SIAM and MPS, Philadelphia, 2009.
- [10] D. Dentcheva, B. Lai, and A. Ruszczyński. Dual methods for probabilistic optimization problems. *Mathematical Methods of Operations Research*, 60:331–346, 2004.
- [11] D. Dentcheva and G. Martinez. Regularization methods for optimization problems with probabilistic constraints. *Mathematical Programming*, 138:223–251, 2013.
- [12] D. Dentcheva, A. Prékopa, and A. Ruszczyński. Concavity and efficient points of discrete distributions in probabilistic programming. *Mathematical Programming*, 89:55–77, 2000.
- [13] J. Elzinga and T.G. Moore. A central cutting plane method for the convex programming problem. *Mathematical Programming*, 8:134–145, 1975.

- [14] Y.M. Ermoliev, T.Y. Ermolieva, G.J. Macdonald, and V.I. Norkin. Stochastic optimization of insurance portfolios for managing exposure to catastrophic risk. *Annals of Operations Research*, 99:207–225, 2000.
- [15] C.I. Fábián and T. Szántai. A randomized method for smooth convex minimization, motivated by probability maximization. *Optimization Online*, 2017. (Posted at the Stochastic Programming area, in March).
- [16] C.A. Floudas and P.M. Pardalos (Eds). *Encyclopedia of Optimization*. Springer - Verlag, 2nd edition, 2009.
- [17] A. Genz. Numerical computation of multivariate normal probabilities. *J. Comp. Graph Stat.*, 1:141–149, 1992.
- [18] A. Genz. Numerical computation of multivariate normal probabilities. *Journal of Computational and Graphical Statistics*, 1:141–150, 1992.
- [19] A. Genz and F. Bretz. *Computation of multivariate normal and t probabilities*. Number 195 in Lecture Notes in Statistics. Springer, Dordrecht, 2009.
- [20] R. Henrion. Introduction to chance constraint programming. Technical report, Weierstrass-Institut für Angewandte Analysis und Stochastik, 2004. [www.wias-berlin.de/people/henrion/ccp.ps](http://www.wias-berlin.de/people/henrion/ccp.ps).
- [21] R. Henrion and A. Möller. Optimization of a continuous distillation process under random inflow rate. *Computer & Mathematics with Applications*, 45:247–262, 2003.
- [22] R. Henrion and A. Möller. A gradient formula for linear chance constraints under Gaussian distribution. *Mathematics of Operations Research*, 37:475–488, 2012.
- [23] R. Henrion and C. Strugarek. Convexity of chance constraints with independent random variables. *Computational Optimization and Applications*, 41:263–276, 2008.
- [24] R. Henrion and C. Strugarek. *Convexity of Chance Constraints with Dependent Random Variables: the use of Copulae*. (Chapter 17 in [2]). Springer New York, 2011.
- [25] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms*. Springer-Verlag, 1993.
- [26] P. Kall and J. Mayer. *Stochastic Linear Programming: Models, Theory, and Computation*. Springer’s International Series in Operations Research and Management Science. Kluwer Academic Publishers, 2005.
- [27] É. Komáromi. A dual method for probabilistic constrained problems. *Mathematical Programming Study*, 28:94–112, 1986. (Special issue on stochastic programming, A. Prékopa and R.J.-B. Wets, editors).
- [28] É. Komáromi. On properties of the probabilistic constrained linear programming problem and its dual. *Journal of Optimization Theory and Applications*, 55:377–390, 1987.

- [29] M. A. Lejeune. Pattern-based modeling and solution of probabilistically constrained optimization problems. *Operations Research*, 60(6):13561372, 2012.
- [30] M. A. Lejeune and F. Margot. Solving chance-constrained optimization problems with stochastic quadratic inequalities. *Operations Research*, page 139, 2016.
- [31] X. Liu, S. Küçükyavuz, and J. Luedtke. Decomposition algorithms for two-stage chance-constrained programs. *Mathematical Programming*, 157(1):219–243, 2016.
- [32] J. Luedtke and S. Ahmed. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19:674–699, 2008.
- [33] J. Luedtke, S. Ahmed, and G.L. Nemhauser. An integer programming approach for linear programs with probabilistic constraints. *Mathematical Programming*, 122(2):247–272, 2010.
- [34] D.G. Luenberger and Y. Ye. *Linear and Nonlinear Programming*. International Series in Operations Research and Management Science. Springer, 2008.
- [35] K. Marti. Differentiation of probability functions : The transformation method. *Computers and Mathematics with Applications*, 30:361–382, 1995.
- [36] K. Marti. Differentiation of probability functions : The transformation method. *Math. Programming*, 75(2):201–220, 1996.
- [37] J. Mayer. *Stochastic Linear Programming Algorithms: A Comparison Based on a Model Management System*. Gordon and Breach Science Publishers, 1998.
- [38] J. Mayer. *On the Numerical solution of jointly chance constrained problems. Chapter 12 in [68]*. Springer, 1st edition, 2000.
- [39] B.L. Miller and H.M. Wagner. Chance constrained programming with joint constraints. *Operations Research*, 13:930–945, 1965.
- [40] M. Minoux. *Programmation Mathématique: Théorie et Algorithmes*. Tec & Doc Lavoisier, 2nd edition, 2007.
- [41] D.R. Morgan, J.W. Eheart, and A.J. Valocchi. Aquifer remediation design under uncertainty using a new chance constraint programming technique. *Water Resources Research*, 29:551–561, 1993.
- [42] A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs. *SIAM Journal of Optimization*, 17(4):969–996, 2006.
- [43] B. Pagnoncelli, S. Ahmed, and A. Shapiro. Sample average approximation method for chance constrained programming: Theory and applications. *J. Optim. Theory Appl*, 142:399–416, 2009.

- [44] A. Prékopa. On probabilistic constrained programming. In H.W. Kuhn, editor, *Proceedings of the Princeton Symposium on Mathematical Programming*, pages 113–138. Princeton University Press, Princeton, New Jersey, 1970.
- [45] A. Prékopa. Logarithmic concave measures with applications to stochastic programming. *Acta Scientiarum Mathematicarum (Szeged)*, 32:301–316, 1971.
- [46] A. Prékopa. Contributions to the theory of stochastic programming. *Mathematical Programming*, 4:202–221, 1973.
- [47] A. Prékopa. On logarithmic concave measures and functions. *Acta Scientiarum Mathematicarum (Szeged)*, 34:335–343, 1973.
- [48] A. Prékopa. Dual method for a one-stage stochastic programming problem with random rhs obeying a discrete probability distribution. *Z. Operations Research*, 34:441–461, 1990.
- [49] A. Prékopa. *Stochastic Programming*. Kluwer Academic Publishers, Dordrecht, 1995.
- [50] A. Prékopa. *Probabilistic programming*. In [58] (Chapter 5). Elsevier, Amsterdam, 2003.
- [51] A. Prékopa. On the relationship between probabilistic constrained, disjunctive and multiobjective programming. Technical Report 7-2007, Rutgers Center for Operations Research, Rutgers University, Piscataway, NJ, 2007. (RUTCOR Research Report).
- [52] A. Prékopa, S. Ganczer, I. Deák, and K. Patyi. The STABIL stochastic programming model and its experimental application to the electrical energy sector of the Hungarian economy. In M.A.H. Dempster, editor, *Stochastic Programming*, pages 369–385. Academic Press, London, 1980.
- [53] A. Prékopa and T. Szántai. Flood control reservoir system design using stochastic programming. *Math. Programming Study*, 9:138–151, 1978.
- [54] A. Prékopa and T. Szántai. A new multivariate gamma distribution and its fitting to empirical streamflow data. *Water Resources Research*, 14:19–24, 1978.
- [55] A. Prékopa and T. Szántai. On optimal regulation of a storage level with application to the water level regulation of a lake. *European Journal of Operations Research*, 3:175–189, 1979.
- [56] A. Prékopa, B. Vizvári, and T. Badics. Programming under probabilistic constraint with discrete random variable. In F. Giannesi, T. Rapcsák, and S. Komlósi, editors, *New Trends in Mathematical Programming*, pages 235–255. Kluwer, Dordrecht, 1998.
- [57] A. Ruszczyński. *Nonlinear Optimization*. Princeton University Press, 2006.

- [58] A. Ruszczyński and A. Shapiro. *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*. Elsevier, Amsterdam, 2003.
- [59] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming. Modeling and Theory*, volume 9 of *MPS-SIAM series on optimization*. SIAM and MPS, Philadelphia, 2009.
- [60] T. Szántai. *Numerical evaluation of probabilities concerning multi-dimensional probability distributions*. PhD thesis, Hungarian Academy of Sciences, 1985.
- [61] T. Szántai. A computer code for solution of probabilistic-constrained stochastic programming problems. In Y.M. Ermoliev and R.J.-B. Wets, editors, *Numerical Techniques for Stochastic Optimization*, pages 229–235. Springer-Verlag, Berlin, 1988.
- [62] T. Szántai. Probabilistic constrained programming and distributions with given marginals. In J. Stepan V. Benes, editor, *Distributions with Given Marginals and Moment Problems*, pages 205–210, 1997.
- [63] E. Tamm. On  $g$ -concave functions and probability measures (russian). *Eesti NSV Teaduste Akademia Toimetised, Füüsika-Matemaatika*, 28:17–24, 1977.
- [64] D.M. Topkis and A.F. Veinott. On the convergence of some feasible direction algorithms for nonlinear programming. *SIAM Journal on Control*, 5(2):268–279, 1967.
- [65] S. Uryas’ev. Derivatives of probability functions and integrals over sets given by inequalities. *Journal of Computational and Applied Mathematics*, 56(1-2):197–223, 1994.
- [66] S. Uryas’ev. Derivatives of probability functions and some applications. *Annals of Operations Research*, 56:287–311, 1995.
- [67] S. Uryas’ev. *Derivatives of probability and Integral functions: General Theory and Examples. Appearing in [16]*. Springer - Verlag, 2nd edition, 2009.
- [68] S. Uryas’ev (ed). *Probabilistic Constrained Optimization: Methodology and Applications*. Kluwer Academic Publishers, 2000.
- [69] W. van Ackooij. Decomposition approaches for block-structured chance-constrained programs with application to hydro-thermal unit commitment. *Mathematical Methods of Operations Research*, 80:227253, 2014.
- [70] W. van Ackooij. Eventual convexity of chance constrained feasible sets. *Optimization (A Journal of Math. Programming and Operations Research)*, 64:1263–1284, 2015.
- [71] W. van Ackooij, V. Berge, W. de Oliveira, and C. Sagastizábal. Probabilistic optimization via approximate  $p$ -efficient points and bundle methods. *Computers & Operations Research*, 77:177–193, 2017.

- [72] W. van Ackooij and W. de Oliveira. Level bundle methods for constrained convex optimization with various oracles. *Computation Optimization and Applications*, 57(3):555–597, 2014.
- [73] W. van Ackooij and R. Henrion. (sub-)gradient formulae for probability functions of random inequality systems under gaussian distribution. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):63–87, 2017.
- [74] W. van Ackooij, R. Henrion, A. Möller, and R. Zorgati. On probabilistic constraints induced by rectangular sets and multivariate normal distributions. *Mathematical Methods of Operations Research*, 71(3):535–549, 2010.
- [75] W. van Ackooij, R. Henrion, A. Möller, and R. Zorgati. On joint probabilistic constraints with Gaussian Coefficient Matrix. *Operations Research Letters*, 39:99–102, 2011.
- [76] W. van Ackooij, R. Henrion, A. Möller, and R. Zorgati. Joint chance constrained programming for hydro reservoir management. *Optimization and Engineering*, 15:509–531, 2014.
- [77] W. van Ackooij and C. Sagastizábal. Constrained bundle methods for upper inexact oracles with application to joint chance constrained energy problems. *Siam Journal on Optimization*, 24(2):733–765, 2014.
- [78] C. van de Panne and W. Popp. Minimum-cost cattle feed under probabilistic protein constraints. *Management Science*, 9:405–430, 1963.
- [79] A.F. Veinott. The supporting hyperplane method for unimodal programming. *Operations Research*, 15:147–152, 1967.
- [80] B. Vízvári. The integer programming background of a stochastic integer programming algorithm of Dentcheva-Prékopa-Ruszczynski. *Optimization Methods and Software*, 17:543–559, 2002.
- [81] G. Zoutendijk. *Methods of Feasible Directions: A Study in Linear and Non-Linear Programming*. Elsevier Publishing Co., Amsterdam, 1960.