http://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Vol. 68, No. 2, March–April 2020, pp. 572–590
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Methods

A Primal–Dual Lifting Scheme for Two-Stage Robust Optimization

Angelos Georghiou,[a] Angelos Tsoukalas,[b] Wolfram Wiesemann[c]

[a] Desautels Faculty of Management, McGill University, Montreal, Quebec H3A 0G4, Canada; [b] Olayan School of Business, American University of Beirut, 1107-2020 Lebanon; [c] Imperial College Business School, Imperial College London, London SW7 2AZ, United Kingdom

Contact: angelos.georghiou@mcgill.ca, https://orcid.org/0000-0003-4490-4020 (AG); at60@aub.edu.lb, https://orcid.org/0000-0003-2679-3953 (AT); https://orcid.org/0000-0003-3076-1591 (WW)

Received: October 10, 2017
Revised: May 17, 2018; January 3, 2019
Accepted: March 22, 2019
Published Online in Articles in Advance: February 4, 2020

Subject Classification: programming: stochastic
Area of Review: Optimization

https://doi.org/10.1287/opre.2019.1873

Abstract. Two-stage robust optimization problems, in which decisions are taken both in anticipation of and in response to the observation of an unknown parameter vector from within an uncertainty set, are notoriously challenging. In this paper, we develop convergent hierarchies of primal (conservative) and dual (progressive) bounds for these problems that trade off the competing goals of tractability and optimality: Although the coarsest bounds recover a tractable but suboptimal affine decision rule approximation of the two-stage robust optimization problem, the refined bounds lift extreme points of the uncertainty set until an exact but intractable extreme point reformulation of the problem is obtained. Based on these bounds, we propose a primal–dual lifting scheme for the solution of two-stage robust optimization problems that accommodates for discrete, here-and-now decisions, infeasible problem instances, and the absence of a relatively complete recourse. The incumbent solutions in each step of our algorithm afford rigorous error bounds, and they can be interpreted as piecewise affine decision rules. We illustrate the performance of our algorithm on illustrative examples and on an inventory management problem.

Funding: The authors gratefully acknowledge funding from the Imperial Business Analytics Centre and the EPSRC [Grants EP/M027856/1, EP/M028240/1, and EP/N020030/1].
Supplemental Material: The online appendix is available at https://doi.org/10.1287/opre.2019.1873.

Keywords: robust optimization • two-stage problems • decision rules • error bounds

1. Introduction

In the last two decades, robust optimization has emerged as a powerful methodology for immunizing mathematical programs against uncertainty in the problem data. Many dynamic optimization problems can be naturally formulated as *two-stage robust optimization problems* of the form

$$\begin{array}{lll} \text{minimize} & \boldsymbol{q}^\top \boldsymbol{x} & \\ \text{subject to} & \boldsymbol{T}(\boldsymbol{\xi})\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}(\boldsymbol{\xi}) \geq \boldsymbol{h}(\boldsymbol{\xi}) & \forall \boldsymbol{\xi} \in \Xi \quad (\mathcal{P}) \\ & \boldsymbol{x} \in \mathbb{R}^{n_1}, \;\; \boldsymbol{y} : \Xi \mapsto \mathbb{R}^{n_2}, & \end{array}$$

where $\boldsymbol{q} \in \mathbb{R}^{n_1}$ and $\boldsymbol{W} \in \mathbb{R}^{m \times n_2}$, and $\boldsymbol{T} : \Xi \overset{A}{\mapsto} \mathbb{R}^{m \times n_1}$ and $\boldsymbol{h} : \Xi \overset{A}{\mapsto} \mathbb{R}^m$ are affine functions of the uncertain parameter vector $\boldsymbol{\xi}$, which is only known to reside in the nonempty and bounded uncertainty set $\Xi = \{\boldsymbol{\xi} \in \mathbb{R}^k : \boldsymbol{F}\boldsymbol{\xi} \leq \boldsymbol{g}\}$, $\boldsymbol{F} \in \mathbb{R}^{l \times k}$, and $\boldsymbol{g} \in \mathbb{R}^l$. Problem $\mathcal{P}$ determines a first-stage decision $\boldsymbol{x}$, which does not depend on the realization of $\boldsymbol{\xi}$, as well as a second-stage policy $\boldsymbol{y}(\boldsymbol{\xi})$, which can adapt to the realization of $\boldsymbol{\xi}$, which are immunized against all parameter realizations $\boldsymbol{\xi} \in \Xi$. Without loss of generality, we may assume that the objective function only involves the first-stage decision $\boldsymbol{x}$ as we can move the second-stage cost to the constraints via an epigraph reformulation.

The two-stage robust optimization problem $\mathcal{P}$ has been used in diverse application domains, ranging from network design and operations, such as network flow problems (Atamtürk and Zhang 2007) and vehicle routing (Gounaris et al. 2013, Subramanyam et al. 2017b), to railway shunting and timetabling (Liebchen et al. 2009), energy systems (Warrington et al. 2013; Lorca and Sun 2015, 2017; Ruiz and Conejo 2015), the strategic (An et al. 2014) and operative (Ardestani-Jaafari and Delage 2016, Lappas and Gounaris 2016) aspects of operations management, and healthcare (Subramanyam et al. 2017b). It is also frequently used to determine approximately optimal solutions to more generic (but, at the same time, computationally more challenging) multistage robust optimization problems (Delage and Iancu 2015). For a detailed review of the applications of problem $\mathcal{P}$, we refer the reader to Bertsimas et al. (2011a), Gabrel et al. (2014), and Yanıkoğlu et al. (2019).

The two-stage robust optimization problem $\mathcal{P}$ is convex, but it contains infinitely many decision variables and constraints. In fact, problem $\mathcal{P}$ is NP-hard

(Guslitser 2002), and the instances that can be solved in polynomial time are both rare and restrictive (Anderson and Moore 1990, Bertsimas et al. 2010, Gounaris et al. 2013). As a result, much of the existing research has focused on developing tractable conservative approximations to problem $\mathcal{P}$. Early attempts in this direction have proposed to restrict the second-stage decision $\boldsymbol{y}$ to affine (Guslitser 2002, Kuhn et al. 2011), segregated affine (Chen et al. 2008, Chen and Zhang 2009, Goh and Sim 2010), piecewise affine (Georghiou et al. 2015), and algebraic as well as trigonometric polynomial functions (Bertsimas et al. 2011c) (so-called *decision rules*) of the parameters $\boldsymbol{\xi}$. Decision rules have recently been extended to incorporate both continuous and discrete second-stage decisions, either by partitioning the uncertainty set Ξ into hyperrectangles (Vayanos et al. 2011, Gorissen et al. 2015) or by resorting to a semi-infinite solution scheme (Bertsimas and Georghiou 2015). By themselves, decision rule approximations only provide a conservative bound on the optimal value of the two-stage robust optimization problem $\mathcal{P}$. To estimate the incurred suboptimality, decision rules are often combined with progressive bounds that emerge from replacing the uncertainty set Ξ in $\mathcal{P}$ with a finite subset of the parameter realizations. Scenario subsets that lead to good progressive bounds can be obtained from the Lagrange multipliers associated with the optimal solution of the decision rule problem (Hadjiyiannis et al. 2011, Bertsimas and de Ruiter 2016). The decision rule approaches naturally extend to robust optimization problems with more than two stages. The suboptimality of decision rules has been investigated in Bertsimas and Goyal (2010), Bertsimas et al. (2010), and Kuhn et al. (2011). For a survey of the decision rule literature, see Delage and Iancu (2015).

Instead of relying on decision rules, the two-stage robust optimization problem $\mathcal{P}$ can also be conservatively approximated by its K-adaptability problem or via its copositive reformulation. The K-adaptability problem selects K candidate second-stage decisions here and now (that is, before observing the realization of $\boldsymbol{\xi}$) and implements the best of these decisions after the realization of $\boldsymbol{\xi}$ is known. Different K-adaptability solution schemes have been proposed in Bertsimas and Caramanis (2010), Hanasusanto et al. (2015), and Subramanyam et al. (2017a), and their suboptimality has been analyzed in Bertsimas et al. (2011b). The copositive reformulation of problem $\mathcal{P}$, on the other hand, is exact and convex, but it is typically difficult to solve. Tractable conservative approximations to this reformulation can be obtained via semidefinite programming (Xu and Burer 2016, Hanasusanto and Kuhn 2017).

All of the solution schemes reviewed so far have in common that they provide a conservative approximation to the two-stage robust optimization problem $\mathcal{P}$. Problem $\mathcal{P}$ can be solved *exactly* through an iterative approximation of its worst-case, second-stage cost function or its uncertainty set. An iterative approximation of the second-stage cost function can be obtained through a variant of Benders' decomposition (Jiang et al. 2010, Thiele et al. 2010, Bertsimas et al. 2013, Zhao et al. 2013). To this end, problem $\mathcal{P}$ is decomposed into a convex master problem involving the first-stage decisions and an outer (progressive) approximation of the worst-case, second-stage cost as well as a nonconvex subproblem that provides cuts for the cost function. The Benders decomposition scheme has been extended to the multistage version of problem $\mathcal{P}$ in Georghiou et al. (2016). Alternatively, Zeng and Zhao (2013) and Ayoub and Poss (2016) propose a column- and constraint-generation scheme, which is based on semi-infinite programming techniques, to iteratively approximate the uncertainty set in problem $\mathcal{P}$. Here, the convex master problem is a relaxation of problem $\mathcal{P}$ that involves finitely many realizations $\boldsymbol{\xi} \in \Xi$, and the nonconvex subproblem identifies parameter realizations $\boldsymbol{\xi} \in \Xi$ to be added to the master problem. Both the Benders decomposition approaches and the semi-infinite programming schemes rely on sequences of progressive approximations to determine an optimal solution to problem $\mathcal{P}$ in finite time. In contrast, the iterative solution schemes presented in Bertsimas and Dunning (2016) and Postek and den Hertog (2016) extend the uncertainty set partitioning approaches of Vayanos et al. (2011) and Gorissen et al. (2015) to construct a sequence of conservative approximations to problem $\mathcal{P}$. Here, the convex master problem determines constant or affine decision rules for each set of the partition, and the subproblem identifies a refined partition for the master problem. Both approaches extend to integer decisions and more than two stages. It has been shown, however, that even asymptotic convergence to an optimal solution of problem $\mathcal{P}$ cannot be guaranteed in general (Bertsimas and Dunning 2016).

Instead of approximating the second-stage decisions, the second-stage cost function, or the uncertainty set of the two-stage robust optimization problem $\mathcal{P}$, it has been proposed in Zhen et al. (2016) to solve $\mathcal{P}$ through an iterative reformulation of the problem itself. The authors use Fourier–Motzkin elimination to reduce the number of second-stage decisions in problem $\mathcal{P}$ at the expense of additional constraints. This results in a hierarchy of increasingly accurate conservative approximations of $\mathcal{P}$, which converge to a static robust optimization problem that is equivalent to $\mathcal{P}$.

In this paper, we develop an alternative solution scheme for the two-stage robust optimization problem $\mathcal{P}$ that aims to provide an attractive trade-off between the conflicting objectives of optimality and

tractability. We summarize our key contributions as follows.

1. We develop convergent hierarchies of primal (conservative) and dual (progressive) bounds to the two-stage robust optimization problem $\mathcal{P}$. Our bounds combine affine decision rules with an extreme point enumeration to trade off the conflicting goals of tractability and optimality. Although the primal bounds apply to any bounded polyhedral uncertainty set, the dual bounds require information about the sum of outer products of the extreme points of the uncertainty set, which can be computed in closed form for several classes of common uncertainty sets.

2. We propose a primal–dual lifting scheme that is inspired by polyhedral combinatorics. Our solution approach accommodates for discrete here-and-now decisions and infeasible problem instances as well as the absence of a relatively complete recourse. The initial bounds are based on affine decision rules and can, thus, be computed efficiently.

3. We highlight the intimate relationship between our bounds and piecewise affine decision rules over simplicial decompositions of the uncertainty set Ξ in problem $\mathcal{P}$.

We believe that the proposed approach fills a gap in the literature: Both the Benders decomposition schemes and the semi-infinite programming approaches reviewed typically consider instances of problem $\mathcal{P}$ with right-hand uncertainty and a relatively complete recourse. Moreover, although the master problems in these approaches provide tractable progressive bounds, the subproblems providing the conservative bounds constitute biaffine or mixed-integer optimization problems that are difficult to solve. Although the adaptive uncertainty partitioning approaches can provide tractable conservative and progressive bounds, they do not offer a convergence guarantee. To our best knowledge, the Fourier–Motzkin elimination scheme presented in Zhen et al. (2016) is the only approach that satisfies all of the properties outlined here. We show in our numerical experiments that, although both the Fourier–Motzkin elimination scheme and our approach are outperformed by the column- and constraint-generation scheme (Zeng and Zhao 2013) on instances of an inventory-management problem with relatively complete recourse, our approach outperforms both alternatives on a variant of the problem that lacks a relatively complete recourse. We, thus, believe that our lifting approach provides a promising complementary method to solve two-stage robust optimization problems.

The remainder of the paper is structured as follows. Section 2 presents our hierarchies of primal and dual bounds to problem $\mathcal{P}$, and Section 3 employs these bounds to develop a primal–dual lifting scheme for the solution of problem $\mathcal{P}$. We discuss the relationship between our bounds and piecewise affine decision rules in Section 4, and we report on numerical experiments in Section 5. The online appendix shows how to compute the sum of outer products of the extreme points of an uncertainty set, which is required for our dual bound, for several classes of common uncertainty sets.

Remark 1 (Discrete Here-and-Now Decisions). Throughout the paper, we discuss how our algorithm extends to a variant of problem $\mathcal{P}$ that contains continuous and discrete first-stage decisions:

$$\begin{array}{ll} \text{minimize} & \boldsymbol{q}_{\mathrm{c}}^{\top}\boldsymbol{x}_{\mathrm{c}} + \boldsymbol{q}_{\mathrm{d}}^{\top}\boldsymbol{x}_{\mathrm{d}} \\ \text{subject to} & \boldsymbol{T}_{\mathrm{c}}(\boldsymbol{\xi})\,\boldsymbol{x}_{\mathrm{c}} + \boldsymbol{T}_{\mathrm{d}}(\boldsymbol{\xi})\,\boldsymbol{x}_{\mathrm{d}} + \boldsymbol{W}\boldsymbol{y}(\boldsymbol{\xi}) \geq \boldsymbol{h}(\boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \in \Xi \\ & \boldsymbol{x}_{\mathrm{c}} \in \mathbb{R}^{nc_1},\ \ \boldsymbol{x}_{\mathrm{d}} \in \mathbb{Z}^{nd_1},\ \ \boldsymbol{y} : \Xi \mapsto \mathbb{R}^{n_2}. \qquad (\mathcal{P}_{\mathrm{d}}) \end{array}$$

We see that our bounds and the lifting scheme naturally extend to this more generic problem.

1.1. Notation

For a finite-dimensional set Ω, we denote by $\operatorname{ext}\Omega$, $\operatorname{conv}\Omega$, and $\operatorname{cl}\Omega$ the set of extreme points, the convex hull, and the closure of Ω, respectively. We define by $\{f : \mathbb{R}^n \overset{A}{\mapsto} \mathbb{R}^m\}$ the set of all affine functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. The vector of ones and the identity matrix are denoted by $\mathbf{e}$ and $\boldsymbol{I}$, respectively, and their dimensions are clear from the context. We use $\langle \boldsymbol{A}, \boldsymbol{B}\rangle = \operatorname{tr}(\boldsymbol{A}^{\top}\boldsymbol{B})$ to denote the trace inner product between two symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$.

2. A Duality Scheme for Two-Stage Robust Optimization

Problem $\mathcal{P}$ is a convex but challenging optimization problem as it involves infinitely many decision variables and constraints. This can be partially resolved if we replace the uncertainty set Ξ in $\mathcal{P}$ with its extreme points $\operatorname{ext}\Xi$:

$$\begin{array}{ll} \text{minimize} & \boldsymbol{q}^{\top}\boldsymbol{x} \\ \text{subject to} & \boldsymbol{T}(\boldsymbol{\xi})\,\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}(\boldsymbol{\xi}) \geq \boldsymbol{h}(\boldsymbol{\xi}) \qquad \forall \boldsymbol{\xi} \in \operatorname{ext}\Xi \quad (\mathcal{P}') \\ & \boldsymbol{x} \in \mathbb{R}^{n_1},\ \ \boldsymbol{y} : \operatorname{ext}\Xi \mapsto \mathbb{R}^{n_2}, \end{array}$$

and in the following, we refer to this problem as the *extreme point reformulation* $\mathcal{P}'$.

Proposition 1 (Extreme Point Reformulation). *Problems $\mathcal{P}$ and $\mathcal{P}'$ are equivalent in the following sense: any feasible solution to one problem can be transformed into a feasible solution to the other problem that attains the same objective value.*

Proof of Proposition 1. One readily verifies that, for every feasible solution $\boldsymbol{x} \in \mathbb{R}^{n_1}$ and $\boldsymbol{y} : \Xi \mapsto \mathbb{R}^{n_2}$ to the robust optimization problem $\mathcal{P}$, the restriction of $\boldsymbol{y}$ to $\operatorname{ext}\Xi$ provides a feasible solution to the extreme point reformulation $\mathcal{P}'$ that attains the same objective value.

Let $\boldsymbol{x} \in \mathbb{R}^{n_1}$ and $\boldsymbol{y}' : \operatorname{ext}\Xi \mapsto \mathbb{R}^{n_2}$ be a feasible solution to the extreme point reformulation $\mathcal{P}'$.

By construction, there is a function $\lambda : \Xi \times \operatorname{ext}\Xi \mapsto \mathbb{R}_+$ that satisfies

$$\sum_{\xi' \in \operatorname{ext}\Xi} \lambda(\xi, \xi') = 1 \qquad \text{and} \qquad \sum_{\xi' \in \operatorname{ext}\Xi} \lambda(\xi, \xi') \cdot \xi' = \xi \qquad \forall \xi \in \Xi.$$

Hence, (x, y) with $y(\xi) := \sum_{\xi' \in \operatorname{ext}\Xi} \lambda(\xi, \xi') \cdot y'(\xi')$, $\xi \in \Xi$, is feasible in problem $\mathcal{P}$ because

$$\begin{aligned} T(\xi)\, x + W y(\xi) &= \sum_{\xi' \in \operatorname{ext}\Xi} \lambda(\xi, \xi') \cdot \left[T(\xi')\, x + W y'(\xi')\right] \\ &\geq \sum_{\xi' \in \operatorname{ext}\Xi} \lambda(\xi, \xi') \cdot h(\xi') = h(\xi) \qquad \forall \xi \in \Xi, \end{aligned}$$

where the first identity follows from the definition of (x, y) and the fact that the mapping T is affine, the inequality follows from the feasibility of (x, y') in the extreme point reformulation $\mathcal{P}'$, and the second identity holds because $\sum_{\xi' \in \operatorname{ext}\Xi} \lambda(\xi, \xi') \cdot \xi' = \xi$. The statement now follows from the fact that (x, y) attains the same objective value in $\mathcal{P}$ as (x, y') does in $\mathcal{P}'$. □

Note that the proof of Proposition 1 does not rely on the characteristics of the first-stage decision x. Hence, the statement immediately extends to the generalized two-stage robust optimization problem $\mathcal{P}_d$ from Remark 1 that contains both continuous and discrete here-and-now decisions.

The extreme point reformulation $\mathcal{P}'$ is finite-dimensional, but it remains computationally burdensome because the number of decision variables and constraints scales with the number of extreme points of Ξ. By the dual upper bound theorem (Seidel 2004, section 26.3), this number satisfies

$$|\operatorname{ext}\Xi| \leq \binom{l - \lceil k/2 \rceil}{\lfloor k/2 \rfloor} + \binom{l - 1 - \lceil (k-1)/2 \rceil}{\lfloor (k-1)/2 \rfloor},$$

and the bound is attained by dual cyclic polytopes. This superpolynomial growth is to be expected as problem $\mathcal{P}$ is NP-hard (Guslitser 2002, theorem 3.4). We note that, for a fixed number k of uncertain problem parameters, the dual upper bound theorem implies that $|\operatorname{ext}\Xi| \in \mathcal{O}(l^{\lfloor k/2 \rfloor})$. Moreover, the *expected* number of vertices of uniformly sampled polyhedra is much lower than this upper bound (Buchta et al. 1985).

We emphasize that the statement of Proposition 1 does not hold if problem $\mathcal{P}$ exhibits random recourse, that is, if W in $\mathcal{P}$ depends on ξ. This has also been observed in Ayoub and Poss (2016).

Example 1 (Random Recourse). Consider the following instance of $\mathcal{P}$ with random recourse:

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & \xi \cdot y(\xi) \geq 1 \qquad \forall \xi \in [-1, 1] \\ & y : [-1, 1] \mapsto \mathbb{R}. \end{array}$$

This problem is infeasible because any second-stage decision y violates the constraint $0 \cdot y(0) \geq 1$, but its extreme point formulation is solved by any y satisfying $y(-1) \leq -1$ and $y(1) \geq 1$. □

The next two sections develop families of primal and dual bounds on the optimal value of the two-stage robust optimization problem $\mathcal{P}$ that combine the extreme point reformulation $\mathcal{P}'$ with affine decision rules.

2.1. Hierarchy of Primal Bounds

A popular conservative approximation to the two-stage robust optimization problem $\mathcal{P}$ replaces the second-stage decision $y : \Xi \mapsto \mathbb{R}^{n_2}$ with an affine decision rule $y : \Xi \overset{A}{\mapsto} \mathbb{R}^{n_2}$. The resulting problem has finitely many decision variables (the intercept and the slopes of y) but infinitely many constraints. Classical robust optimization techniques then employ linear programming duality to reformulate this semi-infinite problem as a linear program that scales polynomially in the size (n_1, n_2, m, k, l) of the input data (Guslitser 2002). The affine decision rule approximation performs surprisingly well on practical problems (Kuhn et al. 2011), and it is even optimal in some problem classes (Anderson and Moore 1990, Bertsimas et al. 2010, Gounaris et al. 2013). In general, however, affine decision rules are suboptimal even in seemingly benign problems (Georghiou et al. 2015, example 4.5).

In the following, we present a family of conservative approximations to $\mathcal{P}$ that includes the highly tractable but usually suboptimal affine decision rules and the optimal but typically intractable extreme point reformulation $\mathcal{P}'$ as special cases. We parameterize our approximation by the *scenario set* $\Xi_S \subseteq \operatorname{ext}\Xi$, which gives rise to a complementary *affine set* $\Xi_A \subseteq \Xi$ defined by $\Xi_A = \operatorname{conv}([\operatorname{ext}\Xi] \setminus \Xi_S)$:

$$\begin{array}{lll} \text{minimize} & q^\top x & \\ \text{subject to} & T(\xi)\, x + W y_S(\xi) \geq h(\xi) & \forall \xi \in \Xi_S \\ & T(\xi)\, x + W y_A(\xi) \geq h(\xi) & \forall \xi \in \Xi_A \\ & x \in \mathbb{R}^{n_1}, \;\; y_S : \Xi_S \mapsto \mathbb{R}^{n_2}, \;\; y_A : \Xi_A \overset{A}{\mapsto} \mathbb{R}^{n_2}. & (\overline{\mathcal{P}}(\Xi_S)) \end{array}$$

Problem $\overline{\mathcal{P}}(\Xi_S)$ optimizes over individual second-stage decisions $y_S : \Xi_S \mapsto \mathbb{R}^{n_2}$ for all realizations ξ in the scenario set Ξ_S and over an affine decision rule $y_A : \Xi_A \overset{A}{\mapsto} \mathbb{R}^{n_2}$ for the affine set Ξ_A. The scenario set Ξ_S is a finite subset of the extreme points of Ξ, whereas the affine set Ξ_A is a polyhedral subset of Ξ. Robust optimization techniques allow us to reformulate problem $\overline{\mathcal{P}}(\Xi_S)$ as a linear program that scales polynomially in the parameters $(n_1, n_2, m, |\Xi_S|, k, \ell)$, where ℓ denotes the number of inequalities required to describe Ξ_A (see next page). The choices $\Xi_S = \emptyset$ (and, thus, $\Xi_A = \Xi$)

and $\Xi_S = \operatorname{ext}\Xi$ (and, thus, $\Xi_A = \emptyset$) recover the affine decision rule approximation and the extreme point reformulation $\mathcal{P}'$, respectively. The requirement that $\Xi_A = \operatorname{conv}([\operatorname{ext}\Xi] \setminus \Xi_S)$ is illustrated in Figure 1, and it ensures that

$$\begin{aligned}\operatorname{conv}(\Xi_S \cup \Xi_A) &= \operatorname{conv}(\Xi_S \cup \operatorname{conv}([\operatorname{ext}\Xi] \setminus \Xi_S)) \\ &= \operatorname{conv}(\Xi_S \cup ([\operatorname{ext}\Xi] \setminus \Xi_S)) \\ &= \operatorname{conv}(\operatorname{ext}\Xi) = \Xi.\end{aligned}$$

In the following, we use $\overline{\mathcal{P}}(\Xi_S)$ both to refer to the bounding problem and to its optimal objective value. We now show that $\overline{\mathcal{P}}(\Xi_S)$ bounds problem $\mathcal{P}$ from above.

Proposition 2 (Primal Bound). *Problem $\overline{\mathcal{P}}(\Xi_S)$ satisfies the following two properties:*

i. *Any feasible solution to $\overline{\mathcal{P}}(\Xi_S)$ can be transformed into a feasible solution to $\overline{\mathcal{P}}(\Xi_S')$, $\Xi_S \subseteq \Xi_S'$ which attains the same objective value.*

ii. *Any feasible solution to $\overline{\mathcal{P}}(\operatorname{ext}\Xi)$ can be transformed into a feasible solution to $\mathcal{P}$ that attains the same objective value and vice versa.*

Proof of Proposition 2. As for the first statement, let $(\boldsymbol{x}, \boldsymbol{y}_S, \boldsymbol{y}_A)$ be a feasible solution to $\overline{\mathcal{P}}(\Xi_S)$. In that case, $(\boldsymbol{x}, \boldsymbol{y}_S', \boldsymbol{y}_A')$ with $\boldsymbol{y}_S'(\boldsymbol{\xi}) := \boldsymbol{y}_S(\boldsymbol{\xi})$ for $\boldsymbol{\xi} \in \Xi_S$; $:= \boldsymbol{y}_A(\boldsymbol{\xi})$ for $\boldsymbol{\xi} \in \Xi_S' \setminus \Xi_S$ and $\boldsymbol{y}_A'(\boldsymbol{\xi}) := \boldsymbol{y}_A(\boldsymbol{\xi})$, $\boldsymbol{\xi} \in \Xi_A'$ is a feasible solution to $\overline{\mathcal{P}}(\Xi_S')$ that attains the same objective value.

The second statement directly follows from Proposition 1. □

The second property states that $\overline{\mathcal{P}}(\operatorname{ext}\Xi)$ is equivalent to problem $\mathcal{P}$. The first property implies that $\overline{\mathcal{P}}(\Xi_S)$ provides a conservative approximation to $\overline{\mathcal{P}}(\Xi_S')$ if $\Xi_S \subseteq \Xi_S'$, and in particular—by virtue of the second statement—to problem $\mathcal{P}$. We remark that the second statement also holds true for $\Xi_S \neq \operatorname{ext}\Xi$ if the implied affine set Ξ_A is a simplex (Bertsimas and Goyal 2012, theorem 1).

The proof of Proposition 2 does not exploit any properties of the first-stage decision $\boldsymbol{x}$, and hence, the result immediately extends to the mixed-integer, two-stage robust optimization problem $\mathcal{P}_\mathrm{d}$.

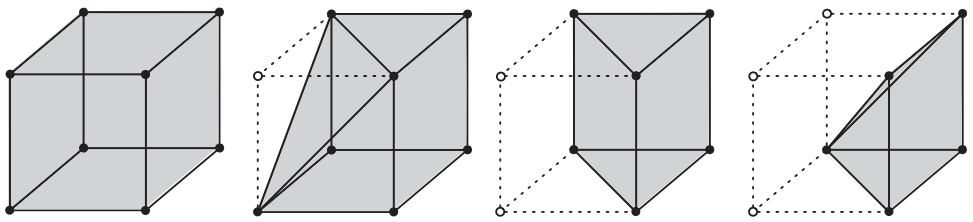
The size of problem $\overline{\mathcal{P}}(\Xi_S)$ depends on the number of scenarios $|\Xi_S|$ and the number of inequalities ℓ required to describe Ξ_A, both of which depend on the choice of the uncertainty set Ξ. If Ξ is the one-norm ball in $\mathbb{R}^k$, $\Xi = \{\boldsymbol{\xi} \in \mathbb{R}^k : \|\boldsymbol{\xi}\|_1 \leq 1\}$, for example, then Ξ_S contains at most $|\operatorname{ext}\Xi| = 2k$ scenarios, and any set Ξ_A can be described efficiently by (subsets of) the projection of $\Xi' = \{(\boldsymbol{\xi}, \boldsymbol{\chi}) \in \mathbb{R}^k \times \mathbb{R}^k : \mathbf{e}^\top \boldsymbol{\chi} \leq 1,\ \boldsymbol{\chi} \geq \pm\boldsymbol{\xi}\}$ onto its first k components. If the uncertainty set is the ∞-norm ball in $\mathbb{R}^k$, $\Xi = \{\boldsymbol{\xi} \in \mathbb{R}^k : \|\boldsymbol{\xi}\|_\infty \leq 1\}$, on the other hand, then Ξ_S can contain up to $|\operatorname{ext}\Xi| = 2^k$ scenarios, and the sets Ξ_A form the class of zero/one polytopes that can have exponentially many facets (Ziegler 2000). The hope is that, in practice, instances of $\overline{\mathcal{P}}(\Xi_S)$ with compact descriptions of Ξ_A and Ξ_S produce tight bounds on the optimal value of problem $\mathcal{P}$. Section 3 presents an iterative procedure that aims to determine such instances, and Section 5 investigates the performance of this procedure in an inventory-management problem.

2.2. Hierarchy of Dual Bounds

Hadjiyiannis et al. (2011) propose a progressive bound for the two-stage robust optimization problem $\mathcal{P}$ that replaces the uncertainty set Ξ with a subset of its extreme points $\operatorname{ext}\Xi$. This bound has found widespread use in both two-stage (Zeng and Zhao 2013, Ayoub and Poss 2016) and multistage (Bertsimas and Georghiou 2015, Bertsimas and Dunning 2016, Postek and den Hertog 2016) robust optimization. In this section, we derive a family of dual bounds that tighten the bound of Hadjiyiannis et al. (2011). Our bounds can be interpreted as an application of the primal approximation from the previous section to a dual extreme point formulation of $\mathcal{P}$. To this end, we consider the linear programming dual of the extreme point reformulation $\mathcal{P}'$:

$$\begin{array}{lll}\text{maximize} & \displaystyle\sum_{\boldsymbol{\xi}\in\operatorname{ext}\Xi} \boldsymbol{h}(\boldsymbol{\xi})^\top \boldsymbol{\lambda}(\boldsymbol{\xi}) & \\ \text{subject to} & \displaystyle\sum_{\boldsymbol{\xi}\in\operatorname{ext}\Xi} \boldsymbol{T}(\boldsymbol{\xi})^\top \boldsymbol{\lambda}(\boldsymbol{\xi}) = \boldsymbol{q} & \\ & \boldsymbol{W}^\top \boldsymbol{\lambda}(\boldsymbol{\xi}) = \mathbf{0} \qquad \forall \boldsymbol{\xi} \in \operatorname{ext}\Xi & \\ & \boldsymbol{\lambda} : \operatorname{ext}\Xi \mapsto \mathbb{R}^m_+. & \end{array} \tag{1}$$

Figure 1. Different Decompositions of a Three-Dimensional Hypercube into Scenario Sets Ξ_S (Hollow Circles) and Implied Affine Sets Ξ_A (Shaded Regions)



Note. In all cases, the convex hull of Ξ_S and Ξ_A recovers the original hypercube.

Contrary to the primal extreme point reformulation $\mathcal{P}'$, problem (1) is a maximization problem. Therefore, a conservative approximation to problem (1) amounts to a progressive approximation to problem $\mathcal{P}'$ and, hence, to problem $\mathcal{P}$. In order to conservatively approximate problem (1), we proceed along the lines of the previous section: we partition the extreme points $\operatorname{ext}\Xi$ into a scenario set Ξ_S, for which we select individual decisions $\lambda_S : \Xi_S \mapsto \mathbb{R}^m_+$ and an implied affine set $\Xi_A = \operatorname{conv}([\operatorname{ext}\Xi] \setminus \Xi_S)$ for which the decisions $\lambda_A : \Xi_A \overset{A}{\mapsto} \mathbb{R}^m_+$ are restricted to an affine function of ξ. Contrary to the primal bounding problem $\overline{\mathcal{P}}(\Xi_S)$, however, the uncertain problem parameters ξ enter the dual extreme point reformulation (1) quadratically, and they, thus, require some additional care. One can then derive the following family of dual bounds to problem $\mathcal{P}$:

$$\begin{array}{lll}
\text{maximize} & \langle \boldsymbol{H}^\top \boldsymbol{\Lambda}, \boldsymbol{\Sigma}_A \rangle + \displaystyle\sum_{\xi\in\Xi_S} h(\xi)^\top \lambda_S(\xi) & \\
\text{subject to} & \langle \boldsymbol{T}_i^\top \boldsymbol{\Lambda}, \boldsymbol{\Sigma}_A \rangle + \displaystyle\sum_{\xi\in\Xi_S} \boldsymbol{T}_i(\xi)^\top \lambda_S(\xi) = q_i & \\
 & & \forall i = 1, \ldots, n_1 \\
 & \boldsymbol{W}^\top \lambda_S(\xi) = \boldsymbol{0} & \forall \xi \in \Xi_S \\
 & \boldsymbol{W}^\top \lambda_A(\xi) = \boldsymbol{0} & \forall \xi \in \Xi_A \\
 & \lambda_S : \Xi_S \mapsto \mathbb{R}^m_+, \quad \lambda_A : \Xi_A \overset{A}{\mapsto} \mathbb{R}^m_+. & \qquad (\underline{\mathcal{P}}(\Xi_S))
\end{array}$$

In this problem, we use the notational shorthands $\boldsymbol{H} = (\boldsymbol{h}_0, \boldsymbol{H}_J)$, $\boldsymbol{T}_i = (\boldsymbol{t}_{i,0}, \boldsymbol{T}_{i,J})$, and $\boldsymbol{\Lambda} = (\lambda_0, \boldsymbol{\Lambda}_J)$, where $\boldsymbol{h}_0, \boldsymbol{t}_{i,0}, \lambda_0 \in \mathbb{R}^m$ denote the intercepts and $\boldsymbol{H}_J, \boldsymbol{T}_{i,J}, \boldsymbol{\Lambda}_J \in \mathbb{R}^{m\times k}$ the slopes (Jacobians) of the affine functions $\boldsymbol{h}(\xi)$, $\boldsymbol{T}_i(\xi) = (T_{1,i}(\xi), \ldots, T_{m,i}(\xi))^\top$, and $\lambda_A(\xi)$, and the *moment matrix*

$$\boldsymbol{\Sigma}_A = \sum_{\xi\in\operatorname{ext}\Xi_A} \begin{pmatrix} 1 \\ \xi \end{pmatrix} \begin{pmatrix} 1 \\ \xi \end{pmatrix}^\top = \begin{pmatrix} |\operatorname{ext}\Xi_A| & \displaystyle\sum_{\xi\in\operatorname{ext}\Xi_A} \xi^\top \\ \displaystyle\sum_{\xi\in\operatorname{ext}\Xi_A} \xi & \displaystyle\sum_{\xi\in\operatorname{ext}\Xi_A} \xi\xi^\top \end{pmatrix}$$

records the sum of the extreme point scenarios in Ξ_A as well as their outer products. Standard robust optimization techniques allow us to reformulate problem $\underline{\mathcal{P}}(\Xi_S)$ as a linear program that scales polynomially in the parameters $(n_1, n_2, m, |\Xi_S|, k, \ell)$, where ℓ denotes the number of inequalities in the description of Ξ_A (see discussion before Remark 3). The choice $\Xi_S = \emptyset$ (and, thus, $\Xi_A = \Xi$) corresponds to a highly tractable but usually suboptimal dual affine decision rule formulation, whereas $\Xi_S = \operatorname{ext}\Xi$ (and, thus, $\Xi_A = \emptyset$) recovers the optimal but typically intractable dual extreme point reformulation (1). We use $\underline{\mathcal{P}}(\Xi_S)$ to refer to both the bounding problem and its optimal objective value.

We now formalize our reasoning that $\underline{\mathcal{P}}(\Xi_S)$ bounds problem $\mathcal{P}$ from the following.

Proposition 3 (Dual Bound). *Problem $\underline{\mathcal{P}}(\Xi_S)$ satisfies the following two properties:*

i. *Any feasible solution to $\underline{\mathcal{P}}(\Xi_S)$ can be transformed into a feasible solution to $\underline{\mathcal{P}}(\Xi'_S)$, $\Xi_S \subseteq \Xi'_S$, which attains the same objective value.*

ii. *$\underline{\mathcal{P}}(\operatorname{ext}\Xi)$ is infeasible if $\mathcal{P}$ is unbounded and vice versa. If both problems are feasible, then their optimal values coincide.*

Proof of Proposition 3. By weak linear programming duality, the dual extreme point reformulation (1) bounds the extreme point reformulation $\mathcal{P}'$—and hence, by virtue of Proposition 1, the two-stage robust optimization problem $\mathcal{P}$—from below. In particular, (1) is infeasible if $\mathcal{P}$ is unbounded and vice versa. Moreover, strong linear programming duality implies that both problems attain the same optimal value if they are both feasible. The second statement of the proposition now follows from the fact that the dual extreme point reformulation (1) is equivalent to $\underline{\mathcal{P}}(\operatorname{ext}\Xi)$.

As for the first statement, fix a feasible solution (λ_S, λ_A) for the problem $\underline{\mathcal{P}}(\Xi_S)$. We claim that the solution (λ'_S, λ'_A) defined through $\lambda'_S(\xi) := \lambda_S(\xi)$ for $\xi \in \Xi_S$; $:= \lambda_A(\xi)$ for $\xi \in \Xi'_S \setminus \Xi_S$ and $\lambda'_A(\xi) := \lambda_A(\xi)$, $\xi \in \Xi'_A$, is feasible in $\underline{\mathcal{P}}(\Xi'_S)$ and attains the same objective value as (λ_S, λ_A) does in $\underline{\mathcal{P}}(\Xi_S)$. Indeed, the objective function value of (λ'_S, λ'_A) in $\underline{\mathcal{P}}(\Xi'_S)$ evaluates to

$$\begin{aligned}
&\langle \boldsymbol{H}^\top \boldsymbol{\Lambda}, \boldsymbol{\Sigma}'_A \rangle + \sum_{\xi\in\Xi'_S} h(\xi)^\top \lambda'_S(\xi) \\
&= \langle \boldsymbol{H}^\top \boldsymbol{\Lambda}, \boldsymbol{\Sigma}'_A \rangle + \sum_{\xi\in\Xi'_S\setminus\Xi_S} h(\xi)^\top \lambda'_S(\xi) + \sum_{\xi\in\Xi_S} h(\xi)^\top \lambda'_S(\xi) \\
&= \langle \boldsymbol{H}^\top \boldsymbol{\Lambda}, \boldsymbol{\Sigma}'_A \rangle + \sum_{\substack{\xi\in\operatorname{ext}\Xi_A: \\ \xi\notin\operatorname{ext}\Xi'_A}} h(\xi)^\top \lambda_A(\xi) + \sum_{\xi\in\Xi_S} h(\xi)^\top \lambda_S(\xi) \\
&= \langle \boldsymbol{H}^\top \boldsymbol{\Lambda}, \boldsymbol{\Sigma}'_A \rangle + \langle \boldsymbol{H}^\top \boldsymbol{\Lambda}, \boldsymbol{\Sigma}_A - \boldsymbol{\Sigma}'_A \rangle + \sum_{\xi\in\Xi_S} h(\xi)^\top \lambda_S(\xi) \\
&= \langle \boldsymbol{H}^\top \boldsymbol{\Lambda}, \boldsymbol{\Sigma}_A \rangle + \sum_{\xi\in\Xi_S} h(\xi)^\top \lambda_S(\xi),
\end{aligned}$$

where $\boldsymbol{\Lambda} = (\lambda_0, \boldsymbol{\Lambda}_J)$ and $\boldsymbol{\Lambda}_J \in \mathbb{R}^{m\times k}$ and $\lambda_0 \in \mathbb{R}^m$ denote the slopes and the intercept of λ_A, respectively, and $\boldsymbol{\Sigma}_A$ and $\boldsymbol{\Sigma}'_A$ are the moment matrices corresponding to the affine sets Ξ_A and Ξ'_A. Here, the second identity follows from the definition of λ'_S and the fact that

$$\operatorname{ext}\Xi_A = \operatorname{ext}(\operatorname{conv}([\operatorname{ext}\Xi] \setminus \Xi_S)) = (\operatorname{ext}\Xi) \setminus \Xi_S,$$

and hence,

$$\begin{aligned}
(\operatorname{ext}\Xi_A) \setminus (\operatorname{ext}\Xi'_A) &= ([\operatorname{ext}\Xi] \setminus \Xi_S) \setminus ([\operatorname{ext}\Xi] \setminus \Xi'_S) \\
&= \Xi'_S \setminus \Xi_S.
\end{aligned}$$

The third identity holds because $\boldsymbol{h}(\xi) = \boldsymbol{H}_J\xi + \boldsymbol{h}_0$ and $\lambda_A(\xi) = \boldsymbol{\Lambda}_J\xi + \lambda_0$, and the last identity follows from the linearity of the inner product operator.

An analogous argument shows that (λ_S', λ_A') satisfies the first constraint in $\underline{\mathcal{P}}(\Xi_S')$, and one readily verifies that the remaining constraints are satisfied as well. Thus, the first statement follows. □

The second property states that the (extended real-valued) optimal values of $\underline{\mathcal{P}}(\text{ext}\Xi)$ and the two-stage robust optimization problem $\mathcal{P}$ coincide if at least one of the problems is feasible. It is possible, however, that both problems are infeasible: this is the case, for example, if $\boldsymbol{T}(\xi) = \boldsymbol{0}$ and $\boldsymbol{W} = \boldsymbol{0}$ but $\boldsymbol{q} \neq \boldsymbol{0}$ and $\boldsymbol{h}(\xi) \not\leq \boldsymbol{0}$ for some $\xi \in \Xi$. Our incremental lifting scheme in Section 3 avoids this pathological case by solely operating on feasible instances of the bounding problems. For such instances, the second statement guarantees that the optimal values of $\underline{\mathcal{P}}(\text{ext}\,\Xi)$ and $\mathcal{P}$ coincide. The first statement implies that $\underline{\mathcal{P}}(\Xi_S)$ provides a progressive approximation to $\underline{\mathcal{P}}(\Xi_S')$ if $\Xi_S \subseteq \Xi_S'$ and, in particular—by virtue of the second statement—to problem $\mathcal{P}$.

Remark 2 (Discrete Here-and-Now Decisions). Perhaps surprisingly, we can derive a convergent hierarchy of dual bounds to the mixed-integer, two-stage robust optimization problem $\mathcal{P}_\text{d}$ as well. To this end, consider the following extreme point reformulation $\mathcal{P}_\text{d}'$ of problem $\mathcal{P}_\text{d}$, where we separately optimize over the discrete and continuous here-and-now decisions:

$$\text{minimize} \quad \boldsymbol{q}_\text{d}^\top \boldsymbol{x}_\text{d} + \left[\begin{array}{ll} \text{minimize} & \boldsymbol{q}_\text{c}^\top \boldsymbol{x}_\text{c} \\ \text{subject to} & \boldsymbol{T}_\text{c}(\xi)\,\boldsymbol{x}_\text{c} + \boldsymbol{W}\boldsymbol{y}(\xi) \\ & \geq [\boldsymbol{h}(\xi) - \boldsymbol{T}_\text{d}(\xi)\,\boldsymbol{x}_\text{d}] \\ & \qquad \forall \xi \in \text{ext}\,\Xi \\ & \boldsymbol{x}_\text{c} \in \mathbb{R}^{nc_1}, \\ & \boldsymbol{y} : \text{ext}\,\Xi \mapsto \mathbb{R}^{n_2} \end{array}\right]$$
$$\text{subject to} \quad \boldsymbol{x}_\text{d} \in \mathbb{Z}^{nd_1}.$$

For any fixed, discrete here-and-now decision $\boldsymbol{x}_\text{d} \in \mathbb{Z}^{nd_1}$, we can replace the embedded minimization problem over the continuous here-and-now decision $\boldsymbol{x}_\text{c} \in \mathbb{R}^{nc_1}$ with its linear programming dual:

$$\text{minimize} \quad \boldsymbol{q}_\text{d}^\top \boldsymbol{x}_\text{d} + \left[\begin{array}{ll} \text{maximize} & \sum_{\xi \in \text{ext}\,\Xi} [\boldsymbol{h}(\xi) \\ & \qquad - \boldsymbol{T}_\text{d}(\xi)\,\boldsymbol{x}_\text{d}]^\top \lambda(\xi) \\ \text{subject to} & \sum_{\xi \in \text{ext}\,\Xi} \boldsymbol{T}_\text{c}(\xi)^\top \lambda(\xi) = \boldsymbol{q}_\text{c} \\ & \boldsymbol{W}^\top \lambda(\xi) = \boldsymbol{0} \\ & \qquad \forall \xi \in \text{ext}\,\Xi \\ & \lambda : \text{ext}\,\Xi \mapsto \mathbb{R}_+^m \end{array}\right]$$
$$\text{subject to} \quad \boldsymbol{x}_\text{d} \in \mathbb{Z}^{nd_1}.$$

Similar arguments as before show that the inner maximization problem can be bounded from below:

$$\text{minimize} \quad \boldsymbol{q}_\text{d}^\top \boldsymbol{x}_\text{d} + \left[\begin{array}{ll} \text{maximize} & \left\langle \left[\boldsymbol{H} - \sum_{i=1}^{nd_1} \boldsymbol{T}_{\text{d},i}\, x_{\text{d},i}\right]^\top \boldsymbol{\Lambda}, \boldsymbol{\Sigma}_A \right\rangle \\ & \qquad + \sum_{\xi \in \Xi_S} [\boldsymbol{h}(\xi) - \boldsymbol{T}_\text{d}(\xi)\,\boldsymbol{x}_\text{d}]^\top \\ & \qquad \cdot \lambda_S(\xi) \\ \text{subject to} & \langle \boldsymbol{T}_{\text{c},i}{}^\top \boldsymbol{\Lambda}, \boldsymbol{\Sigma}_A \rangle \\ & \qquad + \sum_{\xi \in \Xi_S} \boldsymbol{T}_{\text{c},i}(\xi)^\top \lambda_S(\xi) \\ & = q_{\text{c},i} \qquad \forall i = 1, \ldots, nc_1 \\ & \boldsymbol{W}^\top \lambda_S(\xi) = \boldsymbol{0} \quad \forall \xi \in \Xi_S \\ & \boldsymbol{W}^\top \lambda_A(\xi) = \boldsymbol{0} \quad \forall \xi \in \Xi_A \\ & \lambda_S : \Xi_S \mapsto \mathbb{R}_+^m, \\ & \lambda_A : \Xi_A \overset{A}{\mapsto} \mathbb{R}_+^m, \end{array}\right]$$
$$\text{subject to} \quad \boldsymbol{x}_\text{d} \in \mathbb{Z}^{nd_1}.$$

Here, $\boldsymbol{T}_{\text{c},i} = (\boldsymbol{t}_{\text{c},i,0}, \boldsymbol{T}_{\text{c},i,J})$, $\boldsymbol{T}_{\text{d},i} = (\boldsymbol{t}_{\text{d},i,0}, \boldsymbol{T}_{\text{d},i,J})$ with $\boldsymbol{t}_{\text{c},i,0}, \boldsymbol{t}_{\text{d},i,0} \in \mathbb{R}^m$ denote the intercepts and $\boldsymbol{T}_{\text{c},i,J}, \boldsymbol{T}_{\text{d},i,J} \in \mathbb{R}^{m \times k}$ the slopes of the affine functions $\boldsymbol{T}_{\text{c},i}(\xi) = (T_{\text{c},1,i}(\xi), \ldots, T_{\text{c},m,i}(\xi))^\top$ and $\boldsymbol{T}_{\text{d},i}(\xi) = (T_{\text{d},1,i}(\xi), \ldots, T_{\text{d},m,i}(\xi))^\top$. For any fixed, discrete here-and-now decision $\boldsymbol{x}_\text{d} \in \mathbb{Z}^{nd_1}$, the embedded maximization problem is a robust optimization problem that is linear in the decision variables λ_S and λ_A. Using robust optimization techniques to reformulate this semi-infinite problem as a linear program and subsequently invoking strong linear programming duality results in an embedded minimization problem that attains the same optimal value. The resulting min–min problem collapses to the following single-stage, mixed-integer linear program:

$$\begin{array}{ll} \text{minimize} & \boldsymbol{q}_\text{d}{}^\top \boldsymbol{x}_\text{d} + \boldsymbol{q}_\text{c}{}^\top \boldsymbol{\alpha} \\ \text{subject to} & \boldsymbol{h}(\xi) - \boldsymbol{T}_\text{d}(\xi)\,\boldsymbol{x}_\text{d} - \sum_{i=1}^{nc_1} \alpha_i \cdot \boldsymbol{T}_{\text{c},i}(\xi) \\ & \qquad + \boldsymbol{W}\beta(\xi) \leq \boldsymbol{0} \qquad \forall \xi \in \Xi_S \\ & \left[\boldsymbol{H} - \sum_{i=1}^{nd_1} \boldsymbol{T}_{\text{d},i}\, x_{\text{d},i} - \sum_{i=1}^{nc_1} \alpha_i \boldsymbol{T}_{\text{c},i}\right] \boldsymbol{\Sigma}_A \\ & \qquad = \left(\boldsymbol{W}(\gamma^1 - \gamma^2) - \eta \;\; \boldsymbol{W}(\Delta^1 - \Delta^2) - \boldsymbol{Z}\right) \\ & \boldsymbol{F}\delta_\nu^1 \leq \gamma_\nu^1 \cdot \boldsymbol{g},\; \boldsymbol{F}\delta_\nu^2 \leq \gamma_\nu^2 \cdot \boldsymbol{g} \quad \forall \nu = 1, \ldots, n_2 \\ & \boldsymbol{F}\zeta_\mu \leq \eta_\mu \cdot \boldsymbol{g} \qquad \forall \mu = 1, \ldots, m \\ & \boldsymbol{\alpha} \in \mathbb{R}^{nc_1},\; \beta(\xi) \in \mathbb{R}^{n_2},\; \xi \in \Xi_S, \\ & \gamma^1 \in \mathbb{R}_+^{n_2},\; \gamma^2 \in \mathbb{R}_-^{n_2},\; \Delta^1, \Delta^2 \in \mathbb{R}^{n_2 \times k}, \\ & \eta \in \mathbb{R}_+^m,\; \boldsymbol{Z} \in \mathbb{R}^{m \times k} \\ & \boldsymbol{x}_\text{d} \in \mathbb{Z}^{nd_1}. \end{array}$$

In this formulation, we assume that $\Xi_A = \{\boldsymbol{\xi} \in \mathbb{R}^k : \boldsymbol{F}\boldsymbol{\xi} \leq \boldsymbol{g}\}$ with $\boldsymbol{F} \in \mathbb{R}^{l \times k}$ and $\boldsymbol{g} \in \mathbb{R}^l$, the matrices $\boldsymbol{\Delta}^i \in \mathbb{R}^{n_2 \times k}, i = 1, 2$, contain the rows $(\boldsymbol{\delta}^i_\nu)^\top \in \mathbb{R}^k, \nu = 1, \ldots, n_2$, and the matrix $\boldsymbol{Z} \in \mathbb{R}^{m \times k}$ contains the rows $(\boldsymbol{\zeta}_\mu)^\top \in \mathbb{R}^k$, $\mu = 1, \ldots, m$.

We note that a similar reformulation can be derived for the linear, two-stage robust optimization problem $\mathcal{P}$. Because this reformulation lacks the intuitive structure that our dual bounds $\underline{\mathcal{P}}(\Xi_S)$ for the problem $\mathcal{P}$ afford, however, we skip its derivation for the sake of brevity.

As in the previous section, the size of problem $\underline{\mathcal{P}}(\Xi_S)$ depends on the number of scenarios $|\Xi_S|$ and the number of inequalities ℓ required to describe Ξ_A, both of which depend on the choice of the uncertainty set Ξ. In addition, the dual approximation $\underline{\mathcal{P}}(\Xi_S)$ contains the moment matrix $\boldsymbol{\Sigma}_A$, which appears difficult to compute as it involves sums of the (outer products of the) extreme points in the affine set Ξ_A. For several classes of commonly used uncertainty sets, we show in the online appendix how $\boldsymbol{\Sigma}_A$ can be derived analytically when $\Xi_A = \Xi$. In the next section, we present a solution scheme for the two-stage robust optimization problem $\mathcal{P}$, which computes the dual affine decision rule bound $\underline{\mathcal{P}}(\emptyset)$ and, subsequently, removes individual vertices of $\operatorname{ext}\Xi_A$. Thus, if the initial moment matrix $\boldsymbol{\Sigma}_A$ for $\Xi_A = \Xi$ can be computed efficiently, then the subsequent updates of Ξ_A merely require the subtraction of individual extreme points. On the other hand, it remains unclear whether the moment matrices $\boldsymbol{\Sigma}_A$ corresponding to generic polyhedral uncertainty sets can be calculated efficiently, that is, without explicitly enumerating them.

We note that our dual bound $\underline{\mathcal{P}}(\Xi_S)$ is closely related to the lower bound proposed in Hadjiyiannis et al. (2011).

Remark 3 (Relation to the Sampling Bound of Hadjiyiannis et al. (2011)). The progressive bound of Hadjiyiannis et al. (2011) can be interpreted as a relaxation of the extreme point reformulation $\mathcal{P}'$ that only considers the scenarios $\boldsymbol{\xi} \in \Xi_S$ in a single, fixed scenario set $\Xi_S \subseteq \operatorname{ext}\Xi$:

$$\begin{array}{ll} \text{minimize} & \boldsymbol{q}^\top \boldsymbol{x} \\ \text{subject to} & \boldsymbol{T}(\boldsymbol{\xi})\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}(\boldsymbol{\xi}) \geq \boldsymbol{h}(\boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \in \Xi_S \\ & \boldsymbol{x} \in \mathbb{R}^{n_1}, \ \ \boldsymbol{y} : \Xi_S \mapsto \mathbb{R}^{n_2}. \end{array} \qquad (SB)$$

The dual associated with this problem corresponds to an instance of our dual bound $\underline{\mathcal{P}}(\Xi_S)$, where we fix $\boldsymbol{\lambda}_A(\boldsymbol{\xi}) = \boldsymbol{0}$, $\boldsymbol{\xi} \in \Xi_A$, and only optimize over $\boldsymbol{\lambda}_S$. Thus, our bound $\underline{\mathcal{P}}(\Xi_S)$ is at least as tight as the bound of Hadjiyiannis et al. (2011). □

We close this section with a summary of the findings of Propositions 2 and 3:

Theorem 1 (Duality). *The primal and dual bounds $\overline{\mathcal{P}}(\Xi_S)$ and $\underline{\mathcal{P}}(\Xi_S)$ satisfy*

i. *Weak duality: $\underline{\mathcal{P}}(\Xi_S) \leq \mathcal{P} \leq \overline{\mathcal{P}}(\Xi_S)$ for all $\Xi_S \subseteq \operatorname{ext}\Xi$.*

ii. *Strong duality: $\underline{\mathcal{P}}(\Xi_S) = \mathcal{P} = \overline{\mathcal{P}}(\Xi_S)$ for $\Xi_S = \operatorname{ext}\Xi$ if $\mathcal{P}$ is feasible.*

3. Primal–Dual Lifting Scheme

Our solution scheme for problem $\mathcal{P}$ starts with a *feasibility phase*, which determines a feasible solution (or recognizes that no such solution exists), and then proceeds with an *optimality phase*, which computes an optimal solution (or identifies that the problem is unbounded). The algorithm starts with the efficiently computable affine decision rule bounds $\overline{\mathcal{P}}(\emptyset)$ and $\underline{\mathcal{P}}(\emptyset)$ and iteratively transfers extreme points $\boldsymbol{\xi}^\star \in \operatorname{ext}\Xi_A$ of the affine set to the scenario set Ξ_S.

Our algorithm can be summarized as follows.

1. **Initialization.** Set $\Xi_A = \Xi$ and $\Xi_S = \emptyset$.
2. **Feasibility Phase.** Consider the *feasibility problem*

$$\begin{array}{ll} \text{minimize} & v \\ \text{subject to} & \boldsymbol{T}(\boldsymbol{\xi})\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}(\boldsymbol{\xi}) \geq \boldsymbol{h}(\boldsymbol{\xi}) - v\boldsymbol{e} \quad \forall \boldsymbol{\xi} \in \Xi \\ & \boldsymbol{x} \in \mathbb{R}^{n_1}, \ \ \boldsymbol{y} : \Xi \mapsto \mathbb{R}^{n_2}, \ \ v \in \mathbb{R}_+. \end{array}$$

i. Solve $\overline{\mathcal{P}}(\Xi_S)$ and $\underline{\mathcal{P}}(\Xi_S)$ associated with this problem. If the optimal value of $\overline{\mathcal{P}}(\Xi_S)$ is zero, go to step 3: the current solution to $\overline{\mathcal{P}}(\Xi_S)$ is feasible in $\mathcal{P}$. If the optimal value of $\underline{\mathcal{P}}(\Xi_S)$ is strictly positive, terminate: problem $\mathcal{P}$ is infeasible.

ii. Select any $\boldsymbol{\xi}^\star \in \Xi_A^\star$ (defined as follows), update $\Xi_S \leftarrow \Xi_S \cup \{\boldsymbol{\xi}^\star\}$ as well as $\Xi_A \leftarrow \operatorname{conv}([\operatorname{ext}\Xi_A] \setminus \{\boldsymbol{\xi}^\star\})$, and go to back to step 2(i).

3. **Optimality Phase.** Consider the *optimality problem*

$$\begin{array}{ll} \text{minimize} & \boldsymbol{q}^\top \boldsymbol{x} \\ \text{subject to} & \boldsymbol{T}(\boldsymbol{\xi})\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}(\boldsymbol{\xi}) \geq \boldsymbol{h}(\boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \in \Xi \\ & \boldsymbol{x} \in \mathbb{R}^{n_1}, \ \ \boldsymbol{y} : \Xi \mapsto \mathbb{R}^{n_2}. \end{array}$$

i. Solve $\overline{\mathcal{P}}(\Xi_S)$ and $\underline{\mathcal{P}}(\Xi_S)$ associated with this problem. If $\overline{\mathcal{P}}(\Xi_S)$ is unbounded, terminate: problem $\mathcal{P}$ is unbounded. Otherwise, if the optimal values of $\overline{\mathcal{P}}(\Xi_S)$ and $\underline{\mathcal{P}}(\Xi_S)$ coincide, terminate: the current solution to $\overline{\mathcal{P}}(\Xi_S)$ is optimal in $\mathcal{P}$.

ii. Select any $\boldsymbol{\xi}^\star \in \Xi_A^\star$ (defined as follows), update $\Xi_S \leftarrow \Xi_S \cup \{\boldsymbol{\xi}^\star\}$ as well as $\Xi_A \leftarrow \operatorname{conv}([\operatorname{ext}\Xi_A] \setminus \{\boldsymbol{\xi}^\star\})$, and go to back to step 3(i).

The feasibility phase operates on a feasibility variant of problem $\mathcal{P}$ that minimizes the maximum constraint violation. Both bounding problems $\overline{\mathcal{P}}(\Xi_S)$ and $\underline{\mathcal{P}}(\Xi_S)$ associated with this feasibility problem are always feasible because Ξ is bounded. Thus, Propositions 2 and 3 guarantee the validity of the conservative and progressive bounds, and the update step 2(ii) ensures that both bounds converge monotonically. After finitely many iterations, either the conservative bound evaluates to zero (indicating that a feasible solution has been found), or the progressive bound becomes strictly positive (indicating that the problem is infeasible).

If a feasible solution has been found, the algorithm proceeds with the optimality phase and computes conservative and progressive bounds on the problem $\mathcal{P}$ itself. Because a feasible solution has already been determined, these bounds are valid and converge monotonically. After finitely many iterations, either the bounds coincide (indicating that an optimal solution has been found), or the conservative bound evaluates to $-\infty$ (indicating that the problem is unbounded).

The update steps 2(ii) and 3(ii) move one extreme point of the polyhedron Ξ_A to the scenario set Ξ_S. Ideally, we would transfer the extreme point that leads to the greatest improvement of the conservative and progressive bounds. To this end, we define the *binding scenario set* as $\Xi_A^\star = \{\xi \in \operatorname{ext}\Xi_A : \boldsymbol{T}(\xi)\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}(\xi) \not> \boldsymbol{h}(\xi) - v\mathbf{e}\}$ in the feasibility phase and as $\Xi_A^\star = \{\xi \in \operatorname{ext}\Xi_A : \boldsymbol{T}(\xi)\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}(\xi) \not> \boldsymbol{h}(\xi)\}$ in the optimality phase. We now show that we can restrict ourselves to these binding scenarios in the update steps.

Observation 1 (Binding Scenarios). For any scenario set $\Xi_S \subseteq \operatorname{ext}\Xi$, the binding scenario set $\Xi_A^\star$ satisfies the following two properties:

i. If $\xi^\star \notin \Xi_A^\star$, then the update $\Xi_S \leftarrow \Xi_S \cup \{\xi^\star\}$ and $\Xi_A \leftarrow \operatorname{conv}([\operatorname{ext}\Xi_A] \setminus \{\xi^\star\})$ does not improve the objective value of the conservative approximation $\overline{\mathcal{P}}(\Xi_S)$.

ii. If $\Xi_A^\star = \emptyset$, then $\overline{\mathcal{P}}(\Xi_S) = \underline{\mathcal{P}}(\Xi_S)$.

Proof of Observation 1. We prove both statements for the optimality problem in step 3 of our primal–dual lifting scheme; similar arguments apply to the feasibility problem in step 2.

As for the first statement, let $(\boldsymbol{x}, \boldsymbol{y}_A, \boldsymbol{y}_S)$ be an optimal solution to the conservative approximation $\overline{\mathcal{P}}(\Xi_S)$ in step 3(i), and assume that $\boldsymbol{y}_A(\xi) = \boldsymbol{Y}_A\xi + \boldsymbol{y}_A$ for all $\xi \in \Xi_A$. Fix a nonbinding scenario $\xi^\star \in \operatorname{ext}\Xi_A \setminus \Xi_A^\star$ and assume to the contrary that $\overline{\mathcal{P}}(\Xi_S') < \overline{\mathcal{P}}(\Xi_S)$ for the bounding problem $\overline{\mathcal{P}}(\Xi_S')$, $\Xi_S' = \Xi_S \cup \{\xi^\star\}$ that results from lifting the scenario $\xi^\star$. Let $(\boldsymbol{x}', \boldsymbol{y}_A', \boldsymbol{y}_S')$ be an optimal solution to the lifted problem $\overline{\mathcal{P}}(\Xi_S')$ and assume that $\boldsymbol{y}_A'(\xi) = \boldsymbol{Y}_A'\xi + \boldsymbol{y}_A'$ for all $\xi \in \Xi_A'$. Consider the convex combinations $\hat{\boldsymbol{x}} = \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{x}'$, $\hat{\boldsymbol{y}}_A(\xi) = [\lambda\boldsymbol{Y}_A + (1-\lambda)\boldsymbol{Y}_A']\xi + [\lambda\boldsymbol{y}_A + (1-\lambda)\boldsymbol{y}_A']$, $\xi \in \Xi_A$, and $\hat{\boldsymbol{y}}_S(\xi) = \lambda\boldsymbol{y}_S(\xi) + (1-\lambda)\boldsymbol{y}_S'(\xi)$, $\xi \in \Xi_S$. We show that, for $\lambda \uparrow 1$, $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}_A, \hat{\boldsymbol{y}}_S)$ is feasible in the conservative approximation $\overline{\mathcal{P}}(\Xi_S)$. Because $\boldsymbol{q}^\top\hat{\boldsymbol{x}} < \boldsymbol{q}^\top\boldsymbol{x}$, this would contradict the optimality of $(\boldsymbol{x}, \boldsymbol{y}_A, \boldsymbol{y}_S)$ in $\overline{\mathcal{P}}(\Xi_S)$.

We first note that, for all $\xi \in \Xi_S$, we have

$$\begin{aligned}
&\boldsymbol{T}(\xi)\hat{\boldsymbol{x}} + \boldsymbol{W}\hat{\boldsymbol{y}}_S(\xi) \geq \boldsymbol{h}(\xi)\\
&\quad\iff \boldsymbol{T}(\xi)[\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{x}'] + \boldsymbol{W}\big[\lambda\boldsymbol{y}_S(\xi) + (1-\lambda)\boldsymbol{y}_S'(\xi)\big]\\
&\qquad\qquad \geq \boldsymbol{h}(\xi)\\
&\quad\iff \lambda\big[\boldsymbol{T}(\xi)\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}_S(\xi)\big] + (1-\lambda)\big[\boldsymbol{T}(\xi)\boldsymbol{x}' + \boldsymbol{W}\boldsymbol{y}_S'(\xi)\big]\\
&\qquad\qquad \geq \boldsymbol{h}(\xi),
\end{aligned}$$

and the last inequality holds because $\boldsymbol{T}(\xi)\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}_S(\xi) \geq \boldsymbol{h}(\xi)$ and $\boldsymbol{T}(\xi)\boldsymbol{x}' + \boldsymbol{W}\boldsymbol{y}_S'(\xi) \geq \boldsymbol{h}(\xi)$ for all $\xi \in \Xi_S$ by construction. Similarly, we observe that for all $\xi \in \Xi_A$, we have

$$\begin{aligned}
&\boldsymbol{T}(\xi)\hat{\boldsymbol{x}} + \boldsymbol{W}\hat{\boldsymbol{y}}_A(\xi) \geq \boldsymbol{h}(\xi)\\
&\quad\iff \boldsymbol{T}(\xi)[\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{x}'] + \boldsymbol{W}\big[\lambda(\boldsymbol{Y}_A\xi + \boldsymbol{y}_A)\\
&\qquad\qquad + (1-\lambda)(\boldsymbol{Y}_A'\xi + \boldsymbol{y}_A')\big] \geq \boldsymbol{h}(\xi)\\
&\quad\iff \lambda\big[\boldsymbol{T}(\xi)\boldsymbol{x} + \boldsymbol{W}(\boldsymbol{Y}_A\xi + \boldsymbol{y}_A)\big] + (1-\lambda)[\boldsymbol{T}(\xi)\boldsymbol{x}'\\
&\qquad\qquad + \boldsymbol{W}(\boldsymbol{Y}_A'\xi + \boldsymbol{y}_A')\big] \geq \boldsymbol{h}(\xi).
\end{aligned}$$

The last inequality holds for all $\xi \in \operatorname{ext}\Xi_A'$ because $\boldsymbol{T}(\xi)\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}_A(\xi) \geq \boldsymbol{h}(\xi)$ and $\boldsymbol{T}(\xi)\boldsymbol{x}' + \boldsymbol{W}\boldsymbol{y}_A'(\xi) \geq \boldsymbol{h}(\xi)$ for all $\xi \in \operatorname{ext}\Xi_A'$ by construction. We furthermore observe that $\boldsymbol{T}(\xi^\star)\hat{\boldsymbol{x}} + \boldsymbol{W}\hat{\boldsymbol{y}}_A(\xi^\star) \geq \boldsymbol{h}(\xi^\star)$ for all λ sufficiently close to one because $\boldsymbol{T}(\xi^\star)\boldsymbol{x} + \boldsymbol{W}(\boldsymbol{Y}_A\xi^\star + \boldsymbol{y}_A) > \boldsymbol{h}(\xi^\star)$ by assumption. We, thus, conclude that $\boldsymbol{T}(\xi)\hat{\boldsymbol{x}} + \boldsymbol{W}\hat{\boldsymbol{y}}_A(\xi) \geq \boldsymbol{h}(\xi)$ for all $\xi \in \operatorname{ext}\Xi_A$ as long as λ is sufficiently close to one. The linearity of $\boldsymbol{T}$, $\hat{\boldsymbol{y}}_A$, and $\boldsymbol{h}$ then implies that $\boldsymbol{T}(\xi)\hat{\boldsymbol{x}} + \boldsymbol{W}\hat{\boldsymbol{y}}_A(\xi) \geq \boldsymbol{h}(\xi)$ for all $\xi \in \Xi_A$ as desired.

In view of the second statement, we note that, if $\Xi_A^\star = \emptyset$, then we can remove from the conservative approximation $\overline{\mathcal{P}}(\Xi_S)$ all constraints involving the realizations $\xi \in \Xi_A$ without changing the optimal value of the problem. The dual of this reduced problem can be interpreted as an instance of the progressive approximation $\underline{\mathcal{P}}(\Xi_S)$, where $\boldsymbol{\lambda}_A(\xi) = \mathbf{0}$ for all $\xi \in \Xi_A$. Strong linear programming duality holds because $\overline{\mathcal{P}}(\Xi_S)$ is feasible by construction. The statement then follows because the dual of the reduced problem is a restriction of the progressive approximation $\underline{\mathcal{P}}(\Xi_S)$ that attains the same optimal value as the conservative approximation $\overline{\mathcal{P}}(\Xi_S)$.

The first property of Observation 1 implies that the update steps 2(ii) and 3(ii) only need to consider the scenarios $\xi^\star$ of the binding scenario set $\Xi_A^\star$. The second property guarantees that the binding scenario set is never empty: if $\Xi_A^\star$ was empty in step 2(ii), then $\overline{\mathcal{P}}(\Xi_S) = \underline{\mathcal{P}}(\Xi_S)$ in step 2(i), which implies that the algorithm would have either proceeded with step 3 (if both bounds are zero) or terminated (if both bounds are positive). A similar reasoning applies if $\Xi_A^\star$ was empty in step 3(ii). We note that the crucial role of binding scenarios has previously been recognized in the context of uncertainty set partitioning approaches, see Song and Luedtke (2015), Bertsimas and Dunning (2016), and Postek and den Hertog (2016).

Observation 1 does not specify which scenario $\xi^\star \in \Xi_A^\star$ to lift. A natural approach to determine candidate scenarios $\xi^\star \in \Xi_A^\star$ to lift is to fix the optimal solution to the current bounding problem $\overline{\mathcal{P}}(\Xi_S)$ and determine one binding scenario for each constraint by minimizing the constraint's slack over all $\xi \in \Xi_A$. This requires the solution of a linear program for each

constraint. Any of the binding scenarios thus identified is a candidate for the lifting in steps 2(ii) and 3(ii) of the algorithm. We describe a more elaborate selection heuristic in our numerical example in Section 5.2.

The updates of the affine set Ξ_A in steps 2(ii) and 3(ii) guarantee that $\Xi_A = \operatorname{conv}([\operatorname{ext}\Xi] \setminus \Xi_S)$ throughout the algorithm. The update steps are intimately related to the forbidden vertices problem (Angulo et al. 2015), which optimizes a linear function over all but a few designated vertices of a polyhedron. The update steps can be implemented without enumerating all vertices of Ξ_A. To this end, we determine the neighboring extreme points of $\boldsymbol{\xi}^\star$ in Ξ_A, $N(\boldsymbol{\xi}^\star)$, for example, through simplex pivoting steps (Gal 1993). We then determine the half spaces defining $\operatorname{conv} N(\boldsymbol{\xi}^\star)$ via facet enumeration (Seidel 2004), and we add those half spaces of $\operatorname{conv} N(\boldsymbol{\xi}^\star)$ to the description of Ξ_A that do not contain the extreme point $\boldsymbol{\xi}^\star$ to be removed. Figure 2 illustrates our update procedure.

We are now ready to prove the correctness of our iterative solution scheme.

Theorem 2 (Finite Convergence). *The algorithm terminates in finite time, and it either determines an optimal solution to $\mathcal{P}$ or it correctly identifies the infeasibility or unboundedness of the problem.*

Proof of Theorem 2. We first show that the algorithm terminates after at most $|\operatorname{ext}\Xi|$ executions of the solution steps 2(i) and 3(i). Indeed, assume to the contrary that the algorithm would execute steps 2(i) and 3(i) more than $|\operatorname{ext}\Xi|$ times. During each execution of the update steps 2(ii) and 3(ii), one extreme point of the affine set Ξ_A is transferred to the scenario set Ξ_S. Thus, after $|\operatorname{ext}\Xi|$ executions of the update steps, the bounding problems in the solution steps 2(i) and 3(i) become $\underline{\mathcal{P}}(\operatorname{ext}\Xi)$ and $\overline{\mathcal{P}}(\operatorname{ext}\Xi)$. Because the feasibility and optimality bounding problems are feasible by construction, we must have $\underline{\mathcal{P}}(\operatorname{ext}\Xi) = \overline{\mathcal{P}}(\operatorname{ext}\Xi)$. In that case, however, either $\underline{\mathcal{P}}(\operatorname{ext}\Xi) > 0$ in step 2(i) or one of the two termination criteria in step 3(i) is met.

If the algorithm terminates because $\underline{\mathcal{P}}(\Xi_S) > 0$ in step 2(i), then there is no solution $(\boldsymbol{x}, \boldsymbol{y}, v)$ to the progressive bound $\underline{\mathcal{P}}(\Xi_S)$ such that $\boldsymbol{x}$ and $\boldsymbol{y}$ satisfy $\boldsymbol{T}(\boldsymbol{\xi})\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}(\boldsymbol{\xi}) \geq \boldsymbol{h}(\boldsymbol{\xi})$ for all $\boldsymbol{\xi} \in \Xi$. Proposition 3 then implies that there is no solution $(\boldsymbol{x}, \boldsymbol{y})$ to problem $\mathcal{P}$ that satisfies $\boldsymbol{T}(\boldsymbol{\xi})\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}(\boldsymbol{\xi}) \geq \boldsymbol{h}(\boldsymbol{\xi})$ for all $\boldsymbol{\xi} \in \Xi$; that is, problem $\mathcal{P}$ is indeed infeasible.

Assume now that the algorithm terminates because $\overline{\mathcal{P}}(\Xi_S)$ unbounded in step 3(i). By Proposition 2, any feasible solution to $\overline{\mathcal{P}}(\Xi_S)$ can be transformed into a feasible solution to $\mathcal{P}$ that achieves the same objective value. We, thus, conclude that problem $\mathcal{P}$ is unbounded as well.

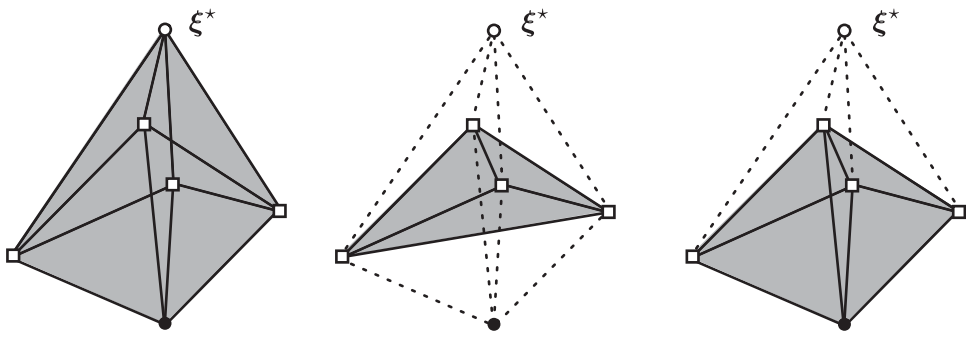
Finally, assume that the algorithm terminates because $\underline{\mathcal{P}}(\Xi_S) = \overline{\mathcal{P}}(\Xi_S)$ in step 3(i). Because the bounding problems are feasible by construction, Propositions 2 and 3 imply that

$$\overline{\mathcal{P}}(\Xi_S) \geq \overline{\mathcal{P}}(\operatorname{ext}\Xi) = \mathcal{P} = \underline{\mathcal{P}}(\operatorname{ext}\Xi) \geq \underline{\mathcal{P}}(\Xi_S),$$

where we denote by $\mathcal{P}$ the optimal value of problem $\mathcal{P}$. Because $\underline{\mathcal{P}}(\Xi_S) = \overline{\mathcal{P}}(\Xi_S)$, however, we know that the optimal values of all these problems coincide. Proposition 2 then allows us to transform any optimal solution to $\overline{\mathcal{P}}(\Xi_S)$ into an optimal solution to problem $\mathcal{P}$. □

The runtime of the algorithm is determined by (i) the number of iterations (i.e., the number of times that steps 2 and 3 are executed), (ii) the size of the bounding problems $\overline{\mathcal{P}}(\Xi_S)$ and $\underline{\mathcal{P}}(\Xi_S)$ in each iteration, (iii) the complexity of selecting the scenario $\boldsymbol{\xi}^\star \in \Xi_A^\star$ to be removed, (iv) the complexity of the update $\Xi_A \leftarrow \operatorname{conv}([\operatorname{ext}\Xi_A] \setminus \{\boldsymbol{\xi}^\star\})$, and (v) the computation of the moment matrices $\boldsymbol{\Sigma}_A$. The algorithm performs up to $|\operatorname{ext}\Xi|$ iterations in the worst case, and an upper bound for this number is provided in Section 2. The complexity of the bounding problems as well as the selection of $\boldsymbol{\xi}^\star \in \Xi_A^\star$ primarily depend on the description of the intermediate sets Ξ_A, which has been discussed

Figure 2. Illustration of the Update Step in $\mathbb{R}^3$



Notes. On the left, the circular vertex $\boldsymbol{\xi}^\star$ to be removed from the distorted diamond has the four square vertices as neighbors. The middle polyhedron illustrates $\operatorname{conv} N(\boldsymbol{\xi}^\star)$, the convex hull of these neighbors. Out of the four half spaces defining this convex hull, the upper two do not contain $\boldsymbol{\xi}^\star$ and are, thus, added to the updated description of Ξ_A on the right.

in Section 2.1. For our lifting method, the time required for the update $\Xi_A \leftarrow \operatorname{conv}([\operatorname{ext}\Xi_A] \setminus \{\xi^\star\})$ is determined by

(i) The size of $N(\xi^\star)$: this quantity depends on the degeneracy of $\xi^\star$ in Ξ_A. If $k+\sigma$ constraints of Ξ_A are binding at $\xi^\star$, then $\xi^\star$ has at most

$$\binom{k+\sigma}{k-1}(\ell-k-\sigma)$$

neighbors, where ℓ denotes the number of constraints that describe Ξ_A. In the worst case, each vertex of Ξ_A is a neighbor of $\xi^\star$, which is the case if Ξ_A is two-neighborly. The primal upper bound theorem implies that $\operatorname{conv} N(\xi^\star)$ can consist of up to

$$\binom{|N(\xi^\star)| - \lceil k/2 \rceil}{\lfloor k/2 \rfloor} + \binom{|N(\xi^\star)| - 1 - \lceil (k-1)/2 \rceil}{\lfloor (k-1)/2 \rfloor}$$

half spaces, and this bound is attained by primal cyclic polytopes (Seidel 2004).

(ii) The complexity of computing the convex hull $\operatorname{conv} N(\xi^\star)$: this convex hull can be computed in time $\mathcal{O}(|N(\xi^\star)|^{\lfloor k/2 \rfloor})$.

It can be shown that, even when a single scenario $\xi^\star$ is lifted, our description of $\Xi_A = \operatorname{conv}([\operatorname{ext}\Xi_A] \setminus \{\xi^\star\}$ can grow exponentially (Angulo et al. 2015, proposition 5). Although there are *lifted* formulations of Ξ_A that only grow *quadratically* if a single scenario is lifted (Angulo et al. 2015, proposition 6), it is known that optimizing a linear function over Ξ_A after a flexible (i.e., not a priori fixed) number of scenarios has been lifted is NP-hard (Angulo et al. 2015, theorem 11) for generic polyhedral uncertainty sets. On the other hand, for special classes of uncertainty sets, such as uncertainty sets with zero–one vertices, the same problem is solvable in polynomial time if a suitable lifted formulation of Ξ_A is used (Angulo et al. 2015, section 3). The initial moment matrix $\mathbf{\Sigma}_A$, finally, can often be computed efficiently—see the online appendix—and updating $\mathbf{\Sigma}_A$ only requires subtracting the (outer product of the) scenario $\xi^\star \in \Xi_A^\star$ that is removed in each iteration.

We close this section by outlining possible extensions of our iterative solution scheme.

Remark 4 (Discrete Here-and-Now Decisions). Theorem 2 does not exploit any specific properties of the here-and-now decisions x, and our primal–dual lifting scheme, therefore, immediately carries over to the mixed-integer, two-stage robust optimization problem $\mathcal{P}_\mathrm{d}$. Observation 1, on the other hand, crucially relies on the convexity in x, and one can construct counterexamples in which lifting a nonbinding scenario results in an improved bound. Thus, in the presence of discrete here-and-now decisions, we can no longer restrict ourselves to lifting scenarios $\xi^\star \in \Xi_A^\star$.

Remark 5 (Other Extensions). The algorithm can be altered in several ways. For example, the progressive and conservative bounding problems in steps 2(i) and 3(i) could involve multiple affine sets $\Xi_{A,1}, \ldots, \Xi_{A,s}$, or they could assign piecewise affine decision rules to the realizations $\xi \in \Xi_A$. If convergence to an optimal solution is not required, then the affine set Ξ_A could be replaced with outer approximations of $\operatorname{conv}([\operatorname{ext}\Xi] \setminus \Xi_S)$ without affecting the validity of the bounds. This is advantageous if these outer approximations have compact descriptions, as is the case for Löwner–John ellipsoids (Henk 2012), for example. One could also transfer multiple extreme points to Ξ_S in steps 2(ii) and 3(ii), and it might be advantageous to transfer extreme points back to Ξ_A. Finally, one could envision lifting different extreme points for the primal and the dual bounds. □

4. Relation to Piecewise Affine Decision Rules

Although a feasible solution (x, y_A, y_S) to the conservative approximation $\overline{\mathcal{P}}(\Xi_S)$ provides an implementable first-stage decision, it only provides implementable recourse decisions for the parameter realizations $\xi \in \Xi_A \cup \Xi_S$. This is of no concern for most applications, in which only the first-stage decision is implemented. In some situations, however, an implementable recourse decision $y(\xi)$ for the two-stage robust optimization problem $\mathcal{P}$ might be required here and now for every $\xi \in \Xi$. This is frequently the case in real-time control applications, in which there is not enough time to solve optimization problems to determine the recourse actions, as well as in embedded systems that lack the processing power or energy supply to solve optimization problems. In this section, we, therefore, elaborate how the second-stage decision (y_A, y_S) of a feasible solution (x, y_A, y_S) to $\overline{\mathcal{P}}(\Xi_S)$ can be transformed into a decision rule $y : \Xi \mapsto \mathbb{R}^{n_2}$ that prescribes implementable recourse decisions to $\mathcal{P}$ for all $\xi \in \Xi$. Along the way, we discover some insightful connections between our lifting scheme and the piecewise affine decision rules studied in Chen et al. (2008), Chen and Zhang (2009), Goh and Sim (2010), Georghiou et al. (2015), and others.

We define the *graph* of a decision rule $y : \Xi \mapsto \mathbb{R}^{n_2}$ as

$$\operatorname{gr} y \;=\; \left\{ \begin{pmatrix} \xi \\ y(\xi) \end{pmatrix} : \xi \in \Xi \right\}.$$

By construction, $\operatorname{gr} y$ is uniquely specified through y and vice versa. We also define the *recourse set* of a second-stage decision (y_A, y_S) in problem $\overline{\mathcal{P}}(\Xi_S)$ as

$$\operatorname{rec}(y_A, y_S) = \operatorname{conv}\left(\left\{ \begin{pmatrix} \xi \\ y_A(\xi) \end{pmatrix} : \xi \in \Xi_A \right\} \cup \left\{ \begin{pmatrix} \xi \\ y_S(\xi) \end{pmatrix} : \xi \in \Xi_S \right\} \right).$$

To economize on notation, we omit the dependence of $\operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S)$ on Ξ_A and Ξ_S. Intuitively speaking, $\operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S)$ constitutes the convex hull of the graphs $\operatorname{gr} \boldsymbol{y}_A$ and $\operatorname{gr} \boldsymbol{y}_S$ restricted to their respective domains Ξ_A and Ξ_S. We now show that $\operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S)$ contains those recourse decisions $\boldsymbol{y}$ to the two-stage problem $\mathcal{P}$ that corresponds to the second-stage decision $(\boldsymbol{y}_A, \boldsymbol{y}_S)$ in $\overline{\mathcal{P}}(\Xi_S)$.

Proposition 4. *If $(\boldsymbol{x}, \boldsymbol{y}_A, \boldsymbol{y}_S)$ is feasible in $\overline{\mathcal{P}}(\Xi_S)$, then, for any decision rule $\boldsymbol{y} : \Xi \mapsto \mathbb{R}^{n_2}$ satisfying $\operatorname{gr} \boldsymbol{y} \subseteq \operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S)$, $(\boldsymbol{x}, \boldsymbol{y})$ is feasible in $\mathcal{P}$ and attains the same objective value as $(\boldsymbol{x}, \boldsymbol{y}_A, \boldsymbol{y}_S)$ in $\overline{\mathcal{P}}(\Xi_S)$.*

Proof of Proposition 4. It is clear that $(\boldsymbol{x}, \boldsymbol{y})$ attains the same objective value in $\mathcal{P}$ as $(\boldsymbol{x}, \boldsymbol{y}_A, \boldsymbol{y}_S)$ does in $\overline{\mathcal{P}}(\Xi_S)$. We now show that $(\boldsymbol{x}, \boldsymbol{y})$ is feasible in $\mathcal{P}$. One readily verifies that

$$\operatorname{ext} \operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S) = \left\{ \begin{pmatrix} \xi' \\ \boldsymbol{y}_A(\xi') \end{pmatrix} : \xi' \in \operatorname{ext} \Xi_A \right\} \cup \left\{ \begin{pmatrix} \xi' \\ \boldsymbol{y}_S(\xi') \end{pmatrix} : \xi' \in \Xi_S \right\},$$

and because $\operatorname{gr} \boldsymbol{y} \subseteq \operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S)$, by construction, for every $\xi \in \Xi$ there is a function $\lambda : \operatorname{ext} \Xi \mapsto \mathbb{R}_+$ that satisfies $\sum_{\xi' \in \operatorname{ext} \Xi} \lambda(\xi') = 1$ and

$$\begin{pmatrix} \xi \\ \boldsymbol{y}(\xi) \end{pmatrix} = \sum_{\xi' \in \operatorname{ext} \Xi_A} \lambda(\xi') \begin{pmatrix} \xi' \\ \boldsymbol{y}_A(\xi') \end{pmatrix} + \sum_{\xi' \in \Xi_S} \lambda(\xi') \begin{pmatrix} \xi' \\ \boldsymbol{y}_S(\xi') \end{pmatrix}. \tag{2}$$

We, thus, obtain that

$$\begin{aligned}
&\boldsymbol{T}(\xi)\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}(\xi) \\
&= \boldsymbol{T}\left(\sum_{\xi' \in \operatorname{ext} \Xi_A} \lambda(\xi')\, \xi' + \sum_{\xi' \in \Xi_S} \lambda(\xi')\, \xi' \right) \boldsymbol{x} \\
&\quad + \boldsymbol{W}\boldsymbol{y}\left(\sum_{\xi' \in \operatorname{ext} \Xi_A} \lambda(\xi')\, \xi' + \sum_{\xi' \in \Xi_S} \lambda(\xi')\, \xi' \right) \\
&= \sum_{\xi' \in \operatorname{ext} \Xi_A} \lambda(\xi')\, \boldsymbol{T}(\xi')\boldsymbol{x} + \sum_{\xi' \in \Xi_S} \lambda(\xi')\, \boldsymbol{T}(\xi')\boldsymbol{x} \\
&\quad + \boldsymbol{W}\left(\sum_{\xi' \in \operatorname{ext} \Xi_A} \lambda(\xi')\, \boldsymbol{y}_A(\xi') + \sum_{\xi' \in \Xi_S} \lambda(\xi')\, \boldsymbol{y}_S(\xi') \right) \\
&= \sum_{\xi' \in \operatorname{ext} \Xi_A} \lambda(\xi') \left[\boldsymbol{T}(\xi')\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}_A(\xi') \right] \\
&\quad + \sum_{\xi' \in \Xi_S} \lambda(\xi') \left[\boldsymbol{T}(\xi')\boldsymbol{x} + \boldsymbol{W}\boldsymbol{y}_S(\xi') \right] \\
&\geq \sum_{\xi' \in \operatorname{ext} \Xi_A} \lambda(\xi')\, \boldsymbol{h}(\xi') + \sum_{\xi' \in \Xi_S} \lambda(\xi')\, \boldsymbol{h}(\xi') \\
&= \boldsymbol{h}\left(\sum_{\xi' \in \operatorname{ext} \Xi_A} \lambda(\xi')\, \xi' + \sum_{\xi' \in \Xi_S} \lambda(\xi')\, \xi' \right) = \boldsymbol{h}(\xi),
\end{aligned}$$

where the first, second, and last identity are due to (2) and the fact that $\boldsymbol{T}$ and $\boldsymbol{h}$ are affine, and the inequality holds because of the feasibility of $(\boldsymbol{x}, \boldsymbol{y}_A, \boldsymbol{y}_S)$ in $\overline{\mathcal{P}}(\Xi_S)$. □

Figure 3 illustrates the recourse set of a one-dimensional recourse decision. Interestingly, the reverse implication of Proposition 4 does not hold in general. In fact, the next example shows that there can be feasible solutions $(\boldsymbol{x}, \boldsymbol{y})$ to problem $\mathcal{P}$ for which $\operatorname{gr} \boldsymbol{y} \not\subseteq \operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S)$ for every feasible solution $(\boldsymbol{x}, \boldsymbol{y}_A, \boldsymbol{y}_S)$ to every conservative approximation $\overline{\mathcal{P}}(\Xi_S)$ of $\mathcal{P}$.

Example 2. Consider the two-stage robust optimization problem

$$\begin{array}{ll} \text{minimize} & x \\ \text{subject to} & x \geq y(\xi) \\ & y(\xi) \geq \xi, \;\; y(\xi) \geq -\xi \end{array} \Bigg\} \;\; \forall \xi \in [-1, 1]. \tag{3}$$

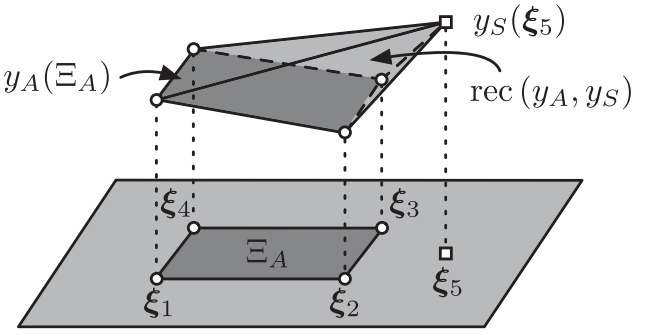
For every scenario set $\Xi_S \subseteq \{-1, 1\}$, the unique optimal solution (x, y_A, y_S) to $\overline{\mathcal{P}}(\Xi_S)$ satisfies $x = 1$ and $\operatorname{rec}(y_A, y_S) = [-1, 1] \times \{1\}$. Problem (3) is optimized, however, by $x = 1$ and every decision rule $y : [-1, 1] \mapsto \mathbb{R}_+$ satisfying $y(\xi) \in [|\xi|, 1]$.

The question naturally arises which decision rule from within $\operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S)$ we should select. A subclass of decision rules that are of special interest are *piecewise affine decision rules* $\boldsymbol{y} : \Xi \mapsto \mathbb{R}^{n_2}$ for which there is a partition of Ξ into finitely many polyhedra $\Xi_1, \ldots, \Xi_L$ such that $\boldsymbol{y}$ is affine on each Ξ_ℓ, $\ell = 1, \ldots, L$. It turns out that piecewise affine decision rules are intimately related to *simplicial decompositions* of Ξ, which are subdivisions of Ξ into finitely many k-dimensional simplices that only intersect at their boundaries and whose union recovers Ξ.

Proposition 5. *Consider a scenario set $\Xi_S \subseteq \operatorname{ext} \Xi$ as well as the second-stage decisions $\boldsymbol{y}_A : \Xi_A \mapsto \mathbb{R}^{n_2}$ and $\boldsymbol{y}_S : \Xi_S \mapsto \mathbb{R}^{n_2}$. Then*

i. *For every piecewise affine decision rule $\boldsymbol{y} : \Xi \mapsto \mathbb{R}^{n_2}$ with $\operatorname{gr} \boldsymbol{y} \in \operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S)$, there is a simplicial decomposition $\{\Xi_\ell\}_\ell$ of Ξ such that $\boldsymbol{y}$ is affine on every simplex Ξ_ℓ.*

Figure 3. For $\Xi_A = \operatorname{conv}\{\xi_1, \ldots, \xi_4\}$ and $\Xi_S = \{\xi_5\}$ in the Shaded Plane, the Recourse Set $\operatorname{rec}(y_A, y_S)$ Is Shown as the Light-Shaded Pyramid



Note. Its darker shaded bottom represents $y_A(\Xi_A) = \{y_A(\xi) : \xi \in \Xi_A\}$.

ii. *For every simplicial decomposition* $\{\Xi_\ell\}_\ell$ *of* Ξ, *there is a decision rule* $\boldsymbol{y} : \Xi \mapsto \mathbb{R}^{n_2}$ *with* $\operatorname{gr} \boldsymbol{y} \in \operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S)$ *such that* $\boldsymbol{y}$ *is affine on every simplex* Ξ_ℓ.

Proof of Proposition 5. As for the first statement, there is a partition $\{\hat{\Xi}_\ell\}_\ell$ of Ξ into polyhedra such that $\boldsymbol{y}$ is affine over each $\hat{\Xi}_\ell$. Fix a simplicial decomposition $\{\Xi_{\ell,\ell'}\}_{\ell'}$ for each polyhedron $\hat{\Xi}_\ell$. Then $\boldsymbol{y}$ is also affine over the simplicial decomposition $\{\Xi_{\ell,\ell'}\}_{\ell,\ell'}$ of Ξ.

In view of the second statement, fix any simplicial decomposition $\{\Xi_\ell\}_\ell$ of Ξ as well as

$$\boldsymbol{y}(\boldsymbol{\xi}') \in \left\{ \boldsymbol{y} \in \mathbb{R}^{n_2} \,:\, \begin{pmatrix} \boldsymbol{\xi}' \\ \boldsymbol{y} \end{pmatrix} \in \operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S) \right\} \quad \text{for} \quad \boldsymbol{\xi}' \in \bigcup_\ell \operatorname{ext} \Xi_\ell, \tag{4}$$

where $\bigcup_\ell \operatorname{ext} \Xi_\ell$ constitutes the set of all corner points of the simplices Ξ_ℓ in the simplicial decomposition. For each simplex Ξ_ℓ, we set $\boldsymbol{y}(\boldsymbol{\xi}) = \sum_{\boldsymbol{\xi}' \in \operatorname{ext} \Xi_\ell} \lambda(\boldsymbol{\xi}; \boldsymbol{\xi}') \cdot \boldsymbol{y}(\boldsymbol{\xi}')$ for all remaining points $\boldsymbol{\xi} \in \Xi_\ell \setminus (\operatorname{ext} \Xi_\ell)$, where $\lambda : \Xi_\ell \times \operatorname{ext} \Xi_\ell \mapsto \mathbb{R}_+$ is the unique weighting function that satisfies

$$\sum_{\boldsymbol{\xi}' \in \operatorname{ext} \Xi_\ell} \lambda(\boldsymbol{\xi}; \boldsymbol{\xi}') = 1 \qquad \text{and}$$
$$\sum_{\boldsymbol{\xi}' \in \operatorname{ext} \Xi_\ell} \lambda(\boldsymbol{\xi}; \boldsymbol{\xi}') \cdot \boldsymbol{\xi}' = \boldsymbol{\xi} \qquad \forall \boldsymbol{\xi} \in \Xi_\ell;$$

that is, $\lambda(\boldsymbol{\xi}; \cdot)$ are the barycentric coordinates of $\boldsymbol{\xi}$ in Ξ_ℓ. The statement now follows if

$$\begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{y}(\boldsymbol{\xi}) \end{pmatrix} \in \operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S) \qquad \forall \boldsymbol{\xi} \in \Xi_\ell.$$

Because of (4), this holds for $\boldsymbol{\xi} \in \operatorname{ext} \Xi_\ell$. Moreover, for $\boldsymbol{\xi} \in \Xi_\ell \setminus (\operatorname{ext} \Xi_\ell)$, we have

$$\begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{y}(\boldsymbol{\xi}) \end{pmatrix} = \sum_{\boldsymbol{\xi}' \in \operatorname{ext} \Xi_\ell} \lambda(\boldsymbol{\xi}; \boldsymbol{\xi}') \cdot \begin{pmatrix} \boldsymbol{\xi}' \\ \boldsymbol{y}(\boldsymbol{\xi}') \end{pmatrix} \in \operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S),$$

where the membership follows from (4) and the convexity of $\operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S)$. Because the simplex Ξ_ℓ was chosen arbitrarily, we, thus, conclude that $\operatorname{gr} \boldsymbol{y} \subseteq \operatorname{rec}(\boldsymbol{y}_A, \boldsymbol{y}_S)$.

Propositions 4 and 5 allow us to complete any feasible solution $(\boldsymbol{x}, \boldsymbol{y}_A, \boldsymbol{y}_S)$ in problem $\overline{\mathcal{P}}(\Xi_S)$ to a feasible solution $(\boldsymbol{x}, \boldsymbol{y})$ in problem $\mathcal{P}$ that attains the same objective value and that is piecewise affine over any fixed simplicial decomposition $\{\Xi_\ell\}_\ell$ of Ξ. To construct a piecewise affine decision rule with a compact description, we propose to combine the affine decision rule $\boldsymbol{y}_A$ over Ξ_A with a piecewise affine decision rule over $\Xi \setminus \Xi_A$ that is affine on every simplex of a simplicial decomposition of $\operatorname{cl}(\Xi \setminus \Xi_A)$. Note that $\operatorname{cl}(\Xi \setminus \Xi_A) = \operatorname{conv}(\Xi_S \cup \bigcup\{N(\boldsymbol{\xi}) : \boldsymbol{\xi} \in \Xi_S\})$, where $N(\boldsymbol{\xi})$ denotes the neighboring extreme points of $\boldsymbol{\xi}$. A simplicial decomposition of $\operatorname{cl}(\Xi \setminus \Xi_A)$ can be found with standard triangulation schemes (Büeler et al. 2000).

5. Numerical Experiments

We now analyze the computational performance of our primal–dual lifting scheme from Section 3 in the context of two illustrative examples (Section 5.1) as well as an inventory-management problem (Section 5.2). In our experiments, we assess the scalability of our algorithm in terms of the problem size (measured by the number of decision variables and constraints) and the number of uncertain problem parameters. We also investigate to which degree the different components of our algorithm add to the algorithm's runtime.

A secondary objective of this section is to compare our algorithm with some of the state-of-the-art solution approaches for two-stage robust optimization problems. To this end, we compare our solution scheme with the conservative bounds offered by linear and piecewise linear decision rules (Ben-Tal et al. 2004, Kuhn et al. 2011, Georghiou et al. 2015), the progressive bounds provided by Hadjiyiannis et al. (2011) and Bertsimas and de Ruiter (2016), the column- and constraint-generation scheme from Zeng and Zhao (2013) as well as the Fourier–Motzkin elimination approach proposed by Zhen et al. (2016). Table 1 compares the guarantees offered by these approaches as well as their underlying assumptions. Although the pitfalls of drawing conclusions from a small test set are well known, we hope to generate some insights

Table 1. Comparison of Our Method ("Primal–Dual") with the Fourier–Motzkin Elimination Approach Proposed by Zhen et al. (2016) ("FME & SB"), the Column- and Constraint-Generation Scheme from Zeng and Zhao (2013) ("Zeng & Zhao"), and the Piecewise Linear Decision Rules (Ben-Tal et al. 2004, Kuhn et al. 2011, Georghiou et al. 2015) Combined with the Progressive Bounds Provided by Hadjiyiannis et al. (2011) or Bertsimas and de Ruiter (2016) ("ADR & SB")

	Technique	Convergence	Recourse	Uncertainty sets	Subproblems
Primal–dual	Lifting of uncertainty set	Finite	Generic	Specific classes	Convex hulls & LPs
FME & SB	Elimination of second-stage decisions	Finite	Generic[a]	Generic polyhedra	Elimination & LPs
Zeng & Zhao	Column- and constraint-generation	Finite	Rel. compl.	Generic polyhedra	MILPs
ADR & SB	(Piecewise) affine decision rules	None	Generic[a]	Generic polyhedra	LPs

[a]With the exception of Bertsimas and de Ruiter (2016), which requires relatively complete recourse.

into the intricate relationship between the characteristics of problem $\mathcal{P}$ and the suitability of the different solution schemes.

All optimization problems in this section were solved in single-threaded mode with the Gurobi 7.5 optimization package (see http://www.gurobi.com) on a 2.9-GHz computer with 8 GB RAM. Our C++ implementation of the primal–dual lifting scheme uses the LRS package for vertex and facet enumeration (see http://cgm.cs.mcgill.ca/~avis/C/lrs.html).

5.1. Illustrative Examples

We adapt two examples from the literature to illustrate the structure of the piecewise affine decision rules of our lifting approach (Section 5.1.1), and we show how the affine decision rules $\boldsymbol{\lambda}_A$ in the dual problem $\underline{\mathcal{P}}(\Xi_S)$ can contribute to the tightness of the lower bounds (Section 5.1.2).

5.1.1. Worst-Case Value of a Sum-of-Max Function. We consider the following adaptation of the example (TOY2) from Gorissen and den Hertog (2013):

$$\begin{array}{lll} \text{minimize} & \tau & \\ \text{subject to} & \tau \geq \mathbf{e}^\top \boldsymbol{y}(\boldsymbol{\xi}) & \\ & y_1(\boldsymbol{\xi}) \geq \max\{x, x+\xi_1+\xi_2\} & \\ & y_2(\boldsymbol{\xi}) \geq \max\{x, x+\xi_1-\xi_2\} & \forall \boldsymbol{\xi} \in \Xi \\ & y_3(\boldsymbol{\xi}) \geq \max\{x, x-\xi_1+\xi_2\} & \\ & y_4(\boldsymbol{\xi}) \geq \max\{x, x-\xi_1-\xi_2\} & \\ & \tau, x \in \mathbb{R}_+, \quad \boldsymbol{y} : \Xi \mapsto \mathbb{R}^4. & \end{array}$$

In this problem, we set the uncertainty set to $\Xi = \{\boldsymbol{\xi} \in [-2,2]^2 : \|\boldsymbol{\xi}\|_1 \leq 3\}$. The problem is optimized by the here-and-now decisions $(\tau^\star, x^\star) = (4, 0)$.

The uncertainty set Ξ satisfies $|\operatorname{ext}\Xi| = 8$. If we lift one of the binding scenarios $\boldsymbol{\xi}^\star \in \Xi_A^\star$ randomly in each iteration, then our lifting scheme converges after five iterations. Figure 4 visualizes the piecewise affine decision rules $y_1(\boldsymbol{\xi})$ corresponding to the optimal primal solutions $(\boldsymbol{x}, \boldsymbol{y}_A, \boldsymbol{y}_S)$ in every iteration of our lifting scheme. The initial dual bound $\underline{\mathcal{P}}(\emptyset)$ has an optimal objective value of 3.33, and the dual bound attains the optimal objective value four of the two-stage robust optimization problem after lifting any single extreme point $\boldsymbol{\xi} \in \operatorname{ext}\Xi_A$.

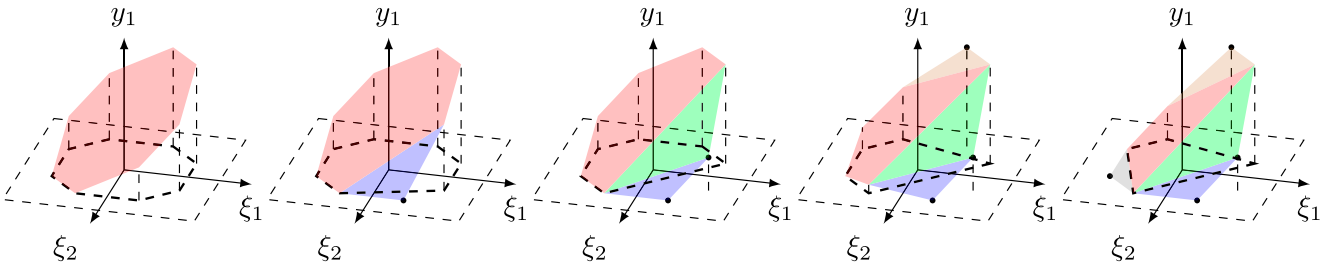
5.1.2. Worst-Case Makespan of a Temporal Network. We now employ the two-stage robust optimization problem $\mathcal{P}$ to estimate the worst-case makespan of a temporal network (e.g., a project, a digital circuit, or a production process). To this end, we define a temporal network as a directed, acyclic graph $G = (V, E)$ whose nodes $V = \{1, \ldots, n\}$ represent the tasks and whose arcs $E \subset V \times V$ denote the temporal precedences between the tasks. We assume that the duration $d_i(\boldsymbol{\xi})$ of each task $i \in V$ depends on the uncertain problem parameters $\boldsymbol{\xi} \in \Xi$. Moreover, we assume that the start times of the tasks $i \in V_0 \subseteq V$ have to be chosen here and now, that is, before the realizations of the uncertain parameters $\boldsymbol{\xi}$ are observed. This could be required to synchronize the network with other projects, circuits, or production processes.

The problem is a variant of model (1) in Wiesemann et al. (2012), and it can be formulated as

$$\begin{array}{lll} \text{minimize} & \tau & \\ \text{subject to} & \tau \geq y_n(\boldsymbol{\xi}) + d_n(\boldsymbol{\xi}) & \\ & y_j(\boldsymbol{\xi}) \geq y_i(\boldsymbol{\xi}) + d_i(\boldsymbol{\xi}) \quad \forall (i,j) \in E & \forall \boldsymbol{\xi} \in \Xi \\ & y_i(\boldsymbol{\xi}) = x_i \quad \forall i \in V_0 & \\ & \tau \in \mathbb{R}_+, \quad x_i \in \mathbb{R}_+, \, i \in V_0, \quad \boldsymbol{y} : \Xi \mapsto \mathbb{R}_+^n. & \end{array}$$

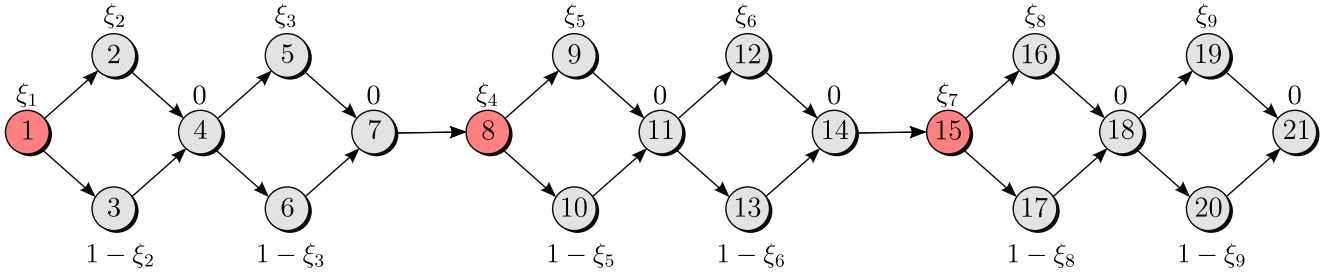
In this formulation, the wait-and-see decisions $y_i(\boldsymbol{\xi})$ capture the start times of the tasks $i \in V$, the epigraphical here-and-now decision τ records the worst-case makespan of the network, and the here-and-now decisions x_i correspond to the static start times of the restricted tasks $i \in V_0$. The presence of the restricted tasks $i \in V_0$ implies that the problem has no relatively complete recourse.

Figure 4. (Color online) Illustration of the Piecewise Affine Decision Rules $y_1(\xi)$ Corresponding to the Optimal Primal Solutions $(\boldsymbol{x}, \boldsymbol{y}_A, \boldsymbol{y}_S)$ in Every Iteration of Our Lifting Scheme



Note. From left to right, the optimal objective values of $\overline{\mathcal{P}}(\emptyset)$ are 6, 6, 5, 4.86, and 4.

Figure 5. (Color online) Temporal Network with 21 Tasks and 26 Precedences



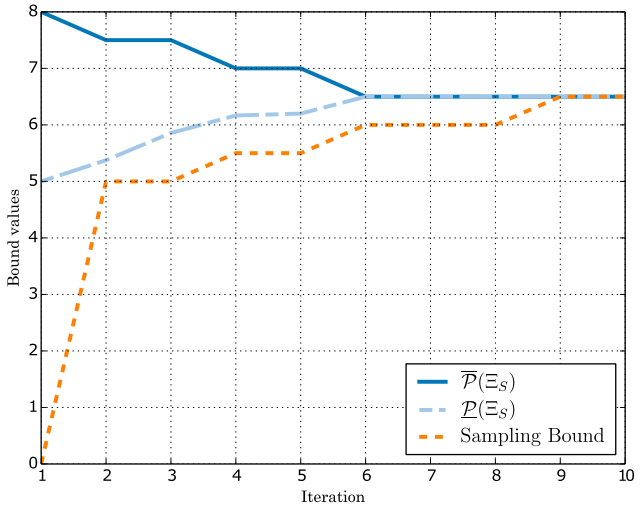
Notes. The durations $d_i(\xi)$ are displayed above the tasks $i \in V$, and the uncertainty set is $\Xi = \{\xi \in \mathbb{R}^9 : \|\xi - \mathbf{e}/2\|_1 \leq 1/2\}$. The start times of tasks 1, 8, and 15 have to be selected here and now.

We apply our primal–dual lifting scheme to the temporal network in Figure 5. We also present a variant of our dual bound $\underline{\mathcal{P}}(\Xi_S)$ in which we fix all affine decisions to $\boldsymbol{\lambda}_A(\xi) = \mathbf{0}$ for all $\xi \in \Xi_A$. This variant, which we refer to as "sampling bound," can be interpreted as an extension of the sampling bound in Hadjiyiannis et al. (2011) that iteratively grows the scenario set Ξ_S (see Remark 3). The results of both methods are shown in Figure 6. Our approach converges to the optimal objective value of the two-stage robust optimization problem $\mathcal{P}$ after six out of $|\text{ext}\,\Xi| = 18$ possible iterations. The less flexible sampling bound requires nine iterations to converge in this example.

5.2. Case Study: Inventory Management

We consider an inventory-management problem with n products whose demands $D_i(\xi)$, $i = 1, \ldots, n$, are governed by the uncertain risk factors $\xi \in \mathbb{R}^k$. The demand for product i can be served through a standard order x_i (with unit cost c_x), which has to be placed before the demand is known, or through an express order y_i (with unit cost $c_y > c_x$), which can be submitted after the demand has been observed. Any excess inventory $h_i(\xi)$ and any backlogged demand $b_i(\xi)$ in the second period incurs unit costs of c_h and c_b, respectively, and the express orders across all products must not exceed B units. The objective is to determine an ordering policy that minimizes the worst-case sum of the ordering, inventory holding, and backlogging costs over all anticipated demand realizations.

Figure 6. (Color online) Results of Our Primal–Dual Lifting Scheme (Top Two Blue Lines) and the Sampling Bound (Bottom Orange Line) Applied to the Temporal Network in Figure 5



The problem can be formulated as the following instance of problem $\mathcal{P}$:

$$\begin{array}{ll} \text{minimize} & \tau \\ \text{subject to} & \left.\begin{array}{l} \tau \geq \mathbf{e}^\top \big[c_x \cdot \boldsymbol{x} + c_y \cdot \boldsymbol{y}(\xi) \\ \qquad + c_h \cdot \boldsymbol{h}(\xi) + c_b \cdot \boldsymbol{b}(\xi)\big] \\ \boldsymbol{h}(\xi) \geq \boldsymbol{x} + \boldsymbol{y}(\xi) - \boldsymbol{D}(\xi) \\ \boldsymbol{b}(\xi) \geq \boldsymbol{D}(\xi) - \boldsymbol{x} - \boldsymbol{y}(\xi) \\ \mathbf{e}^\top \boldsymbol{y}(\xi) \leq B \end{array}\right\} \forall \xi \in \Xi \qquad (5) \\ & \boldsymbol{x} \in \mathbb{R}^n_+, \quad \boldsymbol{y}, \boldsymbol{h}, \boldsymbol{b} : \Xi \mapsto \mathbb{R}^n_+. \end{array}$$

Here, the epigraphical variable τ records the worst-case costs over all demand realizations. Although the existence of this first-stage epigraphical variable implies that problem (5) lacks a relatively complete recourse, we emphasize that an equivalent min–max formulation of the problem (as employed by the column- and constraint-generation scheme) enjoys a relatively complete recourse. Note also that we model the second-stage decisions $\boldsymbol{y}$, $\boldsymbol{h}$, and $\boldsymbol{b}$ as functions of the (typically unobservable) risk factors as opposed to functions of the (eventually observable) product demands $\boldsymbol{D}(\xi)$. This simplification is justified because, by construction, no optimal solution to problem (5) takes different second-stage decisions for different realizations of the risk factors ξ that give rise to the same product demands $\boldsymbol{D}(\xi)$. Thus, we can always convert the optimal second-stage decisions in problem (5) to equivalent implementable decisions that only depend on the product demands $\boldsymbol{D}(\xi)$.

We assume that the n products are grouped into $\lfloor\sqrt{n}\rfloor$ different product categories such that each category contains between $\lfloor\sqrt{n}\rfloor$ and $\lceil\sqrt{n}\rceil + 1$ products. The product demands are governed by a factor model of the form

$$D_i(\boldsymbol{\xi}) = \boldsymbol{\phi}_i^\top \boldsymbol{\xi} + \varphi_i, \; i = 1, \ldots, n, \text{ with } \boldsymbol{\xi} \in \Xi = [-1, 1]^k.$$

For each product $i = 1, \ldots, n$, we choose the factor loading vector $\boldsymbol{\phi}_i \in \mathbb{R}^k$ uniformly at random from $[-1, 1]^k$ and, subsequently, scale it so that $\|\boldsymbol{\phi}_i\|_1 = 1$. We then set $\varphi_i = \|\boldsymbol{\phi}_i\|_1$. This parameter choice implies that $\{\boldsymbol{\phi}_i^\top \boldsymbol{\xi} + \varphi_i : \boldsymbol{\xi} \in \Xi\} = [0, 2]$. Moreover, we ensure that the factor loading vectors $\boldsymbol{\phi}_i$, $\boldsymbol{\phi}_j$ associated with two products i, j of the same category have the same signs (but typically not the same values) for each component; that is, $\operatorname{sgn}(\phi_{il}) = \operatorname{sgn}(\phi_{jl})$ for all $l = 1, \ldots, k$. This expresses the assumption that the demands for products of the same category are positively correlated. We fix $c_x = 0$ and select c_y uniformly at random from $[0, 2]$. Thus, we interpret the standard orders as sunk costs, and we aim to minimize the amount of express deliveries, which carry a per-unit premium of c_y over the standard orders. We select the inventory holding costs uniformly at random from $[3, 5]$, whereas the backlogging costs are chosen uniformly at random from $[0, 1/2]$. The upper bound on the express orders $\boldsymbol{y}(\boldsymbol{\xi})$ is set to $B = n$, which implies that up to 50% of the maximum demand can be covered via express deliveries.

To select the binding scenarios $\boldsymbol{\xi}^\star \in \Xi_A^\star$ to lift in each iteration of our primal–dual lifting scheme, we fix the optimal solution to the bounding problem $\overline{\mathcal{P}}(\Xi_S)$ and determine one binding scenario for each constraint by minimizing the constraint's slack over all $\boldsymbol{\xi} \in \Xi_A$. We then tentatively lift each such scenario $\boldsymbol{\xi}^\star$ and evaluate its impact on the upper bounding problem. Afterward, instead of lifting a single scenario in each iteration (as described in our algorithm outline in Section 3), we lift every scenario that improves the upper bound (see also our discussion in Remark 5). Our modification leads to an algorithm that performs fewer iterations than the one outlined in Section 3, but it is able to better close the gap for larger instances.

We first compare the scalability of our primal–dual lifting scheme with the other approaches from Table 1 as the number of products and the number of uncertain parameters vary. The results are presented in Table 2. Each instance class is identified by the label "*n*-*k*," where n denotes the number of products and k refers to the number of uncertain parameters, respectively. Here and in the following, the optimality gaps are computed relative to the mean value of (i) the smallest upper bound and (ii) the largest lower bound generated by any of the solution approaches listed in the table after six hours runtime. All values are averages over 25 randomly generated instances.

Table 2 shows that the column- and constraint-generation scheme of Zeng and Zhao (2013) clearly outperforms the other three methods.[1] Our primal–dual lifting scheme is able to close between 29.2% (instance 60–40) and 73.3% (instance 40–20) of the optimality gap of the affine decision rules. The Fourier–Motzkin elimination, on the other hand, only closes between 15.1% (instance 60–40) and 54.8% (instance 40–20) of the optimality gap of the affine decision rules. A closer inspection of the algorithms revealed that the majority of the runtime of the column- and constraint-generation scheme is spent on the mixed-integer linear programming subproblems, which do not grow in size and, therefore, guarantee a steady progress of the overall procedure. In contrast, both the scenario selection and the lifting steps of our primal–dual approach require considerable time, and the bounding problems tend to grow larger over time because of the increased complexity of the convex hulls of the remaining extreme points Ξ_A. The Fourier–Motzkin elimination scheme, finally, suffers from a rapid growth of the number of constraints, and the

Table 2. Inventory Management with Relatively Complete Recourse: Optimality Gaps of Our Method ("Primal–Dual"), the Fourier-Motzkin Elimination Proposed by Zhen et al. (2016) ("FME & SB"), the Column- and Constraint-Generation of Zeng and Zhao (2013) ("Zeng & Zhao"), and the Affine Decision Rule Approximation with the Sampling Bound of Hadjiyiannis et al. (2011) ("ADR & SB") for Various Time Limits and Instance Classes

	Primal–dual			FME & SB			Zeng & Zhao			
Instance	10 minutes	One hour	Six hours	10 minutes	One hour	Six hours	10 min	One hour	Six hours	ADR & SB
40–20, %	10.5	9.8	5.6	14.0	9.5	9.5	63.9	1.3	0.9	21.0
40–30, %	18.5	17.0	16.3	21.4	20.7	20.7	76.0	2.3	1.8	26.5
40–40, %	16.8	14.8	9.9	24.7	20.7	20.7	64.4	2.5	2.3	27.4
60–20, %	12.0	11.3	10.3	19.1	16.1	16.1	93.6	2.3	2.2	20.0
60–30, %	16.2	16.0	14.6	21.6	18.7	18.7	82.7	24.3	2.4	23.5
60–40, %	17.4	15.4	15.0	21.2	18.0	18.0	92.7	42.3	2.6	21.2

Note. We do not show the runtimes for the affine decision rule approximation as the results were obtained within one minute for all instances.

Table 3. Inventory Management with Relatively Complete Recourse: Percentages of the Overall Runtime of Our Algorithm That Are Spent in Each Iteration on Solving the Upper Bound Problems $\overline{\mathcal{P}}(\Xi_S)$ ("Upper Bounds"), Solving the Lower Bound Problems $\underline{\mathcal{P}}(\Xi_S)$ ("Lower Bounds"), Finding the Neighbors of the Extreme Points $\xi^\star \in \Xi_A^\star$ to Be Lifted ("Neighbors), Calculating the Convex Hulls conv $([\text{ext}\,\Xi_A] \setminus \{\xi^\star\})$ ("Convex Hulls") and the Remaining Overhead ("Remainder")

	Upper bounds	Lower bounds	Neighbors	Convex hulls	Remainder
40–20, %	28.8	0.3	9.4	58.7	2.7
40–30, %	31.9	0.3	15.7	48.5	3.5
40–40, %	43.2	0.9	17.9	34.8	3.2
60–20, %	45.1	0.5	14.1	37.0	3.3
60–30, %	61.0	0.8	16.5	19.4	2.3
60–40, %	69.6	0.9	19.6	7.7	2.2

constraint elimination procedure outlined in section 4 of Zhen et al. (2016) was only able to identify a small fraction of the constraints as redundant.

We have also solved the inventory management problem with the piecewise affine decision rules proposed in Georghiou et al. (2015). We have not been able to obtain any noticeable improvements relative to the results of the affine decision rules. Moreover, we have attempted to replace the sampling bound of Hadjiyiannis et al. (2011) with the refined bound proposed in Bertsimas and de Ruiter (2016). This has not led to significant improvements either. Finally, we have tried to solve the instances monolithically with the extreme point reformulation $\mathcal{P}'$. For the given time limit of six hours, this reformulation could not be solved for any of the instances.

Table 3 shows the fractions of the runtime that are spent on the different steps of our primal–dual lifting scheme. We observe that the majority of the runtime is spent on (i) solving the upper bound problems $\overline{\mathcal{P}}(\Xi_S)$ and (ii) calculating the convex hulls conv $([\text{ext}\,\Xi_A]\setminus\{\xi^\star\})$. This is not surprising because, in each iteration, our algorithm tentatively lifts every binding scenario $\xi^\star \in \Xi_A^\star$ and evaluates its impact on the upper bounding problem, thus requiring the solution of many upper bounding problems as well as the calculation of many convex hulls. This approach trades off the conflicting goals of lifting sufficiently many extreme points in each iteration (and, thus, not wasting useful computation) and lifting only those extreme points that improve the bounds (and, thus, not unduly growing the representations of the bounding problems $\overline{\mathcal{P}}(\Xi_S)$ and $\underline{\mathcal{P}}(\Xi_S)$).

We now consider a variant of the inventory management problem (5) that additionally imposes an upper bound I on the number of products $\mathbf{e}^\top \boldsymbol{h}(\boldsymbol{\xi})$ held in inventory:

$$\begin{array}{lll} \text{minimize} & \tau & \\ \text{subject to} & \left.\begin{array}{l} \tau \geq \mathbf{e}^\top\big[c_x \cdot \boldsymbol{x} + c_y \cdot \boldsymbol{y}(\boldsymbol{\xi}) \\ \qquad + c_h \cdot \boldsymbol{h}(\boldsymbol{\xi}) + c_b \cdot \boldsymbol{b}(\boldsymbol{\xi})\big] \\ \boldsymbol{h}(\boldsymbol{\xi}) \geq \boldsymbol{x} + \boldsymbol{y}(\boldsymbol{\xi}) - \boldsymbol{D}(\boldsymbol{\xi}) \\ \boldsymbol{b}(\boldsymbol{\xi}) \geq \boldsymbol{D}(\boldsymbol{\xi}) - \boldsymbol{x} - \boldsymbol{y}(\boldsymbol{\xi}) \\ \mathbf{e}^\top \boldsymbol{y}(\boldsymbol{\xi}) \leq B \\ \mathbf{e}^\top \boldsymbol{h}(\boldsymbol{\xi}) \leq I \end{array}\right\} \forall \boldsymbol{\xi} \in \Xi \\ & \boldsymbol{x} \in \mathbb{R}^n_+, \quad \boldsymbol{y}, \boldsymbol{h}, \boldsymbol{b} : \Xi \mapsto \mathbb{R}^n_+. & \end{array} \tag{6}$$

In our experiments, we chose the inventory bound $I = n/2$, which implies that up to 25% of the maximum demand can be held in inventory. We emphasize that even the equivalent min–max formulation of this problem (as employed by the column- and constraint-generation scheme) lacks a relatively complete recourse: excessive first-stage orders $\boldsymbol{x}$ that outstrip

Table 4. Inventory Management Without Relatively Complete Recourse

	Primal–dual			FME & SB			Zeng & Zhao			
Instance	10 minutes	One hour	Six hours	10 minutes	One hour	Six hours	10 minutes	One hour	Six hours	ADR & SB
40–20, %	11.6	9.4	7.9	12.5	12.5	12.5	118.4	118.4	118.4	21.7
40–30, %	20.3	18.1	17.6	23.5	22.8	22.8	125.1	125.1	125.1	29.4
40–40, %	14.9	14.8	13.0	23.8	21.9	21.9	118.7	118.7	118.7	26.9
60–20, %	13.7	13.1	11.7	19.1	16.2	16.2	120.5	120.5	120.5	19.7
60–30, %	19.6	18.0	17.8	21.8	18.9	18.9	119.2	119.2	119.2	22.7
60–40, %	15.4	14.9	11.9	18.9	18.3	18.3	118.6	118.6	118.6	20.7

Note. The columns and table entries have the same interpretation as in Table 2.

the demand by more than I units, that is, first-stage orders satisfying $\mathbf{e}^\top[x - D(\xi)]_+ > I$, lead to an infeasible second stage.

Table 4 compares the scalability of our primal–dual lifting scheme with the other approaches from Table 1 for the problem (6). Interestingly, the column- and constraint-generation scheme is now outperformed by all of the other approaches. A closer inspection of the results revealed that the column- and constraint-generation scheme, which now requires the solution of both feasibility and optimality subproblems, typically gets stuck at an early iteration. Indeed, contrary to the optimality subproblems, which can be terminated early as long as a feasible (but possibly suboptimal) solution has been determined, the feasibility subproblems need to be solved until either the lower bound strictly exceeds zero (indicating that the current first-stage decision is infeasible and producing a violated realization of the uncertain parameter vector ξ) or until the upper bound equals zero (in which case the feasibility subproblem has been solved to optimality, indicating that the current first-stage decision is feasible). It turns out that the solution of these feasibility subproblems becomes very challenging, and the algorithm typically gets stuck at an early iteration when it neither produces a violated realization of ξ nor confirms the feasibility of the candidate first-stage decision x. In contrast, neither our primal–dual lifting scheme nor the Fourier–Motzkin elimination are significantly affected by the lack of a relatively complete recourse: the primal–dual lifting scheme closes between 21.6% (instance 60–30) and 63.6% (instance 40–20) of the optimality gap of the affine decision rules, and the Fourier–Motzkin elimination closes between 11.6% (instance 60–40) and 42.4% (instance 40–20) of the optimality gap of the affine decision rules.

Acknowledgments

The authors are indebted to David Avis and Komei Fukuda for valuable advice on various aspects of polyhedral combinatorics as well as the constructive suggestions of the three anonymous referees.

Endnote

[1] Our implementation of the column- and constraint-generation scheme uses a time limit of 10 minutes for the solution of each subproblem. Although this ensures that the algorithm does not get stuck in the attempt to solve one of the subproblems, it implies that the overall scheme will typically not obtain optimal solutions.

References

An Y, Zeng B, Zhang Y, Zhao L (2014) Reliable *p*-median facility location problem: Two-stage robust models and algorithms. *Transportation Res. Part B: Methodological* 64:54–72.

Anderson BDO, Moore JB (1990) *Optimal Control: Linear Quadratic Methods* (Prentice-Hall, Inc., Upper Saddle River, NJ).

Angulo G, Ahmed S, Dey SS, Kaibel V (2015) Forbidden vertices. *Math. Oper. Res.* 40(2):350–360.

Ardestani-Jaafari A, Delage E (2016) Robust optimization of sums of piecewise linear functions with application to inventory problems. *Oper. Res.* 64(2):474–494.

Atamtürk A, Zhang M (2007) Two-stage robust network flow and design under demand uncertainty. *Oper. Res.* 55(4):662–673.

Ayoub J, Poss M (2016) Decomposition for adjustable robust linear optimization subject to uncertainty polytope. *Comput. Management Sci.* 13(2):219–239.

Ben-Tal A, Goryashko A, Guslitzer E, Nemirovski A (2004) Adjustable robust solutions of uncertain linear programs. *Math. Programming* 99(2):351–376.

Bertsimas D, Caramanis C (2010) Finite adaptibility in multistage linear optimization. *IEEE Trans. Automatic Control* 55(12):2751–2766.

Bertsimas D, de Ruiter FJCT (2016) Duality in two-stage adaptive linear optimization: Faster computation and stronger bounds. *INFORMS J. Comput.* 28(3):500–511.

Bertsimas D, Dunning I (2016) Multistage robust mixed-integer optimization with adaptive partitions. *Oper. Res.* 64(4):980–998.

Bertsimas D, Georghiou A (2015) Design of near optimal decision rules in multistage adaptive mixed-integer optimization. *Oper. Res.* 63(3):610–627.

Bertsimas D, Goyal V (2010) On the power of robust solutions in two-stage stochastic and adaptive optimization problems. *Math. Oper. Res.* 35(2):284–305.

Bertsimas D, Goyal V (2012) On the power and limitations of affine policies in two-stage adaptive optimization. *Math. Programming* 134(2):491–531.

Bertsimas D, Brown DB, Caramanis C (2011a) Theory and applications of robust optimization. *SIAM Rev.* 53(3):464–501.

Bertsimas D, Goyal V, Sun XA (2011b) A geometric characterization of the power of finite adaptability in multistage stochastic and adaptive optimization. *Math. Oper. Res.* 36(1):24–54.

Bertsimas D, Iancu DA, Parrilo PA (2010) Optimality of affine policies in multi-stage robust optimization. *Math. Oper. Res.* 35(2):363–394.

Bertsimas D, Iancu DA, Parrilo PA (2011c) A hierarchy of near-optimal policies for multistage adaptive optimization. *IEEE Trans. Automatic Control* 56(12):2809–2824.

Bertsimas D, Litvinov E, Sun XA, Zhao J, Zheng T (2013) Adaptive robust optimization for the security constrained unit commitment problem. *IEEE Trans. Power Systems* 28(1):52–63.

Buchta C, Müller J, Tichy RF (1985) Stochastical approximation of convex bodies. *Math. Ann.* 271(2):225–235.

Büeler B, Enge A, Fukuda K (2000) Exact volume computation for polytopes: A practical study. Kalai G, Ziegler GM, eds. *Polytopes—Combinatorics and Computation* (Birkhäuser, Basel, Switzerland), 131–154.

Chen X, Zhang Y (2009) Uncertain linear programs: Extended affinely adjustable robust counterparts. *Oper. Res.* 57(6):1469–1482.

Chen X, Sim M, Sun P, Zhang J (2008) A linear decision-based approximation approach to stochastic programming. *Oper. Res.* 56(2):344–357.

Delage E, Iancu DA (2015) Robust multistage decision making. Aleman DM, Thiele AC, eds. *The Operations Research Revolution: Tutorials in Operations Research* (INFORMS, Catonsville, MD), 20–46.

Gabrel V, Murat C, Thiele A (2014) Recent advances in robust optimization: An overview. *Eur. J. Oper. Res.* 235(3):471–483.

Gal T (1993) Degeneracy graphs: Theory and application an updated survey. *Ann. Oper. Res.* 46(1):81–105.

Georghiou A, Tsoukalas A, Wiesemann W (2019) Robust dual dynamic programming. *Oper. Res.* 67(3):599–904.

Georghiou A, Wiesemann W, Kuhn D (2015) Generalized decision rule approximations for stochastic programming via liftings. *Math. Programming* 152(1):301–338.

Goh J, Sim M (2010) Distributionally robust optimization and its tractable approximations. *Oper. Res.* 58(4):902–917.

Gorissen BL, den Hertog D (2013) Robust counterparts of inequalities containing sums of maxima of linear functions. *Eur. J. Oper. Res.* 227(1):30–43.

Gorissen BL, Yanıkoğlu İ, den Hertog D (2015) A practical guide to robust optimization. *Omega* 53:124–137.

Gounaris CE, Wiesemann W, Floudas CA (2013) The robust capacitated vehicle routing problem under demand uncertainty. *Oper. Res.* 61(3):677–693.

Guslitser E (2002) Uncertainty-immunized solutions in linear programming. Unpublished master's thesis, Technion, Haifa, Israel.

Hadjiyiannis MJ, Goulart PJ, Kuhn D (2011) A scenario approach for estimating the suboptimality of linear decision rules in two-stage robust optimization. *Proc. 50th IEEE Conf. Decision Control Eur. Control Conf.* (IEEE, Piscataway, NJ), 7386–7391.

Hanasusanto GA, Kuhn D (2018) Conic programming reformulations of two-Stage distributionally robust linear programs over Wasserstein balls. *Oper. Res.* 66(3):849–869.

Hanasusanto GA, Kuhn D, Wiesemann W (2015) *K*-adaptability in two-stage robust binary programming. *Oper. Res.* 63(4):877–891.

Henk M (2012) Löwner–John ellipsoids. *Documenta Mathematica—Extra Volume: Optimization Stories* (Deutschen Mathematiker-Vereinigung, Berlin), 95–106.

Jiang R, Zhang M, Li G, Guan Y (2014) Two-stage network constrained robust unit commitment problem. *Eur. J. Oper. Res.* 234(3):751–762.

Kuhn D, Wiesemann W, Georghiou A (2011) Primal and dual linear decision rules in stochastic and robust optimization. *Math. Programming* 130(1):177–209.

Lappas NH, Gounaris CE (2016) Multi-stage adjustable robust optimization for process scheduling under uncertainty. *AIChE J.* 62(5):1646–1667.

Liebchen C, Lübbecke M, Möhring RH, Stiller S. (2009) The concept of recoverable robustness, linear programming recovery, and railway applications. Ahuja RK, Möhring RH, Zaroliagis CD, eds. *Robust and Online Large-Scale Optimization: Models and Techniques for Transportation Systems* (Springer, Berlin, Heidelberg), 1–27.

Lorca A, Sun XA (2015) Adaptive robust optimization with dynamic uncertainty sets for multi-period economic dispatch under significant wind. *IEEE Trans. Power Systems* 30(4):1702–1713.

Lorca A, Sun XA (2017) Multistage robust unit commitment with dynamic uncertainty sets and energy storage. *IEEE Trans. Power Systems* 32(3):1678–1688.

Postek K, den Hertog D (2016) Multistage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set. *INFORMS J. Comput.* 28(3):553–574.

Ruiz C, Conejo AJ (2015) Robust transmission expansion planning. *Eur. J. Oper. Res.* 242(2):390–401.

Seidel R (2004) Convex hull computations. Toth CD, O'Rourke J, Goodman JE, eds. *Handbook of Discrete and Computational Geometry* (Chapman and Hall, Boca Raton, FL), 494–512.

Song Y, Luedtke J (2015) An adaptive partition-based approach for solving two-stage stochastic programs with fixed recourse. *SIAM J. Optim.* 25(3):1344–1367.

Subramanyam A, Gounaris C, Wiesemann W (2017a) *K*-adaptability in two-stage mixed-integer robust optimization. Accessed January 22, 2020, http://www.optimization-online.org/DB_FILE/2017/06/6093.pdf.

Subramanyam A, Mufalli F, Pinto JM, Gounaris CE (2017b) Robust multi-period vehicle routing under customer order uncertainty. Accessed January 22, 2020, http://www.optimization-online.org/DB_HTML/2017/04/5947.html.

Thiele A, Terry T, Epelman M (2010) Robust linear optimization with recourse. Technical report, Lehigh University, Bethlehem, PA, and University of Michigan, Ann Arbor.

Vayanos P, Kuhn D, Rustem B (2011) Decision rules for information discovery in multi-stage stochastic programming. *Proc. 50th IEEE Conf. Decision Control Eur. Control Conf.* (IEEE, Piscataway, NJ), 7368–7373.

Warrington J, Goulart P, Mariéthoz S, Morari M (2013) Policy-based reserves for power systems. *IEEE Trans. Power Systems* 28(4):4427–4437.

Wiesemann W, Kuhn D, Rustem B (2012) Robust resource allocations in temporal networks. *Math. Programming* 135(1–2):437–471.

Xu G, Burer S (2018) A copositive approach for two-stagae adjustable robust optimization with uncertain right-hand sides. *Comput. Optim. Appl.* 70(1):33–59.

Yanıkoğlu İ, Gorissen BL, den Hertog D (2019) Adjustable robust optimization—a survey and tutorial. *Eur. J. Oper. Res.* 277(3):799–813.

Zeng B, Zhao L (2013) Solving two-stage robust optimization problems using a column-and-constraint generation method. *Oper. Res. Lett.* 41(5):457–461.

Zhao C, Wang J, Watson J-P, Guan Y (2013) Multi-stage robust unit commitment considering wind and demand response uncertainties. *IEEE Trans. Power Systems* 28(3):2708–2717.

Zhen J, den Hertog D, Sim M (2018) Adjustable robust optimization via Fourier-Motzkin elimination. *Oper. Res.* 66(4):1086–1100.

Ziegler GM. (2000) Lectures on 0/1-polytopes. Kalai G, Ziegler GM, eds. *Polytopes—Combinatorics and Computation* (Birkhäuser, Basel, Switzerland), 1–41.

Angelos Georghiou is an assistant professor of operations management with the Desautels Faculty of Management, McGill University, Montreal, Quebec, Canada. His current research interests include the development of tractable computational methods for the solution of stochastic and robust optimization problems and applications in operations management, healthcare, and energy.

Angelos Tsoukalas is an assistant professor of operations research at Olayan School of Business, American University of Beirut. His research interests lie in optimization theory, algorithms, and applications.

Wolfram Wiesemann is associate professor of management science and operations as well as fellow of the KPMG Centre for Advanced Business Analytics at Imperial College Business School, London. His research interests revolve around the methodological aspects of decision making under uncertainty as well as applications in operations management, energy, and finance.