Approximation in stochastic integer programming

Stougie, L; van der Vlerk, Maarten H.

Approximation in Stochastic Integer Programming

LEEN STOUGIE* and MAARTEN H. VAN DER VLERK†

April 2003

Mathematics Subject Classification: 90C15, 90C11

SOM theme A: *Primary processes within firms*

Electronic version: `http://som.rug.nl`

Abstract

Approximation algorithms are the prevalent solution methods in the field of stochastic programming. Problems in this field are very hard to solve. Indeed, most of the research in this field has concentrated on designing solution methods that approximate the optimal solutions. However, efficiency in the complexity theoretical sense is usually not taken into account. Quality statements mostly remain restricted to convergence to an optimal solution without accompanying implications on the running time of the algorithms for attaining more and more accurate solutions.

However, over the last twenty years also some studies on performance analysis of approximation algorithms for stochastic programming have appeared. In this direction we find both probabilistic analysis and worst-case analysis. There have been studies on performance ratios and on absolute divergence from optimality. Only recently the complexity of stochastic programming problems has been addressed, indeed confirming that these problems are harder than most combinatorial optimization problems.

---

* Eindhoven Technical University and CWI Amsterdam, Netherlands, `leen@win.tue.nl`

† Department of Econometrics & OR, University of Groningen, Netherlands, `m.h.van.der.vlerk@eco.rug.nl` The research of this author has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences (KNAW).

Approximation in the traditional stochastic programming sense will not be discussed in this paper. The reader interested in this issue is referred to surveys on stochastic programming, like the Handbook on Stochastic Programming [31] or the text books [2, 16, 29]. We concentrate on the studies of approximation algorithms which are more similar in nature to those for combinatorial optimization.

Key words: integer recourse, approximation

1. Introduction

Stochastic programming models arise as reformulations or extensions of optimization problems with random parameters. To set the stage for our review of approximation in stochastic (integer) programming, we first introduce the models and give an overview of relevant mathematical properties.

Consider the optimization problem

$$\begin{array}{ll} \min\limits_{x} & cx \\ \text{s.t.} & Ax = b \\ & \boldsymbol{T}x = \boldsymbol{h} \\ & x \in X, \end{array}$$

where $X \subset \mathbb{R}^n$ specifies nonnegativity of and possibly integrality constraints on the decision variables x. In addition to the m_1 deterministic constraints $Ax = b$, there is a set of m constraints $\boldsymbol{T}x = \boldsymbol{h}$, whose parameters $\boldsymbol{T}$ and $\boldsymbol{h}$ depend on information which becomes available only after a decision x is made. The *stochastic programming* approach to such problems is to assume that this uncertainty can be modeled by random variables with known probability distribution, and then to reformulate the model to obtain a meaningful and well-defined optimization problem. In this paper we will use bold face characters for random variables, and plain face to indicate their realizations.

1.1 Stochastic programming models

The first important class of stochastic programming models, known as *recourse models*, is obtained by allowing additional or recourse decisions after observing the realizations of the random variables $(\boldsymbol{T}, \boldsymbol{h})$. Thus, recourse models are dynamic: time is modeled discretely by means of stages, corresponding to the available information. If all uncertainty is dissolved at the same moment, this is captured by a recourse model with two stages: 'present' and 'future'. Given a first-stage decision x, for every possible realization q, T, h of $\boldsymbol{q}, \boldsymbol{T}, \boldsymbol{h}$, infeasibilities $h - Tx$ are compensated at minimal costs by choosing second-stage decisions as an optimal solution of the second-stage problem

$$\begin{array}{ll} \min\limits_{y} & qy \\ \text{s.t.} & Wy = h - Tx, \\ & y \in Y, \end{array}$$

where $\boldsymbol{q}$ is the (random) recourse unit cost vector, the recourse matrix W specifies the available technology, and the set $Y \subset \mathbb{R}^{n_2}_{+}$ is defined analogously to X. We will use $\boldsymbol{\xi} =$

$(\boldsymbol{q}, \boldsymbol{T}, \boldsymbol{h})$ to denote the random object representing all randomness in the problem. The value function of this second-stage problem, specifying the minimal recourse costs as a function of the first-stage decision x and a realization of ξ, will be denoted by $v(x, \xi)$; its expectation $Q(x) := \mathbb{E}_{\boldsymbol{\xi}}[v(x, \boldsymbol{\xi})]$ gives the expected recourse costs associated with a first-stage decision x. Thus, the two-stage recourse model is

$$\begin{array}{ll} \min\limits_{x} & cx + Q(x) \\ \text{s.t.} & Ax = b \\ & x \in X, \end{array} \tag{1}$$

where the objective function $cx + Q(x)$ specifies the total expected costs of a decision x.

Example 1.1 Consider the following production planning problem. Using n production resources, denoted by $x \in \mathbb{R}^n_+$ with corresponding unit cost vector c, a production plan needs to be made such that the uncertain future demand for m products, denoted by $\boldsymbol{h} \in \mathbb{R}^m$, is satisfied at minimal costs. The available production technology suffers from failures: deploying resources x yields uncertain amounts of products $\boldsymbol{T}_i x$, $i = 1, \ldots, m$. Restrictions on the use of x are captured by the constraints $Ax = b$.

We assume that the uncertainty about future demand and the production technology can be modelled by the random vector $(\boldsymbol{T}, \boldsymbol{h})$, whose joint distribution is known, for example based on historical data.

A possible two-stage recourse model for this problem is based on the following extension of the model. For each of the individual products, if the demand h_i turns out to be larger than the production $T_i x$, the demand surplus $h_i - T_i x$ is bought from a competitor at unit costs q^1_i. On the other hand, a demand shortage gives rise to storage costs of q^2_i per unit. The corresponding second-stage problem and its value function are

$$\begin{array}{lll} v(x, \xi) = & \min\limits_{y} & q^1 y^1 + q^2 y^2 \\ & \text{s.t.} & y^1 - y^2 = h - Tx, \qquad \xi \in \Xi, \\ & & y = (y^1, y^2) \in \mathbb{R}^{2m}_+. \end{array}$$

Defining Q as the expectation of this value function, we obtain a two-stage recourse model that fits the general form (1).

This particular model type with recourse matrix $W = (I_m, -I_m)$, where I_m is the m-dimensional identity matrix, is known as a simple recourse model. The integer recourse version of this model, for example corresponding to the case that only batches of fixed size can be bought, will be discussed in Section 3. ◁

So far, we have introduced the recourse concept as a modelling tool to handle random constraints, by means of specifying recourse actions with corresponding recourse costs. There is however another class of problems for which the (two-stage) recourse model is a natural approach, namely *hierarchical planning models* (HPM). Such problems involve decisions at two distinct levels: *strategic* decisions which have a long-term impact, and *operational* decisions which are depending on the strategic decisions. For example, in the hierarchical scheduling problem discussed in Section 4, the strategic decision is the number of machines to be installed, and the operational decisions involve the day-to-day scheduling of jobs on these machines. At the time that the strategic decision needs to be made, only probabilistic information on the operational level problems (e.g. the number of jobs to be scheduled) is available. Hierarchical planning models fit the structure of two-stage recourse models, with strategic and operational decisions corresponding to first-stage and second-stage variables, respectively. Moreover, since strategic decisions are typically fixed for a relatively long period of time, it is natural to use the expectation of the operational costs as a measure of future costs.

Unlike conventional linear recourse models (1), HPM are not necessarily formulated as (mixed-integer) LP problems, see our example in Section 4. Nevertheless, despite these differences in interpretation and formulation, we use the generic name (two-stage) recourse model to refer to both model types, in line with the stochastic programming literature.

In many applications new information becomes available at several distinct moments, say $t = 1, \ldots, H$, where H is the planning horizon. That is, we assume that realizations of random vectors $\boldsymbol{\xi}^t = (\boldsymbol{q}^t, \boldsymbol{T}^t, \boldsymbol{h}^t)$ become known at time t. This can be modelled explicitly in a multistage recourse structure: for each such moment $t = 1, \ldots, H$, a time stage with corresponding recourse decisions is defined. In compact notation, the multistage recourse model is

$$\begin{array}{ll} \min\limits_{x^0} & cx^0 + Q^1(x^0) \\ \text{s.t.} & Ax^0 = b \\ & x^0 \in X, \end{array}$$

where the functions Q^t, $t = 1, \ldots, H$, representing expected recourse costs, are recursively defined as

$$Q^t(x^{t-1}) := \mathbb{E}_{\boldsymbol{\xi}^t}\left[v^t(x^{t-1}, \boldsymbol{\xi}^t) \,\middle|\, \xi^1, \ldots, \xi^{t-1}\right],$$

where the expectation is with respect to the conditional distribution of $\boldsymbol{\xi}^t$ given $\xi^1, \ldots, \xi^{t-1}$,

$$\begin{aligned} v^t(x^{t-1}, \xi^t) := \min_{x^t} \quad & q^t x^t + Q^{t+1}(x^t) \\ \text{s.t.} \quad & W^t x^t = h^t - T^t x^{t-1} \\ & x^t \in X^t, \end{aligned}$$

and $Q^{H+1} \equiv 0$ (or some other suitable choice). In this paper we concentrate on two-stage problems only.

The second main class of stochastic programming problems consists of *probabilistic or chance-constrained problems*, which model random constraints[1] by requiring that they should be satisfied with some prescribed reliability $\alpha \in [0, 1]$; typically, $\alpha \in (.5, 1)$. Thus, the random constraints $\boldsymbol{T}x \geq \boldsymbol{h}$ are replaced by the *joint chance constraint*

$$\Pr\{\boldsymbol{T}x \geq \boldsymbol{h}\} \geq \alpha,$$

or by *m individual chance constraints*

$$\Pr\{\boldsymbol{T}_i x \geq \boldsymbol{h}_i\} \geq \alpha_i, \quad i = 1, \ldots, m.$$

Since we will not consider chance-constrained models in our discussion of approximation results, we do not present them in more detail here.

1.2 Mathematical properties

In this section, we review mathematical properties of recourse models. This provides the background and motivation for the discussion of approximation results.

First we consider properties of continuous recourse models. Some of the results will be used when we discuss the complexity of this problem class, and furthermore they facilitate the subsequent discussion of properties of mixed-integer recourse models. We state all properties here without proof. In the Notes at the end of the paper references to the proofs are given.

Remark 1.1 As before, all models are discussed here in their canonical form, i.e., all constraints are either equalities or nonnegativities. The models in subsequent sections, which also contain inequalities and/or simple bounds, can be written in canonical form using standard transformations.

---

1 Barring uninteresting cases, chance constraints make sense only for inequality constraints.

1.2.1 Continuous recourse

Properties of (two-stage) recourse models follow from those of the recourse function Q. In case all second-stage variables are continuous, properties of the value function v are well-known from duality and perturbation theory for linear programming, and are summarized here for easy reference.

Lemma 1.1 *The function v, defined for $x \in \mathbb{R}^n$ and $\xi = (q, T, h) \in \mathbb{R}^{n_2+m(n+1)}$,*

$$v(x, \xi) = \inf \left\{ qy : Wy = h - Tx,\ y \in \mathbb{R}^{n_2}_+ \right\}$$

takes values in $[-\infty, \infty]$.

It is a convex polyhedral function of x for each $\xi \in \mathbb{R}^{n_2+m(n+1)}$, and it is concave polyhedral in q and convex polyhedral in (h, T) for all $x \in \mathbb{R}^n$. □

If for some x the function v takes on the value $+\infty$ with positive probability, this means that x is extremely unattractive since it has infinitely high expected recourse costs $Q(x)$. From a modelling point of view this is not necessarily a problem, but in practice it may be desirable to exclude this situation.

On the other hand, the situation that $v(x, \xi)$ equals $-\infty$ with positive probability should be excluded altogether. Indeed, the value $-\infty$ indicates that the model does not adequately represent our intention, which is penalization of infeasibilities.

Finiteness of v is often guaranteed by assuming that the recourse is *complete* and *sufficiently expensive.*

Definition 1.1 The recourse is *complete* if $v < +\infty$, i.e., if for all $t \in \mathbb{R}^m$ there exists a $y \in Y$ such that $Wy = t$.

Assuming that $Y = \mathbb{R}^{n_2}_+$, completeness is a property of the recourse matrix W only. Such a matrix is called a *complete recourse matrix.*

Definition 1.2 The recourse is *sufficiently expensive* if $v > -\infty$ with probability 1, i.e., if $\Pr\{\xi \in \Xi : \exists \lambda \in \mathbb{R}^m \text{ such that } \boldsymbol{q} \geq \lambda W\} = 1$.

For example, the recourse is sufficiently expensive if $\Pr\{\boldsymbol{q} \geq 0\} = 1$.

From now on we assume that the recourse is complete and sufficiently expensive. Then the recourse or expected value function $Q(x)$ is finite if the distribution of ξ satisfies the following condition:

For all i, j, k the random functions $\boldsymbol{q}_j \boldsymbol{h}_i$ and $\boldsymbol{q}_j \boldsymbol{T}_{ik}$ have finite expectations.

Sufficiency of this *weak covariance condition* follows from the representation of basic feasible solutions in terms of the problem parameters.

The following properties of the recourse function Q are inherited from the second-stage value function v.

Theorem 1.1 *Consider the continuous recourse function Q, defined by*

$$Q(x) = \mathbb{E}_{\boldsymbol{\xi}} \left[\inf \left\{ \boldsymbol{q} y : W y = \boldsymbol{h} - \boldsymbol{T} x,\ y \in \mathbb{R}^{n_2}_+ \right\} \right], \quad x \in \mathbb{R}^n.$$

Assume that the recourse is complete and sufficiently expensive.

(a) The function Q is convex, finite, and (Lipschitz) continuous.

(b) If $\boldsymbol{\xi}$ follows a finite discrete distribution, then Q is a convex polyhedral function.

(c) The function Q is subdifferentiable, with subdifferential

$$\partial Q(x) = \int_{\Xi} \partial v(x, \xi)\, dF(\xi), \quad x \in \mathbb{R}^n,$$

where F is the cdf of the random vector $\boldsymbol{\xi}$.

If $\boldsymbol{\xi}$ follows a continuous distribution, then Q is continuously differentiable.

□

Consider the special case that $\boldsymbol{\xi}$ follows a finite discrete distribution specified by $\Pr\{\boldsymbol{\xi} = (q^k, T^k, h^k)\} = p^k, k = 1, \ldots, K$. The finitely many possible realizations (q^k, T^k, h^k) of the random parameters are also called *scenarios*. It is easy to see that in this case the two-stage recourse model is equivalent to the large-scale linear programming problem

$$\begin{array}{lllll} \max & cx & + & \sum_{k=1}^{K} p^k q^k y^k & \\ \text{s.t.} & Ax & & & = b \\ & T^k x & + & W y^k & = h^k,\ k = 1, \ldots, K \\ & x \in \mathbb{R}^n_+, & & y^k \in \mathbb{R}^{n_2}_+. & \end{array} \tag{2}$$

Analogously, a mixed-integer recourse problem with finite discrete distribution can be represented as a deterministic large-scale mixed-integer programming problem.

1.2.2 Mixed-integer recourse

Mixed-integer recourse models do not posses such nice mathematical properties; in particular, convexity of the recourse function Q is not guaranteed. Indeed, the underlying second-stage value function v is only lower semicontinuous (assuming rationality of the recourse matrix W), and discontinuous in general.

Also in this setting we are mostly interested in the case that v is finite. To have $v < +\infty$ we will assume complete recourse, see Definition 1.1. For example, this condition is satisfied if $\bar{W}$ is a complete recourse matrix, where $\bar{W}$ consists of the columns of W corresponding to the continuous second-stage variables. On the other hand, $v > -\infty$ if the recourse is sufficiently expensive, see Definition 1.2, i.e., if the dual of the LP relaxation of the second-stage problem is feasible with probability 1.

Theorem 1.2 *Consider the mixed-integer recourse function Q, defined by*

$$Q(x) = \mathbb{E}_{\boldsymbol{\xi}}\left[\inf\{qy : Wy = \boldsymbol{h} - \boldsymbol{T}x,\ y \in Y\}\right], \quad x \in \mathbb{R}^n,$$

where $Y := \mathbb{Z}_+^p \times \mathbb{R}_+^{n_2-p}$. Assume that the recourse is complete and sufficiently expensive, and that $\boldsymbol{\xi} = (\boldsymbol{h}, \boldsymbol{T})$ satisfies a weak covariance condition. Then

(a) The function Q is lower semicontinuous on $\mathbb{R}^n$.

(b) Let $D(x)$, $x \in \mathbb{R}^n$, denote the set containing all $\xi \in \Xi$ such that $h - Tx$ is a discontinuity point of the mixed-integer value function v. Then Q is continuous at x if $\Pr\{\boldsymbol{\xi} \in D(x)\} = 0$.
In particular, if $\boldsymbol{\xi}$ is continuously distributed, then Q is continuous on $\mathbb{R}^n$. □

1.3 Outline

As mentioned above, solving stochastic programming problems is very difficult in general. Indeed, such problems are defined in terms of expectations of value functions of linear (mixed-integer) programming problems or indicator functions (in the case of chance constraints). This calls for the evaluation of multi-dimensional integrals, which is computationally challenging already if the underlying random vector ω has low dimension, and becomes a formidable task for problems of realistic size. Even if the underlying distribution is discrete, the typically huge number of possible realizations may render the frequent evaluation of function values impracticable. In Section 2 the computational complexity of two-stage recourse models is addressed.

It is therefore not surprising that much of the stochastic programming literature is de-

voted to approximation of some sorts. For example, a key issue for recourse models is the construction of suitable discrete approximations of the distribution of the underlying random vector. Such an approximation should have a relatively small number of possible realizations, and at the same time result in a good approximation of the recourse function, at least in a neighborhood of an optimal solution. For chance-constrained problems such discrete approximations of the distribution would destroy convexity of the problem. In this context, fast and accurate approximation of high-dimensional (normal) distribution functions receives much research attention.

We do not discuss these 'typical' stochastic programming approximation issues here. They, as well as related subjects such as convergence and stability, are covered in the Handbook on Stochastic Programming [31]. Instead, we consider approximations as they appear in a number of other ways in stochastic programming and which are in spirit closer to approximation in combinatorial optimization.

Section 3 deals with convex approximations for integer recourse problems. Here the problems themselves are approximated by perturbing the distribution functions such as to achieve convex expected value functions. The strength of this approximation is that a bound on the absolute approximation error can be given, making this an example of *worst-case analysis of approximation algorithms.*

Hierarchical planning problems, which are (integer) recourse problems, are discussed in Section 4. The key idea here is to replace hard second-stage problems by easier ones, which asymptotically still give accurate results. Here the approach is *probabilistic analysis of approximation algorithms.*

In Section 5 we will give one of the scarce examples of an approximation algorithm for a stochastic programming problem for which a constant *worst-case performance ratio* can be proved. The example also shows again that stochastic programming problems are usually more complicated than their deterministic counterparts.

We conclude with a section containing bibliographical notes on approximation in stochastic programming as reviewed in this paper. It also addresses some interesting open problems and new research directions in this field, major parts of which are still unexplored.

2. Complexity of two-stage stochastic programming problems

In this section we study the complexity of two-stage stochastic programming problems. The complexity of a problem, in terms of time or space to solve it, is related to input

size. For each instance a bound on the number of elementary computer operations or on the number of computer storage units required to solve the problem instance as a function of the size of its input indicates, respectively, the time or space complexity of the problem. We will see that the way in which the random parameters in stochastic programming problems are described has a crucial impact on the complexity.

To illustrate this we start by studying problem (2), the deterministic equivalent LP formulation of the two-stage stochastic programming problem.

If in the input of the problem each scenario (q^k, T^k, h^k) and its corresponding probability p^k is specified *separately*, then the input size of the problem is just the size of the binary encoding of all the parameters in this (large-scale) deterministic equivalent problem and hence the problem is polynomially solvable in case the decision variables are continuous and NP-complete if there are integrality constraints on decision variables.

However, consider another extreme in which all parameters are independent identically distributed random variables. For example, if in this case each parameter has value a_1 with probability p and a_2 with probability $1-p$, then there are $K = 2^{n_1+mn+m}$ possible scenarios. Hence, the size of the deterministic equivalent problem is exponential in the dimension of the parameter space, which is essentially the size required to encode the input. The complexity changes correspondingly, as will become clear below.

Let us consider models wherein all random (second-stage) parameters are independently and discretely distributed. We will establish $\sharp P$-hardness of the evaluation of the second-stage expected value function $Q(x)$ for fixed x. The class $\sharp P$ consists of counting problems, for which membership to the set of items to be counted can be decided in polynomial time. We notice that strictly following this definition of $\sharp P$, none of the stochastic programming problems can belong to this complexity class. We will use the term $\sharp P$-hard for an optimization problem in the same way as NP-hardness is used for optimization problems, whose recognition version is NP-complete. For an exposition of the definitions and structures of the various complexity classes we refer to [28].

To prove $\sharp P$-hardness of the evaluation of the second stage expected value function $Q(x)$ we use a reduction from the $\sharp P$-complete problem GRAPH RELIABILITY.

Definition 2.1 GRAPH RELIABILITY. Given a directed graph with m arcs and n vertices, what is the probability that the two given vertices u and v are connected if all edges fail independently with probability 1/2 each.

This is equal to the problem of counting the number of subgraphs, from among all 2^m possible subgraphs, that contain a path from u to v.

Theorem 2.1 *Two-stage stochastic programming with discretely distributed parameters is $\sharp P$-hard.*

PROOF. That the problem is $\sharp P$-*easy* can be seen from the fact that for any realization of the second-stage random parameters a linear program remains to be solved.

To prove $\sharp P$-hardness, take any instance of GRAPH RELIABILITY, i.e., a network $G = (V, A)$ with two prefixed nodes u and v in V. Introduce an extra arc from v to u, and introduce for each arc $(i, j) \in A$ a variable y_{ij}. Give each arc a random weight $\mathbf{q}_{ij}$ except for the arc (v, u) that gets weight 1. Let the weights be independent and identically distributed (i.i.d.) with distribution $Pr\{\mathbf{q} = -2\} = Pr\{\mathbf{q} = 0\} = 1/2$. Denote $A' = A \cup (v, u)$. Now define the two-stage stochastic programming problem

$$\max\{-cx + Q(x) \mid 0 \leq x \leq 1\}$$

with $Q(x) = \mathbb{E}_{\boldsymbol{q}}[v(x, \boldsymbol{q})]$ and

$$\begin{aligned} v(x, \boldsymbol{q}) = \quad & \max \quad \sum_{(i,j)\in A} \mathbf{q}_{ij} y_{ij} + y_{vu} \\ & \text{s.t.} \quad \sum_{i:(i,j)\in A'} y_{ij} - \sum_{k:(j,k)\in A'} y_{jk} = 0 \quad \forall j \in V \\ & \qquad y_{ij} \leq x \qquad\qquad\qquad\qquad \forall (i, j) \in A. \end{aligned}$$

The event $\{q = -2\}$ corresponds to failure of the arc in the GRAPH RELIABILITY instance. For a realization of the failures of the arcs, the network has a path from u to v if and only if in the corresponding realization of the weights there exists a path from u to v consisting of arcs with weight 0. The latter accounts for an optimal solution value x of the corresponding realization of the second-stage problem, obtained by setting all y_{ij}'s corresponding to arcs (i, j) on this path and y_{vu} equal to x, whereas $y_{ij} = 0$ for all (i, j) not on the path. If for a realization the graph does not have a path from u to v, implying in the reduced instance that on each path there is an arc with weight -2 and vice versa, then the optimal solution of the realized second-stage problem is 0, by setting all y_{ij}'s equal to 0, and henceforth also $y_{vu} = 0$). Therefore, the network has reliability R if and only if $Q(x) = Rx$ and hence the objective function of the two-stage problem is $(R - c)x$.

Thus, if $c \leq R$ then the optimal solution is $x = 1$ with value $(R - c)$, and if $c \geq R$ then

the optimal solution is $x = 0$ with value 0. Since R can take only 2^n possible values, bisection allows to solve only m two-stage stochastic programming problems to know the exact value of R. □

By total unimodularity of the restriction coefficients matrix in the proof, the same reduction shows that two-stage integer programming problem with discretely distributed parameters is $\sharp P$-hard.

Given a $\sharp P$-oracle for evaluating Q in any point x, solving two-stage stochastic linear programming problems (with discretely distributed random variables) will require a polynomial number of consultations of the oracle, since Q is a concave function in x, and maximizing a concave function over a convex set is known to be easy [26]. Thus, two-stage stochastic linear programming is in the class $P^{\sharp P} = \sharp P$.

Assuming a $\sharp P$-oracle for evaluating Q in any point x of a two-stage stochastic integer programming problem, makes the decision version of this problem a member of NP. The function Q is not convex in this case, but there are a finite number of points x that are candidate for optimality. Thus, the decision version of two-stage stochastic integer programming is in the class $NP^{\sharp P}$.

In case the random parameters of the two-stage stochastic programming problem are continuously distributed, the evaluation of the function Q in a single point of its domain requires the computation of a multiple integral. Most of the stochastic programming literature on this subclass of problems is concerned with how to get around this obstacle. We give the complexity of this class of problems without proof.

Theorem 2.2 *Two-stage stochastic programming problems with continuously distributed parameters is $\sharp P$-hard, even if all stochastic parameters have the uniform* $[0, 1]$ *distribution.* □

The membership of this problem in $\sharp P$ requires additional conditions on the input distributions, since exact computation may not even be in *PSPACE.*

3. Convex approximations for integer recourse problems

In this section we consider convex approximations for pure integer recourse models. For such problems, the second-stage problem is necessarily defined using only inequal-

ities. Moreover, in all models considered here only the right-hand side vector $\boldsymbol{h}$ is random. The second-stage value function is thus

$$\begin{array}{rll} v(x,\boldsymbol{h}) := & \min\limits_{y} & qy \\ & \text{s.t.} & Wy \geq \boldsymbol{h} - Tx, \qquad x \in \mathbb{R}^n,\ \boldsymbol{h} \in \mathbb{R}^m \\ & & y \in \mathbb{Z}^{n_2}_+, \end{array}$$

where the components of W are assumed to be integers. Assuming complete and sufficiently expensive recourse as before, v is a finite, discontinuous, piecewise constant function; in particular, v is non-convex. It follows from Theorem 1.2 that the integer recourse function $Q(x) = \mathbb{E}_{\boldsymbol{h}}[v(x,\boldsymbol{h})]$, $x \in \mathbb{R}^n$, is continuous if $\boldsymbol{h}$ is continuously distributed, but in general Q is non-convex.

However, for certain integer recourse models, characterized by their recourse matrices W, a class of distributions of $\boldsymbol{h}$ is known such that the corresponding recourse function Q is convex. Thus, for such integer recourse models we can construct convex approximations of the function Q by approximating any given distribution of $\boldsymbol{h}$ by a distribution belonging to this special class.

Below we first apply this approach to the simple integer recourse model. Subsequently, we consider general complete integer recourse models, starting with the case of totally unimodular recourse matrices.

3.1 Simple integer recourse

The simple integer recourse second-stage problem is defined as

$$\begin{array}{ll} \min\limits_{y} & q^+y^+ + q^-y^- \\ \text{s.t.} & y^+ \geq \boldsymbol{h} - Tx, \\ & y^- \geq -(\boldsymbol{h} - Tx), \\ & y^+,\ y^- \in \mathbb{Z}^m_+, \end{array}$$

where the indices $+$ and $-$ are conventionally used to indicate surplus and shortage, respectively. This recourse structure is obviously complete, and it is sufficiently expensive if $q^+ \geq 0$ and $q^- \geq 0$ (componentwise), as will be assumed from now on.

It is trivial to find a closed form for the simple integer recourse value function. Due to the simple recourse structure, this function is separable in the *tender variables* $z := Tx$:

$$v(z,\boldsymbol{h}) = \sum_{i=1}^{m} v_i(z_i,\boldsymbol{h}_i), \qquad z,\boldsymbol{h} \in \mathbb{R}^m,$$

where

$$v_i(z_i, \boldsymbol{h}_i) = q_i^+ \lceil \boldsymbol{h}_i - z_i \rceil^+ + q_i^- \lfloor \boldsymbol{h}_i - z_i \rfloor^-, \tag{3}$$

with $\lceil s \rceil^+ := \max\{0, \lceil s \rceil\}$ and $\lfloor s \rfloor^- := \max\{0, -\lfloor s \rfloor\}$, $s \in \mathbb{R}$. Since all functions v_i have the same structure, we restrict the presentation to one such function, and drop the index. It is straightforward to translate the results below back to the full-dimensional case.

Given the closed form (3), it follows that the one-dimensional generic simple integer recourse function Q equals

$$Q(z) = q^+ \mathbb{E}_{\boldsymbol{h}} \left[\lceil \boldsymbol{h} - z \rceil^+ \right] + q^- \mathbb{E}_{\boldsymbol{h}} \left[\lfloor \boldsymbol{h} - z \rfloor^- \right], \quad z \in \mathbb{R}, \tag{4}$$

where $\boldsymbol{h} \in \mathbb{R}$ is a random variable. Throughout we assume that $\mathbb{E}_{\boldsymbol{h}} [|\boldsymbol{h}|]$ is finite, which is necessary and sufficient for finiteness of the function Q.

Lemma 3.1 *Consider the one-dimensional simple integer recourse function Q defined in (4).*

(a) For all $z \in \mathbb{R}$,

$$Q(z) = q^+ \sum_{k=0}^{\infty} \Pr\{\boldsymbol{h} > z + k\} + q^- \sum_{k=0}^{\infty} \Pr\{\boldsymbol{h} < z - k\}.$$

(b) Assume that $\boldsymbol{h}$ has a pdf f that is of bounded variation. Then the right derivative Q'_+ exists everywhere:

$$Q'_+(z) = -q^+ \sum_{k=0}^{\infty} f_+(z + k) + q^- \sum_{k=0}^{\infty} f_+(z - k), \quad z \in \mathbb{R},$$

where f_+ is the right-continuous version of f. □

Theorem 3.1 *The one-dimensional simple recourse function Q is convex if and only if the underlying random variable $\boldsymbol{h}$ is continuously distributed with a pdf f that is of bounded variation, such that*

$$f_+(s) = G(s + 1) - G(s), \quad s \in \mathbb{R}, \tag{5}$$

where G is an arbitrary cdf with finite mean value. □

Sufficiency of (5) is easy to see, since it implies that

$$Q'_+(z) = -q^+ \big(1 - G(z)\big) + q^- G(z + 1), \quad z \in \mathbb{R}, \tag{6}$$

is non-decreasing. Below we will make extensive use of the following special case.

Corollary 3.1 *Assume that $\boldsymbol{h}$ is continuously distributed with a pdf f whose right-continuous version is constant on every interval $[\alpha + k, \alpha + k + 1)$, $k \in \mathbb{Z}$, for some $\alpha \in [0, 1)$. Then the function Q is piecewise linear and convex, with knots contained in $\{\alpha + \mathbb{Z}\}$.*

PROOF. Immediate from Theorem 3.1 and (6), since $f_+(s) = G(s + 1) - G(s)$ where G is the cdf of a discrete distribution with support contained in $\alpha + \mathbb{Z}$. □

To arrive at convex approximations of the function Q, we will use Corollary 3.1 to construct suitable approximations of the distribution of the random variable $\boldsymbol{h}$. For future reference, we present the multivariate definition of the approximations that we have in mind.

Definition 3.1 Let $\boldsymbol{h} \in \mathbb{R}^m$ be a random vector with arbitrary continuous or discrete distribution, and choose $\alpha = (\alpha_1, \ldots, \alpha_m) \in [0, 1)^m$. Define *the α-approximation* $\boldsymbol{h}_\alpha$ as the random vector with joint pdf f_α that is constant on every hypercube $C_\alpha^k := \prod_{i=1}^m (\alpha_i + k_i - 1, \alpha_i + k_i]$, $k \in \mathbb{Z}^m$, such that $\Pr\{\boldsymbol{h}_\alpha \in C_\alpha^k\} = \Pr\{\boldsymbol{h} \in C_\alpha^k\}$, $k \in \mathbb{Z}^m$.

Returning to the one-dimensional case, it is easy to see that the α-approximations $\boldsymbol{h}_\alpha$, $\alpha \in [0, 1)$, of an arbitrary random variable $\boldsymbol{h}$, satisfy the assumptions of Corollary 3.1. It follows that the *α-approximations* of the function Q, defined for $\alpha \in [0, 1)$,

$$Q_\alpha(z) := q^+ \mathbb{E}_{\boldsymbol{h}_\alpha}\left[\lceil \boldsymbol{h}_\alpha - z \rceil^+\right] + q^- \mathbb{E}_{\boldsymbol{h}_\alpha}\left[\lfloor \boldsymbol{h}_\alpha - z \rfloor^-\right], \quad z \in \mathbb{R},$$

are piecewise linear convex approximation of Q, with knots contained in $\{\alpha + \mathbb{Z}\}$. Moreover, it follows from Lemma 3.1 (a) and Definition 3.1 that

$$Q_\alpha(z) = Q(z), \quad z \in \{\alpha + \mathbb{Z}\}.$$

We conclude that, for each $\alpha \in [0, 1)$, Q_α is the piecewise linear convex function generated by the restriction of Q to $\{\alpha + \mathbb{Z}\}$. See Figure 3.1 for an example of the function Q and one of its α-approximations.

In the discussion above, no assumptions were made on the type of distribution of $\boldsymbol{h}$. However, to establish a non-trivial bound on the approximation error, we need to assume that $\boldsymbol{h}$ is continuously distributed. This loss of generality is acceptable, because for the case with discretely distributed $\boldsymbol{h}$ it is possible to construct the convex hull of the function Q.

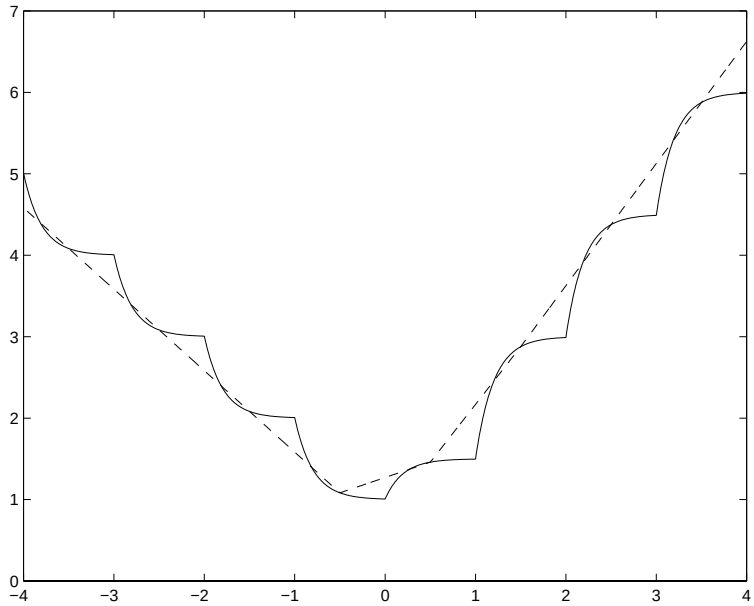

Figure 3.1: The function Q and its α-approximation Q_α (dashed) in case $\boldsymbol{h}$ is exponentially distributed with parameter 5, $q^+ = 1$, $q^- = 1.5$, and $\alpha = 0.5$.

Theorem 3.2 *Assume that $\boldsymbol{h}$ is continuously distributed with a pdf f that is of bounded variation. Then, for all $\alpha \in [0, 1)$,*

$$\|Q_\alpha - Q\|_\infty \leq (q^+ + q^-)\frac{|\Delta| f}{4},$$

where $|\Delta| f$ denotes the total variation of f.

PROOF. We will sketch a proof for the special case that $q^+ = 1$ and $q^- = 0$. The proof for the general case is analogous.

Assume that $q^+ = 1$ and $q^- = 0$. Then the function Q reduces to the expected surplus function $g(z) := \mathbb{E}_{\boldsymbol{h}}\left[\lceil \boldsymbol{h} - z \rceil^+\right]$, $z \in \mathbb{R}$, with α-approximations $g_\alpha(z) := \mathbb{E}_{\boldsymbol{h}_\alpha}\left[\lceil \boldsymbol{h}_\alpha - z \rceil^+\right]$, $\alpha \in [0, 1)$. Since $g(z) = g_\alpha(z)$ if $z \in \{\alpha + \mathbb{Z}\}$, consider an arbitrary fixed $z \notin \{\alpha + \mathbb{Z}\}$, and let $\underline{z} \in \{\alpha + \mathbb{Z}\}$ be such that $\underline{z} < z < \underline{z} + 1$.

Using Lemma 3.1 (b) we find that

$$g(\underline{z}) - g(z) = \int_{\underline{z}}^{z} \sum_{k=0}^{\infty} f(t+k)dt.$$

It follows from Lemma 2.5 in [20] that

$$1 - F(\underline{z}) - \frac{|\Delta| f}{2} \leq \sum_{k=0}^{\infty} f(t+k) \leq 1 - F(\underline{z}) + \frac{|\Delta| f}{2}, \quad t \in (\underline{z}, \underline{z}+1),$$

so that

$$\begin{aligned} \left(1 - F(\underline{z}) - \frac{|\Delta| f}{2}\right)(z - \underline{z}) &\leq g(\underline{z}) - g(z) \\ &\leq \left(1 - F(\underline{z}) + \frac{|\Delta| f}{2}\right)(z - \underline{z}). \end{aligned} \tag{7}$$

On the other hand, using Lemma 3.1 (a) we see that

$$g(s+1) = g(s) - (1 - F(s)), \quad s \in \mathbb{R}.$$

Since the function g_a coincides with g on $\{\alpha + \mathbb{Z}\}$, and moreover g_α is linear on the interval $[\underline{z}, \underline{z}+1]$, it follows that

$$g(\underline{z}) - g_\alpha(z) = (1 - F(\underline{z}))(z - \underline{z}). \tag{8}$$

Together, (7) and (8) imply

$$|g_\alpha(z) - g(z)| \leq (z - \underline{z})\frac{|\Delta| f}{2}, \quad z \in [\underline{z}, \underline{z}+1]. \tag{9}$$

Similarly, by comparing $g(z)$ and $g_\alpha(z)$ to $g(\underline{z}+1)$, one obtains

$$|g_\alpha(z) - g(z)| \leq (\underline{z} + 1 - z)\frac{|\Delta| f}{2}, \quad z \in [\underline{z}, \underline{z}+1]. \tag{10}$$

For α-approximations of expected surplus function g, the claimed error bound now follows from (9) and (10) on the observation that $\min\{(z - \underline{z}), (\underline{z} + 1 - z)\} \leq 1/2$.

Analogously, the same error bound can be derived for the special case with $q^+ = 0$ and $q^- = 1$. The claim for the general case then follows trivially. □

The uniform error bound of Theorem 3.2 can be reduced by a factor 2 if the following combination of α-approximations is used. For $\alpha \in [0, 0.5)$ and $\beta = \alpha + 0.5$, define the

pdf

$$f_{\alpha\beta}(s) = \frac{f_\alpha(s) + f_\beta(s)}{2}, \quad s \in \mathbb{R},$$

where f_α and f_β are density functions of α-approximations as before. The resulting convex approximations $Q_{\alpha\beta}$ of Q satisfy

$$\|Q_{\alpha\beta} - Q\|_\infty \leq (q^+ + q^-)\frac{|\Delta| f}{8}. \tag{11}$$

It can be shown that this error bound can not be reduced by using other convex combinations of pdf of type f_α.

The error bound presented above is proportional to the total variation of the pdf f of $\boldsymbol{h}$. For many distributions, e.g. with unimodal densities, the total variation of a pdf decreases as the variance of the distribution increases. We may therefore expect that the approximation Q_α becomes better as the variance of such distributions becomes higher.

Finally, we remark that convex approximations of the function Q can be represented as (one-dimensional) *continuous* simple recourse functions. The latter functions are defined like (4), except that no rounding operations are involved. In the case of α-approximations, the corresponding modification of the underlying distribution is known in closed form [19].

Lemma 3.2 *Let $\boldsymbol{h}$ be a continuous random variable with cdf F with finite mean value, and $\alpha \in [0, 1)$. Then*

$$Q_\alpha(z) = q^+ \mathbb{E}_{\boldsymbol{\varphi}_\alpha}\left[(\boldsymbol{\varphi}_\alpha - z)^+\right] + q^- \mathbb{E}_{\boldsymbol{\varphi}_\alpha}\left[(\boldsymbol{\varphi}_\alpha - z)^-\right] + \frac{q^+ q^-}{q^+ + q^-}, \quad z \in \mathbb{R},$$

where $\boldsymbol{\varphi}_\alpha$ is a discrete random variable with support in $\alpha + \mathbb{Z}$ and, for $k \in \mathbb{Z}$,

$$\Pr\{\boldsymbol{\varphi}_\alpha = \alpha + k\} \quad = \quad \frac{q^+}{q^+ + q^-}\Pr\{\boldsymbol{h} \in C_\alpha^k\} + \frac{q^-}{q^+ + q^-}\Pr\{\boldsymbol{h} \in C_\alpha^{k+1}\}.$$

□

We conclude that simple *integer* recourse functions can be approximated by *continuous* simple recourse functions with *discretely distributed* right-hand side parameters, simply by dropping the integrality restrictions and a modification of the distribution according to Lemma 3.2. The resulting convex problem can be solved using existing algorithms for continuous simple recourse problems with discrete underlying distributions.

3.2 Complete integer recourse

We now turn to the much more general class of complete integer recourse models. In addition to completeness and sufficiently expensive recourse, so that v is finite, we assume that the recourse matrix W is integer (or rational, so that integrality of W can be obtained by scaling). We will see that also in this case α- approximations of the distribution of $\boldsymbol{h}$ lead to convex approximations of the recourse function Q. In fact, if the recourse matrix is *totally unimodular* (TU) then this approach leads to the convex hull of Q. Below we first derive the results for this special case.

Because W is TU, the extreme points of the feasible set $\{y \in \mathbb{R}^{n_2} : Wy \geq h\}$ are integral for any integer right-hand side h. However, in our recourse problem the right-hand side $\boldsymbol{h} - Tx$ is not an integer vector in general. But since Wy is integral for all $y \in \mathbb{Z}^{n_2}$ we may round up the right-hand-side. Due to the assumption that W is TU, we may now relax the integrality restrictions on y, without changing the optimal value of the problem. That is,

$$\begin{aligned} v(x, \boldsymbol{h}) &:= \min_{y} \; qy \\ &\quad \text{s.t.} \quad Wy \geq \boldsymbol{h} - Tx, \; y \in \mathbb{Z}^{n_2}_{+} \\ &= \min_{y} \; qy \qquad (12) \\ &\quad \text{s.t.} \quad Wy \geq \lceil \boldsymbol{h} - Tx \rceil, \; y \in \mathbb{R}^{n_2}_{+} \\ &= \max_{\lambda} \; \lambda \lceil \boldsymbol{h} - Tx \rceil \qquad (13) \\ &\quad \text{s.t.} \quad \lambda W \leq q, \; \lambda \in \mathbb{R}^{m}_{+}, \end{aligned}$$

where the last equality follows from (strong) LP duality.

Since the recourse structure is complete and sufficiently expensive, it follows that the dual feasible region $\Lambda := \{\lambda \in \mathbb{R}^{m}_{+} : \lambda W \leq q\}$ is a bounded, non-empty polyhedral set. Hence,

$$v(x, \boldsymbol{h}) = \max_{k=1,\ldots,K} \lambda^k \lceil \boldsymbol{h} - Tx \rceil, \quad x \in \mathbb{R}^n, \boldsymbol{h} \in \mathbb{R}^m, \qquad (14)$$

where λ^k, $k = 1, \ldots, K$, are the finitely many extreme points of the dual feasible set Λ.

Thus, v is the maximum of finitely many round up functions, and hence non-convex. However, as we will see below, the recourse function Q is convex if the underlying distribution of $\boldsymbol{h}$ is of a certain type. Analogous to the simple recourse case, this allows the construction of convex approximations of Q by means of special purpose approximations of the distribution.

To set the stage, we first study the expected round up function

$$R(z) := \lambda \mathbb{E}_{\boldsymbol{h}} [\lceil \boldsymbol{h} - z \rceil], \quad z \in \mathbb{R}^m,$$

defined for any fixed $\lambda \in \mathbb{R}^m$.

If $m = 1$, $\lambda = 1$, and $\boldsymbol{h}$ is continuously distributed, then

$$R(z) = \mathbb{E}_{\boldsymbol{h}} \left[\lceil \boldsymbol{h} - z \rceil^+ \right] - \mathbb{E}_{\boldsymbol{h}} \left[\lfloor \boldsymbol{h} - z + 1 \rfloor^- \right], \quad z \in \mathbb{R}, \tag{15}$$

since $\lceil s \rceil = \lceil s \rceil^+ - \lceil s \rceil^-$, $s \in \mathbb{R}$, and $\lceil s \rceil^- = \lfloor s + 1 \rfloor^-$ for all $s \notin \mathbb{Z}$. The right-hand side of (15) is very similar to the one-dimensional simple recourse function with $q^+ = 1$ and $q^- = -1$. Hence, in view of Corollary 3.1 it is not surprising that this one-dimensional function R is convex if $\boldsymbol{h}$ has a piecewise constant pdf of the type specified in that lemma. This result can be generalized to m-dimensional round up functions.

Lemma 3.3 *Let $\boldsymbol{h} \in \mathbb{R}^m$ be a continuous random vector with joint pdf $f_{\boldsymbol{h}}$ that is constant on every hypercube $C^k_\alpha := \prod_{i=1}^m (\alpha_i + k_i - 1, \alpha_i + k_i]$, $k \in \mathbb{Z}^m$, for an arbitrary but fixed $\alpha = (\alpha_1, \ldots, \alpha_m) \in [0, 1)^m$. Then*

$$\mathbb{E}_{\boldsymbol{h}} [\lceil \boldsymbol{h} - z \rceil] = \mathbb{E}_{\boldsymbol{\varphi}_\alpha} \left[\boldsymbol{\varphi}_\alpha - z \right] = \mu_\alpha - z, \quad z \in \mathbb{R}^m,$$

where $\boldsymbol{\varphi}_\alpha = \lceil \boldsymbol{h} - \alpha \rceil + \alpha$ is a discrete random vector with mean value μ_α and support in $\alpha + \mathbb{Z}^m$, with

$$\Pr\{\boldsymbol{\varphi}_\alpha = \alpha + k\} = \Pr\{\boldsymbol{h} \in C^k_\alpha\}, \quad k \in \mathbb{Z}^m.$$

Hence, in this case the round up function $R(z) = \lambda \mathbb{E}_{\boldsymbol{h}} [\lceil \boldsymbol{h} - z \rceil]$, $z \in \mathbb{R}^m$, is affine with gradient $-\lambda$.

PROOF. We use that

$$\mathbb{E}_{\boldsymbol{h}} [\lceil \boldsymbol{h} - z \rceil] = \sum_{k \in \mathbb{Z}^m} \Pr\{\boldsymbol{h} \in C^k_\alpha\} \mathbb{E}_{\boldsymbol{h}} \left[\lceil \boldsymbol{h} - z \rceil \,\middle|\, \boldsymbol{h} \in C^k_\alpha \right], \quad z \in \mathbb{R}^m. \tag{16}$$

For each fixed $k \in \mathbb{Z}^m$, $\Pr\{\boldsymbol{h} \in C^k_\alpha\}$ is either zero or the conditional distribution of $\boldsymbol{h}$ given $\boldsymbol{h} \in C^k_\alpha$ is uniform on C^k_α. In that case, the components of the vector $\boldsymbol{h}$ are independent random variables on C^k_α, with each $\boldsymbol{h}_i$ uniformly distributed on $(\alpha_i + k_i - 1, \alpha_i + k_i]$, $i = 1, \ldots, m$. Hence, writing each component as in (15) and applying Lemma 3.2 to each term individually, it follows that

$$\mathbb{E}_{\boldsymbol{h}} \left[\lceil \boldsymbol{h} - z \rceil \,\middle|\, \boldsymbol{h} \in C^k_\alpha \right] = \alpha + k - z, \quad z \in \mathbb{R}^m. \tag{17}$$

Substitution of (17) in (16) proves the first claim.

The second claim follows trivially from the first one. □

Based on Lemma 3.3, we define α-approximations of the function R: for $\alpha \in [0,1)^n$,

$$R_\alpha(z) := \lambda \mathbb{E}_{\boldsymbol{h}_\alpha} [\lceil \boldsymbol{h}_\alpha - z \rceil], \quad z \in \mathbb{R}^m.$$

In general, an α-approximation is neither a lower bound nor an upper bound. However, since $R(z+k) = R(z) - \lambda k$, $k \in \mathbb{Z}^m$, for every z, we see that $R(z) + \lambda z$ is a periodic function, which repeats itself on every set C_α^k. Thus, defining

$$\alpha^\star \in \operatorname{argmin} \left\{ R(z) + \lambda z : z \in [0,1)^m \right\}, \tag{18}$$

$R_{\alpha^\star}$ is a lower bound for R, which is sharp at every $z \in \alpha^\star + \mathbb{Z}^m$. By construction, the affine function $R_{\alpha^\star}$ is actually the convex hull of R.

The components $\alpha_i^\star$, $i = 1, \ldots, m$, of the parameter vector $\alpha^\star$ can independently be determined analytically in almost all practical cases. If the marginal distribution of $\boldsymbol{h}_i$ is continuous, one-sided derivatives of the function $R_i(z_i) := \lambda_i \mathbb{E}_{\boldsymbol{h}_i} [\lceil \boldsymbol{h}_i - z_i \rceil]$ (analogous to Lemma 3.1) are used; if it is discrete with finitely many different fractional values in its support, the computation of $\alpha_i^\star$ is based on the direct relation between these fractional values and discontinuities of the lower semicontinuous function R_i.

Now we are ready to prove the main result for this class of models with TU recourse matrix. Using the dual representation (14) of the value function v, we have

$$Q(x) = \mathbb{E}_{\boldsymbol{h}} \left[\max_{k=1,\ldots,K} \lambda^k \lceil \boldsymbol{h} - Tx \rceil \right], \quad x \in \mathbb{R}^n.$$

Note that Q is not simply the pointwise maximum of a number of expected round up functions R. However, the results above for the function R play a major role in the proof of Theorem 3.3.

Theorem 3.3 *Consider the integer recourse expected value function $\mathcal{Q}$, defined as*

$$\mathcal{Q}(x) = \mathbb{E}_{\boldsymbol{h}} \left[\min_y qy : Wy \geq \boldsymbol{h} - Tx,\ y \in \mathbb{Z}_+^{n_2} \right], \quad x \in \mathbb{R}^{n_1}, \tag{19}$$

Assume that

- *(i) the recourse is complete and sufficiently expensive, and*
- *(ii) the recourse matrix W is totally unimodular.*

If

- *(iii) the matrix T is of full row rank,*

then the convex hull of Q is the continuous recourse expected value function $\mathcal{Q}_{\alpha^\star}$, defined as

$$\mathcal{Q}_{\alpha^\star}(x) = \mathbb{E}_{\boldsymbol{\varphi}_{\alpha^\star}}\left[\min_y qy : Wy \geq \boldsymbol{\varphi}_{\alpha^\star} - Tx,\ y \in \mathbb{R}^{n_2}_+\right], \quad x \in \mathbb{R}^{n_1}, \tag{20}$$

where $\alpha^\star$ is defined by (18), and $\boldsymbol{\varphi}_{\alpha^\star}$ is the discrete random vector $\boldsymbol{\varphi}_{\alpha^\star} = \lceil \boldsymbol{h} - \alpha^\star \rceil + \alpha^\star$ with support in $\alpha^\star + \mathbb{Z}^m$, and

$$\Pr\{\boldsymbol{\varphi}_{\alpha^\star} = \alpha^\star + k\} = \Pr\{\boldsymbol{h} \in C^k_{\alpha^\star}\}, \quad k \in \mathbb{Z}^m.$$

If condition (iii) is not satisfied, then $\mathcal{Q}_{\alpha^\star}$ is a lower bound for $\mathcal{Q}$.

PROOF. We will prove that $\mathcal{Q}_{\alpha^\star}$ is the convex hull of $\mathcal{Q}$ if T is of full row rank. The other case then follows from Theorem 2.2 in [17].

Assuming that T is of full row rank, we may conveniently consider $\mathcal{Q}$ as a function of the tender variables $z := Tx \in \mathbb{R}^m$.

First we will prove that $Q_{\alpha^\star}$ is a lower bound for Q, and subsequently that $Q_{\alpha^\star}(z) = Q(z)$ for all $z \in \alpha^\star + \mathbb{Z}^m$. This completes the proof, since all vertices of the polyhedral function $Q_{\alpha^\star}$ are contained in $\alpha^\star + \mathbb{Z}^m$.

Using the dual representation (14) of the value function v, we have

$$Q(z) = \mathbb{E}_{\boldsymbol{h}}\left[\max_{k=1,\ldots,K} \lambda^k \lceil \boldsymbol{h} - z \rceil\right], \quad z \in \mathbb{R}^m,$$

and, analogously,

$$Q_{\alpha^\star}(z) = \mathbb{E}_{\boldsymbol{\varphi}_{\alpha^\star}}\left[\max_{k=1,\ldots,K} \lambda^k (\boldsymbol{\varphi}_{\alpha^\star} - z)\right], \quad z \in \mathbb{R}^m.$$

Conditioning on the events $\boldsymbol{h} \in C^l_{\alpha^\star}$, $l \in \mathbb{Z}^m$, we obtain, for $z \in \mathbb{R}^m$,

$$\begin{aligned}
Q(z) &= \sum_{l \in \mathbb{Z}^m} \Pr\{\boldsymbol{h} \in C^l_{\alpha^\star}\}\, \mathbb{E}_{\boldsymbol{h}}\left[\max_{k=1,\ldots,K} \lambda^k \lceil \boldsymbol{h} - z \rceil \;\middle|\; \boldsymbol{h} \in C^l_{\alpha^\star}\right] \\
&\geq \sum_{l \in \mathbb{Z}^m} \Pr\{\boldsymbol{h} \in C^l_{\alpha^\star}\} \max_{k=1,\ldots,K} \lambda^k \mathbb{E}_{\boldsymbol{h}}\left[\lceil \boldsymbol{h} - z \rceil \;\middle|\; \boldsymbol{h} \in C^l_{\alpha^\star}\right] \\
&\geq \sum_{l \in \mathbb{Z}^m} \Pr\{\boldsymbol{h} \in C^l_{\alpha^\star}\} \max_{k=1,\ldots,K} \lambda^k \mathbb{E}_{\boldsymbol{h}_{\alpha^\star}}\left[\lceil \boldsymbol{h}_{\alpha^\star} - z \rceil \;\middle|\; \boldsymbol{h}_{\alpha^\star} \in C^l_{\alpha^\star}\right] \\
&= \sum_{l \in \mathbb{Z}^m} \Pr\{\boldsymbol{h} \in C^l_{\alpha^\star}\} \max_{k=1,\ldots,K} \lambda^k \left(\alpha^\star + l - z\right)
\end{aligned}$$

$$= \sum_{l \in \mathbb{Z}^m} \Pr\{\boldsymbol{\varphi}_{\alpha^\star} = \alpha^\star + l\} \max_{k=1,\ldots,K} \lambda^k \left(\alpha^\star + l - z\right) = Q_{\alpha^\star}(z).$$

The second inequality is valid because each λ^k is nonnegative, so that the α-approximation $\lambda^k \mathbb{E}_{\boldsymbol{h}_{\alpha^\star}} \left[\lceil \boldsymbol{h}_{\alpha^\star} - z \rceil \;\middle|\; \boldsymbol{h}_{\alpha^\star} \in C^l_{\alpha^\star} \right]$ is a lower bound for $\lambda^k \mathbb{E}_{\boldsymbol{h}} \left[\lceil \boldsymbol{h} - z \rceil \;\middle|\; \boldsymbol{h} \in C^l_{\alpha^\star} \right]$ by the choice of $\alpha^\star$. The subsequent equality holds by Lemma 3.3.

It remains to prove that $Q_{\alpha^\star} = Q$ on $\alpha^\star + \mathbb{Z}^m$. Consider a fixed $\bar{z} \in \alpha^\star + \mathbb{Z}^m$ and a fixed $l \in \mathbb{Z}^m$. Then $\lceil \boldsymbol{h} - \bar{z} \rceil = l - \lfloor \bar{z} \rfloor$ is constant for all $\boldsymbol{h} \in C^l_\alpha$, so that there exists a $\lambda(\bar{z}, l)$ satisfying

$$\lambda(\bar{z}, l) \in \operatorname*{argmax}_{k=1,\ldots,K} \lambda^k \lceil \boldsymbol{h} - \bar{z} \rceil \quad \forall \boldsymbol{h} \in C^l_\alpha.$$

Since this is true for every $\bar{z} \in \alpha^\star + \mathbb{Z}^m$ and $l \in \mathbb{Z}^m$, it follows that, for $z \in \alpha^\star + \mathbb{Z}^m$,

$$\begin{aligned}
Q(z) &= \sum_{l \in \mathbb{Z}^m} \Pr\{\boldsymbol{h} \in C^l_{\alpha^\star}\} \; \lambda(z, l) \; \mathbb{E}_{\boldsymbol{h}} \left[\lceil \boldsymbol{h} - z \rceil \;\middle|\; \boldsymbol{h} \in C^l_{\alpha^\star} \right] \\
&= \sum_{l \in \mathbb{Z}^m} \Pr\{\boldsymbol{h} \in C^l_{\alpha^\star}\} \; \lambda(z, l) \; \mathbb{E}_{\boldsymbol{h}_{\alpha^\star}} \left[\lceil \boldsymbol{h}_{\alpha^\star} - z \rceil \;\middle|\; \boldsymbol{h}_{\alpha^\star} \in C^l_{\alpha^\star} \right] \\
&= \sum_{l \in \mathbb{Z}^m} \Pr\{\boldsymbol{h} \in C^l_{\alpha^\star}\} \; \lambda(z, l) \; \left(\alpha^\star + l - z\right) \\
&= \sum_{l \in \mathbb{Z}^m} \Pr\{\boldsymbol{h} \in C^l_{\alpha^\star}\} \max_{k=1,\ldots,K} \lambda^k \left(\alpha^\star + l - z\right).
\end{aligned}$$

The second equality follows from the fact that each α-approximation is sharp on $\alpha + \mathbb{Z}^m$. The last equality follows from the definition of $\lambda(z, l)$ and $\boldsymbol{\varphi}_{\alpha^\star} - z = l - \lfloor z \rfloor$, $z \in \alpha^\star + \mathbb{Z}^m$. □

We conclude that if the recourse matrix W is totally unimodular, then the integer complete recourse problem with recourse function Q can be approximated by the continuous complete recourse problem with recourse function $\mathcal{Q}_{\alpha^\star}$. To construct this approximation, the integer restrictions on the second-stage variables are dropped, and the distribution of the right-hand side parameters is modified according to Theorem 3.3. The resulting continuous complete recourse problem with discretely distributed right-hand side parameters can be solved by existing special purpose algorithms [2, 16].

In particular, if the matrix T is of full row rank, then solving the approximating problem will yield the true optimal solution, at least if the first-stage constraints are not binding.

Finally, we drop the assumption that W is TU. In this case, we will prove that $\mathcal{Q}_{\alpha^\star}$ is

a strictly better convex approximation than the one obtained using the LP relaxation of the second-stage problem. The latter convex function will be denoted by $\mathcal{Q}^{LP}$, defined as

$$\mathcal{Q}^{LP}(x) := \mathbb{E}_{\boldsymbol{h}}\left[\min_{y}\left\{qy : Wy \geq \boldsymbol{h} - Tx,\ y \in \mathbb{R}^{n_2}_{+}\right\}\right], \quad x \in \mathbb{R}^{n_1}. \tag{21}$$

Theorem 3.4 *Consider the functions $\mathcal{Q}_{\alpha^\star}$ and $\mathcal{Q}^{LP}$, defined by (20) and (21) respectively, which both are convex lower bounds for the integer recourse expected value function $\mathcal{Q}$, defined by (19).*

(a) $\mathcal{Q}_{\alpha^\star} \geq \mathcal{Q}^{LP}$

(b) Assume

(i) $q \geq 0$, *so that* 0 *is a trivial lower bound for* v *and* Q*;*

(ii) there exists a subset L *of* $\mathbb{Z}^m$ *such that the support* Ω *is a subset of* $\bigcup_{l\in L}\{\boldsymbol{h} : \boldsymbol{h} \leq \alpha^\star + l\}$ *and* $\Pr\{\boldsymbol{h} < \alpha^\star + l \mid \boldsymbol{h} \in C^l_{\alpha^\star}\} > 0$ *for all* $l \in L$.

Then the function $\mathcal{Q}_{\alpha^\star}$ is a strictly better convex approximation of $\mathcal{Q}$ than $\mathcal{Q}^{LP}$, in the sense that $\mathcal{Q}(x) > 0$ implies $\mathcal{Q}_{\alpha^\star}(x) > \mathcal{Q}^{LP}(x)$.

PROOF. As before, we condition on the events $\boldsymbol{h} \in C^l_{\alpha^\star}$, $l \in \mathbb{Z}^m$, to obtain, for $x \in \mathbb{R}^{n_1}$,

$$\mathcal{Q}_{\alpha^\star}(x) = \sum_{l\in\mathbb{Z}^m} \Pr\{\boldsymbol{h} \in C^l_{\alpha^\star}\} \max_{k=1,\ldots,K} \lambda^k\left(\alpha^\star + l - Tx\right) \tag{22}$$

and

$$\mathcal{Q}^{LP}(x) = \sum_{l\in\mathbb{Z}^m} \Pr\{\boldsymbol{h} \in C^l_{\alpha^\star}\}\, \mathbb{E}_{\boldsymbol{h}}\left[\max_{k=1,\ldots,K} \lambda^k\left(\boldsymbol{h} - Tx\right) \mid \boldsymbol{h} \in C^l_{\alpha^\star}\right]. \tag{23}$$

For each $l \in \mathbb{Z}^m$ it follows from the definition of $C^l_{\alpha^\star} = \prod_{i=1}^m \left(\alpha^\star_i + l_i - 1, \alpha^\star_i + l_i\right]$ that $\alpha^\star + l \geq h$ for all $h \in C^l_{\alpha^\star}$. Thus, for $k = 1, \ldots, K$, $\lambda^k\left(\alpha^\star + l - Tx\right) \geq \lambda^k\left(h - Tx\right)$ for all $h \in C^l_{\alpha^\star}$, since $\lambda^k \geq 0$. Substitution in (22) and (23) proves that $\mathcal{Q}_{\alpha^\star} \geq \mathcal{Q}^{LP}$.

To prove (b), we first show that $\mathcal{Q}(x) > 0$ implies $\mathcal{Q}_{\alpha^\star}(x) > 0$. To this end, define

$$N(x) := \left\{t \in \mathbb{R}^m : v(t - Tx) > 0\right\}, \qquad x \in \mathbb{R}^{n_1}.$$

Then $\mathcal{Q}(x) > 0$ if and only if $\Pr\{\boldsymbol{h} \in N(x)\} > 0$, which is equivalent to $\Pr\{\boldsymbol{h} \in \operatorname{int} N(x)\} > 0$ since $N(x)$ is an open set. By Definition 3.1, it follows that then also $\Pr\{\boldsymbol{h}_\alpha \in N(x)\} > 0$, which implies $\mathcal{Q}_\alpha(x) > 0$ for all $\alpha \in [0, 1)^m$.

Let x be such that $\mathcal{Q}(x) > 0$, implying $\mathcal{Q}_{\alpha^\star}(x) > 0$. Then, since each term of (22) is

non-negative by assumption (i), there exists an $\bar{l} \in L$ such that

$$\max_{k=1,\ldots,K} \lambda^k \left(\alpha^\star + \bar{l} - Tx\right) > 0;$$

obviously, any optimal solution $\bar{\lambda}$ of this problem satisfies $\bar{\lambda} \neq 0$. For an arbitrary but fixed $\bar{h} \in C^l_{\alpha^\star}$ such that $\bar{h} < \alpha^\star + \bar{l}$, it holds

$$\lambda\left(\bar{h} - Tx\right) \leq \lambda\left(\alpha^\star + \bar{l} - Tx\right) \quad \forall \lambda \geq 0,$$

with strict inequality unless $\lambda = 0$. Let $\hat{\lambda}$ be an optimal solution of $\max_k \lambda^k \left(\bar{h} - Tx\right)$. Then there are two possibilities:

(i) $\hat{\lambda} = 0$, so that $\hat{\lambda}\left(\bar{h} - Tx\right) = 0 < \bar{\lambda}\left(\alpha^\star + \bar{l} - Tx\right)$;
(ii) $\hat{\lambda} \neq 0$, so that $\hat{\lambda}\left(\bar{h} - Tx\right) < \hat{\lambda}\left(\alpha^\star + \bar{l} - Tx\right) \leq \bar{\lambda}\left(\alpha^\star + \bar{l} - Tx\right)$.

We conclude that, for all $\bar{h} \in C^{\bar{l}}_{\alpha^\star}$ with $\bar{h} < \alpha^\star + \bar{l}$,

$$\max_{k=1,\ldots,K} \lambda^k \left(\alpha^\star + \bar{l} - Tx\right) > \max_{k=1,\ldots,K} \lambda^k \left(\bar{h} - Tx\right). \tag{24}$$

Since $\Pr\{\boldsymbol{h} < \alpha^\star + \bar{l} \mid \boldsymbol{h} \in C^{\bar{l}}_{\alpha^\star}\} > 0$ by assumption (ii), and (24) holds with weak inequality for all $h \in C^{\bar{l}}_{\alpha^\star}$, it follows that

$$\max_{k=1,\ldots,K} \lambda^k \left(\alpha^\star + \bar{l} - Tx\right) > \mathbb{E}_{\boldsymbol{h}}\left[\max_{k=1,\ldots,K} \lambda^k \left(\boldsymbol{h} - Tx\right) \,\middle|\, \boldsymbol{h} \in C^{\bar{l}}_{\alpha^\star}\right]. \tag{25}$$

Finally, using that (25) holds with weak inequality for all $l \in \mathbb{Z}^m$, we see from (22) and (23) that $\mathcal{Q}_{\alpha^\star}(x) > \mathcal{Q}^{LP}(x)$. □

For example, condition (b) (ii) of Theorem 3.4 is satisfied if $\boldsymbol{h}$ follows a non-degenerated continuous distribution.

Note that the distribution of $\boldsymbol{\varphi}_{\alpha^\star}$ as defined in Theorem 3.3 is always discrete, no matter what kind of distribution $\boldsymbol{h}$ follows. Thus, in particular if $\boldsymbol{h}$ is continuously distributed, $\mathcal{Q}_{\alpha^\star}$ is not only a better approximation of $\mathcal{Q}$, it is also *computationally more tractable* than $\mathcal{Q}^{LP}$ which in this case is defined as an m-dimensional integral.

4. Hierarchical planning models

Consider a two-level decision situation. At the higher level, aggregate decisions are made concerning acquisition of resources. At the lower level, one has to decide on the actual allocation of the resources. The time horizon for aggregate decisions in such

hierarchical decision problems may range from several months to a year. At the time aggregate decisions are made much detailed information of what will ultimately be required of the resources is not yet known with certainty. As mentioned in the introduction, two-stage stochastic programming is the right tool to model the lower level of hierarchical planning problems accurately, using stochastic parameters for which probability distributions are specified. The objective at the higher level is to minimize known costs at that level plus the expected objective value of optimal lower level decisions.

We focus on hierarchical planning problems with detailed-level problems which are of a combinatorial nature. This class of problems includes hierarchical scheduling problems, hierarchical vehicle routing problems, and hierarchical knapsack problems (capital budgeting problems). We will consider the design and analysis of approximation algorithms for such problems. In management science literature such algorithms are often called *hierarchical planning systems.* We diverge here from the previous section by not aiming at approximation algorithms for the stochastic integer programming problem in which all hierarchical combinatorial optimization problems can be formulated, but instead at algorithms tailored to the specific particular hierarchical combinatorial optimization problems. We use as an example a hierarchical scheduling problem and apply probabilistic analysis to measure the performance quality of an approximation algorithm for this problem

Consider the following *hierarchical scheduling problem.* At the time machines are to be installed only probabilistic information is available on the jobs to be processed. The two-stage stochastic programming model of this problem has to select the number or the types of the machines so as to minimize the installation costs of the machines plus the expected cost of processing the jobs optimally on the installed machines.

In this problem, the machines to be installed at the aggregate level are identical and work in parallel. Installation of each machine costs c. A decision is required on the number of machines to be installed. If x denotes this number, then the installation costs are cx.

There are N jobs to be processed and each job j requires processing for a time t_j, $j = 1, \ldots, N$. At the time the machines are purchased, there is only probabilistic information about the processing times of the jobs. A schedule of the jobs on the available machines is feasible if each job gets assigned one time interval of length equal to its processing time on one of the machines, and each machine processes only one job at a time. The makespan of a set of jobs is the time by which the last job is completed in a feasible schedule. The objective is to minimize the sum of installation costs of the machines and the expected makespan of the jobs on the available machines. (To make

dimensions compatible assume, without loss of generality, that the cost per time unit in the second-stage schedule is equal to 1.)

Let $v^*(x, \mathbf{t})$ denote the optimal second-stage costs, which is a random variable, a function of the random processing times of the jobs $\mathbf{t} = (\mathbf{t_1}, \ldots, \mathbf{t_N})$. Let $Q(x) = \mathbb{E}_{\mathbf{t}}[v^*(x, \mathbf{t})]$ denote its expectation. Then the objective is $\min z(x) = cx + Q(x)$. Let x^* denote the optimal solution.

From Section 2 we know that computing $Q(x)$ is a formidable task. Even if the distribution of $\mathbf{t}$ would be given by discrete probabilities over a set of vectors, the deterministic equivalent problem is NP-hard, since computing the optimal makespan of a set of jobs on more than one machine is NP-hard. The approximation algorithm H consists of replacing $Q(x)$ by a simple function $Q^H(x)$ as an approximate.

Obviously, given a realization t of $\mathbf{t}$, $\sum_{j=1}^{N} t_j/x$, the makespan of the schedule in which all x machines have equal workload, is a lower bound on $v^*(x, t)$:

$$v^*(x, t) \geq \sum_{j=1}^{N} t_j/x. \tag{26}$$

We choose to take $Q^H(x) = \mathbb{E}_{\mathbf{t}}\left[\sum_{j=1}^{N} \mathbf{t_j}\right]/x$ as an approximate value for $Q(x)$. If we assuming for simplicity that $\mathbf{t_1}, \ldots, \mathbf{t_N}$ have equal mean μ then $Q^H(x) = N\mu/x$. We solve the approximate problem

$$\min z^H(x) = cx + Q^H(x). \tag{27}$$

$z^H(x)$ is a convex function and $\frac{dz^H}{dx} = 0$ at $x = \sqrt{N\mu/c}$. Since the number of machines must be integer, we use discrete optimization to find that $x^H = \lceil \sqrt{N\mu/c + 1/4} - 1/2 \rceil$ is an optimal solution. The outcome of the approximation algorithm is then $z(x^H)$.

Taking expectations in (26) yields $Q(x) \geq Q^H(x)$. Hence $z(x) \geq z^H(x)$ for all x, and therefore

$$\min z(x) = z(x^*) \geq z^H(x^H) = cx^H + Q^H(x^H) \geq 2\sqrt{cN\mu}. \tag{28}$$

To estimate the quality of the approximation we aim to find an appropriate upper bound on $z^H(x^H)$ in terms of $z(x^*)$. It is well-known [14] that the list scheduling rule, which assigns the jobs in an arbitrary order to the machines and each next job is assigned to the earliest available machine, yields the following upper bound on the makespan for

any number of machines x given any realization t of $\mathbf{t}$:

$$v^*(x,t) \leq \sum_{j=1}^{N} t_j/x + \max_{j=1,\ldots,N} t_j.$$

In particular for x^H, denoting $t_{\max} = \max_{j=1,\ldots,N} t_j$ and taking expectations yields

$$Q(x^H) \leq Q^H(x^H) + \mathbb{E}_{\mathbf{t}}\left[\mathbf{t}_{\max}\right].$$

Hence,

$$z(x^H) \leq cx^H + Q^H(x^H) + \mathbb{E}_{\mathbf{t}}\left[\mathbf{t}_{\max}\right]. \tag{29}$$

Together, (28) and (29) give a bound on the worst-case performance ratio of the approximation algorithm.

Lemma 4.1

$$\frac{z(x^H)}{z(x^*)} \leq 1 + \frac{\mathbb{E}_{\mathbf{t}}\left[\mathbf{t}_{\max}\right]}{2\sqrt{cN\mu}}.$$

Probability theory (see e.g. [34]) tells us that

Lemma 4.2 *If* $\mathbf{t_1}, \ldots, \mathbf{t_N}$ *have finite second moments, then*

$$\lim_{N\to\infty} \mathbb{E}_{\mathbf{t}}\left[\mathbf{t}_{\max}\right]/\sqrt{N} = 0.$$

This probabilistic result applied to Lemma 4.1 yields asymptotic optimality of the approximation algorithm.

Theorem 4.1 *If* $\mathbf{t_1}, \ldots, \mathbf{t_N}$ *have finite second moments then*

$$\lim_{N\to\infty} \frac{z(x^H)}{z(x^*)} = 1.$$

PROOF. The combination of Lemmas 4.1 and 4.2 implies $\lim_{N\to\infty} \frac{z(x^H)}{z(x^*)} \leq 1$. Clearly, $\frac{z(x^H)}{z(x^*)} \geq 1$. □

The above shows something more than asymptotic optimality. In replacing the expected second-stage optimal costs by an estimate, the NP-hardness of the second-stage problem is not taken into consideration. This could imply that we take the estimate of a

quantity that we are unable to compute efficiently once we obtain a realization of the second-stage parameters (the processing times of the jobs).

However, as we have seen, asymptotic optimality is proved by comparing the solution given by the algorithm to a solution based on a polynomial time approximation algorithm for the second-stage problem. It implies that optimality is retained even if we use a simple approximation algorithm to solve the second-stage problem.

We could also assess the quality of an algorithm that uses x^H as the number of machines and list scheduling for the second-stage problem on x^H machines. The solution value is a random variable $z^{LS}(x^H, \mathbf{t}) = cx^H + v^{LS}(x^H, \mathbf{t})$. One could wonder how close this value is to the solution value of an algorithm that selects x^* machines and upon a realization of the processing times is able to select the optimal schedule on those x^* machines. Let denote this solution value by $z^*(x^*) = cx^* + v^*(x^*, \mathbf{t})$.

Or one could even wonder how $z^{LS}(x^H, \mathbf{t})$ compares to the solution value of an optimal clairvoyant algorithm that is able to know the realization of the processing times before deciding the number of machines to be installed. In this case the optimal number of machines becomes a random variable denoted by $x^0(\mathbf{t})$. Let us denote the optimal solution value in this case by $z^0(x^0(\mathbf{t}), \mathbf{t}) = cx^0(\mathbf{t}) + v^*(x^0(\mathbf{t}), \mathbf{t})$. This is the solution of the model that in stochastic programming is called the *distribution model.* In more popular terms it is called the *wait and see* model for obvious reasons opposed to *here and now* model used for the two-stage model. The approximation algorithm presented above appears to have the strong asymptotic optimality property that, again under the assumption that the random processing times have finite second moments,

$$\lim_{N\to\infty} \frac{z^{LS}(x^H, \mathbf{t})}{z^0(x^0(\mathbf{t}), \mathbf{t})} = 1,$$

with probability 1, or *almost surely*. A sequence of random variables $\mathbf{y_1}, \ldots, \mathbf{y_N}$ is said to converge almost surely to a random variable $\mathbf{y}$ if $Pr\{\lim_{N\to\infty} \mathbf{y_N} = \mathbf{y}\} = 1$. The proof is similar to the proof of Theorem 4.1. Under some mild extra conditions on the random variables, which are satisfied for example if the distributions of all random variables have bounded support, this asymptotic quality guarantee implies directly that

$$\lim_{N\to\infty} \frac{z^{LS}(x^H, \mathbf{t})}{z^*(x^*, \mathbf{t})} = 1,$$

almost surely. It also implies the result of Theorem 4.1.

The ideas used above in constructing an asymptotically optimal approximation algorithm for the two-stage stochastic scheduling problem are applicable more generally.

Given a two-stage combinatorial optimization problem replace the value function by an estimate that is asymptotically accurate and use an asymptotically optimal approximation algorithm for the second-stage problem in case this problem is NP-hard.

5. Worst-case performance analysis

As an example of worst-case performance analysis of approximation algorithms for stochastic optimization problems we consider a service provision problem. Actually, to the best of our knowledge it is the only example of this type of analysis in stochastic programming. In that sense, a rich research area lies nearly unexplored.

The problem we study concerns provision of services from a resource. For each of a given set of services there are requests from customers. In order to meet a request for a service, the service has to be installed and once installed, the request has to be served. Both installation and provision of a service requires capacity from the same resource. The resource has limited capacity. Each request served yields a given profit. The problem is to select a subset of the services to be installed and to decide which customer requests to serve, such as to maximize the total profit by serving requests.

If all demands for services are known in advance, the problem is NP-hard *in the ordinary sense* and a fully-polynomial time approximation scheme exists.

We study the problem with uncertain demand for services. The uncertainty is represented by a discrete probability distribution over the demands. The two-stage stochastic programming problem is to select services to be installed such as to maximize expected profit of serving requests for services. We will show that this problem is *strongly* NP-hard. Thus, the complexity of the problem increases by introducing stochasticity.

We analyse the performance of an approximation algorithm for this problem under the restriction that the resource has enough capacity to install all services. It may not be optimal to install all of them since it may leave too little capacity for serving the requests.

We start with formulating the problem as a two-stage stochastic integer programming problem. Let n be the number of services and s the capacity of the single resource. Let q_j be the profit obtained from allocating one resource unit to meeting demand for service, j. Each service j requires a resource capacity r_j to be installed, which is independent of the demand met. Demand is denoted by the random vector $\boldsymbol{d} \in \mathbb{R}^n$, with $\boldsymbol{d}_j$ denoting the demand for service j. Binary decision variables z_j are used to indicate

whether service j is installed ($z_j = 1$), or not ($z_j = 0$), $j = 1, \ldots, n$. Decision variable x_j gives the amount of resource used to meet demand for service j. The two-stage stochastic programming formulation becomes:

$$\begin{array}{lll} \max & \mathbb{E}_{\boldsymbol{d}}\left[v(z, \boldsymbol{d})\right] & \\ \text{s.t.} & \sum_{j=1}^{n} r_j z_j \leq s & \\ & z_j \in \{0, 1\} & j = 1, \ldots, n, \end{array}$$

with

$$\begin{array}{lll} v(z, \boldsymbol{d}) = \max & \sum_{j=1}^{n} q_j x_j & \\ \text{s.t.} & \sum_{j=1}^{n} x_j \leq s - \sum_{j=1}^{n} r_j z_j & \\ & x_j \leq \boldsymbol{d}_j z_j & j = 1, \ldots, n, \\ & x_j \geq 0 & j = 1, \ldots, n. \end{array}$$

The second-stage problem is to set the values of the variables x_j under two constraints: the *capacity constraint* ensuring that resource capacity is not exceeded and the *demand constraint* ensuring that demand is not exceeded and met only for services that have been installed. The constraint in the first stage ensures relatively complete recourse; i.e. for every first stage solution that is feasible with respect to the first stage constraints, the resulting second-stage problem is feasible for every realization of the random parameters.

Let K be the number of scenarios describing the probability distribution on demand, p^k the probability that scenario k occurs, and d_j^k the demand for service j in scenario k. Given the scenarios the the following deterministic equivalent linear mixed integer program can be formulated, in which we use x_{jk} to denote the resource allocated to providing service j in scenario k (we use a subscript instead of superscript for k here because of notational convenience later on).

$$\begin{array}{lll} \max & \sum_{k=1}^{K} p^k \sum_{j=1}^{n} q_j x_{jk} & \\ \text{s.t.} & \sum_{j=1}^{n} (r_j z_j \quad + x_{jk}) \leq s & k = 1, \ldots, K, \\ & d_j^k z_j \quad - x_{jk} \geq 0 & j = 1, \ldots, n, \quad k = 1, \ldots, K, \\ & z_j \in \{0, 1\}, \;\; x_{jk} \;\; \geq 0 & j = 1, \ldots, n, \quad k = 1, \ldots, K. \end{array} \tag{30}$$

Though integrality conditions only hold for the first stage variables z_j, if the data, resource capacity, installation requirements, and demands are integral, the second stage will have an integer solution in every scenario.

Theorem 5.1 *The stochastic single resource service provision problem is strongly NP-hard.*

PROOF. The natural recognition version of this problem obtained by introducing a number Λ and asking if there is a feasible solution with objective value at least Λ is in NP, following directly from the deterministic equivalent formulation. To see that it is strongly NP-Complete consider a reduction from the well-known strongly NP-Complete vertex cover problem (see [13]):

> Given a graph $G = (V, E)$ with $|V|$ vertices and $|E|$ edges and a constant κ, does there exist a subset V' of the vertices, such that each edge in E is incident to at least one vertex in V', and such that $|V'| \leq \kappa$?

For every vertex $j \in V$ introduce a service j with installation requirement $\alpha = \frac{1}{\kappa|E|}$. For every edge introduce a scenario with demand 1 for the two services incident to it and demand 0 for all other services. Let all scenarios have probability $\frac{1}{|E|}$. Define $q_j = |E| \; \forall j \in V$, $s = \kappa\alpha + 1$ (resource capacity), and $\Lambda = |E|$.

If there exists a vertex cover of size at most κ then there is a solution to the instance of the stochastic service provision problem with total expected profit at least $|E|$. Install the services corresponding to the vertices in the vertex cover. Then for each scenario (edge) at least one of the services with demand 1 is installed. The total capacity used by the installation of the services is at most $\kappa\alpha$ leaving at least capacity 1 in each scenario to satisfy demand.

The other direction is a bit more complicated. Suppose there does not exist a vertex cover of size κ or less. Then installing all services corresponding to a vertex cover would use node capacity strictly greater than $\kappa\alpha$ leaving strictly less than 1 for meeting demand in each of the $|E|$ scenarios, making a total expected profit of at least $|E|$ unattainable. Installing any set of services of size $L < \kappa$ would leave $(\kappa - L)\alpha + 1$ node capacity for meeting demand in each scenario. However, at least one edge will remain uncovered, implying that there is at least one scenario in which both services with a positive demand are not installed. With at most $|E|-1$ scenarios the expected profit will be at most $(|E|-1)((\kappa-L)\alpha+1) \leq (|E|-1)(\kappa\alpha+1) = (|E|-1)(\frac{1}{|E|}+1) < |E| = \Lambda$. □

As announced, we assume that $\sum_{j=1}^{n} r_j \leq s$. Moreover, to facilitate the exposition the assumption is made that no demand is higher than the node capacity minus the corresponding installation requirement: For any service j in any scenario k, $d_j^k \in [0, s - r_j]$. If necessary, this can be ensured by preprocessing.

The approximation algorithm that we will present is based on rounding the optimal solution of the LP-relaxation of problem (30), obtained by replacing the binary restrictions on the z-variables by $0 \leq z_j \leq 1$, $j = 1, \ldots, n$. To facilitate the exposition we assume, without loss of generality, that the resource capacity s is equal to 1.

Let $(z^{\mathrm{LP}}, x^{\mathrm{LP}})$ be an optimal basic solution of the LP relaxation. Let ℓ be the number of fractional z_j^{LP} and let ℓ_w of these services have $r_j \leq w$ for some $0 < w < 1$ to be chosen later. Let Z be the set of services with $z_j^{\mathrm{LP}} = 1$. By renumbering the services if necessary, assume that $0 < z_j^{\mathrm{LP}} < 1$ and $r_j \leq w$ for $j = 1, \ldots, \ell_w$ and $0 < z_j^{\mathrm{LP}} < 1$ and $r_j > w$ for $j = \ell_w + 1, \ldots, \ell$. Write the optimal LP value as

$$\pi^{\mathrm{LP}} = \pi_0^{\mathrm{LP}} + \pi_1^{\mathrm{LP}} + \pi_2^{\mathrm{LP}} \tag{31}$$

where

$$\pi_0^{\mathrm{LP}} = \sum_{j \in \mathbb{Z}} \sum_{k=1}^{K} p^k q_j x_{jk}^{\mathrm{LP}},$$

$$\pi_1^{\mathrm{LP}} = \sum_{j=1}^{\ell_w} \sum_{k=1}^{K} p^k q_j x_{jk}^{\mathrm{LP}},$$

and

$$\pi_2^{\mathrm{LP}} = \sum_{j=\ell_w+1}^{\ell} \sum_{k=1}^{K} p^k q_j x_{jk}^{\mathrm{LP}}.$$

Feasible solutions generated from the LP solution constitute the approximation algorithm, which selects from those solutions the best one. The algorithm is therefore a kind of rounding algorithm and we denote its solution value by π^R. Let π^{OPT} denote the optimal solution value of the stochastic integer program.

The first feasible solution is obtained by installing service j if and only if $z_j^{\mathrm{LP}} = 1$; i.e., install all services $j \in \mathbb{Z}$. The remaining capacity is then allocated to serve demand for the installed services in a greedy way, in order of non-increasing q_j values. Denote the

resulting solution by (z^G, x^G) and its value by π^G. Then, obviously,

$$\pi_0^{\mathrm{LP}} \leq \pi^G \leq \pi^R. \tag{32}$$

The next set of feasible solutions is used to bound π_1^{LP}. Define $A = \sum_{j=1}^{\ell} r_j z_j^{\mathrm{LP}}$ and note that $\sum_{j=1}^{\ell_w} x_{jk}^{\mathrm{LP}} \leq 1 - A$ for each $k = 1, \ldots, K$. Partition the set $\{1, \ldots, \ell_w\}$ into I subsets, $\{S_i\}_{i=1}^{I}$, such that

$$\sum_{j \in S_i} r_j \leq \beta + w \qquad i = 1, \ldots, I$$

and

$$\sum_{j \in S_i} r_j \geq \beta \qquad i = 1, \ldots, I-1, \tag{33}$$

for some constant $\beta > 0$ to be chosen later, such that $\beta + w < 1$. Notice that $\sum_{j \in S_I} r_j$ is allowed to be smaller than β. In the algorithm this partition is made in the most simple way, starting filling set S_1 until addition of the next service would make the sum of installation requirements exceed $\beta + w$. This service is then the first one of S_2, etc.

In the optimal solution of the LP relaxation at most $1 - A$ units of capacity are available for the x variables. Installing only the services in one of the sets S_i will leave at least $1 - \beta - w$ units of capacity available. The x-variable values from the LP relaxation solution corresponding to services in S_i may be scaled down, if necessary, to use a total of no more than $1 - \beta - w$ units of capacity in each scenario.

For each $i = 1, \ldots, I$ we obtain a feasible solution (z^{H_i}, x^{H_i}) with $z_j^{H_i} = 1$ for $j \in S_i$, $z_j^{H_i} = 0$ for $j \notin S_i$, $x_{jk}^{H_i} = \gamma x_{jk}^{\mathrm{LP}}$ for $j \in S_i$, $k = 1, \ldots, K$ and $x_{jk}^{H_i} = 0$ for $j \notin S_i$ and all k, where

$$\gamma = \begin{cases} \dfrac{1 - \beta - w}{1 - A} & \text{if } \beta + w \geq A, \\ 1 & \text{otherwise.} \end{cases} \tag{34}$$

The objective value of solution (z^{H_i}, x^{H_i}) is

$$\pi^{H_i} = \sum_{j \in S_i} \sum_{k=1}^{K} p^k q_j x_{jk}^{H_i} = \gamma \sum_{j \in S_i} \sum_{k=1}^{K} p^k q_j x_{jk}^{\mathrm{LP}}.$$

Hence,

$$\pi_1^{\mathrm{LP}} = \sum_{i=1}^{I} \sum_{j \in S_i} \sum_{k=1}^{K} p^k q_j x_{jk}^{\mathrm{LP}} = \frac{1}{\gamma} \sum_{i=1}^{I} \pi^{H_i} \leq \frac{I}{\gamma} \pi^R \tag{35}$$

By the assumption $\sum_{j=1}^{n} r_j \leq s$ and the definition of the sets S_i (specifically (33)) we have

$$1 \geq \sum_{j=1}^{n} r_j \geq \sum_{j=1}^{\ell_w} r_j = \sum_{i=1}^{I} \sum_{j \in S_i} r_j \geq (I-1)\beta. \tag{36}$$

Thus, $I \leq 1 + 1/\beta$ which inserted in (35) implies that

$$\pi_1^{\mathrm{LP}} \leq \frac{\beta+1}{\beta\gamma} \pi^R. \tag{37}$$

The last set of feasible solutions considered by the algorithm consists of installing each service $j = \ell_w + 1, \ldots, \ell$ (having $r_j \geq w$) individually. Since $A = \sum_{j=1}^{\ell} r_j z_j^{\mathrm{LP}} \geq \sum_{j=\ell_w+1}^{\ell} r_j z_j^{\mathrm{LP}} \geq w \sum_{j=\ell_w+1}^{\ell} z_j^{\mathrm{LP}}$, we have

$$\sum_{j=\ell_w+1}^{\ell} z_j^{\mathrm{LP}} \leq \frac{A}{w}.$$

Just installing service j has objective value $q_j E[\delta_j]$, since we have assumed that for any service j in any scenario k. $d_j^k \in [0, s - r_j]$. Satisfying the demand constraints implies that $\sum_{k=1}^{K} p^k x_{jk}^{\mathrm{LP}} \leq E[\delta_j] z_j^{\mathrm{LP}}$. Altogether this yields the following bound.

$$\begin{aligned} \pi_2^{\mathrm{LP}} &= \sum_{j=\ell_w+1}^{\ell} \sum_{k=1}^{K} p^k q_j x_{jk}^{\mathrm{LP}} \\ &\leq \sum_{j=\ell_w+1}^{\ell} q_j E[\delta_j] z_j^{\mathrm{LP}} \\ &\leq \pi^R \sum_{j=\ell_w+1}^{\ell} z_j^{\mathrm{LP}} \leq \frac{A}{w} \pi^R \leq \frac{A}{w} \pi^{\mathrm{OPT}} \end{aligned} \tag{38}$$

Combining (32), (37), and (38) gives

$$\pi^{\mathrm{LP}} \leq \left(1 + \frac{\beta+1}{\beta\gamma} + \frac{A}{w}\right) \pi^R \tag{39}$$

Theorem 5.2 *Under the assumption that* $\sum_{j=1}^{n} r_j \leq 1$, *the approximation algorithm has worst-case performance ratio*

$$\frac{\pi^{OPT}}{\pi^R} \leq (5 + 2\sqrt{3}).$$

PROOF. The choice of w and β depends on A in (39). When $A < \frac{1}{2}$ take $w = 1 - \frac{1}{2}\sqrt{3}$ and $\beta = -\frac{1}{2} + \frac{1}{2}\sqrt{3}$ and when $A \geq \frac{1}{2}$ take $w = \beta = \frac{1}{2}A$. In both cases $w + \beta \geq A$, and therefore $\gamma = \frac{1-\beta-w}{1-A}$. In the former case (39) leads to

$$\begin{aligned} \pi^{\mathrm{OPT}} \leq \pi^{\mathrm{LP}} &\leq \left(1 + \frac{2(1+\sqrt{3})(1-A)}{-1+\sqrt{3}} + \frac{A}{1-\frac{1}{2}\sqrt{3}}\right)\pi^R \\ &= \left(1 + (1+\sqrt{3})^2(1-A) + 4(1+\frac{1}{2}\sqrt{3})A\right)\pi^R \\ &= (5+2\sqrt{3})\pi^R. \end{aligned}$$

In the latter case (39) leads to

$$\pi^{\mathrm{OPT}} \leq \pi^{\mathrm{LP}} \leq \left(4 + \frac{2}{A}\right)\pi^R \leq 8\pi^R \leq (5+2\sqrt{3})\pi^R.$$

□

We notice that so far tightness of the bound has not been established. There exist an instance in which the ratio between the LP-bound and the optimal value is 4 and an instance for which the algorithm has ratio 2. The results show the possibilities of achieving worst-case performance results for approximation algorithms for stochastic integer programming problems. It is worthwhile to stress once more that the deterministic counterpart of the problem, having the same number of binary decision variables, is weakly NP-hard. Thus, the complexity of the problem increases by introducing stochasticity, even if it only means adding continuous decision variables for each scenario of the problem.

6. Notes

Stochastic programming models date back to the fifties [5, 3]. Several surveys on stochastic programming have appeared of which we mention here the introductory book of Kall and Wallace [16] and the comprehensive books by Prekopa [29] and by Birge and Louveaux [2]. For surveys specifically on stochastic integer programming we refer to the chapter by Louveaux and Schultz in the Handbook on Stochastic Programming [31], and the survey papers Klein Haneveld and Van der Vlerk [21], Römisch and Schultz [30], and Stougie and Van der Vlerk [37]. Resources on the Internet are the *Stochastic Programming Community Home Page* [4] and the bibliography [42].

The focus in this paper is on the two-stage recourse model. For a detailed discussion of the multistage model and generalizations (including random recourse matrices and nonlinear models) we refer to the Handbook on Stochastic Programming [31] or to [2, 16, 29].

More about the important class of chance-constrained problems and the related (conditional) value at risk models can be found in the Handbook on Stochastic Programming [31]. This class is of problems is very well surveyed in [29] and [39].

The mathematical properties of two-stage stochastic linear programming problems have been derived by various people and at a rather early stage in the research activities on stochastic programming. In particular we refer to the overview by Wets [45] and the monograph Kall [15].

The mathematical properties of two-stage stochastic integer programming problems have been established much more recently [36, 41, 32]. Schultz [32] proved the properties of the mixed-integer recourse function presented in Theorem 1.2. In addition, Schultz presented rather technical conditions for Lipschitz continuity of the function Q.

The results in Section 2 are selected from [12]. $\sharp P$-completeness of the problem GRAPH RELIABILITY has been proved in [40]. That exact evaluation of the second-stage expected value function may not even be in $PSPACE$ in case random parameters are continuously distributed follows from a result in [23].

Dyer and Stougie [12] also prove $PSPACE$-hardness of a specific non-standard version of a multi-stage stochastic programming problem if the number of stages is considered to be part of the input. The complexity of standard multi-stage stochastic programming remains unsolved.

Kannan *et al.* [11] have designed a polynomial randomized approximation scheme for the two-stage stochastic programming problem with continuously distributed parameters and continuous decision variables, when the input distributions are restricted to be *log-concave.* Their scheme relies heavily on the convexity of Q, and therefore cannot be applied to the two-stage stochastic integer programming problem.

The idea in Section 3 of approximating the expected value function of a stochastic programming problem with integer recourse by a convex function through perturbing the distributions of the random right-hand sides is due Klein Haneveld *et al.* [20, 19]. They implemented this idea for the case of simple integer recourse. See Van der Vlerk [44] for a generalization to multiple simple recourse models, allowing for piecewise-

linear penalty cost functions. The extension to the compete integer recourse case was done by Van der Vlerk [43].

For the problem with simple integer recourse, the formula and properties in Lemma 3.1 have been derived by Louveaux and Van der Vlerk [25], while the characterization of all probability distributions that lead to convex expected value functions in Theorem 3.1 is due to Klein Haneveld *et al.* [20].

The uniform error bounds on the α-approximation in Theorem 3.2 and on the $\alpha\beta$-approximation in (11) are from [19]. There it is also shown that the latter error bound can not be reduced by using other convex combinations of probability density functions of type f_α. The error bounds are derived in case the distributions of the random right hand sides are continuous. For the case with discretely distributed $\boldsymbol{h}$ it is possible to construct the convex hull of the function Q, see [18].

Algorithms for continuous simple recourse problems with discretely distributed right-hand side parameters can be found in e.g. [2, 16]. Using the structure of such problems, they can be represented as relatively *small* deterministic LP problems.

If the matrix W is complete but not TU, then the function $\mathcal{Q}_{\alpha^\star}$ defined in Theorem 3.3 can be used as a convex lower bounding approximation of the function $\mathcal{Q}$, allowing to approximately solve the integer recourse problem by solving a continuous complete recourse model. Although this approach is easy to implement and in many cases will give better results than using the LP lower bound $\mathcal{Q}^{LP}$, no (non-trivial) bound on the approximation error is known. Indeed, in most applications the approximation will not be good enough for this purpose. On the other hand, because of the properties discussed in Section 3, the function $\mathcal{Q}_{\alpha^\star}$ is well-suited as a building block in special-purpose algorithms for integer complete recourse models; several of these algorithms [1, 22, 27, 33] use the LP relaxation $\mathcal{Q}^{LP}$ for bounding purposes.

Hierarchical planning problems appear in many applications in management science. Usually the solution methods consist of solving the problems at the different levels separately and glue them together. Dempster *et al.* [6, 7] gave the first mathematically rigorous analysis of such a hierarchical planning system. They presented the result on the hierarchical scheduling problem exposed in Section 4. Their result has been extended to other hierarchical scheduling problems with different types of machines and common deadlines for the jobs by Stougie [36].

The notion of asymptotic optimality with respect to an optimal clairvoyant algorithm was introduced by Lenstra *et al.* [24]. In the same paper the authors investigated a general framework for the probabilistic analysis of approximation algorithms for hier-

archical planning problems. They show implications between the various asymptotic quality statements. Applications of this framework on routing and location problems appeared in [36], where also an survey of the above mentioned research can be found.

The probabilistic value analysis of combinatorial optimization problems which are used in the estimates for the second-stage costs form a body of literature on its own (see for a survey [35]).

Section 5 is extracted from work by Dye *et al.* [9]. In the same paper a pseudo-polynomial time dynamic programming algorithm is derived if the number of scenarios is fixed. The existence of a fully polynomial time approximation scheme for this case is open. NP-hardness in the ordinary sense of the deterministic counterpart of the problem was proved in [8]. In the same paper a fully-polynomial time approximation scheme has been presented for this deterministic problem. All versions of the problem with multiple resources are strongly NP-hard [10, 8].

The setting of the problem is inspired by an application in telecommunication dealing with provision of *processing based services* on a computer network with distributed processing capabilities [38].

Worst-case performance analysis in stochastic integer programming with discretely distributed second-stage parameters like the one presented in Section 5 is an almost unexplored rich research field with many challenging questions.

References

[1] S. Ahmed, M. Tawarmalani, and N.V. Sahinidis. A finite branch and bound algorithm for two-stage stochastic integer programs. Stochastic Programming E-Print Series, `http://www.speps.info`, 2000.

[2] J.R. Birge and F.V. Louveaux. *Introduction to Stochastic Programming*. Springer Verlag, New York, 1997.

[3] A. Charnes, W.W. Cooper, and G.H. Symonds. Cost horizons and certainty equivalents: An approach to stochastic programming of heating oil. *Management Science*, 4:183–195, 1958.

[4] COSP. Stochastic programming community home page. World Wide Web, `http://stoprog.org`, 2002.

[5] G.B. Dantzig. Linear programming under uncertainty. *Management Science*, 1:197–206, 1955.

[6] M.A.H. Dempster, M.L. Fisher, L. Jansen, B.J. Lageweg, J.K. Lenstra, and A.H.G. Rinnooy Kan. Analytical evaluation of hierarchical planning systems. *Oper. Res.*, 29:707–716, 1981.

[7] M.A.H. Dempster, M.L. Fisher, L. Jansen, B.J. Lageweg, J.K. Lenstra, and A.H.G. Rinnooy Kan. Analysis of heuristics for stochastic programming: results for hierarchical scheduling problems. *Math. Oper. Res.*, 8:525–537, 1983.

[8] S. Dye, L. Stougie, and A. Tomasgard. Approximation algorithms and relaxations for a service provision problem on a telecommunication network. Working paper 2-98, Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, N-7034 Trondheim, Norway, 1998 (to appear in *Discrete Applied Mathematics*).

[9] S. Dye, L. Stougie, and A. Tomasgard. The stochastic single node service provision problem. COSOR-Memorandum 99-04, Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, Eindhoven, 1999.

[10] S. Dye, A. Tomasgard, and S.W. Wallace. Feasibility in transportation networks with supply eating arcs. *Networks*, 31:165–176, 1998.

[11] M. Dyer, R. Kannan, and L. Stougie. A simple randomised algorithm for convex optimisation. SPOR-Report 2002-05, Dept. of Mathematics and Computer Science, Eindhoven Technical University, Eindhoven, 2002.

[12] M. Dyer and L. Stougie. Computational complexity of stochastic programming problems. SPOR-Report 2003-09, Dept. of Mathematics and Computer Science, Eindhoven Technical University, Eindhoven, 2003.

[13] M.R. Garey and D.S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness.* Freeman, New York, 1979.

[14] R.L. Graham. Bounds for certain multiprocessing anomalies. *Bell System Technological Journal*, 45:1563–1581, 1966.

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

[16] P. Kall and S.W. Wallace. *Stochastic Programming.* Wiley, Chichester, 1994. Also available as PDF file at `http://www.unizh.ch/ior/Pages/Deutsch/Mitglieder/Kall/bib/ka-wal-94.pdf%`.

[17] W.K. Klein Haneveld, L. Stougie, and M.H. van der Vlerk. Some theorems on the convex hull of a function with an application to stochastic integer programming. Discussion Paper TI 94-81, Tinbergen Institute, Amsterdam-Rotterdam, 1994.

[18] W.K. Klein Haneveld, L. Stougie, and M.H. van der Vlerk. An algorithm for the construction of convex hulls in simple integer recourse programming. *Ann. Oper. Res.*, 64:67–81, 1996.

[19] W.K. Klein Haneveld, L. Stougie, and M.H. van der Vlerk. Convex approxi-

mations for simple integer recourse models by perturbing the underlying distribution. Research Report 97A19, SOM, University of Groningen, `http://som.rug.nl`, 1997.

[20] W.K. Klein Haneveld, L. Stougie, and M.H. van der Vlerk. Convex simple integer recourse problems. Research Report 97A10, SOM, University of Groningen, `http://som.rug.nl`, 1997.

[21] W.K. Klein Haneveld and M.H. van der Vlerk. Stochastic integer programming: General models and algorithms. *Ann. Oper. Res.*, 85:39–57, 1999.

[22] G. Laporte and F.V. Louveaux. The integer L-shaped method for stochastic integer programs with complete recourse. *Oper. Res. Lett.*, 13:133–142, 1993.

[23] J. Lawrence. Polytope volume computation. *Math. Comput.*, 57:259–271, 1991.

[24] J.K. Lenstra, A.H.G. Rinnooy Kan, and L. Stougie. A framework for the probabilistic analysis of hierarchical planning systems. *Ann. Oper. Res.*, 1:23–42, 1984.

[25] F.V. Louveaux and M.H. van der Vlerk. Stochastic programming with simple integer recourse. *Math. Program.*, 61:301–325, 1993.

[26] Yu. Nesterov and A. Nemirovski. *Interior point polynomial methods in Convex Programming: theory and applications.* SIAM Series in Applied Mathematics. SIAM, 1994.

[27] V.I. Norkin, Yu.M. Ermoliev, and A. Ruszczyński. On optimal allocation of indivisibles under uncertainty. *Oper. Res.*, 46(3):381–395, 1998.

[28] C. H. Papadimitriou. *Computational complexity*. Addison-Wesley, 1994.

[29] A. Prékopa. *Stochastic Programming.* Kluwer Academic Publishers, Dordrecht, 1995.

[30] W. Römisch and R. Schultz. Multistage stochastic integer programs: An introduction. In M. Grötchel, S.O. Krumke, and J. Rambau, editors, *Online Optimization of Large Scale Systems*, pages 581–600. Springer, Berlin, 2001.

[31] A. Ruszczynski and A. Shapiro, editors. *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science.* North-Holland, to appear (2003).

[32] R. Schultz. Continuity properties of expectation functions in stochastic integer programming. *Math. Oper. Res.*, 18:578–589, 1993.

[33] R. Schultz, L. Stougie, and M.H. van der Vlerk. Solving stochastic programs with complete integer recourse: A framework using Gröbner Bases. *Math. Program.*, 83(2):229–252, 1998.

[34] R.J. Serfling. *Approximation theorems of mathematical statistics.* Wiley, New York, 1980.

[35] J.M. Steele. *Probability theory and combinatorial optimization.* SIAM CBMS-NSF Regional Conference Series in Applied Mathematics, 69. SIAM, Philadelphia, 1997.

[36] L. Stougie. *Design and analysis of algorithms for stochastic integer programming*, volume 37 of *CWI Tract.* Centrum voor Wiskunde en Informatica, Amsterdam, 1987.

[37] L. Stougie and M.H. van der Vlerk. Stochastic integer programming. In M. Dell'Amico, F. Maffioli, and S. Martello, editors, *Annotated Bibliographies in Combinatorial Optimization*, chapter 9, pages 127–141. Wiley, 1997.

[38] A. Tomasgard, J.A. Audestad, S. Dye, L. Stougie, M.H. van der Vlerk, and S.W. Wallace. Modelling aspects of distributed processing in telecommunication networks. *Ann. Oper. Res.*, 82:161–184, 1998.

[39] S.P. Uryasev, editor. *Probabilistic constrained optimization.* Kluwer Academic Publishers, Dordrecht, 2000. Methodology and applications.

[40] L. G. Valiant. The complexity of enumeration and reliability problems. *SIAM Journal on Computing*, 8:410–421, 1979.

[41] M.H. van der Vlerk. *Stochastic programming with integer recourse.* PhD thesis, University of Groningen, The Netherlands, 1995.

[42] M.H. van der Vlerk. Stochastic programming bibliography. World Wide Web, `http://mally.eco.rug.nl/spbib.html`, 1996-2002.

[43] M.H. van der Vlerk. Convex approximations for complete integer recourse models. Research Report 02A21, SOM, University of Groningen, `http://som.rug.nl`, 2002. Also available as *Stochastic Programming E-Print Series* 2002–10, `http://www.speps.info`.

[44] M.H. van der Vlerk. On multiple simple recourse models. Research Report 02A06, SOM, University of Groningen, `http://som.rug.nl`, 2002. Also available as *Stochastic Programming E-Print Series* 2002–7, `http://www.speps.info`.

[45] R.J-B. Wets. Stochastic programming: Solution techniques and approximation schemes. In A. Bachem, M. Grötschel, and B. Korte, editors, *Mathematical Programming, The State of the Art*, pages 566–603. Springer-Verlag, Berlin, 1982.