DEPARTMENT OF OPERATIONS RESEARCH
UNIVERSITY OF AARHUS

Working Paper no. 2002/4

Applying the minimax criterion in stochastic recourse programs

Morten Riis
Kim Allan Andersen

ISSN 1600-8987

Department of Mathematical Sciences
Telephone: +45 8942 1111
E-mail: institut@imf.au.dk

Building 530, Ny Munkegade
DK-8000 Aarhus C, Denmark
URL: www.imf.au.dk

Applying the minimax criterion in stochastic recourse programs

Morten Riis*
Department of Operations Research
University of Aarhus, building 530
Ny Munkegade
DK - 8000 Århus C
Denmark

Kim Allan Andersen
Department of Operations Research
University of Aarhus, building 530
Ny Munkegade
DK - 8000 Århus C
Denmark

Abstract

We consider an optimization problem in which some uncertain parameters are replaced by random variables. The minimax approach to stochastic programming concerns the problem of minimizing the worst expected value of the objective function with respect to the set of probability measures that are consistent with the available information on the random data. Only very few practicable solution procedures have been proposed for this problem and the existing ones rely on simplifying assumptions. In this paper, we establish a number of stability results for the minimax stochastic program, justifying in particular the approach of restricting attention to probability measures with support in some known finite set. Following this approach, we elaborate solution procedures for the minimax problem in the setting of two-stage stochastic recourse models, considering the linear recourse case as well as the integer recourse case. Since the solution procedures are modifications of well-known algorithms, their efficacy is immediate from the computational testing of these procedures and we do not report results of any computational experiments.

Keywords: Stochastic Programming; Minimax Techniques; Stability; Algorithms.

1 Introduction

Almost any real-life decision problem involves some degree of uncertainty. The importance of taking this inherent uncertainty into account when formulating such problems as mathematical programs is well-recognized and has received great attention during recent years. In

*Corresponding author. Email: riis@imf.au.dk

this paper we consider an optimization problem in which some parameters are not known with certainty and therefore replaced by random variables. Assuming that an outcome of random parameters constitutes a vector in R^N, we will be concerned with $\mathcal{P}(\mathbb{R}^N)$, the set of all Borel probability measures on R^N. A classical approach to the class of stochastic programming problems is to assume that the distribution $\mu \in \mathcal{P}(\mathbb{R}^N)$ of random parameters is known and then to solve the problem,

$$\min_{x \in X} \mathcal{Q}(x, \mu), \tag{1.1}$$

where X is some closed, convex set, and the function $\mathcal{Q}(x, \mu)$ typically denotes the total expected cost given the decision x and the distribution μ. For an introduction to general classes of stochastic programming models we refer to the textbooks by Birge and Louveaux [4], Kall and Wallace [16], and Prékopa [22].

A major concern in the formulation of the stochastic programming model (1.1) is the fact that the distribution μ of random parameters will hardly ever be directly accessible. In practice, the only information on the probability distribution that is available may be an estimate based on statistical information on the random data. In this situation a possible alternative approach is the following. First a class $\mathcal{A} \subseteq \mathcal{P}(\mathbb{R}^N)$ of possible or conceivable distributions, consistent with the available information, is established. A typical example is the case when $\mathcal{A}$ is the set of all distributions satisfying a number of mathematical constraints, taking the form of upper and lower bounds on mean, variance and other moments, but in general this class may be defined by any information characterizing probability distributions. Given the set of conceivable distributions $\mathcal{A}$ the following problem is solved,

$$MMP(\mathcal{A}) \qquad \min_{x \in X} \Big\{ f(x) := \sup_{\mu \in \mathcal{A}} \mathcal{Q}(x, \mu) \Big\}. \tag{1.2}$$

Problem (1.2) is referred to as the minimax approach to stochastic programming. Starting with Žáčková [40], variations of this approach have been considered by authors such as e.g. Birge and Dulá [2], Birge and Wets [5], Breton and Hachem [6, 7], Dupačová [9, 10], Ermoliev, Gaivoronski and Nedeva [12], Kall [15], and Shapiro and Kleywegt [37]. Only few numerical solution procedures for the minimax problem (1.2) have been presented in these papers. Furthermore, the ones proposed rely on results such as the following. If $\mathcal{A}$ is the class of all probability measures with support in some compact set Ξ, satisfying a number of generalized moment conditions,

$$\int_{\Xi} g_i(\xi) \mu(d\xi) \leq \alpha_i, \quad i = 1, \ldots, L,$$

where $g_i(\cdot), i = 1 \ldots, L$, are bounded continuous functions on Ξ, then $\mathcal{A}$ is a compact, convex set, and its extreme points are discrete measures with finite supports of at most $L+1$ points

(cf. Karr [17, Theorem 2.1], see also Kempermann [18]). Thus attention can be restricted to discrete measures in $\mathcal{A}$ having finite support of at most $L+1$ points. In Breton and Hachem [6, 7] this approach is employed to develop two alternative solution procedures, an extension of the progressive hedging algorithm (see Rockafellar and Wets [27]) and a bundle method, respectively. In fact, Breton and Hachem not only assume that all measures in $\mathcal{A}$ have finite support of at most $L+1$ points, but also that the support is known and identical for all $\mu \in \mathcal{A}$. In Ermoliev, Gaivoronski and Nedeva [12], on the other hand, generalized linear programming techniques are employed to determine the mass points and corresponding probabilities of a "worst" distribution in $\mathcal{A}$ with $L+1$ mass points. These techniques are combined with a projected quasi subgradient approach to determine an optimal solution of the minimax problem. Situations may occur, however, in which the set $\mathcal{A}$ does not fit into the above-mentioned framework, so that the restriction of attention to measures with finite support is not immediately valid. Also, even if attention can be restricted to measures in $\mathcal{A}$ with finite support, the inner maximization in (1.2), taking the form e.g. of a generalized moment problem, may still be intractable. In such situations, stability analysis of the minimax problem, when the set of conceivable measures is subjected to perturbations, becomes relevant. In particular, the results presented in Section 2 justify the approach followed by Breton and Hachem [6, 7], restricting attention to discrete measures with support in some known finite set.

In this paper we will be particularly concerned with two-stage stochastic recourse models for optimization under uncertainty. Such models are based on the assumption that an alternating process of decisions and observations of random data is appropriate. More specifically, we assume that some decisions must be taken in a first stage in which only distributional information on the uncertainties is available. The outcome of random parameters is subsequently observed and some corrective (or *recourse*) actions may be taken in a second stage. A typical approach to this class of problems is to minimize the sum of first-stage cost and expected second-stage, i.e. to let

$$\mathcal{Q}(x,\mu) := cx + \int_{\mathbb{R}^N} \Phi(x,\xi)\mu(d\xi). \tag{1.3}$$

Here the random vector $\xi \in \mathbb{R}^N$ is constituted by the components of a tuple (q,h,T), and the second-stage value function Φ is given accordingly by

$$\Phi(x,\xi) := \min\Big\{qy \mid Wy \geq h + Tx,\ y \in Y\Big\}. \tag{1.4}$$

It is assumed that $c \in \mathbb{R}^{n_1}$ is a known vector, and that W is a known rational matrix of size $m \times n_2$, referred to as the recourse matrix. Furthermore, it is assumed that the

second-stage cost q is a random n_2-vector, the second-stage right-hand h is a random m-vector, and the technology matrix T is a random matrix of size $m \times n_1$. Finally, the set $Y \subseteq \mathbb{R}^{n_2}$ may or may not contain integrality restrictions on some or all of the variables. Under fairly general assumptions, the function $\mathcal{Q}(\cdot, \mu)$ is a real-valued, lower semicontinuous function on $\mathbb{R}^{n_1}$ for any $\mu \in \mathcal{P}(\mathbb{R}^N)$. Also, it is easily seen that the supremum of a family of lower semicontinuous functions is again a lower semicontinuous function (cf. the proof of Proposition 4.1 in Section 4 below). Thus the minimax problem (1.2) is well-defined in the sense that one minimizes a lower semicontinuous function, and hence the optimal value is actually attained, provided that the problem is feasible and bounded.

The reason for focusing on two-stage stochastic recourse problems in the present paper is the fact that much research has gone into stability properties of optimal solutions to these problems, when the underlying probability distribution μ varies in some subset of $\mathcal{P}(\mathbb{R}^N)$. In particular, qualitative and quantitative continuity properties of the recourse function $\mathcal{Q}$ lead to continuity results for the optimal value function and the optimal solution set mapping as functions of μ. For a collection of results on stability in two-stage stochastic linear programming, we refer to Dupačová [10, 11], Kall [14], Robinson and Wets [25], Römisch and Schultz [28, 29, 30], Shapiro [36], and Wang [39], and for results in two-stage stochastic programming with mixed-integer recourse, we refer to Schultz [32, 33, 34]. It turns out that the continuity properties of $\mathcal{Q}$ established in these papers are also sufficient to provide stability results for the minimax problem (1.2), when the set $\mathcal{A}$ is subjected to perturbations. Still, the results of the present paper are applicable for any other class of problems for which similar continuity properties can be established.

Also, solution procedures for the class of two-stage stochastic recourse programs have been the center of extensive research. In 1969 Van Slyke and Wets [38] introduced the L-shaped method, a solution procedure for problems with linear recourse, based on Benders decomposition principle. Since then, several modifications and enhancements of this algorithm have been introduced. Among the most important of these procedures are the multicut version of the L-shaped algorithm presented by Birge and Louveaux [3], the stochastic decomposition procedure introduced by Higle and Sen [13], and the regularized decomposition method introduced by Ruszczynski [31]. Alternative procedures based on decomposition with respect to scenarios rather than stages have been considered by various authors, the most prominent example being the progressive hedging algorithm elaborated by Rockafellar and Wets [27]. For problems with (mixed-) integer recourse a smaller number of general purpose algorithms exist. Still a number of alternatives have been proposed, such as e.g. the branch and bound procedure presented by Ahmed, Tawarmalani and Sahinidis [1], the dual decomposition procedure introduced by Carøe and Schultz [8], the integer L-shaped method

introduced by Laporte and Louveaux [20], and a framework using the concept of Gröbner bases presented by Schultz, Stougie and van der Vlerk [35].

This paper is organized as follows. In Section 2, we establish a number of stability results for the minimax problem (1.2) with a general objective function $\mathcal{Q}$. We discuss their application to the class of two-stage stochastic recourse problems, where the recourse function is given by (1.3), and we discuss the implications of our results for the development of numerical solution procedures for the problem. In particular, the results justify the assumption that all measures in $\mathcal{A}$ are discrete with support in some known finite set. Employing this assumption, we elaborate solution procedures for the minimax problem in the setting of two-stage stochastic programs with linear recourse and integer recourse in Section 3 and Section 4, respectively. The solution procedure for the linear recourse case is a modification of the L-shaped algorithm and its multicut version. We give a proof of finite convergence whereas efficacy of the algorithm is immediate from the extensive application of the original L-shaped algorithm, and therefore no computational experiments are reported. Likewise, the solution procedure for the integer recourse case is a modification of a branch and bound algorithm presented by Ahmed, Tawarmalani and Sahinidis [1], and hence we only prove finite convergence and refer to the computational testing of Ahmed et al. for demonstration of efficacy of the algorithm. Finally we give some conclusions in Section 5.

2 Stability

In this section we establish qualitative and quantitative stability results for the minimax problem (1.2) when the set of conceivable distributions $\mathcal{A}$ is subjected to perturbations. To facilitate the analysis, we endow the set $\mathcal{P}(\mathbb{R}^N)$ of all Borel probability measures on $\mathbb{R}^N$ with the notion of weak convergence. We recall that a sequence $\{\mu_n\}_{n\geq 1}$ of Borel probability measures on $\mathbb{R}^N$ is said to converge weakly to some $\mu \in \mathcal{P}(\mathbb{R}^N)$, if for any bounded continuous function, $g : \mathbb{R}^N \mapsto \mathbb{R}$, we have

$$\int_{\mathbb{R}^N} g(t)\mu_n(dt) \xrightarrow{n\to\infty} \int_{\mathbb{R}^N} g(t)\mu(dt),$$

and in that case we write $\mu_n \xrightarrow{w} \mu$. The qualitative stability results will be based on the assumed joint continuity of $\mathcal{Q}$ with respect to x and μ, whereas the quantitative results rely on quantitative continuity properties of $\mathcal{Q}(x,\cdot)$ for $x \in X$. Sufficient conditions for these properties to hold for two-stage stochastic recourse models are briefly discussed.

The stability results presented here take the form of continuity properties of the optimal value function and the optimal solution set mapping as functions of $\mathcal{A}$. Since the problem is non-convex in several important cases, such as e.g. mixed-integer recourse models, we will

include local optimizers in the analysis. To this end, we define for any non-empty open set, $V \subseteq \mathbb{R}^{n_1}$, a localized version of the optimal value function,

$$\varphi_V(\mathcal{A}) := \min_{x \in X \cap \operatorname{cl} V} \sup_{\mu \in \mathcal{A}} \mathcal{Q}(x, \mu),$$

and a localized version of the optimal solution set mapping,

$$\Psi_V(\mathcal{A}) := \big\{x \in X \cap \operatorname{cl} V \mid \sup_{\mu \in \mathcal{A}} \mathcal{Q}(x, \mu) = \varphi_V(\mathcal{A})\big\},$$

where $\operatorname{cl} V$ denotes the closure of V. Moreover, we will adopt the notion of a complete local minimizing (CLM) set, originally introduced by Robinson [24] and Klatte [19]. Given some set $\mathcal{A} \subseteq \mathcal{P}(\mathbb{R}^N)$ of conceivable distributions, a set $M \subseteq \mathbb{R}^{n_1}$ is called a CLM set for $MMP(\mathcal{A})$ if there exists some open set $V \subseteq \mathbb{R}^{n_1}$ such that $M = \Psi_V(\mathcal{A})$ and $M \subseteq V$. Clearly, the set of global minimizers of problem (1.2), as well as any set of strict local minimizers, is a CLM set. Moreover, the notion of CLM sets precludes pathologies arising when $\mathcal{A} \subseteq \mathcal{P}(\mathbb{R}^N)$ and $V \subseteq \mathbb{R}^{n_1}$ are such that $\Psi_V(\mathcal{A})$ is a set of local minimizers for problem (1.2), while even the slightest perturbation of $\mathcal{A}$ leads to local solution sets with respect to V, which do not contain any local minimizers of (1.2).

Proposition 2.1. *Let $\mathcal{A} \subseteq \mathcal{P}(\mathbb{R}^N)$ and let $\{\mathcal{A}_n\}_{n\geq 1}$ be a sequence of sets of probability measures converging to $\mathcal{A}$ in the sense that $\cup_{n\geq 1}\mathcal{A}_n \subseteq \mathcal{A}$, and $\mu \in \mathcal{A}$ if and only if there exists a sequence $\{\mu_n\}_{n\geq 1}$ such that $\mu_n \in \mathcal{A}_n$ for all $n \geq 1$ and $\mu_n \xrightarrow{w} \mu$ as $n \to \infty$. Also, let $V \subseteq \mathbb{R}^{n_1}$ be some bounded open set such that $\Psi_V(\mathcal{A})$ is a CLM set for $MMP(\mathcal{A})$ with respect to V. If $\mathcal{Q} : \mathbb{R}^{n_1} \times \mathcal{P}(\mathbb{R}^N) \mapsto \mathbb{R}$ is continuous at (x, μ) for all $x \in X$ and $\mu \in \mathcal{A}$, then*

(a) $\varphi_V(\mathcal{A}_n) \to \varphi_V(\mathcal{A})$ as $n \to \infty$;

(b) $\sup_{x \in \Psi_V(\mathcal{A}_n)} \operatorname{dist}(x, \Psi_V(\mathcal{A})) \to 0$ as $n \to \infty$;

(c) There exists $N \geq 1$ such that $\Psi_V(\mathcal{A}_n)$ is a CLM set for $MMP(\mathcal{A}_n)$ with respect to V for all $n \geq N$.

Proof. Define for all $x \in X \cap \operatorname{cl} V$ and $n \geq 1$,

$$f(x) := \sup_{\mu \in \mathcal{A}} \mathcal{Q}(x, \mu) \quad \text{and} \quad f_n(x) := \sup_{\mu \in \mathcal{A}_n} \mathcal{Q}(x, \mu).$$

Now let $x^* \in \Psi_V(\mathcal{A})$ so that we have $\varphi_V(\mathcal{A}) = f(x^*)$, and for $n \geq 1$ let $x_n^* \in \Psi_V(\mathcal{A}_n)$ so that we have $\varphi_V(\mathcal{A}_n) = f_n(x_n^*)$.

The assumption that $\cup_{n\geq 1}\mathcal{A}_n \subseteq \mathcal{A}$ implies that for all $x \in X \cap \operatorname{cl} V$ and $n \geq 1$ we have $f_n(x) \leq f(x)$ so that $\limsup_{n\to\infty} \varphi_V(\mathcal{A}_n) \leq \varphi_V(\mathcal{A})$. To prove (a) by contradiction we assume

that $\liminf_{n\to\infty} \varphi_V(\mathcal{A}_n) < \varphi_V(\mathcal{A}) - \epsilon$ for some $\epsilon > 0$. As a consequence of this assumption, there exists some infinite subset $\mathbb{K}_1 \subseteq \mathbb{N}$ such that $\varphi_V(\mathcal{A}_n) < \varphi_V(\mathcal{A}) - \epsilon$ for all $n \in \mathbb{K}_1$, and hence $\mathcal{Q}(x_n^*, \mu_n) < \varphi_V(\mathcal{A}) - \epsilon$ for all $\mu_n \in \mathcal{A}_n$ and all $n \in \mathbb{K}_1$. Now, compactness of $X \cap \operatorname{cl} V$ implies the existence of some infinite subset $\mathbb{K}_2 \subseteq \mathbb{K}_1$ such that the subsequence $\{x_n^*\}_{n\in\mathbb{K}_2}$ converges to some $\bar{x} \in X \cap \operatorname{cl} V$. Also, for any $\mu \in \mathcal{A}$ we can select a sequence $\{\mu_n\}_{n\in\mathbb{K}_2}$ converging weakly to μ such that $\mu_n \in \mathcal{A}_n$ for all $n \in \mathbb{K}_2$. Joint continuity of $\mathcal{Q}$ now implies that

$$\mathcal{Q}(\bar{x}, \mu) = \lim_{\substack{n\to\infty \\ n\in\mathbb{K}_2}} \mathcal{Q}(x_n^*, \mu_n) < \varphi_V(\mathcal{A}) - \epsilon.$$

Thus $f(\bar{x}) < \varphi_V(\mathcal{A})$, a contradiction. This proves part (a).

To prove part (b), we let $\bar{x}$ be any accumulation point of the sequence $\{x_n^*\}_{n\geq 1}$, i.e. for some infinite subset $\mathbb{K}_1 \subseteq \mathbb{N}$ the sequence $\{x_n^*\}_{n\in\mathbb{K}_1}$ converges to $\bar{x}$. Let $\epsilon > 0$ be given and let $\bar{\mu} \in \mathcal{A}$ be such that $\mathcal{Q}(\bar{x}, \bar{\mu}) > f(\bar{x}) - \epsilon$. For any sequence $\{\bar{\mu}_n\}_{n\geq 1}$ of probability measures such that $\bar{\mu}_n \in \mathcal{A}_n$ and $\bar{\mu}_n \xrightarrow{w} \bar{\mu}$ as $n \to \infty$ we now have

$$\lim_{\substack{n\to\infty \\ n\in\mathbb{K}_1}} \mathcal{Q}(x_n^*, \bar{\mu}_n) = \mathcal{Q}(\bar{x}, \bar{\mu}) > f(\bar{x}) - \epsilon.$$

On the other hand, for any $n \geq 1$ the definition of $f_n(\cdot)$ implies $f_n(x_n^*) \geq \mathcal{Q}(x_n^*, \bar{\mu}_n)$ and from part (a) we have

$$\lim_{n\to\infty} f_n(x_n^*) = \lim_{n\to\infty} \varphi_V(\mathcal{A}_n) = \varphi_V(\mathcal{A}).$$

Thus $f(\bar{x}) < \varphi_V(\mathcal{A}) + \epsilon$ and since ϵ was arbitrary we must have $f(\bar{x}) \leq \varphi_V(\mathcal{A})$ so that $\bar{x} \in \Psi(\mathcal{A})$. This proves part (b).

To prove part (c) we note that $\Psi(\mathcal{A}) \subseteq V$ where $\Psi(\mathcal{A})$ is closed and V is open. Hence for some $\delta > 0$ we must have that $\operatorname{dist}(x, \Psi_V(\mathcal{A})) < \delta$ implies $x \in V$. This completes the proof. □

Remark 2.1. Proposition 2.1 applies for example in the situation when $\mathcal{A}$ is an infinite set of probability measures, rendering numerical solution of the inner maximization problem in (1.2) intractable. In this case Proposition 2.1 justifies the approach of approximating $\mathcal{A}$ by smaller subsets which are easier handled computationally. In particular, if $\mathcal{A}$ is the set of all Borel probability measures with support in some set Ξ, satisfying a number of generalized moment conditions, a possible approach is to restrict attention to those measures in $\mathcal{A}$ having support in some known finite set $\Xi' \subseteq \Xi$. In this way the inner maximization problem is simplified from a generalized moment problem to an ordinary linear program. Moreover, as a consequence of Proposition 2.1 we see that the optimal solution of such a simplified problem converges to the true optimal solution as the approximation of Ξ provided by Ξ' is progressively improved.

In some situations, however, different kinds of approximations of $\mathcal{A}$ than the one expressed in Proposition 2.1 may be more desirable. Consider for example the situation when $\mathcal{A}$ is constituted by a number of specific distributions, rendering the inner maximization problem in (1.2) intractable, e.g. $\mathcal{A}$ being a finite set of distributions that are absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^N$. In this case, we do not want the approximations of $\mathcal{A}$ to be subsets thereof, but rather to be sets of probability measures approximating each of the individual measures in $\mathcal{A}$. In this situation the following proposition applies.

Proposition 2.2. *Let $\mathcal{A} = \{\mu^1, \ldots, \mu^k\} \subseteq \mathcal{P}(\mathbb{R}^N)$ and for $n \geq 1$ let $\mathcal{A}_n = \{\mu_n^1, \ldots, \mu_n^k\}$ where $\mu_n^j \xrightarrow{w} \mu^j$ as $n \to \infty$ for $j = 1, \ldots, k$. Also, let $V \subseteq \mathbb{R}^{n_1}$ be some bounded open set such that $\Psi_V(\mathcal{A})$ is a CLM set for $MMP(\mathcal{A})$ with respect to V. If $\mathcal{Q} : \mathbb{R}^{n_1} \times \mathcal{P}(\mathbb{R}^N) \mapsto \mathbb{R}$ is continuous at (x, μ) for all $x \in X$ and $\mu \in \mathcal{A}$, then*

(a) $\varphi_V(\mathcal{A}_n) \to \varphi_V(\mathcal{A})$ as $n \to \infty$;

(b) $\sup_{x \in \Psi_V(\mathcal{A}_n)} \operatorname{dist}(x, \Psi_V(\mathcal{A})) \to 0$ as $n \to \infty$;

(c) There exists $N \geq 1$ such that $\Psi_V(\mathcal{A}_n)$ is a CLM set for $MMP(\mathcal{A}_n)$ with respect to V for all $n \geq N$.

Proof. Define for all $x \in X \cap \operatorname{cl} V$ and $n \geq 1$

$$f(x) := \max_{\mu \in \mathcal{A}} \mathcal{Q}(x, \mu) \quad \text{and} \quad f_n(x) := \max_{\mu \in \mathcal{A}_n} \mathcal{Q}(x, \mu).$$

Now let $x^* \in \Psi_V(\mathcal{A})$ and let $\mu^* \in \mathcal{A}$ be such that $\varphi_V(\mathcal{A}) = f(x^*) = \mathcal{Q}(x^*, \mu^*)$, and for $n \geq 1$ let $x_n^* \in \Psi_V(\mathcal{A}_n)$ so that we have $\varphi_V(\mathcal{A}_n) = f_n(x_n^*)$.

As in the proof of Proposition 2.1 we may show that $\liminf_{n\to\infty} \varphi_V(\mathcal{A}_n) \geq \varphi_V(\mathcal{A})$. Hence to prove part (a) by contradiction we assume that $\limsup_{n\to\infty} \varphi_V(\mathcal{A}_n) > \varphi_V(\mathcal{A})$. Since we have $f_n(x_n^*) \leq f_n(x^*)$ for all $n \geq 1$, the assumption implies $\limsup_{n\to\infty} f_n(x^*) > \varphi_V(\mathcal{A})$. Thus for some $j \in \{1, \ldots, k\}$ we have

$$\mathcal{Q}(x^*, \mu^j) = \lim_{n\to\infty} \mathcal{Q}(x^*, \mu_n^j) > \varphi_V(\mathcal{A}),$$

a contradiction since $\mathcal{Q}(x^*, \mu^j) \leq \mathcal{Q}(x^*, \mu^*) = \varphi_V(\mathcal{A})$ for all $j \in \{1, \ldots, k\}$. This proves part (a). The proofs of part (b) and (c) are similar to those in Proposition 2.1. □

Next, we turn to quantitative stability results for the minimax problem (1.2). To this end we assume that D is some distance defined on the set $\mathcal{P}(\mathbb{R}^N)$. Also, for $\mu \in \mathcal{P}(\mathbb{R}^N)$ and $\mathcal{A} \subseteq \mathcal{P}(\mathbb{R}^N)$ we let $D(\mu, \mathcal{A}) = \inf_{\nu \in \mathcal{A}} D(\mu, \nu)$ and introduce a Hausdorff-like distance between sets of probability measures, defined for $\mathcal{A}, \mathcal{B} \subseteq \mathcal{P}(\mathbb{R}^N)$ by

$$D_H(\mathcal{A}, \mathcal{B}) = \max\{\sup_{\mu \in \mathcal{A}} D(\mu, \mathcal{B}), \sup_{\mu \in \mathcal{B}} D(\mu, \mathcal{A})\}$$

Proposition 2.3. *Let $\mathcal{A} \subseteq \mathcal{P}(\mathbb{R}^N)$ and let $V \subseteq \mathbb{R}^{n_1}$ be some bounded open set such that $\Psi_V(\mathcal{A})$ is a CLM set for $MMP(\mathcal{A})$ with respect to V. If there exists constants $L, p, \delta > 0$ such that $|\mathcal{Q}(x,\mu) - \mathcal{Q}(x,\nu)| \leq L \cdot D(\mu,\nu)^p$ whenever $x \in X$ and $\mu, \nu \in \mathcal{P}(\mathbb{R}^N)$ with $D(\mu,\nu) < \delta$, then*

$$|\varphi_V(\mathcal{A}) - \varphi_V(\mathcal{B})| \leq L \cdot D_H(\mathcal{A},\mathcal{B})^p$$

whenever $\mathcal{B} \subseteq \mathcal{P}(\mathbb{R}^N)$ with $D_H(\mathcal{A},\mathcal{B}) < \delta$.

Proof. Let $\mathcal{B} \subseteq \mathcal{P}(\mathbb{R}^N)$ with $D_H(\mathcal{A},\mathcal{B}) < \delta$ and define for all $x \in X \cap \operatorname{cl} V$,

$$f(x) = \sup_{\mu \in \mathcal{A}} \mathcal{Q}(x,\mu) \quad \text{and} \quad g(x) = \sup_{\mu \in \mathcal{B}} \mathcal{Q}(x,\mu)$$

and let $x_A^* \in \Psi_V(\mathcal{A})$ and $x_B^* \in \Psi_V(\mathcal{B})$ so that we have $\varphi_V(\mathcal{A}) = f(x_A^*)$ and $\varphi_V(\mathcal{B}) = g(x_B^*)$. Obviously, we have

$$\varphi_V(\mathcal{A}) - \varphi_V(\mathcal{B}) \leq f(x_B^*) - g(x_B^*).$$

Let $\epsilon > 0$ be given and let $\bar{\mu} \in \mathcal{A}$ be such that $\mathcal{Q}(x_B^*, \bar{\mu}) > f(x_B^*) - \epsilon$. Also, let $\bar{\nu} \in \mathcal{B}$ be such that $D(\bar{\mu},\bar{\nu}) \leq D_H(\mathcal{A},\mathcal{B}) < \delta$ and $\mathcal{Q}(x_B^*,\bar{\nu}) \leq g(x_B^*)$. We now have

$$\begin{aligned} f(x_B^*) - g(x_B^*) &< \mathcal{Q}(x_B^*,\bar{\mu}) - \mathcal{Q}(x_B^*,\bar{\nu}) + \mathcal{Q}(x_B^*,\bar{\nu}) - g(x_B^*) + \epsilon \\ &\leq |\mathcal{Q}(x_B^*,\bar{\mu}) - \mathcal{Q}(x_B^*,\bar{\nu})| + \epsilon \\ &\leq L \cdot D_H(\mathcal{A},\mathcal{B})^p + \epsilon. \end{aligned}$$

Since ϵ was arbitrary we get

$$\varphi_V(\mathcal{A}) - \varphi_V(\mathcal{B}) \leq L \cdot D_H(\mathcal{A},\mathcal{B})^p,$$

and in the exact same way we may show that

$$\varphi_V(\mathcal{B}) - \varphi_V(\mathcal{A}) \leq L \cdot D_H(\mathcal{A},\mathcal{B})^p.$$

Thus we have

$$|\varphi_V(\mathcal{A}) - \varphi_V(\mathcal{B})| \leq L \cdot D_H(\mathcal{A},\mathcal{B})^p,$$

and the proof is complete. □

Remark 2.2. Note that if the distance D is such that it metrizes the topology of weak convergence, then the result in Proposition 2.3 quantifies those in Proposition 2.1 (a) and Proposition 2.2 (a).

Remark 2.3. Consider the formulation of a classical two-stage stochastic program with recourse, where the recourse function $\mathcal{Q}$ is given by (1.3)-(1.4). When $Y = \mathbb{R}^{n_2}$, basic assumptions guarantee that the recourse function is a well-defined, real-valued, convex function on $\mathbb{R}^{n_1}$. Since $\mathcal{Q}(\cdot, \mu)$ is convex for $\mu \in \mathcal{P}(\mathbb{R}^N)$, joint continuity of $\mathcal{Q}$ with respect to x and μ is implied by continuity of $\mathcal{Q}(x, \cdot)$ for all $x \in X$. (See e.g. Rockafellar [26, Theorem 10.7].) The continuity of $\mathcal{Q}$ with respect to μ, on the other hand, typically requires some uniform integrability condition to be satisfied. Such an approach was formalized by e.g. Robinson and Wets [25]. To arrive at quantitative continuity results of $\mathcal{Q}(x, \cdot)$ for $x \in X$, on the other hand, it is necessary to equip the set $\mathcal{P}(\mathbb{R}^N)$ with some suitable distance D, preferably metrizing (at least locally) the topology of weak convergence. Now, employing some suitable uniform integrability condition one may establish quantitative continuity results, for example of the form required in Proposition 2.3. For a collection of such results, related in particular to the bounded Lipschitz metric and an L_p-Wasserstein metric, we refer to Römisch and Schultz [28, 29, 30].

Remark 2.4. Consider again the formulation of the recourse function $\mathcal{Q}$ given by (1.3)-(1.4). When Y contains integrality restrictions on some or all of the second-stage variables, the analysis is somewhat impeded by the fact that the recourse function $\mathcal{Q}$ is no longer necessarily convex nor even continuous. Structural properties have mainly been investigated for the case of fixed second-stage cost q. To arrive at continuity of $\mathcal{Q}(\cdot, \cdot)$ at some point (x, μ) in this setting, the above-mentioned uniform integrability condition must be combined with some assumption guaranteeing that the set of those $(h, T) \in \mathbb{R}^{m \times (1+n_1)}$, for which the second-stage value function (1.4) is discontinuous at (x, h, T), has μ-measure zero. Using this line of approach, Schultz [33] established qualitative continuity of $\mathcal{Q}(\cdot, \cdot)$ under the assumption that the conditional distribution of h given T is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m$. The results were quantified in Schultz [34] for problems with fixed technology matrix.

Remark 2.5. Several alternatives to the formulation of the recourse function (1.3) have been proposed by various authors. Riis and Schultz [23], for example, proposed the following formulation, defining the recourse function as the probability of total cost exceeding some given threshold value ϕ,

$$\mathcal{Q}(x, \mu) := \mu\Big(\{\xi \in \mathbb{R}^N \mid cx + \Phi(x, \xi) > \phi\}\Big),$$

where the second-stage value function is still defined by (1.4). Riis and Schultz showed that this formulation is in fact equivalent to a classical two-stage stochastic program with mixed-integer recourse and hence is a special case of (1.3). Moreover, it was shown that the above-mentioned continuity properties remain valid, even under simplified assumptions.

3 Solution Procedure for Linear Recourse Models

In this section, we consider the minimax problem (1.2) in the setting of a two-stage stochastic program with linear recourse, i.e. the recourse function is given by

$$\mathcal{Q}(x,\mu) := cx + \int_{\mathbb{R}^N} \Phi(x,\xi)\mu(d\xi), \tag{3.1}$$

where the random vector $\xi \in \mathbb{R}^N$ is constituted by the components of the second-stage cost q, the second-stage right-hand side h and the technology matrix T, and the second-stage value function is given by

$$\Phi(x,\xi) := \min\Big\{qy \;\Big|\; Wy \geq h + Tx,\ y \in \mathbb{R}^{n_2}_+\Big\}. \tag{3.2}$$

We elaborate a solution procedure for the problem under the following assumptions.

(A1) The set of conceivable distributions $\mathcal{A}$ is defined as the set of all probability measures μ with support in some finite set $\Xi = \{\xi_1, \ldots, \xi_S\}$, satisfying a number of generalized moment conditions,

$$\int_\Xi g_i(\xi)\mu(d\xi) \leq \alpha_i, \quad i = 1, \ldots, L.$$

(A2) For all $t \in \mathbb{R}^m$ there exists $y \in \mathbb{R}^{n_2}_+$ such that $Wy \geq t$.

(A3) For all $\xi \in \Xi$ there exists $\pi \in \mathbb{R}^m_+$ such that $\pi W \leq q$.

According to Assumption (A1) attention is restricted to a finite number of scenarios, each scenario ξ_s corresponding to a possible outcome of random parameters (q_s, h_s, T_s). This assumption is justified by the stability results established in the previous section, since the optimal solution of a minimax problem, employing a more general definition of the set of conceivable distributions, may be approximated to any given accuracy by solutions of minimax problems, employing only sets of probability measures with support in known finite sets (cf. Remark 2.1 and Remark 2.3).

Assumption (A2) is the assumption of complete recourse, ensuring feasibility of the second-stage problem for any right-hand side $t \in \mathbb{R}^m$, whereas Assumption (A3) is the assumption of dual feasibility, employed to ensure boundedness of the second-stage problems. We note that the assumption of complete recourse may be replaced by the weaker assumption of relatively complete recourse, ensuring feasibility of the second-stage problems only for those right-hand sides that may actually occur, i.e. for all $x \in X$ and all $\xi \in \Xi$ there exists $y \in \mathbb{R}^{n_2}_+$ such that $Wy \geq h + Tx$. If relatively complete recourse is not inherent in the problem, it may be established by the inclusion of feasibility cuts (see e.g. Birge and Louveaux [4]).

Employing Assumption (A1) we may reformulate the minimax problem (1.2) in terms of the scenario probabilities $p_1, \ldots, p_S$ as follows

$$\min\Big\{ f(x) := cx + \max_{p \in \mathcal{P}} \sum_{s=1}^{S} p_s \Phi(x, \xi_s) \Big\} \tag{3.3}$$

where

$$\mathcal{P} = \Big\{ p \in \mathbb{R}^S_+ \;\Big|\; \sum_{s=1}^{S} p_s = 1,\ \sum_{s=1}^{S} p_s g_i(\xi_s) \le \alpha_i,\ i = 1 \ldots, L \Big\}. \tag{3.4}$$

Note that the inner maximization problem is now a linear programming problem over a bounded polyhedron and hence the maximum value is always attained.

The following proposition states the relevant structural properties of the function f.

Proposition 3.1. *Assume (A1)-(A3) and let f be defined by (3.2) and (3.3). Then f is a real-valued, piecewise linear and convex function on X.*

Proof. It is well-known that for each $p \in \mathcal{P}$ the function $g_p(x) = cx + \sum_{s=1}^{S} p_s \Phi(x, \xi_s)$ is a real-valued, piecewise linear and convex function on X, cf. e.g. Birge and Louveaux [4]. Noting that $\mathcal{P}$ is a bounded polyhedron, we see that f is the maximum of a finite number of real-valued, piecewise linear and convex functions, corresponding to the extreme points of $\mathcal{P}$. □

The solution procedure that we will propose is based on the following reformulation of problem (3.3),

$$\begin{aligned} &\min\ cx + \theta \\ &\text{s.t. } \theta \ge \sum_{s=1}^{S} p_s \Phi(x, \xi_s), \quad p \in \mathcal{P}, \\ &\qquad x \in X,\ \theta \in \mathbb{R}. \end{aligned} \tag{3.5}$$

As in L-shaped decomposition, the constraints of this problem may be replaced by linear inequalities referred to as optimality cuts. The algorithm progresses by sequentially solving a master problem and adding optimality cuts which are violated at the current solution. In particular, given a solution (x^ν, θ^ν) of the master problem in iteration ν, the second-stage problems are solved to obtain $w_s^\nu = \Phi(x^\nu, \xi_s) = \pi_s^\nu (h_s + T_s x^\nu)$ where π_s^ν are optimal dual solutions for $s = 1, \ldots, S$. Next, denoting by p^ν the optimal solution of the linear programming problem $w^\nu = \max_{p \in \mathcal{P}} \sum_{s=1}^{S} w_s^\nu p_s$, we obtain the following inequality that is valid for all $x \in X$ and binding at $x = x^\nu$,

$$\max_{p \in \mathcal{P}} \sum_{s=1}^{S} p_s \Phi(x, \xi_s) \ge \sum_{s=1}^{S} p_s^\nu \pi_s^\nu (h_s + T_s x).$$

In this way we see that if $\theta^\nu < w^\nu$, the current solution (x^ν, θ^ν) is cut off by including the following constraint in the master problem,

$$\theta \geq \sum_{s=1}^{S} p_s^\nu \pi_s^\nu (h_s + T_s x).$$

Algorithm 1

Step 1 (*Initialization*) Set $\nu = 0$ and $\bar{z} = \infty$ and let the current master problem be $\min\{cx + \theta \mid x \in X,\ \theta \in \mathbb{R}\}$.

Step 2 (*Solve master problem*) Solve the current master problem. Let (x^ν, θ^ν) be an optimal solution vector if one exists; if the problem is unbounded, then let (x^ν, θ^ν) be a feasible solution with $cx^\nu + \theta^\nu < \bar{z}$

Step 3 (*Termination*) If $cx^\nu + \theta^\nu = \bar{z}$, stop; the current solution is optimal.

Step 4 (*Solve subproblems*) For $s = 1, \ldots, S$ solve the second-stage problems to find $w_s^\nu = \Phi(x^\nu, \xi_s)$.

Step 5 (*Solve max-problem*) Solve the problem $w^\nu = \max_{p \in \mathcal{P}} \sum_{s=1}^{S} w_s^\nu p_s$. If $\theta^\nu < w^\nu$ a violated optimality cut is added to the master problem.

Step 6 (*Update bound*) Let $\bar{z} := \min\{\bar{z}, cx^\nu + w^\nu\}$. Go to Step 2.

It is easily seen that Algorithm 1 terminates in a finite number of iterations whenever a solution to the minimax problem exists.

Proposition 3.2. *Assume (A1)-(A3). If problem (3.5) is feasible and bounded, then Algorithm 1 terminates with an optimal solution in a finite number of iterations.*

Proof. Assume that the minimax problem has an optimal solution $\bar{x}$. In any iteration ν of the algorithm we must have $cx^\nu + \theta^\nu \leq c\bar{x} + \max_{p \in \mathcal{P}} \sum_{s=1}^{S} p_s \Phi(\bar{x}, \xi_s)$, since the master problem is a relaxation of problem (3.5). As mentioned above, the current solution (x^ν, θ^ν) is cut off by an optimality cut whenever $\theta^\nu < \max_{p \in \mathcal{P}} \sum_{s=1}^{S} p_s \Phi(x^\nu, \xi_s)$. This can only happen a finite number of times since the number of optimality cuts is finite, cf. Proposition 3.1. Thus we will eventually have $\theta^\nu = \max_{p \in \mathcal{P}} \sum_{s=1}^{S} p_s \Phi(x^\nu, \xi_s)$, at which point the current solution is feasible for problem (3.5) and hence optimal. □

Consider now the following alternative reformulation of problem (3.3),

$$\begin{aligned} \min\ & cx + \theta \\ \text{s.t. } & \theta \geq \sum_{s=1}^{S} p_s \sigma_s, \quad p \in \mathcal{P}, \\ & \sigma_s \geq \Phi(x, \xi_s), \quad s = 1, \ldots, S, \\ & x \in X,\ \theta \in \mathbb{R},\ \sigma \in \mathbb{R}^S. \end{aligned}$$

This reformulation leads directly to the following multicut version of Algorithm 1.

Algorithm 2

Step 1 (*Initialization*) Set $\nu = 0$ and $\bar{z} = \infty$ and let the current master problem be $\min\{cx + \theta \mid x \in X,\ \theta \in \mathbb{R},\ \sigma \in \mathbb{R}^S\}$.

Step 2 (*Solve master problem*) Solve the current master problem. Let $(x^\nu, \theta^\nu, \sigma^\nu)$ be an optimal solution vector if one exists; if the problem is unbounded, then let $(x^\nu, \theta^\nu, \sigma^\nu)$ be a feasible solution with $cx^\nu + \theta^\nu < \bar{z}$.

Step 3 (*Termination*) If $cx^\nu + \theta^\nu = \bar{z}$, stop; the current solution is optimal.

Step 4 (*Solve subproblems*) For $s = 1, \ldots, S$ solve the second-stage problems to find $w_s^\nu = \Phi(x^\nu, \xi_s) = \pi_s^\nu(h_s + T_s x^\nu)$. If $\sigma_s^\nu < w_s^\nu$, add the cut $\sigma_s \geq \pi_s^\nu(h_s + T_s x)$ to the master problem.

Step 5 (*Solve max-problem*) Solve the problem $w^\nu = \max_{p \in \mathcal{P}} \sum_{s=1}^S w_s^\nu p_s$ and let p^ν be an optimal solution. If $\theta^\nu < \sum_{s=1}^S p_s^\nu \sigma_s^\nu$, add the cut $\theta \geq \sum_{s=1}^S p_s^\nu \sigma_s$ to the master problem.

Step 6 (*Update bound*) Let $\bar{z} := \min\{\bar{z}, cx^\nu + w^\nu\}$. Go to Step 2.

Proposition 3.3. *Assume (A1)-(A3). If problem (3.5) is feasible and bounded then Algorithm 2 terminates with an optimal solution in a finite number of iterations.*

Proof. The proof is similar to that of Proposition 3.2. □

Remark 3.1. The multicut approach of Algorithm 2 offers some computational advantages compared to the single-cut approach of Algorithm 1, since more detailed information is passed to the master problem in each iteration. The improved detailing, however, comes at the cost of an increased complexity of the master problem since the size of the problem grows quite rapidly. Also, due to their resemblance to the original L-shaped method and its multicut version, respectively, the algorithms are bound to suffer from some of the same drawbacks. Apart from the growing size of the master problem, such drawbacks include the tendency for early iterations to oscillate heavily, causing slow convergence towards an optimal solution. In the case of L-shaped decomposition some of these drawbacks were circumvented by the regularized decomposition method introduced by Ruszczynski [31]. The idea is to introduce an incumbent solution a^ν and include a quadratic regularizing term of the form $\frac{\alpha}{2}\|x - a^\nu\|^2$ in the objective. Clearly, a similar approach could be used for the algorithm presented here, but we will not go into the details of such an implementation.

Remark 3.2. Riis and Schultz [23] elaborated a solution procedure to find the minimum value of a recourse function given by

$$\mathcal{Q}(x,\mu) := \mu\Big(\big\{\xi \in \mathbb{R}^N \mid cx + \Phi(x,\xi) > \phi\big\}\Big), \tag{3.6}$$

where ϕ is some given threshold value. This procedure is in many ways similar to the multicut version of the L-shaped algorithm. Hence we may modify Algorithm 2 in a similar way to obtain a solution procedure for the minimax problem (1.2) with $\mathcal{Q}$ defined by (3.6).

4 Solution Procedure for Integer Recourse Models

If second-stage variables are restricted to integer values, the solution procedures presented in the previous section break down, since the second-stage value function is no longer convex and piecewise linear, but in fact only lower semicontinuous. To solve the minimax problem (1.2) in this setting, we elaborate an extension of the branch and bound algorithm for two-stage stochastic programs with integer recourse, proposed by Ahmed, Tawarmalani and Sahinidis [1]. As in (3.1) the recourse function is defined as the sum of first-stage cost and expected second-stage cost, whereas the second-stage value function is now given by

$$\Phi(x,\xi) := \min\Big\{qy \;\Big|\; Wy \geq h + Tx,\ y \in \mathbb{Z}^{n_2}_+\Big\}. \tag{4.1}$$

The solution procedure is elaborated under the following assumptions.

(B1) The set of conceivable distributions $\mathcal{A}$ is defined as the set of all probability measures μ with support in some finite set $\Xi = \{\xi_1, \ldots, \xi_S\}$, satisfying a number of generalized moment conditions,

$$\int_\Xi g_i(\xi)\mu(d\xi) \leq \alpha_i, \quad i = 1, \ldots, L.$$

(B2) For all $t \in \mathbb{R}^m$ there exists $y \in \mathbb{Z}^{n_2}$ such that $Wy \geq t$.

(B3) For all $\xi \in \Xi$ there exists $\pi \in \mathbb{R}^m_+$ such that $\pi W \leq q$.

Note that Assumption (B2) is a natural extension of the complete recourse assumption for the integer recourse case, while Assumption (B3) is in fact the same as Assumption (A3). Employing Assumption (B1), we may once again reformulate the minimax problem (1.2) as

$$\min\Big\{f(x) := cx + \max_{p\in\mathcal{P}} \sum_{s=1}^{S} p_s \Phi(x,\xi_s)\Big\}, \tag{4.2}$$

where the set $\mathcal{P}$ is still defined by (3.4). According to the following proposition, problem (4.2) is well-defined in the sense that one minimizes a real-valued, lower semicontinuous function.

Proposition 4.1. *Assume (B1)-(B3) and let f be defined by (4.1) and (4.2). Then f is a real-valued, lower semicontinuous function on X.*

Proof. For each $p \in \mathcal{P}$ the function $g_p(x) = \sum_{s=1}^{S} cx + p_s\Phi(x, \xi_s)$ is a real-valued, lower semicontinuous function on X, cf. e.g. Nemhauser and Wolsey [21]. As in the proof of Proposition 3.1 we see that f is real-valued. Now, let $x \in X$ and let $\{x_n\}_{n\geq 1}$ be some sequence in X converging to x. Assuming that $\bar{p} \in \arg\max_{p\in\mathcal{P}} g_p(x)$, we have

$$\liminf_{n\to\infty} f(x_n) \geq \liminf_{n\to\infty} g_{\bar{p}}(x_n) \geq g_{\bar{p}}(x) = f(x),$$

which completes the proof. □

The algorithm presented in [1] is based on the following assumptions.

(B4) The technology matrix is fixed, i.e. $T_s = T$ for $s = 1, \ldots, S$.

(B5) The first-stage constraint set X is non-empty and compact.

(B6) The recourse matrix W is integral.

Employing Assumption (B4), we may reformulate problem (4.2) by introducing the variable transformation $\chi = Tx$ for $x \in X$ to obtain the following formulation,

$$\min_{\chi\in\mathcal{X}}\Big\{F(\chi) := h(\chi) + H(\chi)\Big\} \tag{4.3}$$

where

$$\begin{aligned} h(\chi) &= \min\big\{cx \mid Tx = \chi,\ x \in X\big\}, \\ H(\chi) &= \max_{p\in\mathcal{P}}\Big\{\overline{\Psi}_p(\chi) := \sum_{s=1}^{S} p_s\Psi_s(\chi)\Big\}, \\ \Psi_s(\chi) &= \min\big\{q_s y \mid Wy \geq h_s + \chi,\ y \in \mathbb{Z}^{n_2}\big\}, \end{aligned}$$

and

$$\mathcal{X} = \big\{\chi \in \mathbb{R}^m \mid \chi = Tx,\ x \in X\big\}.$$

As in the proof of Proposition 4.1 it is easily seen that the function $H(\cdot)$ is real-valued and lower semicontinuous, and since $h(\cdot)$ is clearly real-valued and continuous, we see that problem (4.3) is well-defined, and under Assumption (B5) the optimal value exists and is actually attained for some $\chi^* \in \mathcal{X}$. Furthermore, given an optimal solution $\chi^* \in \mathcal{X}$ of the transformed problem (4.3), it is easily seen that $x^* \in X$ is an optimal solution of the minimax problem (4.2) if $x^* \in \arg\min\big\{cx \mid Tx = \chi^*,\ x \in X\big\}$, cf. [1, Theorem 3.2].

The insight of Ahmed et al. was to note that for any $p \in \mathcal{P}$ the discontinuity points of the function $\overline{\Psi}_p(\cdot)$ are contained in a finite number of hyperplanes that are all orthogonal to the variable axes. In fact, they show that the function is piecewise constant over rectangular regions of $\mathcal{X}$, the boundaries of which are orthogonal to the variable axes. This result leads directly to the following.

Lemma 4.1. *Assume (B1)-(B6), let* $k = (k_1^1, \ldots, k_j^s, \ldots, k_m^S) \in \mathbb{Z}^{mS}$ *be a vector of integers, and let*

$$C(k) := \bigcap_{s=1}^{S} \prod_{j=1}^{m} (k_j^s - h_j^s - 1, k_j^s - h_j^s]$$

and

$$\mathcal{K} := \{k \in \mathbb{Z}^{mS} \mid C(k) \cap \mathcal{X} \neq \emptyset\}.$$

Then $|\mathcal{K}| < \infty$ *and for all* $k \in \mathcal{K}$ *the function* $H(\cdot)$ *is constant over* $C(k)$.

Proof. According to [1, Theorem 4.4 and 4.5] the result is true for the function $\overline{\Psi}_p(\cdot)$ for any $p \in \mathcal{P}$. The proposition follows immediately. □

The branch and bound algorithm, formally stated below, proceeds by partitioning the feasible set $\mathcal{X}$ into regions of the form $\mathcal{X} \cap \Pi_{j=1}^m (l_j, u_j]$, where each l_j, $j = 1, \ldots, m$, is a possible point of discontinuity of $H(\cdot)$, i.e. $l_j + h_j^s$ is integral for some $s = 1, \ldots, S$. This is combined with a specialized bounding procedure which is a simple generalization of the one presented by Ahmed et al.

Algorithm 3

Step 1 (*Initialization*) Set $\bar{z} = \infty$. Let $l^P, u^P \in \mathbb{R}^m$ be such that $\mathcal{X} \subseteq \Pi_{j=1}^m (l_j^P, u_j^P]$ and for all $j = 1, \ldots, m$, $l_j^P + h_j^s$ is integral for some $s = 1, \ldots, S$. Let $\mathcal{L}$ consist of problem P defined by (4.3) with the additional constraints $l^P < \chi \leq u^P$. Also, let $\epsilon \in \mathbb{R}^m$ be such that $H(\cdot)$ is constant over $\Pi_{j=1}^m (l_j, l_j + \epsilon_j]$ whenever $l \in \mathbb{R}^m$ is such that for all $j = 1, \ldots, m$, $l_j + h_j^s$ is integral for some $s = 1, \ldots, S$.

Step 2 (*Termination*) If $\mathcal{L} = \emptyset$, stop. The solution that yielded the upper bound $\bar{z}$ is optimal.

Step 3 (*Node selection*) Select and remove from the list of open problems $\mathcal{L}$ a problem P, defined as $\min\{F(\chi) \mid l^P < \chi \leq u^P,\ \chi \in \mathcal{X}\}$.

Step 4 (*Bounding*) Obtain a lower bound on P by solving the lower bounding problem $z^P := H(l^P + \epsilon) + \min\{cx \mid Tx = \chi,\ l^P \leq \chi \leq u^P,\ x \in X\}$ and let χ^P be an optimal solution. Update the upper bound as $\bar{z} := \min\{\bar{z}, F(\chi^P)\}$ and remove from $\mathcal{L}$ all problems P' with $z^{P'} \geq \bar{z}$. If $z^P \geq \bar{z}$ go to Step 2.

Step 5 (*Branching*) Select an index $j \in \{1, \ldots, m\}$ and a value v_j with $l_j^P < v_j < u_j^P$ and such that $v_j + h_j^s$ is integral for some $s = 1, \ldots, S$. Construct two new problems P' and P'', obtained from P by adding the constraints $\chi_j > v_j$ and $\chi_j \leq v_j$, respectively. Let $z^{P'} = z^{P''} = z^P$ and add the two problems to $\mathcal{L}$.

Before we prove finite termination of Algorithm 3, let us note that Ahmed et al. presented a procedure for the a priori determination of the constant ϵ. The procedure is based on the result in Lemma 4.1 and simply determines the smallest possible width of the non-empty regions $C(k)$, $k \in \mathbb{Z}^{mS}$. Observing the definition of ϵ, it is easily seen that the optimal value z^P of the lower bounding problem in Step 4 of the algorithm is a lower bound on the optimal value of the current problem P. In particular, since $H(\cdot)$ is clearly non-decreasing in χ, the definition of ϵ implies that $H(l^P + \epsilon)$ is a lower bound on $H(\chi)$ for $l^P < \chi \leq u^P$. Moreover, $\min\{cx \mid Tx = \chi,\ l^P \leq \chi \leq u^P\ x \in X\}$ is clearly a lower bound on $h(\chi)$ for $l^P < \chi \leq u^P$.

Proposition 4.2. *Assume (B1)-(B6). Then Algorithm 3 terminates with an optimal solution in a finite number of iterations.*

Proof. Suppose in some iteration of the algorithm that the current problem P is such that $H(\cdot)$ is constant over the set $\{\chi \in \mathcal{X} \mid l^P < \chi \leq u^P\}$. Then $H(\chi^P) \leq H(l^P + \epsilon)$ so that $F(\chi^P) = h(\chi^P) + H(\chi^P) \leq h(\chi^P) + H(l^P + \epsilon) = z^P$ and the current problem is fathomed with no further refinements of the partition. Thus branching only occurs whenever the set $\{\chi \in \mathcal{X} \mid l^P < \chi \leq u^P\}$ contains a discontinuity point of $H(\cdot)$. By Lemma 4.1 and the definition of Step 5, this can only happen a finite number of times and hence the algorithm terminates in a finite number of iterations. Optimality follows from validity of the lower and upper bounding procedures, cf. the proof of [1, Theorem 6.4]. □

For further implementational details, such as e.g. specification of the branching rule and improvements of the lower bounding procedure as well as extension of the algorithm to the case of random technology matrix, we refer to Ahmed, Tawarmalani and Sahinidis [1].

5 Conclusions

In this paper we have considered the minimax approach to stochastic programming. Considering some objective function $\mathcal{Q}$ which is dependent on a decision x and the probability distribution of random parameters μ, the problem is to minimize the function f defined by $f(x) = \sup_{\mu \in \mathcal{A}} \mathcal{Q}(x, \mu)$, where $\mathcal{A}$ is a given set of possible or conceivable distributions. Considering this formulation of the problem, we have established a number of stability properties in the form of continuity properties of local optimal values and local optimal solutions

when the set $\mathcal{A}$ is subjected to perturbations (Proposition 2.1, Proposition 2.2 and Proposition 2.3). The stability results justify in particular the approach of approximating the set $\mathcal{A}$ by simpler sets, consisting only of measures with support in some known finite set. Using this approach, we have elaborated solution procedures for the minimax problem in the setting of stochastic recourse models, considering the linear recourse case as well as the integer recourse case. In the linear recourse case, the objective function f is piecewise linear and convex (Proposition 3.1), and hence modifications of the L-shaped algorithm and its multicut version apply and are easily seen to converge to an optimal solution in a finite number of iterations (Proposition 3.2 and Proposition 3.3). In the integer recourse case the function f is no longer convex but in fact only lower semicontinuous (Proposition 4.1). Introducing a proper variable transformation and observing that the transformed objective is piecewise constant on certain rectangular regions (Lemma 4.1), a branch and bound procedure for two-stage stochastic integer programs may be modified to solve the minimax problem, and once again convergence to an optimal solution in a finite number of iterations is guaranteed (Proposition 4.2).

References

[1] S. Ahmed, M. Tawarmalani, and N.V. Sahinidis. A finite branch and bound algorithm for two-stage stochastic integer programs. Submitted for publication. Available at http://www.isye.gatech.edu/~sahmed, 2000.

[2] J.R. Birge and J.H. Dulá. Bounding separable recourse functions with limited distribution information. *Anaals of Operations Research*, 30:277–298, 1991.

[3] J.R. Birge and F.V. Louveaux. A multicut algorithm for two-stage stochastic linear programs. *European Journal of Operational Research*, 34:384–392, 1988.

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

[5] J.R. Birge and R.J-B Wets. Computing bounds for stochastic programming problems by means of a generalized moment problem. *Mathematics of Operations Research*, 12:149–162, 1987.

[6] M. Breton and S.E. Hachem. Algorithms for the solution of stochastic dynamic minimax problems. *Computational Optimization and Applications*, 4:317–345, 1995.

[7] M. Breton and S.E. Hachem. A scenario aggregation algorithm for the solution of stochastic dynamic minimax problems. *Stochastic and Stochastics Reports*, 53:305–322, 1995.

[8] C.C. Carøe and R. Schultz. Dual decomposition in stochastic integer programming. *Operations Research Letters*, 24:37–45, 1999.

[9] J. Dupačová. The minimax approach to stochastic programming and an illustrative application. *Stochastics*, 20:73–88, 1987.

[10] J. Dupačová. Stochastic programming with incomplete information: A survey of results on postoptimization and sensitivity analysis. *Optimization*, 18:507–532, 1987.

[11] J. Dupačová. Stability and sensitivity analysis for stochastic programming. *Annals of Operations Research*, 27:115–142, 1990.

[12] Y. Ermoliev, A. Gaivoronski, and C. Nedeva. Stochastic optimization problems with partially known distribution functions. *SIAM Journal on Control and Optimization*, 23:697–716, 1985.

[13] J.L. Higle and S. Sen. Stochastic decomposition: An algorithm for two-stage linear programs with recourse. *Mathematics of Operations Research*, 16:650–669, 1991.

[14] P. Kall. On approximations and stability in stochastic programming. In J. Guddat, H.Th. Jongen, B. Kummer, and F. Nožička, editors, *Parametric Optimization and Related Topics*, pages 387–407. Akademie Verlag, Berlin, 1987.

[15] P. Kall. An upper bound for stochastic linear programming using first and total second moments. *Annals of Operations Research*, 30:267–276, 1991.

[16] P. Kall and S.W. Wallace. *Stochastic Programming*. John Wiley and Sons, Chichester, UK, 1994.

[17] A.F. Karr. Extreme points of certain sets of probability measures, with applications. *Mathematics of Operations Research*, 8:74–85, 1983.

[18] J.H.B. Kempermann. The general moment problem, a geometric approach. *Annals of Mathematical Statistics*, 39:93–122, 1968.

[19] D. Klatte. A note on quantitative stability results in nonlinear optimization. In K. Lommatzsch, editor, *Proceedings of the 19. Jahrestagung Mathematische Optimierung*, Seminarbericht Nr. 90, pages 77–86, Sektion Mathematik, Humboldt-Universität Berlin, 1987.

[20] G. Laporte and F.V. Louveaux. The integer L-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters*, 13:133–142, 1993.

[21] G.L. Nemhauser and L.A. Wolsey. *Integer and Combinatorial Optimization*. Wiley Interscience, New York, 1988.

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

[23] M. Riis and R. Schultz. Applying the minimum risk criterion in stochastic recourse programs. Working Paper 2001/3, University of Aarhus, Department of Operations Research, 2001. Available at http://home.imf.au.dk/riis/.

[24] S.M. Robinson. Local epi-continuity and local optimization. *Mathematical Programming*, 37:208–222, 1987.

[25] S.M. Robinson and R.J-B Wets. Stability in two-stage stochastic programming. *SIAM Journal on Control and Optimization*, 25:1409–1416, 1987.

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

[27] R.T. Rockafellar and R.J-B Wets. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16:119–147, 1991.

[28] W. Römisch and R. Schultz. Distribution sensitivity in stochastic programming. *Mathematical Programming*, 50:197–226, 1991.

[29] W. Römisch and R. Schultz. Stability analysis for stochastic programs. *Annals of Operations Research*, 30:241–266, 1991.

[30] W. Römisch and R. Schultz. Stability of solutions for stochastic programs with complete recourse. *Mathematics of Operations Research*, 18:590–609, 1993.

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

[32] R. Schultz. Continuity properties of expectation functions in stochastic integer programming. *Mathematics of Operations Research*, 18:578–589, 1993.

[33] R. Schultz. On structure and stability in stochastic programs with random technology matrix and complete integer recourse. *Mathematical Programming*, 70:73–89, 1995.

[34] R. Schultz. Rates of convergence in stochastic programs with complete integer recourse. *SIAM Journal on Optimization*, 6:1138–1152, 1996.

[35] R. Schultz, L. Stougie, and M.H. van der Vlerk. Solving stochastic programs with integer recourse by enumeration: A framework using Gröbner basis reductions. *Mathematical Programming*, 83:229–252, 1998.

[36] A. Shapiro. Quantitative stability in stochastic programming. *Mathematical Programming*, 67:99–108, 1994.

[37] A. Shapiro and A. Kleywegt. Minimax analysis of stochastic problems. *Optimization Methods and Software* (to appear).

[38] R.M. Van Slyke and R.J-B Wets. L-shaped linear programs with applications to optimal control and stochastic linear programming. *SIAM Journal of Applied Mathematics*, 17:638–663, 1969.

[39] J. Wang. Distribution sensitivity analysis for stochastic programs with complete recourse. *Mathematical Programming*, 31:286–297, 1985.

[40] J. Žáčková. On minimax solutions of stochastic linear programming problems. *Časopsis pro Pěstování Matematiky*, 91:423–430, 1966.