Multi-stage Adjustable Robust Mixed-Integer Optimization via Iterative Splitting of the Uncertainty set

Postek, K.S.; den Hertog, D.

Publication date:
2014

Document Version
Early version, also known as pre-print

Citation for published version (APA):
Postek, K. S., & den Hertog, D. (2014). *Multi-stage Adjustable Robust Mixed-Integer Optimization via Iterative Splitting of the Uncertainty set*. (CentER Discussion Paper; Vol. 2014-056). Operations research.

CentER Discussion Paper

No. 2014-056

MULTI-STAGE ADJUSTABLE ROBUST MIXED-INTEGER OPTIMIZATION VIA ITERATIVE SPLITTING OF THE UNCERTAINTY SET

By

Krzysztof Postek, Dick den Hertog

29 September, 2014

ISSN 0924-7815
ISSN 2213-9532

TILBURG UNIVERSITY

Multi-stage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set

Krzysztof Postek*† Dick den Hertog*

September 25, 2014

Abstract

In this paper we propose a methodology for constructing decision rules for integer and continuous decision variables in multiperiod robust linear optimization problems. This type of problems finds application in, for example, inventory management, lot sizing, and manpower management. We show that by iteratively splitting the uncertainty set into subsets one can differentiate the later-period decisions based on the revealed uncertain parameters. At the same time, the problem's computational complexity stays at the same level as for the static robust problem. This holds also in the non-fixed recourse situation. In the fixed recourse situation our approach can be combined with linear decision rules for the continuous decision variables. We provide theoretical results how to split the uncertainty set by identifying sets of uncertain parameter scenarios to be divided for an improvement in the worst-case objective value. Based on this theory, we propose several splitting heuristics. Numerical examples entailing a capital budgeting and a lot sizing problem illustrate the advantages of the proposed approach.

Keywords: adjustable, decision rules, integer, multi-stage, robust optimization
JEL codes: C61

1 Introduction

Robust optimization (RO, see Ben-Tal et al. (2009)) has become one of the main approaches to optimization under uncertainty. One of its applications are multiperiod problems where, period after period, values of the uncertain parameters are revealed and new decisions are implemented. Adjustable Robust Optimization (ARO, see Ben-Tal et al. (2004)) addresses such problems by formulating the decision variables as functions of the revealed uncertain parameters. Ben-Tal et al.

*CentER and Department of Econometrics and Operations Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands
†Correspondence to: k.postek@tilburguniversity.edu

(2004) prove that without any functional restrictions on the form of adjustability, the resulting problem is NP-hard. For that reason, several functional forms of the decision rules have been proposed, with the most popular being the affinely adjustable decision rules. However, only for a limited class of problems do they yield problems that can be reformulated to a computationally tractable form (see Ben-Tal et al. (2009)). In particular, for problems without fixed recourse, where the later-period problem parameters depend also on the uncertain parameters from earlier periods, it is nontrivial to construct tractable decision rules. The difficulty level grows even more when the adjustable variables are binary or integer. Addressing this problem is the topic of our paper. We propose a method to construct computationally tractable adjustable decision rules, applicable also to problems with integer adjustable variables and to problems without fixed recourse. For problems with fixed recourse our methodology can be combined with linear decision rules for the continuous decision variables.

The contribution of our paper is twofold. First, we propose a methodology of iterative splitting of the uncertainty set into subsets, for each of which a scalar later-period decision shall be determined. A given decision is implemented in the next period if the revealed uncertain parameter belongs to the corresponding subset. Using scalar decisions per subset ensures that the resulting problem has the same complexity as the static robust problem. This approach provides an upper bound on the optimal value of the adjustable robust problem. Next to that, we propose a method of obtaining lower bounds, being a generalization of the approach of Hadjiyiannis et al. (2011).

As a second contribution, we provide theoretical results supporting the decision of how to split the uncertainty set into smaller subsets for problems with continuous decision variables. The theory identifies sets of scenarios for the uncertain parameters that have to be divided. On the basis of these results, we propose set-splitting heuristics for problems including also integer decision variables. As a side result, we prove the reverse of the result of Gorissen et al. (2014). Namely, we show that the optimal KKT vector of the tractable robust counterpart of a linear robust problem, obtained using the results of Ben-Tal et al. (2014), yields an optimal solution to the optimistic dual (see Beck and Ben-Tal (2009)) of the original problem.

ARO was developed to (approximately) solve problems with continuous variables. Ben-Tal et al. (2004) introduce the concept of using affinely adjustable decision rules and show how to apply such rules to obtain (approximate) optimal solutions to multiperiod problems. Their approach has been later extended to other function classes by Chen et al. (2007), Chen et al. (2009), Ben-Tal et al. (2009) and Bertsimas et al. (2011b). Bertsimas et al. (2010) prove that for a specific class of multiperiod control problems the affinely adjustable decision rules result in optimal adjustable solution. Bertsimas and Goyal (2010) show that the static robust solutions perform well also in stochastic programming problems.

Later, developments have been made allowing ARO to (approximately) solve problems involving adjustable integer variables. Bertsimas and Caramanis (2007) propose a sampling method for constructing adjustable robust decision rules ensuring, under certain conditions, that the robust constraints are satisfied with high probability. Bertsimas and Caramanis (2010) introduce the term of finite adaptability in two-period problems, with a fixed number of possible second-period decisions. They also show that finding the best values for these variables is NP-hard. In a later paper, Bertsimas et al. (2011a) characterize the geometric conditions for the uncertainty sets under which finite adaptability provides good approximations of the adjustable robust solutions.

Vayanos et al. (2011) split the uncertainty set into hyper-rectangles, assigning to each of them the corresponding later-period adjustable linear and binary variables. Contrary to this, our method does not impose any geometrical form of the uncertainty subsets. Bertsimas and Georghiou (2014a) propose to use piecewise linear decision rules, both for the continuous and the binary variables (for the binary variables, value 0 is implemented if the piecewise linear decision rule is positive). They use a cutting plane approach that gradually increases the fraction of the uncertainty set that the solution is robust to, reaching complete robustness when their approach terminates. In our approach, the decision rules proposed ensure full robustness after each of the so-called splitting rounds, and the more splitting rounds, the better the value of the objective function. In a recent paper, Bertsimas and Georghiou (2014b) propose a different type of decision rules for binary variables. Since the resulting problems are exponential in the size of the original formulation, authors propose their conservative approximations, giving a systematic tradeoff between computational tractability and level of conservatism. In our approach, instead of imposing a functional form of the decision rules, we focus on splitting the uncertainty set into subsets with different decisions. Also, we ensure robustness precisely against the specified uncertainty set and allow non-binary integer variables.

Hanasusanto et al. (2014) apply finite adaptability to two-period decision problems with binary variables, where the decision maker constructs a fixed number of time-2 policies and implements the best of them after the uncertain parameters are observed. The resulting problems can be transformed to MILP problems of size exponential relative to the number K of policies (in the non-fixed recourse situation - for fixed recourse problems the reformulation is polynomial). They also study the approximation quality provided by such reformulations and complexity issues. Our approach applies to general multi-period problems and allows also explicitly non-binary integer variables.

We test our methodology on problem instances from Bertsimas and Georghiou (2014a) and Hanasusanto et al. (2014). The experiments reveal the our methodology performs worse on small instances, where the 'more exact' approaches of other authors can be solved fast to optimality. However, as the problems grow in size, it is able to provide comparable or better results after a significantly shorter computation.

The composition of the remainder of the paper is as follows. Section 2 introduces the set-splitting methodology for the case of two-period problems with adjustable continuous variables. Section 3 extends the approach to multiperiod problems, and Section 4 extends the multiperiod case to problems with integer decision variables.

Section 5 proposes heuristics to be used as a part of the method. Section 6 gives two numerical examples, showing that the methodology of our paper offers substantial gains in terms of the worst-case objective function improvement. Section 7 concludes and lists the potential directions for future research.

2 Two-period problems

For ease of exposition we first introduce our methodology on the case of two-period problems with continuous decision variables only. The extension to multi-period problems is given in Section 3, and the extension to problems with integer variables is given in Section 4.

2.1 Description

Consider the following two-period optimization problem:

$$\begin{aligned} \min_{x_1,x_2} \quad & c_1^T x_1 + c_2^T x_2 \\ \text{s.t.} \quad & A_1(\zeta)x_1 + A_2(\zeta)x_2 \leq b \quad \forall \zeta \in \mathcal{Z}, \end{aligned} \tag{1}$$

where $c_1 \in \mathbb{R}^{d_1}, c_2 \in \mathbb{R}^{d_2}$, $b \in \mathbb{R}^m$ are fixed parameters, $\zeta \in \mathbb{R}^L$ is the uncertain parameter and $\mathcal{Z} \subset \mathbb{R}^L$ is a compact and convex uncertainty set. Vector $x_1 \in \mathbb{R}^{d_1}$ is the decision implemented at time 1 before the value of ζ is known, and $x_2 \in \mathbb{R}^{d_2}$ is the decision vector implemented at time 2, after the value of ζ is known. It is assumed that the functions $A_1 : \mathbb{R}^L \to \mathbb{R}^{m\times d_1}, A_2 : \mathbb{R}^L \to \mathbb{R}^{m\times d_2}$ are linear. We refer to the rows of matrix A_1 and A_2 as $a_{1,i}^T(\zeta)$ and $a_{2,i}^T(\zeta)$ respectively, with $a_{1,i}(\zeta) = P_{1,i}\zeta$ and $a_{2,i}(\zeta) = P_{2,i}\zeta$, where $P_{1,i} \in \mathbb{R}^{d_1\times L}, P_{2,i} \in \mathbb{R}^{d_2\times L}$ (uncertain parameter can contain a single fixed component, which would result in the intercepts of the affine transformations $A_1(\zeta), A_2(\zeta)$.

The static robust problem (1) where the decision vector x_2 is independent from the value of ζ makes no use of the fact that x_2 can adjust to the revealed ζ. The adjustable version of problem (1) is:

$$\begin{aligned} \min_{x_1,x_2(\zeta),z} \quad & z \\ \text{s.t.} \quad & c_1^T x_1 + c_2^T x_2(\zeta) \leq z, \quad \forall \zeta \in \mathcal{Z} \\ & A_1(\zeta)x_1 + A_2(\zeta)x_2(\zeta) \leq b \quad \forall \zeta \in \mathcal{Z}. \end{aligned} \tag{2}$$

Since this problem is NP-hard (see [3]), the concept of linear decision rules has been proposed. Then, the time 2 decision vector is defined as $x_2 = v + V\zeta$, where $v \in \mathbb{R}^{d_2}, V \in \mathbb{R}^{d_2\times L}$ (see [3]) and the problem is:

$$\begin{aligned} \min_{x_1,v,V} \quad & z \\ \text{s.t.} \quad & c_1^T x_1 + c_2^T (v + V\zeta) \leq z, \quad \forall \zeta \in \mathcal{Z} \\ & A_1(\zeta)x_1 + A_2(\zeta)(v + V\zeta) \leq b \quad \forall \zeta \in \mathcal{Z}. \end{aligned} \tag{3}$$

In the general case such constraints are quadratic in ζ, because of the term $A_2(\zeta)(v + V\zeta)$. Only for special cases the constraint system can be rewritten as

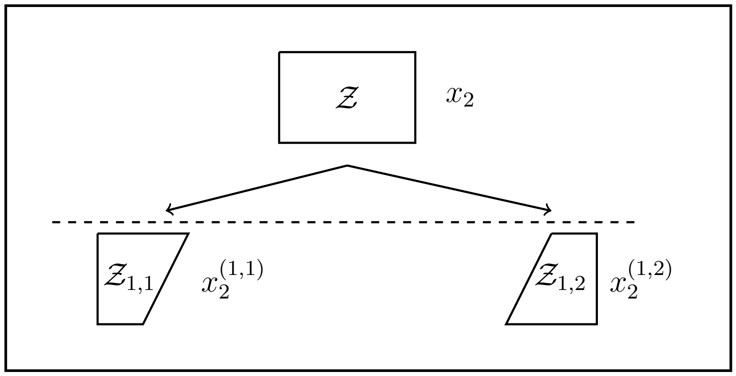

Figure 1: Scheme of the first splitting.

a computationally tractable system of inequalities. Moreover, linear decision rules cannot be used if (part of) the decision vector x_2 is required to be integer.

We propose a different approach. Its idea lies in splitting the set $\mathcal{Z}$ into a collection of subsets $\mathcal{Z}_{r,s}$ where $s \in \mathcal{N}_r$ and $\cup_{s\in\mathcal{N}_r}\mathcal{Z}_{r,s} = \mathcal{Z}$ (r denotes the index of the splitting round and s denotes the set index). For each $\mathcal{Z}_{r,s}$ a different, fixed time 2 decision shall be determined. We split the set $\mathcal{Z}$ in rounds into smaller and smaller subsets using hyperplanes. The following example illustrates this idea.

Example 1. We split the uncertainty set $\mathcal{Z}$ with a hyperplane $g^T\zeta = h$ into the following two sets:

$$\mathcal{Z}_{1,1} = \mathcal{Z} \cap \left\{\zeta : g^T\zeta \leq h\right\} \quad \text{and} \quad \mathcal{Z}_{1,2} = \mathcal{Z} \cap \left\{\zeta : g^T\zeta \geq h\right\}.$$

At time 2 the following decision is implemented:

$$x_2 = \begin{cases} x_2^{(1,1)} & \text{if } \zeta \in \mathcal{Z}_{1,1} \\ x_2^{(1,2)} & \text{if } \zeta \in \mathcal{Z}_{1,2} \\ x_2^{(1,1)} \text{ or } x_2^{(1,2)} & \text{if } \zeta \in \mathcal{Z}_{1,1} \cap \mathcal{Z}_{1,2}. \end{cases}$$

The splitting is illustrated in Figure 1. Now, the following constraints have to be satisfied:

$$\begin{cases} A_1(\zeta)\, x_1 + A_2(\zeta)\, x_2^{(1,1)} \leq b, & \forall \zeta \in \mathcal{Z}_{1,1} \\ A_1(\zeta)\, x_1 + A_2(\zeta)\, x_2^{(1,2)} \leq b, & \forall \zeta \in \mathcal{Z}_{1,2}. \end{cases}$$

Since there are two values for the decision at time 2, there are also two 'objective function' values: $c_1^T x_1 + c_2^T x_2^{(1,1)}$ and $c_1^T x_1 + c_2^T x_2^{(1,2)}$. The worst-case value is:

$$z = \max\left\{c_1^T x_1 + c_2^T x_2^{(1,1)}, c_1^T x_1 + c_2^T x_2^{(1,2)}\right\}.$$

After splitting $\mathcal{Z}$ into two subsets, one is solving the following problem:

$$\begin{aligned} \min \quad & z^{(1)} \\ \text{s.t.} \quad & c_1^T x_1 + c_2^T x_2^{(1,s)} \leq z^{(1)}, \quad s = 1,2 \\ & A_1(\zeta)\, x_1 + A_2(\zeta)\, x_2^{(1,s)} \leq b, \quad \forall \zeta \in \mathcal{Z}_{1,s}, \quad s = 1,2. \end{aligned} \tag{4}$$

Since for each s the constraint system is less restrictive than in (1), an improvement in the optimal value can be expected. Also, the average-case performance is expected to be better than in the case of (1), due to the variety of time 2 decision variants. ☐

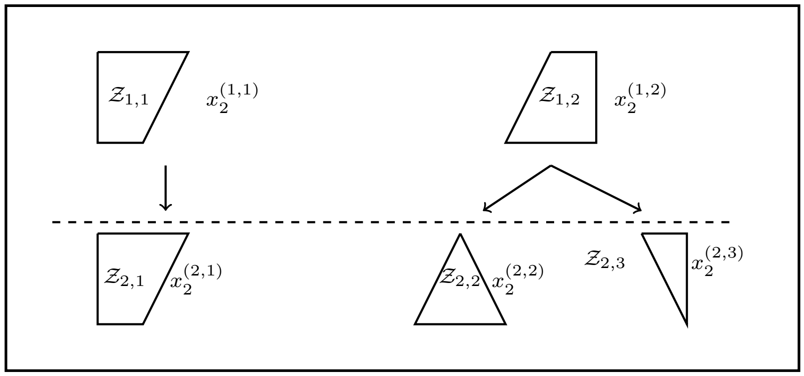

Figure 2: An example of second split for the two-period case.

The splitting process can be continued so that the already existing sets $\mathcal{Z}_{r,s}$ are split with hyperplanes. This is illustrated by the continuation of our example.

Example 1 (continuing from p. 5)**.** Figure 2 illustrates the second splitting round, where the set $\mathcal{Z}_{1,1}$ is left not split, but the set $\mathcal{Z}_{1,2}$ is split with a new hyperplane into two new subsets $\mathcal{Z}_{2,2}$ and $\mathcal{Z}_{2,3}$. Then, a problem results with three uncertainty subsets and three decision variants $x^{(2,s)}$ for time 2. □

In general, after the r-th splitting round there are N_r uncertainty subsets $\mathcal{Z}_{r,s}$ and N_r decision variants $x_2^{(r,s)}$. The problem is then:

$$\begin{aligned} \min \quad & z^{(r)} \\ \text{s.t.} \quad & c_1^T x_1 + c_2^T x_2^{(r,s)} \le z^{(r)}, \quad s \in \mathcal{N}_r \\ & A_1(\zeta)x_1 + A_2(\zeta)x_2^{(r,s)} \le b, \quad \forall \zeta \in \mathcal{Z}_{r,s}, \qquad s \in \mathcal{N}_r = \{1, ..., N_r\}\,. \end{aligned} \tag{5}$$

The finer the splitting of the uncertainty set, the lower optimal value one may expect. In the limiting case, as the maximum diameter of the uncertainty subsets for a given r converges to 0 as $r \to +\infty$, it should hold that the optimal value of (5) converges to $\overline{z}_{\text{adj}}$ - the optimal value of (2). In [8] authors study the question of finding the optimal k time 2 decision variants, and prove under several regularity assumptions that as the number k of variants tends to $+\infty$, the optimal solution to the k-adaptable problem converges to $\overline{z}_{\text{adj}}$.

Determining whether further splitting is needed and finding the proper hyperplanes is crucial for an improvement in the worst-case objective value to occur. The next two subsections provide some theory for determining (1) how far the current optimum is from the best possible value, (2) how to choose the splitting hyperplanes.

2.2 Lower bounds

As the problem becomes larger with subsequent splitting rounds, it is important to know how far the current optimal value is from $\overline{z}_{\text{adj}}$ or its lower bound. We use a lower bounding idea proposed for two-period robust problems in [18], and used also in [11].

Let $\overline{\mathcal{Z}} = \left\{\zeta^{(1)}, \ldots, \zeta^{(|\overline{\mathcal{Z}}|)}\right\} \subset \mathcal{Z}$ be a finite set of scenarios for the uncertain parameter. Consider the problem

$$\begin{array}{lll} \min\limits_{w, x_1, x_2^{(i)}, i=1,\ldots,|\overline{\mathcal{Z}}|} & w & \\ \text{s.t.} & c_1^T x_1 + c_2^T x_2^{(i)} \leq w, \quad i = 1, ..., |\overline{\mathcal{Z}}| & \\ & A_1\left(\zeta^{(i)}\right) x_1 + A_2\left(\zeta^{(i)}\right) x_2^{(i)} \leq b, \quad i = 1, ..., |\overline{\mathcal{Z}}|, & \end{array} \tag{6}$$

where each $x_1 \in \mathbb{R}^{d_1}$ and $x_2^{(i)} \in \mathbb{R}^{d_2}$, for all i. Then, the optimal value of (6) is a lower bound for $\overline{z}_{\text{adj}}$, the optimal value of (2) and hence, to any problem (5).

Since each scenario in $\overline{\mathcal{Z}}$ increases the size of the problem to solve, it is essential to include a possibly small number of scenarios determining the current optimal value of problem (5). The next section indicates a special class of scenarios and based on this, in Section 5 we propose heuristic techniques to construct $\overline{\mathcal{Z}}$.

2.3 How to split

2.3.1 General theorem

To obtain results supporting the decision about splitting the subsets $\mathcal{Z}_{r,s}$, we study the dual of problem (5). We assume that (5) satisfies Slater's condition. By result of Ben-Tal and Beck (2010) the dual of (5) is equivalent to:

$$\begin{array}{ll} \max\limits_{\lambda^{(r)}, \mu^{(r)}, \zeta^{(r)}} & -\sum\limits_{s \in \mathcal{N}_r} \sum\limits_{i=1}^{m} \lambda_{s,i}^{(r)} b_i \\ \text{s.t.} & \sum\limits_{s \in \mathcal{N}_r} \sum\limits_{i=1}^{m} \lambda_{s,i}^{(r)} a_{1,i}\left(\zeta^{(r,s,i)}\right) + \sum\limits_{s \in \mathcal{N}_r} \mu_s^{(r)} c_1 = 0 \\ & \sum\limits_{i=1}^{m} \lambda_{s,i}^{(r)} a_{2,i}\left(\zeta^{(r,s,i)}\right) + \mu_s^{(r)} c_2 = 0, \quad \forall s \in \mathcal{N}_r \\ & \sum\limits_{s \in \mathcal{N}_r} \mu_s^{(r)} = 1 \\ & \lambda^{(r)}, \mu^{(r)} \geq 0 \\ & \zeta^{(r,s,i)} \in \mathcal{Z}_{r,s}, \quad \forall s \in \mathcal{N}_r, \quad \forall 1 \leq i \leq m. \end{array} \tag{7}$$

Because Slater's condition holds, strong duality holds, and for an optimal $\overline{x}^{(r)}$ to problem (5), with objective value $\overline{z}^{(r)}$, there exist $\overline{\lambda}^{(r)}, \overline{\mu}^{(r)}, \overline{\zeta}^{(r)}$, such that the dual optimal value is attained and equal to $\overline{z}^{(r)}$. For each $s \in \mathcal{N}_r$ let us define

$$\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right) = \left\{\overline{\zeta}^{(r,s,i)} \in \mathcal{Z}_{r,s} : \quad \overline{\lambda}_{s,i}^{(r)} > 0\right\},$$

which is a set of worst-case scenarios for ζ determining that the optimal value for (5) cannot be better than $\overline{z}^{(r)}$. The following theorem states that at least one of the sets $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)$ must be split in order for the optimal value $\overline{z}^{(r')}$ of the problem after the subsequent splitting rounds to be better than $\overline{z}^{(r)}$.

Theorem 1. Assume that problem (5) satisfies Slater's condition, $\overline{x}^{(r)}$ is the optimal primal solution, and $\overline{\lambda}^{(r)}$ $\overline{\mu}^{(r)}, \overline{\zeta}^{(r)}$ is the optimal dual solution. Assume that at a splitting round $r' > r$ there exists a sequence of distinct numbers $\{i_1, i_2, ..., i_{N_r}\} \subset \mathcal{N}_{r'}$ such that $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right) \subset \mathcal{Z}_{r',i_s}$ for each $1 \leq s \leq N_r$, that is, each set $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)$ remains not divided, staying a part of some uncertainty subset. Then, it holds that the optimal value $\overline{z}^{(r')}$ after the r'-th splitting round is equal to $\overline{z}^{(r)}$.

Proof. We shall construct a lower bound for the problem after the r'-th round with value $\overline{z}^{(r)}$ by choosing proper $\lambda^{(r')}, \mu^{(r')}, \zeta^{(r')}$. Without loss of generality we assume that $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right) \subset \mathcal{Z}_{r',s}$ for all $s \in N_r$. We take the dual problem of the problem after the r'-th splitting round in the form (7). We assign the following values:

$$\begin{aligned}
\lambda_{s,i}^{(r')} &= \begin{cases} \overline{\lambda}_{s,i}^{(r)} & \text{for } 1 \le s \le N_r \\ 0 & \text{otherwise} \end{cases} \\
\mu_s^{(r')} &= \begin{cases} \overline{\mu}_s^{(r)} & \text{for } 1 \le s \le N_r \\ 0 & \text{otherwise} \end{cases} \\
\zeta^{(r',s,i)} &= \begin{cases} \bar{\zeta}^{(r,s,i)} & \text{if } s \le N_r, \quad \overline{\lambda}_{s,i}^{(r)} > 0 \\ \text{any } \zeta^{(r',s,i)} \in \mathcal{Z}_{r',s} & \text{otherwise.} \end{cases}
\end{aligned}$$

Such variables are dual feasible and give an objective value to the dual equal to $\overline{z}^{(r)}$. Since the dual objective value provides a lower bound on the primal problem after the r'-th round, the theorem follows. □

The above result provides an important insight. If there exists $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)$ with more than one element, then at least one of such sets $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)$ should be divided in the splitting process. On the other hand, if no such $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)$ exists, then splitting should stop since, by Theorem 1, the optimal value cannot improve.

Corollary 1. If for optimal $\overline{\lambda}^{(r)}, \overline{\mu}^{(r)}, \overline{\zeta}^{(r)}$ it holds that:

$$\left|\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)\right| \le 1, \quad \forall s \in \mathcal{N}_r,$$

then $\overline{z}^{(r)} = \overline{z}_{\text{adj}}$, where $\overline{z}_{\text{adj}}$ is the optimal value of (2).

Proof. A lower-bound program with a scenario set $\overline{\mathcal{Z}} = \cup_{s \in \mathcal{N}_r} \overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)$ has an optimal value at most $\overline{z}_{\text{adj}}$. By duality arguments similar to Theorem 1, the optimal value of such a lower bound problem must be equal to $\overline{z}^{(r)}$. This, combined with the fact that $\overline{z}^{(r)} \ge \overline{z}_{\text{adj}}$ gives $\overline{z}^{(r)} = \overline{z}_{\text{adj}}$. □

If there is more than one dual optimal $\overline{\lambda}^{(r)}$, then each of them may imply different sets $\overline{\mathcal{Z}}_{r,s}(\overline{\lambda}^{(r)})$ to be divided. Hence, determining exactly the sets of points to be divided is a difficult task. The next section proposes efficient methods of finding approximate sets of scenarios to be split.

2.3.2 Finding the sets of scenarios to be split

Active constraints. The first method of constructing approximate scenario sets relies on the remark that for a given optimal solution $\overline{x}^{(r)}$ to (5), a $\overline{\lambda}_{s,i}^{(r)} > 0$ corresponds to an active primal constraint. That means, for each $s \in \mathcal{N}_r$ we can define the set:

$$\Phi_{r,s}\left(\overline{x}^{(r)}\right) = \left\{\zeta : \quad \exists_i : a_{1,i}^T(\zeta)\overline{x}_1 + a_{2,i}^T(\zeta)\overline{x}_2^{(r,s)} = b_i\right\}.$$

Though some $\Phi_{r,s}\left(\overline{x}^{(r)}\right)$ may contain infinitely many elements, one can approximate it by finding a single scenario for each constraint, solving the following problem for each s, i:

$$\begin{array}{ll}\min\limits_{\zeta} & b_i - a_{1,i}^T(\zeta)\overline{x}_1 + a_{2,i}^T(\zeta)\overline{x}_2^{(r,s)} \\ & \zeta \in \mathcal{Z}_{r,s}.\end{array} \tag{8}$$

If for given s, i the optimal value of (8) is 0, we add the optimal ζ to the set $\overline{\mathcal{Z}}_{r,s}\left(\overline{x}^{(r)}\right)$. However, such a set could include ζ's for which there exists no $\overline{\lambda}_{s,i}^{(r)} > 0$ being a part of optimal dual solution.

Using the KKT vector of the robust problem. As explained above, it would be beneficial to find a way to obtain the values of $\overline{\lambda}^{(r)}$ to choose only the scenarios $\zeta^{(r,s,i)}$ for which it holds that $\overline{\lambda}_{s,i}^{(r)} > 0$. To do this, one needs to remove the nonconvexity of problem (7) or solve it in some other way. To do this, we shall assume that each $\mathcal{Z}_{r,s}$ is representable by a finite set of convex constraints:

$$\mathcal{Z}_{r,s} = \{\zeta : \quad h_{r,s,j}(\zeta) \leq 0, \quad j = 1, ..., I_{r,s}\}, \quad \forall s \in \mathcal{N}_r, \tag{9}$$

where each $h_{r,s,j}(.)$ is a closed convex function. For an overview of sets representable in this way we refer to [4], mentioning here only that such formulation entails also conic sets. With such a set definition, by results of [15], we can transform (7) to an equivalent convex problem by substituting $\lambda_{s,i}^{(r)}\zeta^{(r,s,i)} = \xi^{(r,s,i)}$. Combining this with the definition of the rows of matrices A_1, A_2, we obtain the following problem, equivalent to (7):

$$\begin{array}{ll}\max\limits_{\substack{\lambda^{(r)},\mu^{(r)}\\ \xi^{(r)}}} & -\sum\limits_{s\in\mathcal{N}_r}\sum\limits_{i=1}^{m}\lambda_{s,i}^{(r)}b_i \\ \text{s.t.} & \sum\limits_{s\in\mathcal{N}_r}\sum\limits_{i=1}^{m}P_{1,i}\xi^{(r,s,i)} + \sum\limits_{s\in\mathcal{N}_r}\mu_s^{(r)}c_1 = 0 \\ & \sum\limits_{i=1}^{m}P_{2,i}\xi^{(r,s,i)} + \mu_s^{(r)}c_2 = 0, \quad \forall s \in \mathcal{N}_r \\ & \sum\limits_{s\in\mathcal{N}_r}\mu_s^{(r)} = 1 \\ & \lambda^{(r)}, \mu^{(r)} \geq 0 \\ & \lambda_{s,i}^{(r)}h_{s,j}\left(\dfrac{\xi^{(r,s,i)}}{\lambda_{s,i}^{(r)}}\right) \leq 0, \quad \forall s \in \mathcal{N}_r, \quad i = 1, \dots, m, \quad j = 1, ..., I_{r,s}.\end{array} \tag{10}$$

Problem (10) is convex in the decision variables. Optimal variables for (10), with substitution

$$\zeta^{(r,s,i)} = \begin{cases} \dfrac{\xi^{(r,s,i)}}{\lambda_{s,i}^{(r)}} & \text{for } \lambda_{s,i}^{(r)} > 0 \\ \zeta^{(r,s,i)} \in \mathcal{Z}_{r,s} & \text{for } \overline{\lambda}_{s,i}^{(r)} = 0, \end{cases}$$

are optimal for (7). Hence, one may construct the sets of points to be split as:

$$\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right) = \left\{\frac{\overline{\xi}^{(r,s,i)}}{\overline{\lambda}_{s,i}^{(r)}} : \quad \overline{\lambda}_{s,i}^{(r)} > 0\right\}.$$

Thus, in order to obtain a set $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)$, one needs the solution to the convex problem (10). It turns out that this solution can be obtained at no extra cost apart from solving (5) if we assume representation (9) and that the tractable robust counterpart of (5) satisfies Slater's condition - one can use then its optimal KKT vector.

Tractable robust counterpart of (5) constructed using methodology of [4] is:

$$
\begin{array}{lll}
\min\limits_{z^{(r)},x_1,x_2^{(r,s)},v^{s,i,j},u_{s,i,j}} & z^{(r)} & \\
\text{s.t.} & c_1^T x_1 + c_2^T x_2^{(r,s)} \leq z^{(r)}, \quad s \in \mathcal{N}_r & \\
& \sum_{j=1}^{I_{r,s}} u_{s,i,j} h_{s,i,j}^* \left(\frac{v^{s,i,j}}{u_{s,i,j}} \right) \leq b_i, \quad \forall s \in \mathcal{N}_r, \forall 1 \leq i \leq m & \\
& \sum_{j=1}^{I_{r,s}} v^{s,i,j} = P_{1,i}^T x_1 + P_{2,i}^T x_2^{(r,s)}, \quad \forall s \in \mathcal{N}_r, \forall 1 \leq i \leq m. &
\end{array}
\tag{11}
$$

Let us denote the Lagrange multipliers of the three subsequent constraint types by $\mu_s^{(r)}, \lambda_{s,i}^{(r)}, \xi^{(r,s,i)}$, respectively. Now we can formulate the theorem stating that the KKT vector of the optimal solution to (11) gives the optimal solution to (10).

Theorem 2. Suppose that problem (11) satisfies Slater's condition. Then, the components of the optimal KKT vector of (11) yield the optimal solution to (10).

Proof. The Lagrangian for problem (11) is:

$$
\begin{aligned}
L(z, x_1, x_2, v, u, \lambda, \mu, \xi) = \; & z^{(r)} + \sum_s \mu_s^{(r)} \left\{ c_1^T x_1 + c_2^T x_2^{(r,s)} - z^{(r)} \right\} + \\
& + \sum_{s,i} \lambda_{s,i}^{(r)} \left(\sum_j u_{s,i,j} h_{s,j}^* \left(\frac{v^{s,i,j}}{u_{s,i,j}} \right) - b_i \right) + \\
& - \sum_{s,i} \left(\xi^{(r,s,i)} \right)^T \left(\sum_j v^{s,i,j} - P_{1,i}^T x_1 - P_{2,i}^T x_2^{(r,s)} \right)
\end{aligned}
$$

We now show that the Lagrange multipliers correspond to the decision variables with the corresponding names in problem (10), by deriving the Lagrange dual problem:

$$
\begin{aligned}
& \max_{\lambda \geq 0, \mu \geq 0, \xi} \min_{\substack{z, x_1, x_2, \\ v^{s,i,j}, u_{s,i,j}}} L(z, x_1, x_2, v, u, \lambda, \mu, \xi) = \\
= & \max_{\lambda \geq 0, \mu \geq 0, \xi} \left\{ \min_{z^{(r)}} \left(1 - \sum_s \mu_s^{(r)} \right) z^{(r)} + \min_{x_1} \left(\sum_s \mu_s^{(r)} c_1 + \sum_{s,i} P_{1,i} \xi^{(r,s,i)} \right)^T x_1 \right. \\
& + \min_{x_2^{(r,s)}} \sum_s \left(\mu_s^{(r)} c_2 + \sum_i P_{2,i} \xi^{(r,s,i)} \right)^T x_2^{(r,s)} \\
& \left. + \sum_{s,i,j} \min_{v^{s,i,j}, u_{s,i,j}} \left\{ \lambda_{s,i}^{(r)} u_{s,i,j} h_{s,j}^* \left(\frac{v^{s,i,j}}{u_{s,i,j}} \right) - \left(\xi^{(r,s,i)} \right)^T v^{s,i,j} \right\} \right\} \\
= & \max_{\lambda \geq 0, \mu \geq 0, \xi} \left\{ - \sum_s \mu_{s,i}^{(r)} b_i \,\middle|\, 1 - \sum_s \mu_s^{(r)} = 0, \quad \sum_s \mu_s^{(r)} c_1 + \sum_{s,i} P_{1,i} \xi^{(r,s,i)} = 0, \right. \\
& \left. \mu_s^{(r)} c_2 + \sum_i P_{2,i} \xi^{(r,s,i)} = 0, \quad \forall s, \lambda_{s,i}^{(r)} h_{s,j} \left(\frac{\xi^{(r,s,i)}}{\lambda_{s,i}^{(r)}} \right) \leq 0 \quad \forall s, i, j \right\}
\end{aligned}
$$

Hence, one arrives at the problem equivalent to (10) and the theorem follows. □

Due to Thorem 2, we know that the optimal solution to (10), and thus to (7), can be obtained at no extra computational effort since most of the solvers produce the KKT vector as a part of output.

Sets $\overline{\mathcal{Z}}_{r,s} \left(\overline{\lambda}^{(r)} \right)$ or $\overline{\mathcal{Z}}_{r,s} \left(\overline{x}^{(r)} \right)$ obtained using the methods above may only be some of many possible sets for a given problem. Hence, there is no guarantee that by splitting $\overline{\mathcal{Z}}_{r,s}$ one splits 'all the ζ scenarios that must be split.' However, these approaches are computationally tractable and may already give a good practical performance, as shown in the numerical examples of Section 6.

3 Multiperiod problems

3.1 Description

In this section we extend the basic two-period methodology to the case with more than two periods, which requires a more extensive notation. The uncertain parameter and the decision vector are:

$$\zeta = \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_{T-1} \end{bmatrix} \in \mathbb{R}^{L_1} \times ... \times \mathbb{R}^{L_{T-1}}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_T \end{bmatrix} \in \mathbb{R}^{d_1} \times .... \times \mathbb{R}^{d_T}.$$

Value of the component ζ_t is revealed at time t. The decision x_t is implemented at time t, after the value of ζ_{t-1} is known but before ζ_t is known. We introduce a special notation for the time-dependent parts of the vectors. The symbol $x_{s:t}$, where $s \leq t$ shall denote the part of the vector x corresponding to periods s through t. We also define $L = \sum\limits_{t=1}^{T-1} L_i$ and $d = \sum\limits_{t=1}^{T} d_t$.

The considered robust multi-period problem is:

$$\begin{array}{ll} \min\limits_{x} & c^T x \\ \text{s.t.} & A(\zeta)x \leq b, \quad \forall \zeta \in \mathcal{Z}, \end{array} \tag{12}$$

where the matrix $A : \mathbb{R}^L \to \mathbb{R}^{m \times d}$ is linear and its i-th row is denoted by a_i^T. In the multi-period case we also split the set $\mathcal{Z}$ into a collection of sets $\mathcal{Z}_{r,s}$ where $\cup_{s \in \mathcal{N}_r} \mathcal{Z}_{r,s} = \mathcal{Z}$ for each r. By $\text{Proj}_t(\mathcal{Z}_{r,s})$ we denote the projection of the set $\mathcal{Z}_{r,s}$ onto the space corresponding to the uncertain parameters from the first t periods:

$$\text{Proj}_t(\mathcal{Z}_{r,s}) = \{\xi : \quad \exists \zeta \in \mathcal{Z}_{r,s}, \quad \xi = \zeta_{1:t}\}.$$

Contrary to the two-period case, every subset $\mathcal{Z}_{r,s}$ shall correspond to a vector $x^{(r,s)} \in \mathbb{R}^d$, i.e. a vector including decisions for all the periods.

In the two-period case, the time 1 decision was common for all the variants of decision variables. In the multi-period notation this condition would be written as $x_1^{(r,s)} = x_1^{(r,s+1)}$ for $1 \leq s \leq N_r - 1$. In the two-period case each of the uncertainty subsets $\mathcal{Z}_{r,s}$ corresponded to a separate variant $x_2^{(r,s)}$, and given a ζ, any of them could be chosen if only it held at time 2 that $\zeta \in \mathcal{Z}_{r,s}$. In this way, it was guaranteed that

$$\forall \zeta \in \mathcal{Z} \quad \exists x_2^{(r,s)} : \quad A_1(\zeta)x_1 + A_2(\zeta)x_2^{(r,s)} \leq b.$$

In the multi-period case the main obstacle is the fact that the information about subsequent components of ζ is revealed period after period, whereas at the same time decisions need to be implemented. In general up to time T one may not know to which $\mathcal{Z}_{r,s}$ the vector ζ will surely belong to.

For instance, suppose that at time 1 the decision $\overline{x}_1$ is implemented. At time 2, knowing only the value ζ_1 there may be many potential sets $\mathcal{Z}_{r,s}$ to which ζ may belong and for which $\overline{x}_1 = x_2^{(r,s)}$ - all the $\mathcal{Z}_{r,s}$ for which $\zeta_1 \in \mathrm{Proj}_1(\mathcal{Z}_{r,s})$. Suppose that a decision $\overline{x}_2 = x_2^{(r,s)}$ is chosen at time 2, for some s. Then, at time 3 there must exist a set $\mathcal{Z}_{r,s}$ such that $\zeta_{1:2} \in \mathrm{Proj}_2(\mathcal{Z}_{r,s})$ and for which $\overline{x}_{1:2} = x_{1:2}^{(r,s)}$, so that its decision for time 3 can be implemented.

In general, at each time period $2 < t \leq T$ there must exist a set $\mathcal{Z}_{r,s}$ such that the vector $\zeta_{1:t-1} \in \mathrm{Proj}_{t-1}(\mathcal{Z}_{r,s})$, and for which it holds that $\overline{x}_{1:t-1} = x_{1:t-1}^{(r,s)}$, where $\overline{x}_{1:t-1}$ stands for the decisions already implemented. This is our version of a requirement that in Ben-Tal et al. (2009) is called 'non-anticipativity restriction.' We propose an iterative splitting strategy ensuring this postulate is satisfied.

In this strategy, the early-period decisions corresponding to various sets $\mathcal{Z}_{r,s}$ are identical, as long as it is not possible to distinguish to which of them the vector ζ will belong. Our strategy facilitates simple determination of these equality constraints between various decisions and is based on the following notion.

Definition 1. A hyperplane defined by a normal vector $g \in \mathbb{R}^L$ and intercept term $h \in \mathbb{R}$ is a time t splitting hyperplane (called later t-SH) if:

$$t = \min\left\{u : g^T\zeta = h \quad \Leftrightarrow \quad g_{1:u}^T\zeta_{1:u} = h, \quad \forall \zeta \in \mathbb{R}^L\right\}.$$

In other words, t is the smallest number such that for any $\zeta \in \mathbb{R}^L$ it is sufficient to know the part $\zeta_{1:t}$ to determine if it holds that $g^T\zeta \leq h$ or $g^T\zeta \geq h$. We shall refer to a hyperplane by the pair (g, h). $\square$

We illustrate with an example how the first splitting can be done and how the corresponding equality structure between decision vectors $x^{(r,s)}$ is determined.

Example 2. We split the uncertainty set $\mathcal{Z}$ with a 1-SH (g, h). Then, two subsets result:

$$\mathcal{Z}_{1,1} = \mathcal{Z} \cap \left\{\zeta : g^T\zeta \leq h\right\} \quad \text{and} \quad \mathcal{Z}_{1,2} = \mathcal{Z} \cap \left\{\zeta : g^T\zeta \geq h\right\}.$$

Now, there are two decision vectors $x^{(1,1)}, x^{(1,2)} \in \mathbb{R}^d$. Their time 1 decisions should be identical since they are implemented before the value of ζ_1 is known, allowing to determine whether $\zeta \in \mathcal{Z}_{1,1}$ or $\zeta \in \mathcal{Z}_{1,2}$. Thus, we add a constraint $x_1^{(1,1)} = x_1^{(1,2)}$. This splitting is illustrated in Figure 3.

The problem to be solved after the first splitting round is analogous to the two-period case, with the equality constraint added:

$$\begin{array}{rl}
\min\limits_{z^{(1)}, x^{(1,s)}} & z^{(1)} \\
\text{s.t.} & c^T x^{(1,i)} \leq z^{(1)}, \quad i = 1, 2 \\
& A(\zeta)\, x^{(i)} \leq b, \quad \forall \zeta \in \mathcal{Z}_{1,s}, \quad s = 1, 2 \\
& x_1^{(1,1)} = x_1^{(1,2)}.
\end{array}$$

$\square$

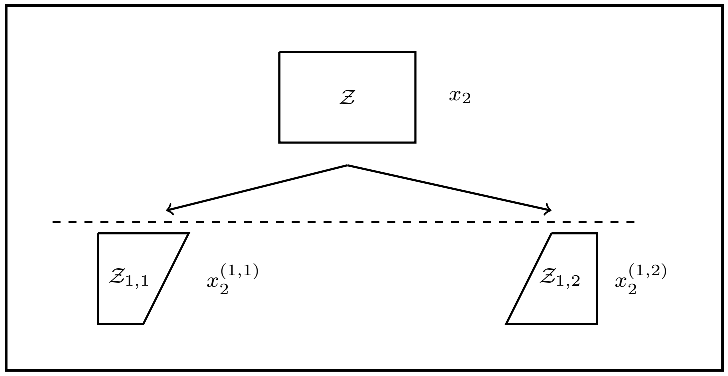

Figure 3: A multi-period problem after a single splitting with a time-2 splitting hyperplane.

The splitting process may be continued and multiple types of t-SHs are possible. To state our methodology formally, we define a parameter $t_{max}(\mathcal{Z}_{r,s})$ for each set $\mathcal{Z}_{r,s}$. If the set $\mathcal{Z}_{r,s}$ is a result of subsequent splits with various t-SH's, the number $t_{max}(\mathcal{Z}_{r,s})$ denotes the largest t of them. By convention, for the set $\mathcal{Z}$ it shall hold that $t_{max}(\mathcal{Z}) = 0$. The following rule defines how the subsequent sets can be split and what the values of the parameter t_{max} for each of the resulting sets are.

Rule 1. A set $\mathcal{Z}_{r,s}$ can be split only with a t-SH such that $t \geq t_{max}(\mathcal{Z}_{r,s})$. For the resulting two sets $\mathcal{Z}_{r+1,s'}, \mathcal{Z}_{r+1,s''}$ we define $t_{max}(\mathcal{Z}_{r+1,s'}) = t_{max}(\mathcal{Z}_{r+1,s''}) = t$. If the set is not split and in the next round it becomes the set $\mathcal{Z}_{r+1,s'}$ then $t_{max}(\mathcal{Z}_{r+1,s'}) = t_{max}(\mathcal{Z}_{r,s})$. □

The next rule defines the equality constraints for the problem after the $(r+1)$-th splitting round, based on the problem after the r-th splitting round.

Rule 2. Let a set $\mathcal{Z}_{r,s}$ be split with a t-SH into sets $\mathcal{Z}_{r+1,s'}, \mathcal{Z}_{r+1,s''}$. Then the constraint $x_{1:t}^{(r+1,s')} = x_{1:t}^{(r+1,s'')}$ is added to the problem after the $(r+1)$-th splitting round.

Assume the problem after splitting round r includes sets $\mathcal{Z}^{r,s}$ and $\mathcal{Z}^{r,u}$ with a constraint $x_{1:k_s}^{(r,s)} = x_{1:k_s}^{(r,u)}$, and the sets $\mathcal{Z}^{r,s}, \mathcal{Z}^{r,u}$ are split into $\mathcal{Z}^{r+1,s'}, \mathcal{Z}^{r+1,s''}$ and $\mathcal{Z}^{r+1,u'}, \mathcal{Z}^{r+1,u''}$, respectively. Then, the constraint $x_{1:k_s}^{(r+1,s')} = x_{1:k_s}^{(r+1,u')}$ is added to the problem after the $(r+1)$-th splitting round. □

The first part of Rule 2 ensures that the decision vectors $x^{(r+1,s')}, x^{(r+1,s'')}$ can differ only from time period $t+1$ on, since only then one can distinguish between the sets $\mathcal{Z}_{r,s'}, \mathcal{Z}_{r,s''}$. The second part of Rule 2 ensures that the dependence structure between decision vectors from stage r is not 'lost' after the splitting. Rule 2 as a whole ensures that $x_{1:k_s}^{(r+1,s')} = x_{1:k_s}^{(r+1,s'')} = x_{1:k_s}^{(r+1,u')} = x_{1:k_s}^{(r+1,u'')}$. We illustrate the application of Rules 1 and 2 with a continuation of our example.

Example 2 (continuing from p. 12)**.** By Rule 1 we have $t_{max}(\mathcal{Z}_{1,1}) = t_{max}(\mathcal{Z}_{1,2}) = 1$. Thus, each of the sets $\mathcal{Z}_{1,1}, \mathcal{Z}_{1,2}$ can be split with a t-SH where $t \geq 1$. We split the set $\mathcal{Z}_{1,1}$ with a 1-SH and the set $\mathcal{Z}_{1,2}$ with a 2-SH. The scheme of the second splitting round is given in Figure 4.

We obtain 4 uncertainty sets $\mathcal{Z}_{2,s}$ and 4 decision vectors $x^{(2,s)}$. The lower part of

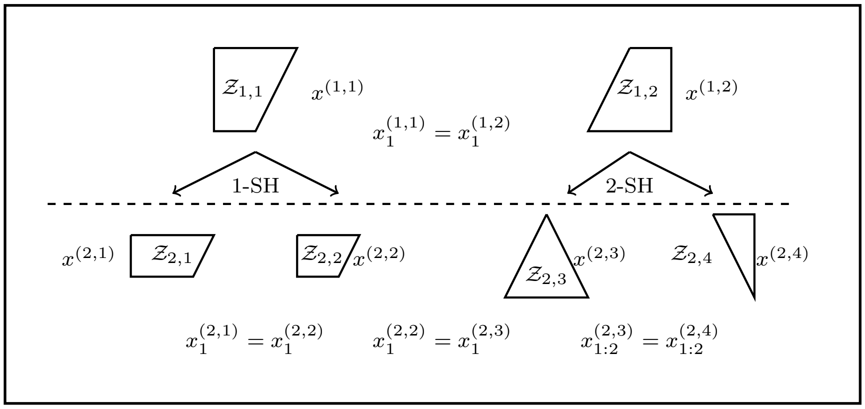

Figure 4: Example of second splitting round for the multi-period case.

Figure 4 includes three equality constraints. The first constraint $x_1^{(2,1)} = x_1^{(2,2)}$ and the third constraint $x_{1:2}^{(2,3)} = x_{1:2}^{(2,4)}$ follow from the first part of Rule 2, whereas the second equality constraint $x_1^{(2,2)} = x_1^{(2,3)}$ is determined by the second part of Rule 2. The equality constraints imply that $x_1^{(2,1)} = x_1^{(2,2)} = x_1^{(2,3)} = x_1^{(2,4)}$.

The problem after the second splitting round is:

$$
\begin{aligned}
\min_{z^{(2)},x^{(2,s)}} \quad & z^{(2)} \\
\text{s.t.} \quad & c^T x^{(2,s)} \le z^{(2)}, \quad s = 1, ..., 4 \\
& A(\zeta)\, x^{(2,s)} \le b, \quad \forall \zeta \in \mathcal{Z}_{2,s}, \quad s = 1, ..., 4 \\
& x_1^{(2,1)} = x_1^{(2,2)} \\
& x_1^{(2,2)} = x_1^{(2,3)} \\
& x_{1:2}^{(2,3)} = x_{1:2}^{(2,4)}.
\end{aligned}
$$

The time structure of decisions for subsequent time periods is illustrated in Figure 5. All the decision variables within a single cell are equal, for instance, all the decisions $x_1^{(2,s)}$. The decision making process goes as follows.

At time 1 there is only one possibility for the first decision. Then, at time 2 the value of the parameter ζ_1 is known and one can determine if ζ is within the set $\mathcal{Z}_{1,1}$ or $\mathcal{Z}_{1,2}$, or both.

If $\zeta \in \mathcal{Z}_{1,1}$, further verification is needed to determine whether $\zeta \in \mathcal{Z}_{2,1}$ or $\zeta \in \mathcal{Z}_{2,2}$, to choose the correct variant of decisions for time 2 and later.

If $\zeta \in \mathcal{Z}_{1,2}$, the time 2 decision $x_2^{(2,3)} = x_2^{(2,4)}$ is implemented. Later, the value of ζ_2 is revealed and based on it, one determines if $\zeta \in \mathcal{Z}_{2,3}$ or $\zeta \in \mathcal{Z}_{2,4}$. In the first case from time 3 on the decisions $x_3^{(2,3)}, x_4^{(2,3)}, ..., x_T^{(2,3)}$ are implemented. If otherwise, the decisions $x_3^{(2,4)}, x_4^{(2,4)}, ..., x_T^{(2,4)}$ are implemented.

If $\zeta \in \mathcal{Z}_{1,1} \cap \mathcal{Z}_{1,2}$, then a further verification is needed to determine whether $\zeta \in \mathcal{Z}_{1,2} \cap \mathcal{Z}_{2,1}$ or $\zeta \in \mathcal{Z}_{1,2} \cap \mathcal{Z}_{2,2}$ (or both).

For example, if $\zeta \in \mathcal{Z}_{1,2} \cap \mathcal{Z}_{2,1}$, then at time 2 one can implement either $x_2^{(2,1)}$ or $x_2^{(2,3)} = x_2^{(2,4)}$. It is best to choose the decision for which the entire decision vector $x^{(r,s)}$ gives the best worst-case objective. If one chooses $x_2^{(2,3)} = x_2^{(2,4)}$, then after time 2 it is known if $\zeta \in \mathcal{Z}_{2,3}$ or $\zeta \in \mathcal{Z}_{2,4}$, and the sequence of decisions for later pe-

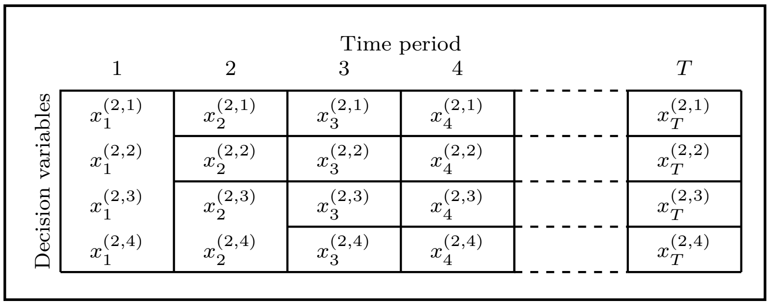

Figure 5: Time structure of the decision variants after the second splitting. Decision vectors within the same cell are equal.

riods is chosen. If one has chosen $x_2^{(2,1)}$ then the sequence of decisions $x_3^{(2,1)}, ..., x_T^{(2,1)}$ is implemented later. Analogous procedure holds for other possibilities. $\square$

In general, the problem after the r-th splitting round has N_r subsets $\mathcal{Z}_{r,s}$ and decision vectors $x^{(r,s)}$. Its formulation is:

$$\begin{aligned} \min_{z^{(r)}, x^{(r,s)}} \quad & z^{(r)} \\ & c^T x^{(r,s)} \leq z^{(r)}, \quad s \in \mathcal{N}_r \\ & A(\zeta)\, x^{(r,s)} \leq b, \quad \forall \zeta \in \mathcal{Z}_{r,s}, \quad s \in \mathcal{N}_r \\ & x_{1:k_s}^{(r,s)} = x_{1:k_s}^{(r,s+1)}, \quad s \in \mathcal{N}_r \setminus \{N_r\}, \end{aligned} \tag{13}$$

where k_s is the number of the first time period decisions that are required to be identical as a result of applying Rule 2.

3.2 Lower bounds

Similar to the two-period case, one can obtain lower bounds for the adjustable robust solution. The lower bound problem differs from the two-period case since the uncertain parameter may have a multi-period equality structure of the components that can be exploited.

Let $\overline{\mathcal{Z}} = \left\{\zeta^{(1)}, \ldots, \zeta^{(|\overline{\mathcal{Z}}|)}\right\} \subset \mathcal{Z}$ be a finite set of scenarios for the uncertain parameter. Then, the optimal solution to

$$\begin{aligned} \min_{w, x_2^{(i)}, i=1,\ldots,|\overline{\mathcal{Z}}|} \quad & w \\ \text{s.t.} \quad & c^T x_2^{(i)} \leq w, \quad i = 1, ..., \left|\overline{\mathcal{Z}}\right| \\ & A\left(\zeta^{(i)}\right) x^{(i)} \leq b, \quad i = 1, ..., \left|\overline{\mathcal{Z}}\right| \\ & x_{1:t}^{(i)} = x_{1:t}^{(j)} \quad \forall_{i,j,t} : \zeta_{1:t}^{(i)} = \zeta_{1:t}^{(j)} \end{aligned} \tag{14}$$

is a lower bound for problem (13).

In the multi-period case it is required that for each decision vectors $x^{(i)}, x^{(j)}$ whose corresponding uncertain scenarios are identical up to time t the corresponding decisions must be the same up to time t as well. This is needed since up to time t one cannot distinguish between $\zeta^{(i)}$ and $\zeta^{(j)}$ and the decisions made should be the same. The equality structure between the decision vectors $x^{(i)}$ can be obtained efficiently (using at most $\left|\overline{\mathcal{Z}}\right| - 1$ vector equalities) if uncertain parameter is one-dimensional in each time period - one achieves it by sorting the set $\overline{\mathcal{Z}}$ lexicographically.

3.3 How to split

3.3.1 General theorem

We assume that (13) satisfies Slater's condition. By the result of Ben-Tal and Beck (2010) the dual of (13) is equivalent to:

$$
\begin{array}{ll}
\max & -\sum\limits_{s\in\mathcal{N}_r}\sum\limits_{i=1}^{m}\lambda_{s,i}^{(r)}b_i \\
\text{s.t.} & \sum\limits_{i=1}^{m}\lambda_{s,i}^{(r)}a_i\left(\zeta^{(r,s,i)}\right)+\mu_s^{(r)}c+\begin{bmatrix}\nu_s^{(r)}\\0\end{bmatrix}-\begin{bmatrix}\nu_{s-1}^{(r)}\\0\end{bmatrix}=0, \quad \forall 1<s<N_r \\
& \sum\limits_{i=1}^{m}\lambda_{1,i}^{(r)}a_i\left(\zeta^{(r,1,i)}\right)+\mu_1^{(r)}c+\begin{bmatrix}\nu_1^{(r)}\\0\end{bmatrix}=0 \\
& \sum\limits_{i=1}^{m}\lambda_{N_r,i}^{(r)}a_i\left(\zeta^{(r,N_r,i)}\right)+\mu_{N_r}^{(r)}c-\begin{bmatrix}\nu_{r,N_r-1}^{(r)}\\0\end{bmatrix}=0 \\
& \sum\limits_{s\in\mathcal{N}_r}\mu_s^{(r)}=1 \\
& \lambda^{(r)},\mu^{(r)}\geq 0 \\
& \zeta^{(r,s,i)}\in\mathcal{Z}_{r,s}, \quad \forall s\in\mathcal{N}_r, \quad \forall 1\leq i\leq m.
\end{array}
\tag{15}
$$

Because of Slater's condition, strong duality holds and for an optimal solution $\overline{x}(r)$ with objective value $\overline{z}^{(r)}$ there exist $\overline{\lambda}^{(r)},\overline{\mu}^{(r)},\overline{\nu}^{(r)},\overline{\zeta}^{(r)}$ such that the optimal value in (7) is attained and equal to $\overline{z}^{(r)}$. For each subset $\mathcal{Z}_{r,s}$ we define:

$$
\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)=\left\{\overline{\zeta}^{(r,s,i)}\in\mathcal{Z}_{r,s}: \quad \overline{\lambda}_{s,i}^{(r)}>0\right\}.
$$

Then, the following result holds, stating that at least one of the sets $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)$, for which $\left|\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)\right|>1$, should be split.

Theorem 3. Assume that problem (13) satisfies Slater's condition, $\overline{x}^{(r)}$ is the the optimal primal solution, and that $\overline{\lambda}^{(r)},\overline{\mu}^{(r)},\overline{\nu}^{(r)},\overline{\zeta}^{(r)}$ are the optimal dual variables. Assume that at any splitting round $r'>r$ there exists a sequence of distinct numbers $\{i_1,i_2,...,i_{N_r}\}\subset\mathcal{N}_{r'}$ such that $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)\subset\mathcal{Z}_{r',i_s}$ and for each $1\leq s\leq N_r$ it holds that $\mathcal{Z}_{r',i_s}$ results from splitting the set $\mathcal{Z}_{r,s}$. Then, the optimal value $\overline{z}^{(r')}$ is the same as $\overline{z}^{(r)}$, that is, $\overline{z}^{(r')}=\overline{z}^{(r)}$.

Proof. We construct a lower bound for the problem after the r'-th round with value $\overline{z}^{(r)}$. Without loss of generality we assume that $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)\subset\mathcal{Z}_{r',s}$ for all $1\leq s\leq N_r$. By Rule 2, problem after the r'-th splitting round implies equality constraints $x_{1:k_s}^{(r',s)}=x_{1:k_s}^{(r',s+1)}$, where $1\leq s\leq N_r-1$. Take the dual (15) of the problem after the r'-th splitting round. We assign the following values for $\lambda^{(r')},\mu^{(r')}$:

$$
\begin{aligned}
\lambda_{s,i}^{(r')} &= \begin{cases}\overline{\lambda}_{s,i}^{(r)} & \text{for } 1\leq s\leq N_r\\ 0 & \text{otherwise}\end{cases}\\
\mu_s^{(r')} &= \begin{cases}\overline{\mu}_s^{(r)} & \text{for } 1\leq s\leq N_r\\ 0 & \text{otherwise}\end{cases}\\
\nu_s^{(r')} &= \begin{cases}\overline{\nu}_s^{(r)} & \text{for } 1\leq s\leq N_r-1\\ 0 & \text{otherwise}\end{cases}\\
\zeta^{(r',s,i)} &= \begin{cases}\bar{\zeta}^{(r,s,i)} & \text{if } 1\leq s\leq N_r, \quad \overline{\lambda}_{s,i}^{(r)}>0\\ \text{any } \zeta^{(r',s,i)}\in\mathcal{Z}_{r',s,i} & \text{otherwise.}\end{cases}
\end{aligned}
$$

These values are dual feasible and give an objective value to the dual problem equal to $\overline{z}^{(r)}$. Since the dual objective value provides a lower bound for the primal problem, the objective function value for the problem after the r'-th round cannot be better than $\overline{z}^{(r)}$. □

Similar to the two-period case, one can prove that if each of the sets $\overline{\mathcal{Z}}_{r,s}$ has at most one element, then the splitting process may stop since the optimal objective value cannot be better than $\overline{z}^{(r)}$.

3.3.2 Finding the sets of scenarios to be split

For the multi-period case, the same observations hold that have been made in the case of the two-period problem. That is, one may construct sets $\overline{\mathcal{Z}}_{r,s}\left(\overline{x}^{(r)}\right)$ by searching for the scenarios ζ corresponding to active primal constraints, or sets $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)$ by using the optimal KKT variables of the tractable counterpart of (13). The latter approach is preferred for its inclusion only of the 'critical scenarios' in the meaning of Theorem 3.

4 Problems with integer variables

4.1 Methodology

A particularly difficult application field for adjustable robust decision rules is when some of the decision variables are integer. Our methodology can be particularly useful since the decisions are fixed numbers for each of the uncertainty subset $\mathcal{Z}_{r,s}$. A general multiperiod robust adjustable problem with integer and continuous variables can be solved through splitting in the same fashion as in Section 2 and 3.

Suppose, taking the notation of Section 3, that the indices of components of the vector x to be integer belong to a set $\mathcal{I}$. Then, the mixed-integer version of problem (13) has only an additional integer condition:

$$\begin{array}{ll}\min\limits_{z^{(r)},x^{(r,s)}} & z^{(r)} \\ & c^T x^{(r,s)} \leq z^{(r)}, \quad s \in \mathcal{N}_r \\ & A\left(\zeta\right) x^{(r,s)} \leq b, \quad \forall \zeta \in \mathcal{Z}_{r,s}, \quad s \in \mathcal{N}_r \\ & x^{(r,s)}_{1:k_s} = x^{(r,s+1)}_{1:k_s}, \quad i \in \mathcal{N}_r \setminus \{N_r\} \\ & x^{(r,s)}_i \in \mathbb{Z}, \quad \forall s \in \mathcal{N}_r, \forall i \in \mathcal{I}. \end{array} \tag{16}$$

To obtain lower bounds, we propose the analogues of the strategies given in Sections 2.2 and 3.2, with the integer condition.

4.2 Finding the sets of scenarios to be split

For mixed integer optimization the available duality tools are substantially weaker than for problems with continuous variables. One can utilize the subadditive duality theorems to derive results 'similar' to the ones from Section 2.3 and 3.3, but they are not applicable in practice. Two approaches that seem intuitively correct are: (1) separating scenarios responsible for constraints that are 'almost active' for the optimal solution $\overline{x}^{(r)}$, (2) separating scenarios found on the basis of the LP relaxation of problem (16). We now discuss these two approaches.

Almost active constraints. In the continuous case, the sets $\overline{\mathcal{Z}}_{r,s}\left(\overline{x}^{(r)}\right)$ were found by identifying ζ's generating active constraints for the optimal primal solution. One can also apply this approach in the mixed-integer case, with a correction due to the fact that in mixed-integer problems the notion of 'active constraints' loses its proper meaning - in general case the worst-case value of a left-hand side is not a continuous function of the decision variable x. For that reason, it may happen that:

$$\sup_{\zeta \in \mathcal{Z}_{r,s}} a_i(\zeta)^T x^{(r,s)} < b_i,$$

even for constraints that are critical - being elements of a set of constraints prohibiting the optimal objective value of (16) from being better than $\overline{z}^{(r)}$. However, for each $s \in \mathcal{N}_r$ one can define an approximate set $\overline{\mathcal{Z}}_{r,s}\left(\overline{x}^{(r)}, \epsilon\right)$ of ζ's corresponding to 'almost active' constraints. To find such ζ's, for a precision level $\epsilon > 0$ and $s \in \mathcal{N}_r, 1 \leq i \leq m$ one solves the following problem:

$$\begin{aligned} \min_{\zeta} \quad & b_i - a_i\left(\zeta\right)^T \overline{x}^{(r,s)} - \epsilon \\ \text{s.t.} \quad & \zeta \in \mathcal{Z}_{r,s}. \end{aligned} \tag{17}$$

If the result is a nonpositive optimal value, then one can add the optimal solution ζ to the set $\overline{\mathcal{Z}}_{r,s}\left(\overline{x}^{(r)}, \epsilon\right)$. However, this strategy may be subject to scaling problems since ϵ may imply a different degree of 'almost activeness' for different constraints. One may try to mitigate this issue by normalizing the coefficients of each constraint before solving problem (17).

KKT vector of the LP relaxation. Another approach for problems with integer variables, less sensitive to scaling issues, is to determine the sets $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}(r)\right)$ corresponding to the LP relaxation of problem (16). This approach is expected to perform well in problems where the optimal mixed integer solution is close to the optimal solution of the LP relaxation.

5 Heuristics

In this section we propose heuristics for choosing the hyperplanes to split sets $\mathcal{Z}_{r,s}$ (by splitting their corresponding sets $\overline{\mathcal{Z}}_{r,s}$) in the $(r+1)$-th splitting round, for constructing the lower bound scenario sets $\overline{\mathcal{Z}}$, and for deciding when to stop the splitting algorithm.

From now on we fix the optimal primal solution after the r-th splitting round $\overline{x}^{(r)}$ and the sets $\overline{\mathcal{Z}}_{r,s}$, making no distinction between the sets $\overline{\mathcal{Z}}_{r,s}\left(\overline{x}^{(r)}\right)$ obtained by using the optimal KKT vector of the problems' LP relaxations and the sets $\overline{\mathcal{Z}}_{r,s}\left(\overline{\lambda}^{(r)}\right)$

obtained by searching constraint-wise for scenarios that make the constraints (almost) active. We only consider splitting of sets $\mathcal{Z}_{r,s}$ for which $\left|\overline{\mathcal{Z}}_{r,s}\right| > 1$.

5.1 Choosing the t for the t-SHs

In multi-period problems one must determine the t for the t-SH, and this choice should balance two factors. Intuitively, the set $\mathcal{Z}_{r,s}$ should be split with a $t \geq t_{max}(\mathcal{Z}_{r,s})$ for which the components ζ_t are most dispersed over $\zeta \in \overline{\mathcal{Z}}_{r,s}$. On the other hand, choosing a high value of t in an early splitting round reduces the range of possible t-SHs in later rounds because of Rule 1.

We propose that each $\mathcal{Z}_{r,s}$ is split with a t-SH for which the components ζ_t show biggest dispersion within the set $\overline{\mathcal{Z}}_{r,s}$ (measured, for example, with variance) and where $t_{max}(\mathcal{Z}_{r,s}) \leq t \leq t_{max}(\mathcal{Z}_{r,s}) + q$, with q being a predetermined number. If the dispersion equals 0 for all $t_{max}(\mathcal{Z}_{r,s}) \leq t \leq t_{max}(\mathcal{Z}_{r,s}) + q$ then we propose to choose the smallest $t \geq t_{max}(\mathcal{Z}_{r,s})$ such that the components ζ_t show a nonzero dispersion within $\overline{\mathcal{Z}}_{r,s}$.

5.2 Splitting hyperplane heuristics

In this subsection we provide propositions for constructing the splitting hyperplanes.

Heuristic 1. The idea of this heuristic is to determine the two most distant scenarios in $\overline{\mathcal{Z}}_{r,s}$ and to choose a hyperplane that separates them strongly.

Find the $\zeta^{(a)}, \zeta^{(b)} \in \overline{\mathcal{Z}}_{r,s}$ maximizing $\left\|\zeta^{(i)}_{1:t} - \zeta^{(j)}_{1:t}\right\|_2$ over $\zeta^{(i)}, \zeta^{(j)} \in \overline{\mathcal{Z}}_{r,s}$. Then, split the set $\mathcal{Z}_{r,s}$ with a t-SH defined by:

$$g_j = \begin{cases} \zeta^{(a)}_j - \zeta^{(b)}_j & \text{if } j \leq t \\ 0 & \text{otherwise} \end{cases}, \quad h = \frac{g^T\left(\zeta^{(a)} + \zeta^{(b)}\right)}{2}.$$

If $\overline{\mathcal{Z}}_{r,s}$ is a result of search for the (almost) active constraints, $\zeta \in \overline{\mathcal{Z}}_{r,s}$ may be an element of the optimal facet for the search problem. Separation of entire facets may then give better results than of single ζ's. Then, the heuristic would separate the two most distant facets with, for example, their bisector hyperplane. □

Heuristic 2. The idea of this heuristic is to divide the set $\overline{\mathcal{Z}}_{r,s}$ into two sets whose cardinalities differ by as little as possible.

Choose an arbitrary normal vector g for the t-SH. Then, determine the intercept term h such that the term $\left|\left|\overline{\mathcal{Z}}^-_{r,s}\right| - \left|\overline{\mathcal{Z}}^+_{r,s}\right|\right|$ is minimized, with

$$\overline{\mathcal{Z}}^-_{r,s} = \overline{\mathcal{Z}}_{r,s} \cap \left\{\zeta : g^T\zeta \leq h\right\}, \quad \overline{\mathcal{Z}}^+_{r,s} = \overline{\mathcal{Z}}_{r,s} \cap \left\{\zeta : g^T\zeta \geq h\right\}.$$

The best h can be found using binary search. □

Heuristic 3. The idea of this heuristic is to split the set $\mathcal{Z}_{r,s}$ with a hyperplane, and to manipulate the late period decisions while keeping the early-period decisions fixed, in such a way that the maximum worst-case 'objective function' for the two sets is minimized. We describe it for the multi-period case.

Choose an arbitrary normal vector g for the t-SH. For a given intercept h define the two sets:

$$\mathcal{Z}_{r+1,s}^{h-} = \mathcal{Z}_{r,s} \cap \left\{ \zeta : g^T \zeta \leq h \right\}, \quad \mathcal{Z}_{r+1,s}^{h+} = \mathcal{Z}_{r,s} \cap \left\{ \zeta : g^T \zeta \geq h \right\}.$$

For a fixed g and h we define the following function:

$$\begin{aligned} \tau(h) = \min_{x^{(r,s')}, x^{(r,s'')}, w} \quad & w \\ \text{s.t.} \quad & c_1^T x_1 + c_2^T x_2^{(r,s')} \leq w \\ & c_1^T x_1 + c_2^T x_2^{(r,s'')} \leq w \\ & A_1(\zeta) x_1 + A_2(\zeta) x_2^{(r,s')} \leq b, \quad \forall \zeta \in \mathcal{Z}_{r+1,s}^{h-} \\ & A_1(\zeta) x_1 + A_2(\zeta) x_2^{(r,s'')} \leq b, \quad \forall \zeta \in \mathcal{Z}_{r+1,s}^{h+} \\ & x_{1:t_{max}(\mathcal{Z}_{r,s})}^{(r,s')} = x_{1:t_{max}(\mathcal{Z}_{r,s})}^{(r,s'')} = \overline{x}_{1:t_{max}(\mathcal{Z}_{r,s})}^{(r,s)}. \end{aligned} \tag{18}$$

Equality constraints ensure that the decision variables related by equality constraints to other decision vectors stay with the same values (not to lose the feasibility of the decision vectors for sets $\mathcal{Z}_{r,p}$, where $p \neq s$). The aim is to minimize $\tau(h)$ over the domain of h for which both $\mathcal{Z}_{r+1,s}^{h-}$ and $\mathcal{Z}_{r+1,s}^{h+}$ are nonempty. Function $\tau(h)$ is quasiconvex in h, which has been noted in a different setting in [8]. $\square$

5.3 Constructing the lower bound scenario sets

The key premise is that the size of the set $\overline{\mathcal{Z}}^{(r)}$ (the lower bound scenario set after the r-th splitting round) should be kept limited since each additional scenario increases the size of the lower bound problem. Hence, it is important that the limited number of scenarios covers set $\mathcal{Z}$ well.

Summing the scenario sets. One approach is to use $\overline{\mathcal{Z}}^{(r)} = \cup_{s \in \mathcal{N}_r} \overline{\mathcal{Z}}_{r,s}$ after each splitting round, since the sets $\overline{\mathcal{Z}}_{r,s}$ approximate the set of the scenarios not allowing the objective value to improve. To reduce the size of $\overline{\mathcal{Z}}^{(r)}$, we propose that $\overline{\mathcal{Z}}$ contains at most k elements of each $\overline{\mathcal{Z}}_{r,s}$, where k is a predetermined number. This approach implies that the lower bound sequence $\left\{ \overline{w}^{(r)} \right\}$, where $\overline{w}^{(r)}$ is the optimal value of the lower bound problem after the r-th splitting round, needs not be nondecreasing in r.

Incremental building of a scenario set. To ensure a nondecreasing lower bound sequence, one can construct the sets incrementally, starting with $\overline{\mathcal{Z}}^{(1)}$ after the first splitting round and enlarging it with new scenarios after each splitting round. We describe a possible variant of this idea for the multi-period case. Assume that problem (14) has been solved after the r-th splitting round, the lower-bounding scenario set is $\overline{\mathcal{Z}}^{(r)}$ and the optimal value of the lower-bounding problem is $\overline{w}^{(r)}$. Suppose that after the $(r+1)$-th splitting round one wants to add $\zeta' \in \overline{\mathcal{Z}}_{r+1,s}$ to the set $\overline{\mathcal{Z}}^{(r+1)}$. We add the scenario ζ' if (1) there is no $1 \leq i \leq |\overline{\mathcal{Z}}^{(r)}|$ such that $A(\zeta')\left(\overline{x}^{(i)}\right) \leq b$ and (2) there is no $x^{(\zeta')}$ such that

$$\left\{x_{1:t}^{(\zeta')} = x_{1:t}^{(i)} \quad \forall 1 \leq i \leq |\overline{\mathcal{Z}}^{(r)}|, \quad \forall t : \zeta'_{1:t} = \zeta_{1:t}^{(i)}\right\}, \qquad c^T x^{(\zeta')} \leq \overline{w}^{(r)},$$

where $\overline{x}^{(i)}$ are the decision vectors from the lower bound problem after the r-th splitting round. Condition (1) excludes the case when $\overline{\mathcal{Z}}^{(r)}$ already contains a scenario whose decision vector in the lower bound problem after the r-th round is robust to ζ'. Condition (2) excludes the case when it is possible to construct a decision vector for ζ' that would have the same decisions as the vectors $x^{(i)}$ for the time periods where ζ is the same as elements of $\overline{\mathcal{Z}}^{(r)}$, and would give at most the same objective value.

Simple heuristic. We propose also an approach that combines approximately the properties of the two propositions above and is fast at the same time. The idea is to build up the lower-bounding set iteratively and add from each $\overline{\mathcal{Z}}_{r,s}$ the k scenarios whose sum of distances from the elements of $\overline{\mathcal{Z}}^{(r-1)}$ is largest. The distance between two vectors is measured by the 2-norm.

5.4 Stopping the algorithm

As the splitting continues, the computational workload related to solving the split problem grows because of the number of variables and uncertainty subsets. We propose three stopping rules for the splitting method: (1) when the objective value is closer to the lower bound than a predetermined threshold level, (2) when the limit of total computational time is reached, (3) when the maximum number of splitting rounds is reached.

6 Numerical experiments

6.1 Capital budgeting

The first numerical experiment involves no fixed recourse and is the capital budgeting problem taken from Hanasusanto et al. (2014). In this problem, a company can allocate an investment budget of B to a subset of projects $i \in \{1, \ldots, N\}$. Each project i has uncertain costs $c_i(\zeta)$ and uncertain profits $r_i(\zeta)$, modelled as affine functions of an uncertain vector ζ of risk factors. The company can invest in a project before or after observing the risk factors ζ. A postponed investment in project i incurs the same costs $c_i(\zeta)$, but yields only a fraction $\theta \in [0, 1)$ of the profits $r_i(\zeta)$.

The problem of maximizing the worst-case return can be formulated as:

$$\begin{array}{ll} \max & R \\ \text{s.t.} & R \leq r(\zeta)^T(x+\theta y), \quad \forall \zeta \in \mathcal{Z} \\ & c(\zeta)^T(x+y) \leq B, \quad \forall \zeta \in \mathcal{Z} \\ & x+y \leq 1 \\ & x, y \in \{0,1\}^N, \end{array}$$

where the decisions x_i and y_i attain value 1 if and only if an early or late investment in project i is undertaken, respectively. The uncertainty set is $\mathcal{Z} = [-1,1]^F$, where F is the number of risk factors.

We adopt the same random data setting as Hanasusanto et al. (2014). In all instances we use $F = 4$. The project costs and profits are modelled as:

$$c_i(\zeta) = (1 + \Phi_i^T \zeta/2)c_i^0, \qquad r_i(\zeta) = (1 + \Psi_i^T \zeta/2)r_i^0, \quad i = 1, \ldots, N.$$

Parameters c_i^0 and r_i^0 are the nominal costs and profits of project i, whereas Φ_i and Ψ_i represent the i-th rows of the factor loading matrices $\Phi, \Psi \in \mathbb{R}^{N \times 4}$ as column vectors. The nominal costs c^0 are sampled uniformly from $[0,10]^N$, and the nominal profits are set to $r^0 = c^0/5$. The components in each row of Φ and Ψ are sampled uniformly from the unit simplex in $\mathbb{R}^4$. The investment budget is set to $B = 1^T c^0/2$, and we set $\theta = 0.8$. Table 1 gives the results of Hanasusanto et al. (2014), who apply a K-adaptability approach and sample 100 instances for each combination of N and K (the number of time-2 decision variants) and try to solve it to optimality within a time limit of 2h per instance.

Table 1: Results of Hanasusanto et al. (2014). K is the number of time-2 decision variants allowed and N is the number of projects. The columns are (1) - percentage of instances solved to optimality within 2h, (2) - average solution time of the instances solved within 2h, (3) - average objective improvements (including the suboptimal solutions from Gurobi for the instances not solved within 2h.

	$K=2$			$K=3$			$K=4$		
N	(1) (%)	(2) s	(3) (%)	(1) (%)	(2) s	(3) (%)	(1) (%)	(2) s	(3) (%)
5	100	<1	48.67	100	1	68.71	100	36	79.50
10	100	4	59.34	74	1210	86.91	0	-	102.48
15	100	512	63.69	0	-	91.75	0	-	106.93
20	2	5232	64.78	0	-	93.20	0	-	108.61
25	0	-	64.85	0	-	93.72	0	-	109.10
30	0	-	64.98	0	-	94.08	0	-	109.42

We sample 50 instances for each N and conduct 8 splitting rounds for $N = 5, 10$, 6 for $N = 15, 20$ and 4 for $N = 25, 30$ (for smaller problems one can allow more splitting rounds to obtain better objectives and still operate within reasonable time limits). To split the uncertainty sets we use the worst-case scenarios coming from the optimal KKT vector of the LP relaxation of the robust MILP problems (see Section 2.3.2). In each splitting round we split all subsets $\mathcal{Z}_{r,s}$ for which $|\overline{\mathcal{Z}}_{r,s}| > 1$. The splitting hyperplanes are constructed using Heuristic 1 (see Section 5.2). The upper bound scenario sets are constructed according to the 'simple heuristic' (see Section 5.3) with $k = 2$. The after-splitting robust MILP problems are solved with

Table 2: Our results for the capital budgeting problem. ‘Splitting rounds’ denotes the number of splitting rounds conducted. ‘Average case improvement’ denotes the increase of the average-case objective value obtained with the adjustable decisions, relative to the one yielded by the static solution. The relative optimality gaps are computed as $\frac{(\text{UB}-\text{LB})}{0.5(\text{UB}+\text{LB})} * 100\%$, where OB is the objective function value and UB is the upper bound value. Remaining terminology is the same as in Table 5.

Splitting rounds	N	Obj improvement (%)	Initial gap (%)	Final gap (%)	Average case improvement (%)	Mean time (s)
8	5	57.89	106.09	39.00	12.11	5.40
	10	93.81	100.15	27.68	20.36	26.81
6	15	102.63	100.00	24.29	23.13	4.72
	20	107.81	100.00	22.24	24.79	5.43
4	25	105.33	100.00	23.48	24.30	3.96
	30	106.88	100.00	22.93	24.80	6.54

Gurobi precision set to 0.5%. All problems were formulated using CVX package and solved with Gurobi solver on an Intel Core 2.66GHz computer.

Apart from the worst-case results, for each instance we conduct a simulation study by sampling from $[-1,1]^4$ uniformly 500 scenarios of the risk factors’ values and computing the objective function values obtained using the static robust solution and our splitting-based adjustable solution.

Table 2 gives the results of our methodology. All the instances have been solved fast, with the largest average time equal to 26.81s. We remark here that, typically for problems with binary variables, the distribution of the solution times is heavy-tailed, and whereas most of the instances are solved within 2-3s, some instances take much more time and influence the average times in this way. Our methodology performs worse on the small instances, which the ‘more exact’ method of Hanasusanto et al. (2014) can solve efficiently in short time. For larger instances our improvements in the objective value are close to the best values of Hanasusanto et al. (2014) for larger instances $N = 20, 25, 30$ - ours being $107.81, 105.33, 106.88\%$ versus their $108.61, 109.10, 109.42\%$, respectively.

We also compare the running time performance of our method to the results of Hanasusanto et al. (2014) though we should mention that the main objective of Hanasusanto et al. (2014) was to find the best solution using a fixed number of time 2 policies. For larger instances ($N \geq 15$) the results of Hanasusanto et al. (2014) are based on suboptimal solutions from Gurobi obtained after 2 hours of computation per instance (see Table 1), whereas our method uses on average less than 27s per instance, with most of the mean times being less than 7s. Upon request, we obtained the Gurobi output of Hanasusanto et al. (2014). It reveals that in majority of instances studied by them, the objective value obtained by the solver after 60s is within 5% of the end objective value obtained after the time limit of 7200s, given in Table 1.

The right part of Table 2 gives the average-case improvements obtained using the adjustable decisions. The improvements are significantly smaller than the worst-case improvements, stabilizing around the level of 25% for larger N.

Figure 6 shows the average (over problem instances for given N) improvements of the worst-case objective functions and the upper bounds for all N. One can see that the relative gap between the upper bound values and the worst-case objective values decreases significantly with the number N of projects.

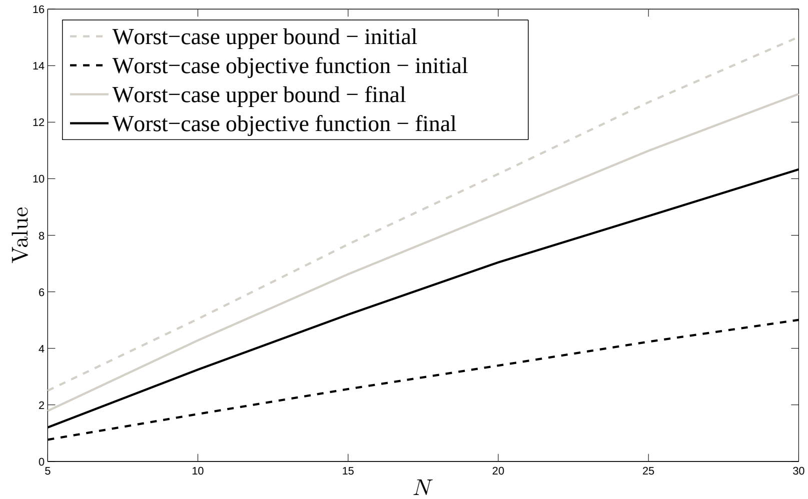

Figure 6: Plots of initial and final upper bound on the worst-case objective function values and the initial and final worst-case objective function values (average over all problem instances for each N).

We summarize now the results of the first numerical example. Hanasusanto et al. (2014) give good worst-case objective value improvements with a small number of time-2 decision variants (at most 4) after a longer computation time, whereas our splitting method gives such improvements after a short computation time, but with more time-2 decision variants. For example, 9 splitting rounds typically result in a division of the uncertainty set $\mathcal{Z}$ into more than 10 parts, each with a corresponding time-2 decision variant. Thus, our methodology is preferred when it is the computation time, and not the number of decision variants, that is to be kept low.

6.2 Lot sizing problem

As the second numerical experiment we consider a multi-stage lot sizing problem taken from Bertsimas and Georghiou (2014a). The problem entails a single product, T time periods, and the following parameters:

- ζ_t, where $t = 1, \ldots, T$, is the uncertain demand in period t
- l_t, where $t = 2, \ldots, T$, is the lowest possible demand in period t
- u_t, where $t = 2, \ldots, T$, is the highest possible demand in period t
- c_{y_n}, where $n = 1, \ldots, N$, is the cost of buying a fixed quantity q_n of the product
- c_x is the ordering cost per product unit for purchases that are delivered in the subsequent period
- c_h is the holding cost per product unit
- $\overline{x}_{\mathrm{tot},t}$, where $t = 2, \ldots, T$, is the cumulative orders limit up to time period t.

The variables are:

- I_t, where $t = 1, \dots, T$, is the level of available inventory after period t
- x_t, where $t = 1, \dots, T-1$ is the product amount ordered in period t, after $\zeta_1, \dots, \zeta_t$ is known, and delivered in period $t+1$, at unit price c_x
- y_{nt}, where $n = 1, \dots, N$, $t = 2, \dots, T$, is a binary decision made after $\zeta_1, \dots, \zeta_t$ is known, whether to buy a fixed quantity q_n of the product in time period t, delivered in the same time period.

The problem is to minimize the worst-case combined ordering and holding costs (referred later to as the 'total cost'), subject to cumulative ordering constraints:

$$
\begin{array}{ll}
\min & z \\
\text{s.t.} & \sum\limits_{t=2}^{T} \Big(c_x x_{t-1}(\zeta_{1:t-1}) + c_h I_t(\zeta_{1:t}) + \sum_{n=1}^{N} c_{y_n} q_n y_{nt}(\zeta_{1:t}) \Big) \le z, \quad \forall \zeta \in \mathcal{Z} \\
& \left.\begin{array}{l}
I_t(\zeta_{1:t}) = I_{t-1}(\zeta_{1:t-1}) + x_{t-1}(\zeta_{1:t-1}) + \sum_{n=1}^{N} q_n y_{nt}(\zeta_{1:t}) - \zeta_t \\
0 \le x_{t-1}(\zeta_{1:t-1}) \\
0 \le I_t(\zeta_{1:t}) \\
\sum\limits_{j=1}^{t-1} x_j(\zeta_{1:j}) \le \overline{x}_{\text{tot},t}
\end{array}\right\} \begin{array}{c} \forall t = 2, \dots, T, \\ \forall \zeta \in \mathcal{Z} \end{array} \\
& y_{nt}(\zeta_{1:t}) \in \{0, 1\}, \quad \forall n, t \\
& x_t(\zeta_{1:t}) \ge 0, \quad \forall t
\end{array}
\tag{19}
$$

where

$$
\mathcal{Z} = \{ \zeta : \zeta_1 = 1, \quad l_t \le \zeta_t \le u_t, \quad t = 2, \dots, T \}.
$$

The above problem is transformed by eliminating the variables I_t for $t = 2, \dots, T$. The adjustable variables are x_t, allowed to depend on $\zeta_{1:t}$ for $t = 1, \dots, T-1$ and y_{nt}, allowed to depend on $\zeta_{1:t}$ for $t = 2, \dots, T$.

Problem parameters are sampled as in Bertsimas and Georghiou (2014a). Ordering costs are chosen from $c_x \in [0; 5]$ and $c_{y_n} \in [0; 10]$, separately for all $n = 1, \dots, N$, such that $c_x < c_{y_n}$. Holding costs are elements of $c_h \in [0; 10]$ with the fixed ordering quantities set to $q_n = 100/N$ for all $n = 1, \dots, N$. The cumulative ordering budget is set to $\overline{x}_{\text{tot},t} = \sum_{s=1}^{t-1} \bar{x}_s$ for $t = 2, \dots, T$, with $\bar{x}_t \in [0; 100]$ and the lower and upper bounds for the demand are sampled uniformly as $l_t \in [0; 25]$ and $u_t \in [75; 100]$, $t = 2, \dots, T$. We assume that the initial inventory level I_1 equals zero. Table 3 gives the results obtained by Bertsimas and Georghiou (2014a) using their methodology of piecewise linear decision rules for the decision variables.

We sample and solve 50 instances of the problem for $N = 2, 3$ and $T = 2, 4, \dots, 10$. Since $q_n = 100/N$ for all n and the splitting method facilitates use of integer non-binary variables, we substitute $z_t(\zeta_{1:t}) = \sum_{n=1}^{N} y_{nt}(\zeta_{1:t})$ for all $t = 2, \dots, T$, such that $0 \le z_t(\zeta_{1:t}) \le N$ for all t. In this way we switch from binary to integer variables in order to reduce the problem size.

Since problem (19) involves fixed recourse only, we study also the impact of using linear decision rules for the continuous variables $x_t(\zeta_{1:t})$. In such case we set $x_t(\zeta_{1:t})$

Table 3: Results of Bertsimas and Georghiou (2014a). The relative optimality gaps are computed as $\frac{(\mathrm{OB}-\mathrm{LB})}{0.5(\mathrm{OB}+\mathrm{LB})} * 100\%$, where OB is the objective function value and LB is the lower bound value. 'Nonadaptive gap' denotes the relative optimality gap computed for the solution where the integer decisions are static and the linear decision rules are implemented for the continuous decision variables. '$\mathcal{PB}_t(1)$ gap' denotes relative optimality gap computed for the solution obtained using the binary adjustability technique used by the authors and where the linear decision rules are implemented for the continuous decision variables. 1% and 5% at the top of the Table are two variants of solver precision used when solving the MILP problems.

		1% optimality			5% optimality		
N	T	$\mathcal{PB}_t(1)$ gap (%)	Nonadaptive gap (%)	Mean time (s)	$\mathcal{PB}_t(1)$ gap (%)	Nonadaptive gap (%)	Mean time (s)
2	2	0	17.6	0.1	0.6	17.6	0.4
	4	24.2	68.6	50.6	27.3	68.6	45.5
	6	37.4	62.0	4833.8	38.9	62.1	956.8
	8	37.9	84.4	27531.1	38.0	84.4	19573.1
	10	39.7	89.9	35716.6	42.0	89.9	31464.1
3	2	0	27.6	0.1	1.2	27.6	0.1
	4	17.2	73.3	3381.8	23.9	73.3	781.6
	6	34.5	66.2	9181.0	38.4	66.1	3298.1
	8	37.6	83.4	28742.7	38.1	83.7	21885.5
	10	-	89.7	-	41.1	90.7	39141.5

to be an affine function of $\zeta_1, \ldots, \zeta_t$:

$$x_t(\zeta_{1:t}) = \alpha_{t,0} + \sum_{j=1}^{t} \alpha_{t,j}\zeta_j, \quad \forall t = 1, \ldots, T-1,$$

where $\alpha_{t,j}$ are then treated as decision variables implemented in period t.

Each problem instance is solved in four ways: 1) applying static decisions to all variables 2) applying linear decision rules to the continuous variables and static decisions to the integer variables 3) applying only the splitting methodology to all variables 4) applying the splitting methodology to all variables, combined with linear decision rules for the continuous decisions (the parameters $\alpha_{t,j}$ can also differ after splitting of the uncertainty set).

For each instance we conduct 4 splitting rounds. For splitting we use the worst-case scenario sets obtained using optimal KKT vectors from the robust counterpart of the LP relaxation of the problem (see Sections 2.3.2 and 3.3.2). In each splitting round we split all subsets $\mathcal{Z}_{r,s}$ for which $|\overline{\mathcal{Z}}_{r,s}| > 1$. Time periods t for the t-SHs are chosen according to the biggest variance of uncertain demands from subsequent periods with $q = 2$ (see Section 5.1). Splitting hyperplanes are constructed using Heuristic 1 (see Section 5.2). The scenario sets for the lower bound problems are constructed according to the 'simple heuristic' (see Section 5.3) with $k = 2$. For $T = 2, 4$ the lower bound scenario sets include also all vertices of the uncertainty set $\mathcal{Z}$. The after-splitting robust MILP problems are solved with Gurobi precision (the relative duality gap when the solver stops) equal to 0.1%. All problems were formulated using CVX package and solved with Gurobi solver on an Intel Core 2.66GHz computer.

Tables 4 and 5 give our results for $N = 2$ and $N = 3$, respectively. All methodologies offer substantial improvements in the objective value compared to the static robust solution. Also, combination of our splitting methodology with linear decision rules

Table 4: Our results for the lot sizing problem for $N = 2$. LDR stands for the solution with linear decision rules for the continuous decision variables and static decisions for the integer variables, S stands for only our splitting methodology applied to all variables, S+LDR stands for a combination of set splitting with linear decision rules for the continuous variables. 'Objective improvement' is the the decrease in the average worst-case objective value reduction, relative to the static robust solution. Optimality gaps are computed as in Table 3. 'Initial gap' is the optimality gap for the static robust solution and the lower bound obtained after the first splitting round. 'Final gap' is the optimality gap computed with the objective value and lower bound after the last splitting round. The asterisk indicates the fact that for $T = 2, 4$ the lower bound scenario sets include also all vertices of the uncertainty set $\mathcal{Z}$. All the static robust problems were solved in less than 2s.

	Objective improvement (%)			Initial gap (%)	Final gap (%)			Mean time (s)	
T	LDR	S	S+LDR		LDR	S	S+LDR	S	S+LDR
2	0	11.39	11.38	51.02∗	51.02∗	15.49∗	15.51∗	2.36	2.77
4	31.64	28.07	42.32	85.78∗	52.46∗	57.34∗	34.04∗	5.67	7.69
6	43.77	30.29	54.94	113.14	69.22	87.51	47.39	5.64	10.09
8	48.91	26.32	61.01	125.59	78.73	107.17	54.68	7.54	15.03
10	52.09	22.43	64.21	134.65	86.16	121.02	61.85	9.23	24.23

Table 5: Our results for the lot sizing problem for $N = 3$. Terminology is the same as in Table 4.

	Objective improvement (%)			Initial gap (%)	Final gap (%)			Mean time (s)	
T	LDR	S	S+LDR		LDR	S	S+LDR	S	S+LDR
2	0	22.94	22.94	61.90∗	61.90∗	17.64∗	17.64∗	2.25	2.61
4	32.66	31.70	47.22	95.06∗	62.30∗	65.09∗	39.14∗	5.24	7.39
6	43.99	29.41	56.86	118.38	78.36	96.55	54.14	5.85	9.80
8	50.14	25.13	62.05	129.06	85.27	113.58	61.48	7.11	14.18
10	53.21	21.42	64.82	136.55	92.22	125.08	68.88	9.18	55.82

(S+LDR) gives a strong combined effect - the objective value improves significantly more than using any of the methods S or LDR separately - by as much as 64.82% for $N = 3, T = 10$, compared to 53.21% for LDR and 21.42% for S. For $T = 2$ the linear decision rules cannot bring any improvement because x_1 is a scalar. One can observe that for problems with larger T our methodology gives better objective improvements. Also, the relative optimality gaps decrease significantly in all cases, mostly due to improvements in the objective function. All problems have been solved fast, with the maximum mean time equal to 55.82s.

We compare now our results to those of Bertsimas and Georghiou (2014a). The main difference between the methods lies in the fact that decision rules proposed by Bertsimas and Georghiou (2014a) satisfy the problem's constraints with a high probability (99%), obtained using Hoeffding bounds, whereas our methodology ensures 100% robustness by design. Comparing the numbers from Tables 3 (column '$\mathcal{PB}_t(1)$ gap'), 4, and 5 (columns 'Final gap (%) - S+LDR'), one can see that our methodology performs worse in terms of the final optimality gap. For example, for $N = 2, T = 4$ our result is 39.16% compared to their 24.2% for $N = 2, T = 4$. This can be partly explained by the difference between the types of robustness, and also by different way of choosing the scenarios for the lower bounding problems. On the other hand, our method provides significantly faster computation times which, combined with full robustness, may be an appealing property. In particular, this is visible on larger instances, with our mean solution times being significantly lower, e.g., our 55.82s compared to 39141.5s for $N = 3, T = 10$.

Table 6: Average-case performance of the solutions obtained using the three methodologies in comparison to the static robust solution. 'Average-case improvement' is the average reduction of the total cost, relative to the total cost obtained with the static solution for the given demand scenario.

	Average-case improvement (%)					
	$N = 2$			$N = 3$		
T	LDR	S	S+LDR	LDR	S	S+LDR
2	0.00	18.55	18.55	0.00	14.59	14.13
4	21.87	22.91	31.90	24.51	26.51	37.65
6	30.80	23.24	41.02	33.69	22.72	45.81
8	35.23	20.00	48.05	40.94	18.83	51.56
10	39.67	16.95	51.68	43.79	15.78	55.07

In addition to the worst-case results, for each solved instance we conduct a simulation study. In each of them we sample uniformly 500 demand scenario realizations $l \leq d_{\text{realized}} \leq u$ and compute the average total cost incurred by 1) the static robust policy 2) the linear decision rules for the continuous variables and static decisions for the integer variables (LDR) 3) splitting-based decisions for all variables (S) 4) combination splitting based decisions with linear decision rules for all variables (S+LDR). Table 6 gives the results on average-case improvements relative to the static robust solution. The table shows that our method not only offers substantial improvements on the worst-case basis, but also in terms of the average-case total cost, in particular when combined with the linear decision rules for the continuous variables.

To sum up the results of this numerical example, the main benefits of our approach have been: 1) fast computation time even for large problems, corresponding to the number of splitting rounds (the more splitting rounds, the better the improvement in the objective, but also the longer computation time), 2) substantial improvements in the objective function value, 3) robustness to the entire uncertainty set after each splitting round. Due to the 100% robustness of our method, one can set the MILP solver precision to a non-default value, obtaining robust feasible solutions with only slightly worse objective value than using higher precision, after a significantly shorter time.

7 Conclusions

In this paper we have introduced the method of iterative splitting of the uncertainty set for multi-period robust mixed-integer linear optimization problems. We have provided theory on how to determine efficiently which scenarios of the uncertain parameter are more important to be separated than others and how to obtain lower bounds for the adjustable worst-case value. Based on these theoretical results, we have proposed several heuristics for each part of the method.

Our approach can be used to a variety of problems. In particular, this applies to problems with a non-fixed recourse and adjustable integer variables (also non-binary), where implementation of other decision rules may not be possible or may involve large computational effort. For adjustable continuous variables in the non-fixed recourse setting, our method bypasses the challenge of dealing with interactions of uncertain parameters, as would be the case with linear or polynomial decision rules.

For fixed recourse problems the splitting method can be combined with other decision rules, such as linear decision rules, allowing them to take different forms over different parts of the uncertainty set. The second numerical experiment reveals that such a combination gives a strong joint effect. Our iterative method guarantees robustness of the decisions to the entire uncertainty set after each of the splitting rounds. Thus, depending on time constraints, the decision maker can set how many splitting rounds to conduct, with each additional round costing additional effort but bringing potentially extra improvements in the objective value.

Numerical experiments conducted on problems from Bertsimas and Georghiou (2014a) and Hanasusanto et al. (2014) have shown our methodology to perform well in practice. In both cases was our method outperformed on small problem instances. However, as the problems grow, our methodology was giving comparable results after only a fraction of the computation time of other authors.

We give now potential directions for further research. First, more theoretical results can be obtained regarding the choice of best splits of the uncertainty sets, and in particular, the 'best' distribution of the splits in time. Secondly, it is important to obtain better lower bound values, possibly by combining our method with results of other authors, e.g., Kuhn et al. (2011). Last, it is interesting to investigate whether our method, combined with the results of Ben-Tal et al. (2014), can be used efficiently in multistage nonlinear robust optimization problems.

Acknowledgements

We are grateful to Angelos Georghiou and Grani Hanasusanto for their explanations and sharing with us the exact results of their papers.

References

[1] Ben-Tal, A., Goryashko, A., Guslitzer, E., & Nemirovski, A. (2004). Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, Vol. 99(2), pp. 351-376.

[2] Beck, A., & Ben-Tal, A. (2009). Duality in robust optimization: primal worst equals dual best. *Operations Research Letters*, Vol. 37(1), pp. 1-6.

[3] Ben-Tal, A., El Ghaoui, L. & Nemirovski, A. (2009). *Robust optimization.* (Princeton University Press)

[4] Ben-Tal, A., Den Hertog, D., & Vial, J.-Ph. (2014). Deriving robust counterparts of nonlinear uncertain inequalities (Online First). *Mathematical Programming.*

[5] Bertsimas, D., & Caramanis, C. (2007). Adaptability via sampling. In *46th IEEE Conference on Decision and Control* (pp. 4717-4722). IEEE.

[6] Bertsimas, D., & Caramanis, C. (2010). Finite adaptability in multistage linear optimization. *IEEE Transactions on Automatic Control*, Vol. 55(12), pp. 2751-2766.

[7] Bertsimas, D., & Goyal, V. (2010). On the power of robust solutions in two-stage stochastic and adaptive optimization problems. *Mathematics of Operations Research*, Vol. 35(2), pp. 284-305.

[8] Bertsimas, D., Iancu, D.A., & Parrilo, P.A. (2010). Optimality of affine policies in multistage robust optimization. *Mathematics of Operations Research*, Vol. 35(2), pp. 363-394.

[9] Bertsimas, D., Goyal, V., & Sun, X.A. (2011a). A geometric characterization of the power of finite adaptability in multistage stochastic and adaptive optimization. *Mathematics of Operations Research*, Vol. 36(1), pp. 24-54.

[10] Bertsimas, D., Iancu, D.A., & Parrilo, P.A. (2011b). A hierarchy of near-optimal policies for multistage adaptive optimization. In *Automatic Control, IEEE Transactions on*, Vol. 56(12), pp. 2809-2824.

[11] Bertsimas, D., & Georghiou, A. (2014a). Design of near optimal decision rules in multistage adaptive mixed-integer optimization. Available at: `http://www.optimization-online.org/DB_FILE/2013/09/4027.pdf`.

[12] Bertsimas, D., & Georghiou, A. (2014b). Binary decision dules for multistage adaptive mixed-integer optimization. Available at: `http://www.optimization-online.org/DB_FILE/2014/08/4510.pdf`.

[13] Chen, X., Sim, M., & Sun, P. (2007). A robust optimization perspective on stochastic programming. *Operations Research*, Vol. 55(6), pp. 1058-1071.

[14] Chen, X., & Zhang, Y. (2009). Uncertain linear programs: Extended affinely adjustable robust counterparts. *Operations Research*, Vol. 57(6), pp. 1469-1482.

[15] Gorissen, B. L., Ben-Tal, A., Blanc, H., & Den Hertog, D. (2014). A new method for deriving robust and globalized robust solutions of uncertain linear conic optimization problems having general convex uncertainty sets (Online First). *Operations Research*.

[16] Grant, M., & Boyd, S. (2014). CVX: Matlab software for disciplined convex programming, version 2.0 beta. `http://cvxr.com/cvx`.

[17] Gurobi Optimization, Inc. (2014). Gurobi Optimizer Reference Manual, `http://www.gurobi.com`.

[18] Hadjiyiannis, M.J., Goulart, P.J., & Kuhn, D. (2011). A scenario approach for estimating the suboptimality of linear decision rules in two-stage robust optimization. *IEEE Conference on Decision and Control and European Control Conference*, Orlando, USA.

[19] Hanasusanto, G.A., Kuhn, D., & Wiesemann, W. (2014). Two-stage robust integer programming. Available online at: `http://www.optimization-online.org/DB_FILE/2014/03/4294.pdf`.

[20] Kuhn, D., Wiesemann, W., & Georghiou, A. (2011). Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming*, Vol. 130(1), pp. 177-209.

[21] Vayanos, P., Kuhn, D., & Rustem, B. (2011). Decision rules for information discovery in multi-stage stochastic programming. In *2011 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC),* (pp. 7368-7373).