

Rüdiger Schultz

Stochastic programming with integer variables

Received: December 2, 2002 / Accepted: April 23, 2003

Published online: May 28, 2003 – © Springer-Verlag 2003

Abstract. Including integer variables into traditional stochastic linear programs has considerable implications for structural analysis and algorithm design. Starting from mean-risk approaches with different risk measures we identify corresponding two- and multi-stage stochastic integer programs that are large-scale block-structured mixed-integer linear programs if the underlying probability distributions are discrete. We highlight the role of mixed-integer value functions for structure and stability of stochastic integer programs. When applied to the block structures in stochastic integer programming, well known algorithmic principles such as branch-and-bound, Lagrangian relaxation, or cutting plane methods open up new directions of research. We review existing results in the field and indicate departure points for their extension.

1. Introduction

During the last decade, integer variables have gained increased attention in stochastic programming. Regarding theoretical analysis and algorithmic treatment of stochastic programs this leads to a considerable change of paradigms. While classical stochastic linear programs in continuous variables benefit from inherent convexity properties, this is no longer the case if integrality enters the model. The present paper aims at discussing achievements and challenges in stochastic integer programming. We will consider well-established topics of integer programming and study their impacts in stochastic integer programming.

2. Models

In stochastic programming, like in mathematical programming in general, integer variables, first of all, serve to extend the modeling abilities by representing issues such as indivisibility, Boolean decisions, disjunctions, or piecewise linearity. These features may already be part of the random optimization problem which stands at the beginning of each stochastic programming model. The way a random optimization problem is turned into a stochastic program, on the one hand, depends on the interplay of making decisions and gaining information. On the other hand, it depends on the statistical parameters that are employed for building the objective function and the constraints of the stochastic program. These parameters may be simply the expected value or, as in

approaches expressing risk aversion, probabilities, dispersions, or conditional expectations. Depending on the risk measure selected, statistical parameters entering the model may be another source of integrality in stochastic programming.

2.1. Two-stage model

The two-stage model reflects the simplest mode of the mentioned interplay of decision and information. It is based on the random mixed-integer linear program

$$\min_{x,y,y'} \left\{ c^\top x + q^\top y + q'^\top y' : Tx + Wy + W'y' = h(\omega), \right. \\ \left. x \in X, y \in \mathbb{Z}_+^{\bar{m}}, y' \in \mathbb{R}_+^{m'} \right\}. \tag{1}$$

All ingredients are supposed to have conformable dimensions. The matrices W, W' are rational, and $X \subseteq \mathbb{R}^m$ is a nonempty polyhedron, possibly involving integer requirements to components of x . The right-hand side $h(\omega) \in \mathbb{R}^s$ is a random vector on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. For ease of exposition we confine ourselves in the present paper to randomness in the right-hand sides of the respective optimization problems. Decisions on the variables x and (y, y') have to be made stage-wise. Before observing the realization of $h(\omega)$ the variable x has to be selected in the first stage. After having decided on x and having observed $h(\omega)$, the second-stage (or recourse) decision (y, y') has to be made. Such a scheme frequently arises in decision making under uncertainty. The first stage corresponds to decisions that have to be made without anticipation of (parts of) the problem data. The second stage allows for corrective actions after uncertainty has been unveiled. The decision problem under uncertainty is to select the first-stage decision in an optimal way. Optimality, of course, has to be made precise in this context. Different statistical parameters then come into play.

Given x and $h(\omega)$, the second-stage decision (y, y') has to be selected to be the best possible. The corresponding optimal value can be expressed as $\Phi(h(\omega) - Tx)$ where

$$\Phi(t) := \min \left\{ q^\top y + q'^\top y' : Wy + W'y' = t, y \in \mathbb{Z}_+^{\bar{m}}, y' \in \mathbb{R}_+^{m'} \right\}. \tag{2}$$

In parametric optimization the above object is well-known as the value function of a mixed-integer program. Studies such as [7, 12] will be very useful when analyzing structure and stability of the stochastic integer programs we are heading for. By integer programming theory ([40]), the function Φ is real-valued on \mathbb{R}^s if $W(\mathbb{Z}_+^{\bar{m}}) + W'(\mathbb{R}_+^{m'}) = \mathbb{R}^s$ and $\{u \in \mathbb{R}^s : W^\top u \leq q, W'^\top u \leq q'\} \neq \emptyset$ (complete and sufficiently expensive recourse) which, therefore, is assumed throughout. Without integer requirements in (2), linear programming duality provides that Φ is a piecewise linear convex function. The breakdown of convexity in (2) in the presence of integer variables is the main source of the challenges that arise when passing from stochastic linear programming to stochastic integer programming.

A first-stage decision x now induces the total (random) costs $c^\top x + \Phi(h(\omega) - Tx)$. The decision problem of finding an optimal x without anticipating future realization of $h(\omega)$ may be formulated as finding a “best” random variable in the indexed family

$$\left(c^\top x + \Phi(h(\omega) - Tx) \right)_{x \in X}. \tag{3}$$

Stochastic orders (see e.g. [39]) provide a rich selection of scalar criteria for choosing an optimal member from the function family in (3). In the stochastic programming context at hand it is reasonable to base this choice on the following guidelines:

- The scalar criterion shall be in tune with basic asymptotic results in stochastics such as the law of large numbers and, if possible, be consistent with ordering principles such as stochastic dominance, see e.g. [42] for a recent exposition.
- The resulting stochastic program shall be well-posed from formal viewpoint. For instance, it makes sense to strive for an objective function that is at least lower semi-continuous in x , such that its infimum over a non-empty compact set is always attained. The latter is less esoteric than it might seem, since the value function Φ has poor analytical properties and is discontinuous in general.
- Finally, and not least importantly, the resulting stochastic program shall be computationally accessible. We will see that, with discretely distributed $h(\omega)$, stochastic integer programs may be equivalently restated as large-scale, block-structured mixed-integer optimization problems. The scalar criterion then determines the block structure, and one is especially interested in criteria inducing decomposable structures.

Not surprisingly, the most popular scalar criterion is taking the expectation. This leads to the objective function

$$Q_{\mathbb{E}}(x) := \int_{\Omega} \left(c^{\top} x + \Phi(h(\omega) - Tx) \right) \mathbb{P}(d\omega) \quad (4)$$

and the traditional expectation-based (two-stage) stochastic program with recourse

$$\min \{ Q_{\mathbb{E}}(x) : x \in X \}. \quad (5)$$

Without integer requirements, this model is the well-studied stochastic linear program with recourse, for details see the textbooks [11, 26, 43] and [51]. Among the existing models in stochastic integer programming, (5) is best understood so far. Therefore, it will have a central role subsequently in the present paper.

2.2. Risk measures

From a stochastics viewpoint, optimizing an expectation tacitly supposes repeatability of the decision process a great number of times under identical conditions. Asymptotic results such as the law of large numbers then provide convergence in stochastic terms of the random quantities to their expected values. On the other hand, safety issues are addressed only inadequately by the expectation framework. This has led to mean-risk models. In the present context, a mean-risk model reads

$$\min \{ Q_{\mathbb{E}}(x) + \rho \cdot Q_{risk}(x) : x \in X \} \quad (6)$$

where $Q_{\mathbb{E}}$ is as in (4), $\rho > 0$ is some preselected weight factor, and Q_{risk} is a functional measuring risk. The model (6) typically arises from using a weighting method to generate the efficient frontier associated with decisions involving mean as one criterion, and risk as the other. Thus, a bi-criteria problem is in the background, and a family of

mean-risk models (6) with ρ varying in \mathbb{R}_+ usually has to be solved. In stochastic integer programming the image set $\{(Q_{\mathbb{E}}(x), Q_{risk}(x)) \in \mathbb{R}^2 : x \in X\}$ is nonconvex in general. Therefore, in general only a subset of the efficient frontier, the so-called supported part, is reached when solving (6) for the ρ in \mathbb{R}_+ , see [8, 29] for some basic facts of multi-criteria optimization, and [54] for an example of tracing supported efficient points by using a specification of (6).

The identification of proper risk measures is a field of active research in stochastics, see for instance the recent papers [3, 42, 45] and the references therein. Here, we will not pursue risk modeling in general but rather introduce some specifications of Q_{risk} and discuss whether they lead to reasonable models in the sense of the guidelines formulated in Subsection 2.1.

Following Markowitz' seminal work, mean-variance models have gained attraction in decision making under uncertainty. In terms of (6) this suggests

$$Q_{risk}(x) := \int_{\Omega} \left[c^{\top}x + \Phi(h(\omega) - Tx) - \int_{\Omega} \left(c^{\top}x + \Phi(h(\omega) - Tx) \right) \mathbb{P}(d\omega) \right]^2 \mathbb{P}(d\omega).$$

Conceptually the variance has several drawbacks. In the minimization framework at hand, deviations below target, if at all, are far less critical than deviations above. So the symmetric penalty imposed by the variance should be abandoned in favour of one-sided measures. Moreover, the square involved drives us away from mixed-integer *linear* programming that will arise as a powerful tool for solving (5) and, if possible, should be used for mean-risk extensions as well. Finally, the square and the discontinuous nature of Φ may lead to a rather ill-posed functional $Q_{risk}(\cdot)$. In [54] an example is given where $Q_{risk}(\cdot)$ fails to be lower semicontinuous and (6), with compact X , has a finite infimum which is not attained. Some other difficulties that arise when including the variance of second-stage objective function values into the objective of a two-stage stochastic linear program are reported in [60].

The outlined deficiencies of variance suggest one-sided risk measures that are defined in terms of piecewise linear expressions, if necessary involving mixed-integer variables. Instances fitting into these requirements are excess probability, conditional value-at-risk, or absolute semideviation.

With a preselected threshold $\varphi_o \in \mathbb{R}$ the excess probability functional

$$Q_{\mathbb{P}}(x) := \mathbb{P} \left(\left\{ \omega \in \Omega : c^{\top}x + \Phi(h(\omega) - Tx) > \varphi_o \right\} \right) \tag{7}$$

measures the probability of facing total random objective values exceeding φ_o . Excess probabilities, sometimes also called aspiration level objectives, arise in quite different modeling situations. For instance as ruin probability if φ_o is a critical cost level, or as probability of interrupting a random network if φ_o is a properly chosen path length [71]. The study of probability objectives in stochastic programming seemingly dates back to [9, 16]. In stochastic programs with recourse (then without integer variables) an early contribution to risk aversion by a probability term like (7) is [10].

Excess probabilities do not quantify the extent to which objective values exceed the threshold. The latter may be achieved by another interesting risk measure, the

conditional value-at-risk [42, 45]. Some prerequisites are needed for its introduction. Let us denote by

$$F(x, \eta) := \mathbb{P}(\{\omega \in \Omega : c^\top x + \Phi(h(\omega)) - Tx \leq \eta\})$$

the distribution function of the random variable $c^\top x + \Phi(h(\omega)) - Tx$. With a preselected probability $0 < \alpha < 1$, the α -Value-at-Risk is given by

$$\eta_\alpha(x) := \min\{\eta : F(x, \eta) \geq \alpha\}.$$

The α -Conditional-Value-at-Risk (α -CVaR) $Q_{CVaR}(x)$ is the expected value of what is called the α -tail distribution of $c^\top x + \Phi(h(\omega)) - Tx$. This distribution is given by the distribution function

$$F_\alpha(x, \eta) := \begin{cases} 0 & , \text{ for } \eta < \eta_\alpha(x) \\ \frac{F(x, \eta) - \alpha}{1 - \alpha} & , \text{ for } \eta \geq \eta_\alpha(x) \end{cases}$$

If the underlying random variable, in our case $c^\top x + \Phi(h(\omega)) - Tx$, has no probability atom at $\eta_\alpha(x)$ then $Q_{CVaR}(x)$ equals a conditional expectation, namely

$$Q_{CVaR}(x) = \mathbb{E}(c^\top x + \Phi(h(\omega)) - Tx \mid c^\top x + \Phi(h(\omega)) - Tx \geq \eta_\alpha(x)).$$

We have preferred the general definition since, due to the discontinuous nature of Φ and the numerical importance of discretely distributed $h(\omega)$, later on, probability atoms rather will be the rule than the exception. The general definition is rather difficult to handle analytically. The following minimization formula from [42, 45] serves well in overcoming this difficulty.

Proposition 1. *It holds*

$$Q_{CVaR}(x) = \min_{\eta \in \mathbb{R}} f_\alpha(x, \eta)$$

where

$$f_\alpha(x, \eta) := \eta + \frac{1}{1 - \alpha} \int_{\Omega} \max\{c^\top x + \Phi(h(\omega)) - Tx - \eta, 0\} \mathbb{P}(d\omega).$$

As a first consequence, Proposition 1 yields that $Q_{CVaR}(\cdot)$ is convex for models with linear recourse, i.e., models where integrality is missing in (2). Indeed, Φ is convex without integrality in (2), and thus f_α jointly convex in (x, η) . The argument is completed by observing that minimizing a jointly convex function with respect to one variable produces a function that is convex in the other variable.

The α -Value-at-Risk (α -VaR) arising in the above construction is a popular risk measure itself that is written into many financial regulations. Here, we will not pursue its discussion, but remark that α -CVaR, clearly, provides an upper bound for α -VaR such that first-stage decisions x leading to a small α -CVaR lead to a small α -VaR as well, see [45] for a thorough discussion of relation between α -CVaR and α -VaR.

For the definitions of $Q_{\mathbb{P}}$ and Q_{CVaR} a static threshold and a static probability must be preselected. Such preselections can be avoided when working with a dynamic threshold, for instance. The absolute semideviation is a one-sided risk measure which is determined

by a piecewise linear expression and uses the expectation as a dynamic threshold. For the random variables from (3) it reads

$$Q_{ASD}(x) := \int_{\Omega} \max \left\{ c^{\top} x + \Phi(h(\omega) - Tx) - \int_{\Omega} (c^{\top} x + \Phi(h(\omega) - Tx)) \mathbb{P}(d\omega), 0 \right\} \mathbb{P}(d\omega).$$

Introducing risk terms into traditional stochastic programming models has regained attention among researchers only recently. Our selection of excess probability, conditional value-at-risk, and absolute semideviation is rather subjective than exhaustive. From an integer programming viewpoint adding risk terms to stochastic integer programs is interesting since it induces further types of computationally challenging large-scale mixed-integer programs.

2.3. Block structures

In many practical situations the dimension of the random vector $h(\omega)$ that underlies the mean-risk model (6) is too big to enable direct computation of the multi-dimensional integrals in the objective of (6) for continuous probability distributions. Therefore, discrete probability measures have a prominent role in stochastic programming computations. This is backed by stability properties we will mention later in this paper. Beside turning integrals into sums, discrete distributions allow for equivalent representations of mean-risk models as block-structured mixed-integer linear programs. Assume that $h(\omega)$ follows a discrete finite distribution with realizations (or scenarios) h_j and probabilities $p_j, j = 1, \dots, J$.

Proposition 2. *With the above discrete distribution of $h(\omega)$ the following is valid.*

1. *The expectation-based model $\min\{Q_{\mathbb{E}}(x) : x \in X\}$ is equivalent to*

$$\begin{aligned} \min_{x, y_j, y'_j} \left\{ c^{\top} x + \sum_{j=1}^J p_j (q^{\top} y_j + q'^{\top} y'_j) : \right. \\ Tx + Wy_j + W'y'_j = h_j, \\ \left. x \in X, y_j \in \mathbb{Z}_+^{\bar{m}}, y'_j \in \mathbb{R}_+^{m'}, j = 1, \dots, J \right\}. \end{aligned} \tag{8}$$

2. *If X is bounded then there exists a constant $M > 0$ such that the mean-risk model $\min\{Q_{\mathbb{E}}(x) + \rho \cdot Q_{\mathbb{P}}(x) : x \in X\}$ is equivalent to*

$$\begin{aligned} \min_{x, y_j, y'_j, \theta_j} \left\{ c^{\top} x + \sum_{j=1}^J p_j (q^{\top} y_j + q'^{\top} y'_j) + \rho \cdot \sum_{j=1}^J p_j \theta_j : \right. \\ Tx + Wy_j + W'y'_j = h_j, \\ c^{\top} x + q^{\top} y_j + q'^{\top} y'_j - (M - \varphi_o) \theta_j \leq \varphi_o, \\ \left. x \in X, y_j \in \mathbb{Z}_+^{\bar{m}}, y'_j \in \mathbb{R}_+^{m'}, \theta_j \in \{0, 1\}, j = 1, \dots, J \right\}. \end{aligned} \tag{9}$$

3. The mean-risk model $\min\{Q_{\mathbb{E}}(x) + \rho \cdot Q_{CVaR}(x) : x \in X\}$ is equivalent to

$$\begin{aligned} \min_{x, \eta, y_j, y'_j, v_j} \{ & c^\top x + \sum_{j=1}^J p_j (q^\top y_j + q'^\top y'_j) + \rho \cdot (\eta + \frac{1}{1-\alpha} \sum_{j=1}^J p_j v_j) : \\ & Tx + Wy_j + W'y'_j = h_j, \\ & c^\top x + q^\top y_j + q'^\top y'_j - \eta \leq v_j, \\ & x \in X, \eta \in \mathbb{R}, y_j \in \mathbb{Z}_+^{\bar{m}}, y'_j \in \mathbb{R}_+^{m'}, v_j \in \mathbb{R}_+, j = 1, \dots, J\}. \end{aligned} \tag{10}$$

4. For $0 < \rho < 1$, the mean-risk model $\min\{Q_{\mathbb{E}}(x) + \rho \cdot Q_{ASD}(x) : x \in X\}$ is equivalent to

$$\begin{aligned} \min_{x, y_j, y'_j, w, v_j} \{ & c^\top x + (1-\rho) \cdot w + \rho \cdot \sum_{j=1}^J p_j v_j : \\ & Tx + Wy_j + W'y'_j = h_j, \\ & \sum_{i=1}^J p_i (q^\top y_i + q'^\top y'_i) = w, \\ & q^\top y_j + q'^\top y'_j \leq v_j, w \leq v_j, \\ & x \in X, y_j \in \mathbb{Z}_+^{\bar{m}}, y'_j \in \mathbb{R}_+^{m'}, w, v_j \in \mathbb{R}, j = 1, \dots, J\}. \end{aligned} \tag{11}$$

Proof. The equivalence in part 1 is already well-known from linear two-stage models without integer requirements, see [11, 26, 43]. The same proof applies here.

The crucial point in part 2 is introducing the Boolean variables θ_j to “count” whether $c^\top x + \Phi(h_j - Tx) > \varphi_0$ for a given x and $j \in \{1, \dots, J\}$. In case $\theta_j = 1$, the “big M” guarantees that the “counting constraint” becomes vacuous. A feasible choice is $M \geq \sup\{c^\top x + \Phi(h_j - Tx) : x \in X, j = 1, \dots, J\}$. Since X is bounded, the supremum on the right is less than $+\infty$, see Theorem 8.1 in [7] or [54], for instance.

Part 3 is a direct consequence of Proposition 1. For part 4 the following reformulation is useful. To compress notation we use the symbols \mathbb{E} for taking the expectation and $g(x, \omega) := c^\top x + \Phi(h(\omega) - Tx)$. Applying $\max\{s - t, 0\} = \max\{s, t\} - t$ for $s, t \in \mathbb{R}$ we obtain:

$$\begin{aligned} & \min\{Q_{\mathbb{E}}(x) + \rho \cdot Q_{ASD}(x) : x \in X\} \\ &= \min_{x \in X} \{ \mathbb{E}[g(x, \omega)] + \rho \mathbb{E}[\max\{g(x, \omega) - \mathbb{E}[g(x, \omega)], 0\}] \} \\ &= \min_{x \in X} \{ (1-\rho) \mathbb{E}[g(x, \omega)] + \rho \mathbb{E}[\max\{g(x, \omega), \mathbb{E}[g(x, \omega)]\}] \} \\ &= \min_{x \in X} \{ c^\top x + (1-\rho) \sum_{j=1}^J p_j \Phi(h_j - Tx) \\ & \quad + \rho \sum_{j=1}^J p_j \max\{\Phi(h_j - Tx), \sum_{i=1}^J p_i \Phi(h_i - Tx)\} \}. \quad \square \end{aligned}$$

The mixed-integer linear program in the above part 1 obeys a dual-block angular structure that is omnipresent in two-stage stochastic linear programming. Through

$Tx + Wy_j + W'y'_j = h_j$ each second-stage variable (y_j, y'_j) , $j = 1, \dots, J$ is linked explicitly with the first-stage variable x . There are no explicit links among second-stage variables belonging to different scenarios. Algorithmically, this structure has been the departure point of very successful decomposition algorithms in stochastic linear programming without integer requirements. Also in the presence of integer variables the structure leads to promising approaches, as will be seen later in this paper.

Parts 2 and 3 of Proposition 2 confirm that the mean-risk models there again induce the principal block structure from part 1. Indeed, no explicit links among second-stage variables for different scenarios occur. All these variables are explicitly linked with first-stage variables. Of course, modeling requirements may produce auxiliary first- and second-stage variables such as η in part 3 or θ_j in part 2. One concludes that excess probability and conditional value-at-risk are algorithmically amenable since they lead to mean-risk models allowing for direct extension of methods applying to the expectation-based model (5).

In contrast with parts 2 and 3, part 4 displays a mixed-integer linear program with explicit coupling among second-stage variables for different scenarios. The resulting block structure does not fit into existing schemes of stochastic programming and, therefore, may be an interesting object of future research.

Part 2 of Proposition 2 is an example how integrality in stochastic programming may be induced by the probabilistic context a model is developed in. The articles [44, 54] contain some first results for this class of problems.

2.4. Multi-stage extension

In the two-stage models introduced so far uncertainty is unveiled at once and decisions subdivide into those before and those after unveiling uncertainty. Multi-stage stochastic programs address the more complex case where uncertainty is unveiled stepwise, with intermediate decisions. This makes the essential difference with deterministic multi-period optimization problems where, as well, decisions are assigned to time stages but the *whole* data information is available in the beginning. The decision $x_t \in \mathbb{R}^{m_t}$ at stage $t \in \{1, \dots, T\}$ must be based on information available up to time t only (nonanticipativity). The information is modeled as a discrete time stochastic process $\{b_t\}_{t=1}^T$ on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with b_t taking values in \mathbb{R}^{s_t} . The underlying random optimization problem reads

$$\begin{aligned} \min \{ & c_1^\top x_1 + c_2^\top x_2 + \dots + c_T^\top x_T : \\ & A_{11}x_1 & & & & = b_1 \\ & A_{21}x_1 + A_{22}x_2 & & & & = b_2(\omega) \\ & \vdots & & \ddots & & \vdots \\ & A_{T1}x_1 + \dots + A_{TT}x_T & & & & = b_T(\omega) \\ & x_1 \in X_1 \quad x_2 \in X_2 \quad \dots \quad x_T \in X_T & & & & \}. \end{aligned}$$

We assume that b_1 is deterministic, that all matrices arising have rational entries, and that $X_t, t = 1, \dots, T$ are nonempty polyhedra, possibly involving integer requirements. Although we have adapted the notation to the more usual one for multi-stage models it is evident that the random optimization problem (1) from Subsection 2.1 fits into the above model for $T = 2$.

The fact that decisions always must be based on the *available* information leads to the following sequence of actions and observations. First, the decision x_1 is made. Then $b_2(\omega)$ is observed, and the decision $x_2 = x_2(x_1, b_2(\omega))$ is made. In general, the decision x_t depends on the vector $(x_1, \dots, x_{t-1}, b_2(\omega), \dots, b_t(\omega))$. Finally, $x_T = x_T(x_1, \dots, x_{T-1}, b_2(\omega), \dots, b_T(\omega))$ is made. As in the two-stage case, the aim is to optimize the first-stage decision. Again, this can be understood as selecting a “best” member from a family of random variables. This family can be expressed by

$$(c_1^\top x_1 + \Phi_{mult}(x_1, b(\omega)))_{x_1 \in X_1, A_{11}x_1 = b_1} \tag{12}$$

with the mixed-integer value function

$$\begin{aligned} \Phi_{mult}(x_1, b(\omega)) := & \min_{\substack{A_{21}x_1 + A_{22}x_2 = b_2(\omega) \\ x_2 \in X_2}} \left\{ c_2^\top x_2 + \min_{\substack{A_{31}x_1 + A_{32}x_2 + A_{33}x_3 = b_3(\omega) \\ x_3 \in X_3}} \{ c_3^\top x_3 + \dots \right. \\ & + \min_{\substack{A_{T-1,1}x_1 + \dots + A_{T-1,T-1}x_{T-1} = b_{T-1}(\omega) \\ x_{T-1} \in X_{T-1}}} \{ c_{T-1}^\top x_{T-1} \\ & \left. + \min_{\substack{A_{T1}x_1 + \dots + A_{TT}x_T = b_T(\omega) \\ x_T \in X_T}} c_T^\top x_T \dots \right\}. \end{aligned} \tag{13}$$

It is the nested nature that makes the value function Φ_{mult} a far more complicated mathematical object than its two-stage counterpart Φ from (2). Relatively little is known about Φ_{mult} in the mixed-integer situation adopted here.

Multi-stage stochastic integer programs arise by imposing scalar criteria for making the “best” selection in (12). The principal ideas developed in Subsections 2.2 and 2.3 for the two-stage case readily extend and provide flexibility for including risk aversion into multi-stage models. Conceptually, this is quite close to the two-stage case. Technically, however, multi-stage models are much more demanding. Beside the mentioned nesting, the nonanticipativity of decisions is more subtle, and the model size is considerably bigger.

As with two-stage models, taking the expectation is the most popular scalar criterion in the multi-stage case. For separability reasons, the expectation then unfolds into a nested sequence of conditional expectations:

$$\begin{aligned} & \min_{\substack{A_{11}x_1 = b_1 \\ x_1 \in X_1}} \left\{ c_1^\top x_1 + \mathbb{E}_{b(\omega)}[\Phi_{mult}(x_1, b(\omega))] \right\} \\ & = \min_{\substack{A_{11}x_1 = b_1 \\ x_1 \in X_1}} \left\{ c_1^\top x_1 + \right. \\ & \quad \mathbb{E}_{b_2(\omega)} \left[\min_{\substack{A_{21}x_1 + A_{22}x_2 = b_2(\omega) \\ x_2 \in X_2}} \left\{ c_2^\top x_2 + \mathbb{E}_{b_3(\omega)|b_{[1,2]}} \left[\min_{\substack{A_{31}x_1 + A_{32}x_2 + A_{33}x_3 = b_3(\omega) \\ x_3 \in X_3}} \{ c_3^\top x_3 + \dots \right. \right. \right. \\ & \quad \left. \left. \left. + \mathbb{E}_{b_{T-1}(\omega)|b_{[1,T-2]}} \left[\min_{\substack{A_{T-1,1}x_1 + \dots + A_{T-1,T-1}x_{T-1} = b_{T-1}(\omega) \\ x_{T-1} \in X_{T-1}}} \{ c_{T-1}^\top x_{T-1} \right. \right. \right. \right. \\ & \quad \left. \left. \left. \left. + \mathbb{E}_{b_T(\omega)|b_{[1,T-1]}} \left[\min_{\substack{A_{T1}x_1 + \dots + A_{TT}x_T = b_T(\omega) \\ x_T \in X_T}} c_T^\top x_T \right] \dots \right] \right] \right] \right\}. \end{aligned}$$

3. Mixed-integer value functions for structure and stability

In Section 2 we had set aside the question whether the models presented are mathematically sound. The expectation and risk functionals there involve different modes of integration, and for the latter to be well-posed certain requirements have to be met. Further important issues are the analytical properties of the mentioned functionals that have crucial impact on well-posedness and difficulty of the resulting mean-risk optimization problems. In (2) and (13) we had already seen that value functions of mixed-integer programs have a central role in the mathematical structures behind our models. This explains the role of parametric (integer) optimization in stochastic (integer) programming. The structure of stochastic programs is essentially determined by an interplay of value functions and facts from probability theory. In what follows we will illustrate this on the functionals $Q_{\mathbb{E}}(\cdot)$ from (4) and $Q_{\mathbb{P}}(\cdot)$ from (7).

We begin with collecting prerequisites about the mixed-integer value function

$$\Phi(t) := \min \left\{ q^\top y + q'^\top y' : Wy + W'y' = t, y \in \mathbb{Z}_+^{\bar{m}}, y' \in \mathbb{R}_+^{m'} \right\}$$

already introduced in (2). The material has been adapted from parametric optimization sources such as [7, 12].

Proposition 3. *Let W, W' be rational matrices and assume that $W(\mathbb{Z}_+^{\bar{m}}) + W'(\mathbb{R}_+^{m'}) = \mathbb{R}^s$ and $\{u \in \mathbb{R}^s : W^\top u \leq q, W'^\top u \leq q'\} \neq \emptyset$. Then it holds:*

1. Φ is real-valued and lower semicontinuous on \mathbb{R}^s .
2. There exists a countable partition $\mathbb{R}^s = \cup_{i=1}^\infty \mathcal{T}_i$ such that the restrictions of Φ to \mathcal{T}_i are piecewise linear and Lipschitz continuous with a uniform constant not depending on i .
3. Each of the sets \mathcal{T}_i has a representation $\mathcal{T}_i = \{t_i + \mathcal{K}\} \setminus \cup_{j=1}^N \{t_{ij} + \mathcal{K}\}$ where \mathcal{K} denotes the polyhedral cone $W'(\mathbb{R}_+^{m'})$ and t_i, t_{ij} are suitable points from \mathbb{R}^s ; moreover, N does not depend on i .
4. There exist positive constants β, γ such that $|\Phi(t_1) - \Phi(t_2)| \leq \beta \|t_1 - t_2\| + \gamma$ whenever $t_1, t_2 \in \mathbb{R}^s$.

For convenience we denote by μ the image measure $\mathbb{P} \circ h^{-1}$ on \mathbb{R}^s . With this notation, the functions $Q_{\mathbb{E}}$ and $Q_{\mathbb{P}}$ become

$$Q_{\mathbb{E}}(x) = \int_{\mathbb{R}^s} (c^\top x + \Phi(h - Tx)) \mu(dh) \quad \text{and}$$

$$Q_{\mathbb{P}}(x) = \mu(\{h \in \mathbb{R}^s : c^\top x + \Phi(h - Tx) > \varphi_0\}).$$

Proposition 4. *Let W, W' be rational matrices and assume that $W(\mathbb{Z}_+^{\bar{m}}) + W'(\mathbb{R}_+^{m'}) = \mathbb{R}^s$ and $\{u \in \mathbb{R}^s : W^\top u \leq q, W'^\top u \leq q'\} \neq \emptyset$. Then it holds:*

1. The function $Q_{\mathbb{P}} : \mathbb{R}^m \rightarrow \mathbb{R}$ is real-valued lower semicontinuous. If μ has a density, then $Q_{\mathbb{P}}$ is continuous.
2. If $\int_{\mathbb{R}^s} \|h\| \mu(dh) < \infty$ then $Q_{\mathbb{E}} : \mathbb{R}^m \rightarrow \mathbb{R}$ is real-valued lower semicontinuous, and continuous if μ has a density.

Proof. The lower semicontinuity from part 1 of Proposition 3 ensures measurability of the relevant integrands and level sets in the definitions of $Q_{\mathbb{E}}$ and $Q_{\mathbb{P}}$. The integral defining $Q_{\mathbb{E}}$ then is always finite thanks to the finiteness assumption in part 2 above and the estimate in part 4 of Proposition 3. Therefore, both $Q_{\mathbb{P}}$ and $Q_{\mathbb{E}}$ are real-valued on \mathbb{R}^m . For all $x \in \mathbb{R}^m$, we introduce the notations $M(x) := \{h \in \mathbb{R}^s : c^\top x + \Phi(h - Tx) > \varphi_o\}$, $M_e(x) := \{h \in \mathbb{R}^s : c^\top x + \Phi(h - Tx) = \varphi_o\}$, and $M_d(x) := \{h \in \mathbb{R}^s : \Phi \text{ is discontinuous at } h - Tx\}$. Moreover, we denote by $\liminf_{x_n \rightarrow x} M(x_n)$ and $\limsup_{x_n \rightarrow x} M(x_n)$ the (set theoretic) limes inferior and limes superior, i.e., the sets of all points belonging to all but a finite number of the sets $M(x_n)$, $n \in \mathbb{N}$, and to infinitely many of the sets $M(x_n)$, respectively. Then it is possible to verify the following inclusions for all $x \in \mathbb{R}^m$, for details see [54],

$$M(x) \subseteq \liminf_{x_n \rightarrow x} M(x_n) \subseteq \limsup_{x_n \rightarrow x} M(x_n) \subseteq M(x) \cup M_e(x) \cup M_d(x).$$

By the semicontinuity of the probability measure on sequences of sets it follows

$$\begin{aligned} Q_{\mathbb{P}}(x) &= \mu(M(x)) \leq \mu(\liminf_{x_n \rightarrow x} M(x_n)) \\ &\leq \liminf_{x_n \rightarrow x} \mu(M(x_n)) = \liminf_{x_n \rightarrow x} Q_{\mathbb{P}}(x_n), \end{aligned}$$

establishing the asserted lower semicontinuity of $Q_{\mathbb{P}}$. Parts 2 and 3 of Proposition 3 provide that the sets $M_e(x)$ and $M_d(x)$ are contained in countably many hyperplanes in \mathbb{R}^s , i.e., in a set of Lebesgue measure zero. For the case that μ has a density this yields $\mu(M_e(x) \cup M_d(x)) = 0$ for all $x \in \mathbb{R}^m$. The last estimate then continues:

$$\begin{aligned} Q_{\mathbb{P}}(x) &= \mu(M(x)) = \mu(M(x) \cup M_e(x) \cup M_d(x)) \geq \mu(\limsup_{x_n \rightarrow x} M(x_n)) \\ &\geq \limsup_{x_n \rightarrow x} \mu(M(x_n)) = \limsup_{x_n \rightarrow x} Q_{\mathbb{P}}(x_n), \end{aligned}$$

implying continuity of $Q_{\mathbb{P}}$.

For the detailed proofs regarding $Q_{\mathbb{E}}$ we refer to [52]. The lower semicontinuity results through Fatou’s Lemma from the lower semicontinuity of Φ in part 1 of Proposition 3. The continuity of $Q_{\mathbb{E}}$ follows from Lebesgue’s Theorem on Dominated Convergence. The almost sure convergence of integrands needed there is another implication of the above argument that $\mu(M_d(x)) = 0$ if μ has a density. For both Fatou’s Lemma and Lebesgue’s Theorem integrable bounds of integrands are needed. These are constructed using part 4 of Proposition 3. □

The above proof is an example for the typical interplay of value function properties and probability facts in the structural analysis of stochastic programs. Beyond Proposition 4 there are many open research problems. These concern structural properties of mean-risk models along those introduced in Subsection 2.2. Much more challenging in this respect are multi-stage models where, before addressing stochastic integer programs, research must be invested into a deeper understanding of the nested value function (13).

In Subsection 2.3 we had already pointed to the issue of stability. Roughly speaking, one is interested in the parametric dependence of the optimal value and the set of optimal

solutions to a stochastic program, say (6), where the underlying probability measure has the role of the parameter. This abstract setting covers important topics related with the measure in (6), the most prominent being numerical approximation, incomplete information, and statistical estimation. Technically, the analysis of the involved functions $Q_{\mathbb{E}}$ and Q_{risk} jointly in the decision variable x and the probability measure μ becomes crucial then. In particular, this requires a suitable convergence framework on the space of probability measures. Such details being beyond the scope of the present paper we refer to the surveys [46, 53]. Although more elaborate probability instruments are used in the stability analysis of stochastic integer programs, thorough understanding of mixed-integer value functions is indispensable still, and their interplay with probability again poses numerous interesting research problems.

4. Branch-and-bound for decomposition

Branch-and-bound is a basic principle for solving optimization problems that consists of three phases. In the branching phase the feasible region is partitioned, usually step-wise with increasing granularity. Bounding aims at finding upper and lower bounds of the optimal objective values on the elements of the current partition. It often relies on relaxations. In addition there is a coordination phase to guide the solution process. Coordination usually involves rules for increasing the granularity (branching rules) and rules for eliminating parts of the feasible region from further consideration (pruning due to infeasibility, optimality, or inferiority).

The stochastic integer programs we had seen in (8), (9), (10), (11), and (15) all are mixed-integer linear programs. Their constraints fall into three categories: integrality, nonanticipativity, and stage constraints. The latter includes all constraints that explicitly link variables within and across the stages and that do not belong to the integrality and nonanticipativity constraints. With the exception of model (11) the stage constraints are separable with respect to the scenarios. The nonanticipativity arises explicitly in the multi-stage model (15). It can be made explicit in the two-stage models (8), (9), (10), (11), too. According to the number J of scenarios one introduces copies x_j , $j = 1, \dots, J$ of x and adds the constraint $x_1 = \dots = x_J$.

In what follows we will consider algorithms for stochastic integer programs that arise through different branching and bounding strategies in the frame sketched above. We will be less specific about the accompanying coordination that may follow established rules or their analogues.

4.1. Relaxing nonanticipativity while maintaining integrality

To minimize notational effort we will confine ourselves to the purely expectation based model $\min\{Q_{\mathbb{E}}(x) : x \in X\}$ arising in (5) and (8). Branching is accomplished by partitioning X with the help of linear constraints (possibly involving tolerances to have disjoint subregions). Let X_k be a current element of the partition. The corresponding subproblem reads

$$\begin{aligned}
& \min \{ Q_{\mathbb{E}}(x) : x \in X_k \} \\
& = \min_{x_j, y_j, y'_j} \left\{ \sum_{j=1}^J p_j (c^\top x_j + q^\top y_j + q'^\top y'_j) : Tx_j + Wy_j + W'y'_j = h_j, \right. \\
& \quad \left. x_j \in X_k, y_j \in \mathbb{Z}_+^{\bar{m}}, y'_j \in \mathbb{R}_+^{m'}, j = 1, \dots, J, \sum_{j=1}^J H_j x_j = 0 \right\}. \quad (16)
\end{aligned}$$

Here, nonanticipativity is represented by the two-stage analogue to the expression in (14). Any relaxation that removes nonanticipativity from the constraints and maintains separability of the objective now decouples the above model into scenario-wise mixed-integer subproblems, and provides a lower bound for $Q_{\mathbb{E}}(x)$ on X_k . Note that this holds in full analogy for the stochastic integer programs (9), (10), (15), but not for (11). In [13] Lagrangian relaxation has been employed for the decoupling. It leads to the Lagrangian

$$L(x, y, y', \lambda) := \sum_{j=1}^J L_j(x_j, y_j, y'_j, \lambda)$$

where

$$L_j(x_j, y_j, y'_j, \lambda) := p_j (c^\top x_j + q^\top y_j + q'^\top y'_j) + \lambda^\top H_j x_j, \quad j = 1, \dots, J.$$

The Lagrangian dual whose optimal value provides the desired lower bound then reads

$$\max \{ D(\lambda) : \lambda \in \mathbb{R}^l \} \quad (17)$$

where

$$\begin{aligned}
D(\lambda) = \min \left\{ \sum_{j=1}^J L_j(x_j, y_j, y'_j, \lambda) : Tx_j + Wy_j + W'y'_j = h_j, \right. \\
\left. x_j \in X_k, y_j \in \mathbb{Z}_+^{\bar{m}}, y'_j \in \mathbb{R}_+^{m'}, j = 1, \dots, J \right\}.
\end{aligned}$$

The above mixed-integer program fully decouples, namely,

$$D(\lambda) = \sum_{j=1}^J D_j(\lambda) \quad (18)$$

where

$$\begin{aligned}
D_j(\lambda) = \min \{ L_j(x_j, y_j, y'_j, \lambda) : Tx_j + Wy_j + W'y'_j = h_j, \\
x_j \in X_k, y_j \in \mathbb{Z}_+^{\bar{m}}, y'_j \in \mathbb{R}_+^{m'} \}.
\end{aligned}$$

The Lagrangian dual (17) is equivalent to a non-smooth convex minimization problem for which bundle methods from convex optimization can be employed, [21, 24, 28]. At each iteration these methods require the objective value and one subgradient of D . Both these entities can be read off an optimal solution to the optimization problem in (18). Computationally, the lower bounding thus reduces to repeated solution of the mixed-integer linear programs behind $D_j(\lambda)$, $j = 1, \dots, J$, for iterated values of λ . Integer

programming theory says that the lower bound obtained by the optimal value of (17) is never worse the bound obtained by LP relaxation of the initial subproblem (16).

Regarding an upper bound, the x_j -parts of optimal solutions to the problems behind $D_j(\lambda)$, $j = 1, \dots, J$, for optimal or nearly optimal λ are useful. To regain the relaxed nonanticipativity one may decide for the most frequent value arising or may average and round. If the resulting x -value is feasible for the initial subproblem it, of course, yields an upper bound for the optimal value.

As long as nonanticipativity is the only constraint across different scenarios, subproblems in the above procedure are single-scenario mixed-integer linear programs. Among the models introduced in Section 2 the same is true for (9), (10), (15), but not for (11). Although formally valid for the multi-stage model (15) the procedure faces specific difficulties there. The nonanticipativity constraints are more complex so that generating feasible points becomes more involved. For the same reason, the dimension the Lagrangian dual lives in is often too excessive for the application of existing subgradient methods.

In the next subsection we will address a decomposition method that does not require the Lagrangian dual. Another idea to circumvent dual optimization is presented in [34] where branch-and-price methodology is employed for solving special structured multi-stage stochastic integer programs.

4.2. Relaxing nonanticipativity and integrality

The approach to be described next has been proposed in [1, 2]. Again our exposition will resort to the expectation-based model from (5) and (8). We assume that the first-stage vector x has components in $\{0, 1\}$ only. When relaxing in (8) both nonanticipativity and integrality the problem reduces to solving the following single-scenario linear programming problems for $j = 1, \dots, J$

$$\min \left\{ c^\top x_j + q^\top y_j + q'^\top y'_j : Tx_j + Wy_j + W'y'_j = h_j, \right. \\ \left. x_j \in X_{rel}, y_j \in \mathbb{R}_+^m, y'_j \in \mathbb{R}_+^{m'} \right\}.$$

The notation X_{rel} indicates that integer requirements on x are replaced by membership in $[0, 1]^m$. The weighted sum of the optimal values of the above problems, clearly, provides a lower bound for the optimal value of (8).

Branching is accomplished by an interplay of two modes of partition. On the one hand, each of the single-scenario problems undergoes the well-known LP-based branch-and-bound of integer programming. For each problem a branching tree is developed where the root node corresponds to the j -th problem above and further nodes arise by fixing components of x to either 0 or 1. The second mode of partition leads to step-wise regaining of nonanticipativity. It rests on fixing variables in the different single-scenario problems to identical values if this is claimed by a nonanticipativity constraint. In our situation this concerns the first-stage variables only, and branching leads to fixing these variables across *all* scenarios to either 0 or 1, but never partially to 0 and partially to 1. From the viewpoint of the full model this corresponds to what is known as logical or constraint branching.

Using suitable coordination routines the method works through the individual branching trees. When processing an active node corresponding to a first-stage variable, the logical branching is employed, leading to a respective fixing of variables in the other trees. Via the weighted sum of optimal values lower bounds are obtained for the elements of the partition. Upper bounds arise as soon as there are solutions fulfilling all integrality and nonanticipativity constraints.

The characteristic feature of the above method is that nonanticipativity is handled implicitly. This increases the logical coordination effort but avoids the dimensionality problem mentioned at the end of Subsection 4.1. In addition, subproblems are merely linear instead of integer programs. Altogether, the method seems an interesting option for multi-stage problems.

Special care has to be taken in the mixed-integer case where fixing continuous variables is problematic. A remedy could be to relax (beside integrality) nonanticipativity for the integer variables only. In the relaxed problem the remaining nonanticipativity of the continuous variables then maintains coupling among scenarios. However, the relaxed problem now can be understood as a stochastic linear program without integer requirements for which powerful decomposition methods exist, see e.g. [11, 23, 26, 43, 49, 68].

Conceptually, the present method again applies to stochastic programs where nonanticipativity is the only explicit constraint interlinking scenarios. As in Subsection 4.1 these are the models (8), (9), (10), (15), but not (11).

4.3. Relaxing stage constraints

The stage constraints in the stochastic integer programs of Section 2 fall into two categories: constraints interlinking different time stages and reflecting the dynamics of the problem, and constraints acting within the individual stages. Although relaxation of stage constraints has gained some interest in the existing literature, so far it has never been imbedded into the branch-and-bound framework adopted here. In other words, it has been carried out in the root node only with emphasis on the bounding and without considering the possibility of partitioning the decision space. Nevertheless, some of the most successful practical applications of stochastic integer programming are based on relaxing stage constraints.

After relaxing stage constraints the problem remains a stochastic program since nonanticipativity is present still. The approach thus only makes sense if the resulting stochastic program is much simpler than the initial one. This can be achieved by relaxing the dynamics constraints interlinking the time stages. In [48], a paper dealing with multi-stage stochastic programs in continuous variables, this is called nodal decomposition. For stochastic integer programs nodal decomposition is a widely open field. For initial results see [19, 47].

In many practical applications one meets complex systems consisting of independent components that are loosely coupled through constraints reflecting interaction. Quite often the interaction follows identical principles throughout the stages such that all coupling constraints act within individual stages. Relaxing these constraints then usually leads to substantial decoupling. Nonanticipativity again being untouched the

partial models are stochastic integer programs still. A prominent field of application of this approach is power generation where the partial models are multi-stage stochastic integer programs assigned to the individual generation units of the system. For these single-unit subproblems highly efficient special purpose solvers have been developed making the initial problems the largest stochastic integer programs solved so far, see [20, 41, 61–63].

The mode of relaxation in the contributions reported in the present subsection is Lagrangian relaxation. This has initiated research into efficient subgradient methods for solving the Lagrangian dual and into adapted heuristics for generating upper bounds through feasible points. In [19] the duality gaps occurring with the relaxations in Subsections 4.1 and 4.3 are compared.

5. Disjunctive cuts for convexification and decomposition

With a discrete probability distribution of $h(\omega)$ and with Φ as in (2) the expectation-based model (5) may be represented as

$$\min \left\{ c^\top x + \sum_{j=1}^J p_j \Phi(h_j - Tx) : x \in X \right\}. \tag{19}$$

An algorithmic alternative to the full-size model (8) is working with the above model which is explicit in x only, but includes the implicit function Φ . As a spurring example let us consider the counterpart model to (19) in continuous variables where X is a nonempty polyhedron and Φ is replaced by

$$\Phi_{cont}(t) := \min \left\{ q^\top y + q'^\top y' : Wy + W'y' = t, (y, y') \in \mathbb{R}_+^{\bar{n}+m'} \right\}.$$

Linear programming duality provides

$$\Phi_{cont}(t) = \max \left\{ t^\top u : W^\top u \leq q, W'^\top u \leq q' \right\}.$$

Under the general assumptions from Subsection 2.1 the above linear program has a nonempty and bounded feasible region, with vertices $d_k, k = 1, \dots, K$, say. Therefore, $\Phi_{cont}(t) = \max_k d_k^\top t$ and the counterpart to (19) in continuous variables can be written as

$$\min \left\{ c^\top x + \sum_{j=1}^J p_j \max_{k=1, \dots, K} d_k^\top (h_j - Tx) : x \in X \right\}. \tag{20}$$

Solving (20) amounts to minimizing a sum of convex polyhedral functions, which is well-studied in convex optimization. However, (20) is implicit still, since the vertices d_k are not available initially and typically unveiled in the course of the iteration only. The method of choice is to iterate x in a Benders-type scheme by successively updating and minimizing lower approximations

$$\min \left\{ c^\top x + \sum_{j=1}^J p_j \max_{k \in \mathcal{K}_j} d_k^\top (h_j - Tx) : x \in X \right\} \tag{21}$$

where $\mathcal{K}_j, j = 1, \dots, J$, are subsets of $\{1, \dots, K\}$. The update is accomplished by generating vertices d_k through dual solutions of the single-scenario linear programs

$$\min \left\{ q^\top y + q'^\top y' : Wy + W'y' = h_j - Tx^v, (y, y') \in \mathbb{R}_+^{m+m'} \right\} \quad (22)$$

where x^v is the current iterate of x . The minimization in (21) is essentially carried out by linear programming, usually involving enhancements such as (quadratic) regularization, cut deletion, or cut aggregation.

The above account is a quick lesson on a powerful solution method in (continuous) stochastic linear programming, see [23, 49, 68] for further details. In the core of the method there is a piecewise linear convex function, cf. (20), whose approximations are updated by cuts of the type $d_k^\top (h_j - Tx)$ and minimized successively, cf. (21). The cuts are made available through linear programming duality from single-scenario second-stage linear programs with stochasticity $h(\omega)$ and first-stage decision x in the right-hand side, cf. (22). Beside the cuts themselves their functional dependence on (x, ω) is explicitly available. Indeed, the cut coefficients, although obtained locally at $x = x^v$, enter the global expressions $\max_{k \in \mathcal{K}_j} d_k^\top (h_j - Tx)$ that successively shape the convex objective in (20).

Returning to (19) the first observation is that single-scenario programs like (22) now are mixed-integer, and LP duality is no longer available for cut generation. If explicit convex hull descriptions for the feasible regions of all single-scenario mixed-integer programs for all $h(\omega)$ and all x were available, then, in theory, one could proceed according to the above method. Complete descriptions in general being out of reach, one convexifies feasible regions of subproblems and the epigraph of Φ only where necessary, and exploits structure wherever possible. This is the principal approach in [56] from where we have taken the material to be presented next. The topic is research in progress, with many open questions. So far, disjunctive cutting planes, see [4–6, 25], provide the major integer programming input into this recent line of stochastic programming research.

Let us assume that X is a nonempty polyhedron, for simplicity without integer requirements, and that the mixed-integer program defining $\Phi(t)$ is a mixed- $\{0, 1\}$ linear program of the form

$$\min \left\{ q^\top y + q'^\top y' : Wy + W'y' \geq t, y \in \{0, 1\}^m, y' \in \mathbb{R}_+^{m'} \right\}. \quad (23)$$

For $\omega \in \Omega := \{1, \dots, J\}$ and $x \in X$ consider the feasible sets

$$Y(x, \omega) := \left\{ (y, y') \in \{0, 1\}^m \times \mathbb{R}_+^{m'} : Wy + W'y' \geq h(\omega) - Tx \right\}$$

and assume for simplicity that

$$\left\{ y' \in \mathbb{R}_+^{m'} : W'y' \geq h(\omega) - Tx - Wy \right\} \neq \emptyset$$

for all $\omega \in \Omega, x \in X, y \in \{0, 1\}^m$.

Suppose that $Y(x, \omega)$ is contained in a disjunctive set $S(x, \omega)$, i.e., with suitable matrices C_l and vectors $\gamma_l, l = 1, \dots, L$, we have

$$Y(x, \omega) \subseteq S(x, \omega) := \cup_{l=1}^L S_l(x, \omega)$$

with nonempty sets

$$S_l(x, \omega) := \left\{ (y, y') \in [0, 1]^{\tilde{m}} \times \mathbb{R}_+^{m'} : Wy + W'y' \geq h(\omega) - Tx, C_l(y, y') \geq \gamma_l \right\}.$$

For convenience we rewrite

$$S_l(x, \omega) = \left\{ \tilde{y} \in \mathbb{R}_+^{\tilde{m}} : \tilde{W}_l \tilde{y} \geq r_l(x, \omega) \right\}.$$

An inequality $\pi^\top \tilde{y} \geq \pi_o$ then is valid for $S(x, \omega)$ if and only if for all $l = 1, \dots, L$

$$\begin{aligned} \pi_o &\leq \min \left\{ \pi^\top \tilde{y} : \tilde{W}_l \tilde{y} \geq r_l(x, \omega), \tilde{y} \geq 0 \right\} \\ &= \max \left\{ \lambda^\top r_l(x, \omega) : \lambda^\top \tilde{W}_l \leq \pi, \lambda \geq 0 \right\} \end{aligned}$$

if and only if for all $l = 1, \dots, L$ there exist $\lambda_l \geq 0$ such that

$$\lambda_l^\top \tilde{W}_l \leq \pi \quad \text{and} \quad \lambda_l^\top r_l(x, \omega) \geq \pi_o. \tag{24}$$

Now fix $(\bar{x}, \bar{\omega})$, and consider π, λ_l, π_o fulfilling (24). Then, for *arbitrary* $(x, \omega) \in X \times \Omega$ the inequality $\pi^\top \tilde{y} \geq \pi_o(x, \omega)$ with

$$\pi_o(x, \omega) := \min_{l=1, \dots, L} \lambda_l^\top r_l(x, \omega) \tag{25}$$

is valid for $S(x, \omega)$.

Hence, valid inequalities can be formulated with common cut coefficients and with individual right-hand sides that are adjusted according to (25) for each (x, ω) . Following the theory of disjunctive cutting planes the cut coefficient vector π and the right-hand side $\pi_o(x, \omega)$ can be computed by solving a suitable linear program.

One observes that $\pi_o(\cdot, \omega)$ is a concave function on X . Facets of the convex hull of its epigraph

$$\text{epi } \pi_o(\cdot, \omega) := \{(\alpha, x) : \alpha \geq \pi_o(x, \omega), x \in X\}$$

can be computed by starting from suitable disjunctions and solving linear programs, for details see [56]. This leads to an explicit convex hull approximation $\pi_c(\cdot, \omega)$ of $\pi_o(\cdot, \omega)$ that can be represented as

$$\pi_c(x, \omega) = \max_{i=1, \dots, I} \left\{ \varepsilon_i(\omega) + \delta_i(\omega)^\top x \right\}, \quad (x, \omega) \in X \times \Omega.$$

For all $(x, \omega) \in X \times \Omega$ we obtain

$$\pi^\top \tilde{y} \geq \pi_o(x, \omega) \quad \text{iff} \quad \pi^\top \tilde{y} \geq \varepsilon_i(\omega) + \delta_i(\omega)^\top x \quad \text{for all } i = 1, \dots, I.$$

This relation is essential for convexifying the feasible sets $Y(x, \omega)$ and thus for updating LP relaxations that come closer and closer to the single-scenario problems in the stochastic integer program. The dual solutions to these LP relaxations provide cuts that, analogously to those in (21), successively shape the (convex hull) of the objective of the stochastic integer program. Again it is crucial that the functional dependence of these cuts on (x, ω) is available explicitly.

The above considerations lead to a decomposition algorithm for solving (19) that is derived in detail in [56]. Here we will only sketch key ingredients and basic steps. The problem to be solved can be represented as

$$\min \left\{ c^\top x + \sum_{\omega \in \{1, \dots, J\}} p(\omega) \Phi(h(\omega) - Tx) : x \in X \right\} \quad (26)$$

where Φ is as in (23). Consider the LP relaxation to the single-scenario problem

$$\min \left\{ q^\top y + q'^\top y' : Wy + W'y' \geq h(\omega) - Tx, y \in [0, 1]^{\tilde{m}}, y' \in \mathbb{R}_+^{m'} \right\}, \quad \omega \in \Omega,$$

and rewrite it as

$$\min \left\{ \tilde{q}^\top \tilde{y} : \tilde{W}\tilde{y} \geq \tilde{h}(\omega) - \tilde{T}x, \tilde{y} \in \mathbb{R}_+^{\tilde{m}} \right\}.$$

STEP 0 – INITIALIZATION.

Let $v := 1$, $x^v \in X$, $W^v := \tilde{W}$, $h^v(\omega) := \tilde{h}(\omega)$, $T^v(\omega) := \tilde{T}$, $\mathcal{K}^v(\omega) := \emptyset$, $V^v := +\infty$, $v^v := -\infty$.

STEP 1 – SOLUTION OF LP SUBPROBLEMS.

Find an optimal solution $y^v(\omega)$ of $\min\{\tilde{q}^\top \tilde{y} : W^v \tilde{y} \geq h^v(\omega) - T^v(\omega)x^v, \tilde{y} \in \mathbb{R}_+^{\tilde{m}}\}$ for each $\omega \in \Omega$. If $y^v(\omega)$ meets all integer requirements for all $\omega \in \Omega$ then update $V^v := \min\{c^\top x^v + \sum_{\omega} p(\omega)\Phi(h(\omega) - Tx^v), V^v\}$.

STEP 2 – CUT GENERATION AND UPDATES.

If for some $\omega \in \Omega$ a relevant component $y_i^v(\omega)$ fails to be integer then formulate the disjunction $S_v(x^v, \omega) := S_{oi}(x^v, \omega) \cup S_{li}(x^v, \omega)$ where

$$\begin{aligned} S_{oi}(x^v, \omega) &:= \left\{ \tilde{y} \in \mathbb{R}_+^{\tilde{m}} : W^v \tilde{y} \geq h^v(\omega) - T^v(\omega)x^v, \tilde{y}_i = 0 \right\}, \\ S_{li}(x^v, \omega) &:= \left\{ \tilde{y} \in \mathbb{R}_+^{\tilde{m}} : W^v \tilde{y} \geq h^v(\omega) - T^v(\omega)x^v, \tilde{y}_i = 1 \right\}. \end{aligned}$$

According to the procedure outlined above, this disjunction leads to a coefficients vector π^v and to quantities $\varepsilon_i^v(\omega)$, $\delta_i^v(\omega)$ such that the inequality $\pi^{v\top} \tilde{y} \geq \varepsilon_i^v(\omega) + \delta_i^v(\omega)^\top x$ is valid for the single-scenario subproblem belonging to ω , for all $x \in X$. Both π^v and $\varepsilon_i^v(\omega)$, $\delta_i^v(\omega)$ are computed by solving suitable linear programs. The procedure is carried out for all $\omega \in \Omega$ for which a relevant component of $y^v(\omega)$ fails an integer requirement. As a benefit that is specific to our two-stage stochastic programming setting we have that the coefficients vector π^v does not depend on ω and has to be computed only once, whereas individual LPs have to be solved for the $(\varepsilon_i^v(\omega), \delta_i^v(\omega))$, $\omega \in \Omega$.

Subsequently, the cut $\pi^{v\top} \tilde{y} \geq \varepsilon_i^v(\omega) + \delta_i^v(\omega)^\top x$ is appended to the system $W^v \tilde{y} \geq h^v(\omega) - T^v(\omega)x$ such that the updated quantities W^{v+1} , $h^{v+1}(\omega)$, and $T^{v+1}(\omega)$ arise. If $\omega \in \Omega$ is such that $y^v(\omega)$ fulfils all integer requirements then $W^{v+1} := W^v$, $h^{v+1}(\omega) := h^v(\omega)$, and $T^{v+1}(\omega) := T^v(\omega)$.

STEP 3 – SOLUTION OF UPDATED LP SUBPROBLEMS.

For each $\omega \in \Omega$, solve $\min\{\tilde{q}^\top \tilde{y} : W^{v+1} \tilde{y} \geq h^{v+1}(\omega) - T^{v+1}(\omega)x^v, \tilde{y} \in \mathbb{R}_+^{\tilde{m}}\}$, and let $d^v(\omega)$ be an optimal dual multiplier. Update $\mathcal{K}^{v+1}(\omega) := \mathcal{K}^v(\omega) \cup \{d^v(\omega)\}$.

STEP 4 – SOLUTION OF THE MASTER PROBLEM.

Solve

$$\min \left\{ c^\top x + \sum_{\omega} p(\omega) \max_{d(\omega) \in \mathcal{K}^{\nu+1}(\omega)} d(\omega)^\top (h^{\nu+1}(\omega) - T^{\nu+1}(\omega)x) : x \in X \right\}, \quad (27)$$

and let $x^{\nu+1}$ be its optimal solution and v^{ν} its optimal value. If $V^{\nu} - v^{\nu}$ drops below some preselected optimality tolerance, then stop. Otherwise, update $\nu := \nu + 1$ and repeat from Step 1.

The above method is a purely cutting plane based approach to solving two-stage stochastic integer programs by decomposition. Its essential feature is that common cut coefficients are shared among different subproblems given by different scenarios and first-stage feasible solutions. The number of these subproblems being quite big in general, sharing cut coefficients is a substantial shortcut.

To guarantee convergence of the method one has to ensure that the approximations given by the objective functions in the master (27) achieve accuracy at least locally around an optimal solution to (26). This requires some theoretical backbone that complete descriptions of the convex hulls involved can be computed if necessary. For disjunctive cuts this is provided by the fact, see e.g. [5], that the convex hull of a mixed- $\{0, 1\}$ feasible region can be obtained by successively taking the convex hull with respect to one disjunctive variable at a time. In the above method this would correspond to computing in Step 2 all instead of just one of the disjunctive cuts induced by the variable $y_i^{\nu}(\omega)$. But even then, issues such as efficient cut management or moving towards branch-and-cut remain open. Some first results to make the approach more implementable can be found in [56].

The principles behind the above method remain valid for the two-stage mean-risk models from Proposition 2 as long as there is no explicit coupling of second-stage variables from different scenarios, i.e., for the problems (9) and (10).

6. Further issues

We conclude with a brief discussion of further integer programming issues relevant for stochastic programming, but, due to space constraints, not covered in more detail in the main body of the text.

Specific models and algorithms. In the present paper the focus has been on some general integer programming principles and their impact in stochastic integer programming. Similar to integer programming, specific models and algorithms have their role in stochastic integer programming, too. Beside the work in power optimization already mentioned in Subsection 4.3 applications were developed in areas such as routing [31, 69] and supply chains [1]. For two-stage models with binary first-stage variables the integer L-shaped method is a finite cutting plane method that can be made efficient for hard problems if lower bounding functionals are available, see [30, 31], and [14, 15] for a mixed-integer extension. Stochastic programs with simple integer recourse [33, 65] can be analyzed in dimension one and are particularly well-understood, both structurally and algorithmically.

Disjunctions – chance constraints. In chance constrained (or probabilistic) programming, constraints involve probabilistic statements about first-stage solutions, see [11, 26, 43]. It is well-known that, for discrete random variables, chance constraints can be expressed as disjunctions. This has given rise to some initial studies, see [17, 55]. First results about chance constrained stochastic programs with integer variables can be found in [18, 50].

Test sets. In recent years, test set methods have regained interest in integer programming. These methods proceed by augmenting feasible points to optimality with the help of a finite test set. In each iteration the test set is screened for an improving vector. If the outcome is positive an improved iterate is created by adding the direction to the current iterate. In the opposite case, the current iterate is optimal. Test sets have strong conceptual relations to lattice bases, monomial bases, and completion procedures from computational algebra. In [22] it is shown that test set vectors belonging to the pure-integer version of the model (8) decompose into building blocks that are determined by the constraint matrix of a single-scenario problem only and that do not depend on the number of scenarios. Building blocks can be computed directly by an algebraic completion procedure, without advance knowledge of the test set. Augmentation to optimality can be accomplished at the building block level already.

Total unimodularity – approximations. Aiming at the interaction between recourse structure and distribution properties, in [67] a lower bound to the expected-value function $Q_{\mathbb{E}}$ from (4) is constructed under the assumption that the optimization problem behind Φ is a pure-integer linear program. The bound can be determined analytically by resorting to principles known from simple integer recourse models. It is motivated by stochastic programs with totally unimodular recourse matrix, for which it is exact. The bound is never worse than the bound obtained by relaxing integrality which makes it attractive for enhancing algorithms working with LP bounds.

Stochastic scheduling. A crucial assumption, tacitly made in all the models we have introduced, is that the probability distribution that underlies the stochastic program does not depend on the decisions. This makes the essential distinction with stochastic scheduling, which addresses a class of integer programs under uncertainty where, typically, scheduling decisions influence the probability distributions of subsequent random events, see [35–38, 64, 70].

Although the present paper is an attempt to give a broad view into stochastic integer programming it is not intended to be comprehensive. For additional reading the surveys [27, 32, 59] and the annotated bibliography [58] are recommended. Resources on the Internet are the COSP website [57] and the bibliography [66].

Acknowledgement. The author wishes to thank Maarten H. van der Vlerk (University of Groningen) for beneficial discussions. Further thanks are due to two anonymous referees for valuable comments that helped improving the final version of this paper.

References

1. Alonso-Ayuso, A., Escudero, L.F., Garín, A., Ortuño, M.T., Pérez, G.: An approach for strategic supply chain planning under uncertainty based on stochastic 0–1 programming. *J. Global Optim.* **26**, 97–124 (2003)
2. Alonso-Ayuso, A., Escudero, L.F., Ortuño, M.T.: BFC, A branch-and-fix coordination algorithmic framework for solving some types of stochastic pure and mixed 0–1 programs. *Eur. J. Operational Res.* (to appear)
3. Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. *Math. Finance* **9**, 203–228 (1999)
4. Balas, E.: Disjunctive programs: cutting planes from logical conditions. In: O.L. Mangasarian, R.R. Meyer, S.M. Robinson (eds.) *Nonlinear Programming*, Vol. 2 (Academic Press, New York 1975) 279–312
5. Balas, E.: Disjunctive programming. *Ann. Discrete Math.* **5**, 3–31 (1979)
6. Balas, E., Ceria, S., Cornuéjols, G.: A lift-and-project cutting plane algorithm for mixed 0–1 programs. *Math. Programming* **58**, 295–324 (1993)
7. Bank, B., Mandel, R.: *Parametric Integer Optimization* (Akademie-Verlag, Berlin 1988)
8. Benson, H.P.: Multi-objective optimization: Pareto optimal solutions, properties. In: C.A. Floudas, P.M. Pardalos (eds.) *Encyclopedia of Optimization*, Vol. III (Kluwer, Dordrecht 2001) 489–493
9. Boreau, B.: Programme de risque minimal en programmation linéaire stochastique. *Comptes Rendus de l'Académie des Sciences Paris* **259**, 981–983 (1964)
10. Boreau, B.: Minimum risk criterion in stochastic optimization. *Econ. Computation and Econ. Cybern. Studies and Res.* **2**, 31–39 (1981)
11. Birge, J.R., Louveaux, F.: *Introduction to Stochastic Programming* (Springer, New York 1997)
12. Blair, C.E., Jeroslow, R.G.: The value function of a mixed integer program: I, *Discrete Math.* **19**, 121–138 (1977)
13. Carøe, C.C., Schultz, R.: Dual decomposition in stochastic integer programming. *Operations Res. Lett.* **24**, 37–45 (1999)
14. Carøe, C.C., Tind, J.: A cutting plane approach to mixed 0–1 stochastic integer programs. *Eur. J. Operational Res.* **101**, 306–316 (1997)
15. Carøe, C.C., Tind, J.: *L*-shaped decomposition of two-stage stochastic programs with integer recourse. *Math. Programming* **83**, 139–152 (1998)
16. Charnes, A., Cooper, W.W.: Deterministic equivalents for optimizing and satisficing under chance constraints. *Operations Res.* **11**, 18–38 (1963)
17. Dentcheva, D., Prékopa, A., Ruszczyński, A.: Concavity and efficient points for discrete distributions in stochastic programming. *Math. Programming* **89**, 55–79 (2000)
18. Dentcheva, D., Prékopa, A., Ruszczyński, A.: Bounds for integer stochastic programs with probabilistic constraints. *Discrete Appl. Math.* **124**, 55–65 (2002)
19. Dentcheva, D., Römisich, W.: Duality gaps in nonconvex stochastic optimization, Preprint 02–5, Institut für Mathematik, Humboldt-Universität zu Berlin, 2002. <http://www.speps.info>
20. Gröwe-Kuska, N., Kiwiel, K.C., Nowak, M.P., Römisich, W., Wegner, I.: Power management in a hydrothermal system under uncertainty by Lagrangian relaxation. In: C. Greengard, A. Ruszczyński (eds.) *Decision Making under Uncertainty: Energy and Power*, IMA Volumes in Mathematics and its Applications, Vol. 128 (Springer, New York 2002) 39–70
21. Helmberg, C., Kiwiel, K.C.: A spectral bundle method with bounds. *Math. Programming* **93**, 173–194 (2002)
22. Hemmecke, R., Schultz, R.: Decomposition of test sets in stochastic integer programming. *Math. Programming* **94**, 323–341 (2003)
23. Higle, J.L., Sen, S.: *Stochastic Decomposition: A Statistical Method for Large Scale Stochastic Linear Programming* (Kluwer, Dordrecht 1996)
24. Hiriart-Urruty, J.B., Lemaréchal, C.: *Convex Analysis and Minimization Algorithms* (Springer, Berlin 1993)
25. Jeroslow, R.G.: Cutting plane theory: disjunctive methods. *Ann. Discrete Math.* **1**, 293–330 (1977)
26. Kall, P., Wallace, S.W.: *Stochastic Programming* (Wiley, Chichester 1994)
27. Klein Haneveld, W.K., van der Vlerk, M.H.: Stochastic integer programming: General models and algorithms. *Ann. Operations Res.* **85**, 39–57 (1999)
28. Kiwiel, K. C.: Proximity control in bundle methods for convex nondifferentiable optimization. *Math. Programming* **46**, 105–122 (1990)
29. Korhonen, P.: Multiple objective programming support. In: C.A. Floudas, P.M. Pardalos (eds.) *Encyclopedia of Optimization*, Vol. III (Kluwer, Dordrecht 2001) 566–574
30. Laporte, G., Louveaux, F.V.: The integer *L*-shaped method for stochastic integer programs with complete recourse. *Operations Res. Lett.* **13**, 133–142 (1993)

31. Laporte, G., Louveaux, F.V., Van Hamme, L.: An integer L -shaped algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Res.* **50**, 415–423 (2002)
32. Louveaux, F.V., Schultz, R.: Stochastic integer programming. In: A. Ruszczyński, A. Shapiro (eds.) *Handbooks in Operations Research and Management Science*, 10: Stochastic Programming (Elsevier Science, Amsterdam) (to appear)
33. Louveaux, F.V., van der Vlerk, M.H.: Stochastic programming with simple integer recourse. *Math. Programming* **61**, 301–325 (1993)
34. Lulli, G., Sen, S.: A branch-and-price algorithm for multi-stage stochastic integer programming with application to stochastic batch-sizing problems. *Stochastic Programming E-Print Series (SPEPS)*, 17-2002. <http://www.speps.info>
35. Möhring, R.H.: Scheduling under uncertainty: Optimizing against a randomizing adversary. In: K. Jansen, S. Khuller (eds.) *Approximation Algorithms for Combinatorial Optimization* (Springer, Berlin 2000) 15–26
36. Möhring, R.H., Radermacher, F.J., Weiss, G.: Stochastic scheduling problems I – General strategies. *Z. Operations Res. Ser. A* **28**, 193–260 (1984)
37. Möhring, R.H., Radermacher, F.J., Weiss, G.: Stochastic scheduling problems II – Set strategies. *Z. Operations Res. Ser. A* **29**, 65–104 (1985)
38. Möhring, R.H., Schulz, A.S., Uetz, M.: Approximation in stochastic scheduling: the power of LP-based priority policies. *J. Association for Computing Machinery* **46**, 924–942 (1999)
39. Müller, A., Stoyan, D.: *Comparison Methods for Stochastic Models and Risks* (Wiley, New York 2002)
40. Nemhauser, G.L., Wolsey, L.A.: *Integer and Combinatorial Optimization* (Wiley, New York 1988)
41. Nowak, M.P., Römisich, W.: Stochastic Lagrangian relaxation applied to power scheduling in a hydro-thermal system under uncertainty. *Ann. Operations Res.* **100**, 251–272 (2000)
42. Ogryczak, W., Ruszczyński, A.: Dual stochastic dominance and related mean-risk models. *SIAM J. Optim.* **13**, 60–78 (2002)
43. Prékopa, A.: *Stochastic Programming* (Kluwer, Dordrecht 1995)
44. Riis, M., Schultz, R.: Applying the minimum risk criterion in stochastic recourse programs. *Computational Optim. Applications* **24**, 267–287 (2003)
45. Rockafellar, R.T., Uryasev, S.: Conditional value-at-risk for general loss distributions. *J. Banking & Finance* **26**, 1443–1471 (2002)
46. Römisich, W.: Stability of stochastic programming problems. In: A. Ruszczyński, A. Shapiro (eds.) *Handbooks in Operations Research and Management Science*, 10: Stochastic Programming (Elsevier Science, Amsterdam) (to appear)
47. Römisich, W., Schultz, R.: Multistage stochastic integer programs: an introduction. In: M. Grötschel, S.O. Krumke, J. Rambau (eds.) *Online Optimization of Large Scale Systems* (Springer, Berlin 2001) 581–600
48. Rosa, C.H., Ruszczyński, A.: On augmented Lagrangian decomposition methods for multistage stochastic programs. *Ann. Operations Res.* **64**, 289–309 (1996)
49. Ruszczyński, A.: A regularized decomposition method for minimizing a sum of polyhedral functions. *Math. Programming* **35**, 309–333 (1986)
50. Ruszczyński, A.: Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. *Math. Programming* **93**, 195–215 (2002)
51. Ruszczyński, A., Shapiro, A. (eds.): *Handbooks in Operations Research and Management Science*, 10: Stochastic Programming (Elsevier Science, Amsterdam) (to appear)
52. Schultz, R.: On structure and stability in stochastic programs with random technology matrix and complete integer recourse. *Math. Programming* **70**, 73–89 (1995)
53. Schultz, R.: Some aspects of stability in stochastic programming. *Ann. Operations Res.* **100**, 55–84 (2000)
54. Schultz, R., Tiedemann, S.: Risk aversion via excess probabilities in stochastic programs with mixed-integer recourse. *SIAM Journal on Optimization* (to appear). previous version <http://www.speps.info>
55. Sen, S.: Relaxations for probabilistically constrained programs with discrete random variables. *Operations Res. Lett.* **11**, 81–86 (1992)
56. Sen, S., Hige, J.L.: The C^3 theorem and a D^2 algorithm for large scale stochastic integer programming: set convexification, Preprint, SIE Department, University of Arizona, Tucson, 2000. <http://www.speps.info>
57. Stochastic Programming Community Homepage, sponsored by the Committee on Stochastic Programming (COSP) of the Mathematical Programming Society, <http://stoprog.org>
58. Stougie, L., van der Vlerk, M.H.: Stochastic integer programming. In: M. Dell’Amico, F. Maffioli, S. Martello (eds.) *Annotated Bibliographies in Combinatorial Optimization* (Wiley, Chichester 1997) 127–141
59. Stougie, L., van der Vlerk, M.H.: Approximations in stochastic integer programming. In: D.S. Johnson, J.K. Lenstra, D.B. Shmoys (eds.) *Handbooks in Operations Research and Management Science*, Approximation and Heuristics (Elsevier Science, Amsterdam) (to appear)

60. Takriti, S., Ahmed, S.: On robust optimization of two-stage systems. *Math. Programming* (to appear)
61. Takriti, S., Birge, J.R.: Lagrangian solution techniques and bounds for loosely coupled mixed-integer stochastic programs. *Operations Res.* **48**, 91–98 (2000)
62. Takriti, S., Birge, J.R., Long, E.: A stochastic model for the unit commitment problem. *IEEE Trans. on Power Systems* **11**, 1497–1508 (1996)
63. Takriti, S., Krasenbrink, B., Wu, L.S.-Y.: Incorporating fuel constraints and electricity spot prices into the stochastic unit commitment problem. *Operations Res.* **48**, 268–280 (2000)
64. Uetz, M.: *Algorithms for Deterministic and Stochastic Scheduling* (Cuvillier, Göttingen 2002)
65. van der Vlerk, M.H.: *Stochastic programming with integer recourse* (Labyrinth Publication, Capelle a/d IJssel 1995)
66. van der Vlerk, M.H.: *Stochastic Programming Bibliography*. <http://mally.eco.rug.nl/spbib.html>, 1996–2003
67. van der Vlerk, M.H.: Convex approximations for complete integer recourse models. *Math. Programming* (to appear). <http://www.speps.info>
68. Van Slyke, R.M., Wets, R.J.-B.: L-shaped linear programs with application to optimal control and stochastic programming. *SIAM J. Appl. Math.* **17**, 638–663 (1969)
69. Verweij, B., Ahmed, S., Kleywegt, A.J., Nemhauser, G.L., Shapiro, A.: The sample average approximation method applied to stochastic routing problems: a computational study. *Computational Optim. Applications* **24**, 289–333 (2003)
70. Weiss, G.: A tutorial in stochastic scheduling. In: P. Chrétienne, E.G. Coffman Jr., J.K. Lenstra, Z. Liu (eds.) *Scheduling Theory and its Applications* (Wiley, Chichester 1995) 33–64
71. Woodruff, D.L.: (ed.): *Network Interdiction and Stochastic Integer Programming* (Kluwer, Boston 2002)