Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Robust Optimization of Sums of Piecewise Linear Functions with Application to Inventory Problems

Amir Ardestani-Jaafari, Erick Delage

OPERATIONS RESEARCH
Vol. 64, No. 2, March–April 2016, pp. 474–494
ISSN 0030-364X (print) | ISSN 1526-5463 (online)
http://dx.doi.org/10.1287/opre.2016.1483

Robust Optimization of Sums of Piecewise Linear Functions with Application to Inventory Problems

Amir Ardestani-Jaafari, Erick Delage
Department of Decision Sciences, HEC Montréal, Montreal, Quebec H3T 2A7, Canada
{amir.ardestani-jaafari@hec.ca, erick.delage@hec.ca}

Robust optimization is a methodology that has gained a lot of attention in the recent years. This is mainly due to the simplicity of the modeling process and ease of resolution even for large scale models. Unfortunately, the second property is usually lost when the cost function that needs to be "robustified" is not concave (or linear) with respect to the perturbing parameters. In this paper we study robust optimization of sums of piecewise linear functions over polyhedral uncertainty set. Given that these problems are known to be intractable, we propose a new scheme for constructing conservative approximations based on the relaxation of an embedded mixed-integer linear program and relate this scheme to methods that are based on exploiting affine decision rules. Our new scheme gives rise to two tractable models that, respectively, take the shape of a linear program and a semidefinite program, with the latter having the potential to provide solutions of better quality than the former at the price of heavier computations. We present conditions under which our approximation models are exact. In particular, we are able to propose the first exact reformulations for a robust (and distributionally robust) multi-item newsvendor problem with budgeted uncertainty set and a reformulation for robust multiperiod inventory problems that is exact whether the uncertainty region reduces to a L_1-norm ball or to a box. An extensive set of empirical results will illustrate the quality of the approximate solutions that are obtained using these two models on randomly generated instances of the latter problem.

Keywords: robust optimization; piecewise linear; linear programming relaxation; semidefinite program; tractable approximations; newsvendor problem; inventory problem.
Subject classifications: inventory/production: uncertainty; programming: integer: applications; programming: integer: algorithms: relaxation; programming: infinite dimensional.
Area of review: Optimization.
History: Received July 2013; revisions received March 2015, November 2015; accepted December 2015. Published online in *Articles in Advance* March 24, 2016.

1. Introduction

Since the seminal work of Ben-Tal and Nemirovski (1998), robust optimization is a methodology that has attracted a large amount of attention. Such attention has stemmed in application fields that range from engineering problems like structural design (Ben-Tal and Nemirovski 1997) and circuit design (Boyd et al. 2005), management problems such as portfolio optimization (Goldfarb and Iyengar 2002) and supply chain management (Ben-Tal et al. 2005), to an array of data mining applications such as classification (Xu et al. 2009), regression (El Ghaoui and Lebret 1997) and parameter estimation (Calafiore and El Ghaoui 2001) (see Bertsimas et al. 2011a for a detailed review of such applications). Two important factors that have contributed to this success are (1) the simplicity of the modeling paradigm, and (2) the tractability of many resulting formulations, thus enabling the resolution of problems of scales that can match the practical needs. Unfortunately, the second property is usually lost when the cost function that needs to be "robustified" is not concave (or linear) with respect to the perturbing parameters.

This paper focuses on the following robust optimization problem:

$$\underset{\mathbf{x}\in\mathscr{X}}{\text{minimize}} \max_{\boldsymbol{\zeta}\in\mathscr{Z}} \sum_{i=1}^{N} h_i(\mathbf{x}, \boldsymbol{\zeta}), \tag{1}$$

where $\mathscr{X} \subseteq \mathbb{R}^n$ is a bounded polyhedral set of feasible solutions for the decision $\mathbf{x}$, $\mathscr{Z} \subseteq \mathbb{R}^m$ is the set containing the possible perturbation $\boldsymbol{\zeta}$, and for each i the cost function $h_i(\mathbf{x}, \boldsymbol{\zeta})$ is piecewise linear and convex in both $\mathbf{x}$ and $\boldsymbol{\zeta}$ (although not necessarily jointly convex). In particular, this means that the cost function can be expressed as follows:

$$h_i(\mathbf{x}, \boldsymbol{\zeta}) := \max_k \{\mathbf{c}_x^{i,k}(\boldsymbol{\zeta})^T \mathbf{x} + d_x^{i,k}(\boldsymbol{\zeta})\} := \max_k \{\mathbf{c}_\zeta^{i,k}(\mathbf{x})^T \boldsymbol{\zeta} + d_\zeta^{i,k}(\mathbf{x})\},$$

for some affine mappings $\mathbf{c}_x^{i,k}: \mathbb{R}^m \to \mathbb{R}^n$, $d_x^{i,k}: \mathbb{R}^m \to \mathbb{R}$, $\mathbf{c}_\zeta^{i,k}: \mathbb{R}^n \to \mathbb{R}^m$ and $d_\zeta^{i,k}: \mathbb{R}^n \to \mathbb{R}$. Although objective functions that take the form of sums of piecewise linear function abound in practice, exact solutions to the robust version of these problems are often considered impossible to obtain because of the computational difficulties that arise in solving the inner maximization problem (a.k.a. adversarial problem). Besides the two inventory problems that will be discussed later, such structured functions also play an important role in multiobjective optimization and machine learning (see the electronic companion at http://dx.doi.org/10.1287/opre.2016.1483 for details).

Recently, Gorissen and den Hertog (2013) have made a valuable effort at presenting a comprehensive overview of three families of solution methods that can be employed for this problem: namely, exact methods, tractable conservative approximations,[1] and cutting plane methods. Unfortunately, while there are a few very special cases for which finding an exact solution is known to be tractable, still very little is known theoretically about the quality of conservative approximations that are available. In this paper, we attempt to reduce this gap by bringing the following contributions:

1. We propose a novel scheme for deriving tractable conservative approximations that connects for the first time the suboptimality of an approximate solution directly to the integrality gap of an associated mixed integer linear program (MILP). This allows us to identify fairly general conditions under which the concept of total unimodularity can be used to establish that the approximate solution obtained by solving a linear program of reasonable size is exactly optimal. The connection to MILP optimization also naturally allows us to propose a tighter conservative approximation model that takes the shape of a semidefinite program. This is perhaps surprising given that it is well known that, while schemes that are based on quadratic decision rules will lead to semidefinite program (SDP) approximation models when the uncertainty set is ellipsoidal, such adjustment functions lead in general to optimization problems that are computationally intractable, for instance when the uncertainty set is polyhedral (see Ben-Tal et al. 2009a, p. 372). Indeed, we show for the first time how to obtain SDP approximation models for such uncertainty sets by employing affine decision rules on a clever reformulation of the objective function.

2. We provide for the first time an exact tractable reformulation for a robust multi-item newsvendor problem with demand uncertainty that is nonrectangular, namely where it takes the shape of a budgeted uncertainty set with an integer budget. A novel tractable reformulation is also presented for the distributionally robust version of this problem in which the distribution information includes a budgeted uncertainty set for the support, the mean vector, and a list of lower bounds on first order partial moments. To the best of our knowledge, this appears to be first exact tractable reformulation for instances of multi-item newsvendor problem where there exists information about how the demand for different items behave jointly, a problem that was left open since the early work of Scarf (1958).

3. We propose a new conservative approximation model for a robust multiperiod inventory problem where all orders must be made initially. We prove that this model produces an exact solution when facing a budgeted uncertainty set with a budget equal to one or to the total size of the horizon. Although exact reformulations exist for each of these extreme cases, this is the first model known to be exact for both cases simultaneously. Our empirical study also provides evidence that the suboptimality gap is relatively small with our new model (less than 0.3% gap on average with a maximum observed gap of 5%) when the budget takes on intermediate values. Finally, we present extensive empirical evidence that this model can be used to identify ordering strategies that make better trade-off between performance and robustness in comparison to strategies obtained using existing tractable method in the literature.

The paper is organized as follows. We start in §2 with a brief review of related work and currently available methods for solving problem (1). Section 3 presents our notation. In §4 we introduce our new approximation scheme for the robust optimization problem (1) that is based on the fractional relaxation of an associated mixed integer linear program. In §5 we present implications of our results for a robust and distributionally robust multi-item newsvendor problem. In §6 we apply the new models to a robust multiperiod inventory problem. Section 7 presents experiments on an inventory problem that attempt to evaluate the relative tightness of different approximation schemes and illustrate how one can employ these schemes to explore the trade-offs between expected performance and robustness in choosing an order policy. Finally, we conclude and provide some directions future research in §8.

2. Background and Prior Work

Our work follows very closely the initiative of Gorissen and den Hertog (2013), who were interested in solving problems of the form (1) and where a comprehensive overview of available methods is presented. In Gorissen and den Hertog (2013), the robust optimization of the sums of maxima of linear functions takes the shape of

$$\underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} \max_{\boldsymbol{\zeta}\in\mathcal{Z}} \left\{ \ell(\boldsymbol{\zeta},\mathbf{x}) + \sum_{i=1}^{N} \max_k \{\ell_{i,k}(\boldsymbol{\zeta},\mathbf{x})\} \right\},$$

where ℓ and $\ell_{i,k}$ are bi-affine functions in the uncertain parameter $\boldsymbol{\zeta}$ and the decision variable $\mathbf{x}\in\mathbb{R}^n$, and where $\mathcal{Z}$ is the uncertainty set. Given that $\mathcal{Z}$ is convex, the authors first describe an exact solution approach that is based on reducing the worst-case analysis to a search over the vertices of $\mathcal{Z}$ since the objective function is convex in $\boldsymbol{\zeta}$. This leads to the equivalent finite formulation

$$\begin{aligned} &\underset{\mathbf{x},\mathbf{y}}{\text{minimize}} \quad \max_{v\in\mathcal{V}} \left\{ \ell(\boldsymbol{\zeta}^v,\mathbf{x}) + \sum_{i=1}^{N} y_i^v \right\} \\ &\text{subject to} \quad y_i^v \geqslant \ell_{i,k}(\boldsymbol{\zeta}^v,\mathbf{x}), \quad \forall i,\forall k,\forall v\in\mathcal{V}, \end{aligned}$$

where $\mathcal{V}=\{1,2,\ldots,V\}$ with V the number of vertices of $\mathcal{Z}$, and $\{\boldsymbol{\zeta}^v\}_{v=1}^V$ is the finite set of such vertices. The authors do warn their reader that computational complexity of this approach grows exponentially with respect to the number of constraints that define $\mathcal{Z}$.

Gorissen and den Hertog also propose using cutting plane methods to solve these problems exactly (especially when enumerating the vertices becomes unthinkable). In fact, there is empirical evidence that seems to indicate that such methods are particularly effective in practice for solving two-stage robust optimization problems (see Zeng and Zhao 2013). In each iteration of a cutting plane method, there is a need to establish the worst-case $\boldsymbol{\zeta}$ for some fixed $\mathbf{x}$ to produce a cutting plane: i.e.,

$$\max_{\boldsymbol{\zeta}\in\mathcal{Z}}\left\{\ell(\boldsymbol{\zeta},\mathbf{x})+\sum_{i=1}^{N}\max_{k}\{\ell_{i,k}(\boldsymbol{\zeta},\mathbf{x})\}\right\}.$$

Although there exist some special cases where an efficient procedure might be identified (see Bienstock and Özbay 2008 for an example), the authors suggest that in general this problem can be solved by solving a MILP similar to

$$\begin{aligned}
\underset{\boldsymbol{\zeta}\in\mathcal{Z},\,\mathbf{y},\,\mathbf{z}}{\text{maximize}}\quad & \ell(\boldsymbol{\zeta},\mathbf{x})+\sum_{i=1}^{N}y_i\\
\text{subject to}\quad & y_i\leqslant \ell_{i,k}(\boldsymbol{\zeta},\mathbf{x})+M(1-z_{i,k}),\quad \forall\, i,\ \forall\, k,\\
& \sum_{k=1}^{K}z_{i,k}=1,\quad \forall\, i\\
& \boldsymbol{\zeta}\in\mathcal{Z},\quad \mathbf{z}\in\{0,1\}^{N\times K}.
\end{aligned}$$

Unfortunately, although software products that handle such models are well developed, solving this problem is generally NP-hard (see NP-hardness discussion in §4), thus making this approach prohibitive for large problems. In particular, the experiments we conduct in §4.4 identified instances of such mixed integer linear programs reformulations that could not be solved in less than a day of computation already when $N=m=64$. Finally, polynomial-time solvability of cutting plane methods is not guaranteed except for the ellipsoid method, which is rarely used in practice.

Gorissen and den Hertog (2013) finally explain how the theory of affinely adjustable robust counterpart (AARC) proposed by Ben-Tal et al. (2004) can be used to obtain a conservative approximation method. In this case, each convex term of the objective is replaced with an affine function that is adjusted optimally while ensuring that the objective function upper bounds the true objective. The resulting model takes the following shape:

$$\begin{aligned}
\underset{\mathbf{x},\,\mathbf{v},\,\mathbf{w}}{\text{minimize}}\quad & \max_{\boldsymbol{\zeta}\in\mathcal{Z}}\left\{\ell(\boldsymbol{\zeta},\mathbf{x})+\sum_{i=1}^{N}\{v_i+\mathbf{w}_i^T\boldsymbol{\zeta}\}\right\}\\
\text{subject to}\quad & v_i+\mathbf{w}_i^T\boldsymbol{\zeta}\geqslant \ell_{i,k}(\boldsymbol{\zeta},\mathbf{x}),\quad \forall\, i,\ \forall\, k,\ \forall\,\boldsymbol{\zeta}\in\mathcal{Z},
\end{aligned}$$

for which one can easily formulate a finite dimensional linear programming reformulation using duality theory. It is mentioned that this approach can be improved by using a lifting of the uncertainty space (Chen and Zhang 2009) or by involving quadratic decision rules if the uncertainty set is ellipsoidal. Although the AARC approach is often tractable, very little is theoretically known about the suboptimality of the obtained approximate solution. (We refer the reader to Iancu et al. 2013 for the most general results to date on this topic.)

As mentioned in the introduction, many robust inventory problems can be considered a special case of problem (1). In Bertsimas and Thiele (2006), the authors seem to have been the first to propose an approximation method to solve such problems. Their approach relies on finding the worst-case cost of each period individually before summing the results over all periods. Effectively, they replace problem (1) with the following:

$$\underset{\mathbf{x}\in\mathcal{X}}{\text{minimize}}\sum_{i=1}^{N}\max_{\boldsymbol{\zeta}\in\mathcal{Z}}h_i(\mathbf{x},\boldsymbol{\zeta}).$$

Interestingly, when the uncertainty set takes the shape of the budgeted uncertainty set (see Bertsimas and Sim 2004), they show that the optimal robust policy is equivalent to the optimal policy of the nominal problem under a specifically designed demand vector. In spite of having been used in many occasions (e.g., José Alem and Morabito 2012, Wei et al. 2011), as noted in Gorissen and den Hertog (2013), this conservative approximation does not impose any relation between the worst-case $\boldsymbol{\zeta}$ used to evaluate the different periods.

It appears that Ben-Tal et al. (2004) were the first to address a robust inventory problem in which there is a possibility to make adjustments to future orders as information about demand becomes available. They provide conservative approximations of the problem by applying the concept of affine decision rules. In Ben-Tal et al. (2005), similar ideas are applied to a supply chain problem. Interestingly, the empirical experiments presented there seem to indicate that AARC can perform surprisingly well. A similar success was achieved in Ben-Tal et al. (2009b, §3.2).

3. Notation

We briefly review some notation that is used in the remaining sections. First, let $\mathbf{e}_i$ be the ith column of the identity matrix, and let $\mathbf{1}$ be the vector of all ones, both of their dimensions should be clear from context. Given two matrices of same sizes $\mathbf{A}\cdot\mathbf{B}$ refers to the Frobenius inner product which returns $\sum_{i,j}A_{i,j}B_{i,j}$. We use $\mathbf{A}_{i,:}$ to refer to the ith row of $\mathbf{A}$, and $\mathbf{A}_{:,j}$ would refer to the jth column of $\mathbf{A}$. For the sake of clarity, given a vector $\mathbf{b}$ we might use $(\mathbf{b})_i$, instead of b_i, to refer to the ith component of the vector $\mathbf{b}$.

4. Mixed-Integer Linear Programming Based Approximation

In this section we seek to obtain a conservative approximation of problem (1) using linearization schemes that are used in the field of mixed-integer linear programming. In particular, it is well known that the inner maximization problem

$$\underset{\boldsymbol{\zeta}\in\mathcal{Z}}{\text{maximize}}\sum_{i=1}^{N}\max_{k}\{\mathbf{c}_{i,k}^T\boldsymbol{\zeta}+d_{i,k}\},\tag{2}$$

where $\mathcal{Z}$ is polyhedral and where we dropped the dependence of $\mathbf{c}_{i,k}$ and $d_{i,k}$ on $\mathbf{x}$ for clarity, is NP-hard (see the e-companion for a proof). Given that $\mathcal{Z}$ is polyhedral and bounded, we can assume without loss of generality[2] that it is represented as

$$\mathcal{Z} := \{\boldsymbol{\zeta} \in [-1, 1]^m \mid \mathbf{A}\boldsymbol{\zeta} \leqslant \mathbf{b},\ \|\boldsymbol{\zeta}\|_1 \leqslant \Gamma\},$$

for some $\mathbf{A} \in \mathbb{R}^{p\times m}$ and $\mathbf{b} \in \mathbb{R}^p_+$, and some $0 \leqslant \Gamma \leqslant m$, and with $\mathbf{0} \in \mathcal{Z}$ capturing the "nominal" (i.e., most likely) scenario for $\boldsymbol{\zeta}$. Note that this representation reduces to the budgeted uncertainty set when $\mathbf{A} := \mathbf{0}$ and $\mathbf{b} := \mathbf{0}$, which is the most natural way of capturing that each ζ_i is a perturbation of similar magnitudes while one does not expect too many terms of $\boldsymbol{\zeta}$ being perturbed simultaneously (see Bertsimas and Sim 2004 and its ubiquitous use in robust optimization applications). In the more general case, we expect this representation to be especially relevant in problems where one wishes to emphasize that the uncertainty region is roughly symmetrical around the nominal scenario $\boldsymbol{\zeta}_0 := \mathbf{0}$. Hence, we are left with the following adversarial problem:

$$\underset{\boldsymbol{\zeta}\in\mathbb{R}^m}{\text{maximize}} \quad \sum_{i=1}^{N}\Big\{\max_k \mathbf{c}_{i,k}^T\boldsymbol{\zeta} + d_{i,k}\Big\} \tag{3a}$$

$$\text{subject to} \quad \mathbf{A}\boldsymbol{\zeta} \leqslant \mathbf{b}, \tag{3b}$$

$$\|\boldsymbol{\zeta}\|_\infty \leqslant 1, \tag{3c}$$

$$\|\boldsymbol{\zeta}\|_1 \leqslant \Gamma. \tag{3d}$$

We will initially present two approximation models that will trade off between computational requirements and quality of the solution. We will then relate these models to the important family of approximation schemes known as AARCs.

4.1. Linear Programming Approximation Model

Our first step is to convert this convex maximization problem to a mixed-integer quadratic program by replacing the objective function with

$$\max_{\{\mathbf{z}\in\{0,1\}^{N\times K} \mid \sum_{k=1}^K z_{i,k}=1,\forall i\}} \sum_{i=1}^N\sum_{k=1}^K z_{i,k}(\mathbf{c}_{i,k}^T\boldsymbol{\zeta} + d_{i,k}),$$

where we introduced additional adversarial binary decision variables $z_{i,k}$. As often done for mixed-integer quadratic programs, we will circumvent the difficulty of maximizing the terms that are quadratic in $\mathbf{z}$ and $\boldsymbol{\zeta}$ by linearizing the objective function, yet only after replacing the perturbation variables by the sum of their positive and negative parts (i.e., $\boldsymbol{\zeta} := \boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-$). Specifically, this linearization is obtained by replacing instances of $z_{i,k}\cdot\boldsymbol{\zeta}^+$ by $\boldsymbol{\Delta}^+_{i,k}$ and $z_{i,k}\cdot\boldsymbol{\zeta}^-$ by $\boldsymbol{\Delta}^-_{i,k}$. In steps, the objective becomes

$$\begin{aligned}\sum_{i=1}^N\sum_{k=1}^K & z_{i,k}(\mathbf{c}_{i,k}^T(\boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-) + d_{i,k}) \\ &= \sum_{i=1}^N\sum_{k=1}^K \mathbf{c}_{i,k}^T\boldsymbol{\zeta}^+ z_{i,k} - \mathbf{c}_{i,k}^T\boldsymbol{\zeta}^- z_{i,k} + d_{i,k}z_{i,k} \\ &= \sum_{i=1}^N\sum_{k=1}^K \mathbf{c}_{i,k}^T\boldsymbol{\Delta}^+_{ik} - \mathbf{c}_{i,k}^T\boldsymbol{\Delta}^-_{ik} + d_{i,k}z_{i,k}.\end{aligned}$$

As for the constraints, one can first make explicit the relation between $\boldsymbol{\Delta}^+$ and $\boldsymbol{\zeta}^+$ by imposing $\sum_{k=1}^K \boldsymbol{\Delta}^+_{i,k} = \sum_{k=1}^K z_{i,k}\boldsymbol{\zeta}^+ = \boldsymbol{\zeta}^+$ and similarly the relation between $\boldsymbol{\Delta}^-$ and $\boldsymbol{\zeta}^-$. One can also add to the model what is implied by every linear constraint $\mathbf{a}^T\boldsymbol{\zeta} \leqslant b$ on the $\boldsymbol{\Delta}^+$ and $\boldsymbol{\Delta}^-$, in other words that $\mathbf{a}^T(\boldsymbol{\Delta}^+_{i,k} - \boldsymbol{\Delta}^-_{i,k}) = \mathbf{a}^T\boldsymbol{\zeta}\cdot z_{i,k} \leqslant b z_{i,k}$. Overall, it is easy to show that the following MILP is equivalent to problem (3):[3]

$$\underset{\mathbf{z},\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-,\boldsymbol{\Delta}^+,\boldsymbol{\Delta}^-}{\text{maximize}} \quad \sum_{i=1}^N\sum_{k=1}^K\{\mathbf{c}_{i,k}^T(\boldsymbol{\Delta}^+_{i,k} - \boldsymbol{\Delta}^-_{i,k}) + d_{i,k}z_{i,k}\} \tag{4a}$$

subject to

$$\mathbf{A}(\boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-) \leqslant \mathbf{b} \tag{4b}$$

$$\boldsymbol{\zeta}^+ \geqslant 0 \text{ and } \boldsymbol{\zeta}^- \geqslant 0 \text{ and } \zeta_j^+ + \zeta_j^- \leqslant 1, \quad \forall j \tag{4c}$$

$$\mathbf{1}^T(\boldsymbol{\zeta}^+ + \boldsymbol{\zeta}^-) = \Gamma \tag{4d}$$

$$\sum_{k=1}^K z_{i,k} = 1, \quad \forall i \tag{4e}$$

$$\sum_{k=1}^K \boldsymbol{\Delta}^+_{i,k} = \boldsymbol{\zeta}^+ \text{ and } \sum_{k=1}^K \boldsymbol{\Delta}^-_{i,k} = \boldsymbol{\zeta}^-, \quad \forall i \tag{4f}$$

$$\mathbf{A}(\boldsymbol{\Delta}^+_{i,k} - \boldsymbol{\Delta}^-_{i,k}) \leqslant \mathbf{b}z_{i,k}, \quad \forall i, \forall k \tag{4g}$$

$$\boldsymbol{\Delta}^+_{i,k} \geqslant 0 \text{ and } \boldsymbol{\Delta}^-_{i,k} \geqslant 0 \text{ and } \boldsymbol{\Delta}^+_{i,k} + \boldsymbol{\Delta}^-_{i,k} \leqslant z_{i,k}, \quad \forall i, \forall k \tag{4h}$$

$$\sum_{j=1}^m (\boldsymbol{\Delta}^+_{i,k})_j + (\boldsymbol{\Delta}^-_{i,k})_j = \Gamma z_{i,k}, \quad \forall i, \forall k \tag{4i}$$

$$z_{i,k} \in \{0, 1\}, \quad \forall i, \forall k. \tag{4j}$$

Based on the observation that the fractional relaxation of problem (4) provides an upper bound for the mixed-integer version, we can already conclude that replacing the adversarial problem in (1) with this fractional relaxation will provide us with a conservative approximation for problem (1).

Proposition 1. *The optimization model*

$$\underset{\mathbf{x}\in\mathcal{X},\boldsymbol{\lambda}^+,\boldsymbol{\lambda}^-,\boldsymbol{\Delta},\nu,\boldsymbol{\gamma},\boldsymbol{\rho},\mathbf{w},\boldsymbol{\psi},\boldsymbol{\theta}}{\text{minimize}} \quad \mathbf{1}^T\boldsymbol{\Delta} + \Gamma\nu + \mathbf{b}^T\boldsymbol{\rho} + \mathbf{1}^T\boldsymbol{\gamma} \tag{5a}$$

subject to

$$\nu \geqslant \sum_{i=1}^N \boldsymbol{\lambda}_i^+ - \mathbf{A}^T\boldsymbol{\rho} - \boldsymbol{\Delta} \tag{5b}$$

$$\nu \geqslant \sum_{i=1}^N \boldsymbol{\lambda}_i^- + \mathbf{A}^T\boldsymbol{\rho} - \boldsymbol{\Delta} \tag{5c}$$

$$\gamma_i \geqslant \mathbf{b}^T\mathbf{w}_{i,k} + \mathbf{1}^T\boldsymbol{\psi}_{i,k} + \Gamma\theta_{i,k} + d_{i,k}(\mathbf{x}), \quad \forall i, \forall k \tag{5d}$$

$$\theta_{i,k} \geqslant -\boldsymbol{\lambda}_i^+ - \mathbf{A}^T\mathbf{w}_{i,k} - \boldsymbol{\psi}_{i,k} + \mathbf{c}_{i,k}(\mathbf{x}), \quad \forall i, \forall k \tag{5e}$$

$$\theta_{i,k} \geqslant -\boldsymbol{\lambda}_i^- + \mathbf{A}^T\mathbf{w}_{i,k} - \boldsymbol{\psi}_{i,k} - \mathbf{c}_{i,k}(\mathbf{x}), \quad \forall i, \forall k \tag{5f}$$

$$\boldsymbol{\rho} \geqslant 0,\ \boldsymbol{\Delta} \geqslant 0,\ \mathbf{w}_{i,k} \geqslant 0,\ \boldsymbol{\psi}_{i,k} \geqslant 0, \quad \forall i, \forall k, \tag{5g}$$

where $\boldsymbol{\rho} \in \mathbb{R}^p$, $\boldsymbol{\Delta} \in \mathbb{R}^m$, $\nu \in \mathbb{R}$, $\boldsymbol{\gamma} \in \mathbb{R}^N$, $\boldsymbol{\lambda}_i^+ \in \mathbb{R}^m$, $\boldsymbol{\lambda}_i^- \in \mathbb{R}^m$, $\mathbf{w}_{i,k} \in \mathbb{R}^p$, $\boldsymbol{\psi}_{i,k} \in \mathbb{R}^m$, and $\theta_{i,k} \in \mathbb{R}$, is a conservative approximation of problem (1). Specifically, let $\hat{\mathbf{x}}^*$ and $\hat{v}^*$ be, respectively, the optimal solution and optimal value of this problem, $\hat{v}^*$ is an optimized upper bound for the best achievable worst-case cost, as measured by problem (1), and $\hat{\mathbf{x}}^*$ is guaranteed to achieve a lower worst-case cost than $\hat{v}^*$.

Proof of Proposition 1. This is simply obtained by constructing the dual of problem (4). Specifically, we identified the dual variables of constraints (4b) to (4i) respectively as $\boldsymbol{\rho}$, $\boldsymbol{\Delta}$, ν, $\boldsymbol{\gamma}$, $(\boldsymbol{\lambda}_i^+, \boldsymbol{\lambda}_i^-)$, $\mathbf{w}_{i,k}$, $\boldsymbol{\psi}_{i,k}$, and $\theta_{i,k}$. Since problem (4) is a linear program for which we can show that there is always a feasible solution, one can confirm that duality is strict. After combining the dual problem to the outer minimization in $\mathbf{x}$ we obtain problem (5). A feasible solution for problem (4) can be identified using $\boldsymbol{\zeta}_0 := \mathbf{0}$. We first assign $\zeta_j^+ := \epsilon_j$ and $\zeta_j^+ := -\epsilon_j$ for some $\boldsymbol{\epsilon} \in \mathbb{R}^m$ chosen so that constraints (4b)–(4d) are satisfied (see Endnote 3). Given any binary assignment for $z_{i,k}$ that satisfies constraint (4e), one can complete the solution by setting $(\boldsymbol{\Delta}_{i,k}^+)_j := \boldsymbol{\zeta}_j^+ z_{i,k}$ and $(\boldsymbol{\Delta}_{i,k}^-)_j := \boldsymbol{\zeta}_j^- z_{i,k}$. □

At this point, one should wonder how good this approximation scheme is and whether it can be compared to other schemes that have been proposed in the literature. Although we will later establish valuable connections to existing approximation methods, we will first shed light on how the quality of our approximation is related to the notion of integrality gap of mixed-integer programs and whether we can bound it.

Definition 1. The *integrality gap* for a class of mixed-integer programs, where the objective function is maximized and the optimal value is known to be positive, is the supremum of the ratio between the optimal value achieved by a fractional solution and the optimal value achieved by an integer one. Specifically, if we seek $\max_{\boldsymbol{\zeta} \in \mathcal{U} \cap \mathcal{I}} f(\boldsymbol{\zeta})$ for instances described by $(\mathcal{U}, \mathcal{I}, f) \in \mathbb{F}$, where $\mathcal{U}$ is polyhedron, $\mathcal{I}$ imposes that a set of terms of $\boldsymbol{\zeta}$ be integer valued, and $\mathbb{F}$ refers to a certain family of problems that is being studied, then

$$\text{integrality gap} = \sup_{(\mathcal{U}, \mathcal{I}, f) \in \mathbb{F}} \frac{f_{\mathcal{U}}^*}{f_{\mathcal{U} \cap \mathcal{I}}^*},$$

where $f_{\mathcal{Z}}^* := \max_{\boldsymbol{\zeta} \in \mathcal{Z}} f(\boldsymbol{\zeta})$.

Proposition 2. *Given that for each i and for all $\mathbf{x} \in \mathcal{X}$, $h_i(\mathbf{x}, \cdot)$ is positive definite on $\mathcal{Z}$, let $\hat{\mathbf{x}}$ and $\hat{v}(\hat{\mathbf{x}})$, respectively, be the optimal solution and optimal value obtained from problem (5), then both the true worst-case value for $\hat{\mathbf{x}}$ and the value $\hat{v}(\hat{\mathbf{x}})$ are less than a factor of γ away from the optimal value of problem (1), where γ is the integrality gap for problem (4).*

Proof of Proposition 2. The integrality gap of problem (4) with positive optimal value is defined as a bound on the largest achievable ratio between the optimal value obtained by a fractional solution and the one obtained by an integer solution for this type of problem instances. Given that the integrality gap for problem (4) is γ, this indicates that the worst-case value achieved by any feasible $\mathbf{x}$ is bounded between $\hat{v}(\mathbf{x})$ and $\gamma \hat{v}(\mathbf{x})$. Hence,

$$\max_{\boldsymbol{\zeta} \in \mathcal{Z}} \sum_{i=1}^{N} h_i(\hat{\mathbf{x}}, \boldsymbol{\zeta}) \leqslant \hat{v}(\hat{\mathbf{x}}) \leqslant \hat{v}(\mathbf{x}^*) \leqslant \gamma \max_{\boldsymbol{\zeta} \in \mathcal{Z}} \sum_{i=1}^{N} h_i(\mathbf{x}^*, \boldsymbol{\zeta}),$$

where $\mathbf{x}^*$ is the optimal solution for problem (1), and where we used the fact that $\hat{\mathbf{x}}$ is the minimizer for $\hat{v}(\cdot)$ over $\mathcal{X}$, and that γ is the integrality gap for $\max_{\boldsymbol{\zeta} \in \mathcal{Z}} \sum_{i=1}^{N} h_i(\mathbf{x}^*, \boldsymbol{\zeta})$. The arguments for linking $\hat{v}(\hat{\mathbf{x}})$ to the true optimal value is exactly the same. □

Our main attempt toward bounding the integrality gap consists of identifying three sets of conditions on problem (4) under which there is no integrality gap for problem (4).

Proposition 3. *For any fixed $\mathbf{x} \in \mathbb{R}^n$, given that $\mathbf{A} := \mathbf{0}$ and $\mathbf{b} := \mathbf{0}$, Problem (4) and its fractional relaxation have the same optimal value under either of the following sets of conditions:*

1. *The budget Γ is equal to one (L_1-norm ball).*
2. *The budget Γ is equal to m (Box uncertainty set) and there exists $\alpha_{i,k}$: $\mathbb{R}^n \to \mathbb{R}$ and β_l: $\mathbb{R}^n \to \mathbb{R}$ such that for every (i, k) pair $\mathbf{c}_{i,k}(\mathbf{x}) = \alpha_{i,k}(\mathbf{x}) \sum_{l<i} (\beta_l(\mathbf{x}) \mathbf{e}_l)$.*
3. *The budget Γ is integer and there exists $\alpha_{i,k}$: $\mathbb{R}^n \to \mathbb{R}$ such that for every (i, k) pair $\mathbf{c}_{i,k}(\mathbf{x}) = \alpha_{i,k}(\mathbf{x}) \mathbf{e}_i$.*

The proof of this proposition is deferred to the appendix (available as supplemental material at http://dx.doi.org/10.1287/opre.2016.1483) of this paper and relies, in the cases of conditions 1 and 3, on verifying total unimodularity of a matrix that defines an associated polytope to confirm that this polytope has integer vertices. Based on the above result, we can right away conclude about an important property of problem (1).

Corollary 1. *Given that the set of subfunctions $\{h_i(\mathbf{x}, \boldsymbol{\zeta})\}_{i=1}^N$ satisfies one of the sets of conditions described in Proposition 3, then problem (5) is equivalent to problem (1).*

Hence, we have in hand a conservative approximation of problem (1) that is known to be exact under a fairly general set of conditions. Reading through the three sets of conditions, we might first recognize that Condition 1 reduces the uncertainty set to $\{\boldsymbol{\zeta} \mid \|\boldsymbol{\zeta}\|_1 \leqslant 1\}$, a case for which a tractable robust counterpart can also be obtained through vertex enumeration. For the second set, the uncertainty set reduces to $\{\boldsymbol{\zeta} \mid \|\boldsymbol{\zeta}\|_\infty \leqslant 1\}$, a set for which the number of vertices is exponential, yet in this case the adversarial problem reduces to a model that was well studied in Bertsimas et al. (2010). One might consider the more important contribution to be related to the third condition, which imposes that each term of the objective function involves a different term of $\boldsymbol{\zeta}$ and that Γ be integer. Actually, the fact that this result requires the integrality of Γ indicates that it cannot be explained through any of the special cases identified in Gorissen and den Hertog (2013) and Ben-Tal et al. (2009a, Chapter 12). Furthermore, it extends in a nontrivial way the result of Denton et al. (2010), where $h_i(\mathbf{x}, \boldsymbol{\zeta}) := \max\{0, \mathbf{c}_\zeta^i(\mathbf{x})\zeta_i + d_\zeta^i(\mathbf{x})\}$ to capture delays that are caused by uncertain duration of surgeries in operating rooms.

Remark 1. Note that while one might be able to design a tractable oracle for providing the value and subgradient in $\mathbf{x}$ of the objective function for problems that satisfy these conditions and thus rely on a cutting plane method to achieve optimality, the proposed linear programming reformulation has better worst-case convergence rate and can easily be modified to handle binary decision variables.

Remark 2. Note that many different MILP formulations could have been used to replace problem (4). Namely, the MILP proposed in Gorissen and den Hertog (2013) takes the following form:

$$\begin{aligned}
\underset{z,\boldsymbol{\zeta},\mathbf{y}}{\text{maximize}}\quad & \sum_{i=1}^{N} y_i\\
\text{subject to}\quad & y_i \leqslant \mathbf{c}_{i,k}^T\boldsymbol{\zeta} + d_{i,k} + M(1-z_{i,k}), \quad \forall i, \forall k,\\
& \mathbf{A}\boldsymbol{\zeta} \leqslant \mathbf{b},\\
& -1 \leqslant \boldsymbol{\zeta} \leqslant 1 \text{ and } \sum_{j=1}^{m} |\zeta_j| \leqslant \Gamma,\\
& \sum_{k=1}^{K} z_{i,k} = 1 \text{ and } z_{i,k} \in \{0,1\}, \quad \forall i, \forall k.
\end{aligned}$$

Our particular choice of formulation is our own best attempt at strategically tightening the integrality gap of the resulting model without paying too much of a price in terms of model size. A side product of our analysis will be to present a perhaps surprising connection to a family of affine approximations used in robust optimization problem where decisions are adjustable.

Remark 3. In recent years, total unimodularity has been somewhat of a fruitful tool for identifying simpler reformulation of risk aware decision problems. In van der Vlerk (2004), the authors show how a two-stage stochastic linear program with binary recourse variables can, in some cases, be reformulated as a two-stage problem with continuous recourse yet under a different probability measure. In Candia-Véjar et al. (2011), it is a maximum regret minimization problem involving binary variables (e.g., assignment problem) that is reformulated as a simple mixed-integer linear program. In the context of robust optimization, Düzgün and Thiele (2010) introduced an extension to the budgeted uncertainty set that allows parameters to take on values in different sets of intervals and show that the convex hull of possible realizations has a tractable representation. In Mak et al. (2015), the authors exploit some "hidden convexity" to identify a tractable reformulation for a distributionally robust appointment scheduling problem with marginal moment information. Unlike our work, which employs the budgeted uncertainty set, the model that is analysed does not capture any correlation between parameters. However, the authors do identify a clever representation of the $\sum_i h_i(\mathbf{x},\boldsymbol{\zeta})$ that allows them to handle additional terms that are nonlinear function of some $\boldsymbol{\zeta}_i$.

4.2. Semidefinite Programming Approximation Model

Although §§5 and 6 will present important applications for which one of the three conditions laid out in Proposition 3 is satisfied, there are still many instances of robust optimization model for which problem (5) is inexact. For those instances, there might be a need to dedicate additional computing resources to get a better approximation. Drawing from the techniques used to solve or bound the value of mixed-integer quadratic programs, we explore the use of semidefinite programming formulations that might help tighten the integrality gap.

Following the ideas presented in Lovász and Schrijver (1991), we first introduce additional quadratic constraints that are redundant for the mixed-integer program:

$$z_{i,k}^2 = z_{i,k} \quad \text{and} \quad z_{i,k} z_{i,k'} = 0, \quad \forall i, \forall k \neq k', \forall i,$$
$$0 \leqslant (\zeta_j^+)^2 \leqslant \zeta_j^+ \quad \text{and} \quad 0 \leqslant (\zeta_j^-)^2 \leqslant \zeta_j^-, \quad \forall j.$$

Our next step is to introduce a set of N matrices $\boldsymbol{\Lambda}_i \in \mathbb{R}^{K\times K}$, with $i = 1, 2, \ldots, N$, and two matrices $\boldsymbol{\Lambda}^+, \boldsymbol{\Lambda}^- \in \mathbb{R}^{m\times m}$ as new decision variables that will help characterize the quadratic interactions in the model through $\boldsymbol{\Lambda}_i = \mathbf{z}_{i,:}\mathbf{z}_{i,:}^T$, $\boldsymbol{\Lambda}^+ = \boldsymbol{\zeta}^+\boldsymbol{\zeta}^{+T}$, and $\boldsymbol{\Lambda}^- = \boldsymbol{\zeta}^-\boldsymbol{\zeta}^{-T}$. Indeed, we would need that the following constraints be satisfied:

$$\begin{bmatrix} \boldsymbol{\Lambda}_i & \operatorname{mat}(\boldsymbol{\Delta}_{i,:}^+) \\ \operatorname{mat}(\boldsymbol{\Delta}_{i,:}^+)^T & \boldsymbol{\Lambda}^+ \end{bmatrix} = \begin{bmatrix} \mathbf{z}_{i,:}^T \\ \boldsymbol{\zeta}^+ \end{bmatrix} \begin{bmatrix} \mathbf{z}_{i,:} & \boldsymbol{\zeta}^{+T} \end{bmatrix}, \quad \forall i, \tag{6a}$$

$$\begin{bmatrix} \boldsymbol{\Lambda}_i & \operatorname{mat}(\boldsymbol{\Delta}_{i,:}^-) \\ \operatorname{mat}(\boldsymbol{\Delta}_{i,:}^-)^T & \boldsymbol{\Lambda}^- \end{bmatrix} = \begin{bmatrix} \mathbf{z}_{i,:}^T \\ \boldsymbol{\zeta}^- \end{bmatrix} \begin{bmatrix} \mathbf{z}_{i,:} & \boldsymbol{\zeta}^{-T} \end{bmatrix}, \quad \forall i, \tag{6b}$$

$$\boldsymbol{\Lambda}_i = \operatorname{diag}(\mathbf{z}_{i,:}^T), \quad \forall i, \tag{6c}$$

$$\Lambda_{j,j}^+ \leqslant \zeta_j^+ \quad \text{and} \quad \Lambda_{j,j}^- \leqslant \zeta_j^-, \quad \forall j, \tag{6d}$$

$$\boldsymbol{\Lambda}^+ \geqslant 0 \quad \text{and} \quad \boldsymbol{\Lambda}^- \geqslant 0, \tag{6e}$$

where $\operatorname{mat}(\boldsymbol{\Delta}_{i,:}^+)$ refers to the K by m matrix composed of the terms of $\boldsymbol{\Delta}_{i,:}^+$ organized such that $(\operatorname{mat}(\boldsymbol{\Delta}_{i,:}^+))_{k,j} = (\boldsymbol{\Delta}_{i,k}^+)_j$. On the other hand, $\operatorname{diag}(\cdot)$ is an operator that creates a diagonal matrix from a vector; e.g., $(\operatorname{diag}(\mathbf{z}_{i,:}^T))_{k,k} = z_{ik}$ while $(\operatorname{diag}(\mathbf{z}_{i,:}^T))_{k,k'} = 0$ for all $k \neq k'$.

Unfortunately, equality constraints (6a) and (6b) are not acceptable in a convex optimization model; hence, we relax them using a matrix inequality:

$$\begin{bmatrix} \boldsymbol{\Lambda}_i & \operatorname{mat}(\boldsymbol{\Delta}_{i,:}^+) \\ \operatorname{mat}(\boldsymbol{\Delta}_{i,:}^+)^T & \boldsymbol{\Lambda}^+ \end{bmatrix} \succeq \begin{bmatrix} \mathbf{z}_{i,:}^T \\ \boldsymbol{\zeta}^+ \end{bmatrix} \begin{bmatrix} \mathbf{z}_{i,:} & \boldsymbol{\zeta}^{+T} \end{bmatrix}, \quad \forall i,$$

$$\begin{bmatrix} \boldsymbol{\Lambda}_i & \operatorname{mat}(\boldsymbol{\Delta}_{i,:}^-) \\ \operatorname{mat}(\boldsymbol{\Delta}_{i,:}^-)^T & \boldsymbol{\Lambda}^- \end{bmatrix} \succeq \begin{bmatrix} \mathbf{z}_{i,:}^T \\ \boldsymbol{\zeta}^- \end{bmatrix} \begin{bmatrix} \mathbf{z}_{i,:} & \boldsymbol{\zeta}^{-T} \end{bmatrix}, \quad \forall i,$$

which can easily be reformulated as linear matrix inequalities and leads to the following mixed-integer semidefinite program:

$$\underset{\mathbf{z},\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-,\boldsymbol{\Delta}^+,\boldsymbol{\Delta}^-,\boldsymbol{\Lambda}_i,\boldsymbol{\Lambda}^+,\boldsymbol{\Lambda}^-}{\text{maximize}} \sum_{i=1}^{N}\sum_{k=1}^{K}\{\mathbf{c}_{i,k}^T(\boldsymbol{\Delta}_{i,k}^+ - \boldsymbol{\Delta}_{i,k}^-) + d_{i,k} z_{i,k}\} \tag{7a}$$

subject to

$$\mathbf{A}(\boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-) \leqslant \mathbf{b}, \tag{7b}$$

$$\boldsymbol{\zeta}^+ \geqslant 0 \text{ and } \boldsymbol{\zeta}^- \geqslant 0 \text{ and } \zeta_j^+ + \zeta_j^- \leqslant 1, \quad \forall j, \tag{7c}$$

$$\mathbf{1}^T(\boldsymbol{\zeta}^+ + \boldsymbol{\zeta}^-) = \Gamma, \tag{7d}$$

$$\sum_{k=1}^{K} z_{i,k} = 1, \quad \forall i, \tag{7e}$$

$$\sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k}^+ = \boldsymbol{\zeta}^+ \text{ and } \sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k}^- = \boldsymbol{\zeta}^-, \quad \forall i, \tag{7f}$$

$$\mathbf{A}(\boldsymbol{\Delta}_{i,k}^+ - \boldsymbol{\Delta}_{i,k}^-) \leqslant \mathbf{b} z_{i,k}, \quad \forall i, \forall k, \tag{7g}$$

$$\boldsymbol{\Delta}_{i,k}^+ \geqslant 0 \text{ and } \boldsymbol{\Delta}_{i,k}^- \geqslant 0 \text{ and } \boldsymbol{\Delta}_{i,k}^+ + \boldsymbol{\Delta}_{i,k}^- \leqslant z_{i,k}, \quad \forall i, \forall k, \tag{7h}$$

$$\sum_{j=1}^{m} (\boldsymbol{\Delta}_{i,k}^+)_j + (\boldsymbol{\Delta}_{i,k}^-)_j = \Gamma z_{i,k}, \quad \forall i, \forall k, \tag{7i}$$

$$\begin{bmatrix} \boldsymbol{\Lambda}_i & \operatorname{mat}(\boldsymbol{\Delta}_{i,:}^+) & \mathbf{z}_{i,:}^T \\ \operatorname{mat}(\boldsymbol{\Delta}_{i,:}^+)^T & \boldsymbol{\Lambda}^+ & \boldsymbol{\zeta}^+ \\ \mathbf{z}_{i,:} & \boldsymbol{\zeta}^{+T} & 1 \end{bmatrix} \succeq 0, \quad \forall i, \tag{7j}$$

$$\begin{bmatrix} \boldsymbol{\Lambda}_i & \operatorname{mat}(\boldsymbol{\Delta}_{i,:}^-) & \mathbf{z}_{i,:}^T \\ \operatorname{mat}(\boldsymbol{\Delta}_{i,:}^-)^T & \boldsymbol{\Lambda}^- & \boldsymbol{\zeta}^- \\ \mathbf{z}_{i,:} & \boldsymbol{\zeta}^{-T} & 1 \end{bmatrix} \succeq 0, \quad \forall i, \tag{7k}$$

$$\boldsymbol{\Lambda}_i = \operatorname{diag}(\mathbf{z}_{i,:}^T), \quad \forall i, \tag{7l}$$

$$\Lambda_{j,j}^+ \leqslant \zeta_j^+ \text{ and } \Lambda_{j,j}^- \leqslant \zeta_j^-, \quad \forall j, \tag{7m}$$

$$\boldsymbol{\Lambda}^+ \geqslant 0 \text{ and } \boldsymbol{\Lambda}^- \geqslant 0, \tag{7n}$$

$$z_{i,k} \in \{0,1\}, \quad \forall i, \forall k. \tag{7o}$$

Since this mixed-integer semidefinite program is equivalent to problem (3) yet contains additional constraints compared to problem (4), we can expect that its fractional relaxation will lead to a tighter conservative approximation for problem (1).

Actually, there are a number of different ways one might choose to tighten the relaxation of problem (4) through the addition of linear cuts or lifting in the space of positive semidefinite cones. Problem (7) is one such example that leads to a somewhat concise semidefinite program. We refer the interested reader to Lasserre (2002) and Ghaddar et al. (2011) for a hierarchy of polynomial size semidefinite programming relaxation of mixed-integer quadratic programs for which the integrality gap is known to converge to 1. Based on Proposition 2, it is therefore theoretically possible to find a semidefinite programming model of polynomial size that will generate a solution within a constant factor of the optimal one. Unfortunately, this might often be of little practical relevance. First, this would require us to assess the integrality gap for each of the models in this hierarchy, which can be hard if not impossible to do. Second, the model that is found to achieve a given factor of optimality might be of a size that cannot be solved in a reasonable amount of time. To help resolve this issue, we actually show that one can potentially confirm after solving a conservative approximation model of smaller size that the approximate solution obtained is indeed optimal for problem (3). As shown in the following proposition, this is done by verifying whether there is an optimal assignment for the associated relaxed adversarial problem that lies in the convex hull of integer solutions. We refer the reader to the e-companion for a complete proof.

Proposition 4. *Given a robust optimization problem of the form*

$$\underset{\mathbf{x}\in\mathcal{X}}{\text{minimize}} \max_{\boldsymbol{\zeta}\in\mathcal{U}\cap\mathcal{I}} h(\mathbf{x}, \boldsymbol{\zeta}),$$

where $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{U} \subset \mathbb{R}^m$ are both bounded convex sets, $\mathcal{I} = \{\boldsymbol{\zeta} \in \mathbb{R}^m \mid \zeta_i$ is integer $\forall i \leqslant q\}$ for some $q \leqslant m$, hence imposing that a set of terms of $\boldsymbol{\zeta}$ be integer valued, and $h(\mathbf{x}, \boldsymbol{\zeta})$ is real valued, convex in $\mathbf{x}$, and linear in $\boldsymbol{\zeta}$. Let $\hat{\mathbf{x}}$ be the solution of the conservative approximation

$$\underset{\mathbf{x}\in\mathcal{X}}{\text{minimize}} \max_{\boldsymbol{\zeta}\in\mathcal{U}} h(\mathbf{x}, \boldsymbol{\zeta}).$$

If there exists a $\hat{\boldsymbol{\zeta}} \in \arg\max_{\boldsymbol{\zeta}\in\mathcal{U}} \min_{\mathbf{x}\in\mathcal{X}} h(\mathbf{x}, \boldsymbol{\zeta})$ that is a member of the convex hull of $\mathcal{U}\cap\mathcal{I}$, denoted as $\mathcal{P}(\mathcal{U}\cap\mathcal{I})$, then $\hat{\mathbf{x}}$ is optimal according to the original robust optimization problem.

When $\mathcal{X}$ does not impose integer constraints, this proposition allows us to imagine a solution scheme in which one generates progressively, based on the current pair $(\hat{\mathbf{x}}, \hat{\boldsymbol{\zeta}})$, new tightening constraints based on cutting planes that separates the current $\hat{\boldsymbol{\zeta}}$ from $\mathcal{P}(\mathcal{U}\cap\mathcal{I})$, adds them to problem (4), and re-solves the associated conservative approximation.[4] This process has reached optimality whenever it is impossible to separate $\hat{\boldsymbol{\zeta}}$ from $\mathcal{P}(\mathcal{U}\cap\mathcal{I})$. Note that if $\mathcal{U}$ only has integer vertices, although one might not be aware of it, then optimality of $\mathbf{x}$ is confirmed instantly when failing to separate the first proposal for $\hat{\boldsymbol{\zeta}}$.

Corollary 2. *If $\mathcal{P}(\mathcal{U}\cap\mathcal{I}) = \mathcal{U}$, then the optimality of $\hat{x} \in \arg\min_{\mathbf{x}\in\mathcal{X}} \max_{\boldsymbol{\zeta}\in\mathcal{U}} h(\mathbf{x}, \boldsymbol{\zeta})$ is necessarily confirmed when verifying that $\hat{\boldsymbol{\zeta}} \in \arg\max_{\boldsymbol{\zeta}\in\mathcal{U}} \min_{\mathbf{x}\in\mathcal{X}} h(\mathbf{x}, \boldsymbol{\zeta})$ is a member of the convex hull of $\mathcal{U}\cap\mathcal{I}$.*

Remark 4. For completeness, we briefly outline an algorithm that can be used to determine whether $\hat{\boldsymbol{\zeta}}$ is in the convex hull of $\mathcal{U}\cap\mathcal{I}$. First, let us recall an equivalent definition for convex hull

$$\mathcal{P}(\mathcal{U}\cap\mathcal{I}) = \left\{ \boldsymbol{\zeta} \in \mathbb{R}^m \,\middle|\, \mathbf{c}^T\boldsymbol{\zeta} \leqslant \sup_{\boldsymbol{\zeta}'\in\mathcal{U}\cap\mathcal{I}} \mathbf{c}^T\boldsymbol{\zeta}', \forall \mathbf{c} \in \mathcal{B}(1) \right\},$$

where $\mathcal{B}(1) = \{\mathbf{c} \in \mathbb{R}^m \mid \|\mathbf{c}\|_2 \leqslant 1\}$. Based on this definition, verifying membership of $\hat{\boldsymbol{\zeta}}$ to the convex hull reduces to validating whether $\min_{\mathbf{c}\in\mathcal{B}(1)} \sup_{\boldsymbol{\zeta}'\in\mathcal{U}\cap\mathcal{I}} \mathbf{c}^T(\boldsymbol{\zeta}' - \hat{\boldsymbol{\zeta}})$ is greater or equal to zero or not (if not the argument that minimizes this expression can be used to generate a cutting plane). Finding the minimum of such an expression can be done using a cutting plane algorithm as long as one has an efficient algorithm to solve $\sup_{\boldsymbol{\zeta}'\in\mathcal{U}\cap\mathcal{I}} \mathbf{c}^T(\boldsymbol{\zeta}' - \hat{\boldsymbol{\zeta}})$ when $\mathbf{c}$ is fixed. In practice, one might use CPLEX to do so.

4.3. Relation to Affinely Adjustable Robust Counterparts

We now provide an explicit connection between our approach and affinely adjustable robust counterparts methods. In fact, we demonstrate below that any model that is obtained by exploiting affine decision rules can also be motivated using our mixed-integer linear programming based approximation scheme. This is interesting since it indicates that our new scheme is somewhat more flexible and, perhaps more importantly, implies the possibility of generalizing the techniques discussed in §4.2 so that they can be used to improve the quality of solutions obtained from any AARC approximation of robust multistage problems.

PROPOSITION 5. *Given an adversarial problem of the type*

$$\underset{\boldsymbol{\zeta}\in\mathcal{A}}{\text{maximize}} \sum_{i=1}^{N} \max_k \{\mathbf{c}_{i,k}^T\boldsymbol{\zeta} + d_{i,k}\},$$

where $\mathcal{A} = \{\boldsymbol{\zeta} \mid \mathbf{A}\boldsymbol{\zeta} \leqslant \mathbf{b}\}$ *is a bounded polyhedron, the optimal value of its affinely adjustable robust counterpart*

$$\underset{\boldsymbol{\lambda},\boldsymbol{\gamma}}{\text{minimize}} \ \max_{\boldsymbol{\zeta}\in\mathcal{A}} \sum_{i=1}^{N} \boldsymbol{\lambda}_i^T\boldsymbol{\zeta} + \gamma_i \tag{8a}$$

$$\text{subject to } \boldsymbol{\lambda}_i^T\boldsymbol{\zeta} + \gamma_i \geqslant \mathbf{c}_{i,k}^T\boldsymbol{\zeta} + d_{i,k}, \quad \forall i, \forall k, \forall \boldsymbol{\zeta} \in \mathcal{A}, \tag{8b}$$

is equal to the optimal value of the fractional relaxation of the mixed-integer linear programming problem

$$\underset{\mathbf{z},\boldsymbol{\zeta},\boldsymbol{\Delta}}{\text{maximize}} \ \sum_{i=1}^{N}\sum_{k=1}^{K} \{\mathbf{c}_{i,k}^T\boldsymbol{\Delta}_{i,k} + d_{i,k}z_{i,k}\} \tag{9a}$$

$$\text{subject to } \mathbf{A}\boldsymbol{\zeta} \leqslant \mathbf{b}, \tag{9b}$$

$$\sum_{k=1}^{K} z_{i,k} = 1, \quad \forall i, \tag{9c}$$

$$\sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k} = \boldsymbol{\zeta}, \quad \forall i, \tag{9d}$$

$$\mathbf{A}\boldsymbol{\Delta}_{i,k} \leqslant \mathbf{b}z_{i,k}, \quad \forall i,k, \tag{9e}$$

$$z_{i,k} \in \{0,1\}, \quad \forall i,k. \tag{9f}$$

PROOF OF PROPOSITION 5. The optimal value of the fractional relaxation of problem (9) can be presented in the form

$$\underset{\boldsymbol{\zeta}\in\mathcal{A}}{\text{maximize}} \ \max_{\mathbf{z},\boldsymbol{\Delta}} \sum_{i=1}^{N}\sum_{k=1}^{K} \{\mathbf{c}_{i,k}^T\boldsymbol{\Delta}_{i,k} + d_{i,k}z_{i,k}\} \tag{10a}$$

$$\text{subject to } \sum_{k=1}^{K} z_{i,k} = 1, \quad \forall i, \tag{10b}$$

$$\sum_{k=1}^{K} \boldsymbol{\Delta}_{i,k} = \boldsymbol{\zeta}, \quad \forall i, \tag{10c}$$

$$\mathbf{A}\boldsymbol{\Delta}_{i,k} \leqslant \mathbf{b}z_{i,k}, \quad \forall i,k, \tag{10d}$$

$$z_{i,k} \geqslant 0 \quad \forall i,k. \tag{10e}$$

For any fixed $\boldsymbol{\zeta} \in \mathcal{A}$, since the inner problem is linear and has a feasible solution, strict duality applies and its optimal value is equal to the optimal value of its dual problem. Hence, this inner problem can be reformulated as

$$\underset{\boldsymbol{\gamma},\boldsymbol{\lambda},\boldsymbol{\psi}}{\text{minimize}} \ \sum_{i=1}^{N} \{\gamma_i + \boldsymbol{\lambda}_i^T\boldsymbol{\zeta}\} \tag{11a}$$

subject to

$$\gamma_i - \mathbf{b}^T\boldsymbol{\psi}_{i,k} \geqslant d_{i,k}, \quad \forall i,k, \tag{11b}$$

$$\mathbf{A}^T\boldsymbol{\psi}_{i,k} + \boldsymbol{\lambda}_i = \mathbf{c}_{i,k}, \quad \forall i,k, \tag{11c}$$

$$\boldsymbol{\psi}_{i,k} \geqslant 0, \quad \forall i,k, \tag{11d}$$

where $\gamma_i \in \mathbb{R}$, $\boldsymbol{\lambda}_i \in \mathbb{R}^m$, and $\boldsymbol{\psi}_{i,k} \in \mathbb{R}^p$ are, respectively, the dual variables associated to constraints (10b), (10c), and (10d). Since $\mathcal{A}$ is bounded and convex and the objective function (11a) is bilinear in $\boldsymbol{\zeta}$ and the set of variable $\{\boldsymbol{\gamma}, \boldsymbol{\lambda}, \boldsymbol{\psi}\}$, Sion's minimax theorem guarantees that the maximin formulation is equal to the minimax one, hence the optimal value of problem (10) is equal to the optimal value of

$$\underset{\boldsymbol{\gamma},\boldsymbol{\lambda},\boldsymbol{\psi}}{\text{minimize}} \ \max_{\boldsymbol{\zeta}\in\mathcal{A}} \sum_{i=1}^{N} \{\gamma_i + \boldsymbol{\lambda}_i^T\boldsymbol{\zeta}\} \tag{12a}$$

$$\text{subject to } \gamma_i - \mathbf{b}^T\boldsymbol{\psi}_{i,k} \geqslant d_{i,k}, \quad \forall i,k, \tag{12b}$$

$$\mathbf{A}^T\boldsymbol{\psi}_{i,k} + \boldsymbol{\lambda}_i = \mathbf{c}_{i,k}, \quad \forall i,k. \tag{12c}$$

It can be shown that constraints (12b) and (12c) are equivalent to robust constraint (8b). Specifically, for all index pair (i,k) the right-hand side of the robust constraint of (8b) can be formulated as

$$\underset{\boldsymbol{\zeta}}{\text{maximize}} \ \{\mathbf{c}_{i,k}^T\boldsymbol{\zeta} - \boldsymbol{\lambda}_i^T\boldsymbol{\zeta}\} \tag{13a}$$

$$\text{subject to } \mathbf{A}\boldsymbol{\zeta} \leqslant \mathbf{b}. \tag{13b}$$

Since strict duality applies once again, problem (13) gives the same optimal value as

$$\underset{\boldsymbol{\psi}_{i,k}}{\text{minimize}} \ b^T\boldsymbol{\psi}_{i,k}$$

$$\text{subject to } \mathbf{A}^T\boldsymbol{\psi}_{i,k} = \mathbf{c}_{i,k} - \boldsymbol{\lambda}_i, \quad \forall i,k,$$

$$\boldsymbol{\psi}_{i,k} \geqslant 0, \quad \forall i,k,$$

where each $\boldsymbol{\psi}_{i,k} \in \mathbb{R}^p$ contains the dual variables associated to constraint (13b). Therefore, it is clear that robust constraint (8b) can be reformulated as

$$\gamma_i - d_{i,k} \geqslant \mathbf{b}^T\boldsymbol{\psi}_{i,k}, \quad \forall i,k,$$

$$\mathbf{A}^T\boldsymbol{\psi}_{i,k} = \mathbf{c}_{i,k} - \boldsymbol{\lambda}_i, \quad \forall i,k.$$

We conclude that formulation (9) gives the same optimal value as problem (8). $\square$

This result is interesting as it establishes that any robust counterpart that is obtained using an affine decision rule scheme can be thought of in terms of replacing the adversarial problem with the fractional relaxation of an equivalent MILP. In particular, one can easily verify that using affine decision rules under the particular lifting $\boldsymbol{\zeta} := \boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-$, with $\boldsymbol{\zeta}^+ \geqslant 0$ and $\boldsymbol{\zeta}^- \geqslant 0$, is equivalent to problem (5).

COROLLARY 3. *The fractional relaxation of problem (4) is equivalent to the affinely adjusted approximation of problem (3) applied to the lifted set of perturbation $\boldsymbol{\zeta} = \boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-$, where $\boldsymbol{\zeta}^+$ and $\boldsymbol{\zeta}^-$ are the positive and negative parts of $\boldsymbol{\zeta}$. Specifically, it achieves the same optimal value as the problem*

$$\underset{\boldsymbol{\lambda}^+,\,\boldsymbol{\lambda}^-,\,\boldsymbol{\gamma}}{\text{minimize}} \quad \max_{(\boldsymbol{\zeta}^+,\,\boldsymbol{\zeta}^-)\in\mathcal{Z}'} \sum_{i=1}^{N}\left\{{\boldsymbol{\lambda}_i^+}^T\boldsymbol{\zeta}^+ + {\boldsymbol{\lambda}_i^-}^T\boldsymbol{\zeta}^- + \gamma_i\right\}$$

$$\text{subject to} \quad {\boldsymbol{\lambda}_i^+}^T\boldsymbol{\zeta}^+ + {\boldsymbol{\lambda}_i^-}^T\boldsymbol{\zeta}^- + \gamma_i \geqslant \mathbf{c}_{i,k}^T(\boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-) + d_{i,k}, \quad \forall\, i,\, k,\ \forall\,(\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-)\in\mathcal{Z}',$$

where

$$\mathcal{Z}' = \left\{(\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-)\ \middle|\ \begin{array}{l} \boldsymbol{\zeta}^+ \geqslant 0, \quad \boldsymbol{\zeta}^- \geqslant 0 \\ \zeta_j^+ + \zeta_j^- \leqslant 1 \quad \forall\, j \\ \sum_j \zeta_j^+ + \zeta_j^- = \Gamma \\ \mathbf{A}(\boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-) \leqslant \mathbf{b} \end{array}\right\}.$$

While there is a lot of empirical evidence supporting the strength of affinely adjusted approximation schemes (see, for instance, Ben-Tal et al. 2005), very little is actually known theoretically about the quality of the approximations that are obtained with these methods, either in terms of optimal value or optimal solution. To the best of our knowledge, the authors of Iancu et al. (2013) are the ones that have identified to this date the most general class of problems for which the approximation was exact. We refer the readers in particular to Theorem 3 in their article, which could potentially provide an alternative method for deriving our Proposition 3 given the connection established in Corollary 3. Note, however, that while Iancu et al. do identify problem instances where conditions (P1) and (P2) of Theorem 3 in their paper are satisfied under box uncertainty or the simplex (given the implicit sublattice structure of these uncertainty sets), they left open the question of identifying such instances for a general budgeted uncertainty set with integer budget.

Overall, our new interpretation of AARC methods is particularly interesting as it states that asking whether AARC methods are exact is equivalent to asking whether an associated MILP (problem (4) for example) has an integrality gap or not. Given the extensive efforts that have been dedicated in the last few decades both to developing approximation methods for MILP that are based on fractional relaxation schemes and to measuring the quality of these bounds, we have good hope to find innovative ways of improving the performance of these AARC models.

We conclude this section with a result that establishes the connection between the semidefinite programming based conservative approximation presented in §4.2 and the theory of AARCs. The proof of this final connection can be found in the e-companion.

PROPOSITION 6. *The optimal value of the fractional relaxation of problem (7) is equal to the optimal value of the affinely adjustable robust counterpart of*

$$\underset{\substack{\{\mathbf{v}_i\}_{i=1}^N,\mathbf{w}^+,\mathbf{w}^-,\\ \{\mathbf{Q}_i^+,\mathbf{V}_i^+,\mathbf{q}_i^+,\mathbf{S}_i^+,\mathbf{p}_i^+,r_i^+\}_{i=1}^N\\ \{\mathbf{Q}_i^-,\mathbf{V}_i^-,\mathbf{q}_i^-,\mathbf{S}_i^-,\mathbf{p}_i^-,r_i^-\}_{i=1}^N}}{\text{minimize}} \quad \max_{(\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-)\in\mathcal{Z}'} \Big\{(\mathbf{w}^+)^T\boldsymbol{\zeta}^+ + (\mathbf{w}^-)^T\boldsymbol{\zeta}^- + \sum_{i=1}^N 2{\mathbf{p}_i^+}^T\boldsymbol{\zeta}^+ + 2{\mathbf{p}_i^-}^T\boldsymbol{\zeta}^- + \max_k\big\{\mathbf{c}_{i,k}^T(\boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-) + d_{i,k} + 2(\mathbf{V}_i^+)_{k,:}\boldsymbol{\zeta}^+ + 2(\mathbf{V}_i^-)_{k,:}\boldsymbol{\zeta}^- + (\mathbf{v}_i)_k\big\}\Big\} \tag{14a}$$

subject to

$$(\mathbf{v}_i)_k = (\mathbf{Q}_i^+ + \mathbf{Q}_i^-)_{k,k} + 2(\mathbf{q}_i^+ + \mathbf{q}_i^-)_k + r_i^+ + r_i^-, \quad \forall\, i,\, k, \tag{14b}$$

$$\begin{bmatrix} \mathbf{Q}_i^+ & \mathbf{V}_i^+ & \mathbf{q}_i^+ \\ {\mathbf{V}_i^+}^T & \mathbf{S}_i^+ & \mathbf{p}_i^+ \\ {\mathbf{q}_i^+}^T & {\mathbf{p}_i^+}^T & r_i^+ \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \mathbf{Q}_i^- & \mathbf{V}_i^- & \mathbf{q}_i^- \\ {\mathbf{V}_i^-}^T & \mathbf{S}_i^- & \mathbf{p}_i^- \\ {\mathbf{q}_i^-}^T & {\mathbf{p}_i^-}^T & r_i^- \end{bmatrix} \succeq 0, \quad \forall\, i, \tag{14c}$$

$$\sum_{i=1}^N \mathbf{S}_i^+ \leqslant \operatorname{diag}(\mathbf{w}^+), \quad \sum_{i=1}^N \mathbf{S}_i^- \leqslant \operatorname{diag}(\mathbf{w}^-), \quad \forall\, i, \tag{14d}$$

$$\mathbf{w}^+ \geqslant 0 \ \mathbf{w}^- \geqslant 0, \tag{14e}$$

where $\mathbf{w}^+\in\mathbb{R}^m$, $\mathbf{w}^-\in\mathbb{R}^m$, *while for each* i, $\mathbf{v}_i\in\mathbb{R}^K$, $\mathbf{Q}_i^+\in\mathbb{R}^{K\times K}$, $\mathbf{Q}_i^-\in\mathbb{R}^{K\times K}$, $\mathbf{V}_i^+\in\mathbb{R}^{K\times m}$, $\mathbf{V}_i^-\in\mathbb{R}^{K\times m}$, $\mathbf{q}_i^+\in\mathbb{R}^K$, $\mathbf{q}_i^-\in\mathbb{R}^K$, $\mathbf{S}_i^+\in\mathbb{R}^{m\times m}$, $\mathbf{S}_i^-\in\mathbb{R}^{m\times m}$, $\mathbf{p}_i^+\in\mathbb{R}^m$, $\mathbf{p}_i^-\in\mathbb{R}^m$, $r_i^+\in\mathbb{R}$, $r_i^-\in\mathbb{R}$, *and finally where*

$$\mathcal{Z}' = \left\{(\boldsymbol{\zeta}^+,\boldsymbol{\zeta}^-)\ \middle|\ \begin{array}{l} \boldsymbol{\zeta}^+ \geqslant 0, \quad \boldsymbol{\zeta}^- \geqslant 0 \\ \zeta_j^+ + \zeta_j^- \leqslant 1, \quad \forall\, j \\ \sum_j \zeta_j^+ + \zeta_j^- = \Gamma, \\ \mathbf{A}(\boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-) \leqslant \mathbf{b} \end{array}\right\}.$$

One might actually notice that in problem (14), when all the variables that are minimized over are set to zero, which is a feasible assignment, the problem reduces to problem (3) with the lifted set of perturbation $\boldsymbol{\zeta} = \boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-$. Intuitively, the SDP model is able to obtain a tighter bound by adding some affine perturbations that have a positive net effect on the evaluation of the objective function yet might allow to reach a lower amount when affine decision rules are introduced. In particular, in the case where $\boldsymbol{\zeta}^- := \mathbf{0}$, which we assume for simplicity of exposure, and $0 \leqslant \boldsymbol{\zeta}^+ \leqslant 1$, when constraints (14b) to (14e) are satisfied, then the objective function can be shown to satisfy

$$(\mathbf{w}^+)^T\boldsymbol{\zeta}^+ + \sum_{i=1}^N \Big\{2{\mathbf{p}_i^+}^T\boldsymbol{\zeta}^+ + \max_k\big\{\mathbf{c}_{i,k}^T\boldsymbol{\zeta}^+ + d_{i,k} + 2(\mathbf{V}_i^+)_{k,:}\boldsymbol{\zeta}^+ + (\mathbf{Q}_i^+)_{kk} + 2(\mathbf{q}_i^+)_k + r_i^+\big\}\Big\}$$

$$\geqslant (\mathbf{w}^+)^T(\boldsymbol{\zeta}^+)^2 + \sum_{i=1}^N \Big\{2{\mathbf{p}_i^+}^T\boldsymbol{\zeta}^+ + \max_k\big\{\mathbf{c}_{i,k}^T\boldsymbol{\zeta}^+ + d_{i,k} + 2(\mathbf{V}_i^+)_{k,:}\boldsymbol{\zeta}^+ + (\mathbf{Q}_i^+)_{kk} + 2(\mathbf{q}_i^+)_k + r_i^+\big\}\Big\}$$

$$= (\mathbf{w}^+)^T(\boldsymbol{\zeta}^+)^2 + \sum_{i=1}^{N}\Bigg\{2\mathbf{p}_i^{+T}\boldsymbol{\zeta}^+ + \max_{\substack{z_k\in\{0,1\}\\ \sum_{k=1}^K z_k=1}} \sum_{k=1}^{K} z_k\{\mathbf{c}_{i,k}^T\boldsymbol{\zeta}^+ + d_{i,k} + 2(\mathbf{V}_i^+)_{k,:}\boldsymbol{\zeta}^+ + (\mathbf{Q}_i^+)_{kk} + 2(\mathbf{q}_i^+)_k + r_i^+\}\Bigg\}$$

$$\geqslant \sum_{i=1}^{N} \max_{\substack{z_k\in\{0,1\}\\ \sum_{k=1}^K z_k=1}} \Bigg\{\boldsymbol{\zeta}^{+T}\mathbf{S}_i\boldsymbol{\zeta}^+ + 2\mathbf{p}_i^{+T}\boldsymbol{\zeta}^+ + \sum_{k=1}^{K} z_k\big(\mathbf{c}_{i,k}^T\boldsymbol{\zeta}^+ + d_{i,k} + 2(\mathbf{V}_i^+)_{k,:}\boldsymbol{\zeta}^+ + (\mathbf{Q}_i^+)_{kk} + 2(\mathbf{q}_i^+)_k + r_i^+\big)\Bigg\}$$

$$\geqslant \sum_{i=1}^{N} \max_{\substack{z_k\in\{0,1\}\\ \sum_{k=1}^K z_k=1}} \sum_{k=1}^{K} z_k(\mathbf{c}_{i,k}^T\boldsymbol{\zeta}^+ + d_{i,k}) = \sum_{i=1}^{N}\max_k\{\mathbf{c}_{i,k}^T\boldsymbol{\zeta}^+ + d_{i,k}\},$$

where we used the fact that $\boldsymbol{\zeta}^+ \in [0,1]^m$ to get the first inequality, and constraint (14d) to get the second. We finally used the fact that $(\mathbf{Q}_i^+)_{kk}z_k = \sum_{k'}(\mathbf{Q}_i^+)_{kk}z_k z_{k'}$ over the feasible region that is considered and used constraint (14c), which implies that

$$\boldsymbol{\zeta}^{+T}\mathbf{S}_i\boldsymbol{\zeta}^+ + 2\mathbf{p}_i^{+T}\boldsymbol{\zeta}^+ + 2\mathbf{z}^T(\mathbf{V}_i^+)_{k,:}\boldsymbol{\zeta}^+ + \mathbf{z}^T\mathbf{Q}_i^+\mathbf{z}^T + 2\mathbf{z}^T\mathbf{q}_i^+ + r_i^+ \geqslant 0.$$

A similar argument can be made when we involve the $\boldsymbol{\zeta}^- > 0$. Hence, problem (14) necessarily provides a tight upper bound to problem (3).

Overall, the connection established in Proposition 6 indicates that the scheme we adopt in this paper allows to identify tractable conservative approximations that provide tighter bounds than the well-known applications of affine decision rules. Indeed, while schemes that are based on quadratic decision rules can, in some cases, lead to SDP approximation models, such adjustment functions lead, in general, to optimization problems that are computationally intractable when the uncertainty set is polyhedral (see Ben-Tal et al. 2009a, p. 372). Although one might be able to further approximate those models by applying sum-of-squares techniques as were proposed in Bertsimas et al. (2011b), such an approach leads to SDP models of much larger size than the model presented here.

4.4. Empirical Evaluation of Integrality Gap

We briefly present a set of empirical experiments that illustrates the trade-off that needs to be made between computational effort and quality of the upper bound obtained for problem (3) with $\mathbf{A} := \mathbf{0}$ and $\mathbf{b} := \mathbf{0}$ using three different fractional relaxation schemes. The first bound is obtained by applying affine decision rules directly on $\boldsymbol{\zeta}$; this method will be referred as AARC. We also compare the two improved bounds based on linear program (4), and semidefinite program (7). These methods compete on a set of 100 randomly generated instances of problem (3), which we solved exactly using CPLEX. Each problem instance is generated by sampling each parameter of the objective function uniformly between −1 and 1, and then ensuring that the optimal value is positive by adding a constant term that makes $\sum_i \max_k\{\mathbf{c}_{i,k}^T\mathbf{0} + d_{i,k}\} = 0$; furthermore, a random integer budget of Γ is generated uniformly between 1 to N. Based on the results presented in Table 1, we see that the quality of each bound degrades as N increases, yet an approach based on semidefinite programming will achieve significant improvement in tightness. On the other hand, there is a heavier computational price to pay for the semidefinite programming model. It is also observed that LP (4) provides better results than AARC. Note that all linear programming models were solved using CPLEX 12.4 and the SDP was solved using DSDP 5.8 (Benson et al. 2000).

Table 1. Empirical evaluation of integrality gap and resolution time for a set of randomly generated convex maximization problems of form (3).

Size		AARC	LP (4)	SDP (7)
$N = 8$	CPU time (sec)	0.062	0.064	1.564
	Gap = 1 instances (%)	14	29	59
	Largest gap	1.70	1.58	1.09
	Average gap	1.26	1.14	1.01
$N = 16$	CPU time (sec)	0.17	0.18	27.89
	Gap = 1 instances (%)	3	6	11
	Largest gap	2.32	2.12	1.14
	Average gap	1.82	1.49	1.05
$N = 32$	CPU time	10 sec	10 sec	28.9 min
	Gap = 1 instances (%)	0	0	0
	Largest gap	2.94	2.80	1.22
	Average gap	2.61	1.96	1.10
$N = 64$	CPU time	34 sec	70 sec	19 h
	Min improvement (%)	—	0	43
	Max improvement (%)	—	54	72
	Avg. improvement (%)	—	26	68

Notes. All problems have size $N = m = n$ and $K = 2$. Note that in the case of $N = 64$, it took longer than a full day to solve the MILP with CPLEX. Therefore, we choose to report the minimum, maximum, and average relative improvement (over 100 instances) of the bounds obtained by each model compared to the bound obtained with AARC. Note that an integrality gap of one is optimal.

5. Robust Multi-Item Newsvendor Problem

We now pay closer attention to the multi-item newsvendor problem as described in a general form through the following model:

$$\underset{\mathbf{x}\in\mathscr{X},\, 0\leqslant\mathbf{y}\leqslant\min(\mathbf{x},\mathbf{w})}{\text{maximize}} \sum_{i=1}^{m}\{v_i y_i - c_i x_i + g_i(x_i - y_i) - b_i(w_i - y_i)\},$$

where x_i represents the amount of item i ordered, w_i the demand for this item while $\mathscr{X}$ captures the set of feasible orders and y_i is the (second-stage) amount sold once the demand is known. We also denote the following terms: $c_i \in \mathbb{R}^n$ and $v_i \in \mathbb{R}^n$ are, respectively, the per unit ordering cost and retail prices, $b_i(\cdot)$ is a piecewise linear convex increasing stock-out cost, and $g_i(\cdot)$ is a piecewise linear decreasing concave salvage prices. We assume that for each item $\partial g_i(z)/\partial z < c_i < v_i$ whenever the derivative exists, i.e., that the unit ordering cost is always larger than the marginal salvage price and always lower than the retail price. One can, therefore, not make profits out of salvaging his products. Since $y_i^* = \min(x_i, w_i)$, the two-stage model is equivalent to:

$$\underset{x\in\mathscr{X}}{\text{minimize}} \sum_{i=1}^{m} \{c_i x_i - v_i \min(x_i, w_i) - g_i(x_i - \min(x_i, w_i)) + b_i(w_i - \min(x_i, w_i))\},$$

which is presented in terms of minimizing negative profits to be coherent with problem (1).

Further manipulations of the model will lead to a form that makes the connection with problem (1) explicit:

$$\begin{aligned}
&(c_i x_i - v_i \min(x_i, w_i)) - g_i(x_i - \min(x_i, w_i)) \\
&\quad + b_i(w_i - \min(x_i, w_i)) \\
&= (c_i - v_i)w_i + c_i(x_i - w_i)^+ - (c_i - v_i)(w_i - x_i)^+ \\
&\quad - g_i((x_i - w_i)^+) + b_i((w_i - x_i)^+) \\
&= (c_i - v_i)w_i + (c_i(x_i - w_i) - g_i(x_i - w_i))^+ \\
&\quad + ((v_i - c_i)(w_i - x_i) + b_i(w_i - x_i))^+ \\
&= \max(c_i x_i - v_i w_i - g_i(x_i - w_i), (c_i - v_i)x_i + b_i(w_i - x_i)) \\
&= \max\Big(\max_{k\in\{1,2,\ldots,K^g\}} c_i x_i - v_i w_i - \alpha^g_{i,k}(x_i - w_i) - \beta^g_{i,k}, \\
&\qquad \max_{k\in\{1,2,\ldots,K^b\}} (c_i - v_i)x_i + \alpha^b_{i,k}(w_i - x_i) + \beta^b_{i,k}\Big) \\
&= \max_k \alpha^x_{i,k} x_i + \alpha^w_{i,k} w_i + \beta_{i,k},
\end{aligned}$$

where we exploited the piecewise linear concave and convex structures of $g_i(y) = \min_{k\in\{1,2,\ldots,K^g\}}\{\alpha^g_{i,k} y + \beta^g_{i,k}\}$ and $b_i(y) = \max_{k\in\{1,2,\ldots,K^b\}}\{\alpha^b_{i,k} y + \beta^b_{i,k}\}$, respectively, and later combined the indexes of the two layers of maximum operators so that

$$\alpha^x_{i,k} = \begin{cases} c_i - \alpha^g_{i,k} & \text{if } k \leqslant K^g, \\ c_i - v_i - \alpha^b_{i,k-K^g} & \text{if } k > K^g, \end{cases}$$

$$\alpha^w_{i,k} = \begin{cases} \alpha^g_{i,k} - v_i & \text{if } k \leqslant K^g, \\ \alpha^b_{i,k-K^g} & \text{if } k > K^g, \end{cases}$$

$$\beta_{i,k} = \begin{cases} -\beta^g_{i,k} & \text{if } k \leqslant K^g, \\ \beta^b_{i,k-K^g} & \text{if } k > K^g. \end{cases}$$

When considering robustness in the multi-item newsvendor problem, we introduce a budgeted uncertainty set for the demand vector. Specifically, we assume that the nominal demand vector takes the form $\bar{w}$, that each term is known to lie in the interval $w_i \in [\bar{w}_i - \hat{w}_i, \bar{w}_i + \hat{w}_i]$ and that we do not expect the total perturbation to exceed a budget of Γ. Hence, the robust model takes the form

$$\underset{\mathbf{x}\in\mathscr{X}}{\text{minimize}} \max_{\boldsymbol{\zeta}\in\mathscr{Z}(\Gamma)} \sum_{i=1}^{m} \max_k \{\alpha^x_{i,k} x_i + \alpha^w_{i,k}(\bar{w}_i + \hat{w}_i \zeta_i) + \beta_{i,k}\}, \quad (15)$$

where $\mathscr{Z}(\Gamma) := \{\boldsymbol{\zeta} \in \mathbb{R}^m \mid \|\boldsymbol{\zeta}\|_\infty \leqslant 1,\ \|\boldsymbol{\zeta}\|_1 \leqslant \Gamma\}$. One might easily recognize in this form that each term of the objective function depends on a different component of $\boldsymbol{\zeta}$. Therefore, it is clear, based on Corollary 1, that using the robust counterpart presented in problem (5) will provide an exact solution. The proof of the following corollary, presented in the e-companion, serves to justify how to obtain a more compact robust counterpart.

Corollary 4. *Given that Γ is a strictly positive integer, then the robust multi-item newsvendor problem (15) is equivalent to the following linear program*:

$$\begin{aligned}
\underset{\mathbf{x}\in\mathscr{X},\,\nu,\,\boldsymbol{\gamma},\,\boldsymbol{\psi}}{\text{minimize}} \quad & \Gamma\nu + \mathbf{1}^T\boldsymbol{\gamma} \\
\text{subject to} \quad & \gamma_i \geqslant \psi_{i,k} + \alpha^x_{i,k} x_i + \alpha^w_{i,k}\bar{w}_i + \beta_{i,k}, \quad \forall i, \forall k, \\
& \psi_{i,k} + \nu \geqslant \alpha^w_{i,k}\hat{w}_i, \quad \forall i, \forall k, \\
& \psi_{i,k} + \nu \geqslant -\alpha^w_{i,k}\hat{w}_i, \quad \forall i, \forall k, \\
& \psi_{i,k} \geqslant 0, \quad \forall i, \forall k,
\end{aligned}$$

where $\nu \in \mathbb{R}$, $\boldsymbol{\gamma} \in \mathbb{R}^m$, and $\psi_{i,k} \in \mathbb{R}$.

Given that the distribution-free version of the newsvendor problem has received so much attention over the last 50 years (see Scarf 1958, Gallego and Moon 1993, Moon and Silver 2000, Wang et al. 2010, Hanasusanto et al. 2014, Wiesemann et al. 2014 for some examples), we provide below an exact reformulation for a model that seeks the distributionally robust newspaper quantities when the only information that is available about the distribution includes that the support is $\mathscr{Z}(\Gamma)$, the mean vector is μ, and a list of lower bounds on first order partial moments. To the best of our knowledge, this appears to be the first tractable exact reformulation for such a problem when there exists information about how the demands for different items behave jointly.

Proposition 7. *The distributionally robust optimization model*

$$\underset{\mathbf{x}\in\mathscr{X}}{\text{minimize}} \max_{F\in\mathscr{D}} \mathbb{E}_F\left[\sum_{i=1}^{m} \max_k \{\alpha^x_{i,k} x_i + \alpha^w_{i,k}(\bar{w}_i + \hat{w}_i\zeta_i) + \beta_{i,k}\}\right], \quad (16)$$

where

$$\mathscr{D} = \left\{ F \in \mathscr{M} \;\middle|\; \begin{array}{l} \mathbb{P}_F(\boldsymbol{\zeta} \in \mathscr{Z}(\Gamma)) = 1 \\ \mathbb{E}_F[\boldsymbol{\zeta}] = \boldsymbol{\mu} \\ \mathbb{E}_F[(\boldsymbol{\zeta} - \boldsymbol{\mu})^+] \geqslant \mathbf{r}^+ \\ \mathbb{E}_F[(\boldsymbol{\mu} - \boldsymbol{\zeta})^+] \geqslant \mathbf{r}^- \end{array} \right\},$$

is equivalent to the following linear program

$$\underset{\mathbf{x}\in\mathscr{X},\,t,\,\mathbf{q},\,\boldsymbol{\lambda}^+,\,\boldsymbol{\lambda}^-,\,\nu,\,\boldsymbol{\gamma},\,\boldsymbol{\psi}^+,\,\boldsymbol{\psi}^-}{\text{minimize}} \quad t + \boldsymbol{\mu}^T\mathbf{q} - (\mathbf{r}^+)^T\boldsymbol{\lambda}^+ - (\mathbf{r}^-)^T\boldsymbol{\lambda}^- \quad (17a)$$

subject to

$$t \geqslant \Gamma\nu + \mathbf{1}^T\boldsymbol{\gamma}, \quad (17b)$$

$$\gamma_i \geqslant \psi^+_{i,k} + \alpha^x_{i,k} x_i + \alpha^w_{i,k}\bar{w}_i + \beta_{i,k} - \lambda^+_i\mu_i, \quad \forall i, \forall k, \quad (17c)$$

$$\gamma_i \geqslant \psi^-_{i,k} + \alpha^x_{i,k} x_i + \alpha^w_{i,k}\bar{w}_i + \beta_{i,k} + \lambda^-_i \mu_i, \quad \forall i, \forall k, \tag{17d}$$

$$\psi^+_{i,k} + \nu \geqslant \alpha^w_{i,k}\hat{w}_i - q_i + \lambda^+_i, \quad \forall i, \forall k, \tag{17e}$$

$$\psi^+_{i,k} + \nu \geqslant -\alpha^w_{i,k}\hat{w}_i + q_i - \lambda^+_i, \quad \forall i, \forall k, \tag{17f}$$

$$\psi^-_{i,k} + \nu \geqslant \alpha^w_{i,k}\hat{w}_i - q_i - \lambda^-_i, \quad \forall i, \forall k, \tag{17g}$$

$$\psi^-_{i,k} + \nu \geqslant -\alpha^w_{i,k}\hat{w}_i + q_i + \lambda^+_i, \quad \forall i, \forall k, \tag{17h}$$

$$\psi^+_{i,k} \geqslant 0, \quad \psi^-_{i,k} \geqslant 0, \quad \forall i, \forall k, \tag{17i}$$

$$\boldsymbol{\lambda}^+ \geqslant 0, \quad \boldsymbol{\lambda}^- \geqslant 0. \tag{17j}$$

Proof of Proposition 7. Applying duality theory for semi-infinite linear programs to the inner problem of the distributionally robust problem (16), we obtain the following reformulation (see Delage and Ye 2010 for derivation details):

$$\underset{\mathbf{x}\in\mathscr{X},\, t,\, \mathbf{q},\, \boldsymbol{\lambda}^+,\, \boldsymbol{\lambda}^-}{\text{minimize}} \; t + \boldsymbol{\mu}^T\mathbf{q} - (\mathbf{r}^+)^T\boldsymbol{\lambda}^+ - (\mathbf{r}^-)^T\boldsymbol{\lambda}^-$$

subject to

$$t \geqslant \sum_{i=1}^m \Big\{ \max_k \{\alpha^x_{i,k}x_i + \alpha^w_{i,k}(\bar{w}_i + \hat{w}_i\zeta_i) + \beta_{i,k}\} - q_i\zeta_i + \lambda^+_i \max(0, \zeta_i - \mu_i) + \lambda^-_i \max(0, \mu_i - \zeta_i)\Big\}, \quad \forall \boldsymbol{\zeta} \in \mathscr{Z}(\Gamma),$$

$$\boldsymbol{\lambda}^+ \geqslant 0, \quad \boldsymbol{\lambda}^- \geqslant 0.$$

One might realize that the right-hand side equation of the infinite set of constraint indexed by $\boldsymbol{\zeta}$ is the sum of piecewise linear convex functions in $\boldsymbol{\zeta}$ with $2K$ pieces; for each i, each affine piece gives the highest value over k between either of the two following functions:

$$\alpha^x_{i,k}x_i + \alpha^w_{i,k}(\bar{w}_i + \hat{w}_i\zeta_i) + \beta_{i,k} - q_i\zeta_i + \lambda^+_i(\zeta_i - \mu_i)$$

or

$$\alpha^x_{i,k}x_i + \alpha^w_{i,k}(\bar{w}_i + \hat{w}_i\zeta_i) + \beta_{i,k} - q_i\zeta_i + \lambda^-_i(\mu_i - \zeta_i).$$

Applying similar steps as provided in the proof of Corollary 4, we obtain a conservative approximation of the right-hand side equation

$$t \geqslant \min_{\nu,\, \boldsymbol{\gamma},\, \boldsymbol{\psi}^+,\, \boldsymbol{\psi}^-} \Gamma\nu + \mathbf{1}^T\boldsymbol{\gamma}$$

subject to

$$\gamma_i \geqslant \psi^+_{i,k} + \alpha^x_{i,k}x_i + \alpha^w_{i,k}\bar{w}_i + \beta_{i,k} - \lambda^+_i\mu_i, \quad \forall i, \forall k,$$

$$\gamma_i \geqslant \psi^-_{i,k} + \alpha^x_{i,k}x_i + \alpha^w_{i,k}\bar{w}_i + \beta_{i,k} + \lambda^-_i\mu_i, \quad \forall i, \forall k,$$

$$\psi^+_{i,k} + \nu \geqslant \alpha^w_{i,k}\hat{w}_i - q_i + \lambda^+_i, \quad \forall i, \forall k,$$

$$\psi^+_{i,k} + \nu \geqslant -\alpha^w_{i,k}\hat{w}_i + q_i - \lambda^+_i, \quad \forall i, \forall k,$$

$$\psi^-_{i,k} + \nu \geqslant \alpha^w_{i,k}\hat{w}_i - q_i - \lambda^-_i, \quad \forall i, \forall k,$$

$$\psi^-_{i,k} + \nu \geqslant -\alpha^w_{i,k}\hat{w}_i + q_i + \lambda^+_i, \quad \forall i, \forall k,$$

$$\psi^+_{i,k} \geqslant 0, \quad \psi^-_{i,k} \geqslant 0, \quad \forall i, \forall k.$$

This constraint can then easily be reinserted in the main problem to obtain the model presented in (17). Furthermore, since for each i, each affine pieces only depend on ζ_i, we conclude that this approximation is exact based on Condition 3 of Corollary 1 being satisfied. □

We refer the reader to the e-companion for an additional exact reformulation of a distributionally robust multi-item newsvendor problem in which one instead imposes lower bounds on the probability that the realization occurs in each of a set of nested budgeted uncertainty regions: $\mathscr{Z}(1)$, $\mathscr{Z}(2)$, $\mathscr{Z}(3)$, etc.

6. Robust Multiperiod Inventory Problem

In robust multiperiod inventory problem (RMIP), the inventory manager's objective is to minimize the long term cost of inventory over a horizon of T periods. This long term cost might be composed for each period t of an ordering cost of c_t per unit, a fixed cost of K_t if an order is delivered at time t, a shortage cost of p_t per units of unsatisfied demand, and a holding cost h_t per unit held in storage. In each period, the ordered stocks are first used to satisfy the back-orders and then the current demand if possible. Any extra inventory is held until the next period after paying the associated holding cost. Unfortunately, since future demand is usually not fully determined at the time of making orders, one might require that orders are made such that the worst-case long term cost is as low as possible. This gives rise to the following robust optimization model:

$$\underset{\mathbf{u},\mathbf{v}}{\text{minimize}} \; \max_{\boldsymbol{\zeta}\in\mathscr{Z}} \sum_{t=1}^T \big\{ c_t u_t + K_t v_t + \max\big(h_t x_{t+1}(\mathbf{u},\boldsymbol{\zeta}), -p_t x_{t+1}(\mathbf{u},\boldsymbol{\zeta})\big)\big\} \tag{18a}$$

subject to

$$0 \leqslant u_t \leqslant M v_t, \; v_t \in \{0,1\}, \quad \forall t, \tag{18b}$$

where $\mathbf{v}\in\{0,1\}^T$ and $\mathbf{u}\in\mathbb{R}^T$ represent, respectively, for each t the decision of making an order or not that will be delivered at time t, and the amount to be delivered, and where

$$x_{t+1}(\mathbf{u},\boldsymbol{\zeta}) = x_1 + \sum_{j=1}^t (u_j - (\bar{w}_j + \hat{w}_j\zeta_j)), \quad \forall t.$$

Problem (18) can be considered as a special case of problem (1) where $\mathbf{c}^{t,k}_\zeta(\mathbf{u},\mathbf{v}) = \alpha_{t,k}\sum_{j\leqslant t} e_j\hat{w}_j$ and $d^{t,k}_\zeta(\mathbf{u},\mathbf{v}) = c_t u_t + K_t v_t - \alpha_{t,k}(x_1 + \sum_{j\leqslant t}(u_j - \bar{w}_j))$ where $\alpha_{t,1} = -h_t$ and $\alpha_{t,2} = p_t$. Therefore, we can easily obtain a conservative approximation based on Proposition 1:

$$\text{(LP-RC)} \quad \underset{\substack{\mathbf{u},\mathbf{v},\boldsymbol{\gamma},\boldsymbol{\Delta},\nu\\ \boldsymbol{\theta},\boldsymbol{\lambda}^+,\boldsymbol{\lambda}^-,\boldsymbol{\psi}}}{\text{minimize}} \sum_{t=1}^T \{c_t u_t + K_t v_t + \boldsymbol{\gamma}_t + \Delta_t\} + \Gamma\nu \tag{19a}$$

subject to

$$0 \leqslant u_t \leqslant M v_t, \quad v_t \in \{0,1\}, \quad \forall t, \tag{19b}$$

$$\nu + \boldsymbol{\Delta} \geqslant \sum_{t=1}^{T} \boldsymbol{\lambda}_t^+, \tag{19c}$$

$$\nu + \boldsymbol{\Delta} \geqslant \sum_{t=1}^{T} \boldsymbol{\lambda}_t^-, \tag{19d}$$

$$\gamma_t \geqslant \mathbf{1}^T \boldsymbol{\psi}_{t,k} + \Gamma\theta_{t,k} - \alpha_{t,k}\left(x_1 + \sum_{j=1}^{t}(u_j - \bar{w}_j)\right), \quad \forall t, \forall k, \tag{19e}$$

$$(\boldsymbol{\psi}_{t,k})_j + \theta_{t,k} \geqslant -(\boldsymbol{\lambda}_t^+)_j + \alpha_{t,k}\hat{w}_j, \quad \forall t, \forall j \leqslant t, \forall k, \tag{19f}$$

$$(\boldsymbol{\psi}_{t,k})_j + \theta_{t,k} \geqslant -(\boldsymbol{\lambda}_t^-)_j - \alpha_{t,k}\hat{w}_j, \quad \forall t, \forall j \leqslant t, \forall k, \tag{19g}$$

$$(\boldsymbol{\psi}_{t,k})_j + \theta_{t,k} \geqslant -(\boldsymbol{\lambda}_t^+)_j, \quad \forall t, \forall j > t, \forall k = 1,2, \tag{19h}$$

$$(\boldsymbol{\psi}_{t,k})_j + \theta_{t,k} \geqslant -(\boldsymbol{\lambda}_t^-)_j, \quad \forall t, \forall j > t, \forall k = 1,2, \tag{19i}$$

$$\boldsymbol{\psi}_{t,k} \geqslant 0, \quad \forall t, \forall j \leqslant t, \forall k = 1,2, \tag{19j}$$

where $\boldsymbol{\gamma} \in \mathbb{R}^T$, $\boldsymbol{\Delta} \in \mathbb{R}^T$, $\nu \in \mathbb{R}$, $\boldsymbol{\theta} \in \mathbb{R}^{K\times T}$, $\boldsymbol{\lambda}_t^+ \in \mathbb{R}^T$, $\boldsymbol{\lambda}_t^- \in \mathbb{R}^T$, and $\boldsymbol{\psi}_{t,k} \in \mathbb{R}^T$. Note that the expression $c_t u_t + K_t v_t$ would initially appear in the fourth constraint based on model (5) but was carried to the objective function given that it is independent of k. Interestingly, based on Corollary 1, we have conditions under which this approximation scheme returns an optimal robust solution.

Corollary 5. *The conservative approximation model* (19) *is equivalent to problem* (18) *when* $\Gamma = 1$ *or* $\Gamma = T$.

Following the spirit of Theorem 3.2 in Bertsimas and Thiele (2006), one can also relate the solution of this approximation model to a solution that would be obtained for a specific sequence of deterministic orders.

Proposition 8. (*Optimal robust policy*) *Let* $\boldsymbol{\psi}^*$ *and* $\boldsymbol{\theta}^*$ *be optimal assignments in an optimal solution of problem* (19). *The optimal robust policy of the problem* (19) *is equivalent to the optimal policy of the deterministic version of problem* (18) *with demand set to* $w'_t = \bar{w}_t + \Upsilon_t - \Upsilon_{t-1}$ *where* $\Upsilon_0 = 0$ *and* $\Upsilon_t := (B_{t,2} - B_{t,1})/(h_t + p_t)$ *for* $B_{t,k} = \mathbf{1}^T \boldsymbol{\psi}^*_{t,k} + \Gamma\theta^*_{t,k}$.

Proof of Proposition 8. Given an optimal solution tuple $(\mathbf{u}^* \cdot \mathbf{v}^*, \boldsymbol{\gamma}^*, \boldsymbol{\Delta}^*, \nu^*, \boldsymbol{\theta}^*, \boldsymbol{\lambda}^{+*}, \boldsymbol{\lambda}^{-*}, \boldsymbol{\psi}^*)$ for problem (19), it is clear that $\mathbf{u}^*$, $\mathbf{v}^*$, and $\boldsymbol{\gamma}^*$ would also be the optimal solution of problem (19) if the remaining variable were fixed to $\boldsymbol{\Delta}^*$, ν^*, $\boldsymbol{\theta}^*$, $\boldsymbol{\lambda}^{+*}$, $\boldsymbol{\lambda}^{-*}$, and $\boldsymbol{\psi}^*$. Therefore, problem (19) is equivalent to

$$\underset{u}{\text{minimize}} \sum_{t=1}^{T}\Big\{c_t u_t + K_t 1_{\{u_t>0\}} + \max(h_t \bar{x}_{t+1} + B_{t,1}, -p_t\bar{x}_{t+1} + B_{t,2}) + \Delta_t^*\Big\} + \Gamma\nu^*, \tag{20}$$

where $\bar{x}_{t+1} = x_1 + \sum_{j\leqslant t}(u_j - \bar{w}_j)$, $B_{t,k} = \mathbf{1}^T\boldsymbol{\psi}^*_{t,k} + \Gamma\theta^*_{t,k}$, and where we use $1_{u_t>0}$ as the indicator function that returns one if u_t is strictly positive and zero otherwise. Let us define variable x'_t according to the linear equation $x'_{t+1} = \bar{x}_{t+1} + (B_{t,1} - B_{t,2})/(h_t + p_t)$. This way we have that

$$\max(h_t\bar{x}_{t+1} + B_{t,1}, -p_t\bar{x}_{t+1} + B_{t2}) = \max(h_t x'_{t+1}, -p_t x'_{t+1}) + \frac{h_t B_{t,2} + p_t B_{t,1}}{h_t + p_t};$$

therefore, the problem (20) can be shown equivalent to

$$\underset{\mathbf{u}}{\text{minimize}} \sum_{t=1}^{T}\left\{c_t u_t + K_t 1_{u_t>0} + \max(h_t x'_{t+1}, -p_t x'_{t+1}) + \frac{h_t B_{t,2} + p_t B_{t,1}}{h_t + p_t} + \Delta_t^*\right\} + \Gamma\nu^*.$$

Based on the equation $x'_{t+1} = x'_t + u_t - (\bar{w}_t + \Upsilon_t - \Upsilon_{t-1})$ where $\Upsilon_t := (B_{t,2} - B_{t,1})/(h_t + p_t)$, we can conclude that the optimal robust policy of problem (20) is equivalent to the optimal policy of nominal problem with demand $w'_t = \bar{w}_t + \Upsilon_t - \Upsilon_{t-1}$. □

Remark 5. The optimal cost of the problem (19) is equal to the optimal cost for the nominal problem with the modified demand, w'_t, added to

$$\sum_{t=1}^{T}\left\{\frac{h_t B_{2,t} + p_t B_{t,1}}{h_t + p_t} + \Delta_t^*\right\} + \Gamma\nu.$$

Remark 6. This robust inventory problem was first addressed in Bertsimas and Thiele (2006), where the authors proposed a conservative approximation that relies on reversing the order of the maximization over $\boldsymbol{\zeta}$ and summation over t. The resulting model, which we will later refer to as BT-RC, can be reformulated as

$$\text{(BT-RC)} \quad \underset{\mathbf{u},\mathbf{v},\mathbf{y},\mathbf{q},\mathbf{r}}{\text{minimize}} \sum_{t=1}^{T}\{c_t u_t + K_t v_t + y_t\}$$

subject to

$$y_t \geqslant h_t\left(x_1 + \sum_{j=1}^{t}(u_j - \bar{w}_j) + q_t\Gamma + \sum_{j=1}^{t} r_{j,t}\right), \quad \forall t,$$

$$y_t \geqslant p_t\left(-x_1 - \sum_{j=1}^{t}(u_j - \bar{w}_j) + q_t\Gamma + \sum_{j=1}^{t} r_{j,t}\right), \quad \forall t,$$

$$q_t + r_{j,t} \geqslant \hat{w}_j, \quad \forall t, \forall j \leqslant t,$$

$$q_t \geqslant 0, \; r_{j,t} \geqslant 0, \quad \forall t \forall j \leqslant t,$$

$$0 \leqslant u_t \leqslant M v_t, \quad v_t \in \{0,1\}, \quad \forall t,$$

where $\mathbf{y} \in \mathbb{R}^T$, $\mathbf{q} \in \mathbb{R}^T$, and $\mathbf{r} \in \mathbb{R}^{T\times T}$. Note that to reduce the conservativeness of their approach, the authors use a different budget for each time period, which we choose to omit for the sake of comparing similar models.

This model can actually be interpreted as an AARC of problem (18), where the affine decision rule is constrained to be constant over $\boldsymbol{\zeta}$. As recognized in Gorissen and den

Hertog (2013), this indicates that their approximation can already be tightened by using affine decision rules, and, based on the results of §4.3, tightened even further by using problem (19) since the latter is equivalent to applying AARC on a lifting involving $\boldsymbol{\zeta}^+$ and $\boldsymbol{\zeta}^-$. For completeness, we present AARC applied directly on $\boldsymbol{\zeta}$ for this inventory problem:

$$\text{(AARC)} \quad \underset{\mathbf{u},\mathbf{v},\boldsymbol{\gamma},\boldsymbol{\Delta},\nu,\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\psi}}{\text{minimize}} \sum_{t=1}^{T}\{c_t u_t + K_t v_t + \boldsymbol{\gamma}_t + \Delta_t\} + \Gamma\nu$$

subject to

$$0 \leqslant u_t \leqslant M v_t, \quad v_t \in \{0,1\}, \quad \forall t$$

$$\nu + \Delta_j \geqslant \left|\sum_{t=1}^{T}\lambda_{j,t}\right|, \quad \forall j$$

$$\gamma_t \geqslant \sum_{j=1}^{T}\psi_{j,t,k} + \Gamma\theta_{t,k} - \alpha_{t,k}\left(x_1 + \sum_{j=1}^{t}(u_j - \bar{w}_j)\right), \quad \forall t, \forall k = 1,2,$$

$$\psi_{j,t,k} + \theta_{t,k} \geqslant |\lambda_{j,t} - \alpha_{t,k}\hat{w}_j|, \quad \forall t, \forall j \leqslant t, \forall k = 1,2,$$

$$\psi_{j,t,k} + \theta_{t,k} \geqslant |\lambda_{j,t}|, \quad \forall t, \forall j > t, \forall k = 1,2,$$

$$v \geqslant 0, \quad \psi_{j,t,k} \geqslant 0, \quad \theta_{t,k} \geqslant 0, \quad \forall t, \forall j \leqslant t, \forall k = 1,2,$$

where $\boldsymbol{\gamma} \in \mathbb{R}^T$, $\boldsymbol{\Delta} \in \mathbb{R}^T$, $\nu \in \mathbb{R}$, $\boldsymbol{\theta} \in \mathbb{R}^{T\times 2}$, $\boldsymbol{\lambda} \in \mathbb{R}^{T\times T}$, $\boldsymbol{\psi} \in \mathbb{R}^{T\times T\times 2}$, and $\alpha_{t,1} = -h_t$ $\forall t$ and $\alpha_{t,2} = p_t$ $\forall t$.

Remark 7. One might notice that in this section we focused on an inventory problem where all ordering decisions must be made at time zero and there is no room for adjustment as time unfolds. Although this might appear a bit limiting, our reasons for doing so are twofold. First, we believe this static version of the robust inventory problem is interesting in its own right based on the fact that in some contexts delivery contracts give no freedom to make adjustments to the orders as time evolves; even if there is some freedom, then the formulation studied in this section still gives a meaningful initial ordering plan that can later be improved by solving an updated version of the model. Second, besides the special cases described in Bertsimas et al. (2010), very little is actually known about how to get exact solutions to the static or dynamic version of this robust model. Our hope is that by focusing on the static version of the problem, we might understand what are the tools that can provide better near-optimal solutions.

7. Numerical Experiments

In this section we present numerical experiments for the robust multiperiod inventory problem discussed in §6. We initially present the performance of four different approximation methods for the instance that was studied in Bertsimas and Thiele (2006). This will illustrate how the worst-case bound can be gradually improved by using more computationally demanding models. In order from most tractable to most precise, we have the following list of formulation: the BT-RC model, the AARC model, the LP-RC model, a conservative approximate robust counterpart based on the SDP bound presented in (7) and referred to as SDP-RC, and the exact robust model solved using a cutting-plane method.[5] To study extent to which these conclusions can be generalized, we later extend the numerical analysis to a set of randomly generated instances of the multiperiod inventory problem, where every parameter (e.g., ordering cost, holding cost, amount of uncertainty, etc.) is nonstationary. In doing so, we also explore what is the "price of robustness" (as coined in Bertsimas and Sim 2004) in this class of problems.

7.1. Instance Studied by Bertsimas and Thiele

The instance studied in Section 5.2 of Bertsimas and Thiele (2006) is an inventory problem with $T = 20$, $c_t = 1$, $K_t = 0$, $h_t = 4$, $p_t = 6$, $\bar{w}_t = 100$ and $\hat{w}_t = 40$. Note that this problem is stationary in the sense that the above parameters do not depend on t. Under this context, Table 2 presents the optimal worst-case bound obtained with each method, the true worst-case cost achieved by their respective approximate solution, and their respective suboptimality gap. As expected, when $\Gamma = 0$ all four methods give the same optimal bound and solution, which is the optimal solution of the nominal problem. The performance starts to differ as Γ is increased. We can first confirm that, for any value of Γ, suboptimality gap is always reduced as we move away from the BT-RC model and use more sophisticated versions of our mixed-integer linear programming based approach. We can also confirm that, when Γ equals 1 or 20, the LP-RC and SDP-RC approach are exact, which was guaranteed by Theorem 5. This is also

Table 2. Comparison of the optimal worst-case bound, true worst-case cost, and suboptimality gap for different solution methods to the inventory problem presented in Bertsimas and Thiele (2006).

	Budget of uncertainty				
Method	0	1	10	15	20
		Worst-case bound			
BT-RC	2,000	5,848	31,840	39,560	42,480
AARC	2,000	5,800	31,457	39,306	41,818
LP-RC	2,000	5,800	31,360	38,976	41,818
SDP-RC	2,000	5,800	31,360	38,940	41,818
		Worst-case cost			
BT-RC	2,000	5,848	31,840	39,560	42,480
AARC	2,000	5,800	31,457	39,306	41,818
LP-RC	2,000	5,800	31,360	38,976	41,818
SDP-RC	2,000	5,800	31,360	38,940	41,818
Exact	2,000	5,800	31,360	38,933	41,818
		Suboptimality gap			
BT-RC (%)	0	0.83	1.53	1.61	1.58
AARC (%)	0	0.00	0.31	1.10	0
LP-RC (%)	0	0	0.00	0.11	0
SDP-RC (%)	0	0	0.00	0.02	0

Table 3. Comparison of the optimal worst-case bound, true worst-case cost, and suboptimality gap for different solution methods *averaged* over a set of 1,000 randomly generated instances of multiperiod inventory problems with 10 periods.

	Budget of uncertainty						
Method	1	2	3	4	5	6	10
			Average worst-case bound				
BT-RC	4,455	6,050	7,135	7,871	8,355	8,657	8,976
AARC	3,620	4,861	5,725	6,321	6,720	6,972	7,252
LP-RC	3,477	4,597	5,420	6,031	6,481	6,798	7,252
SDP-RC	3,477	4,592	5,412	6,024	6,476	6,794	7,252
			Average worst-case cost				
BT-RC	4,049	5,400	6,363	7,071	7,581	7,936	8,438
AARC	3,619	4,832	5,677	6,272	6,681	6,946	7,252
LP-RC	3,477	4,597	5,420	6,031	6,481	6,798	7,252
SDP-RC	3,477	4,591	5,411	6,023	6,475	6,794	7,252
True	3,477	4,585	5,407	6,020	6,474	6,794	7,252
			Average suboptimality gap				
BT-RC (%)	16.6	18.0	18.0	17.8	17.5	17.2	16.8
AARC (%)	4.4	5.7	5.2	4.3	3.3	2.3	0
LP-RC (%)	0	0.3	0.3	0.2	0.1	0.1	0
SDP-RC (%)	0	0.1	0.1	0.1	0.0	0.0	0

the case for AARC for this instance but will not be the case in general (see Table 3 for some evidence). In the case of $\Gamma = 10$, we see that using the SDP formulation allows us to reduce the suboptimality to a negligible amount, while it is slightly insufficient for $\Gamma = 15$. Although the suboptimality gap of all approximations methods are somewhat small in this example and the improvements obtained are rather limited, these results already illustrate the key differences between the four different approximation schemes. We expect these differences to be magnified in a richer experimental context.

Finally, it is worth noting that, although some of the solutions obtained are suboptimal, the worst-case cost achieved by a suboptimal solution often indicates exactly the approximate bound returned by the approximation model. For instance, one can observe in Table 2 that for the solutions provided by BT-RC, the worst-case bound approximations that BT-RC provides for $\Gamma \in \{0, 1, 10, 15, 20\}$ are exact. Yet, BT-RC does not return truly optimal solutions except when $\Gamma = 0$. Therefore, one should treat with care the fact that worst-case cost is equal to the worst-case bound for a solution that is returned by a conservative approximation scheme as there might actually exist other solutions that achieve better worst-case cost but for which the worst-case bound is very inaccurate.

7.2. Robust Performances on Randomly Generated Instances

In this second set of experiments, we consider a family of randomly generated instances of the robust multiperiod inventory problem for a horizon of $T = 10$ and $T = 100$. Specifically, each problem instance is created by randomly generating for each period the values for c_t, h_t, p_t based on a uniform distribution between 0 and 10, and the values for $\bar{w}_t$ and $\hat{w}_t$ from a uniform distribution over the interval $[0, 100]$ and $[0, \bar{w}_t]$, respectively. All fixed ordering costs K_t are again considered to be null.

As was done similarly in the previous analysis, we compare in Table 3 the optimal worst-case bound, true worst-case cost, and suboptimality gap for different solution models, yet this time the table presents the average of each indicator over a set of 1,000 randomly generated problem instances. A quick glance at the table should convince the reader that there are obvious gains in terms of worst-case cost for employing either the LP-RC or SDP-RC method instead of BT-RC or AARC. For instance, note that when $\Gamma = 6$, in these experiments the average worst-case cost was 2.2% larger with the AARC method compared to the LP-RC and the SDP-RC methods. More importantly, two valuable insights can be obtained from this table. First, we can estimate from this table that suboptimality of the approximate solution is reduced by a factor of about 4 when replacing the BT-RC method with AARC, by an additional factor of about 10 when replacing the AARC method with LP-RC, and by a final factor of at least 2 when replacing LP-RC with SDP-RC. Second, for general inventory problems it is uncommon for the worst-case bound obtained by any of these methods to be exactly equal to the true worst-case cost for the retrieved approximate solution. This can be observed by comparing the average worst-case bound to the average actual worst-case cost. The only exception is for LP-RC and SDP-RC when $\Gamma = 1$ and $\Gamma = 10$ as our theory previously established. Perhaps surprisingly, in the case of LP-RC, it also appears that for other integer Γ's it is quite rare that the worst-case bound is inexact at the proposed approximate solution, even when this solution is suboptimal.

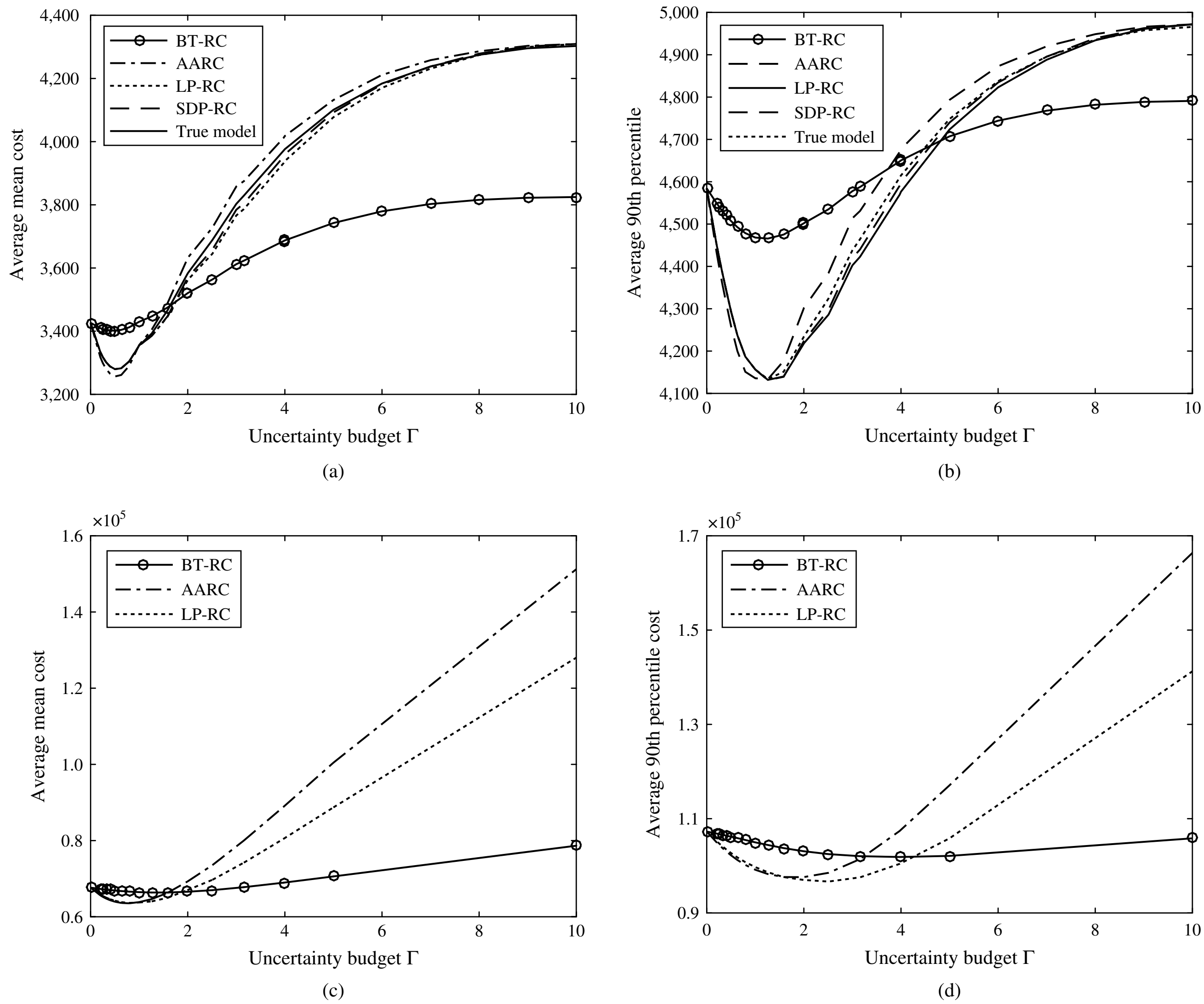
Table 4 presents further statistics regarding the suboptimality of each method. Specifically, for each value of Γ

Table 4. Proportion of random instances (in 1,000 trials) where methods achieved a set of target levels with respect to suboptimality for different values of Γ.

	Subopt. gap	Method (%)			
Γ	Interval	BT-RC	AARC	LP-RC	SDP-RC
1	$\leqslant 0.0001$	0.0	14.1	100	100
	$\leqslant 1$	0.0	26.1	100	100
	$\leqslant 10$	19.9	88.7	100	100
	Maximum gap	56.7	24.9	0	0
3	$\leqslant 0.0001$	0.0	0.4	52.6	56.6
	$\leqslant 1$	0.2	6.2	90.4	98.1
	$\leqslant 10$	15.7	91.1	100.0	100.0
	Maximum gap	54.6	23.0	4.6	2.0
5	$\leqslant 0.0001$	0.0	0.2	57.3	63.5
	$\leqslant 1$	0.0	9.3	96.6	99.7
	$\leqslant 10$	16.5	98.6	100.0	100.0
	Maximum gap	52.2	14.9	2.6	1.3

Note. The worst suboptimality gap attained by each method is also reported.

Figure 1. Average expected cost and average 90th percentile (over 1,000 problem instances) achieved by robust solutions given different level of conservativeness.



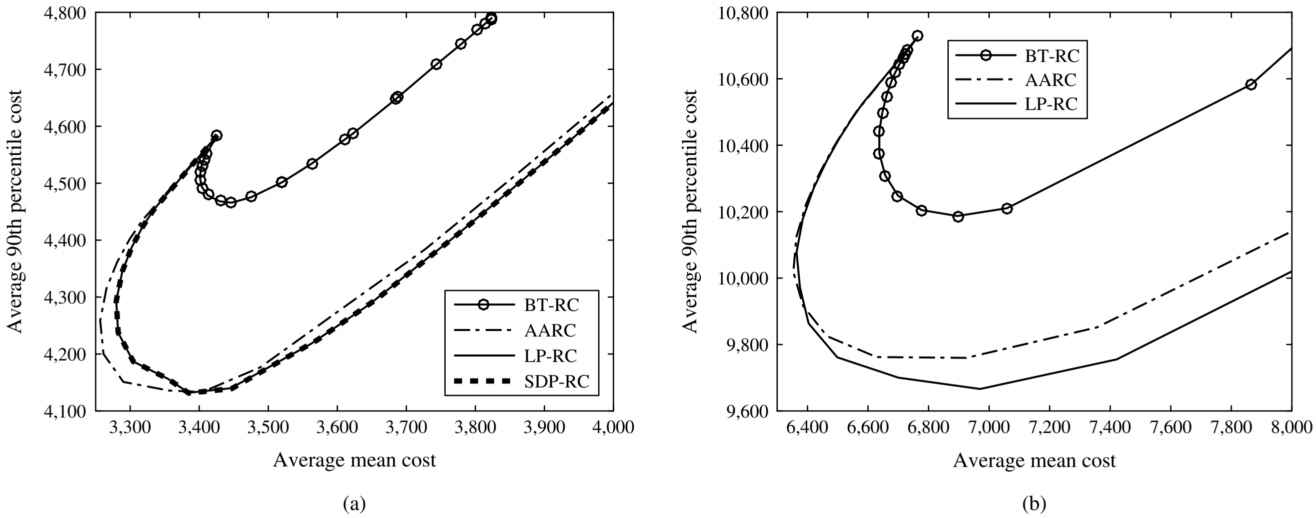
Note. Figure (a) presents average expected cost and (b) presents average 90th percentile for a horizon of 10 periods, (c) and (d) present the same statistics for a horizon of 100 periods.

studied, the table indicates for what proportion of random instances each method was able to recover a solution that was below either 0.0001%, 1%, and 10% of optimality. One can easily see that optimality is significantly improved by using the LP-RC or SDP-RC methods. In terms of worst suboptimality gap obtained over these instances, one can remark an improvement of a factor of 4, 6, and 2 for migrating from the BT-RC approximation to the AARC, and to the LP-RC and SDP-RC, respectively, when the budget is set to 5.

In practice, it is often the case that robust optimization, and especially with the budgeted uncertainty set, is used in a context where uncertain parameters are considered to be drawn from a distribution, e.g., the distribution of historical values. For this reason, we next attempt to measure the inherent trade-off that can be observed between expected cost and value at risk when using different levels of budgets for uncertainty. Namely, given a problem instance, taking the shape of a set of parameters $\{c_t, h_t, p_t, \bar{w}_t, \hat{w}_t\}_{t=1}^T$, we assume that each interval $[\bar{w}_t - \hat{w}_t, \bar{w}_t + \hat{w}_t]$ describes the support of a uniform distribution for the random demand at time period t and that demand is independent between each period. We then compare the performance of the different robust solutions in terms of expected value and the 90th percentile of the total cost achieved when different values of Γ are used.

Figure 1 presents the average expected cost and average 90th percentile[6] achieved over 1,000 problem instances by the five types of approximately robust solutions given different level of conservativeness expressed through Γ. We also present these same results in Figure 2, where the averaged (expected cost, value at risk) pair is plotted for each level of

Figure 2. Average expected cost versus average 90th percentile achieved by the different robust methods when adjusting the level of conservativeness Γ.



Note. Figure (a) presents the achieved risk-return trade-offs for a 10-day horizon, and (b) presents them for a 100-day horizon.

the budget, thus allowing us to identify the general structure of the Pareto frontier identified using each approximate model. These figures clearly present how the solution obtained from the nominal problem, when $\Gamma = 0$, can be improved both with respect to expected cost and to value at risk by considering a robust alternative. Although this behavior might seem surprising, it is in agreement with Delage et al. (2014, Remark 3.2) that claims that the mean value problem, where one replaces every random variable with its expected value, actually provides an optimistic solution (i.e., solution based on best-case distribution) to stochastic programs when the objective function is convex with respect to uncertain parameters. For this family of problems, it also appears that there is a threshold above which the budget Γ leads to solutions that cannot be statistically motivated (i.e., dominated in terms of both expected value and 90th percentile). This threshold appears to be, respectively, 1.5 and 3 in the case of problems with 10 periods and 100 periods. Based on Figures 1(a) and 1(b), it also appears that although there is a lot to gain from using a more sophisticated model than BT-RC, the statistical performances of models that obtain the robust solution with greater precision than AARC are highly comparable for a short horizon. The difference between AARC and LP-RC is a bit more noticeable when the horizon is larger as portrayed in Figures 1(c) and 1(d) where we see that the LP-RC dominates AARC for nearly all values of Γ. Note that in our experiments with $T = 100$, we did not include the performance of SDP-RC since it was too computationally demanding and since the performance seemed highly comparable to LP-RC.

8. Conclusion

In this article we proposed a new scheme that can be used to generate conservative approximations of robust optimization problems involving the sum of piecewise linear functions and a polyhedral uncertainty set. This scheme exploits the fractional relaxation of a MILP known to be equivalent to the adversarial problem and can be used to identify two specific approximation models that, respectively, take the shape of a linear program and a SDP. Although the linear programming model is shown to be equivalent to an application of AARC on a lifting of the parameter space, the SDP model clearly departs from previously known approximation techniques. Our approximation scheme also allows us to exploit the concept of total unimodularity to establish new conditions under which our LP-RC (and implicitly the AARC approach) model provides exact solutions. In particular, we identified the first exact reformulations for a robust (and distributionally robust) multi-item newsvendor problem with budgeted uncertainty set and a reformulation for robust multiperiod inventory problems that is exact whether the uncertainty region reduces to a L_1-norm ball or to a box. An extensive set of empirical results finally illustrates the quality of the solutions obtained from different approximation schemes on randomly generated instances of the latter field of application.

Although very relevant to the discussion of this paper, we leave open the question of how to extend our approximation scheme to uncertainty sets that are not polyhedral. In this regard, it appears that one might be able to reach interesting conclusions in contexts where the uncertainty set takes the shape of $\mathcal{Z} := \{\boldsymbol{\zeta} \mid g_j(\boldsymbol{\zeta}) \leqslant b_j, \forall j = 1, 2, \ldots, J\}$ using a set of convex positive homogeneous $g_j(\cdot)$ functions, i.e., functions for which $g_j(\alpha\boldsymbol{\zeta}) = \alpha g_j(\boldsymbol{\zeta})$ for all $\alpha \geqslant 0$. This is, for example, the case when using an ellipsoidal set where $g_j(\boldsymbol{\zeta}) := \|\boldsymbol{\zeta}\|_2$. Under these conditions, the MILP representation of the adversarial problem studied in §4 will reduce to problem (4) with (4b) and (4g) replaced with

$$g_j(\boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-) \leqslant b_j, \quad \forall j,$$

$$g_j(\boldsymbol{\Delta}^+_{i,k} - \boldsymbol{\Delta}^-_{i,k}) \leqslant b_j z_{i,k}, \quad \forall i, j, k,$$

This is because we can now exploit that for all i, j, and k, we have that

$$g_j(\boldsymbol{\Delta}^+_{i,k}-\boldsymbol{\Delta}^-_{i,k})=g_j(z_{i,k}(\boldsymbol{\zeta}^+-\boldsymbol{\zeta}^-))=z_{i,k}g_j(\boldsymbol{\zeta}^+-\boldsymbol{\zeta}^-)\leqslant b_j z_{i,k}.$$

Hence, the second constraint is necessarily redundant in the mixed integer program but will lead to a tighter conservative approximation after applying fractional relaxation.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/opre.2016.1483.

Acknowledgments

The authors thank the associate editor and referee for their constructive criticism. They also thank Dan Iancu for valuable discussions. The second author gratefully acknowledges the support of the Canadian Natural Sciences and Engineering Research Council [Grant 386416-2010].

Appendix. Proof of Proposition 3

We present this proof in four steps. We first introduce in §A the notion of integral polytope and total unimodularity as these concepts are fundamental components of our proof. We then go through each of the three sets of conditions and present arguments for our conclusions in §B–D.

A.1. Integral Property of Polytope

Let the matrix $\mathbf{A}\in\mathbb{R}^{m\times n}$ and vector $\mathbf{b}\in\mathbb{R}^m$ define the convex polytope $\{\mathbf{z}\in\mathbb{R}^n \mid \mathbf{z}\geqslant 0, \mathbf{A}\mathbf{z}=\mathbf{b}\}$.

Definition 2. A matrix $\mathbf{A}$ is called *totally unimodular* if the determinant of every submatrix of $\mathbf{A}$ is equal to $+1$ or -1.

Lemma 1. (See Truemper 1978 as cited in Grady and Polimeni 2010, p. 313.) *Let the matrix $\mathbf{A}\in\mathbb{R}^{m\times n}$ and vector $\mathbf{b}\in\mathbb{R}^m$ define the convex polytope $\{\mathbf{z}\in\mathbb{R}^n \mid \mathbf{z}\geqslant 0, \mathbf{A}\mathbf{z}=\mathbf{b}\}$. All the vertices of this convex polytope are integer valued if $\mathbf{A}$ is totally unimodular and $\mathbf{b}$ is integer valued.*

Hoffman and Kruskal (1956) proposed the following sufficient conditions for a matrix to be totally unimodular.

Lemma 2. *Let $\mathbf{A}$ be an $m\times n$ matrix containing only elements in the set $\{-1,0,1\}$, then $\mathbf{A}$ is a totally unimodular matrix if both of the following conditions are satisfied*:

1. *Each column of $\mathbf{A}$ contains at most two nonzero elements.*
2. *The rows of $\mathbf{A}$ can be partitioned into two sets $\mathbf{A}_1$ and $\mathbf{A}_2$ such that two nonzero entries in a column are in the same set of rows if they have different signs and in different sets of rows if they have the same sign.*

In what follows we will use the following result, which we present as an example of application of Lemma 2.

Lemma 3. *The polytope defined by*

$$\boldsymbol{\zeta}^+\geqslant 0 \quad \textit{and} \quad \boldsymbol{\zeta}^-\geqslant 0 \quad \textit{and} \quad \zeta^+_j+\zeta^-_j\leqslant 1,\quad \forall j, \tag{21a}$$

$$\mathbf{1}^T(\boldsymbol{\zeta}^+ +\boldsymbol{\zeta}^-)=\Gamma, \tag{21b}$$

$$\sum_{k=1}^K z_{i,k}=1,\quad \forall i, \tag{21c}$$

$$\sum_{k=1}^K \Delta^+_{i,k}=\zeta^+_i \quad \textit{and} \quad \sum_{k=1}^K \Delta^-_{i,k}=\zeta^-_i,\quad \forall i, \tag{21d}$$

$$\Delta^+_{i,k}\geqslant 0 \quad \textit{and} \quad \Delta^-_{i,k}\geqslant 0 \quad \textit{and} \quad \Delta^+_{i,k}+\Delta^-_{i,k}\leqslant z_{i,k},\quad \forall i,k, \tag{21e}$$

where $\boldsymbol{\zeta}^+\in\mathbb{R}^m$, $\boldsymbol{\zeta}^-\in\mathbb{R}^m$, $\mathbf{z}\in\mathbb{R}^{m\times K}$, $\boldsymbol{\Delta}^+\in\mathbb{R}^{m\times K}$, and $\boldsymbol{\Delta}^-\in\mathbb{R}^{m\times K}$, has integer vertices.

Proof of Lemma 3. First, let us realize that constraints (21c)–(21e) make constraint (21a) redundant. Hence, constraints (21a) to (21e) are reduced to

$$\mathbf{1}^T(\boldsymbol{\zeta}^+ +\boldsymbol{\zeta}^-)=\Gamma, \tag{22a}$$

$$\sum_{k=1}^K z_{i,k}=1,\quad \forall i, \tag{22b}$$

$$\sum_{k=1}^K \Delta^+_{i,k}-\zeta^+_i=0 \quad \text{and} \quad \sum_{k=1}^K \Delta^-_{i,k}-\zeta^-_i=0,\quad \forall i, \tag{22c}$$

$$\Delta^+_{i,k}+\Delta^-_{i,k}+s_{i,k}-z_{i,k}=0,\quad \forall i,\,k, \tag{22d}$$

$$\Delta^+_{i,k}\geqslant 0 \quad \text{and} \quad \Delta^-_{i,k}\geqslant 0, \quad \text{and} \quad s_{i,k}\geqslant 0, z_{ik}\geqslant 0,\quad \forall i,\,k, \tag{22e}$$

where constraint (22d) is presented in standard form with some additional decision variables $s_{i,k}\in\mathbb{R}$ for all (i,k) pairs.

In constraints (22a) to (22d), matrix of coefficients $\mathbf{A}$ contains only elements in the set $\{-1,0,1\}$. The columns of matrix $\mathbf{A}$ contain two nonzero elements, except for the columns associated with $s_{i,k}$, which only have one. In the columns associated to Δ^+_{ik} and Δ^-_{ik}, these coefficients are equal to 1 because of constraints (22c) and (22d), and in columns associated with z_{ik}, each column holds a $+1$ and -1 coefficient based on constraints (22b) and (22d). Moreover, in columns associated with ζ^+_i and ζ^-_i, each column holds a $+1$ and -1 coefficient based on constraints (22a) and (22c). Condition 2 of Lemma 2 is easily satisfied. Condition 2 is satisfied if matrix $\mathbf{A}$ is partitioned so that $\mathbf{A}_1$ contains the coefficients associated with constraints (22a) and (22c), and $\mathbf{A}_2$ contains entries of the other rows. $\square$

A.2. Condition 1

We start with the case of $\Gamma=1$. First, one can show that constraints (4b) to (4i) reduce to the following:

$$\boldsymbol{\zeta}^+\geqslant 0 \quad \text{and} \quad \boldsymbol{\zeta}^-\geqslant 0,$$

$$\mathbf{1}^T(\boldsymbol{\zeta}^+ +\boldsymbol{\zeta}^-)=1,$$

$$\sum_{k=1}^K z_{i,k}=1,\quad \forall i,$$

$$\sum_{k=1}^K \boldsymbol{\Delta}^+_{i,k}-\boldsymbol{\zeta}^+=0 \quad \text{and} \quad \sum_{k=1}^K \boldsymbol{\Delta}^-_{i,k}-\boldsymbol{\zeta}^-=0,\quad \forall i,$$

$$\boldsymbol{\Delta}^+_{i,k}\geqslant 0 \quad \text{and} \quad \boldsymbol{\Delta}^-_{i,k}\geqslant 0,\quad \forall i,\,k,$$

$$z_{ik}-\sum_{j=1}^m(\boldsymbol{\Delta}^+_{i,k})_j-(\boldsymbol{\Delta}^-_{i,k})_j\geqslant 0,\quad \forall i,\,k.$$

Together, these constraints further imply that $z_{i,k} = \sum_{j=1}^m (\boldsymbol{\Delta}^+_{i,k})_j + (\boldsymbol{\Delta}^-_{i,k})_j$ since if it was any larger, say by a positive amount $\boldsymbol{\Delta}_{i,k} > 0$ then the sum of $z_{i,k}$ would lead to

$$\sum_{k=1}^K z_{i,k} = \sum_{k=1}^K \left\{ \sum_j \left\{ (\boldsymbol{\Delta}^+_{i,k})_j + (\boldsymbol{\Delta}^-_{i,k})_j \right\} + \boldsymbol{\Delta}_{i,k} \right\}$$
$$= \sum_{j=1}^m \{\zeta_j^+ + \zeta_j^-\} + \sum_{k=1}^K \boldsymbol{\Delta}_{i,k} = 1 + \sum_{k=1}^K \boldsymbol{\Delta}_{i,k} > 1,$$

which contradicts the fact that they should sum to one.

This allows us to say that there is always an optimal solution of the fractional relaxation of problem (4) that lies at one of the vertices of the polytope described by

$$\sum_{j=1}^m \zeta_j^+ + \zeta_j^- = 1,$$
$$\sum_{k=1}^K \boldsymbol{\Delta}^+_{1,k} - \boldsymbol{\zeta}^+ = 0,$$
$$\sum_{k=1}^K \boldsymbol{\Delta}^+_{i,k} - \sum_{k=1}^K \boldsymbol{\Delta}^+_{i-1,k} = 0, \quad \forall i = 2, \ldots, N,$$
$$\sum_{k=1}^K \boldsymbol{\Delta}^-_{1,k} - \boldsymbol{\zeta}^- = 0,$$
$$\sum_{k=1}^K \boldsymbol{\Delta}^-_{i,k} - \sum_{k=1}^K \boldsymbol{\Delta}^-_{i-1,k} = 0, \quad \forall i = 2, \ldots, N,$$
$$z_{i,k} - \sum_{j=1}^m (\boldsymbol{\Delta}^+_{i,k})_j + (\boldsymbol{\Delta}^-_{i,k})_j = 0, \quad \forall i, k,$$
$$\boldsymbol{\zeta}^+ \geqslant 0 \quad \text{and} \quad \boldsymbol{\zeta}^- \geqslant 0,$$
$$\boldsymbol{\Delta}^+_{i,k} \geqslant 0 \quad \text{and} \quad \boldsymbol{\Delta}^-_{i,k} \geqslant 0, \quad \forall i, k.$$

One can easily show that the polytope defined by the first five constraints has integral vertices by confirming that, in the representation $\mathbf{A}\mathbf{y} = \mathbf{b}$, where $\mathbf{y}$ stands for the vector formed by appending all $\boldsymbol{\zeta}^+$, $\boldsymbol{\zeta}^-$, $\boldsymbol{\Delta}^+_{i,k}$ $\boldsymbol{\Delta}^+_{i,k}$ variables, the $\mathbf{A}$ matrix is totally unimodular. Indeed, total unimodularity is directly verified on the $\mathbf{A}$ matrix capturing the first five constraints. For each of these integer vertices, the projection $z_{i,k} := \sum_{j=1}^m \{(\boldsymbol{\Delta}^+_{i,k})_j + (\boldsymbol{\Delta}^-_{i,k})_j\}$ makes $z_{i,k}$ also an integer. This completes the proof that there always exists an optimal solution of the fractional relaxation of problem (4) that is integer; thus, the optimal value of this relaxation is exact.

A.3. Condition 2

For the second case, we will assume without loss of generality that $m = N$, i.e., that there is one perturbation variable per convex subfunction in the objective function. Then we observe that the description of the polytope reduces to the following set of constraints.

$$\boldsymbol{\zeta}^+ \geqslant 0 \quad \text{and} \quad \boldsymbol{\zeta}^- \geqslant 0 \quad \text{and} \quad \zeta_j^+ + \zeta_j^- \leqslant 1, \quad \forall j,$$
$$\sum_{k=1}^K z_{i,k} = 1, \quad \forall i,$$
$$\sum_{k=1}^K \boldsymbol{\Delta}^+_{i,k} = \boldsymbol{\zeta}^+ \quad \text{and} \quad \sum_{k=1}^K \boldsymbol{\Delta}^-_{i,k} = \boldsymbol{\zeta}^-, \quad \forall i,$$
$$\boldsymbol{\Delta}^+_{i,k} \geqslant 0 \quad \text{and} \quad \boldsymbol{\Delta}^-_{i,k} \geqslant 0 \quad \text{and} \quad \boldsymbol{\Delta}^+_{i,k} + \boldsymbol{\Delta}^-_{i,k} \leqslant z_{i,k}, \quad \forall i, k.$$

Since an integer solution is feasible with respect to these constraints, we know that the fractional relaxation must necessarily achieve a higher value than problem (3). Yet, we now show that this relaxed problem is also upper bounded by problem (3) when $\mathbf{c}_{i,k}$ has the given structure.

We get the upper bounding problem by first relaxing the feasible set to the following:

$$\boldsymbol{\zeta}^+ \geqslant 0 \quad \text{and} \quad \boldsymbol{\zeta}^- \geqslant 0 \quad \text{and} \quad \zeta_j^+ + \zeta_j^- \leqslant 1, \quad \forall j,$$
$$\sum_{k=1}^K z_{i,k} = 1, \quad \forall i,$$
$$\sum_{k=1}^K \boldsymbol{\Delta}_{i,k} = (\boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-), \quad \forall i,$$
$$\|\boldsymbol{\Delta}_{i,k}\|_\infty \leqslant z_{i,k}, \quad \forall i, k.$$

For any fixed $\boldsymbol{\zeta}^+$ and $\boldsymbol{\zeta}^-$, the maximum over $\boldsymbol{\Delta}$ and $\mathbf{z}$ is upper bounded by its dual problem. Specifically, the optimal value of the problem

$$\begin{aligned} \underset{\boldsymbol{\Delta},\mathbf{z}}{\text{maximize}} \quad & \sum_{i=1}^N \sum_{k=1}^K \{\mathbf{c}_{i,k}^T \boldsymbol{\Delta}_{i,k} + d_{i,k} z_{i,k}\} & \text{(23a)} \\ \text{subject to} \quad & \sum_{k=1}^K z_{i,k} = 1, \ \forall i, & \text{(23b)} \\ & \sum_{k=1}^K \boldsymbol{\Delta}_{i,k} = \boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-, \quad \forall i, & \text{(23c)} \\ & \|\boldsymbol{\Delta}_{i,k}\|_\infty \leqslant z_{i,k}, \quad \forall i, k, & \text{(23d)} \end{aligned}$$

is upper bounded by the optimal value of

$$\begin{aligned} \underset{\boldsymbol{\gamma},\boldsymbol{\lambda}}{\text{minimize}} \quad & \sum_{i=1}^N \{\gamma_i + \boldsymbol{\lambda}_i^T(\boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-)\} \\ \text{subject to} \quad & \gamma_i \geqslant d_{i,k} + \|\mathbf{c}_{i,k} - \boldsymbol{\lambda}_i\|_1, \quad \forall i, k, \end{aligned}$$

where $\boldsymbol{\gamma} \in \mathbb{R}^N$ and $\boldsymbol{\lambda} \in \mathbb{R}^{N \times m}$ are the dual variables for constraint (23b) and (23c), respectively.

Based on the observation that reversing the order of $\max_{\boldsymbol{\zeta}}$ and $\min_{\boldsymbol{\gamma},\boldsymbol{\lambda}}$ can only lead to a further upper bound, we are left with

$$\begin{aligned} \underset{\boldsymbol{\gamma},\boldsymbol{\lambda}}{\text{minimize}} \quad & \max_{\boldsymbol{\zeta}:\, \|\boldsymbol{\zeta}\|_\infty \leqslant 1} \sum_{i=1}^N \{\gamma_i + \boldsymbol{\lambda}_i^T \boldsymbol{\zeta}\} & \text{(24a)} \\ \text{subject to} \quad & \gamma_i + \boldsymbol{\lambda}_i^T \boldsymbol{\zeta} \geqslant d_{i,k} + \mathbf{c}_{i,k}^T \boldsymbol{\zeta}, \quad \forall \|\boldsymbol{\zeta}\|_\infty \leqslant 1, \forall i, k. & \text{(24b)} \end{aligned}$$

Following some recent results presented in Bertsimas et al. (2010), this last optimization model can be shown to be equivalent to problem (3) when $\mathbf{c}_{i,k}$ has the given structure. Specifically, given the structure of $\mathbf{c}_{i,k}$, problem (3) can be seen as evaluating the worst-case cost of linear dynamic system with bounded independent perturbations under a fixed policy

$$\begin{aligned} \underset{\zeta_1,\ldots,\zeta_N,\, x_1,\ldots,x_N}{\text{maximize}} \quad & \sum_{t=1}^N \left\{ \max_k \alpha_{t,k} x_t + d_{t,k} \right\} \\ \text{subject to} \quad & x_1 = \beta_1 \zeta_1, \\ & x_t = x_{t-1} + \beta_t \zeta_t, \quad t = 2, \ldots, N, \\ & -1 \leqslant \zeta_t \leqslant 1, \quad \forall t. \end{aligned}$$

In Bertsimas et al. (2010, Theorem 3.1), the authors proved that an optimized affine running cost can be used instead of a convex one to measure exactly the total cost incurred by such a policy. Since problem (24) optimizes over all upper bounding affine running cost, it will necessarily achieve as optimal value the optimal value of problem (3). Hence, the fractional relaxation of problem (4) is tight.

A.4. Condition 3

For the third case, we can first realize that the objective function reduces to

$$\sum_{i=1}^{N}\sum_{k=1}^{K}\big\{\alpha_{i,k}((\boldsymbol{\Delta}_{i,k}^{+})_i-(\boldsymbol{\Delta}_{i,k}^{-})_i)+d_{i,k}z_{i,k}\big\}.$$

Hence, it is invariant to our choice for the variables $(\boldsymbol{\Delta}_{i,k}^{+})_j$ and $(\boldsymbol{\Delta}_{i,k}^{-})_j$ for all $j\neq i$. Therefore, we are interested in studying the vertices of the projection of the polytope defined by constraints (4b) to (4i) over the space spanned by the variables $(\boldsymbol{\Delta}_{i,k}^{+})_i$ and $(\boldsymbol{\Delta}_{i,k}^{+})_i$, $z_{i,k}$, $\boldsymbol{\zeta}^{+}$, and $\boldsymbol{\zeta}^{-}$.

When we remove from constraints (4b) to (4i), any constraint that involves $(\boldsymbol{\Delta}_{i,k}^{+})_j$ with $j\neq i$, we get the following set of constraints:

$$\boldsymbol{\zeta}^{+}\geqslant 0 \quad\text{and}\quad \boldsymbol{\zeta}^{-}\geqslant 0 \quad\text{and}\quad \zeta_j^{+}+\zeta_j^{-}\leqslant 1,\quad \forall j,$$
$$\mathbf{1}^T(\boldsymbol{\zeta}^{+}+\boldsymbol{\zeta}^{-})=\Gamma,$$
$$\sum_{k=1}^{K}z_{i,k}=1,\quad \forall i,$$
$$\sum_{k=1}^{K}(\boldsymbol{\Delta}_{i,k}^{+})_i=\zeta_i^{+} \quad\text{and}\quad \sum_{k=1}^{K}(\boldsymbol{\Delta}_{i,k}^{-})_i=\zeta_i^{-},\quad \forall i,$$
$$(\boldsymbol{\Delta}_{i,k}^{+})_i\geqslant 0 \quad\text{and}\quad (\boldsymbol{\Delta}_{i,k}^{-})_i\geqslant 0 \quad\text{and}\quad (\boldsymbol{\Delta}_{i,k}^{+})_i+(\boldsymbol{\Delta}_{i,k}^{-})_i\leqslant z_{i,k},\quad \forall i,\,k.$$

By construction, the polytope defined by these constraints must include the projection that we seek to define. An important property of this polytope is that it has integer vertices when Γ is an integer (see Lemma 3). Yet, when Γ is integer, it is also a subset of the projection that we are interested in. This can be confirmed by verifying that for any feasible solution of these constraints $(\boldsymbol{\zeta}^{+},\boldsymbol{\zeta}^{-},(\boldsymbol{\Delta}_{i,k}^{+})_i,(\boldsymbol{\Delta}_{i,k}^{-})_i,z_{i,k})$ one can create a solution $(\boldsymbol{\zeta}^{+},\boldsymbol{\zeta}^{-},\boldsymbol{\Delta}_{i,k}^{+},\boldsymbol{\Delta}_{i,k}^{-},z_{i,k})$ with $(\boldsymbol{\Delta}_{i,k}^{+})_j:=z_{i,k}(\sum_{k'=1}^{K}(\boldsymbol{\Delta}_{j,k'}^{+})_j)$ and $(\boldsymbol{\Delta}_{i,k}^{-})_j:=z_{i,k}(\sum_{k'=1}^{K}(\boldsymbol{\Delta}_{j,k'}^{-})_j)$ that is feasible according to constraints (4b) to (4i). Specifically, we have for all $i=1,\ldots,N$ and $j=1,\ldots,m$

$$\sum_{k=1}^{K}(\boldsymbol{\Delta}_{i,k}^{+})_j=\sum_{k=1}^{K}z_{i,k}\left(\sum_{k'=1}^{K}(\boldsymbol{\Delta}_{j,k'}^{+})_j\right)=\sum_{k'=1}^{K}(\boldsymbol{\Delta}_{j,k'}^{+})_j=\zeta_j^{+},$$
$$(\boldsymbol{\Delta}_{i,k}^{+})_j+(\boldsymbol{\Delta}_{i,k}^{-})_j=z_{i,k}\left(\sum_{k'=1}^{K}(\boldsymbol{\Delta}_{j,k'}^{+})_j+(\boldsymbol{\Delta}_{j,k'}^{-})_j\right)\leqslant z_{i,k}.$$

For constraints (4i), we verify the two cases. If, for some (i',k'), $z_{i',k'}=0$, then

$$\sum_{j=1}^{m}(\boldsymbol{\Delta}_{i',k'}^{+})_j+(\boldsymbol{\Delta}_{i',k'}^{-})_j=0\leqslant 0.$$

Otherwise $x_{i',k'}=1$ and, since the sum of $z_{i',k}$ over the k's is equal to one, it must be that $\sum_{k=1}^{K}(\boldsymbol{\Delta}_{i',k}^{+})_{i'}=(\boldsymbol{\Delta}_{i',k'}^{+})_{i'}=\zeta_{i'}^{+}$ and similarly for $(\boldsymbol{\Delta}_{i',k'}^{-})_{i'}=\zeta_{i'}^{-}$. Hence, we have that

$$\sum_{j=1}^{m}\big\{(\boldsymbol{\Delta}_{i',k'}^{+})_j+(\boldsymbol{\Delta}_{i',k'}^{-})_j\big\}=z_{i',k'}\sum_{j=1}^{m}\big\{\zeta_j^{+}+\zeta_j^{-}\big\}=\Gamma z_{i',k'}.$$

We are left to conclude that since the projected polytope has integer vertices, it must be that an optimal solution of problem (4) has integer vertices at least for the variables $(\boldsymbol{\zeta}^{+},\boldsymbol{\zeta}^{-},(\boldsymbol{\Delta}_{i,k}^{+})_i,(\boldsymbol{\Delta}_{i,k}^{-})_i,z_{i,k})$. Completing this part of the solution with the suggestion above will give an optimal solution of the problem that is completely integer.

Endnotes

1. In other words, the approximate solution is guaranteed to achieve a worst-case cost that is bounded by the optimal value of the approximate optimization problem.
2. Note that to obtain such a reformulation one might need to identify some $\bar{\boldsymbol{\zeta}}\in\mathcal{Z}$ and $\hat{\boldsymbol{\zeta}}\in\mathbb{R}^m$ such that $\mathcal{Z}\subseteq[\bar{\zeta}_1-\hat{\zeta}_1,\bar{\zeta}_1+\hat{\zeta}_1]\times\cdots\times[\bar{\zeta}_m-\hat{\zeta}_m,\bar{\zeta}_m+\hat{\zeta}_m]$ and reformulate the problem in terms of $\zeta_j':=(\zeta_j-\bar{\zeta}_j)/\hat{\zeta}_j$ for all $j=1,\ldots,m$ with $\boldsymbol{\zeta}'\in\mathcal{Z}'\subseteq[-1,1]^m$. This would lead to reformulating the objective function according to $\mathbf{c}_\zeta^{i,k'}(\mathbf{x}):=\operatorname{diag}(\hat{\boldsymbol{\zeta}})\mathbf{c}_\zeta^{i,k}(\mathbf{x})$ and $d_\zeta^{i,k'}(\mathbf{x}):=\mathbf{c}_\zeta^{i,k}(\mathbf{x})^T\bar{\boldsymbol{\zeta}}+d_\zeta^{i,k}(\mathbf{x})$.
3. Note that when replacing $\boldsymbol{\zeta}:=\boldsymbol{\zeta}^{+}-\boldsymbol{\zeta}^{-}$, we can replace $\|\boldsymbol{\zeta}^{+}+\boldsymbol{\zeta}^{-}\|_1\leqslant\Gamma$ with $\|\boldsymbol{\zeta}^{+}+\boldsymbol{\zeta}^{-}\|_1=\Gamma$ because Γ is assumed smaller or equal to m. In particular, if a candidate solution does not use all the budget, it is always possible to find an index for which $\zeta_i^{+}+\zeta_i^{-}<1$ and add the same amount to both positive and negative term without affecting $\boldsymbol{\zeta}^{+}-\boldsymbol{\zeta}^{-}$.
4. New variables, such as positive semidefinite matrices, can also be added to the adversarial model, in a lifting and projection fashion, as long as these variables are known to be bounded.
5. The exact robust model is solved using an analytic center cutting-plane method up to a precision of 10^{-6} where cuts are generated by using CPLEX to solve the inner mixed-integer linear program.
6. Specifically, each statistic is estimated using 100 samples of random demand vector and averaged over 1,000 problem instances.

References

Ben-Tal A, Nemirovski A (1997) Robust truss topology design via semidefinite programming. *SIAM J. Optim.* 7(4):991–1016.
Ben-Tal A, Nemirovski A (1998) Robust convex optimization. *Math. Oper. Res.* 23(4):769–805.
Ben-Tal A, El Ghaoui L, Nemirovski A (2009a) *Robust Optimization* (Princeton University Press, Princeton, NJ).
Ben-Tal A, Golany B, Shtern S (2009b) Robust multiechelon multiperiod inventory control. *Eur. J. Oper. Res.* 199(3):922–935.
Ben-Tal A, Golany B, Nemirovski A, Vial JP (2005) Retailer-supplier flexible commitments contracts: A robust optimization approach. *Manufacturing Service Oper. Management* 7(3):248–271.
Ben-Tal A, Goryashko A, Guslitzer E, Nemirovski A (2004) Adjustable robust solutions of uncertain linear programs. *Math. Programming* 99(2):351–376.
Benson SJ, Ye Y, Zhang X (2000) Solving large-scale sparse semidefinite programs for combinatorial optimization. *SIAM J. Optim.* 10(2): 443–461.
Bertsimas D, Sim M (2004) The price of robustness. *Oper. Res.* 52(1):35–53.
Bertsimas D, Thiele A (2006) A robust optimization approach to inventory theory. *Oper. Res.* 54(1):150–168.
Bertsimas D, Brown DB, Caramanis C (2011a) Theory and application of robust optimization. *SIAM Rev.* 3(3):464–501.

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

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

Bienstock D, Özbay N (2008) Computing robust basestock levels. *Discrete Optim.* 5(2):389–414.

Boyd S, Kim S-J, Patil D, Horowitz M (2005) Digital circuit optimization via geometric programming. *Oper. Res.* 53(6):899–932.

Calafiore G, El Ghaoui L (2001) Robust maximum likelihood estimation in the linear model. *Automatica* 37(4):573–580.

Candia-Véjar A, Álvarez-Miranda E, Maculan N (2011) Minmax regret combinatorial optimization problems: An algorithmic perspective. *RAIRO-Oper. Res.* 45(2):101–129.

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

Delage E, Ye Y (2010) Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Oper. Res.* 58(3):569–612.

Delage E, Ye Y, Arroyo S (2014) The value of stochastic modeling in two-stage stochastic programs with cost uncertainty. *Oper. Res.* 62(6): 1377–1393.

Denton BT, Miller AJ, Balasubramanian HJ, Huschka TR (2010) Optimal allocation of surgery blocks to operating rooms under uncertainty. *Oper. Res.* 58(4):802–816.

Düzgün R, Thiele A (2010) Robust optimization with multiple ranges: Theory and application to R&D project selection. Technical report, Industrial and Systems Engineering Department, Lehigh University, Bethlehem, PA.

El Ghaoui L, Lebret H (1997) Robust solutions to least-squares problems with uncertain data. *SIAM J. Matrix Anal. Appl.* 18(4):1035–1064.

Gallego G, Moon I (1993) The distribution free newsboy problem: Review and extensions. *J. Oper. Res. Soc.* 44(8):825–834.

Ghaddar B, Vera JC, Anjos MF (2011) Second-order cone relaxations for binary quadratic polynomial programs. *SIAM J. Optim.* 21(1):391–414.

Goldfarb D, Iyengar G (2002) Robust portfolio selection problems. *Math. Oper. Res.* 28(1):1–38.

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.

Grady LJ, Polimeni J (2010) *Discrete Calculus: Applied Analysis on Graphs for Computational Science* (Springer , London).

Hanasusanto G, Kuhn D, Wallace SW, Zymler S (2014) Distributionally robust multi-item newsvendor problems with multimodal demand distributions. *Math. Programming* 1–32.

Hoffman AJ, Kruskal JB (1956) Integral boundary points of convex polyhedra. *Linear Inequalities and Related Systems* 38:223–246.

Iancu DA, Sharma M, Sviridenko M (2013) Supermodularity and affine policies in dynamic robust optimization. *Oper. Res.* 61(4):941–956.

José Alem D, Morabito R (2012) Production planning in furniture settings via robust optimization. *Comput. Oper. Res.* 39(2):139–150.

Lasserre JB (2002) An explicit equivalent positive semidefinite program for nonlinear 0-1 programs. *SIAM J. Optim.* 12(3):756–769.

Lovász L, Schrijver A (1991) Cones of matrices and set-functions and 0-1 optimization. *SIAM J. Optim.* 1(2):166–190.

Mak HY, Rong Y, Zhang J (2015) Appointment scheduling with limited distributional information. *Management Sci.* 61(2):316–334.

Moon I, Silver EA (2000) The multi-item newsvendor problem with a budget constraint and fixed ordering costs. *J. Oper. Res. Soc.* 51(5):602–608.

Scarf H (1958) A min-max solution of an inventory problem. Scarf HE, Arrow KS, Karlin S, eds. *Studies in the Mathematical Theory of Inventory and Production* (Stanford University Press, Stanford, CA), 201–209.

Truemper K (1978) Algebraic characterizations of unimodular matrices. *SIAM J. Appl. Math.* 35(2):328–332.

van der Vlerk MH (2004) Convex approximations for complete integer recourse models. *Math. Programming* 99(2):297–310.

Wang Z, Glynn PW, Ye Y (2010) Likelihood robust optimization for data-driven problems. Working paper, Stanford University, Stanford, CA.

Wei C, Li Y, Cai X (2011) Robust optimal policies of production and inventory with uncertain returns and demand. *Internat. J. Production Econom.* 134(2):357–367.

Wiesemann W, Kuhn D, Sim M (2014) Distributionally robust convex optimization. *Oper. Res.* 62(6):1358–1376.

Xu H, Caramanis C, Mannor S (2009) Robustness and regularization of support vector machines. *J. Machine Learning Res.* 10:1485–1510.

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.

Amir Ardestani-Jaafari is a Ph.D. student in the Department of Decision Sciences at HEC Montréal. His current research interests include robust optimization and stochastic programming especially when applied to logistics and healthcare problems.

Erick Delage is an associate professor in the Department of Decision Sciences at HEC Montréal and chairholder of the Canada Research Chair in decision making under uncertainty. His research interests span the areas of robust and stochastic optimization, decision analysis, applied statistics, and risk management with applications to portfolio optimization, inventory management, and transportation problems.