On Stochastic Integer Programming under Probabilistic Constraints

Darinka Dentcheva[1] András Prékopa[2]
Andrzej Ruszczyński[2]

[1] *Humboldt University Berlin, Institute of Mathematics*
10099 Berlin, Germany
[2] *RUTCOR, Rutgers Center for Operations Research*
Rutgers University, Piscataway, NJ 08854, U.S.A.

Abstract

We consider stochastic programming problems with probabilistic constraints involving integer-valued random variables. The concept of p-efficient points of a probability distribution is used to derive various equivalent problem formulations. Next we modify the concept of r-concave discrete probability distributions and analyse its relevance for problems under consideration. These notions are used to derive new lower and upper bounds for the optimal value of probabilistically constrained stochastic programming problems with integer random variables. We also show how limited information about the distribution can be used to construct such bounds.

1 Introduction

Probabilistic constraints remain one of main challenges of modern stochastic programming. Their motivation is clear: if in the linear program

$$\begin{array}{rl} \min & c^T x \\ \text{subject to} & Tx \geq \xi, \\ & Ax \geq b, \\ & x \geq 0, \end{array}$$

the vector ξ is random, we require that $Tx \geq \xi$ shall hold at least with some prescribed probability $p \in (0, 1)$, rather than *for all* possible realizations of the right hand side. This leads to the following problem formulation:

$$\begin{array}{rl} \min & c^T x \\ \text{subject to} & \mathbb{P}(Tx \geq \xi) \geq p, \\ & Ax \geq b, \\ & x \geq 0, \end{array} \tag{1.1}$$

where the symbol $I\!P$ denotes probability.

Programming under probabilistic constraints was initiated by Charnes, Cooper and Symonds in [4]. They formulated probabilistic constraints individually for each stochastic constraint. Joint probabilistic constraints for independent random variables were used first by Miller and Wagner in [9]. The general case was introduced and first studied by the second author of the present paper in [13, 16].

Much is known about problem (1.1) in the case where ξ has a continuous probability distribution (see [20] and the references therein). However, only a few papers handle the case of a discrete distribution. In [18] a dual type algorithm for solving problem (1.1) has been proposed. Bounds for the optimal value of this problem, based on disjunctive programming, were analyzed in [26]. The case when the matrix T is random, while ξ is not, has been considered in [29]. Recently, in [22], a cutting plane method for solving (1.1) has been presented.

Even though the literature for handling probabilistic constraints with discrete random variables is scarce, the number of potential applications is large. Singh at al.in [27] consider a microelectronic wafer design problem that arises in semiconductor manufacturing. The problem was to maximize the probability rather than to optimize an objective function subject to a probabilistic constraint, but other formulations are possible as well. Another application area are communication and transportation network capacity expansion problems, where arc and node capacities are restricted to be integers [12, 21]. Bond portfolio problems with random integer-valued liabilities can be formalized as (1.1) see [5]. Many production planning problems involving random indivisible demands fit to our general setting as well.

Although we concentrate on integer random variables, all our results easily extend to other discrete distributions with non-uniform grids, under the condition that a uniform lower bound on the distance of grid points can be found.

To fix some notation we assume that in the problems above A is an $m \times n$ matrix, T is an $s \times n$ matrix; $c, x \in I\!R^n$, $b \in I\!R^m$ and ξ is a random vector with values in $I\!R^s$. We use $\mathbb{Z}$ and $\mathbb{Z}_+$ to denote the set of integers and nonnegative integers, respectively. The inequality '$\geq$' for vectors is always understood coordinate-wise.

2 p-Efficient Points

Let us define the sets:

$$\mathcal{D} = \{x \in \mathbb{R}^n : Ax \geq b, \quad x \geq 0\} \tag{2.1}$$

and

$$\mathcal{Z}_p = \{y \in I\!R^s : I\!P(\xi \leq y) \geq p\}. \tag{2.2}$$

Clearly, problem (1.1) can be compactly rewritten as

$$\begin{array}{rl} \min & c^T x \\ \text{subject to} & Tx \in \mathcal{Z}_p, \\ & x \in \mathcal{D}. \end{array} \tag{2.3}$$

While the set $\mathcal{D}$ is a convex polyhedron, the structure of $\mathcal{Z}_p$ needs to be analysed in more detail.

Let F denote the probability distribution function of ξ, and F_i be the marginal probability distribution function of the ith component ξ_i. By assumption, the set $\mathcal{Z}$ of all possible values of the random vector ξ is included in $\mathbb{Z}^s$.

We shall use the concept of p-efficient points, introduced in [18].

Definition 2.1. *Let $p \in [0,1]$. A point $v \in \mathbb{R}^s$ is called a p-*efficient *point of the probability distribution function F, if $F(v) \geq p$ and there is no $y \leq v$, $y \neq v$ such that $F(y) \geq p$.*

Obviously, for a scalar random variable ξ and for every $p \in (0,1)$ there is exactly one p-efficient point: the smallest v such that $F(v) \geq p$. This leads to the following observation.

Remark 2.2. *Let $p \in (0,1)$ and let l_i be the p-efficient point of ξ_i, $i = 1, \dots, s$. Then every $v \in \mathbb{R}^s$ such that $F(v) \geq p$ must satisfy the inequality $v \geq l = (l_1, \dots, l_s)$.*

Proof. For a p-efficient point v we have

$$p \leq F(v) = \mathbb{P}\{\xi \leq v\} \leq \mathbb{P}\{\xi_i \leq v_i\} = F_i(v_i),$$

and, by the definition of l_i, we must have $v_i \geq l_i$. □

Since rounding down to the nearest integer does not change the value of the distribution function, p-efficient points of an integer random vector must be integer. We can thus use Remark 2.2 to get the following interesting fact (noticed earlier in [28] for non-negative integer random variables).

Theorem 2.3. *For each $p \in (0,1)$ the set of p-efficient points of an integer random variable is nonempty and finite.*

Proof. We shall at first show that p-efficient points exist. Since $p < 1$, there must exist y such that $F(y) \geq p$. By Remark 2.2, all v such that $F(v) \geq p$ must satisfy $v \geq l$. Therefore, if y is not p-efficient, one of finitely many integer points v such that $l \leq v \leq y$ must be p-efficient.

We shall now prove the finiteteness of the set of p-efficient points. Suppose that there exisits an infinite sequence of different p-efficient points v^j, $j = 1, 2, \dots$. Since they are integer, and the first coordinate v_1^j is bounded below by l_1, with no loss of generality we may select a subsequence which is non-decreasing in the first coordinate. By a similar token, we can select further subsequences which are non-decreasing in the first k coordinates ($k = 1, \dots, s$). Since the dimension s is finite, we obtain a subsequence of different p-efficient points which is non-decreasing in all coordinates. This contradicts the definition of a p-efficient point. □

Let $p \in (0,1)$ and let v^j, $j \in J$, be *all* p-efficient points of ξ. By Theorem 2.3, J is a finite index set. Let us define the cones

$$K_j = v^j + \mathbb{R}^s_+, \quad j \in J.$$

Remark 2.4. $\mathcal{Z}_p = \bigcup_{j\in J} K_j$.

Proof. If $y \in \mathcal{Z}_p$ then either y is p-efficient or there exists an integer $v \le y$, $v \ne y$, $v \in \mathcal{Z}_p$. By Remark 2.2, one must have $l \le v$. Since there are only finitely many integer points $l \le v \le y$ one of them, v_j, must be p-efficient, and so $y \in K_j$. □

Thus, we obtain (for $0 < p < 1$) the following *disjunctive* formulation of (2.3):

$$\begin{array}{rl} \min & c^T x \\ \text{subject to} & Tx \in \bigcup_{j\in J} K_j, \\ & x \in \mathcal{D}. \end{array} \tag{2.4}$$

Its main advantage is an insight into the nature of the non-convexity of the feasible set. In particular, we can formulate the following necessary and sufficient condition for the existence of an optimal solution of (2.4).

Assumption 2.5. *The set* $\Lambda := \{(u,w) \in \mathbb{R}_+^{m+s} \mid A^T w + T^T u \le c\}$ *is nonempty.*

Theorem 2.6. *Assume that the feasible set of (2.4) is nonempty. Then (2.4) has an optimal solution if and only if Assumption 2.5 holds.*

Proof. If (2.4) has an optimal solution, then for some $j \in J$ the linear program

$$\begin{array}{rl} \min & c^T x \\ \text{subject to} & Tx \ge v^j, \\ & A \ge b, \\ & x \ge 0, \end{array} \tag{2.5}$$

has an optimal solution. By duality in linear programming, its dual

$$\begin{array}{rl} \max & (v^j)^T u + b^T w \\ \text{subject to} & T^T u + A^T w \le c, \\ & u, w \ge 0, \end{array} \tag{2.6}$$

has an optimal solution and the optimal values of both programs are equal. Thus, Assumption 2.5 must hold. On the other hand, if Assumption 2.5 is satisfied, all dual programs (2.6) for $j \in J$ have nonempty feasible sets, so the objective values of all primal problems (2.5) are bounded from below. Since one of them has a nonempty feasible set by assumption, an optimal solution must exist. □

3 r-Concave Discrete Distribution Functions

Since the set $\mathcal{Z}_p$ need not be convex, it is essential to analyse its properties and to find equivalent formulations with more convenient structures. To this end we shall recall and adapt the notion of r-concavity of a distribution function. It uses the *generalized mean function* $m_r : \mathbb{R}_+ \times \mathbb{R}_+ \times [0,1] \to \mathbb{R}$ defined as follows:

$$m_r(a,b,\lambda) = \begin{cases} a^\lambda b^{1-\lambda} & \text{if } r = 0, \\ \max\{a,b\} & \text{if } r = \infty, \\ \min\{a,b\} & \text{if } r = -\infty, \\ 0 & \text{if } r \in (-\infty, 0), ab = 0, \\ (\lambda a^r + (1-\lambda) b^r)^{1/r} & \text{otherwise.} \end{cases} \tag{3.1}$$

Definition 3.1. *A distribution function $F : \mathbb{R}^n \to [0,1]$ is called r-concave, where $r \in [-\infty, \infty]$, if*

$$F(\lambda x + (1-\lambda)y) \geq m_r(F(x), F(y), \lambda)$$

for all $x, y \in \mathbb{R}^s$ and all $\lambda \in [0,1]$,

If $r = -\infty$ we call F *quasi-concave*, for $r = 0$ it is known as *log-concave*, and for $r = 1$ the function F is concave in the usual sense.

Another general concept of r-concavity can be introduced for measures, by considering probabilities of Minkowski sums of sets. In this paper, however, we shall only consider r-concave disribution functions.

The concept of a log-concave probabiltiy measure (the case $r = 0$) was intruduced and studied in [14, 15]. The notion of r-concavity and corresponding results were given in [2, 3]. For detailed description and proofs, see [20].

By monotonicity, r-concavity of a distribution function is equivalent to the inequality

$$F(z) \geq m_r(F(x), F(y), \lambda)$$

for all $z \geq \lambda x + (1-\lambda)y$.

Clearly, distribution functions of integer random variables are not continuous, and cannot be r-concave in the sense of the above definition. Therefore, we relax Definition 3.1 in the following way.

Definition 3.2. *A distribution function F is called* r-concave on the set $\mathcal{A}$ *with $r \in [-\infty, \infty]$, if*

$$F(z) \geq m_r(F(x), F(y), \lambda)$$

for all $z, x, y \in \mathcal{A}$ and $\lambda \in [0,1]$ such that $z \geq \lambda x + (1-\lambda)y$.

The concept of r-concavity on a set can be used to find an equivalent representation of the set $\mathcal{Z}_p$ given by (2.2).

Theorem 3.3. *Let $\mathcal{Z}$ be the set of all values of an integer random vector ξ. If the distribution function F of ξ is r-concave on $\mathcal{Z} + \mathbb{Z}^s_+$ for some $r \in [-\infty, \infty]$, then for every $p \in (0,1)$ one has*

$$\mathcal{Z}_p = \{y \in \mathbb{R}^s : \, y \geq z \geq \sum_{j \in J} \lambda_j v^j, \, \sum_{j \in J} \lambda_j = 1, \, \lambda_j \geq 0, \, z \in \mathbb{Z}^s\},$$

where v^j, $j \in J$, are the p-efficient points of F.

Proof. By the monotonicity of F we have $F(y) \geq F(z)$ if $y \geq z$. It is, therefore, sufficient to show that $\mathbb{P}(\xi \leq z) \geq p$ for all $z \in \mathbb{Z}^s$ such that $z \geq \sum_{j\in J} \lambda_j v^j$ with $\lambda_j \geq 0$, $\sum_{j\in J} \lambda_j = 1$. We consider four cases with respect to r.

Case 1: $r = \infty$. It follows from the definition of r-concavity that $F(z) \geq \max\{F(v^j), j \in J : \lambda_j \neq 0\} \geq p$.

Case 2: $r = -\infty$. Since $F(v^j) \geq p$ for each index $j \in J$ such that $\lambda_j \neq 0$, the assertion follows as in Case 1.

Case 3: $r = 0$. By the definition of r-concavity,

$$F(z) \geq \prod_{j\in J} [F(v^j)]^{\lambda_j} \geq \prod_{j\in J} p^{\lambda_j} = p.$$

Case 4: $r \in (-\infty, 0)$. By the definition of r-concavity,

$$[F(z)]^r \leq \sum_{j\in J} \lambda_j [F(v^j)]^r \leq \sum_{j\in J} \lambda_j p^r = p^r.$$

Since $r < 0$, we obtain $F(z) \geq p$.

Case 5: $r \in (0, \infty)$. By the definition of r-concavity,

$$[F(z)]^r \geq \sum_{j\in J} \lambda_j [F(v^j)]^r \geq \sum_{j\in J} \lambda_j p^r = p^r.$$

□

Under the conditions of Theorem 3.3, problem (2.4) can be formulated in the following equivalent way:

$$
\begin{aligned}
\min \quad & c^T x && (3.2)\\
\text{subject to} \quad & x \in \mathcal{D} && (3.3)\\
& Tx \geq z, && (3.4)\\
& z \in \mathbb{Z}^s, && (3.5)\\
& z \geq \sum_{j\in J} \lambda_j v^j && (3.6)\\
& \sum_{j\in J} \lambda_j = 1 && (3.7)\\
& \lambda_j \geq 0,\ j \in J. && (3.8)
\end{aligned}
$$

So, the probabilistic constraint has been replaced by linear equations and inequalites, together with the integrality requirement (3.5). This condition cannot be dropped, in general. However, if other conditions of the problem imply that Tx is integer (for example, we have an additional constraint in the definition of $\mathcal{D}$ that $x \in \mathbb{Z}^n$, and T has integer entries), we may dispose of z totally, and replace constraints (3.4)–(3.6) with

$$Tx \geq \sum_{j\in J} \lambda_j v^j.$$

The difficulty comes from the implicitly given p-efficient points $v_j, j \in J$. Our objective will be to avoid their enumeration and to develop an approach that generates them only when needed.

We end this section with sufficient conditions for the r-concavity of the joint distribution function in the case of integer-valued independent components. Our assertion, presented in the next proposition is the discrete version of an observation from [11].

Proposition 3.4. *Assume that the components ξ_i of ξ, $i = 1, \dots, s$, are independent, and that the marginal distribution functions F_i are r_i-concave on sets $\mathcal{A}_i \subset \mathbb{Z}$.*

(i) If $r_i > 0$, $i = 1, \dots, s$, then F is r-concave on $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_s$ with $r = [\sum_{i=1}^{s} r_i^{-1}]^{-1}$.

(ii) If $r_i = 0$, $i = 1, \dots, s$, then F is log-concave on $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_s$.

Proof. Assertion (i) is a simple consequence of Hölder's inequality. Assertion (ii) is obvious. $\square$

4 Lagrangian Relaxation

Let us split variables in problem (2.3):

$$\begin{aligned} \min \quad & c^T x \\ & Tx = z, \\ & x \in \mathcal{D}, \\ & z \in \mathcal{Z}_p. \end{aligned} \tag{4.1}$$

Associating Lagrange multipliers $u \in \mathbb{R}^s$ with constraints (4.1) we obtain the Lagrangian function:

$$L(x, z, u) = c^T x + u^T (z - Tx).$$

Owing to the structure of $\mathcal{Z}_p$ (Remark 2.2), we could have replaced equality $Tx = z$ in (4.1) by an inequality $Tx \geq z$, and use $u \geq 0$ in the Lagrangian. However, formal splitting (4.1) leads to the same conclusion. The dual functional has the form

$$\Psi(u) = \inf_{(x,z) \in \mathcal{D} \times \mathcal{Z}_p} L(x, z, u) = h(u) + d(u),$$

where

$$h(u) = \inf\{(c - T^T u)^T x \mid x \in \mathcal{D}\}, \tag{4.2}$$

$$d(u) = \inf\{u^T z \mid z \in \mathcal{Z}_p\}. \tag{4.3}$$

Lemma 4.1. $\operatorname{dom} \Psi = \{u \in \mathbb{R}^s_+ :$ *there exists* $w \in \mathbb{R}^m_+$ *such that* $A^T w + T^T u \leq c\}$.

Proof. Clearly, $\operatorname{dom}\Psi = \operatorname{dom} h \cap \operatorname{dom} d$. Let us calculate $\operatorname{dom} h$. The recession cone of $\mathcal{D}$,

$$C = \{y \in \mathbb{R}^n : Ay \geq 0,\ y \geq 0\},$$

has the dual cone

$$C^* = \{v \in \mathbb{R}^m : v^T y \geq 0 \text{ for all } y \in C\} = \{v \in \mathbb{R}^m : v \geq A^T w,\ w \geq 0\},$$

as follows from Farkas lemma. Thus

$$\operatorname{dom} h = \{u \in \mathbb{R}^s : c - T^T u \in C^*\} = \{u \in \mathbb{R}^s : T^T u + A^T w \leq c,\ w \geq 0\}.$$

On the other hand, by Lemma 2.2, $\operatorname{dom} d = \mathbb{R}^s_+$, and the result follows. □

It follows that Assumption 2.5, which is necessary and sufficient for the existence of solutions, is also necessary and sufficient for the nonemptiness of the domain of the dual functional.

For any $u \in \mathbb{R}^s_+$ the value of $\Psi(u)$ is a lower bound on the optimal value F^* of the original problem. Consequently, the best lower bound will be given by

$$D^* = \sup \Psi(u). \tag{4.4}$$

If an optimal solution of (2.3) exists, then Assumption 2.5 holds, so, by Lemma 4.1,

$$-\infty < D^* \leq F^*.$$

We shall show that the supremum D^* is attained. Indeed, $h(u) = -\delta^*_{\mathcal{D}}(-c + T^T u)$, where $\delta^*_{\mathcal{D}}(\cdot)$ is the support function of $\mathcal{D}$. Thus $h(\cdot)$ is concave and polyhedral (see [23], Corollary 19.2.1). By Remark 2.4, for $u \geq 0$ the minimization in (4.3) may be restricted to finitely many p-efficient points v^j, $j \in J$. For $u \ngeq 0$ one has $d(u) = -\infty$. Therefore, $d(\cdot)$ is concave and polyhedral as well. Consequently, $\Psi(\cdot)$ is concave and polyhedral. Since it is bounded from above by F^*, it must attain its maximum.

Another lower bound may be obtained from the continuous relaxation of problem (2.3)

$$F^*_{\text{co}} = \min\{c^T x \mid Tx = z,\ x \in \mathcal{D},\ z \in \operatorname{co} \mathcal{Z}_p\}. \tag{4.5}$$

It is known (see [10]) that

$$F^*_{\text{co}} = D^* \leq F^*.$$

We now analyse in more detail the structure of the dual functional Ψ. Let us start from $h(\cdot)$.

Fact 4.2. *Let Condition 2.5 be fulfilled. Then for each* $u \in \mathbb{R}^s$

$$h(u) = \sup\{b^T w \mid T^T u + A^T w \leq c,\ w \geq 0\}.$$

Proof. The result follows form the duality theory in linear programming. $\square$

This allows us to reformulate the dual problem (4.4) in a more explicit way:

$$\max \quad d(u) + b^T w \tag{4.6}$$
$$T^T u + A^T w \le c, \tag{4.7}$$
$$u \ge 0, \; w \ge 0. \tag{4.8}$$

Let us observe that we may write 'max' instead of 'sup' because we already know that the supremum is attained. We may also add the constraint '$u \ge 0$' explicitly, since it defines the domain of d.

Properties of $d(\cdot)$ can also be analysed in a more explicit way.

Lemma 4.3. *For every $u \ge 0$ the solution set of the subproblem*

$$\min_{z \in \mathcal{Z}_p} u^T z \tag{4.9}$$

is nonempty and has the following form:

$$\hat{Z}(u) = \bigcup_{j \in \hat{J}(u)} \{v^j\} + C(u),$$

where $\hat{J}(u)$ is the set of p-efficient solutions of (4.9), and

$$C(u) = \{d \in \mathbb{R}^s_+ : d_i = 0 \text{ if } u_i > 0, \; i = 1, \ldots, s\}. \tag{4.10}$$

Proof. The result follows from Remark 2.4. Let us at first consider the case $u > 0$. Suppose that a solution z to (4.9) is not a p-efficient point. Then there is a p-efficient $v \in \mathcal{Z}_p$ such that $v \le z$, so $u^T v < u^T z$, a contradiction. Thus, for all $u \ge 0$ all solutions to (4.9) are p-efficient. In the general case $u \ge 0$, if a solution z is not p-efficient, we must have $u^T v = u^T z$ for all p-efficient $v \le z$. This is equivalent to $z \in \{v\} + C(u)$, as required. $\square$

The last result allows us to calculate the subdifferential of d in a closed form.

Lemma 4.4. *For every $u \ge 0$ one has $\partial d(u) = \operatorname{co}\{v^j, \; j \in \hat{J}(u)\} + C(u)$.*

Proof. From (4.3) it follows that

$$d(u) = -\delta^*_{\mathcal{Z}_p}(-u),$$

where $\delta^*_{\mathcal{Z}_p}(\cdot)$ is the support function of $\mathcal{Z}_p$, and, thus, of $\operatorname{co} \mathcal{Z}_p$. This fact follows from the structure of $\mathcal{Z}_p$ (Remark 2.4) by virtue of Corolarry 16.5.1 in [23]. By [23, Thm 23.5], $g \in \partial \delta^*_{\mathcal{Z}_p}(-u)$ if and only if $\delta^*_{\mathcal{Z}_p}(-u) + \delta_{\operatorname{co} \mathcal{Z}_p}(g) = -g^T u$, where $\delta_{\operatorname{co} \mathcal{Z}_p}(\cdot)$ is the indicator function of $\operatorname{co} \mathcal{Z}_p$. It follows that $g \in \operatorname{co} \mathcal{Z}_p$ and $\delta^*_{\mathcal{Z}_p}(-u) = -g^T u$. Thus, g is a convex combination of solutions to (4.9) and the result follows from Lemma 4.3. $\square$

Therefore the following necessary and sufficient optimality conditions for problem (4.6)–(4.8) can be formulated.

Theorem 4.5. *A pair $(u, w) \in \Lambda$ is an optimal solution of (4.6)–(4.8) if and only if there exists a point $x \in \mathbb{R}^n_+$ such that:*

$$Ax \geq b, \quad w^T(Ax - b) = 0, \tag{4.11}$$

and

$$Tx \in \text{co}\,\{v^j : j \in \hat{J}(u)\} + C(u), \tag{4.12}$$

where $\hat{J}(u)$ is the set of p-efficient solutions of (4.9), and $C(u)$ is given by (4.10).

Proof. The vector x plays the role of the Lagrange multiplier associated with the constraint (4.7). Let us decipher the relation

$$\partial\Big(b^T w + d(u) + x^T(c - T^T u - A^T w)\Big) \cap K(u, w) \neq \emptyset,$$

where $K(u, w)$ is the normal cone to $\mathbb{R}^{m+s}_+$ at (u, w). Using the closed-form expression for the subdifferential of d from Lemma 4.4, we obtain:

$$\partial\Big(b^T w + d(u) + x^T(c - T^T u - A^T w)\Big) = \begin{pmatrix} \text{co}\,\{v^j : j \in \hat{J}(u)\} + C(u) - Tx \\ b - Ax \end{pmatrix}.$$

On the other hand:

$$K(u, w) = \{(u^*, w^*) : u^* \leq 0, w^* \leq 0, \langle u^*, u\rangle = 0, \langle w^*, w\rangle = 0\} = \begin{pmatrix} -C(u) \\ -C(w) \end{pmatrix}.$$

Consequently, the condition $\text{co}\,\{v^j : j \in \hat{J}(u)\} + C(u) - Tx \cap -C(u) \neq \emptyset$ implies the existence of elements $v \in \text{co}\,\{v^j : j \in \hat{J}(u)\}$ and $c_1, c_2 \in C(u)$ such that: $v + c_1 - Tx = -c_2$, which is equivalent to the condition (4.12). Furthermore, we obtain that $b - Ax \cap -C(w) \neq \emptyset$. The definition of $C(w)$ implies condition (4.11). □

It follows that the optimal Lagrangian bound is associated with a certain primal solution x which is feasible with respect to the deterministic constraints and such that $Tx \in \text{co}\,\mathcal{Z}_p$. Moreover, since $(u, w) \in \Lambda$, the point x is optimal for the convex hull problem:

$$\min \quad c^T x \tag{4.13}$$

$$Ax \geq b, \tag{4.14}$$

$$Tx \geq \sum_{j \in J} \lambda_j v^j, \tag{4.15}$$

$$\sum_{j \in J} \lambda_j = 1, \tag{4.16}$$

$$x \geq 0, \ \lambda \geq 0. \tag{4.17}$$

Indeed, associating with (4.14) multipliers w, with (4.15) multipliers u, and with (4.16) a multiplier $\mu = d(u)$, we can show that $(x, \bar{\lambda})$ is optimal for (4.13)-(4.17) provided that $\bar{\lambda}_j$ are the coefficients at v^j in the convex combination in (4.12).

Since the set of p-efficient points is not known, we need a numerical method for solving (4.6)–(4.8) or its dual (4.13)–(4.17).

5 Cone generation methods

The idea of a numerical method for calculating Lagrangian bounds is embedded in the convex hull formulation (4.13)–(4.17). We can easily adapt to it the classical column generation scheme known from large scale linear and integer programming [6, 1].

Cone Generation Method

Step 0: Select a p-efficient point v^0. Set $J_0 = \{0\}$, $k = 0$.

Step 1: Solve the *master problem*

$$\min \quad c^T x \tag{5.1}$$

$$Ax \geq b, \tag{5.2}$$

$$Tx \geq \sum_{j \in J_k} \lambda_j v^j, \tag{5.3}$$

$$\sum_{j \in J_k} \lambda_j = 1, \tag{5.4}$$

$$x \geq 0, \ \ \lambda \geq 0. \tag{5.5}$$

Let u^k be the vector of simplex multipliers associated with the constraint (5.3).

Step 2: Calculate an upper bound for the dual functional:

$$\overline{d}(u^k) = \min_{j \in J_k} (u^k)^T v^j.$$

Step 3: Find a p-efficient solution v^{k+1} of the subproblem:

$$\min_{z \in \mathcal{Z}_p} (u^k)^T z$$

and calculate

$$d(u^k) = (v^{k+1})^T u^k.$$

Step 4: If $d(u^k) = \overline{d}(u^k)$ then stop; otherwise set $J_{k+1} = J_k \cup \{k+1\}$, increase k by one and go to Step 1.

Few comments are in order. The first p-efficient point v^0 can be found by solving (4.9) for an arbitrary $u \geq 0$. All master problems will be solvable, if the first one is solvable, i.e., if the set $\{x \in \mathbb{R}^n_+ : Ax \geq b,\ Tx \geq v^0\}$ is nonempty. If not, adding a penalty term $M\mathbb{1}^T t$ to the objective, and replacing (5.3) by

$$Tx + t \geq \sum_{j \in J_k} \lambda_j v^j,$$

with $t \geq 0$ and a very large M, is the usual remedy ($\mathbb{1}^T = [1\ 1\ \ldots\ 1]$). The calculation of the upper bound at Step 2 is easy, because one can simply select $j_k \in J_k$ with $\lambda_{j_k} > 0$ and set $\overline{d}(u^k) = (u^k)^T v^{j_k}$. At Step 3 one may search for p-efficient solutions only, due to Lemma 4.3.

Convergence of the algorithm follows from a standard argument. The set J_k cannot grow indefinitely, because there are finitely many p-efficient points (Theorem 2.3). If the stopping test of Step 4 is satisfied, optimality conditions for (4.13)–(4.17) are satisfied. Moreover $\hat{J}_k = \{j \in J_k : \langle v^j, u^k \rangle = d(u^k)\} \subseteq \hat{J}(u)$.

Our cone generation method shares its drawbacks with other column generation schemes. Initial iterations are inefficient. The number of p-efficient points grows and there is no reliable way for deleting them. For these reasons, especially when the dimension of x is large and the number of rows of T small, an attractive alternative is provided by *bundle methods* applied directly to the dual problem

$$\max_{u \geq 0} \Big[h(u) + d(u)\Big],$$

because at any $u \geq 0$ subgradients of h and d are readily available. For a comprehensive description of bundle methods the reader is refereed to [7, 8]. It may be interesting to note that in our case they correspond to a version of the augmented Lagrangian method (see [24, 25]).

Let us now focus our attention on solving the auxiliary problem (4.9), which is explicitly written as:

$$\min\{u^T z \mid F(z) \geq p\}, \tag{5.6}$$

where $F(\cdot)$ denotes the distribution function of ξ.

Assume that the components ξ_i, $i = 1, \ldots, s$, are independent. Then we can write the probability constraint in the following form:

$$\ln(F(z)) = \sum_{i=1}^{s} \ln(F_i(z_i)) \geq \ln p.$$

Since we know that one of the solutions is a p-efficient point, with no loss of generality we may restrict the search to integer vectors z. Furthermore, by Remark 2.2, we have $z_i \geq l_i$, where l_i are p-efficient points of ξ_i. We obtain the problem:

$$\min \sum_{i=1}^{s} u_i z_i$$

$$\sum_{i=1}^{s} \ln(F_i(z_i)) \geq \ln p,$$

$$z_i \geq l_i, \quad z_i \in \mathbb{Z}, \quad i = 1, \ldots, s.$$

This is a knapsack problem that can be solved by efficient methods, like dynamic programming (for an appropriately discretized approximation) or branch-and-bound schemes [10].

For log-concave marginals, a 0–1 formulation may be convenient. Let $l_i + b_i$ be an upper bound on z_i. Setting

$$z_i = l_i + \sum_{j=1}^{b_i} z_{ij},$$

where $z_{ij} \in \{0, 1\}$ we can reformulate the problem as follows:

$$\min \sum_{i=1}^{s} \sum_{j=1}^{b_i} u_i z_{ij}$$

$$\sum_{i=1}^{s} \sum_{j=1}^{b_i} a_{ij} z_{ij} \geq r,$$

where $a_{ij} = \ln(F_i(l_i + j)) - \ln(F_i(l_i + j - 1))$, and $r = \ln p - \ln F(l)$. Indeed, by the log-concavity, we have $a_{i,j+1} \leq a_{ij}$, so there is always a solution with nonincreasing z_{ij}, $j = 1, \ldots, b_i$.

Clearly, these simplifications are due to the independence of the components of ξ. If they are dependent, bounding techniques from the next section may be employed.

6 Bounds via binomial moments

If the components of ξ are dependent, subproblem (4.3) may be difficult to solve exactly. Still, some bounds on its optimal solution may prove useful. We shall develop a number of bounds using only partial information on the distribution function of ξ in the form of the marginal distributions:

$$F_{i_1 \ldots i_k}(z_{i_1}, \ldots, z_{i_k}) = \mathbb{P}\{\xi_{i_1} \leq z_{i_1}, \ldots \xi_{i_k} \leq z_{i_k}\}, \quad 1 \leq i_1 < \ldots < i_k \leq s.$$

Since for each marginal distribution one has $F_{i_1 \ldots i_k}(z_{i_1}, \ldots, z_{i_k}) \geq F(z)$ the following relaxation of $\mathcal{Z}_p$ (defined by (2.2)) can be obtained.

Fact 6.1. *For each $z \in \mathcal{Z}_p$ and for every $1 \leq i_1 < \ldots < i_k \leq s$ the following inequality must hold:*

$$F_{i_1 \ldots i_k}(z_{i_1}, \ldots, z_{i_k}) \geq p.$$

We shall base further developments on the following result of [17].

Theorem 6.2. *For any distributon function $F : \mathbb{R}^s \to [0,1]$ and any $1 \leq k \leq s$, at every $z \in \mathbb{R}^s$ the optimal value of the following linear programming problem*

$$
\begin{array}{rcl}
 & \max & v_s \\
v_0 + v_1 + v_2 + v_3 + \cdots + v_s & = & 1 \\
v_1 + 2v_2 + 3v_3 + \cdots + rv_r & = & \sum_{1\leq i\leq s} F_i(z_i) \\
v_2 + \binom{3}{2}v_3 + \cdots + \binom{s}{2}v_s & = & \sum_{1\leq i_1<i_2\leq s} F_{i_1 i_2}(z_{i_1}, z_{i_2}) \\
 & \vdots & \\
v_k + \binom{k+1}{k}v_{k+1} + \cdots + \binom{s}{k}v_s & = & \sum_{1\leq i_1<\ldots<i_k\leq s} F_{i_1\ldots i_k}(z_{i_1}, \ldots, z_{i_k}) \\
v_0 \geq 0,\ v_1 \geq 0,\ \ldots,\ v_s \geq 0. & &
\end{array}
\tag{6.1}
$$

provides an upper bound for $F(z_1, \ldots, z_s)$.

We can use this result to bound our auxiliary problem (4.3).

Proposition 6.3. *Let $\xi = (\xi_1, \ldots, \xi_s)$ be an integer random vector and let $F_{i_1,\ldots,i_k}$ denote its marginal distribution functions. Then for every $p \in (0,1)$ and for every $1 \leq k \leq s$ the optimal value of the problem*

$$
\begin{array}{rcl}
 & \min & u^T z \\
v_0 + v_1 + v_2 + v_3 + \cdots + v_s & = & 1 \\
v_1 + 2v_2 + 3v_3 + \cdots + rv_r & = & \sum_{1\leq i\leq s} F_i(z_i) \\
v_2 + \binom{3}{2}v_3 + \cdots + \binom{s}{2}v_s & = & \sum_{1\leq i_1<i_2\leq s} F_{i_1 i_2}(z_{i_1}, z_{i_2}) \\
 & \vdots & \\
v_k + \binom{k+1}{k}v_{k+1} + \cdots + \binom{s}{k}v_s & = & \sum_{1\leq i_1<\ldots<i_k\leq s} F_{i_1\ldots i_k}(z_{i_1}, \ldots, z_{i_k}) \\
v_0 \geq 0,\ v_1 \geq 0,,\ \ldots,\ v_{s-1} \geq 0,\ v_s \geq p, & & z_1 \geq l_1,\ z_2 \geq l_2,\ \ldots,\ z_s \geq l_s, \\
 & z \in \mathbb{Z}^s &
\end{array}
\tag{6.2}
$$

provides a lower bound on the optimal value $d(u)$ given by (4.3).

Proof. If $z \in \mathcal{Z}_p$, that is, $F(z) \geq p$, then the optimal value of (6.1) satisfies $v_s \geq p$. Thus z and the solution v of (6.1) are feasible for (6.2). Since the objective functions of (4.3) and (6.2) are the same, the result follows. □

Problem (6.2) is a nonlinear mixed-integer problem. Its advantage over the original formulation is that it uses marginal functions in an explicit way which allows for the development of specialized solution methods.

7 Primal feasible solution and upper bounds

Let us consider the optimal solution x^{low} of the convex hull problem (4.13)–(4.17) and the corresponding multipliers λ_j. Define $J^{\text{low}} = \{j \in J : \lambda_j > 0\}$.

If J^{low} contains only one element the point x^{low} is feasible, and therefore optimal, for the disjunctive formulation (2.4). If, however, there are more positive λ's, we need to generate a feasible point. A natural possibility is to consider the *restricted disjunctive* formulation:

$$\begin{array}{rl} \min & c^T x \\ \text{subject to} & Tx \in \bigcup_{j \in J^{\text{low}}} K_j, \\ & x \in \mathcal{D}. \end{array} \tag{7.1}$$

It can be solved by simple enumeration of all cases for $j \in J^{\text{low}}$:

$$\begin{array}{rl} \min & c^T x \\ \text{subject to} & Tx \geq v^j, \\ & x \in \mathcal{D}. \end{array} \tag{7.2}$$

In general, it is not guaranteed that any of these problems has a nonempty feasible set, as the following example shows. Let $n = 3$, $T = I$, and let there will be only three p-efficient points: $v^1 = (1,0,0)$, $v^2 = (0,1,0)$, $v^3 = (0,0,1)$, and two additional deterministic constraints: $x_1 \leq 1/2$, $x_2 \leq 1/2$, and $c = (0,0,1)$. The convex hull problem has $\lambda_1 = \lambda_2 = 1/2$, $\lambda_3 = 0$, but both problems (7.2) for $j = 1,2$ have empty feasible sets.

To ensure that problem (7.1) has a solution it is sufficient that the following stronger version of Assumption 2.5 holds.

Assumption 7.1. *The set* $\Lambda := \{(u,w) \in \mathbb{R}_+^{m+s} \mid A^T w + T^T u \leq c\}$ *is nonempty and bounded.*

Indeed, each od the dual problems (2.6) has an optimal solution, so by duality in linear programming each of the subproblems (7.2) has an optimal solution. We can, therefore, solve all of them and choose the best solution. An alternative strategy would be to solve the corresponding upper bounding problem (7.2) every time a new p-efficient point is generated. This may be computationally efficient, especially if we solve the dual problem (2.6), in which only the objective function changes from iteration to iteration.

If the distribution function of ξ is r-concave on the set of possible values of ξ, Theorem

3.3 provides an alternative formulation of the upper bound problem (7.1):

$$\begin{aligned} \min \quad & c^T x \\ \text{subject to} \quad & x \in \mathcal{D} \\ & Tx \geq z, \\ & z \in \mathbb{Z}^s, \\ & z \geq \sum_{j \in J^{\text{low}}} \lambda_j v^j, \\ & \sum_{j \in J^{\text{low}}} \lambda_j = 1 \\ & \lambda_j \geq 0,\ j \in J^{\text{low}}. \end{aligned}$$

It may be easier to deal with if the number of p-efficient points in L^{low} is large.

References

[1] C. Barnhart, E.L. Johnson, G.L. Nemhauser, M.V.P. Savelsbergh and P.H. Vance, Branch-and-Price: Column Generation for Solving Huge Integer Programs, *Operations Research* 46 (1998) 316–329.

[2] C. Borell, Convex Set Functions in d-Space, *Periodica Mathematica Hungarica* 6 (1975) 111–136.

[3] 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 Equation, *Journal of Functional Analysis* 22 (1976) 366–389.

[4] 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 (1958) 235–263.

[5] C.L. Dert, Asset–Liability Management for Pension Funds: A Multistage Chance-Constrained Programming Approach, PhD Thesis, Erasmus University, Rotterdam, 1995.

[6] L.R Ford and D.R. Fulkerson, A Suggested Computation for Maximal Multicommodity Network Flows, *Management Science* 5 (1958) 97–101.

[7] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms I and II*, Springer-Verlag, Berlin 1993.

[8] K.C. Kiwiel, *Methods of Descent for Nondifferentiable Optimization*, Lecture Notes in Mathematics Vol. 1133, Springer-Verlag, Berlin 1985.

[9] L.B. Miller and H. Wagner, Chance-Constrained Programming with Joint Constraints. *Operations Research* **13** (1965) 930–945.

[10] G.L. Nemhauser and L.A. Wolsey, *Integer and Combinatorial Optimization*, John Wiley & Sons, New York, 1988.

[11] V.I. Norkin and N.V. Roenko, α-Concave Functions and Measures and Their Applications. *Kibernet. Sistem. Anal.* 189 (1991) 77–88 (in Russian); translation in: *Cybernet. Systems Anal.* 27 (1991) 860–869.

[12] V.I. Norkin, Y. Ermoliev and A. Ruszczyński: On optimal allocation of indivisibles under uncertainty, Operations Research 46 (1998) 381–395.

[13] A. Prékopa, On Probabilistic Constrained Programming. *Proceedings of the Princeton Symposium on Mathematical Programming.* Princeton University Press, 1970, pp. 113–138.

[14] A. Prékopa, Logarithmic Concave Functions with Applications to Stochastic Programming, *Acta Sci. Math. (Szeged)* 32 (1971) 301–316.

[15] A. Prékopa, On Logarithmic Concave Measures and Functions, *Acta Sci. Math. (Szeged)* 34 (1973) 339–343.

[16] A. Prékopa, Contributions to the Theory of Stochastic Programming. *Mathematical Programming* 4 (1973) 202–221.

[17] A. Prékopa, Boole-Bonferroni Inequalities and Linear Programming.*Operation Research* 36 (1988) 145–162.

[18] A. Prékopa, Dual Method for a One-Stage Stochastic programming Problem with Random RHS Obeying a Discrete Probability Distribution, *Zeitschrift für Operations Research* 34 (1990) 441–461.

[19] A. Prékopa,Sharp Bounds on Probabilities Using Linear Programming.*Operation Research* 38 (1990) 227-239.

[20] A. Prékopa: *Stochastic Programming*, Kluwer, Dordrecht, Boston, 1995.

[21] A. Prékopa, The Use of Discrete Moment Bounds in Probabilistic Constrained Stochastic Programming Models, *Annals of Operations Research*, to appear.

[22] A. Prékopa, B. Vizvári, and T. Badics, Programming Under Probabilistic Constraint with Discrete Random Variable, in: *New Trends in Mathematical Programming*, L. Grandinetti et al. (Ed.), Kluwer, Dordrecht, Boston, 1998, pp. 235–255.

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

[24] A. Ruszczyński: A regularized decomposition method for minimizing a sum of polyhedral functions, *Mathematical Programming* 35 (1986) 309-333.

[25] A. Ruszczyński, An augmented Lagrangian decomposition method for block diagonal linear programming problems, *Operations Research Letters* 8 (1989) 287–294.

[26] S. Sen, Relaxations for the Probabilistically Constrained Programs with Discrete Random Variables, *Operations Research Letters* 11 (1992) 81–86.

[27] M.R. Singh, C.T. Abraham and R. Akella, A Wafer Design Problem in Semiconductor manufacturing for Reliable Customer Service, *IEEE Transactions on Components, Hybrids and Manufacturing Technology* 13 (1990) 103–108.

[28] B. Vizvári, *Beiträge zum Frobenius Problem*, D.Sc.Nat. Dissertation, Technische Hohschule Carl Schorlemmer, Leuna–Merseburg, Germany, 1987.

[29] S.L. Tayur, R.L. Themas and N.R. Natraj, An Algebraic Geometry Algorithm for Scheduling in the Presence of Setups and Correlated Demands, *Mathematical Programming* 69 (1995) 369–401.