Mathematical Programming
https://doi.org/10.1007/s10107-023-02053-0

FULL LENGTH PAPER

Series A

Designing tractable piecewise affine policies for multi-stage adjustable robust optimization

Simon Thomä[1] · **Grit Walther**[1] · **Maximilian Schiffer**[2]

Received: 11 January 2023 / Accepted: 19 December 2023
© The Author(s) 2024

Abstract
We study piecewise affine policies for multi-stage adjustable robust optimization (ARO) problems with non-negative right-hand side uncertainty. First, we construct new dominating uncertainty sets and show how a multi-stage ARO problem can be solved efficiently with a linear program when uncertainty is replaced by these new sets. We then demonstrate how solutions for this alternative problem can be transformed into solutions for the original problem. By carefully choosing the dominating sets, we prove strong approximation bounds for our policies and extend many previously best-known bounds for the two-staged problem variant to its multi-stage counterpart. Moreover, the new bounds are—to the best of our knowledge—the first bounds shown for the general multi-stage ARO problem considered. We extensively compare our policies to other policies from the literature and prove relative performance guarantees. In two numerical experiments, we identify beneficial and disadvantageous properties for different policies and present effective adjustments to tackle the most critical disadvantages of our policies. Overall, the experiments show that our piecewise affine policies can be computed by orders of magnitude faster than affine policies, while often yielding comparable or even better results.

Keywords Robust optimization · Multi-stage problem · Adjustable optimization · Approximation bounds

Mathematics Subject Classification 90C47 · 90C57 · 90C59 · 49K35

✉ Simon Thomä
simon.thomae@om.rwth-aachen.de

Grit Walther
walther@om.rwth-aachen.de

Maximilian Schiffer
schiffer@tum.de

[1] Chair of Operations Management, RWTH Aachen University, Aachen, Germany

[2] School of Management and Munich Data Science Institute, Technical University of Munich, Munich, Germany

Published online: 03 February 2024

1 Introduction

In practice, most decision-making problems have to be solved in view of uncertain parameters. In the operations research domain, two fundamental frameworks exist to inform such decisions. The *stochastic optimization* framework captures uncertainty by using probability distributions and aims to optimize an expected objective. Initially introduced in the seminal work of Dantzig [26], the framework has been intensively studied and scholars used it to solve a vast variety of problems including production planning [6, 55], relief networks [49], expansion planning [62, 64], and newsvendor problems [52]. While stochastic optimization performs well on many problem classes, finding tractable formulations is oftentimes challenging. Additionally, the data needed to approximate probability distributions might not always be available, and gathering data is often a costly and time-consuming process. The *robust optimization* (RO) framework overcomes many of these shortcomings by capturing uncertainty through distribution-free uncertainty sets instead of probability distributions. By choosing well-representable uncertainty sets, RO often offers computationally tractable formulations that scale well on a variety of optimization problems. In recent years, these favorable properties have led to a steep increase in research interest, see, e.g. [8, 11, 20, 36, 63]. The flexibility of uncertainty sets and scaleability of solution methods also make RO very attractive for applicational purposes and it has been widely applied to many operations management problems [45].

In traditional RO, all decisions must be made before the uncertainty realization is revealed. However, oftentimes some decisions can be delayed until after (part of) the uncertainty realization is known in real-world situations. As a consequence, RO may lead to excessively conservative solutions. To remedy this drawback, Ben-Tal et al. [9] introduced the concept of *adjustable robust optimization* (ARO) where some decisions can be delayed until the uncertainty realization is (partly) known. In general ARO, uncertainty realizations are revealed over multiple stages and decisions can be made after each reveal. A decision made in stage t can thus be modeled as a function of all uncertainties associated with previous stages $t' \leq t$.

While ARO improves decision-making in theory, ARO is equivalent to RO on some special problem instances, where static decision policies yield optimal adjustable solutions [9, 22]. Using similar arguments to the ones used for the optimality of static solutions, Marandi and den Hertog [46] identified conditions where optimal adjustable decisions are independent of some uncertain parameters. In general, static policies do not yield optimal solutions in the ARO setting. Elucidating this, Haddad-Sisakht and Ryan [37] identified a collection of sufficient conditions that imply the suboptimality of static policies and a strict improvement of ARO over RO.

In general, even the task of finding optimal adjustable solutions in the special case of two-stage ARO proves to be computationally intractable [9]. Accordingly, recent works developed many approximation schemes for ARO that often yield good and sometimes even optimal results in practice, see, e.g. [9, 19, 42, 58]. In the special case with only two decision stages, the first stage decisions are fixed before any uncertain parameters are known and the second stage decisions use full knowledge of the uncertainty realization. Two-stage ARO already has many applications in practice and has widely been studied in the literature, see, e.g. [14, 16, 18, 22, 23, 39, 40, 59].

Still, many real-world problems show inherent multi-stage characteristics and cannot be modeled by two-stage ARO. Examples include variants of inventory management [12], humanitarian relief [13], and facility location [5]. The transition from two-stage to multi-stage ARO introduces two main challenges. First, multi-stage ARO problems entail *nonanticipativity restrictions* that disallow decisions to utilize future information. Second, many approaches that solve the problem by iteratively splitting the uncertainty space, like scenario trees [41], and adaptive partitioning [16, 53], grow exponentially in the number of stages. As a consequence, many results found for two-stage ARO do not readily generalize to multi-stage scenarios. Against this background, we design piecewise affine policies for multi-stage ARO that overcome the previously mentioned challenges and extend, although it is not straightforward, many of the best know approximation bounds for two-stage problems to a multi-stage setting.

In the remainder of this section, we formally introduce our problem (Sect. 1.1), discuss closely related work (Sect. 1.2), and summarize our main contributions (Sect. 1.3).

1.1 Problem description

In this work we study multi-stage adjustable robust optimization with covering constraints and a positive affine uncertain right hand side. Specifically, we consider the following problem:

$$\begin{aligned} Z_{AR}(\mathcal{U}) = \min_{\boldsymbol{x}(\boldsymbol{\xi})} \max_{\boldsymbol{\xi}\in\mathcal{U}} \quad & \boldsymbol{c}^\intercal \boldsymbol{x}(\boldsymbol{\xi}) \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x}(\boldsymbol{\xi}) \geq \boldsymbol{D}\boldsymbol{\xi} + \boldsymbol{d} \quad \forall \boldsymbol{\xi} \in \mathcal{U} \end{aligned} \tag{1}$$

with $\boldsymbol{A} \in \mathbb{R}^{l\times n}$, $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{D} \in \mathbb{R}_+^{l\times m}$, $\boldsymbol{d} \in \mathbb{R}_+^l$, and compact $\mathcal{U} \subset \mathbb{R}_+^m$. Here, m is the number of uncertain parameters, n is the number of decisions, and l is the number of constraints. To model the problem's T stages, we split the uncertainty vector $\boldsymbol{\xi}$ into T sub-vectors $\boldsymbol{\xi} = \left(\boldsymbol{\xi}^1, \ldots, \boldsymbol{\xi}^T\right)$ with $\boldsymbol{\xi}^t$ being the uncertainty vector realized in stage t. In the following, we denote by $\underline{\boldsymbol{\xi}}^t := \left(\boldsymbol{\xi}^1, \ldots, \boldsymbol{\xi}^t\right)$ the vector of all uncertainties with known realization in stage t. Similarly, the adjustable decision vector $\boldsymbol{x}(\boldsymbol{\xi})$ divides into $\boldsymbol{x}(\boldsymbol{\xi}) := \left(\boldsymbol{x}^1(\underline{\boldsymbol{\xi}}^1), \ldots, \boldsymbol{x}^T(\underline{\boldsymbol{\xi}}^T)\right)$, where the decision $\boldsymbol{x}^t$ made in stage t has to preserve *nonanticipativity* and may only depend on those uncertainties $\underline{\boldsymbol{\xi}}^t$ whose realization is known in stage t. We explicitly allow $\boldsymbol{\xi}^1$ to be zero-dimensional making the initial decision $\boldsymbol{x}^1$ non-adjustable. Finally, we denote by $\underline{\boldsymbol{x}}^t(\boldsymbol{\xi}) := \left(\boldsymbol{x}^1(\underline{\boldsymbol{\xi}}^1), \ldots, \boldsymbol{x}^t(\underline{\boldsymbol{\xi}}^t)\right)$ the vector of all decisions in the first t stages. Figure 1 visualizes the multi-stage decision process with nonanticipativity restrictions.

Unless explicitly stated otherwise, we assume w.l.o.g. that the following assumption holds throughout the paper.

Assumption 1 $\mathcal{U} \subseteq [0, 1]^m$ is convex, full-dimensional with $e_i \in \mathcal{U}$ for all $i \in \{1, \ldots, m\}$, and *down-monotone*, i.e., $\forall \boldsymbol{\xi} \in \mathcal{U}, \mathbf{0} \leq \boldsymbol{\xi}' \leq \boldsymbol{\xi} : \boldsymbol{\xi}' \in \mathcal{U}$.

Down-monotonicity holds because $\boldsymbol{D}, \boldsymbol{d}, \boldsymbol{\xi}$ are all non-negative, and thus constraints become less restrictive for smaller values of $\boldsymbol{\xi}$. Convexity holds due to the linearity

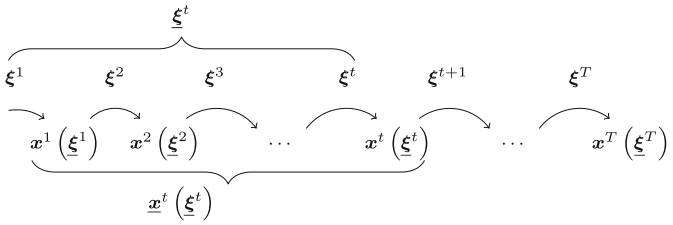


Fig. 1 Illustration of multi-stage decision making over T stages. In each stage t a fraction $\boldsymbol{\xi}^t$ of the uncertainty is realized and decisions $\boldsymbol{x}^t$ are made. Here, decisions $\boldsymbol{x}^t$ may only depend on those uncertainties $\underline{\boldsymbol{\xi}}^t$ whose realization is known in stage t

of the problem, and $\boldsymbol{e}_i \in \mathcal{U} \subseteq [0,1]^m$ holds as $\mathcal{U}$ is compact and $\boldsymbol{D}$ can be re-scaled appropriately. We note that the non-negativity assumption of the right-hand side does restrict the problem space. As Bertsimas and Goyal [18] point out, this assumption prevents the introduction of uncertain or constant upper bounds. However, upper bounds in other decision variables are still possible as $\boldsymbol{A}$ is not restricted, and $\boldsymbol{D}, \boldsymbol{d}$ can be zero. Overall, Problem (1) covers a wide range of different problem classes including network design [50, 61], capacity planning [44, 48, 51], as well as versions of inventory management [12, 60] where capacities are unbounded or subject to the decision makers choice.

In the context of multi-stage decision making, some works require *stagewise* uncertainty, i.e., that the uncertainty set $\mathcal{U}$ consists of uncertainty sets $\mathcal{U}_1, \ldots, \mathcal{U}_T$ for each stage, see, e.g., [19, 20, 32]. Like other approaches based on decision rules [17, 43], we do not need these restricting assumptions. However, we show how to utilize the existence of such a structure in Sect. 3.

1.2 Related work

Feige et al. [30] show that already the two-stage version of Problem (1) with $\boldsymbol{D} = \boldsymbol{1}, \boldsymbol{d} = \boldsymbol{0}$ and $\boldsymbol{A}$ being a 0-1-matrix is hard to approximate with a factor better than $\Omega(\log m)$, even for budgeted uncertainty sets. As it is thereby impossible to find general solutions for $\boldsymbol{x}$, a common technique to get tractable formulations is to restrict the function space.

In this context, Ben-Tal et al. [9] consider $\boldsymbol{x}$ to be affine in $\boldsymbol{\xi}$. Specifically, they propose $\boldsymbol{x}^t$ to be of the form $\boldsymbol{x}^t(\underline{\boldsymbol{\xi}}^t) = \boldsymbol{P}^t\underline{\boldsymbol{\xi}}^t + \boldsymbol{q}^t$. Affine policies have been found to deliver good results in practice [1, 10] and are even optimal for some special problems [19, 42, 58]. Further popular decision rules include segregate affine [24, 25], piecewise constant [20], piecewise affine [14, 31], and polynomial [4, 21] policies, as well as combinations of these [54]. For surveys on adjustable policies we refer to Delage and Iancu [28] and Yanıkoğlu et al. [63].

A key question that arises when using policies to solve ARO problems is how good the solutions are compared to an optimal unrestricted solution. To answer this, many approximation schemes for a priori and a posteriori bounds have been proposed. In the context of a posteriori bounds the focus lies on finding tight upper and lower approximation problems. Hadjiyiannis et al. [38] estimate the suboptimality of affine decision

rules using sample scenarios from the uncertainty set. Similar sample lower bounds are used by Bertsimas and Georghiou [17] to bound the performance of piecewise affine policies. Kuhn et al. [43] investigate the optimality of affine policies by using the gap between affine solutions on the primal and the dual of the problem. Georghiou et al. [31] generalize this primal-dual approach to affine policies on lifted uncertainty sets. Building on both of the previous approaches, Georghiou et al. [33] propose a convergent hierarchy of policies that combine affine policies with extreme point scenario samples. Daryalal et al. [27] construct lower bounds by relaxing nonanticipativity and stage-connecting constraints in multi-stage ARO. They then use these lower bounds to construct primal solutions in a rolling horizon manner.

In the context of a priori bounds, most approximation schemes have been proposed for the two-staged version of Problem (1). For general uncertainty sets on the two-stage version of (1), Bertsimas and Goyal [18] show that affine policies yield an $O(\sqrt{m})$ approximation if $\boldsymbol{c}$ and $\boldsymbol{x}$ are non-negative. They further construct a set of instances where this bound is tight, showing that no better general bounds for affine policies exist. Using geometric properties of the uncertainty sets, Bertsimas and Bidkhori [15] improve on these bounds for some commonly used sets including budgeted uncertainty, norm balls, and intersections of norm balls. Ben-Tal et al. [14] propose new piecewise affine decision rules for the two-stage problem that on some sets improve these bounds even further. In addition to strong theoretical bounds, this new approach also yields promising numerical results that can be found by orders of magnitude faster than solutions for affine adjustable policies. For budgeted uncertainty sets and some generalizations thereof, Housni and Goyal [40] show that affine policies even yield optimal approximations with an *asymptotic bound*, i.e., asymptotic behavior of the approximation bound, see, e.g. [56], of $O\left(\frac{\log m}{\log \log m}\right)$. This bound was shown to be tight by Feige et al. [30] for reasonable complexity assumptions, namely 3SAT cannot be solved in $2^{O(\sqrt{m})}$ time on instances of size m. We present an overview of known a priori approximation bounds for some commonly used uncertainty sets on two-stage ARO in Table 4 of "Appendix A".

To the best of our knowledge, Bertsimas et al. [20] are the only ones that provide a priori bounds for multi-stage ARO so far. They show these bounds for piecewise constant policies using geometric properties of the uncertainty sets. More specifically, they consider multi-stage uncertainty networks, where the uncertainty realization is taken from one of multiple independent uncertainty sets in each stage. While the choice of the set selected in each stage may depend on the sets selected before, the uncertainty sets are otherwise independent. Although this assumption is fairly general, it still leaves many commonly used sets where uncertainty is dependent over multiple stages uncovered. Among others, uncovered sets include widely used hypersphere and budgeted uncertainty.

As can be seen, previous work has predominantly focused on providing tighter approximation bounds for two-stage ARO, inevitably raising the question of whether similar bounds hold for multi-stage ARO as well. This work contributes towards answering this question by extending many of the currently best-known a priori bounds on two-stage ARO to its multi-stage setting. By so doing, we are—to the best of our knowledge—the first ones to provide a priori approximation bounds for the multi-

stage ARO Problem (1), where uncertainty sets can range over multiple stages. Unless explicitly stated otherwise, we will always refer to a priori approximation bounds when we discuss approximation bounds in the remainder of this paper.

1.3 Our contributions

With this work, we extend the existing literature in multiple ways, where our main contributions are as follows.

Tractable piecewise affine policies for multi-stage ARO: Motivated by piecewise affine policies for two-stage ARO [14], we present a framework to construct policies that can be used to efficiently find good solutions for the multi-stage ARO Problem (1). Instead of solving the problem directly for uncertainty $\mathcal{U}$, we first approximate $\mathcal{U}$ by a dominating set $\hat{\mathcal{U}}$. To do so, we define the concept of nonanticipative multi-stage domination and show that this new definition of domination fulfills similar properties to two-stage domination. Based on this new definition, we then construct dominating sets $\hat{\mathcal{U}}$ such that solutions on $\hat{\mathcal{U}}$ can be found efficiently. More specifically, we choose $\hat{\mathcal{U}}$ to be a polytope for which worst-case solutions can be computed by a linear program (LP) over its vertices. In order to ensure nonanticipativity, which is the main challenge of this construction, we introduce a new set of constraints on the vertices that guarantee the existence of nonanticipative extensions from the vertex solutions to the full set $\hat{\mathcal{U}}$. Finally, we show how to use the solution on the dominating set $\hat{\mathcal{U}}$ to construct a valid solution for the original uncertainty set $\mathcal{U}$.

Approximation bounds: To the best of our knowledge, we provide the first approximation bounds for the multi-stage Problem (1) with general uncertainty sets. More specifically, we show that our policies yield $O(\sqrt{m})$ approximations of fully adjustable policies. While this bound is tight for our type of policies in general, we show that better bounds hold for many commonly used uncertainty sets.

While our main contribution is to extend approximation bounds to multi-stage ARO, Problem (1) is further less restrictive than problems previously discussed in the literature on approximation bounds. In addition to being restricted to two-stage ARO, previous work often assumed $\boldsymbol{c}$ and $\boldsymbol{x}$ to be non-negative [14, 15, 18]. Ben-Tal et al. [14] additionally restricted parts of $\boldsymbol{A}$, i.e., they require the parts of $\boldsymbol{A}$ associated with the first stage decision to be non-negative. Our policies do not need this assumption. However, we show that Problem (1) is unbounded whenever there is a feasible $\boldsymbol{x}$ with $\boldsymbol{c}^\intercal \boldsymbol{x} < 0$, due to the non-negativity of the right-hand side. As a consequence, our policies do not readily extend to general maximization problems.

From a theoretical perspective, mainly asymptotic bounds are of interest. In practice, however, also the exact factors of the approximation are important. Throughout the paper, we thus always give the asymptotic, as well as the exact bounds. We compare all our bounds to the previously best-known bounds for the two-stage setting given in

Ben-Tal et al. [14] and show that our constructions yield both constant factor, as well as asymptotic improvements.

Comparison with affine policies: Using the newly found bounds, we show that no approximation bound for affine policies on hypersphere uncertainty exists that is better than the bound we show for our policies. For budgeted uncertainty, on the other hand, we show that affine policies strictly dominate our piecewise affine policies. These findings confirm results that have been reported for the two-stage variant, where affine policies do not perform well for hypersphere uncertainty [18], but very well for budgeted uncertainty [40].

Improvement heuristic: Due to inherent properties of our policy construction, resulting solutions are overly pessimistic on instances where the impact on the objective varies significantly between different uncertainty dimensions. To diminish this effect, we introduce an improvement heuristic that performs at least as well as our policies and that can be integrated into the LP used to construct our policies. While these modifications come at the cost of higher solution times, they allow for significant objective improvements on some instance classes.

Tightening piecewise affine policies via lifting: We show that in the context of Problem (1) the piecewise affine policies via lifting presented by Georghiou et al. [31] yield equivalent solutions to affine policies. To prevent this from happening, we construct tightened piecewise affine policies via lifting using insights from our piecewise affine policies. These new policies integrate the approximative power of affine policies, and our piecewise affine policies and are guaranteed to perform at least as well as the individual policies they combine.

Numerical evidence: Finally, we present two sets of numerical experiments showing that our policies solve by orders of magnitude faster than the affine adjustable policies presented by Ben-Tal et al. [9], the piecewise affine policies via lifting presented by Georghiou et al. [31], and the near-optimal piecewise affine policies by Bertsimas and Georghiou [17], while often yielding comparable or improving results. First, we study a slightly modified version of the tests presented in Ben-Tal et al. [14], allowing us to demonstrate our policies' scalability and the impact of our improvement heuristic. Second, we focus on demand covering instances to demonstrate good performances of our policies for a problem that resembles a practical application. We refer to our git repository (https://github.com/tumBAIS/piecewise-affine-ARO) for all material necessary to reproduce the numerical results outlined in this paper.

Comparison against closely related work: Compared to the closely related work by Ben-Tal et al. [14], who first introduced the concept of domination in the context of ARO, our contributions are multifold. First, we extend domination-based piecewise affine policies to a wider class of problems by switching from a two-stage to a multi-stage setting and relaxing assumptions. We discuss the structural reasons that make this extension non-trivial at the beginning of Sect. 2. In addition to showing stronger approximation guarantees for our policies, we conduct comprehensive theoretical and numerical comparisons with other adaptable policies. Based on these comparisons, we construct two new policies that mitigate weaknesses of domination and integrate

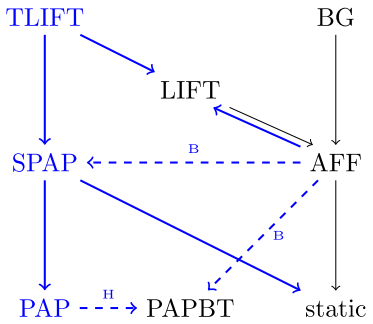


Fig. 2 Relations between multi-stage ARO policies compared in this paper. An arc from a policy P to another policy P' states $Z_P \leq Z_{P'}$, where Z_P and $Z_{P'}$ are optimal objective values for the ARO Problem (1) solved with policy P, P' respectively. Dashed arcs only hold for hypersphere (H) or budgeted (B) uncertainty. Relations proved for the first time in this paper are highlighted (blue, bold). The compared policies are: static policies (static); piecewise affine policies via domination by Ben-Tal et al. [14] (PAPBT); our piecewise affine policies via domination (PAP), c.f., Sects. 2 and 3; affine policies [9] (AFF); our piecewise affine policies with rescaling (SPAP), c.f., Sect. 4; near-optimal piecewise affine policies [17] (BG); piecewise affine policies via lifting [31] (LIFT); our tightened piecewise affine policies via lifting (TLIFT), c.f., Sect. 5

its strength with the strength of other policies. More specifically, the first policy integrates finding a good outer approximation of the uncertainty set in the optimization process. The second policy integrates structural results from domination into lifting policies, c.f., [31]. As a result, we get a hierarchy of piecewise affine adjustable policies with provable relative performance guarantees. We give an overview of all policies constructed in our work and their relative performance guarantees compared to other policies in Fig. 2.

The rest of this paper is structured as follows. In Sect. 2, we introduce our policies and elaborate on their construction. In Sect. 3, we present our approximation bounds for the multi-stage ARO Problem (1). We present an improvement heuristic for our policies in Sect. 4. By using the results of Sects. 2 and 3, we construct tightened piecewise affine policies via lifting in Sect. 5. Finally, we provide numerical evidence for the performance of our policy compared to other state of the art policies in Sect. 6. Section 7 concludes this paper with a brief reflection of our work and avenues for future research. To keep the paper concise, we defer proofs that could possibly interrupt the reading flow to "Appendices B–O".

2 Framework for piecewise affine multi-stage policies

In this section, we present our piecewise affine framework for the multi-stage ARO Problem (1). The main rationale of our framework is to construct new uncertainty sets $\hat{\mathcal{U}}$ that dominate the original uncertainty sets $\mathcal{U}$. With this, our framework follows a similar rationale as the two-stage framework from Ben-Tal et al. [14]. For a problem $Z_{AR}(\mathcal{U})$ we construct $\hat{\mathcal{U}}$ in such a way that $Z_{AR}(\hat{\mathcal{U}})$ can be efficiently solved, and a solution of $Z_{AR}(\hat{\mathcal{U}})$ can be used to generate solutions for $Z_{AR}(\mathcal{U})$.

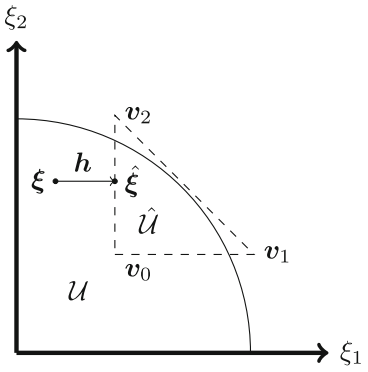


Fig. 3 Two dimensional hypersphere uncertainty set $\mathcal{U}$ with (dashed) dominating set $\hat{\mathcal{U}}$ (2) induced by the convex combination of vertices $\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{v}_2$ and dominating function $\boldsymbol{h}$ (4) that maps a point $\boldsymbol{\xi} \in \mathcal{U}$ to a point $\hat{\boldsymbol{\xi}} \in \hat{\mathcal{U}}$

In this context, we note that one cannot straightforwardly apply the construction scheme used by Ben-Tal et al. [14] due to nonanticipativity requirements. More specifically, Ben-Tal et al. [14] construct $\hat{\mathcal{U}}$ as polytopes, where it is well known that worst-case solutions always occur on extreme points, as any solution can be represented by convex combinations of extreme point solutions. The construction of these convex combinations, however, is not guaranteed to be nonanticipative in the multi-stage setting. To overcome this challenge, we incorporate nonanticipativity in the concept of uncertainty set domination and extend it to a multi-stage setting.

Definition 1 (*Domination*) Given an uncertainty set $\mathcal{U} \subseteq \mathbb{R}^m_+$, we say that $\mathcal{U}$ is *dominated* by $\hat{\mathcal{U}} \subseteq \mathbb{R}^m_+$ if there is a *domination function* $\boldsymbol{h} \colon \mathcal{U} \to \hat{\mathcal{U}}$ with $\boldsymbol{h}(\boldsymbol{\xi}) \geq \boldsymbol{\xi}$, and $\boldsymbol{h}$ can be expressed as $\boldsymbol{h}(\boldsymbol{\xi}) = \left(\boldsymbol{h}^1(\underline{\boldsymbol{\xi}}^1), \ldots, \boldsymbol{h}^T(\underline{\boldsymbol{\xi}}^T)\right)$ where $\boldsymbol{h}^t$ maps to the uncertainties in stage t and depends on uncertainties up to that stage.

Intuitively, an uncertainty set $\hat{\mathcal{U}}$ dominates another set $\mathcal{U}$ if for every point $\boldsymbol{\xi} \in \mathcal{U}$ there is a point $\hat{\boldsymbol{\xi}} \in \hat{\mathcal{U}}$ that is at least as large in each component, i.e., $\hat{\boldsymbol{\xi}} \geq \boldsymbol{\xi}$. Later we show that the dominating set $\hat{\mathcal{U}}$ can be constructed as the convex combination of $m+1$ vertices $\boldsymbol{v}_0, \ldots, \boldsymbol{v}_m$. We also show how to construct dominating functions $\boldsymbol{h}$ for these vertex induced dominating sets. Figure 3 illustrates the hypersphere uncertainty set $\mathcal{U} = \left\{\boldsymbol{\xi} \in \mathbb{R}^m_+ \middle| \|\boldsymbol{\xi}\|_2^2 \leq 1\right\}$ together with our dominating set $\hat{\mathcal{U}}$ (2) and the dominating function $\boldsymbol{h}$ (4) for $m = 2$.

Due to the non-negativity of the problem's right-hand side, domination at most restricts the set of feasible solutions. As a consequence, each feasible solution for a realization $\hat{\boldsymbol{\xi}} \in \hat{\mathcal{U}}$ is also a feasible solution for all realizations $\boldsymbol{\xi} \in \mathcal{U}$ that are dominated by $\hat{\boldsymbol{\xi}}$. Using this property, we can derive piecewise affine policies for $Z_{AR}(\mathcal{U})$ from solutions of $Z_{AR}(\hat{\mathcal{U}})$. Since $\mathcal{U}$ is full-dimensional and down-monotone by Assumption 1, there always exists a factor $\beta \geq 0$ such that scaling $\mathcal{U}$ by β contains $\hat{\mathcal{U}}$. Theorem 1 shows that with this factor β, solutions of problem $Z_{AR}(\hat{\mathcal{U}})$ are β-approximations for problem $Z_{AR}(\mathcal{U})$. It also shows that $Z_{AR}(\hat{\mathcal{U}})$ is unbounded exactly when $Z_{AR}(\mathcal{U})$ is unbounded. Thus, we assume for the remainder of this paper w.l.o.g. that both $Z_{AR}(\mathcal{U})$ and $Z_{AR}(\hat{\mathcal{U}})$ are bounded.

Theorem 1 *Consider an uncertainty set $\mathcal{U}$ from Problem (1) and a dominating set $\hat{\mathcal{U}}$. Let $\beta \geq 1$ be such that $\forall \hat{\boldsymbol{\xi}} \in \hat{\mathcal{U}}\colon \frac{1}{\beta}\hat{\boldsymbol{\xi}} \in \mathcal{U}$. Moreover, let $Z_{AR}(\mathcal{U})$ and $Z_{AR}(\hat{\mathcal{U}})$ be optimal values of Problem (1). Then, either $Z_{AR}(\mathcal{U})$ and $Z_{AR}(\hat{\mathcal{U}})$ are unbounded or*

$$0 \leq Z_{AR}(\mathcal{U}) \leq Z_{AR}(\hat{\mathcal{U}}) \leq \beta \cdot Z_{AR}(\mathcal{U}).$$

We present the proof for Theorem 1 in "Appendix B".

In the remainder of this section, we demonstrate how the results of Theorem 1 can be used to efficiently construct β-approximations for $Z_{AR}(\mathcal{U})$. Therefore, we show in Sect. 2.1 how to construct dominating polytopes $\hat{\mathcal{U}}$ and efficiently find solutions $Z_{AR}(\hat{\mathcal{U}})$ that comply with nonanticipativity requirements. Then, we construct the dominating function $\boldsymbol{h}\colon \mathcal{U} \to \hat{\mathcal{U}}$, which allows us to extend these solutions to solutions for $Z_{AR}(\mathcal{U})$ in Sect. 2.2.

2.1 Construction of the dominating set

In the following, we construct a dominating set in the form of a polytope for which the worst-case solution can be efficiently found by solving a linear program on its vertices. Specifically, for an uncertainty set $\mathcal{U}$, we consider dominating sets $\hat{\mathcal{U}}$ of the form

$$\hat{\mathcal{U}} := \operatorname{conv}(\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) \tag{2}$$

where for all $i \in \{0, \ldots, m\}$: $\frac{1}{\beta}\boldsymbol{v}_i \in \mathcal{U}$ and for all $i \in \{1, \ldots, m\}$: $\boldsymbol{v}_i = \boldsymbol{v}_0 + \rho_i \boldsymbol{e}_i$ for some $\rho_i \in \mathbb{R}_+$. We postpone the construction of the domination function $\boldsymbol{h}$, the base vertex $\boldsymbol{v}_0$, and parameters $\rho_1, \ldots, \rho_m$ to Sect. 2.2 and first focus on the construction of solutions for $Z_{AR}(\hat{\mathcal{U}})$. Here, we extend the notation on $\boldsymbol{x}$ and $\boldsymbol{\xi}$ introduced in Sect. 1.1 to $\boldsymbol{x}_i$ and $\boldsymbol{v}_i$. Consequently, $\boldsymbol{x}_i^t$ is the sub-vector of $\boldsymbol{x}_i$ corresponding to decisions made in stage t, and $\underline{\boldsymbol{v}_i}^t$ is the sub-vector of $\boldsymbol{v}_i$ corresponding to uncertainties up to stage t. Then, the key component for our construction is LP (3)

$$Z_{LP}(\hat{\mathcal{U}}) = \min_{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m} z \tag{3a}$$

$$\text{s.t.} \quad z \geq \boldsymbol{c}^\mathsf{T}\boldsymbol{x}_i \qquad \forall i \in \{0, \ldots, m\} \tag{3b}$$

$$\boldsymbol{A}\boldsymbol{x}_i \geq \boldsymbol{D}\boldsymbol{v}_i + \boldsymbol{d} \qquad \forall i \in \{0, \ldots, m\} \tag{3c}$$

$$\boldsymbol{x}_i^t = \boldsymbol{x}_j^t \qquad \forall i, j \in \{0, \ldots, m\}, t \in \{1, \ldots, T\}, \underline{\boldsymbol{v}_i}^t = \underline{\boldsymbol{v}_j}^t. \tag{3d}$$

Intuitively, the Objective (3a) together with Constraints (3b) minimize the maximal cost over all vertex solutions $\boldsymbol{x}_i$. Constraints (3c) ensure that each $\boldsymbol{x}_i$ is a feasible solution for the respective uncertainty vertex $\boldsymbol{v}_i$ of $\hat{\mathcal{U}}$. Finally, Constraints (3d) ensure nonanticipativity by forcing vertex solutions to be equal unless different uncertainties

were observed. With these constraints, we construct LP (3) such that it is sufficient to find an optimal solution for $Z_{LP}(\hat{\mathcal{U}})$ in order to find an optimal solution for $Z_{AR}(\hat{\mathcal{U}})$.

Lemma 2 *Let $\hat{\mathcal{U}}$ be a dominating set as described in (2), $Z_{LP}(\hat{\mathcal{U}})$ be the solution of LP (3), and $Z_{AR}(\hat{\mathcal{U}})$ be the solution of Problem (1). Then the LP solution $(\boldsymbol{x}_i)$ on the vertices of $\hat{\mathcal{U}}$ can be extended to a solution on the full set $\hat{\mathcal{U}}$ and we find:*

$$Z_{LP}(\hat{\mathcal{U}}) = Z_{AR}(\hat{\mathcal{U}}).$$

We present the proof for Lemma 2 in "Appendix C".

2.2 Construction of the domination function

In the previous section, we showed how to construct dominating sets $\hat{\mathcal{U}}$ such that $Z_{AR}(\hat{\mathcal{U}})$ can be solved efficiently. In order for $\hat{\mathcal{U}}$ to be a valid dominating set for some uncertainty set $\mathcal{U}$, we additionally have to construct a nonanticipative dominating function $\boldsymbol{h}\colon \mathcal{U} \to \hat{\mathcal{U}}$ according to Definition 1. Specifically, we use

$$\boldsymbol{h}(\boldsymbol{\xi}) := (\boldsymbol{\xi} - \boldsymbol{v}_0)_+ + \boldsymbol{v}_0 \tag{4}$$

where $(\cdot)_+$ is the element-wise maximum with 0 and $\boldsymbol{v}_0$ is the base vertex from Definition (2). By construction, $\boldsymbol{h}$ maps each uncertainty realization $\boldsymbol{\xi}$ to its element-wise maximum with $\boldsymbol{v}_0$. It directly follows that $\boldsymbol{h}$ is nonanticipative, as each element in $\boldsymbol{h}(\boldsymbol{\xi})$ solely depends on the corresponding element in $\boldsymbol{\xi}$.

Finally, we have to ensure that $\boldsymbol{h}(\boldsymbol{\xi}) \in \hat{\mathcal{U}}$ for all $\boldsymbol{\xi} \in \mathcal{U}$. We do so by choosing the base vertex $\boldsymbol{v}_0$ and parameters $\rho_1, \ldots, \rho_m$ during the construction of $\hat{\mathcal{U}}$ appropriately. Using $\lambda_i(\boldsymbol{\xi}) := \frac{((\boldsymbol{\xi}-\boldsymbol{v}_0)_+)_i}{\rho_i}$ with the convention $\frac{0}{0} = 0$, we find

$$\boldsymbol{h}(\boldsymbol{\xi}) = \sum_{i=1}^{m} \lambda_i(\boldsymbol{\xi})\boldsymbol{v}_i + \left(1 - \sum_{i=1}^{m} \lambda_i(\boldsymbol{\xi})\right)\boldsymbol{v}_0.$$

By definition, any convex combination of $\boldsymbol{v}_i$ is contained in $\hat{\mathcal{U}}$. Thus, $\boldsymbol{h}$ is a valid domination function if and only if

$$\max_{\boldsymbol{\xi}\in\mathcal{U}} \sum_{i=1}^{m} \lambda_i(\boldsymbol{\xi}) \leq 1. \tag{5}$$

Condition (5) gives a compact criterion to check the validity of dominating sets. By doing so, it lays the basis for our optimal selection of the base vertex $\boldsymbol{v}_0$ and parameters $\rho_0, \ldots, \rho_m$. Checking Condition (5) generally requires solving a convex optimization problem. However, in Sect. 3 we show that for many commonly used special uncertainty sets, this problem can be significantly simplified, leading to low dimensional unconstrained minimization problems or even analytical solutions.

Recall that we showed how to extend a solution $(\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m)$ of LP (3) to the full set $\hat{\mathcal{U}}$ in the proof of Lemma 2. Combining this with $\boldsymbol{h}$ and using Theorem 1 we get a piecewise affine solution for $Z_{AR}(\mathcal{U})$ by

$$\boldsymbol{x}(\boldsymbol{\xi}) = \sum_{i=1}^{m} \lambda_i(\boldsymbol{\xi}) \boldsymbol{x}_i + \left(1 - \sum_{i=1}^{m} \lambda_i(\boldsymbol{\xi})\right) \boldsymbol{x}_0 \tag{6}$$

that has an optimality bound of β.

2.3 Limitations

While our policies overcome the nontrivial challenge of nonanticipativity on extreme point solutions, they still rely on the ability to form convex combinations. As integrality is not preserved by convex combinations, there is no natural way to extend our approach to integer or binary recourse decisions $\boldsymbol{x}$. However, including non-adjustable integer or binary first-stage decisions in our framework is straightforward. Also, it is not straightforwardly possible to incorporate uncertain recourse decisions, i.e., dependence of $\boldsymbol{A}$ on $\boldsymbol{\xi}$, into our approach, as worst-case realizations for problems with uncertain recourse are not necessarily extreme points of $\mathcal{U}$, see, e.g., [3, 33].

3 Optimality bounds for different uncertainty sets

In the previous section, we demonstrated how to construct nonanticipative piecewise affine policies for the multi-stage Problem (1). On this basis, proving approximation bounds mainly depends on geometric properties of the uncertainty sets $\mathcal{U}$. We first show approximation bounds for some commonly used permutation invariant uncertainty sets. On these sets, the dominating sets are permutation invariant and we give closed-form constructions. We then give approximation bounds for our piecewise affine policies on general uncertainty sets. Finally, we demonstrate how the bounds of an uncertainty set $\mathcal{U}$ generalize to transformations of that set. While in theory, mostly asymptotic bounds are of interest, in practice constant factors are important as well. Thus we always state exact, as well as asymptotic bounds. Table 1 gives an overview of all bounds that are explicitly proven in Propositions 5, 7, 9, 10 and 11 of this section. We compare all our results against the results for the two-stage setting in Ben-Tal et al. [14] and show constant factor, as well as asymptotic improvements. For budgeted and hypersphere uncertainty sets, we further compare the theoretical performance of our piecewise affine policies with affine adjustable policies.

For permutation invariant uncertainty sets there exists an optimal choice of $\hat{\mathcal{U}}$ that is also permutation invariant. More specifically, $\boldsymbol{v}_i$ simplifies to $\boldsymbol{v}_0 = \mu \boldsymbol{e}$ and $\boldsymbol{v}_i = \boldsymbol{v}_0 + \rho \boldsymbol{e}_i$, for some μ, ρ.

Lemma 3 *Let $\mathcal{U}$ be a permutation invariant uncertainty set. Then there exist μ, ρ such that for the dominating uncertainty set $\hat{\mathcal{U}}$ spanned by $\boldsymbol{v}_0 = \mu \boldsymbol{e}$ and $\boldsymbol{v}_i = \boldsymbol{v}_0 + \rho \boldsymbol{e}_i$ for $i \in \{1, \ldots, m\}$ there is no other dominating set $\hat{\mathcal{U}}'$ constructed as in (2) with a smaller approximation factor β.*

Table 1 Performance bounds of the piecewise affine policy for different uncertainty sets

No.	Uncertainty set $\mathcal{U}$	Bound	Asymptotic bound
I	$\left\{\boldsymbol{\xi} \in \mathbb{R}_+^m \middle\| \|\boldsymbol{\xi}\|_2^2 \leq 1\right\}$	$\sqrt[4]{m}\frac{\sqrt{m-1}}{\sqrt{2(m-\sqrt{m})}}$	$O(\sqrt[4]{m})$
II	$\left\{\boldsymbol{\xi} \in [0,1]^m \middle\| \|\boldsymbol{\xi}\|_1 \leq k\right\}$	$\frac{k(m-1)}{m+k(k-2)}$	$O\left(\min\left\{k, \frac{m}{k}\right\}\right)$
III	$\left\{\boldsymbol{\xi} \in \mathbb{R}_+^m \middle\| \|\boldsymbol{\xi}\|_p \leq 1\right\}$	$\left(2(m-1)+2^p\right)^{\frac{p-1}{p^2}} p^{\frac{1-p}{p}} (p-1)^{\left(\frac{p-1}{p}\right)^2}$	$O\left(m^{\frac{p-1}{p^2}}\right)$
IV	$\left\{\boldsymbol{\xi} \in \mathbb{R}_+^m \middle\| \boldsymbol{\xi}^\intercal \boldsymbol{\Sigma} \boldsymbol{\xi} \leq 1\right\}$	$\begin{cases} \sqrt{\frac{1}{2} + \frac{am+\sqrt{(1-a)m+am^2}}{2(1-a)}} & a \leq m^{-\frac{2}{3}} \\ \frac{1}{\sqrt{a}} & a > m^{-\frac{2}{3}} \end{cases}$	$O\left(m^{\frac{1}{3}}\right)$
V	$\mathcal{U} \subset \mathbb{R}_+^m$	$2\sqrt{m}+1$	$O\left(\sqrt{m}\right)$

We prove specific bounds for uncertainty sets of the forms I) hypersphere uncertainty; II) budgeted uncertainty; III) p-norm ball uncertainty, with $p \geq 1$; IV) ellipsoid uncertainty, with $\boldsymbol{\Sigma} := (1-a)\mathbf{1} + a\boldsymbol{J}$ where $\mathbf{1}$ is the unity matrix and $\boldsymbol{J}$ the matrix of all ones; V) general uncertainty sets

We present the proof for Lemma 3 in "Appendix D".

With these simplifications, Condition (5) becomes

$$\frac{1}{\rho} \max_{\boldsymbol{\xi} \in \mathcal{U}} \sum_{i=1}^{m} (\xi_i - \mu)_+ \leq 1. \tag{7}$$

With the permutation invariance of the problem, Ben-Tal et al. [14] show that for any μ there exists a $j \leq m$, such that the maximization problem in (7) has a solution that is constant on the first j components and zero on all others components.

Lemma 4 (Lemma 4 in Ben-Tal et al. [14]) *Let $\gamma(j)$ be the maximal average value of the first j components of any $\boldsymbol{\xi} \in \mathcal{U}$*

$$\gamma(j) := \frac{1}{j} \max_{\boldsymbol{\xi} \in \mathcal{U}} \sum_{i=1}^{j} \xi_i.$$

Then for each μ there exists an optimal solution $\boldsymbol{\xi}^*$ for the maximization problem in Eq. (7) that has the form

$$\boldsymbol{\xi}^* = \sum_{i=1}^{j} \gamma(j) \boldsymbol{e}_i,$$

for some $j \leq m$.

Hypersphere uncertainty: We first use Lemma 4 to find a new dominating set for hypersphere uncertainty. By doing so, we find a new approximation bound that improves the bound of $\sqrt[4]{m}$ provided in Ben-Tal et al. [14] by a factor of $\sqrt{\frac{\sqrt{m}+1}{2\sqrt{m}}}$,

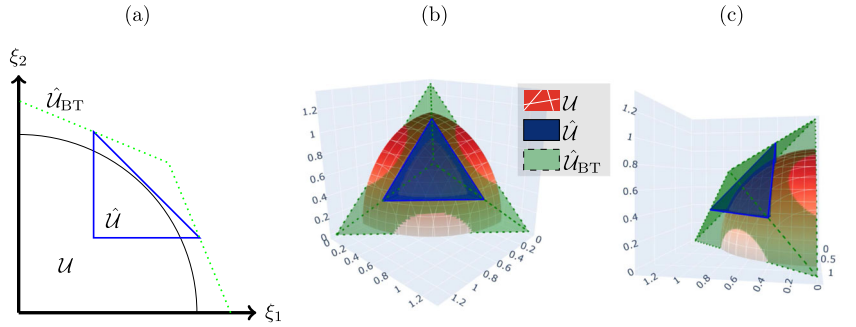


Fig. 4 Comparison of our dominating set $\hat{\mathcal{U}}$ (blue, solid frame) and the dominating set $\hat{\mathcal{U}}_{\mathrm{BT}}$ proposed in Ben-Tal et al. [14] (green, dashed frame) for the hypersphere uncertainty set $\mathcal{U}$ in $m = 2$ (**a**) and $m = 3$ (**b**), (**c**) uncertainty dimensions (color figure online)

which for large m converges towards $\frac{1}{\sqrt{2}}$. While this improvement is irrelevant for the asymptotic complexity of the problem, the new formulation of $\hat{\mathcal{U}}$ does make a difference in practice. In Fig. 4 we illustrate the improvement of our dominating set $\hat{\mathcal{U}}$ over the dominating set $\hat{\mathcal{U}}_{\mathrm{BT}}$ proposed in Ben-Tal et al. [14] for hypersphere uncertainty sets in two and three uncertainty dimensions. Note, that our sets $\hat{\mathcal{U}}$ are fully contained in the sets $\hat{\mathcal{U}}_{\mathrm{BT}}$, and all extreme points of $\hat{\mathcal{U}}_{\mathrm{BT}}$ are located outside of $\hat{\mathcal{U}}$. This implies $Z_{AR}(\hat{\mathcal{U}}) \leq Z_{AR}(\hat{\mathcal{U}}_{\mathrm{BT}})$ for hypersphere uncertainty. The formal proof follows from straightforward convex containment and is left for brevity.

Proposition 5 (Hypersphere) *Consider the hypersphere uncertainty set* $\mathcal{U} = \left\{\boldsymbol{\xi} \in \mathbb{R}^m_+ \middle| \|\boldsymbol{\xi}\|_2^2 \leq 1\right\}$. *Then a solution for* $Z_{AR}(\hat{\mathcal{U}})$ *where* $\hat{\mathcal{U}}$ *is constructed using Criterion* (7) *with*

$$\mu = \frac{1}{2\sqrt[4]{m}}, \qquad \rho = \frac{\sqrt[4]{m}}{2},$$

gives a $\beta = \sqrt{\frac{\sqrt{m}+1}{2}}$ *approximation for problem* $Z_{AR}(\mathcal{U})$.

We present the proof for Proposition 5 in "Appendix E".

We can also use this improved performance bound to show that affine adjustable policies cannot yield better bounds than piecewise affine adjustable policies for $m \geq 153$. This is because there are instances of Problem (1) with hypersphere uncertainty where affine adjustable policies perform at least $\frac{4}{5}\left(\sqrt[4]{m} - \frac{1}{\sqrt[4]{m}}\right)$ worse than an optimal policy. We formalize these results in Proposition 6. Note, that these better bounds do not imply that piecewise affine adjustable policies always yield better results than affine adjustable policies for hypersphere uncertainty.

Proposition 6 *Affine adjustable policies cannot achieve better performance bounds than* $\frac{4}{5}\left(\sqrt[4]{m} - \frac{1}{\sqrt[4]{m}}\right)$ *for Problem* (1) *with hypersphere uncertainty, even for* $\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{A}$ *being non negative and* $\boldsymbol{A}$ *being a 0,1-matrix.*

We present the proof for Proposition 6 in "Appendix F".

Budgeted uncertainty: Next, we tighten the bounds for budgeted uncertainty sets. Proposition 7 shows that our new bound is given by $\beta = \frac{k(m-1)}{m+k(k-2)}$. Using

$$\frac{\beta}{k} = \frac{k(m-1)}{k(m+k(k-2))} = \frac{m-1}{m-1+k^2-2k+1} = \frac{m-1}{m-1+(k-1)^2} \leq 1$$

and

$$\frac{\beta}{\frac{m}{k}} = \frac{k^2(m-1)}{m(m+k(k-2))} = \frac{k^2(m-1)}{k^2(m-1)+k^2+m^2-2km}$$
$$= \frac{k^2(m-1)}{k^2(m-1)+(m-k)^2} \leq 1,$$

we show $\beta \leq \min(k, \frac{m}{k})$, which matches the bound for the two-stage problem variant in Ben-Tal et al. [14]. As $\frac{\beta}{k}$ is decreasing in k and $\frac{\beta}{\frac{m}{k}}$ is increasing in k, we obtain a maximum improvement for $k = \frac{m}{k} \Leftrightarrow k = \sqrt{m}$. At this point the improvement of the bound reaches a factor of $\frac{1}{2}$.

Proposition 7 (Budget) *Consider the budgeted uncertainty set* $\mathcal{U} = \left\{\boldsymbol{\xi} \in [0,1]^m \middle| \|\boldsymbol{\xi}\|_1 \leq k\right\}$ *for some* $k \in \{1, \ldots, m\}$. *Then a solution for* $Z_{AR}(\hat{\mathcal{U}})$ *where* $\hat{\mathcal{U}}$ *is constructed using Criterion* (7) *with*

$$\mu = \frac{k(k-1)}{m+k(k-2)}, \qquad \rho = \frac{k(m-k)}{m+k(k-2)},$$

gives a $\beta = \frac{k(m-1)}{m+k(k-2)}$ *approximation for problem* $Z_{AR}(\mathcal{U})$.

We present the proof for Proposition 7 in "Appendix G".

Note, that there is no result analogous to Proposition 6 for budgeted uncertainty as Housni and Goyal [40] showed that affine policies are in $O\left(\frac{\log(m)}{\log\log(m)}\right)$ for two-stage problems with non-negative $\boldsymbol{c}, \boldsymbol{x}, \boldsymbol{A}$. Furthermore, our piecewise affine policies are strictly dominated by affine policies for integer budgeted uncertainty.

Proposition 8 *Consider Problem* (1) *with budgeted uncertainty and an integer budget. Let* Z_{PAP} *be the optimal value found by our piecewise affine policy and* Z_{AFF} *be the optimal value found by an affine policy. Then*

$$Z_{AFF} \leq Z_{PAP}.$$

We present the proof for Proposition 8 in "Appendix H".

Norm ball uncertainty: In a similar manner as before, we construct new dominating sets for p-norm ball uncertainty and tighten the bound in Ben-Tal et al. [14] by a factor of $2^{-1+\frac{1}{p}-\frac{1}{p^2}} p^{\frac{1}{p}} (p-1)^{\frac{1}{p^2}-\frac{1}{p}}$ for sufficiently large m. This factor is always smaller than one and converges to $\frac{1}{2}$ for large p.

Proposition 9 (p-norm ball) *Consider the p-norm ball uncertainty set* $\mathcal{U} = \left\{\boldsymbol{\xi} \in \mathbb{R}_+^m \middle| \|\boldsymbol{\xi}\|_p \leq 1\right\}$ *with* $p > 1$. *Then a solution for* $Z_{AR}(\hat{\mathcal{U}})$ *where* $\hat{\mathcal{U}}$ *is constructed using Criterion* (7) *with*

$$\mu = 2^{\frac{1}{p}} \left(2(m-1) + 2^p\right)^{-\frac{1}{p^2}} p^{-1} (p-1)^{\frac{1}{p} + \left(1-\frac{1}{p}\right)^2},$$
$$\rho = 2^{\frac{1}{p}-1} \left(2(m-1) + 2^p\right)^{\frac{1}{p}-\frac{1}{p^2}} p^{-1} (p-1)^{\left(1-\frac{1}{p}\right)^2}$$

gives a

$$\begin{aligned}\beta &\leq \left(2(m-1) + 2^p\right)^{\frac{1}{p}-\frac{1}{p^2}} p^{\frac{1}{p}-1} (p-1)^{\left(1-\frac{1}{p}\right)^2} \\ &\leq \left((2m)^{\frac{1}{p}-\frac{1}{p^2}} + 2^{1-\frac{1}{p}}\right) p^{\frac{1}{p}-1} (p-1)^{\left(1-\frac{1}{p}\right)^2} = O\left(m^{\frac{1}{p}-\frac{1}{p^2}}\right)\end{aligned}$$

approximation for problem $Z_{AR}(\mathcal{U})$.

We present the proof for Proposition 9 in "Appendix I".

Ellipsoid uncertainty: For the permutation invariant ellipsoid uncertainty set $\left\{\boldsymbol{\xi} \in \mathbb{R}_+^m \middle| \boldsymbol{\xi}^\intercal \boldsymbol{\Sigma} \boldsymbol{\xi} \leq 1\right\}$ with $m > 1$, $\boldsymbol{\Sigma} := \mathbf{1} + a(\boldsymbol{J} - \mathbf{1})$, $a \in [0, 1]$, $\mathbf{1}$ being the unity matrix, and $\boldsymbol{J}$ being the matrix of all ones, we construct dominating sets via a case distinction on the size of a. While for large a already a scaled simplex gives a good approximation, we construct the dominating set for small a more carefully. By doing so, we improve the previously best known asymptotic bound for the two-stage problem variant of $O(m^{\frac{2}{5}})$ [14] to $O(m^{\frac{1}{3}})$. Note, that for $a = 0$ our bounds converge to the bounds of hypersphere uncertainty in Proposition 5 and for $a = 1$ towards an exact representation.

Proposition 10 (Ellipsoid) *Consider the ellipsoid uncertainty set* $\mathcal{U} = \left\{\boldsymbol{\xi} \in \mathbb{R}_+^m \middle| \boldsymbol{\xi}^\intercal \boldsymbol{\Sigma} \boldsymbol{\xi} \leq 1\right\}$ *with* $m > 1$ *and* $\boldsymbol{\Sigma} := \mathbf{1} + a(\boldsymbol{J} - \mathbf{1})$ *for* $a \in [0, 1]$. *Here* $\mathbf{1}$ *is the unity matrix and* $\boldsymbol{J}$ *is the matrix of all ones. Then a solution for* $Z_{AR}(\hat{\mathcal{U}})$ *where* $\hat{\mathcal{U}}$ *is constructed using Criterion* (7) *with*

$$\mu = \frac{1}{2\sqrt[4]{(1-a)^3 m + (1-a)^2 a m^2}}, \qquad \rho = \frac{1}{4(1-a)\mu} \qquad \text{if } a \leq m^{-\frac{2}{3}},$$
$$\mu = 0, \qquad \rho = \frac{1}{\sqrt{a}} \qquad \text{if } a > m^{-\frac{2}{3}},$$

gives a

$$\beta = \begin{cases} \sqrt{\frac{1}{2}\left(1 + \frac{1}{1-a}\left(am + \sqrt{(1-a)m + am^2}\right)\right)} & \text{if } a \leq m^{-\frac{2}{3}} \\ \frac{1}{\sqrt{a}} & \text{if } a > m^{-\frac{2}{3}} \end{cases} = O(m^{\frac{1}{3}})$$

approximation for problem $Z_{AR}(\mathcal{U})$.

We present the proof for Proposition 10 in "Appendix J".

General uncertainty sets: After having shown specific bounds for some commonly used permutation invariant uncertainty sets, we now give a general bound that holds for all uncertainty sets that fulfill the assumptions of Problem (1). We show that any uncertainty set can be dominated within an approximation factor of $\beta = 2\sqrt{m} + 1$, which improves the bound in Ben-Tal et al. [14] by a factor of $\frac{1}{2}$. As shown in Ben-Tal et al. [14] this approximation bounds is asymptotically tight, when using pure domination techniques. More precisely, for any polynomial number of vertices the budgeted uncertainty set with $k = \sqrt{m}$ cannot be dominated with some β better than $O(\sqrt{m})$.

Proposition 11 (General uncertainty) *Consider any uncertainty set* $\mathcal{U} \subseteq [0, 1]^m$ *that is convex, full-dimensional with* $\boldsymbol{e}_i \in \mathcal{U}$ *for all* $i \in \{1, \ldots, m\}$ *and down-monotone. Then, there always exists a dominating uncertainty set* $\hat{\mathcal{U}}$ *of the form in* (2) *that dominates* $\mathcal{U}$ *by at most a factor of* $\beta = 2\sqrt{m} + 1$.

We present the proof for Proposition 11 in "Appendix K".

Stagewise uncertainty sets: In general, our policies do not require stagewise uncertainty. However, the existence of such a structure can be utilized in the construction of dominating uncertainty sets, leading to approximation bounds that depend linearly on the stagewise approximation bounds.

Proposition 12 (Stagewise) *Let* $\mathcal{U} = \mathcal{U}_1 \times \cdots \times \mathcal{U}_T$ *be a stagewise independent uncertainty set and for each* $\mathcal{U}_t$, *let* $\hat{\mathcal{U}}_t$ *be a dominating set constructed as in* (2). *Let* β_t *be the approximation factor for* $\hat{\mathcal{U}}_t$, *and let* $\beta'_t = \min\{\beta' : \frac{1}{\beta'}\boldsymbol{e} \in \mathcal{U}_t\}$ *be the constant approximation factor for set* $\mathcal{U}_t$. *Then for any partition* $\mathcal{T}_1 \cup \mathcal{T}_2 = \{1, \ldots, T\}$, $\mathcal{T}_1 \cap \mathcal{T}_2 = \emptyset$ *of the stages, there exists a dominating set* $\hat{\mathcal{U}}$ *for* $\mathcal{U}$ *with approximation factor*

$$\beta \leq \max\left(\sum_{t\in\mathcal{T}_1} \beta_t,\ \max_{t\in\mathcal{T}_2} \beta'_t\right).$$

We present the proof for Proposition 12 in "Appendix L".

Transformed uncertainty sets: We note that by the right-hand side $\boldsymbol{D}\boldsymbol{\xi} + \boldsymbol{d}$ of Problem (1) any positive affine transformation of uncertainty sets $\mathcal{U}$ can be dominated by the same affine transformation of the dominating set $\hat{\mathcal{U}}$. As the approximation bounds do not depend on $\boldsymbol{D}$ and $\boldsymbol{d}$, the bounds for the transformed set are the same as for the original set. One well-known uncertainty type covered by these transformations is scaled ellipsoidal uncertainty $\left\{\boldsymbol{\xi} \middle| \sum_{i=1}^m w_i \xi_i^2 \leq 1\right\}$, which was first proposed by Ben-Tal and Nemirovski [7]. These sets can be constructed via transformations from hypersphere uncertainty sets with a diagonal matrix $\boldsymbol{D}$ with $D_{ii} = \frac{1}{\sqrt{w_i}}$. Scaled ellipsoidal uncertainty has been applied to many robust optimization problems, including portfolio optimization [7], supply chain contracting [10], network design [47], and facility location [5].

Another widely used class that is partially covered by these positive affine transformations are factor-based uncertainties given by sets $\mathcal{U} = \{\boldsymbol{D}\boldsymbol{z} + \boldsymbol{d} | \boldsymbol{z} \in \mathcal{U}^z\}$. In these sets, uncertainties affinely depend on a set of factors $\boldsymbol{z}$ that are drawn from a factor uncertainty set $\mathcal{U}^z$. Problems that were solved using such uncertainty sets include, among others, portfolio optimization [34, 35] and multi-period inventory management [2, 57]. In contrast to the general factor sets, that have no limitations on $\boldsymbol{D}$ and $\boldsymbol{d}$, our approach is restricted to positive factor matrices which allows only for positive correlations between uncertainties. Nevertheless, even this subset of factor-based uncertainties has wide applicational use. As an intuitive example, one could consider component demands where the factors are demands for finished products.

4 Re-scaling expensive vertices

On some instances, uncertainties do not have a uniform impact on the objective. While accounting for high values in some of the uncertainty dimensions might drastically influence the objective, accounting for high values in others might barely have an impact. This effect can be crucial in our problem setting, because the creation of dominating sets may overemphasize single uncertainty dimensions by up to a factor of β by design. Accordingly, it might thus be beneficial to dominate these uncertainty dimensions more carefully on instances where a few critical uncertainties cause almost all the cost. In this context, we show that it is possible to shrink critical vertices $\boldsymbol{v}_i$ at the cost of slightly shifting all other vertices towards their direction.

We illustrate the re-scaling process following from Lemma 13 in Fig. 5. In the depicted example we shrink the critical vertex $\boldsymbol{v}_1$ and shift the two remaining vertices $\boldsymbol{v}_0, \boldsymbol{v}_2$ towards uncertainty dimension ξ_1. As a consequence, the cost of the vertex solution $\boldsymbol{x}_1$ decreases, while the costs of the solutions $\boldsymbol{x}_0, \boldsymbol{x}_2$ increase. The cost reduction on the critical vertex $\boldsymbol{v}_1$ leads to a reduction of the worst case vertex cost z which by the construction of LP (3) corresponds to an overall improvement of the objective function. Note, that the dominating set $\hat{\mathcal{U}}$ used in the example is not an optimal choice for the hypersphere uncertainty set depicted. However, the effect would barely be visible in two dimensions without this sub-optimal choice.

Lemma 13 *Let* $\hat{\mathcal{U}} := \text{conv}(\boldsymbol{v}_0, \ldots, \boldsymbol{v}_m)$ *be a dominating set for* $\mathcal{U}$. *Let* $\boldsymbol{s} \in [0, 1]^m$ *be a vector of scales. Then* $\hat{\mathcal{U}}' := \text{conv}(\boldsymbol{v}'_0, \ldots, \boldsymbol{v}'_m)$ *with* $v'_{ij} := s_j(1 - v_{ij}) + v_{ij}$ *is also a dominating set for* $\mathcal{U}$.

We present the proof for Lemma 13 in "Appendix M".

The two extreme cases for the modified dominating sets from Lemma 13 are given by $\boldsymbol{s} = \boldsymbol{0}$ and $\boldsymbol{s} = \boldsymbol{e}$. While the dominating set does not change for $\boldsymbol{s} = \boldsymbol{0}$, all vertices become the unit vector consisting of ones in every component for $\boldsymbol{s} = \boldsymbol{e}$. Intuitively, increasing s_i leads to shifting the ith component towards one. In the same way as we constructed our dominating sets in (2) this shift of the ith component towards one increases the value for all $\boldsymbol{v}_j$ with $j \neq i$ and decrease it for $j = i$. Note that this Lemma is not limited to uncertainty sets that are constructed as described in (2), but holds for any dominating set created as a convex combination of vertices.

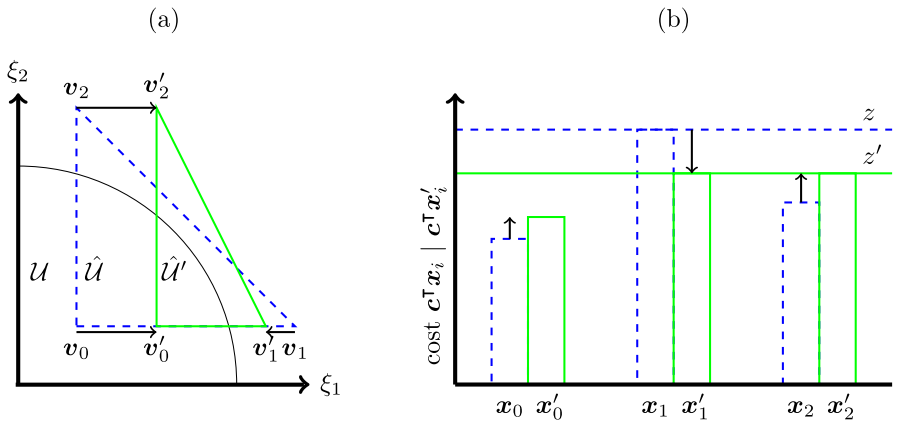


Fig. 5 Re-scaling of the expensive vertex $\boldsymbol{v}_1$ in a dominating set $\hat{\mathcal{U}}$ for uncertainty set $\mathcal{U}$. **a** Shows the change of dominating set $\hat{\mathcal{U}}$ (blue, dashed) and it's vertices $\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{v}_2$ to the new re-scaled dominating set $\hat{\mathcal{U}}'$ (green, solid) with vertices $\boldsymbol{v}_0', \boldsymbol{v}_1', \boldsymbol{v}_2'$ for $s_1 = 0.5, s_2 = 0$. **b** Shows the costs for the vertex solutions $\boldsymbol{x}_i$ with maximal cost z (blue, dashed) compared to the costs for the re-scaled vertex solutions $\boldsymbol{x}_i'$ with maximal cost z' (green, solid) (color figure online)

The transformation used in Lemma 13 is linear, which allows us to add $\boldsymbol{s}$ as a further decision variable to the second constraint of LP (3) for a given $\hat{\mathcal{U}}$. As $\boldsymbol{s} = \boldsymbol{0}$ gives the original dominating set, any optimal solution found with these additional decision variables is at least as good as a non-modified solution. Thus, all performance bounds shown in Sect. 3 also hold for these re-scaled piecewise affine policies. Note that Proposition 8 also extends to the re-scaled uncertainty set; thus, all re-scaled piecewise affine policies are strictly dominated by affine policies for integer budgeted uncertainty.

Adding $\boldsymbol{s}$ to the LP increases its size, which in practice will often lead to an increase of solution times. To limit the increase of model size, it is possible to add only those s_i where one expects $s_i > 0$, as not adding a variable s_i is equivalent to fixing $s_i = 0$. Those s_i with $s_i > 0$ correspond to the critical uncertainty dimensions, and an experienced decision maker with sufficient knowledge of the problem might be able to identify them a priori.

5 Piecewise affine policies via liftings

Georghiou et al. [31] propose piecewise affine policies via liftings. In this section, we strengthen these policies by using the insights from the policies constructed in Sects. 2 and 3.

To construct piecewise affine policies via liftings, in the context of Assumption 1, we first choose $r_i - 1$ breakpoints

$$0 < z_1^i < z_2^i < \cdots < z_{r_i-1}^i < 1,$$

for each uncertainty dimension $i \in \{1, 2, \ldots, m\}$. For ease of notation let $z^i_0 := 0$, $z^i_{r_i} := 1$. With these breakpoints, we define the lifting operator $L\colon \mathbb{R}^m \to \mathbb{R}^{m^L}$,where $m^L := \sum_{i=1}^m r_i$, componentwise by

$$L_{i,j}(\boldsymbol{\xi}) := \Big(\min(\xi_i - z^i_{j-1}, z^i_j - z^i_{j-1})\Big)_+.$$

Further, we define the linear retraction operator $R\colon \mathbb{R}^{m^L} \to \mathbb{R}^m$ componentwise by

$$R_i(\boldsymbol{\xi}^L) := \sum_{j=1}^{r_i} \xi^L_{i,j}.$$

Here $\xi^L_{i,j}$ are the components of the m^L dimensional vector

$$\boldsymbol{\xi}^L = (\xi^L_{1,1}, \xi^L_{1,2}, \ldots, \xi^L_{1,r_1}, \xi^L_{2,1}, \ldots, \xi^L_{m,1}, \ldots, \xi^L_{m,r_m})^\intercal \in \mathbb{R}^{m^L}.$$

Note that $R \circ L\colon \mathcal{U} \to \mathcal{U}$ is the identity. Finally, Georghiou et al. [31] construct a lifted uncertainty set $\mathcal{U}^L$ via

$$\mathcal{U}^L := \left\{ \begin{array}{ll} \boldsymbol{\xi}^L \in \mathbb{R}^{m^L}_+ \colon R(\boldsymbol{\xi}^L) \in \mathcal{U}, & \\ \dfrac{\xi^L_{i,j+1}}{z^i_{j+1} - z^i_j} \le \dfrac{\xi^L_{i,j}}{z^i_j - z^i_{j-1}} & \forall i \in \{1, \ldots, m\},\ j \in \{1, \ldots, r_i - 1\} \\ \xi^L_{i,1} \le z^i_1 & \forall i \in \{1, \ldots, m\} \end{array} \right\}. \quad (8)$$

Uncertainty set (8) is an outer approximation of the lifting $L(\mathcal{U}) \subseteq \mathcal{U}^L$ and omits $R(\mathcal{U}^L) = \mathcal{U}$. Replacing the uncertainty in Problem (1) with this lifted uncertainty set yields the lifted adjustable problem

$$\begin{array}{lll} Z^L_{AR}(\mathcal{U}^L) = \min\limits_{\boldsymbol{x}(\boldsymbol{\xi}^L)} \max\limits_{\boldsymbol{\xi}^L \in \mathcal{U}^L} & \boldsymbol{c}^\intercal \boldsymbol{x}(\boldsymbol{\xi}^L) & \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x}(\boldsymbol{\xi}^L) \ge \boldsymbol{D}R(\boldsymbol{\xi}^L) + \boldsymbol{d} & \forall \boldsymbol{\xi}^L \in \mathcal{U}^L. \end{array} \quad (9)$$

Limiting $\boldsymbol{x}$ to affine policies in the lifted space, yields piecewise affine policies in the original space, which give tighter approximations than affine policies in the original space, i.e., $Z^L_{AFF}(\mathcal{U}^L) \le Z_{AFF}(\mathcal{U})$ [31]. However, $\mathcal{U}^L$ is not a tight outer approximation of $L(\mathcal{U})$, leading to little or no improvements over affine policies on some instances [17, 31]. In fact, we show that in the framework of ARO, the piecewise affine policies induced by lifted uncertainty (8) are equivalent to classical affine policies, in the sense that for any optimal feasible lifted affine policy, there is an affine policy with the same objective value and vice versa.

Proposition 14 *Let Z_{AFF} be the optimal objective for affine policies on $Z_{AR}(\mathcal{U})$ and let Z_{LIFT} be the optimal objective value for lifted affine policies on $Z^L_{AR}(\mathcal{U}^L)$. Then*

$$Z_{AFF} = Z_{LIFT}.$$

We present the proof for Proposition 14 in "Appendix N".

We overcome this shortcoming in the construction of lifted affine policies using our results on dominating uncertainty sets from Sect. 2. Consider the lifting with one break-point v_{0i} per uncertainty dimension. Here v_{0i} is the ith component of the base vector $\boldsymbol{v}_0$ from Sect. 2. Then the lifting operator L becomes

$$L_{ij}(\boldsymbol{\xi}) = \begin{cases} \min(\xi_i, v_{0i}) & \text{if } j = 1 \\ (\boldsymbol{h}(\boldsymbol{\xi}) - \boldsymbol{v}_0)_i & \text{if } j = 2. \end{cases}$$

With this construction of L it is easy to verify that by Condition (5) each $\boldsymbol{\xi}^L \in L(\mathcal{U})$ satisfies $\sum_{i=1}^m \frac{\xi^L_{i,2}}{\rho_i} \leq 1$. Accordingly, we can tighten the lifted uncertainty set $\mathcal{U}^L$ and get the new lifted uncertainty set

$$\hat{\mathcal{U}}^L := \left\{ \boldsymbol{\xi}^L \in \mathcal{U}^L, \sum_{i=1}^m \frac{\xi^L_{i,2}}{\rho_i} \leq 1 \right\}. \tag{10}$$

By construction $\hat{\mathcal{U}}^L$ is an outer approximation of $L(\mathcal{U})$ and we have $L(\mathcal{U}) \subseteq \hat{\mathcal{U}}^L \subseteq \mathcal{U}^L$ and $R(\hat{\mathcal{U}}^L) = \mathcal{U}$. Thus, affine policies on the lifted problem with uncertainty set $\hat{\mathcal{U}}^L$ yield valid piecewise affine policies for the original problem. Furthermore, the construction of $\hat{\mathcal{U}}^L$ guarantees that the lifted policies yield tighter approximations than our piecewise affine policies via domination and classical lifted policies with the same breakpoints. Consequently, all approximation bounds for piecewise affine policies via domination also hold for the strengthened piecewise affine policies via lifting.

Proposition 15 *Let Z_{LIFT} be the optimal objective value found by the lifting policies with breakpoints $\boldsymbol{v}_0$ and lifted uncertainty set $\mathcal{U}^L$ defined in (8), Z_{TLIFT} be the optimal objective value found by the lifting policies with breakpoints $\boldsymbol{v}_0$ and tightened lifted uncertainty set $\hat{\mathcal{U}}^L$ defined in (10), and Z_{SPAP} be the optimal objective value found by the piecewise affine policies with re-scaling described in Sect. 4. Then*

$$Z_{TLIFT} \leq Z_{LIFT}, \qquad Z_{TLIFT} \leq Z_{SPAP}.$$

We present the proof for Proposition 15 in "Appendix O".

6 Numerical experiments

In this section, we present two numerical experiments to compare the performance of our piecewise affine policies for the different constructions of $\hat{\mathcal{U}}$ and our tightened

Table 2 Overview of policies compared in the experiments

Policy	Dominating set or policy function	
BOX	$\hat{\mathcal{U}}_{\text{BOX}} = \{\boldsymbol{e}\}$	Constant policy
AFF	$\boldsymbol{x}^t(\underline{\boldsymbol{\xi}}^t) = \boldsymbol{P}^t\underline{\boldsymbol{\xi}}^t + \boldsymbol{q}^t$	Affine adjustable policy [9]
BG	$\boldsymbol{x}^t(\underline{\boldsymbol{\xi}}^t) = \max_i \overline{\boldsymbol{P}}_i^t\underline{\boldsymbol{\xi}}^t + \overline{\boldsymbol{q}}_i^t - \max_i \underline{\boldsymbol{P}}_i^t\underline{\boldsymbol{\xi}}^t + \underline{\boldsymbol{q}}_i^t$	Near-optimal pap [17]
PAPBT	$\hat{\mathcal{U}}_{\text{BT}} = \beta\,\text{conv}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m, \boldsymbol{v})$	Literature dominating set [14]
PAP	$\hat{\mathcal{U}} = \text{conv}(\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$	Our dominating set (c.f., Sect. 2.1)
SPAP	$\hat{\mathcal{U}} = \text{conv}(\boldsymbol{v}'_0, \boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_m)$	PAP with re-scaling (c.f., Sect. 4)
TLIFT	AFF on $\hat{\mathcal{U}}^L$	Tightened pap via lifting (c.f., Sect. 5)

piecewise affine policies via lifting with the performance of other policies from the literature. We compare the performance in terms of both objective value and computational time.

We run both of the following tests with hypersphere uncertainty sets and budgeted uncertainty sets. In the experiments of Ben-Tal et al. [14] piecewise affine policies performed particularly well compared to affine adjustable policies for hypersphere uncertainty and relatively bad for budgeted uncertainty. Accordingly, considering these two uncertainty types gives a good impression of the benefits and limitations of piecewise affine policies. Additionally, this experimental design allows us to analyze whether or not the new formulations with the tighter bounds presented in Propositions 5 and 7 have a significant impact in practice.

In our studies, we compare the following policies: the affine policies described in Ben-Tal et al. [9] (AFF), the constant policies resulting from a dominating set $\hat{\mathcal{U}} = \{\boldsymbol{e}\}$ with only a single point which by down-monotonicity corresponds to a box (BOX), the near-optimal piecewise affine policies with two pieces proposed in Bertsimas and Georghiou [17] (BG), our piecewise affine policies constructed as described in Propositions 5 and 7 (PAP), the piecewise affine policies constructed as described in Propositions 1 and 5 in Ben-Tal et al. [14] (PAPBT), our piecewise affine policies with the vertex re-scaling heuristic described in Sect. 4 (SPAP), and our tightened piecewise affine policies via lifting described in Sect. 5 (TLIFT). Note, that piecewise affine policies via lifting from Georghiou et al. [31] are implicitly included in the comparison by Proposition 14. In Table 2 we give an overview of all policies compared in our experiments.

For all studies, we used Gurobi Version 9.5 on a 6 core 3.70 GHz i7 8700K processor using a single core per instance.

6.1 Gaussian instances

We base our first set of benchmark instances on the experiments of Ben-Tal et al. [14] and Housni and Goyal [40]. Accordingly, we generate instances of Problem (1) by choosing $m = l = n$, $\boldsymbol{d} = \boldsymbol{0}$, $\boldsymbol{D} = \boldsymbol{1}_m$ and generate $\boldsymbol{c}$, $\boldsymbol{A}$ randomly as

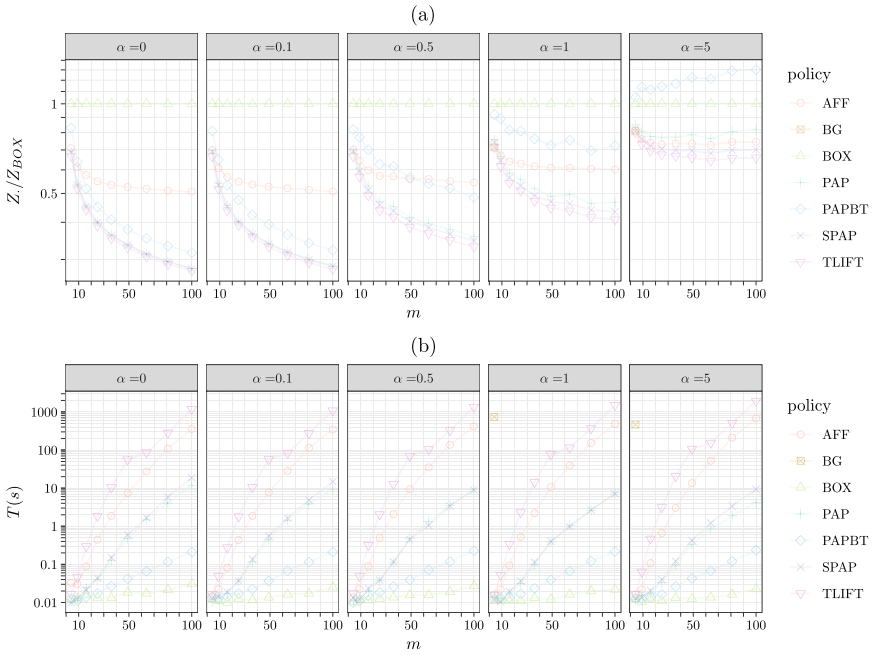


Fig. 6 Relative objective values (**a**) and solution times (**b**) for different policies on gaussian instances with hypersphere uncertainty

$$\boldsymbol{c} = \boldsymbol{e} + \alpha \boldsymbol{g},$$
$$\boldsymbol{A} = \boldsymbol{1} + \boldsymbol{G}.$$

Here, $\boldsymbol{e}$ is the vector of all ones, $\boldsymbol{1}$ is the identity matrix, $\boldsymbol{g}$ and $\boldsymbol{G}$ are randomly generated by independent and identically distributed (i.i.d.) standard gaussians, and α is a parameter that increases the asymmetry of the problem. More specifically, $\boldsymbol{G}$ is given by $G_{ij} = |Y_{ij}|/\sqrt{m}$ and $\boldsymbol{g}$ is given by $g_i = |y_i|$, where Y_{ij} and y_i are i.i.d. standard gaussians. Uncertainties $\boldsymbol{\xi}$ and decision variables $\boldsymbol{x}$ are split into $\lfloor\sqrt{m}\rfloor$ stages where the ith decision always belongs to the same stage as the ith uncertainty. For the budgeted uncertainty sets we use a budget of $k = \sqrt{m}$. We consider values of $m = i^2$ for $i \in \{2, \ldots, 10\}$ and values of α in $\{0, 0.1, 0.5, 1, 5\}$. For each pair of m, α, we consider 30 instances. To make the results more comparable, we scale all objective values presented by the constant policies results (i.e., $Z./Z_{BOX}$) and report averages over all solved instances. For each parameter pair m, α, we only consider those policies that found solutions on at least 75% of instances within a hard solution time limit of 4 hours. Additionally, we present all results on a logarithmic scale and artificially lower bound the scale for solution times by $0.01s$ to make the effects on higher solution times more visible.

Figure 6 shows the performances and solution time results on hypersphere uncertainty sets for the different policies. First, note that BG only finds solutions within the time limit for the smallest instances yielding objective values comparable or marginally

better to TLIFT. For the other policies we observe that piecewise adjustable policies perform significantly better than affine adjustable policies for small values of α. The improvement increases for larger values of m reaching almost a factor of 2 for $m = 100$ on our policies PAP, SPAP, and TLIFT. As expected, the performance of PAP and SPAP for small values of α is almost indistinguishable due to the construction. Additionally, we find that TLIFT only yields marginal improvements over PAP and SPAP for small values of α. For larger values of α the improvements of the piecewise affine adjustable policies vanish and TLIFT starts to improve over SPAP. The two policies without re-scaling (PAP and PAPBT) perform even worse than AFF for $\alpha = 5$. More severely, PAPBT even performs worse than BOX, which already is a worst-case policy. Only SPAP and TLIFT achieve better results than affine adjustable policies for all values of α.

While solution times for all policies except BOX grow exponentially in the instance size, domination-based piecewise affine adjustable policies are by orders of magnitude faster than classical affine adjustable policies (AFF) and piecewise affine polices via lifting (TPAP). These solution time improvements exceed a factor of 100 for piecewise affine policies PAP and SPAP and a factor of 1000 for PAPBT. Also, solution times of domination-based piecewise affine adjustable policies are barely influenced by values of α. This is not the case for AFF and TLIFT, which take longer to solve for increasing α. While this effect is not easily visible in Fig. 6 due to the logarithmic scale, the solution time difference for AFF and TLIFT between $\alpha = 0$ and $\alpha = 5$ reaches up to a factor of two on large instances.

Figure 7 shows the performances and solution time results on budgeted uncertainty sets. We observe that for budgeted uncertainty sets domination-based piecewise affine policies perform slightly worse than affine policies throughout all instances. This observation nicely demonstrates that our theoretical and experimental results are aligned, as the worse performance is perfectly explained by Proposition 8, which shows that affine policies strictly dominate our piecewise affine policies. Again, we observe that for higher values of α, PAP and PAPBT perform even worse than BOX. However, the solution values for SPAP stay within 5% of the affine solution values throughout all instances. We further observe that TLIFT yields the same objective values as AFF throughout all instances. Only BG yields slightly better solutions than AFF on some instances. Solution times behave similarly to solution times on hypersphere uncertainty, confirming that piecewise affine policies are found by orders of magnitude faster on different instances and uncertainty types. Only for BG solution times improve significantly, suggesting that BG is highly dependent on the shape of the uncertainty sets.

Solution time and performance results for $\alpha = 0$ align with the results found by Ben-Tal et al. [14] for the two-stage problem variant. This demonstrates that our generalized piecewise affine policies do not only extend all theoretical performance bounds, but also achieve comparable numerical results in a multi-stage setting. However, by breaking the symmetry by increasing α, we show that pure domination-based piecewise adjustable policies perform poorly on highly asymmetric instances and re-scaling (SPAP) or tightened lifting (TLIFT) constitute good techniques to overcome this shortcoming.

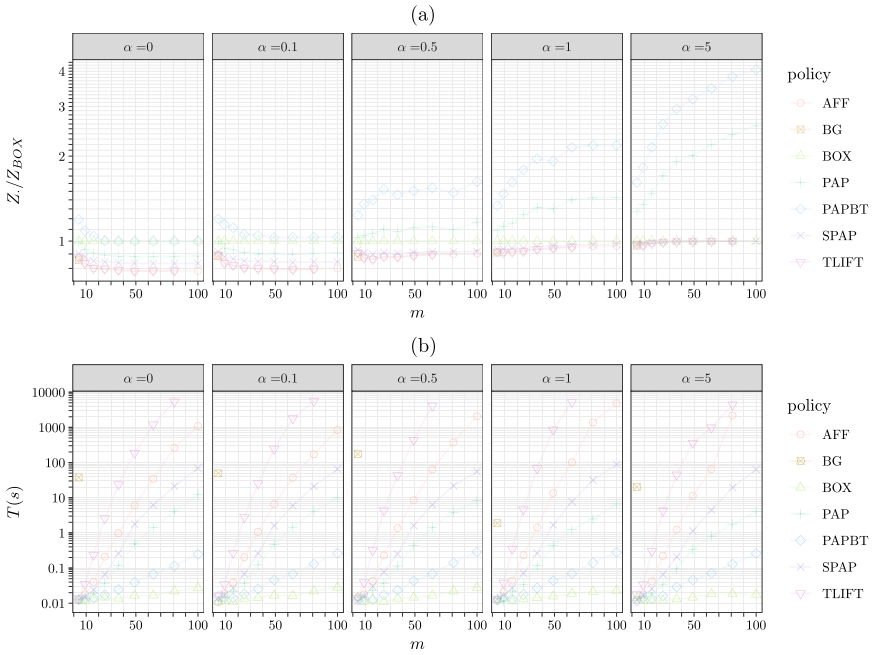


Fig. 7 Relative objective values (**a**) and solution times (**b**) for different policies on gaussian instances with budgeted uncertainty

6.2 Demand covering instances

For the second set of test instances, we consider the robust demand covering problem with non-consumed resources and uncertain demands. The problem has various applications, among others in the domains of appointment scheduling, production planning, and dispatching and is especially relevant for the optimization of service levels. Our instances consist of m^l locations, m^p planning periods, and m^e execution periods per planning period. In each execution period t, an uncertain demand ξ_{lt} has to be covered at each location l. To do so, the decision maker can buy a fixed number of resources R at a unit cost of c^R in the first stage and then distributes these R resources among the locations at the beginning of each planning period. If a demand cannot be met with the resources assigned to a location, the decision maker will either delay the demand to the next period or redirect it to another location. In either case a fraction $q^d_{tl} \in [0, 1]$ or $q^r_{tll'} \in [0, 1]$ of the demand is lost. Each unit of lost demand causes costs of c^D. Mathematically, the robust demand covering problem with non-consumed resources and uncertain demands is given by the robust LP (11), where parameters and variables are summarized in Table 3.

$$\min \quad c^R R + c^D \left(\sum_{t\in\mathcal{T}, l\in\mathcal{L}} q^d_t s^d_{tl} + \sum_{t\in\mathcal{T}, l\neq l'\in\mathcal{L}} q^r_{ll'} s^r_{tll'} \right) \tag{11a}$$

Table 3 Notation for the robust demand covering problem with non-consumed resources and uncertain demands

Sets & elements	
$l \in \mathcal{L}$	Locations
$p \in \mathcal{P}$	Planning periods
$t \in \mathcal{T}$	All ($m^p m^e$) execution periods
Parameters	
c^R	Cost of buying one unit of resources
c^D	Cost of loosing one unit of demand
$q_t^d \in [0, 1]$	Fraction of demand lost when delayed in period t
$q_{ll'}^r \in [0, 1]$	Fraction of demand lost when redirected form location l to l'
Uncertainties	
$p(t)$	Planning period containing execution period t
ξ_{tl}	Demand at location $l \in \mathcal{L}$ in execution period $t \in \mathcal{T}$
Decision variables	
R	Number of resources bought (decided in the very first stage)
r_{pl}	Number of resources assigned to location $l \in \mathcal{L}$ in planning period $p \in \mathcal{P}$ (decided at the beginning of planning period p)
$s_{tll'}^r$	Demand redirected from location l to location l'(decided in decision stage t)
s_{tl}^d	Demand delayed to the next period (decided in decision stage t)

$$\text{s.t.} \quad r_{p(t)l} + s_{tl}^d - (1 - q_{t-1}^d) s_{(t-1)l}^d + \sum_{l' \in \mathcal{L}, l' \neq l} \left(s_{tll'}^r - (1 - q_{ll'}^r) s_{tl'l}^r \right) \geq \xi_{tl} \qquad \forall \boldsymbol{\xi} \in \mathcal{U}, \forall t \in \mathcal{T}, l \in \mathcal{L} \tag{11b}$$

$$\sum_{l \in \mathcal{L}} r_{pl} \leq R \qquad \forall p \in \mathcal{P} \tag{11c}$$

$$R, \boldsymbol{r}, \boldsymbol{s}^r, \boldsymbol{s}^d \geq 0 \tag{11d}$$

Here Objective (11a) minimizes the sum of resource costs and lost demand costs due to delay and relocation. Constraints (11b) ensure that all demands are fulfilled, delayed, or relocated, and Constraints (11c) upper bound the allocated resources in each planning period by the total number of available resources R.

In most real-world applications of the demand covering problem, some of the demand will be revealed before the actual demand occurs, e.g. due to already existing contracts, sign-ups, orders, or due to forecasting. To incorporate the increase in knowledge over time, we assume an uncertainty vector of the form $\boldsymbol{\xi} = \boldsymbol{\xi}^c + \boldsymbol{\xi}^p + \boldsymbol{\xi}^e$. Here, $\boldsymbol{\xi}^c$ is constant and known before the first stage decision, $\boldsymbol{\xi}^p$ is revealed before each planning period and $\boldsymbol{\xi}^e$ accounts for the short-term uncertainties revealed before

each execution period. Specifically, we assume that demands are given by

$$\xi_{lt} = d_{lt}\left(1 + \xi^p_{lt} + \frac{1}{2}\xi^e_{lt}\right), \tag{12}$$

where d_{lt} is the base demand for location l in execution period t. Here the uncertainty vector $(\boldsymbol{\xi}^p, \boldsymbol{\xi}^e)$ is taken from one base uncertainty set $\mathcal{U}^B$ of dimension $m = 2m^l m^p m^e$. Note, that short-term uncertainties have a less severe effect on demands than uncertainties known in advance. The resulting demand uncertainty set

$$\mathcal{U} = \left\{\boldsymbol{\xi} \;\middle|\; \xi_{lt} = d_{lt}\left(1 + \xi^p_{lt} + \frac{1}{2}\xi^e_{lt}\right), (\boldsymbol{\xi}^p, \boldsymbol{\xi}^e) \in \mathcal{U}^B\right\}$$

is $m/2$ dimensional, where in our experiments $\mathcal{U}^B$ is either a hypersphere uncertainty set, or a budgeted uncertainty set with budget $\sqrt{m}$.

To construct our instances we draw $m^l \in \{2, 4, 6, 8, 10\}$ locations at uniform random integer positions in the square $\left[0, 2\lfloor\sqrt{m^l}\rfloor + 1\right]^2$. In each of the $m^p \in \{1, 3, 5, 7\}$ planning periods, we consider $m^e = 8$ execution periods corresponding to the hours in a working day. We assume that a fraction $q^d_{tl} = 0.1$ of the demand is lost when deferred to a later execution period and consider a doubled loss rate ($q^d_{tl} = 0.2$) when demand is deferred to another planning period. Similarly, a fraction of the demand is lost when assigned to another location. We assume this fraction to be correlated to the distance and given by $q^r_{ll'} := \min\left(1, 0.02 \cdot \text{dist}(l, l')\right)$. We draw the base demands d_{lt} uniformly from the normal distribution $\mathcal{N}(10, 4)$. Finally, we set $c^R = 1$ and choose $c^D \in \{0.1, 0.25, 0.5\}$. For each combination of m^l, m^p we consider 45 instances, where we use each possible value for $c^D \in \{0.1, 0.25, 0.5\}$ in a third of these instances.

To analyze practical expected objectives, we also report a simulated average objective that the respective policies achieved on 500 randomly drawn uncertainty realizations $\boldsymbol{\xi}$, in addition to the robust objective value. We give a detailed description of how uniform uncertainty realizations can efficiently be sampled from the budgeted and hypersphere uncertainty sets in "Appendix P". We again scale the results by the results achieved by constant policies and use logarithmic scales. For each instance size, we only consider those policies that found solutions on at least 75% of instances within a hard solution time limit of 2 hours.

First, we observe that BG did not solve any instance within the time limit which can be attributed to the fact that our demand covering instances are significantly larger than our gaussian instances.

For the remaining policies, Fig. 8 shows the performance and solution time results on demand covering instances with hypersphere uncertainty. Compared to our previous experiment (see Sect. 6.1) we no longer observe the strong objective improvements of piecewise affine policies over affine adjustable policies. Still, our piecewise affine formulations give similar results as affine policies, and we observe small improvements on instances with a larger number of planning periods, with TLIFT yielding strict improvements for $m^p \geq 3$. On the simulated realizations, improvements of PAP and

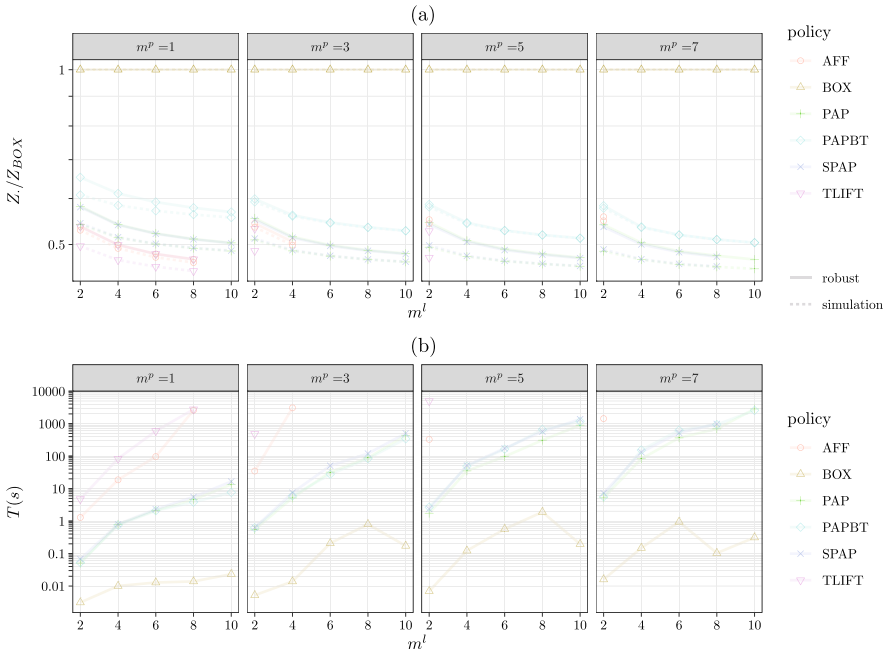


Fig. 8 Relative objective values (**a**) and solution times (**b**) for different policies on demand covering instances with hypersphere uncertainty

SPAP over affine adjustable policies can already be seen for $m^p \geq 3$, which might be of interest for a decision maker with practical interest beyond worst-case solutions.

For the solution times, we observe similar improvements to the ones observed on the gaussian instances. Still, all domination-based piecewise affine policies can be found by orders of magnitude faster than affine policies and TLIFT. Interestingly, PAP is solved similarly fast as PAPBT on these instances, while still achieving up to 15% better objective values on all instances. The largest instance that could be solved by affine adjustable policies within two hours consisted of 320 uncertainty variables and 700 decision variables, while the largest instance solved by PAPs was more than three times larger with 1, 120 uncertainty variables and 6, 230 decision variables.

Figure 9 shows the results on demand covering instances with budgeted uncertainty sets. For budgeted uncertainty sets, domination-based piecewise affine policies perform worse than affine adjustable policies throughout all instances, which again can be explained by Proposition 8. Notably, PAPBT even performs worse than constant policies on most instances. In this setting, domination-based piecewise affine policies remain better only from a solution time perspective, as they still solve by orders of magnitude faster than affine policies. Also, TLIFT does not yield any improvements over AFF.

While in this set of experiments piecewise affine policies do not show the same improvements in the objective over affine adjustable policies, they still perform slightly better with hypersphere uncertainty on larger instances. Also, they still solve by orders

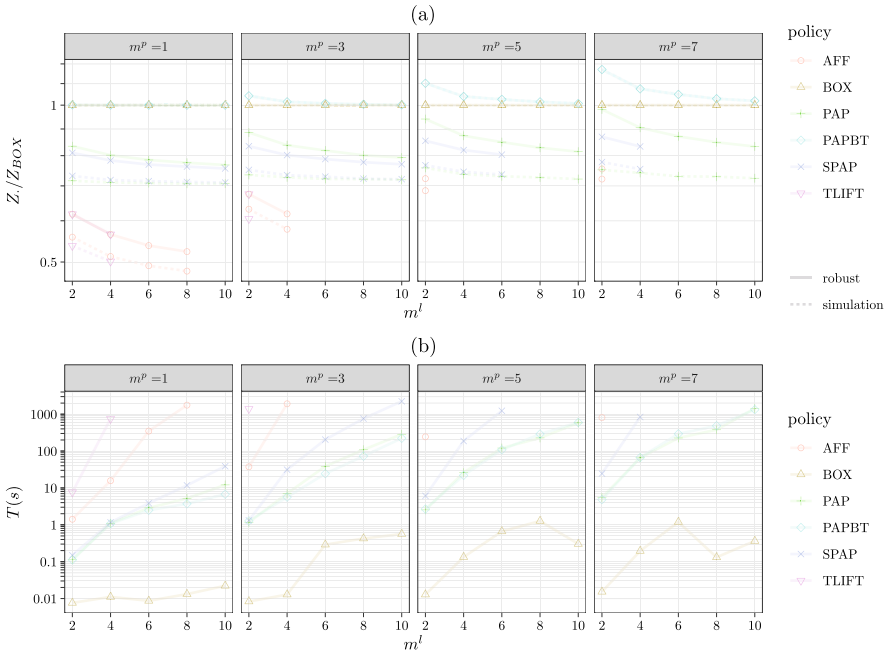


Fig. 9 Relative objective values (**a**) and solution times (**b**) for different policies on demand covering instances with budgeted uncertainty

of magnitude faster, which makes them an attractive alternative for large-scale optimization in practice.

6.3 Discussion

In the experiments presented in Sects. 6.1 and 6.2, some results, e.g. the solution time improvements of piecewise affine policies over affine policies, are consistent throughout all instances. However, other results strongly depend on the set of benchmark instances used. In the following, we give an explanation for the strong solution time improvements, discuss two of the main deviations that we observe between our results on gaussian instances and demand covering instances, and give intuitions of why these differences occur.

Size of robust counterparts: Throughout all experiments, we see strong solution time improvements of piecewise affine policies over affine policies—including affine policies on lifted uncertainty TLIFT. These can be explained by their respective robust counterparts. The robust counterpart for piecewise affine policies is given by LP (3). For the robust counterparts of affine policies, we refer to Ben-Tal et al. [9]. Counterparts of piecewise affine policies have $O(nm)$ variables compared to $O(nm + lm)$ variables for affine policies and both counterparts have $O(lm)$ constraints. More critically, constraints for piecewise affine policies contain at most $O(n)$ variables each

and feature a block structure, which solvers use to significantly speed up the solution process. This block structure is only connected by the nonanticipativity constraints. In the robust counterpart of affine policies on the other hand, $O(l)$ constraints have up to $O(nm)$ variables resulting in a denser constraint matrix and the lack of a block structure. Moreover, the robust counterpart for affine policies on hypersphere uncertainty sets is no longer linear. Instead, a quadratic program has to be solved, which tends to be computationally more challenging.

Solution times of PAPBT: In the experiments we see that PAPBT solves by a factor of 10 to 50 faster than PAP on gaussian instances, but both find solutions similarly fast on-demand covering instances. This can be explained by the construction of PAPBT and the structure of the instances' constraints. On the gaussian instances, $\boldsymbol{A}$ and $\boldsymbol{x}$ are non-negative and $\boldsymbol{D}$ is the unit matrix. In the construction of dominating sets $\hat{\mathcal{U}}$ for PAPBT most of the vertices are chosen to be scaled unit vectors. As a consequence, most Constraints (3c) in LP (3) have a zero right-hand side, such that they trivially hold. Consequently, these constraints can be eliminated, which reduces the total number of constraints by a factor of $O(m)$. On demand covering instances, however, $\boldsymbol{A}$ contains negative entries. Thus, no constraints trivially hold and no constraints can be removed. For a decision maker who is primarily interested in fast policies, this gives a good criterion on when PAPBT can improve solution times and when no such improvements can be expected.

Performance differences between gaussian and demand covering instances: We observe that the strong performance improvements of piecewise affine policies over affine policies on gaussian instances with hypersphere uncertainty do not transfer to our demand covering instances. This suggests that the relative performance between piecewise affine policies and affine policies significantly depends on the structure of the problem at hand. An intuitive explanation for this lies in the policies' construction.

Recall that piecewise affine policies derive solutions by finding vertex solutions $\boldsymbol{x}_i$ that can be extended to a full solution. Thereby, $\boldsymbol{x}_0$ focuses on finding a good solution for uncertainty realizations where all uncertainties take equal values, and each $\boldsymbol{x}_i$ focuses on finding a good recourse to uncertainty ξ_i. Thus, good results can be expected when there are (a) synergy effects that can be utilized by $\boldsymbol{x}_0$, and (b) good universal recourse decisions for each uncertainty ξ_i that do not depend on the realization of other uncertainty dimensions and can be exploited by $\boldsymbol{x}_i$. On the other hand, affine policies directly find solutions on the original uncertainty set. In doing so, they do not depend as strongly on good universal recourse decisions as piecewise affine policies do. However, they also lack the ability to use synergy effects in the way vertex solutions $\boldsymbol{x}_0$ do.

The gaussian instances used fulfill both of the properties that are favorable for piecewise affine policies. By being based on a unity matrix, $\boldsymbol{A}$ has relatively large values along the diagonal, leading to the existence of good universal recourse decisions. Additionally, the relatively small non-negative entries on the non-diagonals lead to synergy effects for uncertainty realizations with many small values. Demand covering instances, however, do not fulfill these properties. The question of how to redirect demand optimally heavily depends on the demand observed at other locations. Also,

the only synergistic effects that can be used solely emerge when multiple demands occur at the same location in a single planning period.

On general instances in practice, we would thus not expect to see the same performance improvements that could be observed on our gaussian benchmark instance. Still, piecewise affine policies find solutions by orders of magnitude faster than affine policies and achieve good results throughout all benchmark instances with hypersphere uncertainty. Additionally, Properties (a) and (b) give intuitive criteria on when strong objective improvements over affine policies can be expected.

7 Conclusion

In this work, we presented piecewise affine policies for multi-stage adjustable robust optimization. We construct these policies by carefully approximating uncertainty sets with a dominating polytope, which yields a new problem that we efficiently solve with a linear program. By making use of the problem's structure, we then extend solutions for the new problem with approximated uncertainty to solutions for the original problem. We show strong approximation bounds for our policies that extend many previously best-known bounds for two-stage ARO to its multi-stage counterpart. By doing so, we contribute towards closing the gap between the state of the art for two-stage and multi-stage ARO. To the best of our knowledge, the bounds we give are the first bounds shown for the general multi-stage ARO Problem. Furthermore, our bounds yield constant factor as well as asymptotic improvements over the state-of-the-art bounds for the two-stage problem variant.

In two numerical experiments, we find that our policies find solutions by a factor of 10 to 1000 faster than affine adjustable policies, while mostly yielding similar or even better results. Especially for hypersphere uncertainty sets our new policies perform well and sometimes even outperform affine adjustable policies up to a factor of two. We observe particularly high improvements on instances that exhibit certain synergistic effects and allow for universal recourse decisions. However, on some instances where few uncertainty dimensions have a high impact on the objective, pure piecewise affine policies perform particularly badly by design, sometimes even worse than constant policies. To mitigate this shortcoming, we present an improvement heuristic that significantly improves the solution quality by re-scaling the critical uncertainty dimension. Furthermore, we construct new tightened piecewise affine policies via lifting that integrate the two frameworks of piecewise affine policies via domination and piecewise affine policies via lifting and combine their approximative power.

While this work extends most of the best-known approximation results for a relatively general class of ARO problems from the two-stage to the multi-stage setting, it remains an open question whether other strong two-stage ARO results can be generalized to multi-stage ARO in a similar manner. Answering this question remains an interesting area for further research. In this context, binary and uncertain recourse decisions remain particularly relevant challenges. Our analysis in Sect. 2.3 has shown that the extension of our policies to encompass these recourse decision types is not straightforward. Nevertheless, exploring the integration of our methodology into the established approaches of piecewise constant policies and k-adaptability, which have

proven to be effective in these cases, appears as a promising starting point for future work. Another interesting area for future research is the extension of piecewise affine policies and the concept of domination to adjustable data-driven and distributionally robust optimization. More specifically, we believe that one can obtain tractable data-driven policies by directly fitting the polyhedral uncertainty sets used to construct our policies from data.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix A Comparison of literature approximation bounds

Table 4 summarizes existing approximation bounds for some commonly used uncertainty sets in two-stage ARO. We want to point out that in addition to being less restrictive, our bounds for multi-stage ARO presented in Table 1 outperform all these bounds except the ones by Housni and Goyal [40], which require the significant restrictive assumption of $\boldsymbol{A}, \boldsymbol{x}, \boldsymbol{c}$ being non-negative.

Appendix B Proof of Theorem 1

Proof We split the proof in two parts. First, we handle the cases where at least $Z_{AR}(\mathcal{U})$ or $Z_{AR}(\hat{\mathcal{U}})$ is negative and show that this already implies that both $Z_{AR}(\mathcal{U})$ and $Z_{AR}(\hat{\mathcal{U}})$ are unbounded. Second, we assume $Z_{AR}(\mathcal{U}), Z_{AR}(\hat{\mathcal{U}}) \geq 0$ are bounded and prove that in this case, the desired performance bounds hold.

Part 1: Let $\tilde{\mathcal{U}} \in \{\mathcal{U}, \hat{\mathcal{U}}\}$ such that $Z_{AR}(\tilde{\mathcal{U}}) < 0$. Then, there exists a solution $\tilde{\boldsymbol{x}}$ such that

$$\max_{\boldsymbol{\xi} \in \tilde{\mathcal{U}}} \boldsymbol{c}^{\mathsf{T}} \tilde{\boldsymbol{x}}(\boldsymbol{\xi}) < 0.$$

Now, assume that $Z_{AR}(\mathcal{U})$ or $Z_{AR}(\hat{\mathcal{U}})$ is bounded and let $\bar{\mathcal{U}} \in \{\mathcal{U}, \hat{\mathcal{U}}\}$ such that $Z_{AR}(\bar{\mathcal{U}})$ is bounded with an optimal solution $\bar{\boldsymbol{x}}$. Arbitrarily fix one $\tilde{\boldsymbol{\xi}} \in \bar{\mathcal{U}}$ and consider the constant vector $\tilde{\boldsymbol{x}}(\tilde{\boldsymbol{\xi}})$. Then $\bar{\boldsymbol{x}}(\tilde{\boldsymbol{\xi}}) + \tilde{\boldsymbol{x}}(\tilde{\boldsymbol{\xi}})$ is a feasible solution for $Z_{AR}(\bar{\mathcal{U}})$ as for

Table 4 Approximation bounds for two-stage ARO

No	[20]	[18]	[15]	[14]	[40]
I	$1+\sqrt{m}$	–	$\frac{1}{2}(1+\sqrt{m})$	$\sqrt[4]{m}$	–
II	$1+\frac{m}{k}$	–	$\frac{k^2+km}{k^2+m}$	$\min(k, \frac{m}{k})$	$\frac{8\log m}{\log\log m}+1$
III	$1+\sqrt[p]{m}$	–	$\frac{m^{\frac{p-1}{p}}+m}{m^{\frac{p-1}{p}}+m^{\frac{1}{p}}}$	$\frac{2}{p}(p-1)^{\frac{p-1}{p}} m^{\frac{p-1}{p^2}}$	–
IV	–	–	–	$\left(\frac{a}{2}+\frac{\sqrt{1-a}}{\sqrt[4]{am^2+(1-a)m}}\right)^{-1}$	–
V	m	$4\sqrt{m}$	–	$4\sqrt{m}$	–
Policy	Static	Affine	Affine	Piecewise affine	Affine

We compare bounds for uncertainty sets of the forms I) hypersphere uncertainty $\left\{\boldsymbol{\xi}\in\mathbb{R}^m_+ \middle| \|\boldsymbol{\xi}\|_2^2\le 1\right\}$; II) budgeted uncertainty $\left\{\boldsymbol{\xi}\in[0,1]^m \middle| \|\boldsymbol{\xi}\|_1\le k\right\}$; III) p-norm ball uncertainty $\left\{\boldsymbol{\xi}\in\mathbb{R}^m_+ \middle| \|\boldsymbol{\xi}\|_p\le 1\right\}$, with $p\ge 1$; IV) ellipsoid uncertainty $\left\{\boldsymbol{\xi}\in\mathbb{R}^m_+ \middle| \boldsymbol{\xi}^\intercal\boldsymbol{\Sigma}\boldsymbol{\xi}\le 1\right\}$, with $\boldsymbol{\Sigma}:=(1-a)\mathbf{1}+a\boldsymbol{J}$ where $\mathbf{1}$ is the unity matrix and $\boldsymbol{J}$ the matrix of all ones; V) general uncertainty sets $\mathcal{U}\subset\mathbb{R}^m_+$. Note, that some of the results require more restrictive assumptions on $\boldsymbol{A},\boldsymbol{x},\boldsymbol{c}$ than our Problem (1). Most prominently [40] requires $\boldsymbol{A},\boldsymbol{x},\boldsymbol{c}$ to be non-negative

all $\bar{\boldsymbol{\xi}}\in\bar{\mathcal{U}}$:

$$\boldsymbol{A}\left(\bar{\boldsymbol{x}}(\bar{\boldsymbol{\xi}})+\tilde{\boldsymbol{x}}(\tilde{\boldsymbol{\xi}})\right)\overset{(a)}{\ge}\boldsymbol{D}\left(\bar{\boldsymbol{\xi}}+\tilde{\boldsymbol{\xi}}\right)+2\boldsymbol{d}\overset{(b)}{\ge}\boldsymbol{D}\bar{\boldsymbol{\xi}}+\boldsymbol{d}.$$

Here, (a) holds because $\tilde{\boldsymbol{x}},\bar{\boldsymbol{x}}$ are feasible solutions and (b) holds as $\tilde{\boldsymbol{\xi}},\boldsymbol{D},\boldsymbol{d}$ are non-negative. For the objective we then find

$$\max_{\tilde{\boldsymbol{\xi}}\in\bar{\mathcal{U}}}\boldsymbol{c}^\intercal\left(\bar{\boldsymbol{x}}(\bar{\boldsymbol{\xi}})+\tilde{\boldsymbol{x}}(\tilde{\boldsymbol{\xi}})\right)\le\max_{\bar{\boldsymbol{\xi}}\in\bar{\mathcal{U}}}\boldsymbol{c}^\intercal\bar{\boldsymbol{x}}(\bar{\boldsymbol{\xi}})+\max_{\boldsymbol{\xi}\in\bar{\mathcal{U}}}\boldsymbol{c}^\intercal\tilde{\boldsymbol{x}}(\boldsymbol{\xi})<Z_{AR}(\bar{\mathcal{U}}).$$

This is a contradiction to $\bar{\boldsymbol{x}}$ being a minimal solution. Thus, $Z_{AR}(\bar{\mathcal{U}})$ cannot be bounded. We have thus shown that if any of $Z_{AR}(\mathcal{U})$, $Z_{AR}(\hat{\mathcal{U}})$ are negative, both have to be unbounded.

Part 2: Assume $Z_{AR}(\mathcal{U}),Z_{AR}(\hat{\mathcal{U}})\ge 0$ are bounded. Let $\hat{\boldsymbol{x}}$ be an optimal solution for $Z_{AR}(\hat{\mathcal{U}})$. Furthermore, let $\boldsymbol{h}:\mathcal{U}\to\hat{\mathcal{U}}$ be the domination function from Definition 1. We claim that $\tilde{\boldsymbol{x}}:=\hat{\boldsymbol{x}}\circ\boldsymbol{h}$ is a feasible solution for $Z_{AR}(\mathcal{U})$. First, we see that by the definition of $\boldsymbol{h}$ the solution $\tilde{\boldsymbol{x}}$ fulfills the nonanticipativity requirements. Specifically, we have

$$\tilde{\boldsymbol{x}}(\boldsymbol{\xi})=\hat{\boldsymbol{x}}(\boldsymbol{h}(\boldsymbol{\xi}))=\left(\hat{\boldsymbol{x}}^1\left(\boldsymbol{h}^1(\underline{\boldsymbol{\xi}}^1)\right),\ldots,\hat{\boldsymbol{x}}^T\left(\boldsymbol{h}^T(\underline{\boldsymbol{\xi}}^T)\right)\right)$$

where the decisions in stage t depend only on the uncertainty revealed up to that stage. For the constraints, we find

$$\boldsymbol{A}\tilde{\boldsymbol{x}}(\boldsymbol{\xi}) = \boldsymbol{A}\hat{\boldsymbol{x}}(\boldsymbol{h}(\boldsymbol{\xi})) \overset{(a)}{\geq} \boldsymbol{D}\boldsymbol{h}(\boldsymbol{\xi}) + \boldsymbol{d} \overset{(b)}{\geq} \boldsymbol{D}\boldsymbol{\xi} + \boldsymbol{d}.$$

Here, (a) follows from the feasibility of $\hat{\boldsymbol{x}}$ and (b) follows from the definition of $\boldsymbol{h}$ and $\boldsymbol{D}$ being non-negative. Thus, $\tilde{\boldsymbol{x}}$ is a well-defined feasible solution for $Z_{AR}(\mathcal{U})$ and we have

$$Z_{AR}(\mathcal{U}) \leq \max_{\boldsymbol{\xi}\in\mathcal{U}} \boldsymbol{c}^{\mathsf{T}}\tilde{\boldsymbol{x}}(\boldsymbol{\xi}) \leq \max_{\hat{\boldsymbol{\xi}}\in\hat{\mathcal{U}}} \boldsymbol{c}^{\mathsf{T}}\hat{\boldsymbol{x}}(\hat{\boldsymbol{\xi}}) = Z_{AR}(\hat{\mathcal{U}}).$$

For the other direction let $\boldsymbol{x}^*$ be an optimal solution for $Z_{AR}(\mathcal{U})$. Then, for all $\hat{\boldsymbol{\xi}} \in \hat{\mathcal{U}}$ we have $\frac{1}{\beta}\hat{\boldsymbol{\xi}} \in \mathcal{U}$ by definition of β. We define $\tilde{\boldsymbol{x}}(\hat{\boldsymbol{\xi}}) := \beta\boldsymbol{x}^*\left(\frac{1}{\beta}\hat{\boldsymbol{\xi}}\right)$ and find

$$\boldsymbol{A}\tilde{\boldsymbol{x}}(\hat{\boldsymbol{\xi}}) = \beta\boldsymbol{A}\boldsymbol{x}^*\left(\frac{1}{\beta}\hat{\boldsymbol{\xi}}\right) \overset{(a)}{\geq} \beta\left(\frac{1}{\beta}\boldsymbol{D}\hat{\boldsymbol{\xi}} + \boldsymbol{d}\right) \overset{(b)}{\geq} \boldsymbol{D}\hat{\boldsymbol{\xi}} + \boldsymbol{d}$$

where (a) follows from the feasibility of $\boldsymbol{x}^*$ and (b) from the non-negativity of $\boldsymbol{d}$. Thus, $\tilde{\boldsymbol{x}}$ is a well-defined feasible solution for $Z_{AR}(\hat{\mathcal{U}})$ and we have

$$Z_{AR}(\hat{\mathcal{U}}) \leq \max_{\hat{\boldsymbol{\xi}}\in\hat{\mathcal{U}}} \boldsymbol{c}^{\mathsf{T}}\tilde{\boldsymbol{x}}(\hat{\boldsymbol{\xi}}) \leq \beta \max_{\boldsymbol{\xi}\in\mathcal{U}} \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x}^*(\boldsymbol{\xi}) = \beta Z_{AR}(\mathcal{U}).$$

Having shown both inequalities this concludes the proof. □

Appendix C Proof of Lemma 2

Proof We begin by showing $Z_{AR}(\hat{\mathcal{U}}) \geq Z_{LP}(\hat{\mathcal{U}})$. Let $\boldsymbol{x}$ be an optimal solution of $Z_{AR}(\hat{\mathcal{U}})$. Then, $(\boldsymbol{x}_i, z)$ is a valid solution for $Z_{LP}(\hat{\mathcal{U}})$, where $\boldsymbol{x}_i, z$ are defined by

$$\begin{aligned} \boldsymbol{x}_i &:= \boldsymbol{x}(\boldsymbol{v}_i) && \forall i \in \{0, \ldots, m\}, \\ z &:= \max_{i\in\{0,\ldots,m\}} c^{\mathsf{T}}\boldsymbol{x}_i. \end{aligned}$$

The first set of LP constraints (3b) holds by definition of z. The second set of constraints (3c) holds as $\boldsymbol{x}$ is a solution of $Z_{AR}(\hat{\mathcal{U}})$ and $\boldsymbol{v}_i \in \hat{\mathcal{U}}$, and the last set of constraints (3d) holds by nonanticipativity of $\boldsymbol{x}$ and the definition of $\boldsymbol{v}_i$. Furthermore, we find

$$Z_{LP}(\hat{\mathcal{U}}) \leq z = \max_{i\in\{0,\ldots,m\}} c^{\mathsf{T}}\boldsymbol{x}_i \leq \max_{\boldsymbol{\xi}\in\hat{\mathcal{U}}} c^{\mathsf{T}}\boldsymbol{x}(\boldsymbol{\xi}) = Z_{AR}(\hat{\mathcal{U}}).$$

For the other direction let $(\boldsymbol{x}_i, z)$ be an optimal solution for $Z_{LP}(\hat{\mathcal{U}})$. Define $\lambda_i(\boldsymbol{\xi})$ for each $\boldsymbol{\xi} \in \hat{\mathcal{U}}$ by:

$$\lambda_i(\boldsymbol{\xi}) := \frac{(\boldsymbol{\xi} - \boldsymbol{v}_0)_i}{\rho_i}.$$

For ease of notation we will in the following drop the explicit dependence on $\boldsymbol{\xi}$ and use λ_i. We directly find

$$\sum_{i=1}^{m} \lambda_i \leq 1$$

for all $\boldsymbol{\xi} \in \hat{\mathcal{U}}$, as by definition of $\hat{\mathcal{U}}$ each $\boldsymbol{\xi}$ is a convex combination of $\boldsymbol{v}_0, \ldots, \boldsymbol{v}_m$. Also, note that λ_i only depends on the ith component of the uncertainty vector $\boldsymbol{\xi}$. Using this we can define a nonanticipative decision vector $\boldsymbol{x}(\boldsymbol{\xi}) = \left(\boldsymbol{x}^1(\underline{\xi}^1), \ldots, \boldsymbol{x}^T(\underline{\xi}^T)\right)$ by

$$\boldsymbol{x}^t(\underline{\xi}^t) := \sum_{i \in \underline{I}^t} \lambda_i \boldsymbol{x}_i^t + \left(1 - \sum_{i \in \underline{I}^t} \lambda_i\right) \boldsymbol{x}_0^t$$

where $\underline{I}^t := \left\{i \in \{1, \ldots, m\} \middle| \underline{\boldsymbol{e}_i}^t \neq \boldsymbol{0}\right\}$ is the index set corresponding to the uncertainties up to stage t.

By the nonanticipativity constraints (3d) we have $\boldsymbol{x}_i^t = \boldsymbol{x}_0^t$ for all $1 \leq t \leq T$ and $i \in \{1, \ldots, m\} \setminus \underline{I}^t$. Thus,

$$\boldsymbol{x}^t(\underline{\xi}^t) = \sum_{i \in \underline{I}^t} \lambda_i \boldsymbol{x}_i^t + \left(1 - \sum_{i \in \underline{I}^t} \lambda_i\right) \boldsymbol{x}_0^t = \sum_{i=1}^{m} \lambda_i \boldsymbol{x}_i^t + \left(1 - \sum_{i=1}^{m} \lambda_i\right) \boldsymbol{x}_0^t,$$

which shows that $\boldsymbol{x}$ is a convex combination of $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m$ by

$$\boldsymbol{x}(\boldsymbol{\xi}) = \sum_{i=1}^{m} \lambda_i \boldsymbol{x}_i + \left(1 - \sum_{i=1}^{m} \lambda_i\right) \boldsymbol{x}_0.$$

Using this, we find

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x}(\boldsymbol{\xi}) &= \sum_{i=1}^{m} \lambda_i \boldsymbol{A}\boldsymbol{x}_i + \left(1 - \sum_{i=1}^{m} \lambda_i\right) \boldsymbol{A}\boldsymbol{x}_0 \\ &\overset{(a)}{\geq} \sum_{i=1}^{m} \lambda_i \boldsymbol{D}\boldsymbol{v}_i + \left(1 - \sum_{i=1}^{m} \lambda_i\right) \boldsymbol{D}\boldsymbol{v}_0 + \boldsymbol{d} \overset{(b)}{=} \boldsymbol{D}\boldsymbol{\xi} + \boldsymbol{d}. \end{aligned}$$

Here, (a) holds because all x_i are valid solutions for (3), and (b) holds as by definition of λ_i we have $\boldsymbol{\xi} = \sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i + \left(1 - \sum_{i=1}^{m} \lambda_i\right) \boldsymbol{v}_0$.

Now we find the missing inequality

$$Z_{AR}(\hat{\mathcal{U}}) \overset{(a)}{\leq} \max_{\boldsymbol{\xi}\in\hat{\mathcal{U}}} \boldsymbol{c}^\mathsf{T}\boldsymbol{x}(\boldsymbol{\xi}) = \max_{\boldsymbol{\xi}\in\hat{\mathcal{U}}} \sum_{i=1}^{m} \lambda_i \boldsymbol{c}^\mathsf{T}\boldsymbol{x}_i + \left(1 - \sum_{i=1}^{m}\lambda_i\right)\boldsymbol{c}^\mathsf{T}\boldsymbol{x}_0$$
$$\overset{(b)}{\leq} \max_{\boldsymbol{\xi}\in\hat{\mathcal{U}}} \max_{i\in\{0,\dots,m\}} \boldsymbol{c}^\mathsf{T}\boldsymbol{x}_i = \max_{i\in\{0,\dots,m\}} \boldsymbol{c}^\mathsf{T}\boldsymbol{x}_i = Z_{LP}(\hat{\mathcal{U}}),$$

where we use for (a) that $\boldsymbol{x}$ is a valid solution for $Z_{AR}(\hat{\mathcal{U}})$. For (b), we use that $\boldsymbol{x}$ is a convex combination of $\boldsymbol{x}_0, \dots, \boldsymbol{x}_m$ with convex coefficients $(1 - \sum_{i=1}^{m}\lambda_i), \lambda_1, \dots, \lambda_m$. The convex sum is upper bounded by its largest summand $\boldsymbol{c}^\mathsf{T}\boldsymbol{x}_i$. We can thus take the maximum over all these $\boldsymbol{c}^\mathsf{T}\boldsymbol{x}_i$, which equals the LP solution value $Z_{LP}(\hat{\mathcal{U}})$. □

Appendix D Proof of Lemma 3

Proof Let $\mathfrak{S}$ be the set of all permutations on $\{1, \dots, m\}$ and for any vector $\boldsymbol{\xi}$ let $\sigma(\boldsymbol{\xi})$ be the vector with components permuted according to σ. Let $\mathfrak{S}_{ji} \subseteq \mathfrak{S}$ be the subset of all permutations mapping component j to component i. Let $\hat{\mathcal{U}}$ be a dominating uncertainty set for $\mathcal{U}$ of the form in (2), such that the approximation factor β is minimal. Let $\boldsymbol{v}_0, \rho_1, \dots, \rho_m$ be the parameters defining the vertices of $\hat{\mathcal{U}}$. Define $\mu := \frac{1}{m}\boldsymbol{e}^\mathsf{T}\boldsymbol{v}_0$, $\rho := \frac{1}{m}\sum_{i=1}^{m}\rho_i$. Then for each $i \in \{1, \dots, m\}$

$$\mu\boldsymbol{e} + \rho\boldsymbol{e}_i \overset{(a)}{=} \frac{1}{|\mathfrak{S}|}\sum_{\sigma\in\mathfrak{S}}\sigma(\boldsymbol{v}_0) + \frac{1}{m}\sum_{j=1}^{m}\rho_j\boldsymbol{e}_i \overset{(b)}{=} \sum_{j=1}^{m}\left(\frac{1}{|\mathfrak{S}|}\sum_{\sigma\in\mathfrak{S}_{ji}}\sigma(\boldsymbol{v}_0) + \frac{1}{m}\rho_j\boldsymbol{e}_i\right)$$
$$\overset{(c)}{=} \sum_{j=1}^{m}\left(\frac{1}{|\mathfrak{S}|}\sum_{\sigma\in\mathfrak{S}_{ji}}\sigma(\boldsymbol{v}_0) + \frac{1}{m|\mathfrak{S}_{ji}|}\sum_{\sigma\in\mathfrak{S}_{ji}}\rho_j\boldsymbol{e}_i\right)$$
$$\overset{(d)}{=} \frac{1}{|\mathfrak{S}|}\sum_{j=1}^{m}\sum_{\sigma\in\mathfrak{S}_{ji}}\left(\sigma(\boldsymbol{v}_0) + \rho_j\boldsymbol{e}_i\right) \overset{(e)}{=} \frac{1}{|\mathfrak{S}|}\sum_{j=1}^{m}\sum_{\sigma\in\mathfrak{S}_{ji}}\sigma\left(\boldsymbol{v}_0 + \rho_j\boldsymbol{e}_j\right),$$

where (a) follows from symmetry and the invariance of the average component under permutations, (b) follows from $\mathfrak{S} = \bigcup_{j=1}^{m}\mathfrak{S}_{ji}$ for all i, (c) follows from $\sum_{\sigma\in\mathfrak{S}_{ji}}\frac{1}{|\mathfrak{S}_{ji}|} = 1$, (d) follows from $|\mathfrak{S}_{ji}| = \frac{1}{m}|\mathfrak{S}|$, and (e) follows from the linearity of permutations. For each $i \in \{1, \dots, m\}$ we now conclude

$$\forall j \in \{1, \dots, m\}\colon \frac{1}{\beta}\left(\boldsymbol{v}_0 + \rho_j\boldsymbol{e}_j\right) \overset{(a)}{\in} \mathcal{U}$$
$$\Rightarrow \forall j \in \{1, \dots, m\}\colon \sigma\left(\frac{1}{\beta}\left(\boldsymbol{v}_0 + \rho_j\boldsymbol{e}_j\right)\right) \overset{(b)}{\in} \mathcal{U}$$
$$\Rightarrow \frac{1}{\beta}\left(\mu\boldsymbol{e} + \rho\boldsymbol{e}_i\right) = \frac{1}{|\mathfrak{S}|}\sum_{j=1}^{m}\sum_{\sigma\in\mathfrak{S}_{ji}}\sigma\left(\frac{1}{\beta}\left(\boldsymbol{v}_0 + \rho_j\boldsymbol{e}_j\right)\right) \overset{(c)}{\in} \mathcal{U}.$$

Here (a) follows from $\hat{\mathcal{U}}$ being a valid dominating set with approximation factor β, (b) follows from permutation invariance of $\mathcal{U}$, and (c) from convexity of $\mathcal{U}$. Furthermore, $\frac{1}{\beta}\mu\boldsymbol{e} \in \mathcal{U}$ as $\mathcal{U}$ is down-monotone. Thus the convex set induced by vertices $\mu\boldsymbol{e}, \mu\boldsymbol{e} + \rho\boldsymbol{e}_1, \ldots, \mu\boldsymbol{e} + \rho\boldsymbol{e}_m$ is also a valid dominating set for $\mathcal{U}$ with approximation factor β. □

Appendix E Proof of Proposition 5

Proof Our main idea is to find μ, ρ such that (7) is always fulfilled and β is minimized. Recall that β is given by the minimal value such that $\frac{1}{\beta}\boldsymbol{v}_i \in \mathcal{U}$ for all $\boldsymbol{v}_i$. In the case of the hypersphere uncertainty set, this is given when $\|\boldsymbol{v}_i\|_2^2 \leq \beta^2$. As for each $i \in \{1, \ldots, m\}$, $\boldsymbol{v}_i$ is given by $\boldsymbol{v}_i = \boldsymbol{v}_0 + \rho\boldsymbol{e}_i$ and $\boldsymbol{v}_0 = \mu\boldsymbol{e}$ we find:

$$\|\boldsymbol{v}_0\|_2^2 \leq \|\boldsymbol{v}_i\|_2^2 = (m-1)\mu^2 + (\mu+\rho)^2 = m\mu^2 + 2\mu\rho + \rho^2. \tag{E1}$$

This is the term for β^2 that we want to minimize in the following.

Using Lemma 4 the left-hand side in (7) becomes:

$$\begin{aligned}\frac{1}{\rho}\max_{j\in\{1,\ldots,m\}}\sum_{i=1}^{j}(\gamma(j)-\mu)_+ &\overset{(a)}{=} \frac{1}{\rho}\max_{j\in\{1,\ldots,m\}} j(\gamma(j)-\mu)_+ \\ &\overset{(b)}{=} \frac{1}{\rho}\max_{j\in\{1,\ldots,m\}} j(\gamma(j)-\mu).\end{aligned}$$

For (a) we use that all summands are the same and for (b) we assume w.l.o.g. that $\max_{j\in\{1,\ldots,m\}}(\gamma(j)-\mu) \geq 0$, as otherwise (7) trivially holds.

Using the property $\|\boldsymbol{\xi}\|_2^2 \leq 1$ of the hypersphere uncertainty set, we find $\gamma(j) = \frac{1}{\sqrt{j}}$ for all j. Substituting $\gamma(j)$ in the above equation we find Property (7) to become:

$$\max_{j\in\{1,\ldots,m\}} \sqrt{j} - j\mu \leq \rho.$$

The maximum of the left hand side is taken for $j = \frac{1}{4\mu^2}$ and for any optimal choice of μ, ρ the inequality will be tight. Thus, we find $\rho = \frac{1}{4\mu}$.

Substituting ρ in Eq. (E1), we get:

$$\beta = \sqrt{m\mu^2 + \frac{1}{2} + \frac{1}{16\mu^2}}.$$

This is minimized by $\mu = \frac{1}{2\sqrt[4]{m}}$, which gives $\rho = \frac{\sqrt[4]{m}}{2}$ and concludes the proof. □

Appendix F Proof of Proposition 6

Proof Consider the following two-stage instance of Problem (1):

$$\begin{aligned} Z(\mathcal{U}) = \min_{\alpha, \boldsymbol{x}(\boldsymbol{\xi})} \max_{\boldsymbol{\xi} \in \mathcal{U}} \quad & \sqrt{m}\alpha + \boldsymbol{e}^\mathsf{T}\boldsymbol{x}(\boldsymbol{\xi}) \\ \text{s.t.} \quad & \alpha \boldsymbol{e} + \boldsymbol{x} \geq \boldsymbol{\xi} \qquad \forall \boldsymbol{\xi} \in \mathcal{U} \\ & \alpha, \boldsymbol{x} \geq 0, \end{aligned}$$

where $\mathcal{U} = \{\boldsymbol{\xi} | \, \|\boldsymbol{\xi}\|_2^2 \leq 1\} \subset [0,1]^m$ is the hypersphere uncertainty set, $\alpha \in \mathbb{R}$ is the first stage decision variable and $\boldsymbol{x} \in \mathbb{R}^m$ are the second stage decisions that may depend on the uncertainty realization $\boldsymbol{\xi}$.

We begin by finding a lower bound for affine policies on this problem class. By symmetry of the problem there always exists an optimal affine adjustable solution $\boldsymbol{x}(\boldsymbol{\xi})$ of the form

$$x_i = a\xi_i + b\sum_{j \neq i} \xi_j + c,$$

for some $a, b, c \in \mathbb{R}$. To see this let $\boldsymbol{x}^*, \alpha$ be an optimal affine solution to the above problem and define

$$\boldsymbol{x}(\boldsymbol{\xi}) := \frac{1}{|\mathfrak{S}|} \sum_{\sigma \in \mathfrak{S}} \sigma \left(\boldsymbol{x}^* \left(\sigma^{-1}(\boldsymbol{\xi}) \right) \right),$$

where $\mathfrak{S}$ is the set of all permutations on m elements. Then

$$\begin{aligned} \alpha \boldsymbol{e} + \boldsymbol{x}(\boldsymbol{\xi}) = \alpha \boldsymbol{e} + \frac{1}{|\mathfrak{S}|} \sum_{\sigma \in \mathfrak{S}} \sigma \left(\boldsymbol{x}^* \left(\sigma^{-1}(\boldsymbol{\xi}) \right) \right) &\overset{(a)}{=} \frac{1}{|\mathfrak{S}|} \sum_{\sigma \in \mathfrak{S}} \sigma \left(\alpha \boldsymbol{e} + \boldsymbol{x}^* \left(\sigma^{-1}(\boldsymbol{\xi}) \right) \right) \\ &\overset{(b)}{\geq} \frac{1}{|\mathfrak{S}|} \sum_{\sigma \in \mathfrak{S}} \sigma \left(\sigma^{-1}(\boldsymbol{\xi}) \right) = \boldsymbol{\xi}, \end{aligned}$$

where (a) follows from the permutation invariance of $\alpha \boldsymbol{e}$ and (b) follows from permutation invariance of $\mathcal{U}$ and feasibility of $\boldsymbol{x}^*$. Furthermore,

$$\begin{aligned} \max_{\boldsymbol{\xi} \in \mathcal{U}} \boldsymbol{e}^\mathsf{T}\boldsymbol{x}(\boldsymbol{\xi}) = \frac{1}{|\mathfrak{S}|} \max_{\boldsymbol{\xi} \in \mathcal{U}} \boldsymbol{e}^\mathsf{T} \sum_{\sigma \in \mathfrak{S}} \sigma \left(\boldsymbol{x}^* \left(\sigma^{-1}(\boldsymbol{\xi}) \right) \right) &\overset{(a)}{=} \frac{1}{|\mathfrak{S}|} \max_{\boldsymbol{\xi} \in \mathcal{U}} \sum_{\sigma \in \mathfrak{S}} \boldsymbol{e}^\mathsf{T} \boldsymbol{x}^* \left(\sigma^{-1}(\boldsymbol{\xi}) \right) \\ &\overset{(b)}{\leq} \frac{1}{|\mathfrak{S}|} \sum_{\sigma \in \mathfrak{S}} \max_{\boldsymbol{\xi} \in \mathcal{U}} \boldsymbol{e}^\mathsf{T} \boldsymbol{x}^* \left(\sigma^{-1}(\boldsymbol{\xi}) \right) \overset{(c)}{=} \frac{1}{|\mathfrak{S}|} \sum_{\sigma \in \mathfrak{S}} \max_{\boldsymbol{\xi} \in \mathcal{U}} \boldsymbol{e}^\mathsf{T} \boldsymbol{x}^* (\boldsymbol{\xi}) = \max_{\boldsymbol{\xi} \in \mathcal{U}} \boldsymbol{e}^\mathsf{T} \boldsymbol{x}^* (\boldsymbol{\xi}). \end{aligned}$$

Here, (a) follows from permutation invariance of $\boldsymbol{e}$, (b) follows from the subadditivity of a maximization function, and (c) follows from permutation invariance of $\mathcal{U}$.

Non-negativity of $\boldsymbol{x}$ follows from $\mathbf{0} \in \mathcal{U}$. Thus $\boldsymbol{x}, \alpha$ is an optimal feasible solution. Additionally for any $\overline{\sigma} \in \mathfrak{S}$,

$$\begin{aligned}\boldsymbol{x}(\overline{\sigma}(\boldsymbol{\xi})) &= \frac{1}{|\mathfrak{S}|}\sum_{\sigma\in\mathfrak{S}} \sigma\left(\boldsymbol{x}^*\left(\sigma^{-1}\circ\overline{\sigma}(\boldsymbol{\xi})\right)\right) = \overline{\sigma}\left(\frac{1}{|\mathfrak{S}|}\sum_{\sigma\in\mathfrak{S}} \overline{\sigma}^{-1}\circ\sigma\left(\boldsymbol{x}^*\left(\sigma^{-1}\circ\overline{\sigma}(\boldsymbol{\xi})\right)\right)\right)\\ &\overset{(a)}{=} \overline{\sigma}\left(\frac{1}{|\mathfrak{S}|}\sum_{\sigma\in\mathfrak{S}} \sigma\left(\boldsymbol{x}^*\left(\sigma^{-1}(\boldsymbol{\xi})\right)\right)\right) = \overline{\sigma}(\boldsymbol{x}(\boldsymbol{\xi})),\end{aligned}$$

where (a) follows from $(\overline{\sigma}^{-1}\circ\cdot)$ being an automorphism on $\mathfrak{S}$. Let σ_{ij} be the permutation switching components i and j. Then $\boldsymbol{x}(\sigma_{ij}(\mathbf{0})) = \sigma_{ij}(\boldsymbol{x}(\mathbf{0}))$ implies $x_i(\mathbf{0}) = x_j(\mathbf{0})$, $\boldsymbol{x}(\sigma_{ij}(\boldsymbol{e}_j)) = \sigma_{ij}(\boldsymbol{x}(\boldsymbol{e}_j))$ implies $x_i(\boldsymbol{e}_i) = x_j(\boldsymbol{e}_j)$ and $x_k(\boldsymbol{e}_i) = x_k(\boldsymbol{e}_j)$ for $i \neq k \neq j$, and $\boldsymbol{x}(\sigma_{ij}(\boldsymbol{e}_k)) = \sigma_{ij}(\boldsymbol{x}(\boldsymbol{e}_k))$ for $i \neq k \neq j$ implies $x_i(\boldsymbol{e}_k) = x_j(\boldsymbol{e}_k)$. This guarantees the claimed structure.

Let $\boldsymbol{\xi} \in \mathcal{U}$ and for some $i \leq m$ define the vector $\boldsymbol{\xi}'$ by $\boldsymbol{\xi}' := \boldsymbol{\xi} - \xi_i \boldsymbol{e}_i$. Then $\boldsymbol{\xi}'$ is also in $\mathcal{U}$ and by $x_i(\boldsymbol{\xi}') \geq 0$ we find that every feasible solution has to fulfill

$$b\sum_{j\neq i}\xi_j + c \geq 0 \tag{F1}$$

for all $\boldsymbol{\xi} \in \mathcal{U}$. Additionally, by $\alpha\boldsymbol{e} + \boldsymbol{x}(\boldsymbol{e}_i) \geq \boldsymbol{e}_i$ every feasible solution also fulfills

$$\alpha + a + c \geq 1. \tag{F2}$$

Using this we obtain a lower bound for the maximum over $\mathcal{U}$ in the objective function as follows

$$\begin{aligned}&\max_{\boldsymbol{\xi}\in\mathcal{U}} \sqrt{m}\alpha + \sum_i \left(a\xi_i + b\sum_{j\neq i}\xi_j + c\right)\\ &\overset{(a)}{\geq} \max_{\boldsymbol{\xi}\in\mathcal{U}} \sqrt{m}\alpha + a\sum_i \xi_i = \max_{\boldsymbol{\xi}\in\mathcal{U}} \sqrt{m}\alpha + a\boldsymbol{e}^\intercal\boldsymbol{\xi}\\ &\overset{(b)}{=} \sqrt{m}\alpha + \sqrt{m}a \overset{(c)}{\geq} \sqrt{m}\alpha + \sqrt{m}(1-\alpha-c) = \sqrt{m}(1-c).\end{aligned}$$

Here (a) follows from Eq. (F1), (b) follows as the maximum is taken for $\boldsymbol{\xi} = \frac{1}{\sqrt{m}}\boldsymbol{e}$, and (c) follows from Eq. (F2). On the other hand, using $\mathbf{0} \in \mathcal{U}$ the maximum can also be lower bounded by

$$\sqrt{m}\alpha + \sum_i c = \sqrt{m}\alpha + mc \geq mc.$$

Thus we find

$$Z_{AFF}(\mathcal{U}) = \min_{\alpha,a,b,c} \max_{\boldsymbol{\xi}\in\mathcal{U}} \sqrt{m}\alpha + \sum_i \left(a\xi_i + b\sum_{j\neq i} \xi_j + c \right)$$
$$\geq \min_c \max\left(\sqrt{m}(1-c), mc\right) \overset{(a)}{=} \sqrt{m} - \frac{m}{m+\sqrt{m}} \geq \sqrt{m} - 1,$$

where (a) holds as the maximum is taken at equality of the two terms, which is given by $c = \frac{\sqrt{m}}{m+\sqrt{m}}$.

Having found a lower bound for affine policies, we will now give an upper bound for optimal policies. Consider the policy given by $\alpha = \frac{1}{\sqrt[4]{m}}$ and $\boldsymbol{x} = (\boldsymbol{\xi} - \alpha \boldsymbol{e})_+$. By construction this is feasible and we find

$$\begin{aligned} Z_{OPT}(\mathcal{U}) &\leq \max_{\boldsymbol{\xi}\in\mathcal{U}} \sqrt{m}\alpha + \boldsymbol{e}^\mathsf{T}(\boldsymbol{\xi} - \alpha\boldsymbol{e})_+ \\ &\overset{(a)}{=} \max_{0\leq k\leq m} \sqrt[4]{m} + k\left(\frac{1}{\sqrt{k}} - \frac{1}{\sqrt[4]{m}}\right) \\ &= \max_{0\leq k\leq m} \sqrt[4]{m} + \sqrt{k} - \frac{k}{\sqrt[4]{m}} \overset{(b)}{=} \frac{5}{4}\sqrt[4]{m}. \end{aligned}$$

Here (a) follows from Lemma 4 and (b) follows as the maximum over k is taken for $k = \frac{\sqrt{m}}{4}$. To finish the proof, we find the optimality ratio, which is given by

$$\frac{Z_{AFF}(\mathcal{U})}{Z_{OPT}(\mathcal{U})} \geq \frac{\sqrt{m}-1}{5\sqrt[4]{m}/4} = \frac{4}{5}\left(\sqrt[4]{m} - \frac{1}{\sqrt[4]{m}}\right).$$

□

Appendix G Proof of Proposition 7

Proof We proceed analogously as in the proof for Proposition 5. Again, the idea is to find μ, ρ such that (7) is always fulfilled and β is minimized. Using the definition of $\mathcal{U}$ and $\boldsymbol{v}_i$ we find $\frac{1}{\beta}\boldsymbol{v}_i \in \mathcal{U}$ to be fulfilled when

$$\beta = \min\left(\frac{m}{k}\mu + \frac{1}{k}\rho, \mu + \rho\right). \tag{G1}$$

Here the first term of the maximization follows from $\frac{1}{\beta}\boldsymbol{e}^\mathsf{T}\boldsymbol{v}_i \leq k$ and the second term from $\frac{1}{\beta}\boldsymbol{e}_i^\mathsf{T}\boldsymbol{v}_i \leq 1$. By Lemma 4 the left hand side in (7) becomes:

$$\frac{1}{\rho} \max_{j\in\{1,\ldots,m\}} j(\gamma(j) - \mu).$$

Using the properties $\|\boldsymbol{\xi}\|_1 \leq k$ and $\xi_i \leq 1$ of the budgeted uncertainty set we find $\gamma(j) = \frac{k}{j}$ for all $j \geq k$ and $\gamma(j) = 1$ for all $j \leq k$. Inserting $\gamma(j)$ into the above, we find Property (7) to become:

$$\max_{j\in\{1,\dots,m\}} j\left(\min\left(1, \frac{k}{j}\right) - \mu\right) \leq \rho.$$

As for $\mu \geq 1$ the vector $\boldsymbol{v}_0$ would already dominate every uncertainty realization in $\mathcal{U}$, we can assume $\mu \leq 1$. With this, the maximum on the left-hand side is taken for $j = k$ and we find the minimal possible ρ to be $\rho = k - k\mu$.

Using this ρ together with β from Eq. (G1), we obtain:

$$\beta = \min\left(\left(\frac{m}{k} - 1\right)\mu + 1, k - (k-1)\mu\right).$$

As the left term in the minimization is linearly increasing in μ and the right term is linearly decreasing in μ, the minimum is taken at equality. This is the case when $\mu = \frac{k(k-1)}{m+k(k-2)}$, which gives $\rho = \frac{k(m-k)}{m+k(k-2)}$. □

Appendix H Proof of Proposition 8

To prove Proposition 8 we first introduce a more general form of piecewise affine policies. Recall that affine policies are given by $\boldsymbol{x}^t(\underline{\boldsymbol{\xi}}^t) = \boldsymbol{P}^t\underline{\boldsymbol{\xi}}^t + \boldsymbol{q}^t$, with decision variables $\boldsymbol{P}^t$ and $\boldsymbol{q}^t$. By introducing additional decision variables $\boldsymbol{\zeta}$ of the same dimension as the uncertainty set, we extend affine policies to *fully piecewise affine policies* (FPAP)

$$\boldsymbol{x}^t(\underline{\boldsymbol{\xi}}^t) = \boldsymbol{P}^t\left(\underline{\boldsymbol{\xi}}^t - \underline{\boldsymbol{\zeta}}^t\right)_+ + \boldsymbol{q}^t,$$

where we use the same notation for $\boldsymbol{\zeta}$, as for $\boldsymbol{\xi}$. By extending the decision matrices $\boldsymbol{P}^t$ with zeroes and concatenating them to a single decision matrix

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{P}^1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ & \boldsymbol{P}^2 & \mathbf{0} & \cdots & \mathbf{0} \\ & & \vdots & & \\ & & \boldsymbol{P}^T & & \end{pmatrix},$$

we get the compact expressions $\boldsymbol{x}(\boldsymbol{\xi}) = \boldsymbol{P}\boldsymbol{\xi} + \boldsymbol{q}$ and $\boldsymbol{x}(\boldsymbol{\xi}) = \boldsymbol{P}\,(\boldsymbol{\xi} - \boldsymbol{\zeta})_+ + \boldsymbol{q}$ for affine and fully piecewise affine policies, respectively.

By our initial assumption $\mathcal{U} \subseteq [0,1]^m$, we can w.l.o.g. assume $\mathbf{0} \leq \boldsymbol{\zeta} \leq \boldsymbol{e}$. First, consider a policy $\boldsymbol{P}, \boldsymbol{\zeta}, \boldsymbol{q}$ with $\zeta_i < 0$. Then replacing ζ_i by 0 and $\boldsymbol{q}$ by $\boldsymbol{q}' := \boldsymbol{q} - \zeta_i \boldsymbol{P}_{:i}$, where $\boldsymbol{P}_{:i}$ is the ith column of $\boldsymbol{P}$, yields an identical policy on the uncertainty set, as $(\xi_i - \zeta_i)_+ = \xi_i - \zeta_i$ for all $\boldsymbol{\xi} \in \mathcal{U}$. Similarly, if $\zeta_i > 1$, we can replace it by 1 without changing the policy, as $(\xi_i - \zeta_i)_+ = (\xi_i - 1)_+ = 0$ for all $\boldsymbol{\xi} \in \mathcal{U}$.

We can show that the fully piecewise affine adjustable policies defined above are in fact a generalization of the piecewise affine policies we introduce in this work.

Lemma 16 *Consider Problem (1). Let Z_{PAP} be the optimal value found by our piecewise affine policy and let Z_{FPAP} be the optimal value found by a fully piecewise affine policy where the right hand side uncertainty is replaced by $\boldsymbol{D}\left((\boldsymbol{\xi}-\boldsymbol{\zeta})_+ + \boldsymbol{\zeta}\right) + \boldsymbol{d}$. Then*

$$Z_{FPAP} \leq Z_{PAP}.$$

Proof We prove the inequality by showing that every feasible piecewise affine policy is a feasible fully piecewise affine policy. Let $\boldsymbol{v}_0, \ldots, \boldsymbol{v}_m$ be the chosen vertices of the piecewise affine policy and let $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m$ be the vertex solutions received from solving LP (3). Then the piecewise affine policy is given by Eq. (6) and we find

$$\begin{aligned}
\boldsymbol{x}(\boldsymbol{\xi}) &= \sum_{i=1}^{m} \lambda_i(\boldsymbol{\xi})\boldsymbol{x}_i + \left(1 - \sum_{i=1}^{m} \lambda_i(\boldsymbol{\xi})\right)\boldsymbol{x}_0 \\
&= \sum_{i=1}^{m} \lambda_i(\boldsymbol{\xi})\left(\boldsymbol{x}_i - \boldsymbol{x}_0\right) + \boldsymbol{x}_0 \\
&\overset{(a)}{=} \sum_{i=1}^{m} \frac{\left((\boldsymbol{\xi}-\boldsymbol{v}_0)_+\right)_i}{\rho_i}\left(\boldsymbol{x}_i - \boldsymbol{x}_0\right) + \boldsymbol{x}_0 \\
&\overset{(b)}{=} \left(\frac{1}{\rho_1}\left(\boldsymbol{x}_1 - \boldsymbol{x}_0\right), \ldots, \frac{1}{\rho_m}\left(\boldsymbol{x}_m - \boldsymbol{x}_0\right)\right)(\boldsymbol{\xi}-\boldsymbol{v}_0)_+ + \boldsymbol{x}_0.
\end{aligned}$$

Here (a) follows from the definition of λ_i in Sect. 2.2 and (b) follows by replacing the sum with a matrix–vector multiplication. By the nonanticipativity Constraint (3d) the vector $\boldsymbol{x}_i - \boldsymbol{x}_0$ is zero for all uncertainty dimensions associated to later stages. Thus setting

$$\begin{aligned}
\boldsymbol{P} &:= \left(\frac{1}{\rho_1}\left(\boldsymbol{x}_1 - \boldsymbol{x}_0\right), \ldots, \frac{1}{\rho_m}\left(\boldsymbol{x}_m - \boldsymbol{x}_0\right)\right) \\
\boldsymbol{\zeta} &:= \boldsymbol{v}_0 \\
\boldsymbol{q} &:= \boldsymbol{x}_0
\end{aligned}$$

yields a valid fully piecewise affine policy that is identical to the initial piecewise affine policy. Even with the modified right hand side, as demanded in the Lemma, the constructed fully piecewise affine policy fulfills all constraints, as by construction of the policy, and definition of the dominating function $\boldsymbol{h}$:

$$\boldsymbol{A}\boldsymbol{x}(\boldsymbol{\xi}) \geq \boldsymbol{D}\boldsymbol{h}(\boldsymbol{\xi}) + \boldsymbol{d} = \boldsymbol{D}\left((\boldsymbol{\xi}-\boldsymbol{\zeta})_+ + \boldsymbol{\zeta}\right) + \boldsymbol{d}.$$

□

Even though fully piecewise affine policies are slightly more flexible than affine policies in general, both are equivalent on Problem (1) under some additional assumptions.

Lemma 17 *Consider Problem (1) with budgeted uncertainty and integer budget. Let Z_{FPAP} be the optimal value found by a fully piecewise affine policy with modified right hand side uncertainty as in Lemma 16 and let Z_{AFF} be the optimal value found by an affine policy. Then*

$$Z_{AFF} = Z_{FPAP}.$$

For the proof of Lemma 17 we first need the following helpful result that allows us to construct a robust counterpart of Problem (1) with fully piecewise affine adjustable policies in a similar manner as robust counterparts for affine policies are constructed.

Lemma 18 *Let $\boldsymbol{c} \in \mathbb{R}^m$, $\boldsymbol{\zeta} \in [0,1]^m$ and $k \in \mathbb{N}_+$. Then the following two problems yield the same objective.*

$$\begin{aligned} \max_{\boldsymbol{\xi}} \quad & \boldsymbol{c}^\intercal (\boldsymbol{\xi} - \boldsymbol{\zeta})_+ \\ \text{s.t.} \quad & \boldsymbol{e}^\intercal \boldsymbol{\xi} \leq k \\ & \boldsymbol{0} \leq \boldsymbol{\xi} \leq \boldsymbol{e} \end{aligned} \tag{H1}$$

and

$$\begin{aligned} \min_{\alpha, \boldsymbol{\beta}} \quad & \boldsymbol{e}^\intercal \boldsymbol{\beta} + k\alpha \\ \text{s.t.} \quad & \boldsymbol{\beta} + \alpha \boldsymbol{e} \geq \operatorname{diag}(\boldsymbol{e} - \boldsymbol{\zeta})\boldsymbol{c} \\ & \alpha \geq 0, \quad \boldsymbol{\beta} \geq \boldsymbol{0}. \end{aligned} \tag{H2}$$

Proof Let I^k be the indices of the largest k values $c_i(1-\zeta_i)$ and let $\boldsymbol{e}_{I^k}$ be the indicator vector of I^k. Then for Problem (H1) we find the following set of inequalities.

$$\max \left\{ \boldsymbol{c}^\intercal (\boldsymbol{\xi} - \boldsymbol{\zeta})_+ \middle| \boldsymbol{e}^\intercal \boldsymbol{\xi} \leq k, \boldsymbol{0} \leq \boldsymbol{\xi} \leq \boldsymbol{e} \right\} \tag{H3}$$

$$\overset{(a)}{\leq} \max \left\{ \boldsymbol{c}^\intercal \operatorname{diag}(\boldsymbol{e} - \boldsymbol{\zeta})\, \boldsymbol{\xi} \middle| \boldsymbol{e}^\intercal \boldsymbol{\xi} \leq k, \boldsymbol{0} \leq \boldsymbol{\xi} \leq \boldsymbol{e} \right\} \tag{H4}$$

$$\overset{(b)}{\leq} \boldsymbol{c}^\intercal \operatorname{diag}(\boldsymbol{e} - \boldsymbol{\zeta})\, \boldsymbol{e}_{I^k}$$

$$\overset{(c)}{\leq} \boldsymbol{c}^\intercal \left(\boldsymbol{e}_{I^k} - \boldsymbol{\zeta}\right)_+$$

$$\overset{(d)}{\leq} \max \left\{ \boldsymbol{c}^\intercal (\boldsymbol{\xi} - \boldsymbol{\zeta})_+ \middle| \boldsymbol{e}^\intercal \boldsymbol{\xi} \leq k, \boldsymbol{0} \leq \boldsymbol{\xi} \leq \boldsymbol{e} \right\} \tag{H5}$$

Here (a) follows from $(\boldsymbol{\xi} - \boldsymbol{\zeta})_+ \leq \operatorname{diag}(\boldsymbol{e} - \boldsymbol{\zeta})\, \boldsymbol{\xi}$. Inequality (b) follows as e_{I^k} is an optimal choice for Problem (H4). For (c), observe that $\operatorname{diag}(\boldsymbol{e} - \boldsymbol{\zeta})\, \boldsymbol{e}_{I^k} = \left(\boldsymbol{e}_{I^k} - \boldsymbol{\zeta}\right)_+$. Finally, (d) follows, as $\boldsymbol{e}_{I^k}$ is a feasible choice for Problem (H5). As (H3) and (H5) are the same problem all inequalities are in fact equalities.

To conclude the proof note that Problem (H4) is a linear optimization problem, where taking the dual yields Problem (H2). □

Having this result, we now show Lemma 17.

Proof of Lemma 17 To prove this result we show that the robust counterparts of the two policies on Problem (1) with integer budget uncertainty are equivalent. Let $k \in \mathbb{N}_+$ be the budget.

Then Problem (1) with affine policies becomes

$$\begin{aligned} Z_{AFF}(\mathcal{U}) = \min_{\boldsymbol{P},\boldsymbol{q}} \quad & \max_{\boldsymbol{\xi}} \left(\boldsymbol{c}^\mathsf{T}\boldsymbol{P}\boldsymbol{\xi} \middle| \boldsymbol{0} \le \boldsymbol{\xi} \le \boldsymbol{e}, \boldsymbol{e}^\mathsf{T}\boldsymbol{\xi} \le k\right) + \boldsymbol{c}^\mathsf{T}\boldsymbol{q} \\ \text{s.t.} \quad & \max_{\boldsymbol{\xi}} \left((\boldsymbol{D}_i - \boldsymbol{A}_i\boldsymbol{P})\,\boldsymbol{\xi} \middle| \boldsymbol{0} \le \boldsymbol{\xi} \le \boldsymbol{e}, \boldsymbol{e}^\mathsf{T}\boldsymbol{\xi} \le k\right) \le \boldsymbol{A}_i\boldsymbol{q} - d_i \quad \forall i \in \{1,\dots,l\}, \end{aligned}$$

where $\boldsymbol{A}_i$ and $\boldsymbol{D}_i$ are the ith rows of $\boldsymbol{A}$ and $\boldsymbol{D}$, respectively, and d_i is the ith entry of $\boldsymbol{d}$. Dualizing the suproblems we find this to be equivalent to

$$\begin{aligned} Z_{AFF}(\mathcal{U}) = \min_{\boldsymbol{P},\boldsymbol{q},\alpha^0,\dots,\alpha^l,\boldsymbol{\beta}^0,\dots,\boldsymbol{\beta}^l} \quad & \boldsymbol{e}^\mathsf{T}\boldsymbol{\beta}^0 + k\alpha^0 + \boldsymbol{c}^\mathsf{T}\boldsymbol{q} \\ \text{s.t.} \quad & \boldsymbol{\beta}^0 + \alpha^0\boldsymbol{e} \ge \left(\boldsymbol{c}^\mathsf{T}\boldsymbol{P}\right)^\mathsf{T} \\ & \boldsymbol{e}^\mathsf{T}\boldsymbol{\beta}^i + k\alpha^i \le \boldsymbol{A}_i\boldsymbol{q} - d_i && \forall i \in \{1,\dots,l\} \\ & \boldsymbol{\beta}^i + \alpha^i\boldsymbol{e} \ge (\boldsymbol{D}_i - \boldsymbol{A}_i\boldsymbol{P})^\mathsf{T} && \forall i \in \{1,\dots,l\} \\ & \alpha^i \ge 0, \quad \boldsymbol{\beta}^i \ge \boldsymbol{0} && \forall i \in \{1,\dots,l\}. \end{aligned} \tag{H6}$$

Similarly Problem (1) with fully piecewise affine policies and modified right hand side becomes

$$\begin{aligned} Z_{FPAP}(\mathcal{U}) = \min_{\boldsymbol{P},\boldsymbol{q},\boldsymbol{\zeta}} \quad & \max_{\boldsymbol{\xi}} \left(\boldsymbol{c}^\mathsf{T}\boldsymbol{P}(\boldsymbol{\xi} - \boldsymbol{\zeta})_+ \middle| \boldsymbol{0} \le \boldsymbol{\xi} \le \boldsymbol{e}, \boldsymbol{e}^\mathsf{T}\boldsymbol{\xi} \le k\right) + \boldsymbol{c}^\mathsf{T}\boldsymbol{q} \\ \text{s.t.} \quad & \max_{\boldsymbol{\xi}} \left((\boldsymbol{D}_i - \boldsymbol{A}_i\boldsymbol{P})\,(\boldsymbol{\xi} - \boldsymbol{\zeta})_+ \middle| \boldsymbol{0} \le \boldsymbol{\xi} \le \boldsymbol{e}, \boldsymbol{e}^\mathsf{T}\boldsymbol{\xi} \le k\right) \\ & \qquad \le \boldsymbol{A}_i\boldsymbol{q} - \boldsymbol{D}_i\boldsymbol{\zeta} - d_i && \forall i \in \{1,\dots,l\} \\ & \boldsymbol{0} \le \boldsymbol{\zeta} \le \boldsymbol{e}. \end{aligned}$$

Using Lemma 18 on the subproblems we find this to be equivalent to

$$\begin{aligned} Z_{FPAP}(\mathcal{U}) = \min_{\boldsymbol{P},\boldsymbol{q},\boldsymbol{\zeta},\alpha^0,\dots,\alpha^l,\boldsymbol{\beta}^0,\dots,\boldsymbol{\beta}^l} \quad & \boldsymbol{e}^\mathsf{T}\boldsymbol{\beta}^0 + k\alpha^0 + \boldsymbol{c}^\mathsf{T}\boldsymbol{q} && \text{(H7a)} \\ \text{s.t.} \quad & \boldsymbol{\beta}^0 + \alpha^0\boldsymbol{e} \ge \left(\boldsymbol{c}^\mathsf{T}\boldsymbol{P}\,\mathrm{diag}(\boldsymbol{e} - \boldsymbol{\zeta})\right)^\mathsf{T} && \text{(H7b)} \\ & \boldsymbol{e}^\mathsf{T}\boldsymbol{\beta}^i + k\alpha^i \le \boldsymbol{A}_i\boldsymbol{q} - \boldsymbol{D}_i\boldsymbol{\zeta} - d_i && \forall i \in \{1,\dots,l\} \quad \text{(H7c)} \\ & \boldsymbol{\beta}^i + \alpha^i\boldsymbol{e} \ge ((\boldsymbol{D}_i - \boldsymbol{A}_i\boldsymbol{P})\,\mathrm{diag}(\boldsymbol{e} - \boldsymbol{\zeta}))^\mathsf{T} && \forall i \in \{1,\dots,l\} \quad \text{(H7d)} \\ & \alpha^i \ge 0, \quad \boldsymbol{\beta}^i \ge \boldsymbol{0} && \forall i \in \{1,\dots,l\} \quad \text{(H7e)} \\ & \boldsymbol{0} \le \boldsymbol{\zeta} \le \boldsymbol{e}. && \text{(H7f)} \end{aligned}$$

We now show that for each solution of Problem (H7) there is an equivalent solution with $\boldsymbol{\zeta} = \mathbf{0}$. Let $\boldsymbol{P}, \boldsymbol{q}, \boldsymbol{\zeta}, \alpha^0, \ldots, \alpha^l, \boldsymbol{\beta}^0, \ldots, \boldsymbol{\beta}^l$ be a feasible solution to Problem (H7). We replace $\boldsymbol{P}$ by $\boldsymbol{P}' := \boldsymbol{P}\,\mathrm{diag}(\boldsymbol{e} - \boldsymbol{\zeta})$, $\boldsymbol{\beta}^i$ by $\boldsymbol{\beta}'^i := \boldsymbol{\beta}^i + \mathrm{diag}(\boldsymbol{\zeta})\boldsymbol{D}_i^\intercal$ for $i > 0$ and $\boldsymbol{\zeta}$ by $\mathbf{0}$ and leave all other decision variables untouched. As none of the variables in the objective function was changed, it is sufficient to show that the new solution is again feasible. Feasibility of Constraint (H7b) directly follows from $\boldsymbol{P}'\,\mathrm{diag}(\boldsymbol{e}) = \boldsymbol{P}\,\mathrm{diag}(\boldsymbol{e} - \boldsymbol{\zeta})$. For Constraint (H7c), we find

$$\begin{aligned}\boldsymbol{e}^\intercal \boldsymbol{\beta}'^i + k\alpha^i &\overset{\text{(a)}}{=} \boldsymbol{e}^\intercal \mathrm{diag}(\boldsymbol{\zeta})\boldsymbol{D}_i^\intercal + \boldsymbol{e}^\intercal \boldsymbol{\beta}^i + k\alpha^i \\ &\overset{\text{(b)}}{\leq} \boldsymbol{e}^\intercal \mathrm{diag}(\boldsymbol{\zeta})\boldsymbol{D}_i^\intercal + \boldsymbol{A}_i\boldsymbol{q} - \boldsymbol{D}_i\boldsymbol{\zeta} - d_i \overset{\text{(c)}}{=} \boldsymbol{A}_i\boldsymbol{q} - d_i,\end{aligned}$$

where (a) follows from the definition of $\boldsymbol{\beta}'^i$, (b) from the feasibility of the original solution and (c) from $\boldsymbol{e}^\intercal \mathrm{diag}(\boldsymbol{\zeta})\boldsymbol{D}_i^\intercal = \boldsymbol{D}_i\boldsymbol{\zeta}$. Similarly, Constraint (H7d) is fulfilled by

$$\begin{aligned}\boldsymbol{\beta}'^i + k\alpha^i\boldsymbol{e} &\overset{\text{(a)}}{=} \mathrm{diag}(\boldsymbol{\zeta})\boldsymbol{D}_i^\intercal + \boldsymbol{\beta}^i + k\alpha^i\boldsymbol{e} \\ &\overset{\text{(b)}}{\geq} \mathrm{diag}(\boldsymbol{\zeta})\boldsymbol{D}_i^\intercal + ((\boldsymbol{D}_i - \boldsymbol{A}_i\boldsymbol{P})\,\mathrm{diag}(\boldsymbol{e} - \boldsymbol{\zeta}))^\intercal \overset{\text{(c)}}{=} (\boldsymbol{D}_i - \boldsymbol{A}_i\boldsymbol{P}')^\intercal,\end{aligned}$$

where (a) and (b) again follow by the definition of $\boldsymbol{\beta}'^i$ and feasibility of the initial problem and (c) follows by definition of $\boldsymbol{P}'$. Finally, $\boldsymbol{\beta}'^i = \boldsymbol{\beta}^i + \mathrm{diag}(\boldsymbol{\zeta})\boldsymbol{D}_i^\intercal \geq \boldsymbol{\beta}^i \geq \mathbf{0}$ holds as $\boldsymbol{\zeta}$ and $\boldsymbol{D}$ are both non-negative. Thus we can w.l.o.g. set $\boldsymbol{\zeta} = \mathbf{0}$ in Problem (H7). In doing so, we transform Problem (H7) into Problem (H6). Thus the robust counterparts of affine policies and fully piecewise affine policies with modified right hand side are equivalent. □

Having shown these results, Proposition 8 directly follows by combining Lemmas 16 and 17. Note, that Lemma 18, which is crucial for the construction of tractable reformulations of fully piecewise affine policies, heavily depends on the linear structure of budgeted uncertainty and the existence of integer optimal solutions. Thus the results cannot easily be extended to other settings and in general one cannot hope for tractable reformulations of FPAPs.

Appendix I Proof of Proposition 9

Proof Analogously to the proofs of Propositions 5 and 7, we use Condition (7) and Lemma 4 to find a valid choice of μ and ρ. Using the property $\|\boldsymbol{e}\|_p \leq 1$ of the p-norm ball uncertainty set, we find $j\gamma(j)^p \leq 1$ for all $j \in \{1, \ldots, m\}$ and thus

$$\gamma(j) = j^{-\frac{1}{p}}.$$

Inserting $\gamma(j)$ in Condition (7), a valid choice of ρ is given by

$$\max_{j\in\{1,\dots,m\}} j\left(j^{-\frac{1}{p}}-\mu\right) \overset{(a)}{\leq} \mu^{1-p}p^{-p}(p-1)^{p-1} =: \rho, \tag{I1}$$

where (a) follows, as the maximum is taken at $j = (p\mu)^{-p}(p-1)^p$. Using the definition of $\mathcal{U}$ and $\boldsymbol{v}_i$ we find $\frac{1}{\beta}\boldsymbol{v}_i \in \mathcal{U}$ to be fulfilled when

$$\begin{aligned}\beta^p &= (m-1)\mu^p + (\mu+\rho)^p \overset{(a)}{\leq} \left(m-1+2^{p-1}\right)\mu^p + 2^{p-1}\rho^p \\ &\overset{(b)}{=} \left(m-1+2^{p-1}\right)\mu^p + 2^{p-1}\mu^{p-p^2}p^{-p^2}(p-1)^{p^2-p}.\end{aligned}$$

Here (a) follows from Jensen's inequality, as x^p is a convex function for $x \geq 0$, $p > 1$, and (b) holds by our choice of ρ in Eq. (I1). We now choose μ to minimize the right-hand side upper bound for β and find the minimum to be realized at $\mu = 2^{\frac{1}{p}}\left(2(m-1)+2^p\right)^{-\frac{1}{p^2}} p^{-1}(p-1)^{\frac{1}{p}+\left(1-\frac{1}{p}\right)^2}$. Using this μ in Eq. (I1), we find the desired ρ and we get the desired upper bound for β by taking the pth root of

$$\begin{aligned}\beta^p &\leq \left(2(m-1)+2^p\right)^{1-\frac{1}{p}} p^{-p}(p-1)^{1+p\left(1-\frac{1}{p}\right)^2} \\ &\quad + \left(2(m-1)+2^p\right)^{1-\frac{1}{p}} p^{p^2-p}(p-1)^{1-p+(p-p^2)\left(1-\frac{1}{p}\right)^2} p^{-p^2}(p-1)^{p^2-p} \\ &= \left(2(m-1)+2^p\right)^{1-\frac{1}{p}} p^{-p}\left((p-1)^{1+p\left(1-\frac{1}{p}\right)^2} + (p-1)^{1-2p+(p-p^2)\left(1-\frac{1}{p}\right)^2+p^2}\right) \\ &= \left(2(m-1)+2^p\right)^{1-\frac{1}{p}} p^{-p}\left((p-1)^{p-1+\frac{1}{p}} + (p-1)^{p-2+\frac{1}{p}}\right) \\ &= \left(2(m-1)+2^p\right)^{1-\frac{1}{p}} p^{-p}(p-1)^{p-2+\frac{1}{p}}\left(1+(p-1)\right) \\ &= \left(2(m-1)+2^p\right)^{1-\frac{1}{p}} p^{1-p}(p-1)^{p-2+\frac{1}{p}}.\end{aligned}$$

□

Appendix J Proof of Proposition 10

Proof We first show that our choice of μ, ρ yields a valid dominating set using Condition (7) and Lemma 4. Using the properties $\boldsymbol{\xi}^\mathsf{T}\boldsymbol{\Sigma}\boldsymbol{\xi} \leq 1$ and $\boldsymbol{\Sigma} = \mathbf{1} + a(\boldsymbol{J}-\mathbf{1})$ of the ellipsoid uncertainty set, we find $(1-a)j\gamma(j)^2 + aj^2\gamma(j)^2 \leq 1$ and thus

$$\gamma(j) = \frac{1}{\sqrt{aj^2+(1-a)j}}.$$

Inserting $\gamma(j)$ in Condition (7) we have to choose ρ such that

$$\rho \geq \max_{j\in\{1,\dots,m\}} j\left(\frac{1}{\sqrt{aj^2+(1-a)j}}-\mu\right).$$

For the right-hand side term we find

$$j\left(\frac{1}{\sqrt{aj^2+(1-a)j}}-\mu\right) \overset{\text{(a)}}{\leq} j\left(\frac{1}{\sqrt{aj^2}}\right) = \frac{1}{\sqrt{a}}$$

and

$$j\left(\frac{1}{\sqrt{aj^2+(1-a)j}}-\mu\right) \leq j\left(\frac{1}{\sqrt{(1-a)j}}-\mu\right) \overset{\text{(b)}}{\leq} \frac{1}{4(1-a)\mu},$$

where for (a) we use $\mu \geq 0, a \leq 1$ and (b) follows as the term is maximized for $j = \frac{1}{4(1-a)\mu^2}$. Thus choosing $\rho = \frac{1}{\sqrt{a}}$ or $\rho = \frac{1}{4(1-a)\mu}$ always yields a valid dominating set. Using the definition of $\mathcal{U}$ and $\boldsymbol{v}_i$ we find $\frac{1}{\beta}\boldsymbol{v}_i \in \mathcal{U}$ to be fulfilled when

$$\beta^2 = \rho^2 + 2(1-a+am)\rho\mu + (am^2+(1-a)m)\mu^2.$$

For $\mu = 0$, $\rho = \frac{1}{\sqrt{a}}$ this results in $\beta = \rho = \frac{1}{\sqrt{a}}$. Fixing $\rho = \frac{1}{4(1-a)\mu}$ yields

$$\beta^2 = \frac{1}{16(1-a)^2\mu^2} + \frac{1}{2}\left(1+\frac{a}{1-a}m\right) + (am^2+(1-a)m)\mu^2,$$

which is minimized for $\mu = \frac{1}{2\sqrt[4]{(1-a)^3m+(1-a)^2am^2}}$ resulting in

$$\beta^2 = \frac{1}{2}\left(1+\frac{1}{1-a}\left(am+\sqrt{(1-a)m+am^2}\right)\right).$$

Finally, we close the proof by showing the desired asymptotic approximation bounds. The case $a > m^{-\frac{2}{3}}$ directly follows from $\beta = \frac{1}{\sqrt{a}} \leq \frac{1}{\sqrt{m^{-\frac{2}{3}}}} = O(m^{\frac{1}{3}})$ and for $a \leq m^{-\frac{2}{3}}$ we find

$$\begin{aligned}\beta &= \sqrt{\frac{1}{2}\left(1+\frac{1}{1-a}\left(am+\sqrt{(1-a)m+am^2}\right)\right)} \\ &\overset{\text{(a)}}{\leq} \sqrt{\frac{1}{2}\left(1+\frac{1}{1-m^{-\frac{2}{3}}}\left(m^{\frac{1}{3}}+\sqrt{m+m^{\frac{4}{3}}}\right)\right)}\end{aligned}$$

$$\overset{\text{(b)}}{=} O(m^{\frac{1}{3}}),$$

where for (a) we use $a \leq m^{-\frac{2}{3}}$ and for (b) we use $m \geq 2$. □

Appendix K Proof of Proposition 11

We prove Proposition 11 by explicitly constructing a dominating set fulfilling the desired properties. To find a good choice of $\rho_1, \ldots, \rho_m$, $\boldsymbol{v}_0$ for the construction of $\hat{\mathcal{U}}$ in (2), we use the iterative approach described in Algorithm 1. Intuitively, we increase the base vertex $\boldsymbol{v}_0$, as well as an upper bound for ρ_i, in each iteration. We bound the approximation factor β, by bounding the maximal number of iterations in the algorithm. Algorithm 1 is a refinement of Algorithm 1 in Ben-Tal et al. [14] and uses a different updating step. This modified updating step leads to a less aggressive increase of the base vertex $\boldsymbol{v}_0$ ultimately improving the approximation bound by a factor of $\frac{1}{2}$.

Algorithm 1 Generating $\boldsymbol{v}_0$ for general uncertainty sets $\mathcal{U}$

$j = 0$
$\boldsymbol{v}^0 = \mathbf{0}$
while $\max_{\boldsymbol{\xi} \in \mathcal{U}} \boldsymbol{e}^\mathsf{T}(\boldsymbol{\xi} - \boldsymbol{v}^j)_+ > j + 1$ **do**
  $\boldsymbol{\xi}^j \in \operatorname{argmax}_{\boldsymbol{\xi} \in \mathcal{U}} \boldsymbol{e}^\mathsf{T}(\boldsymbol{\xi} - \boldsymbol{v}^j)_+$
  $\boldsymbol{v}^{j+1} = \boldsymbol{v}^j + (\boldsymbol{\xi}^j - \boldsymbol{v}^j)_+$
  $j = j + 1$
end while
return $\beta = 2j + 1$, $\boldsymbol{v}_0 = \boldsymbol{v}^j$

Lemma 19 *Let β, $\boldsymbol{v}_0$ be the output of Algorithm 1. Then a solution for $Z_{AR}(\hat{\mathcal{U}})$, where $\hat{\mathcal{U}}$ is constructed according to (2) with $\boldsymbol{v}_0$ and $\rho_i = \frac{\beta+1}{2}$ is a β approximation for $Z_{AR}(\mathcal{U})$ and $\beta \leq 2\sqrt{m} + 1$.*

Proof In this proof let $J := \frac{\beta-1}{2}$ be the index j at termination of Algorithm 1. First, we show that $\hat{\mathcal{U}}$ is a valid dominating set for $\mathcal{U}$ using Criterion (5). Let $\boldsymbol{\xi} \in \mathcal{U}$. Then

$$\sum_{i=1}^{m} \frac{((\boldsymbol{\xi} - \boldsymbol{v}_0)_+)_i}{\rho_i} \overset{\text{(a)}}{=} \frac{1}{J+1} \boldsymbol{e}^\mathsf{T}(\boldsymbol{\xi} - \boldsymbol{v}_0)_+ \overset{\text{(b)}}{\leq} 1,$$

where (a) follows from our choice of ρ_i and (b) from the termination criterion of Algorithm 1. Next, we show $\frac{1}{\beta}\boldsymbol{v}_i \in \mathcal{U}$ for all $i \in \{1, \ldots, m\}$ by

$$\frac{1}{\beta}\boldsymbol{v}_i = \frac{J+1}{\beta}\boldsymbol{e}_i + \frac{1}{\beta}\sum_{j=0}^{J-1}(\boldsymbol{\xi}^j - \boldsymbol{v}^j)_+ \leq \frac{1}{\beta}\left((J+1)\boldsymbol{e}_i + \sum_{j=0}^{J-1}\boldsymbol{\xi}^j\right) \overset{\text{(a)}}{\in} \mathcal{U}.$$

Here, the containment (a) follows from down-monotinicity, convexity and $\boldsymbol{e}_i \in \mathcal{U}$. Finally, we show that Algorithm 1 terminates after at most $\frac{\beta-1}{2} = J \leq \sqrt{m}$ iterations,

which implies $\beta \leq 2\sqrt{m} + 1$. Let $J' := J - 1$ be the index of the last iteration that fulfills the criterion of the loop. Then

$$\begin{aligned} J^2 &= \sum_{j=0}^{J'} J \overset{\text{(a)}}{\leq} \sum_{j=0}^{J'} \boldsymbol{e}^\mathsf{T}(\boldsymbol{\xi}^{J'} - \boldsymbol{v}^{J'})_+ \overset{\text{(b)}}{\leq} \sum_{j=0}^{J'} \boldsymbol{e}^\mathsf{T}(\boldsymbol{\xi}^{J'} - \boldsymbol{v}^{j})_+ \\ &\overset{\text{(c)}}{\leq} \sum_{j=0}^{J'} \max_{\boldsymbol{\xi}\in\mathcal{U}} \boldsymbol{e}^\mathsf{T}(\boldsymbol{\xi} - \boldsymbol{v}^{j})_+ \overset{\text{(d)}}{=} \sum_{j=0}^{J'} \boldsymbol{e}^\mathsf{T}(\boldsymbol{\xi}^{j} - \boldsymbol{v}^{j})_+ \\ &= \sum_{i=1}^{m}\sum_{j=0}^{J'} (\xi_i^j - v_i^j)_+ \overset{\text{(e)}}{=} \sum_{i=1}^{m} v_i^J \overset{\text{(f)}}{=} \sum_{i=1}^{m} \max_{j\in\{0,\ldots,J'\}} \xi_i^j \overset{\text{(g)}}{\leq} \sum_{i=1}^{m} 1 = m. \end{aligned}$$

Here, inequality (a) follows from the termination criterion of Algorithm 1, (b) follows as $\boldsymbol{v}^j$ only increases through the algorithm and thus $\forall j \leq J' \colon \boldsymbol{v}^{J'} \geq \boldsymbol{v}^j$, (c) follows as by $\boldsymbol{\xi}^{J'} \in \mathcal{U}$ the maximum over $\mathcal{U}$ is an upper bound, (d) follows from the choice of $\boldsymbol{\xi}^j$ in Algorithm 1, (e) follows from the construction of $\boldsymbol{v}^j$, (f) follows as by induction $v_i^{j+1} = \max(v_i^j, \xi_i^j) = \max_{j'\leq j} \xi_i^{j'}$, and (g) follows by $\boldsymbol{e}^j \in \mathcal{U} \subseteq [0,1]^m$. □

Appendix L Proof of Proposition 12

Proof For each t, let m_t be the dimension $\mathcal{U}_t$. Then $m = \sum_{t=1}^{T} m_t$, and an uncertainty vector $\boldsymbol{\xi} \in \mathcal{U}$ can be expressed as $\boldsymbol{\xi} = (\boldsymbol{\xi}^1, \ldots, \boldsymbol{\xi}^T) = (\xi_1^1, \ldots, \xi_{m_1}^1, \ldots, \xi_{m_T}^T)$, where $\boldsymbol{\xi}^t \in \mathcal{U}_t$. For each t, let $\boldsymbol{v}_0^t$ be the base vertex and let $\rho_1^t, \ldots, \rho_{m_t}^t$ be the parameters inducing $\hat{\mathcal{U}}_t$ via Construction (2). Let $\beta = \sum_{t\in\mathcal{T}_1} \beta_t$, and construct the dominating set $\hat{\mathcal{U}}$ for $\mathcal{U}$ with base vertex $\boldsymbol{v}_0' := (\boldsymbol{v}_0'^1, \ldots, \boldsymbol{v}_0'^T)$, where

$$\boldsymbol{v}_0'^t := \begin{cases} \boldsymbol{v}_0^t & \text{if } t \in \mathcal{T}_1 \\ \boldsymbol{e} & \text{if } t \in \mathcal{T}_2 \end{cases}$$

and parameters

$$\rho_i'^t := \begin{cases} \frac{\beta}{\beta_t}\rho_i^t & \text{if } t \in \mathcal{T}_1 \\ 0 & \text{if } t \in \mathcal{T}_2 \end{cases}$$

using Construction (2). This construction corresponds to dominating each $\mathcal{U}_t$ with $t \in \mathcal{T}_2$ by the unit vector $\boldsymbol{e}$. Here $\{\boldsymbol{e}\}$ is dominating $\mathcal{U}_t$ by Assumption 1. We dominate the remaining $\mathcal{U}_t$ for $t \in \mathcal{T}_1$ with a combined polytope. Then the maximal sum of convex factors for the vertices is given by

$$\max_{\boldsymbol{\xi}\in\mathcal{U}} \sum_{t\in\mathcal{T}_1}\sum_{i=1}^{m_t} \frac{\left((\boldsymbol{\xi} - \boldsymbol{v}_0')_+\right)_i^t}{\rho_i'^t} \overset{\text{(a)}}{=} \frac{1}{\beta} \max_{\boldsymbol{\xi}\in\mathcal{U}} \sum_{t\in\mathcal{T}_1} \beta_t \sum_{i=1}^{m_t} \frac{\left((\boldsymbol{\xi}^t - \boldsymbol{v}_0^t)_+\right)_i}{\rho_i^t}$$

$$\overset{\text{(b)}}{=} \frac{1}{\beta} \sum_{t \in \mathcal{T}_1} \beta_t \max_{\boldsymbol{\xi}^t \in \mathcal{U}_t} \sum_{i=1}^{m_t} \frac{\left((\boldsymbol{\xi}^t - \boldsymbol{v}_0^t)_+\right)_i}{\rho_i^t} \overset{\text{(c)}}{\leq} \frac{1}{\beta} \sum_{t \in \mathcal{T}_1} \beta_t \overset{\text{(d)}}{=} 1,$$

where (a) follows from the definition of $\boldsymbol{v}_0'$ and $\rho_i'^t$, (b) follows from $\mathcal{U}$ being stagewise independent, (c) follows from Condition (5) and $\hat{\mathcal{U}}_t$ being a dominating set for $\mathcal{U}_t$, and (d) follows from the definition of β. Thus $\hat{\mathcal{U}}$ fulfills Condition (5) and is a valid dominating set for $\mathcal{U}$.

Let $\beta' := \max_{t \in \mathcal{T}_2} \beta_t'$. To see that $\max(\beta, \beta')$ is indeed an upper bound for the approximation factor, we find

$$\begin{aligned}
&\frac{1}{\max(\beta, \beta')} \boldsymbol{v}_0'^t \overset{\text{(a)}}{\leq} \frac{1}{\beta}(\boldsymbol{v}_0'^t + \rho_i'^t \boldsymbol{e}_i) \overset{\text{(b)}}{=} \frac{1}{\beta}(\boldsymbol{v}_0^t + \frac{\beta}{\beta_t} \rho_i^t \boldsymbol{e}_i) \overset{\text{(c)}}{\leq} \frac{1}{\beta_t}(\boldsymbol{v}_0^t + \rho_i^t \boldsymbol{e}_i) \overset{\text{(d)}}{\in} \mathcal{U}_t \qquad \forall t \in \mathcal{T}_1 \\
&\frac{1}{\max(\beta, \beta')} \boldsymbol{v}_0'^t \overset{\text{(e)}}{\leq} \frac{1}{\beta'} \boldsymbol{e} \overset{\text{(f)}}{\leq} \frac{1}{\beta_t'} \boldsymbol{e} \overset{\text{(g)}}{\in} \mathcal{U}_t \qquad \forall t \in \mathcal{T}_2,
\end{aligned}$$

where (a) follows from $\beta \leq \max(\beta, \beta')$ and $\rho_i'^t \geq 0$, (b) follows from the definitions of $\boldsymbol{v}_0'^t$ and $\rho_i'^t$, (c) follows from $\beta_t \leq \beta$ for all $t \in \mathcal{T}_1$, (d) follows as β_t is a valid approximation factor for $\hat{\mathcal{U}}_t$, (e) follows from $\beta' \leq \max(\beta, \beta')$ and the definition of $\boldsymbol{v}_0'^t$, (f) follows from $\beta_t' \leq \beta'$ for all $t \in \mathcal{T}_2$, and (g) follows from the definition of β_t'. Using the stagewise structure and down monotinicity of $\mathcal{U}$ this implies

$$\begin{aligned}
&\frac{1}{\max(\beta, \beta')} \boldsymbol{v}_0' = \frac{1}{\max(\beta, \beta')} (\boldsymbol{v}_0'^1, \ldots, \boldsymbol{v}_0'^T) \in \mathcal{U}, \\
&\frac{1}{\max(\beta, \beta')} (\boldsymbol{v}_0' + \rho_i'^t \boldsymbol{e}_i^t) = \frac{1}{\max(\beta, \beta')} (\boldsymbol{v}_0'^1, \ldots, \boldsymbol{v}_0'^t + \rho_i'^t \boldsymbol{e}_i, \ldots, \boldsymbol{v}_0'^T) \in \mathcal{U}.
\end{aligned}$$

□

Appendix M Proof of Lemma 13

Proof Let $\boldsymbol{h} : \mathcal{U} \to \hat{\mathcal{U}}$ be the domination function of $\hat{\mathcal{U}}$. Then define $\boldsymbol{h}' : \mathcal{U} \to \hat{\mathcal{U}}'$ by

$$h_j'(\boldsymbol{\xi}) := h_j(\boldsymbol{\xi}) + s_j(1 - h_j(\boldsymbol{\xi})). \tag{M1}$$

This new dominating function $\boldsymbol{h}'$ is nonanticipative, as $\boldsymbol{h}$ is nonanticipative and the jth component of $\boldsymbol{h}'$ only depends on the jth component of $\boldsymbol{h}$. To see that (M1) is also a domination function, consider the two cases $h_j(\boldsymbol{\xi}) \geq 1$ and $h_j(\boldsymbol{\xi}) \leq 1$.

1. For $h_j(\boldsymbol{\xi}) \geq 1$ the second summand on the right hand side of (M1) is negative for all $s_j \in [0, 1]$. Thus $h_j'(\boldsymbol{\xi})$ is minimal for $s_j = 1$ and we find $h_j'(\boldsymbol{\xi}) \geq h_j(\boldsymbol{\xi}) + 1(1 - h_j(\boldsymbol{\xi})) = 1$.
2. On the other hand, for $h_j'(\boldsymbol{\xi}) \leq 1$ the second summand is positive and $h_j'(\boldsymbol{\xi})$ is minimal for $s_j = 0$. Thus, we have $h_j'(\boldsymbol{\xi}) \geq h_j(\boldsymbol{\xi})$.

In any case, we find $\boldsymbol{h}'(\boldsymbol{\xi}) \geq \boldsymbol{\xi}$ by $h'_j \geq \min(1, h_j(\boldsymbol{\xi})) \geq \xi_j$ where we used $\mathcal{U} \subseteq [0,1]^m$ from Assumption 1.

Next, we show $\boldsymbol{h}'(\boldsymbol{\xi}) \in \hat{\mathcal{U}}'$. For this let $\boldsymbol{h}(\boldsymbol{\xi}) = \sum_{i=0}^m \lambda_i \boldsymbol{v}_i$ be the convex combination of $\boldsymbol{h}(\boldsymbol{\xi}) \in \hat{\mathcal{U}}$. Then,

$$h'_j(\boldsymbol{\xi}) = h_j(\boldsymbol{\xi}) + s_j(1 - h_j(\boldsymbol{\xi})) \overset{(a)}{=} \sum_{i=0}^m \big(\lambda_i v_{ij} + s_j(1 - v_{ij})\big) = \sum_{i=0}^m \lambda_i v'_{ij}.$$

Note that in (a) we used $\sum_{i=0}^m \lambda_i = 1$. This shows $\boldsymbol{h}'(\boldsymbol{\xi}) = \sum_{i=0}^m \lambda_i \boldsymbol{v}'_i$ and thus $\boldsymbol{h}'(\boldsymbol{\xi}) \in \hat{\mathcal{U}}'$. □

Appendix N Proof of Proposition 14

Proof As $Z_{LIFT} \leq Z_{AFF}$ was already shown by Georghiou et al. [31], we are left to show $Z_{LIFT} \geq Z_{AFF}$. We do so by showing that for any optimal affine solution for $Z^L_{AR}(\mathcal{U}^L)$, there is an affine solution for $Z_{AR}(\mathcal{U})$ with the same objective value. First, we define the average lifting operator $\bar{L} : \mathcal{U} \to \mathcal{U}^L$ via

$$\bar{L}_{ij}(\boldsymbol{\xi}) := (z^i_j - z^i_{j-1})\xi_i.$$

Note, that $R_i(\bar{L}(\boldsymbol{\xi})) = \sum_{j=1}^{r_i}(z^i_j - z^i_{j-1})\xi_i = (z^i_{r_i} - z^i_0)\xi_i = \xi_i$, which implies $R(\bar{L}(\boldsymbol{\xi})) = \boldsymbol{\xi}$. Thus, $\bar{L}(\boldsymbol{\xi})$ fulfills the first condition of (8). Next, we verify the second condition by

$$\frac{\bar{L}_{i,j+1}(\boldsymbol{\xi})}{z^i_{j+1} - z^i_j} = \frac{(z^i_{j+1} - z^i_j)\xi_i}{z^i_{j+1} - z^i_j} = \xi_i = \frac{(z^i_j - z^i_{j-1})\xi_i}{z^i_j - z^i_{j-1}} = \frac{\bar{L}_{ij}(\boldsymbol{\xi})}{z^i_j - z^i_{j-1}}.$$

Finally, the third condition holds by $\bar{L}_{i1} = z^i_1 \xi_i \leq z^i_1$ which follows from $\xi_i \leq 1$ in Assumption 1. Thus $\bar{L}(\boldsymbol{\xi}) \in \mathcal{U}^L$. Let $\boldsymbol{x}^L$ be an optimal affine solution for $Z^L_{AR}(\mathcal{U}^L)$. As $\bar{L}$ is a nonanticipative affine map, the concatenation $\boldsymbol{x}^L \circ \bar{L}$ is also nonanticipative and affine. By $\bar{L}(\mathcal{U}) \subseteq \mathcal{U}^L$ the decision $\boldsymbol{x} := \boldsymbol{x}^L \circ \bar{L}$ is a feasible solution for $Z_{AR}(\mathcal{U})$. Finally, we find

$$Z_{AFF} \leq \max_{\boldsymbol{\xi} \in \mathcal{U}} \boldsymbol{c}^\intercal \boldsymbol{x}(\boldsymbol{\xi}) = \max_{\boldsymbol{\xi}^L \in \bar{L}(\mathcal{U})} \boldsymbol{c}^\intercal \boldsymbol{x}^L(\boldsymbol{\xi}^L) \leq \max_{\boldsymbol{\xi}^L \in \mathcal{U}^L} \boldsymbol{c}^\intercal \boldsymbol{x}^L(\boldsymbol{\xi}^L) = Z_{LIFT}.$$

□

Appendix O Proof of Proposition 15

Proof To prove $Z_{TLIFT} \leq Z_{LIFT}$ and $Z_{TLIFT} \leq Z_{SPAP}$ we show that any affine solution for $Z^L_{AR}(\mathcal{U}^L)$ and any picecwise affine solution with re-scaling for $Z(\hat{\mathcal{U}})$

can be transformed to a feasible affine solution for $Z^L_{AR}(\hat{\mathcal{U}}^L)$ with at most the same objective value.

For $Z_{TLIFT} \leq Z_{LIFT}$, let $\boldsymbol{x}^L \colon \mathcal{U}^L \to \mathbb{R}^n$ be an optimal affine solution to $Z^L_{AR}(\mathcal{U}^L)$. Then $\boldsymbol{x}^L$ is also a feasible affine solution for $Z^L_{AR}(\hat{\mathcal{U}}^L)$ by $\hat{\mathcal{U}}^L \subseteq \mathcal{U}^L$. For the objective value, we find

$$Z_{TLIFT} \leq \max_{\boldsymbol{\xi}^L \in \hat{\mathcal{U}}^L} \boldsymbol{c}^\intercal \boldsymbol{x}^L(\boldsymbol{\xi}^L) \overset{\text{(a)}}{\leq} \max_{\boldsymbol{\xi}^L \in \mathcal{U}^L} \boldsymbol{c}^\intercal \boldsymbol{x}^L(\boldsymbol{\xi}^L) = Z_{LIFT},$$

where (a) also follows from $\hat{\mathcal{U}}^L \subseteq \mathcal{U}^L$.

For $Z_{TLIFT} \leq Z_{SPAP}$ let $\hat{\boldsymbol{x}}(\hat{\boldsymbol{\xi}})$ given by vertex solutions $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m$ and scaling factor $\boldsymbol{s}$ be a an optimal solution to $Z_{LP}(\hat{\boldsymbol{\xi}})$ with re-scaling as described in Sect. 4. We define the map $\boldsymbol{h}^L_{\boldsymbol{s}} \colon \hat{\mathcal{U}}^L \to \hat{\mathcal{U}}'$ from the tightened lifted uncertainty set $\hat{\mathcal{U}}^L$ defined in Eq. (10) to the re-scaled dominating uncertainty set $\hat{\mathcal{U}}'$ defined in Lemma 13 by

$$h^L_{si}(\hat{\boldsymbol{\xi}}^L) := (1 - s_i)(v_{0i} + \hat{\xi}^L_{i2}) + s_i.$$

This mapping is nonanticipative. Furthermore, $\boldsymbol{h}^L_{\boldsymbol{s}}(\hat{\boldsymbol{\xi}}^L)$ dominates $R(\hat{\boldsymbol{\xi}}^L)$ by

$$R(\hat{\boldsymbol{\xi}}^L)_i = \hat{\xi}^L_{i1} + \hat{\xi}^L_{i2} \overset{\text{(a)}}{\leq} (1 - s_i)(\hat{\xi}^L_{i1} + \hat{\xi}^L_{i2}) + s_i \overset{\text{(b)}}{\leq} (1 - s_i)(v_{0i} + \hat{\xi}^L_{i2}) + s_i = h^L_{si}(\hat{\boldsymbol{\xi}}^L).$$

Here, (a) follows as $\hat{\xi}^L_{i1} + \hat{\xi}^L_{i2} \leq 1$ by Assumption 1, and (b) follows from $\hat{\xi}^L_{i1} \leq v_{0i}$ as v_{0i} is the break-point. Using the definition of the re-scaled vertices $\boldsymbol{v}'_i$ from Lemma 13, we find

$$\boldsymbol{h}^L_{\boldsymbol{s}}(\hat{\boldsymbol{\xi}}^L) = \sum_{i=1}^{m} \frac{\hat{\xi}^L_{i2}}{\rho_i} \boldsymbol{v}'_i + \left(1 - \sum_{i=1}^{m} \frac{\hat{\xi}^L_{i2}}{\rho_i}\right) \boldsymbol{v}'_0. \tag{O1}$$

By the tightening constraint in Definition (10) this is a valid convex combination for all $\hat{\boldsymbol{\xi}}^L \in \hat{\mathcal{U}}^L$ and thus $\boldsymbol{h}^L_{\boldsymbol{s}}\left(\hat{\mathcal{U}}^L\right) \subseteq \hat{\mathcal{U}}'$ as claimed. Thus $\boldsymbol{x}^L := \hat{\boldsymbol{x}} \circ \boldsymbol{h}^L_{\boldsymbol{s}}$ is a valid solution for $Z^L_{AR}(\hat{\mathcal{U}}^L)$ and

$$\max_{\hat{\boldsymbol{\xi}}^L \in \hat{\mathcal{U}}^L} \boldsymbol{c}^\intercal \boldsymbol{x}^L(\hat{\boldsymbol{\xi}}^L) = \max_{\hat{\boldsymbol{\xi}} \in \boldsymbol{h}^L_{\boldsymbol{s}}\left(\hat{\mathcal{U}}^L\right)} \boldsymbol{c}^\intercal \hat{\boldsymbol{x}}(\hat{\boldsymbol{\xi}}) \leq \max_{\hat{\boldsymbol{\xi}} \in \hat{\mathcal{U}}} \boldsymbol{c}^\intercal \hat{\boldsymbol{x}}(\hat{\boldsymbol{\xi}}) = Z_{SPAP}.$$

Using that $\hat{\boldsymbol{x}}$ is given by the vertex solutions $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m$ together with (O1) we find

$$\boldsymbol{x}^L(\hat{\boldsymbol{\xi}}^L) = \hat{\boldsymbol{x}} \circ \boldsymbol{h}^L_{\boldsymbol{s}}(\hat{\boldsymbol{\xi}}^L) = \sum_{i=1}^{m} \frac{\hat{\xi}^L_{i2}}{\rho_i} \boldsymbol{x}_i + \left(1 - \sum_{i=1}^{m} \frac{\hat{\xi}^L_{i2}}{\rho_i}\right) \boldsymbol{x}_0 = \sum_{i=1}^{m} \left(\frac{\hat{\xi}^L_{i2}}{\rho_i} (\boldsymbol{x}_i - \boldsymbol{x}_0)\right) + \boldsymbol{x}_0,$$

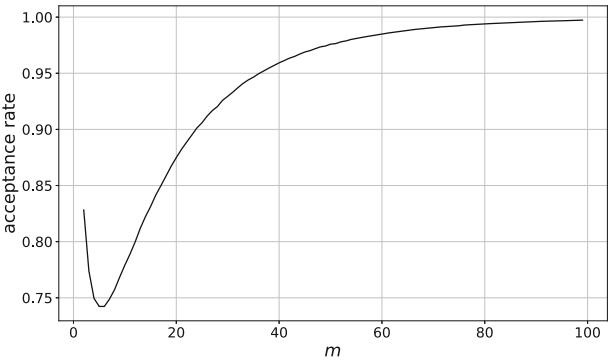


Fig. 10 Acceptance rate of our proposed sampling method for budgeted uncertainty with budget $k = \sqrt{m}$

which is affine in $\hat{\mathcal{U}}^L$ and thus

$$Z_{TLIFT} \leq \max_{\hat{\boldsymbol{\xi}}^L \in \hat{\mathcal{U}}^L} \boldsymbol{c}^\intercal \boldsymbol{x}^L(\hat{\boldsymbol{\xi}}^L).$$

□

Appendix P Sampling Uniform Uncertainty Realizations

In this section, we describe how to generate uniform samples from the uncertainty sets used in our numerical experiments efficiently.

Hypersphere uncertainty: To sample uniformly from the hypersphere uncertainty set $\mathcal{U} = \left\{\boldsymbol{\xi} \in \mathbb{R}^m_+ \middle| \|\boldsymbol{\xi}\|_2^2 \leq 1\right\}$, we first draw i.i.d. samples $\boldsymbol{\xi}^B$ from the m-dimensional unit ball. It is commonly known, that this can efficiently be done by drawing m i.i.d. normal variables that are normalized and rescaled by the radius appropriately, see, e.g. [29]. We now map the samples from the unit ball to the positive orthant by taking the absolute value in each component, i.e., we get $\boldsymbol{\xi} \in \mathcal{U}$ by $\xi_i := |\xi_i^B|$. By rotational invariance of the unit ball, this yields a uniform sample on the hypersphere uncertainty set.

Budgeted uncertainty: For the budgeted uncertainty set $\mathcal{U} = \left\{\boldsymbol{\xi} \in [0,1]^m \middle| \|\boldsymbol{\xi}\|_1 \leq k\right\}$ with budged $k = \sqrt{m}$, we first draw i.i.d. samples $\boldsymbol{\xi}^S$ from the m-dimensional unit simplex $\text{conv}(\mathbf{0}, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_m)$. This can be done efficiently by taking the first m of $m+1$ i.i.d. exponentially distributed variables that were normalized with respect to the one norm, see, e.g. [29]. By scaling the samples from the unit simplex with the budget k, and rejecting all samples that do not fulfill $\boldsymbol{\xi} := k\boldsymbol{\xi}^S \in [0,1]^m$, we get a uniform sample on the budgeted uncertainty set. In Fig. 10, we show the acceptance rate of the samples, which shows that the proposed rejection method is efficient for our application.

References

1. Adida, E., Perakis, G.: A robust optimization approach to dynamic pricing and inventory control with no backorders. Math. Program. **107**(1–2), 97–129 (2006). https://doi.org/10.1007/s10107-005-0681-5
2. Ang, M., Lim, Y.F., Sim, M.: Robust storage assignment in unit-load warehouses. Manag. Sci. **58**(11), 2114–2130 (2012). https://doi.org/10.1287/mnsc.1120.1543
3. Ayoub, J., Poss, M.: Decomposition for adjustable robust linear optimization subject to uncertainty polytope. Comput. Manag. Sci. **13**(2), 219–239 (2016). https://doi.org/10.1007/s10287-016-0249-2
4. Bampou, D., Kuhn, D.: Scenario-free stochastic programming with polynomial decision rules. In: IEEE Conference on Decision and Control and European Control Conference. IEEE, Orlando (2011). https://doi.org/10.1109/cdc.2011.6161150
5. Baron, O., Milner, J., Naseraldin, H.: Facility location: a robust optimization approach. Prod. Oper. Manag. **20**(5), 772–785 (2011). https://doi.org/10.1111/j.1937-5956.2010.01194.x
6. Beebe, J.H., Beightler, C.S., Stark, J.P.: Stochastic optimization of production planning. Oper. Res. **16**(4), 799–818 (1968). https://doi.org/10.1287/opre.16.4.799
7. Ben-Tal, A., Nemirovski, A.: Robust solutions of uncertain linear programs. Oper. Res. Lett. **25**(1), 1–13 (1999). https://doi.org/10.1016/s0167-6377(99)00016-4
8. Ben-Tal, A., Nemirovski, A.: Robust optimization—methodology and applications. Math. Program. **92**(3), 453–480 (2002). https://doi.org/10.1007/s101070100286
9. Ben-Tal, A., Goryashko, A., Guslitzer, E., et al.: Adjustable robust solutions of uncertain linear programs. Math. Program. **99**(2), 351–376 (2004). https://doi.org/10.1007/s10107-003-0454-y
10. Ben-Tal, A., Golany, B., Nemirovski, A., et al.: Retailer-supplier flexible commitments contracts: a robust optimization approach. Manuf. Serv. Oper. Manag. **7**(3), 248–271 (2005). https://doi.org/10.1287/msom.1050.0081
11. Ben-Tal, A., Ghaoui, L.E., Nemirovski, A.: Robust Optimization. Princeton University Press, Princeton (2009)
12. Ben-Tal, A., Golany, B., Shtern, S.: Robust multi-echelon multi-period inventory control. Eur. J. Oper. Res. **199**(3), 922–935 (2009). https://doi.org/10.1016/j.ejor.2009.01.058
13. Ben-Tal, A., Chung, B.D., Mandala, S.R., et al.: Robust optimization for emergency logistics planning: risk mitigation in humanitarian relief supply chains. Transp. Res. Part B Methodol. **45**(8), 1177–1189 (2011). https://doi.org/10.1016/j.trb.2010.09.002
14. Ben-Tal, A., Housni, O.E., Goyal, V.: A tractable approach for designing piecewise affine policies in two-stage adjustable robust optimization. Math. Program. **182**(1–2), 57–102 (2020). https://doi.org/10.1007/s10107-019-01385-0
15. Bertsimas, D., Bidkhori, H.: On the performance of affine policies for two-stage adaptive optimization: a geometric perspective. Math. Program. **153**(2), 577–594 (2015). https://doi.org/10.1007/s10107-014-0818-5
16. Bertsimas, D., Dunning, I.: Multistage robust mixed-integer optimization with adaptive partitions. Oper. Res. **64**(4), 980–998 (2016). https://doi.org/10.1287/opre.2016.1515
17. Bertsimas, D., Georghiou, A.: Design of near optimal decision rules in multistage adaptive mixed-integer optimization. Oper. Res. **63**(3), 610–627 (2015). https://doi.org/10.1287/opre.2015.1365
18. Bertsimas, D., Goyal, V.: On the power and limitations of affine policies in two-stage adaptive optimization. Math. Program. **134**(2), 491–531 (2012). https://doi.org/10.1007/s10107-011-0444-4
19. Bertsimas, D., Iancu, D.A., Parrilo, P.A.: Optimality of affine policies in multistage robust optimization. Math. Oper. Res. **35**(2), 363–394 (2010). https://doi.org/10.1287/moor.1100.0444
20. Bertsimas, D., Goyal, V., Sun, X.A.: A geometric characterization of the power of finite adaptability in multistage stochastic and adaptive optimization. Math. Oper. Res. **36**(1), 24–54 (2011). https://doi.org/10.1287/moor.1110.0482
21. Bertsimas, D., Iancu, D.A., Parrilo, P.A.: A hierarchy of near-optimal policies for multistage adaptive optimization. IEEE Trans. Autom. Control **56**(12), 2809–2824 (2011). https://doi.org/10.1109/tac.2011.2162878
22. Bertsimas, D., Goyal, V., Lu, B.Y.: A tight characterization of the performance of static solutions in two-stage adjustable robust linear optimization. Math. Program. **150**(2), 281–319 (2015). https://doi.org/10.1007/s10107-014-0768-y
23. Bertsimas, D., Shtern, S., Sturt, B.: Technical note—two-stage sample robust optimization. Oper. Res. (2021). https://doi.org/10.1287/opre.2020.2096

24. Chen, X., Zhang, Y.: Uncertain linear programs: extended affinely adjustable robust counterparts. Oper. Res. **57**(6), 1469–1482 (2009). https://doi.org/10.1287/opre.1080.0605
25. Chen, X., Sim, M., Sun, P., et al.: A linear decision-based approximation approach to stochastic programming. Oper. Res. **56**(2), 344–357 (2008). https://doi.org/10.1287/opre.1070.0457
26. Dantzig, G.B.: Linear programming under uncertainty. Manag. Sci. **1**(3–4), 197–206 (1955). https://doi.org/10.1287/mnsc.1.3-4.197
27. Daryalal, M., Bodur, M., Luedtke, J.R.: Lagrangian dual decision rules for multistage stochastic mixed-integer programming. Oper. Res. (2022). https://doi.org/10.1287/opre.2022.2366
28. Delage, E., Iancu, D.A.: Robust multistage decision making. In: The Operations Research Revolution. INFORMS, Catonsville, pp. 20–46 (2015). https://doi.org/10.1287/educ.2015.0139
29. Devroye, L.: Non-Uniform Random Variate Generation. Springer, New York (1986). https://doi.org/10.1007/978-1-4613-8643-8
30. Feige, U., Jain, K., Mahdian, M., et al.: Robust combinatorial optimization with exponential scenarios. In: Integer Programming and Combinatorial Optimization, pp. 439–453. Springer, Berlin (2007). https://doi.org/10.1007/978-3-540-72792-7-33
31. Georghiou, A., Wiesemann, W., Kuhn, D.: Generalized decision rule approximations for stochastic programming via liftings. Math. Program. **152**(1–2), 301–338 (2015). https://doi.org/10.1007/s10107-014-0789-6
32. Georghiou, A., Tsoukalas, A., Wiesemann, W.: Robust dual dynamic programming. Oper. Res. **67**(3), 813–830 (2019). https://doi.org/10.1287/opre.2018.1835
33. Georghiou, A., Tsoukalas, A., Wiesemann, W.: A primal–dual lifting scheme for two-stage robust optimization. Oper. Res. (2020). https://doi.org/10.1287/opre.2019.1873
34. Ghaoui, L.E., Oks, M., Oustry, F.: Worst-case value-at-risk and robust portfolio optimization: a conic programming approach. Oper. Res. **51**(4), 543–556 (2003). https://doi.org/10.1287/opre.51.4.543.16101
35. Goldfarb, D., Iyengar, G.N.: Robust portfolio selection problems. Math. Oper. Res. **28**(1), 1–38 (2003). https://doi.org/10.1287/moor.28.1.1.14260
36. Gorissen, B.L., Yanıkoğlu, İ, den Hertog, D.: A practical guide to robust optimization. Omega **53**, 124–137 (2015). https://doi.org/10.1016/j.omega.2014.12.006
37. Haddad-Sisakht, A., Ryan, S.M.: Conditions under which adjustability lowers the cost of a robust linear program. Ann. Oper. Res. **269**(1–2), 185–204 (2018). https://doi.org/10.1007/s10479-018-2954-4
38. Hadjiyiannis, M.J., Goulart, P.J., Kuhn, D.: A scenario approach for estimating the suboptimality of linear decision rules in two-stage robust optimization. In: IEEE Conference on Decision and Control and European Control Conference. IEEE (2011) https://doi.org/10.1109/cdc.2011.6161342
39. Hanasusanto, G.A., Kuhn, D., Wiesemann, W.: K-adaptability in two-stage robust binary programming. Oper. Res. **63**(4), 877–891 (2015). https://doi.org/10.1287/opre.2015.1392
40. Housni, O.E., Goyal, V.: On the optimality of affine policies for budgeted uncertainty sets. Math. Oper. Res. **46**(2), 674–711 (2021). https://doi.org/10.1287/moor.2020.1082
41. Høyland, K., Wallace, S.W.: Generating scenario trees for multistage decision problems. Manag. Sci. **47**(2), 295–307 (2001). https://doi.org/10.1287/mnsc.47.2.295.9834
42. Iancu, D.A., Sharma, M., Sviridenko, M.: Supermodularity and affine policies in dynamic robust optimization. Oper. Res. **61**(4), 941–956 (2013). https://doi.org/10.1287/opre.2013.1172
43. Kuhn, D., Wiesemann, W., Georghiou, A.: Primal and dual linear decision rules in stochastic and robust optimization. Math. Program. **130**(1), 177–209 (2011). https://doi.org/10.1007/s10107-009-0331-4
44. Laguna, M.: Applying robust optimization to capacity expansion of one location in telecommunications with demand uncertainty. Manag. Sci. **44**(11–part–2), S101–S110 (1998). https://doi.org/10.1287/mnsc.44.11.s101
45. Lu, M., Shen, Z.J.M.: A review of robust operations management under model uncertainty. Prod. Oper. Manag. **30**(6), 1927–1943 (2021). https://doi.org/10.1111/poms.13239
46. Marandi, A., den Hertog, D.: When are static and adjustable robust optimization problems with constraint-wise uncertainty equivalent? Math. Program. **170**(2), 555–568 (2018). https://doi.org/10.1007/s10107-017-1166-z
47. Mudchanatongsuk, S., Ordóñez, F., Liu, J.: Robust solutions for network design under transportation cost and demand uncertainty. J. Oper. Res. Soc. **59**(5), 652–662 (2008). https://doi.org/10.1057/palgrave.jors.2602362
48. Mulvey, J.M., Vanderbei, R.J., Zenios, S.A.: Robust optimization of large-scale systems. Oper. Res. **43**(2), 264–281 (1995). https://doi.org/10.1287/opre.43.2.264

49. Noyan, N., Balcik, B., Atakan, S.: A stochastic optimization model for designing last mile relief networks. Transport. Sci. **50**(3), 1092–1113 (2016). https://doi.org/10.1287/trsc.2015.0621
50. Ordóñez, F., Zhao, J.: Robust capacity expansion of network flows. Networks **50**(2), 136–145 (2007). https://doi.org/10.1002/net.20183
51. Paraskevopoulos, D., Karakitsos, E., Rustem, B.: Robust capacity planning under uncertainty. Manag. Sci. **37**(7), 787–800 (1991). https://doi.org/10.1287/mnsc.37.7.787
52. Porteus, E.L.: Stochastic inventory theory. In: Handbooks in Operations Research and Management Science, pp. 605–652. Elsevier, Amsterdam (1990). https://doi.org/10.1016/s0927-0507(05)80176-8
53. Postek, K., den Hertog, D.: Multistage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set. INFORMS J. Comput. **28**(3), 553–574 (2016). https://doi.org/10.1287/ijoc.2016.0696
54. Rahal, S., Papageorgiou, D.J., Li, Z.: Hybrid strategies using linear and piecewise-linear decision rules for multistage adaptive linear optimization. Eur. J. Oper. Res. **290**(3), 1014–1030 (2021). https://doi.org/10.1016/j.ejor.2020.08.054
55. Schmidt, R.L.: A stochastic optimization model to improve production planning and R&D resource allocation in biopharmaceutical production processes. Manag. Sci. **42**(4), 603–617 (1996). https://doi.org/10.1287/mnsc.42.4.603
56. Sedgewick, R., Flajolet, P.: An Introduction to the Analysis of Algorithms. Addison-Wesley, Boston (2013)
57. See, C.T., Sim, M.: Robust approximation to multiperiod inventory management. Oper. Res. **58**(3), 583–594 (2010). https://doi.org/10.1287/opre.1090.0746
58. Simchi-Levi, D., Trichakis, N., Zhang, P.Y.: Designing response supply chain against bioattacks. Oper. Res. **67**(5), 1246–1268 (2019). https://doi.org/10.1287/opre.2019.1862
59. Simchi-Levi, D., Wang, H., Wei, Y.: Constraint generation for two-stage robust network flow problems. INFORMS J. Optim. **1**(1), 49–70 (2019). https://doi.org/10.1287/ijoo.2018.0003
60. Solyalı, O., Cordeau, J.F., Laporte, G.: The impact of modeling on robust inventory management under demand uncertainty. Manag. Sci. **62**(4), 1188–1201 (2016). https://doi.org/10.1287/mnsc.2015.2183
61. Wang, Z., Qi, M.: Robust service network design under demand uncertainty. Transport. Sci. **54**(3), 676–689 (2020). https://doi.org/10.1287/trsc.2019.0935
62. Xie, F., Huang, Y.: A multistage stochastic programming model for a multi-period strategic expansion of biofuel supply chain under evolving uncertainties. Transport. Res. Part E Logist. Transport. Rev. **111**, 130–148 (2018). https://doi.org/10.1016/j.tre.2018.01.015
63. Yanıkoğlu, İ, Gorissen, B.L., den Hertog, D.: A survey of adjustable robust optimization. Eur. J. Oper. Res. **277**(3), 799–813 (2019). https://doi.org/10.1016/j.ejor.2018.08.031
64. Zou, J., Ahmed, S., Sun, X.A.: Partially adaptive stochastic optimization for electric power generation expansion planning. INFORMS J. Comput. **30**(2), 388–401 (2018). https://doi.org/10.1287/ijoc.2017.0782

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.