Les Cahiers du GERAD ISSN: 0711–2440

Linearized robust counterparts of two-stage robust optimization problems with applications in operations management

A. Ardestani-Jaafari,
E. Delage

G–2016–69

September 2016
Second revision: April 2019

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2019
– Bibliothèque et Archives Canada, 2019

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2019
– Library and Archives Canada, 2019

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Linearized robust counterparts of two-stage robust optimization problems with applications in operations management

Amir Ardestani-Jaafari
Erick Delage

GERAD & Département de sciences de la décision, HEC Montréal, Montréal (Québec), Canada, H3T 2A7

amir.ardestani-jaafari@hec.ca
erick.delage@hec.ca

September 2016
Second revision: April 2019
Les Cahiers du GERAD
G–2016–69

Abstract: In this article, we discuss an alternative method for deriving conservative approximation models for two-stage robust optimization problems. The method mainly relies on a linearization scheme employed in bilinear programming, therefore we will say that it gives rise to the "linearized robust counterpart" models. We identify a close relation between this linearized robust counterpart model and the popular affinely adjustable robust counterpart model. We also describe methods of modifying both types of models to make these approximations less conservative. These methods are heavily inspired by the use of valid linear and conic inequalities in the linearization process for bilinear models. We finally demonstrate how to employ this new scheme in location-transportation and multi-item newsvendor problems to improve the numerical efficiency and performance guarantees of robust optimization.

Keywords: Two-stage adjustable robust optimization, affinely adjustable robust counterpart, linear programming relaxation, bilinear programming

1 Introduction

Classical robust optimization (RO) assumes that all decisions are here-and-now, i.e., they must be made before the realization of uncertainty. However, this assumption is not realistic in many real-world problems. Take for instance a location-transportation problem (as studied in Atamtürk and Zhang (2007), Zeng and Zhao (2013), Bertsimas and de Ruiter (2016)) for which, while the locations of the production/storage facilities need to be decided as quickly as possible, the decision of how many goods to transport to each customer can be delayed until their respective uncertain demand is revealed. To address the uncertainty in such problems, Ben-Tal et al. (2004) introduced an adjustable robust optimization (ARO) problem that takes the following form in a two-stage setting where the uncertainty is limited to the right-hand side of the constraint set:

$$\text{(ARO)} \quad \underset{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y}(\boldsymbol{\zeta})}{\text{maximize}} \quad \min_{\boldsymbol{\zeta}\in\mathcal{U}} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top \boldsymbol{y}(\boldsymbol{\zeta}) \tag{1a}$$

$$\text{subject to} \quad A\boldsymbol{x} + B\boldsymbol{y}(\boldsymbol{\zeta}) \le \Psi(\boldsymbol{x})\boldsymbol{\zeta}\,,\ \forall\,\boldsymbol{\zeta}\in\mathcal{U}\,, \tag{1b}$$

where $\boldsymbol{x} \in \mathbb{R}^{n_x}$ identifies decisions that must be made "here and now" within a feasible set $\mathcal{X}$ while $\boldsymbol{y} : \mathbb{R}^{n_\zeta} \to \mathbb{R}^{n_y}$ identifies decisions that can adapt to the realization of the uncertain vector of parameters $\boldsymbol{\zeta} \subset \mathbb{R}^{n_\zeta}$, with $\mathcal{U}$ as the "uncertainty set". Furthermore, the ARO model will have $A \in \mathbb{R}^{\bar{m}\times n_x}$, $B \in \mathbb{R}^{\bar{m}\times n_y}$, $\boldsymbol{c} \in \mathbb{R}^{n_x}$, $\boldsymbol{d} \in \mathbb{R}^{n_y}$, and $\Psi : \mathbb{R}^{n_x} \to \mathbb{R}^{\bar{m}\times n_\zeta}$ such that $\Psi(\boldsymbol{x})$ is an affine function of variable $\boldsymbol{x}$. Since finding a solution to problem (1) is computationally intractable, Ben-Tal et al. suggested instead solving the affinely adjustable robust counterpart (AARC) of the problem, wherein adjustable decisions are forced to be affine functions of the observed uncertain vector $\boldsymbol{\zeta}$, i.e., the so called affine decision rule $\boldsymbol{y}(\boldsymbol{\zeta}) := \boldsymbol{y} + Y\boldsymbol{\zeta}$, for some $Y \in \mathbb{R}^{n_y\times n_\zeta}$ and $\boldsymbol{y} \in \mathbb{R}^{n_y}$; therefore, problem (1) is conservatively approximated with

$$\text{(AARC)} \quad \underset{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y},Y}{\text{maximize}} \quad \min_{\boldsymbol{\zeta}\in\mathcal{U}} \ \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top (\boldsymbol{y} + Y\boldsymbol{\zeta}) \tag{2a}$$

$$\text{subject to} \quad A\boldsymbol{x} + B(\boldsymbol{y} + Y\boldsymbol{\zeta}) \le \Psi(\boldsymbol{x})\boldsymbol{\zeta},\ \forall\,\boldsymbol{\zeta}\in\mathcal{U}. \tag{2b}$$

In recent years, the AARC framework has been successfully applied in a number of fields of practice, such as energy planning (Jabr et al. 2015), production planning (Melamed et al. 2016, Kim and Do Chung 2017), power systems (Li et al. 2015, Lorca and Sun 2017, Dehghan et al. 2017), location-transportation and lot sizing problems (Bertsimas and de Ruiter 2016), reservoir management (Gauvin et al. 2017), and supply chain management (Simchi-Levi et al. 2016, Buhayenko and den Hertog 2017). While it appears that affine decision rules can perform well in many applications, situations can still emerge in which AARC will propose overly conservative solutions. We present the following example of a location-transportation problem to illustrate this point.

Example 1 (Location-Transportation Problem) *Consider the following robust location-transportation problem:*

$$\underset{\boldsymbol{x},\boldsymbol{v},\boldsymbol{y}(\boldsymbol{\zeta})}{\text{maximize}} \quad \min_{\boldsymbol{\zeta}\in\mathcal{U}} \ -\sum_{i\in\mathcal{I}} \left(cx_i + kv_i\right) + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \eta_{ij} y_{ij}(\boldsymbol{\zeta}) \tag{3a}$$

$$\text{subject to} \quad \sum_{i\in\mathcal{I}} y_{ij}(\boldsymbol{\zeta}) \le \zeta_j\,,\ \forall\,j\in\mathcal{J}\,,\ \forall\,\boldsymbol{\zeta}\in\mathcal{U} \tag{3b}$$

$$\sum_{j\in\mathcal{J}} y_{ij}(\boldsymbol{\zeta}) \le x_i\,,\ \forall\,i\in\mathcal{I}\,,\ \forall\,\boldsymbol{\zeta}\in\mathcal{U} \tag{3c}$$

$$\boldsymbol{y}(\boldsymbol{\zeta}) \ge 0\,,\ \forall\,\boldsymbol{\zeta}\in\mathcal{U} \tag{3d}$$

$$0 \le x_i \le Mv_i\,,\ \forall\,i\in\mathcal{I} \tag{3e}$$

$$v_i \in \{0,1\}\,,\ \forall\,i\in\mathcal{I}\,. \tag{3f}$$

In this problem, variable v_i indicates whether one opens a facility at location i for each $i \in \mathcal{I} := \{1,\ldots,m\}$, variable x_i denotes the production capacity of the facility i, and variable y_{ij} denotes how many goods are shipped from facility i to customers at location j, with $j \in \mathcal{J} := \{1,\ldots,n\}$. The demand for location j is characterized by ζ_j. Parameter $\eta_{ij} > 0$ denotes the unit revenue of goods shipped from facility i to *customer j, while c and k denote variable and fixed capacity cost for each facility respectively. Let us now consider a special case with 2 facility and 3 customer locations. Specifically, we will let the parameters c, k, and M be respectively equal to 0.6, 100,000, and 100,000, while the matrix η is defined as $\eta = \begin{bmatrix} 5.9 & 5.6 & 4.9 \\ 5.6 & 5.9 & 4.9 \end{bmatrix}$. Furthermore, the uncertain demand ζ is assumed to lie in the following budgeted uncertainty set $\mathcal{U}$:*

$$\mathcal{U} = \{\boldsymbol{\zeta} \in \mathbb{R}^n \big| \exists\, \boldsymbol{\delta} \in [0,1]^n,\, \zeta_j = \bar{\zeta}_j - \hat{\zeta}_j \delta_j\,,\, \forall\, j \in \mathcal{J}\,,\, \sum_{j\in\mathcal{J}} \delta_j \leq \Gamma\}\,,$$

where the nominal demand, $\bar{\zeta}_j$, is set to 20,000 while the maximum deviation $\hat{\zeta}_j := 18,000$, for all j, and where the budget $\Gamma := 2$. In this example, one can numerically establish that the most robust solution is to open only one facility, e.g. at first location, with a capacity of 24,000 units in order to achieve in the worst-case a profit of 6,600 while under the nominal demand the profit reaches as much as 10,600. Yet, the AARC model is unable to identify this solution and instead conservatively prescribes not to open any facility in this region. We refer the reader to Appendix A for additional details on distinction between the two approaches in this problem.

In this paper, we introduce a new scheme for constructing conservative approximation models, named Linearized Robust Counterpart (LRC) models, of ARO problems as presented in (1). While similar linearized models have been previously proposed for a robust surgery block allocation problem in Denton et al. (2010), a robust multi-item newsvendor problem in Ardestani-Jaafari and Delage (2016), and more recently for the robust validation of a network design in Chang et al. (2017)[1], we present for the first time a comprehensive investigation of how it can be employed on a general ARO problem with both polyhedral and non-polyhedral uncertainty sets. In doing so, we offer the following contributions:

- Our proposed scheme allows us to re-interpret AARC as the model that is obtained when applying the linearization scheme found in Sherali and Alameddine (1992) to a certain bilinear program derived from the worst-case analysis (a.k.a. adversarial problem). Through this new lens, we are able to exploit valid linear and conic inequalities that are used in bilinear programs to identify tractable conservative approximations (called MLRC and SDP-LRC) that always improve (in some case strictly) upon AARC.
- We establish for the first time that affine decision rules can be applied on a modified version of the ARO problem where constraint violation is allowed but only if the right penalty is paid, namely through:

$$\text{(Penalized ARO)} \quad \underset{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y}(\boldsymbol{\zeta}),\boldsymbol{z}(\boldsymbol{\zeta})}{\text{maximize}} \quad \min_{\boldsymbol{\zeta}\in\mathcal{U}} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top \boldsymbol{y}(\boldsymbol{\zeta}) - \boldsymbol{u}^\top \boldsymbol{z}(\boldsymbol{\zeta})$$
$$\text{subject to} \quad A\boldsymbol{x} + B\boldsymbol{y}(\boldsymbol{\zeta}) \leq \Psi(\boldsymbol{x})\boldsymbol{\zeta} + \boldsymbol{z}(\boldsymbol{\zeta})\,,\, \forall\, \boldsymbol{\zeta} \in \mathcal{U}\,,$$

  where $\boldsymbol{z} : \mathbb{R}^{n_\zeta} \to \mathbb{R}^{\bar{m}}$ is an adjustable variable that measures the amount of violation while $\boldsymbol{u} \in \mathbb{R}^{\bar{m}}$ is a vector of "large enough" (to be clarified later) marginal violation penalties. While the Penalized ARO model is equivalent to ARO in terms of optimal value and set of optimal solutions for $\boldsymbol{x}$, the conservative approximation model obtained when using affine decision rules (i.e., MLRC) is necessarily of better quality (in terms of first stage decisions) than AARC and guaranteed to be feasible even when AARC fails to be feasible.
- We present an extensive set of numerical experiments performed on location-transportation problems that provide evidence that MLRC can significantly improve the quality of first stage decisions when the problem suffers from mid-sized to large potential perturbations. This improvement nearly comes at no cost in terms of computation time when compared to the approach proposed in Ardestani-Jaafari and Delage (2017).
- We present an extensive set of numerical experiments performed on multi-item newsvendor problems providing further evidence that adding conic inequalities to LRC can significantly improve the quality

[1] To the best of our knowledge, the work of Chang et al. (2017) was developed independently from the results presented in this paper and mostly focuses its analysis on the network design application. For this application, the authors claim (without a formal proof) that LRC leads to a tighter conservative approximation than AARC leaving the question open of whether the two approaches are equivalent. They also hint at the possibility of generalizing their approach to the ARO model, which is done in full detail for the first time in this paper.

of the bounds on optimal worst-case profit when the demand of items is correlated or suffers from large potential perturbations. This improvement also comes at a reasonable cost in terms of computation time.

The remainder of the paper is organized as follows. In Section 2, we introduce the linearized robust counterpart model associated to an ARO problem with a polyhedral uncertainty set and establish the equivalence between the LRC and AARC models. Section 3 describes the two methods that can be used to tighten the LRC/AARC approximation. Section 4 briefly describes how one extends our results to general convex uncertainty sets. An exhaustive set of numerical experiments involving the location-transportation problem are presented in Section 5, and involving the multi-item newsvendor problem in Section 6. Finally, we conclude in Section 7 and discuss in Appendix G interesting connections to recent literature that proposes copositive programming reformulations to the ARO problem. In particular, we show for the first time how exact copositive programming reformulations are related to AARC when ARO satisfies the relatively complete recourse property.

2 The Linearized Robust Counterpart model

In order to present the LRC model, we need to make the following three assumptions.

Assumption 1 *Let $\mathcal{U}$ be a bounded and non-empty polyhedral set defined as $\mathcal{U} := \{\zeta | P\boldsymbol{\zeta} \leq \boldsymbol{q}\}$ where $P \in \mathbb{R}^{n_\mathcal{U} \times n_\zeta}, \boldsymbol{q} \in \mathbb{R}^{n_\mathcal{U}}$.*

Assumption 2 *Let the ARO model possess relatively complete recourse, namely, that*

$$\forall\, \boldsymbol{x} \in \mathcal{X}\,,\, \exists\, \boldsymbol{y}(\boldsymbol{\zeta}):\; A\boldsymbol{x} + B\boldsymbol{y}(\boldsymbol{\zeta}) \leq \Psi(\boldsymbol{x})\boldsymbol{\zeta} \quad \forall\, \boldsymbol{\zeta} \in \mathcal{U}\,.$$

Assumption 3 *For all $\boldsymbol{x} \in \mathcal{X}$, there exists a feasible $\boldsymbol{\zeta}$, such that the recourse problem is bounded. In other words, let problem (1) be bounded above.*

The three assumptions described above should not be considered limiting. Considering Assumption 1, it is typically the case that $\mathcal{U}$ includes, at very least, a nominal, or most-likely, scenario, and that all possible scenarios reside in a bounded set. Satisfying Assumption 2 is mostly a matter of formulating $\mathcal{X}$ so that it does not include any solutions for which there might be no feasible second-stage solutions. Finally, it is reasonable to assume that problem (1) is bounded above in realistic practical problems.

Let us now consider the fact that the ARO model can be formulated as

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{maximize}} \qquad g(\boldsymbol{x}) \tag{4}$$

where $g(\boldsymbol{x})$ is defined as

$$\begin{aligned} g(\boldsymbol{x}) := \min_{\boldsymbol{\zeta}\in\mathcal{U}} \quad \max_{\boldsymbol{y}} \qquad & \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top \boldsymbol{y} && \text{(5a)}\\ \text{subject to} \qquad & A\boldsymbol{x} + B\boldsymbol{y} \leq \Psi(\boldsymbol{x})\boldsymbol{\zeta}\,. && \text{(5b)} \end{aligned}$$

Based on Assumption 2, one can apply duality theory on the inner maximization problem to show that $g(\boldsymbol{x})$ is exactly equal to

$$\begin{aligned} g(\boldsymbol{x}) = \min_{\boldsymbol{\zeta},\boldsymbol{\lambda}} \qquad & \boldsymbol{c}^\top \boldsymbol{x} + (\Psi(\boldsymbol{x})\boldsymbol{\zeta})^\top \boldsymbol{\lambda} - (A\boldsymbol{x})^\top \boldsymbol{\lambda} && \text{(6a)}\\ \text{subject to} \qquad & B^\top \boldsymbol{\lambda} = \boldsymbol{d} && \text{(6b)}\\ & P\boldsymbol{\zeta} \leq \boldsymbol{q} && \text{(6c)}\\ & \boldsymbol{\lambda} \geq 0\,, && \text{(6d)} \end{aligned}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^{\bar{m}}$ is the dual variable associated with constraint (5b). Note that problem (6) is known as a bilinear program given that the objective is linear in both $\boldsymbol{\zeta}$ and $\boldsymbol{\lambda}$ when the other is considered fixed.

Lemma 1 *Problem (6) possesses a feasible solution and its optimal value is finite.*

Proof. Assumption 3 guarantees that, for all $\boldsymbol{x} \in \mathcal{X}$, there exists a feasible $\bar{\boldsymbol{\zeta}}$ for which the maximization problem in $\boldsymbol{y}$ has a finite optimal value. By the strong duality property, this indicates that, for this same $\bar{\boldsymbol{\zeta}}$, the minimization problem in $\boldsymbol{\lambda}$ also has a finite optimal value and must therefore have a feasible solution $\bar{\boldsymbol{\lambda}}$. Together, the pair $(\bar{\boldsymbol{\zeta}}, \bar{\boldsymbol{\lambda}})$ constitutes a feasible solution for problem (6). □

In Sherali and Alameddine (1992), the authors propose a so-called Reformulation-Linearization Technique (RLT) that produced linearized versions of bilinear programs. The RLT consists of two sequential phases. The first phase adds a list of valid quadratic inequalities to the problem using pairwise products of inequality constraints, or products of equality constraints with decision variables. The second phase consists in linearizing the problem by replacing all products of decision variables with new decision variables. This produces a linear programming relaxation that is tightened by the valid inequalities that were initially introduced.

In the context that we study here, the first phase leads us to constructing the following constraints

$$
\begin{aligned}
&(6b) \times \boldsymbol{\zeta}^\top \Rightarrow (B^\top \boldsymbol{\lambda} - \boldsymbol{d})\boldsymbol{\zeta}^\top = 0 \Rightarrow \boldsymbol{\zeta}\boldsymbol{\lambda}^\top B = \boldsymbol{\zeta}\boldsymbol{d}^\top ,\\
&(6c) \times (6d) \Rightarrow (\boldsymbol{q} - P\boldsymbol{\zeta})\boldsymbol{\lambda}^\top \geq 0 \Rightarrow P\boldsymbol{\zeta}\boldsymbol{\lambda}^\top \leq \boldsymbol{q}\boldsymbol{\lambda}^\top ,\\
&(6b) \times \boldsymbol{\lambda}^\top \Rightarrow (B^\top \boldsymbol{\lambda} - \boldsymbol{d})\boldsymbol{\lambda}^\top = 0 \Rightarrow \boldsymbol{\lambda}\boldsymbol{\lambda}^\top B = \boldsymbol{\lambda}\boldsymbol{d}^\top ,\\
&(6c) \times (6c) \Rightarrow (\boldsymbol{q} - P\boldsymbol{\zeta})(\boldsymbol{q} - P\boldsymbol{\zeta})^\top \geq 0 \Rightarrow P\boldsymbol{\zeta}\boldsymbol{\zeta}^\top P^\top + \boldsymbol{q}\boldsymbol{q}^\top \geq P\boldsymbol{\zeta}\boldsymbol{q}^\top + \boldsymbol{q}\boldsymbol{\zeta}^\top P^\top ,\\
&(6d) \times (6d) \Rightarrow \boldsymbol{\lambda}\boldsymbol{\lambda}^\top \geq 0 .
\end{aligned}
$$

and to adding them to problem (6) as

$$
\begin{aligned}
g(\boldsymbol{x}) = \min_{\boldsymbol{\zeta},\boldsymbol{\lambda}} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \mathrm{tr}(\Psi(\boldsymbol{x})\boldsymbol{\zeta}\boldsymbol{\lambda}^\top) - (A\boldsymbol{x})^\top \boldsymbol{\lambda} && \text{(7a)}\\
\text{subject to} \quad & (6b) - (6d) \\
& \boldsymbol{\zeta}\boldsymbol{\lambda}^\top B = \boldsymbol{\zeta}\boldsymbol{d}^\top && \text{(7b)}\\
& P\boldsymbol{\zeta}\boldsymbol{\lambda}^\top \leq \boldsymbol{q}\boldsymbol{\lambda}^\top && \text{(7c)}\\
& \boldsymbol{\lambda}\boldsymbol{\lambda}^\top B = \boldsymbol{\lambda}\boldsymbol{d}^\top && \text{(7d)}\\
& P\boldsymbol{\zeta}\boldsymbol{\zeta}^\top P^\top + \boldsymbol{q}\boldsymbol{q}^\top \geq P\boldsymbol{\zeta}\boldsymbol{q}^\top + \boldsymbol{q}\boldsymbol{\zeta}^\top P^\top && \text{(7e)}\\
& \boldsymbol{\lambda}\boldsymbol{\lambda}^\top \geq 0 , && \text{(7f)}
\end{aligned}
$$

where $\mathrm{tr}(\cdot)$ stands for the trace operator, and where, for any two matrices A and B of the same dimension $n \times m$, a constraint $A \leq B$ stands for $A_{ij} \leq B_{ij}$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$, and similarly for the constraint $A = B$. The second phase of RLT is linearization. To do so, we introduce in problem (7) the variables $\Delta \in \mathbb{R}^{n_\zeta \times \bar{m}}$, $\Xi \in \mathbb{R}^{n_\zeta \times n_\zeta}$, and $\Lambda \in \mathbb{R}^{\bar{m} \times \bar{m}}$, respectively defined as $\Delta := \boldsymbol{\zeta}\boldsymbol{\lambda}^\top$, $\Xi := \boldsymbol{\zeta}\boldsymbol{\zeta}^\top$, and $\Lambda := \boldsymbol{\lambda}\boldsymbol{\lambda}^\top$, such that

$$
\begin{aligned}
g(\boldsymbol{x}) = \min_{\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta,\Lambda,\Xi} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \mathrm{tr}(\Psi(\boldsymbol{x})\Delta) - (A\boldsymbol{x})^\top \boldsymbol{\lambda} && \text{(8a)}\\
\text{subject to} \quad & (6b) - (6d) \\
& \Delta B = \boldsymbol{\zeta}\boldsymbol{d}^\top && \text{(8b)}\\
& P\Delta \leq \boldsymbol{q}\boldsymbol{\lambda}^\top && \text{(8c)}\\
& \Lambda B = \boldsymbol{\lambda}\boldsymbol{d}^\top && \text{(8d)}\\
& \Lambda \geq 0 && \text{(8e)}\\
& P\Xi P^\top + \boldsymbol{q}\boldsymbol{q}^\top \geq P\boldsymbol{\zeta}\boldsymbol{q}^\top + \boldsymbol{q}\boldsymbol{\zeta}^\top P^\top && \text{(8f)}\\
& \Lambda = \boldsymbol{\lambda}\boldsymbol{\lambda}^\top && \text{(8g)}\\
& \Xi = \boldsymbol{\zeta}\boldsymbol{\zeta}^\top && \text{(8h)}\\
& \Delta = \boldsymbol{\zeta}\boldsymbol{\lambda}^\top . && \text{(8i)}
\end{aligned}
$$

A simple relaxation of constraints (8g)–(8i) will lead to a conservative approximation model of problem (1) which we refer to as the linearized robust counterpart (LRC) model.

Proposition 1 *The following linearized robust counterpart model is a conservative approximation of problem* (1)*:*

$$\text{(LRC)} \quad \underset{\boldsymbol{x}\in\mathcal{X},Y,\boldsymbol{y},S,\boldsymbol{s}}{\text{maximize}} \quad \boldsymbol{c}^\top\boldsymbol{x} + \boldsymbol{d}^\top\boldsymbol{y} - \boldsymbol{q}^\top\boldsymbol{s} \tag{9a}$$

$$\text{subject to} \quad P^\top S = Y^\top B^\top - \Psi(\boldsymbol{x})^\top \tag{9b}$$

$$A\boldsymbol{x} + B\boldsymbol{y} + S^\top\boldsymbol{q} \leq 0 \tag{9c}$$

$$P^\top\boldsymbol{s} = -Y^\top\boldsymbol{d} \tag{9d}$$

$$\boldsymbol{s} \geq 0\,,\; S \geq 0\,, \tag{9e}$$

where $Y \in \mathbb{R}^{n_y \times n_\zeta}, \boldsymbol{y} \in \mathbb{R}^{n_y}, S \in \mathbb{R}^{n_\mathcal{U} \times \bar{m}}$*, and* $\boldsymbol{s} \in \mathbb{R}^{n_\mathcal{U}}$*.*

Proof. First, we relax problem (8) by removing constraints (8g)–(8i) to get a lower bound for $g(\boldsymbol{x})$. Next, we consider that since, when constraints (6b)–(6d), and (8c) are satisfied, one can simply let $\hat{\Lambda} := \hat{\boldsymbol{\lambda}}\hat{\boldsymbol{\lambda}}^\top$ in order to satisfy constraints (8d) and (8e), the problem stays the same when disregarding Λ and these two constraints. In a similar way, one can disregard Ξ and constraint (8f). Hence, we obtain a lower bound for $g(\boldsymbol{x})$ in the form

$$g(\boldsymbol{x}) \geq g_{LRC}(\boldsymbol{x}) := \min_{\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta} \quad \boldsymbol{c}^\top\boldsymbol{x} + \operatorname{tr}(\Psi(\boldsymbol{x})\Delta) - (A\boldsymbol{x})^\top\boldsymbol{\lambda} \tag{10a}$$

$$\text{subject to} \quad B^\top\boldsymbol{\lambda} = \boldsymbol{d} \tag{10b}$$

$$P\boldsymbol{\zeta} \leq \boldsymbol{q} \tag{10c}$$

$$\boldsymbol{\lambda} \geq 0 \tag{10d}$$

$$\Delta B = \boldsymbol{\zeta}\boldsymbol{d}^\top \tag{10e}$$

$$P\Delta \leq \boldsymbol{q}\boldsymbol{\lambda}^\top\,. \tag{10f}$$

Since, based on Lemma 1, there exists a solution $(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\lambda}})$ that satisfies constraints (10b), (10c), and (10d), one can confirm that the triplet $(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\lambda}}, \hat{\Delta})$, with $\hat{\Delta} := \hat{\boldsymbol{\zeta}}\hat{\boldsymbol{\lambda}}^\top$, is a feasible solution to problem (10). Hence, strong duality applies for problem (10) so that it can be equivalently represented as

$$g_{LRC}(\boldsymbol{x}) = \max_{Y,\boldsymbol{y},S,\boldsymbol{s}} \quad \boldsymbol{c}^\top\boldsymbol{x} + \boldsymbol{d}^\top\boldsymbol{y} - \boldsymbol{q}^\top\boldsymbol{s} \tag{11a}$$

$$\text{subject to} \quad P^\top S = Y^\top B^\top - \Psi(\boldsymbol{x})^\top \tag{11b}$$

$$A\boldsymbol{x} + B\boldsymbol{y} + S^\top\boldsymbol{q} \leq 0 \tag{11c}$$

$$-P^\top\boldsymbol{s} = Y^\top\boldsymbol{d} \tag{11d}$$

$$\boldsymbol{s} \geq 0\,,\; S \geq 0\,, \tag{11e}$$

where the variables $\boldsymbol{y} \in \mathbb{R}^{n_y}$, $\boldsymbol{s} \in \mathbb{R}^{n_\mathcal{U}}$, $Y \in \mathbb{R}^{n_y \times n_\zeta}$, and $S \in \mathbb{R}^{n_\mathcal{U} \times \bar{m}}$ are the dual variables associated with constraints (10b), (10c), (10e), and (10f) respectively. We next combine problem (11) with the maximization over variable $\boldsymbol{x} \in \mathcal{X}$, which leads to LRC (9). □

Given that both LRC and AARC give rise to conservative approximation that take the form of linear programming models when $\mathcal{X}$ is polyhedral, one might wonder how different the solutions obtained from both approaches are. In fact, we conclude this section by demonstrating that the LRC and AARC models can be considered equivalent to each other. This equivalence will in turn be used in Section 3 to identify ways of modifying LRC (and AARC) to improve the quality of these approximation schemes.

Proposition 2 *LRC* (9) *is equivalent to AARC* (2)*. Namely, the optimal value of problems* (2) *and* (9) *are the same, and so is the set of optimal solutions for the variables* $(\boldsymbol{x}, \boldsymbol{y}, Y)$*.*

Proof. As a first step, we reformulate problem (10) as presented in problem (11) where we isolate the role of $(\boldsymbol{s}, S)$ from $(\boldsymbol{y}, Y)$:

$$g_{LRC}(\boldsymbol{x}) = \max_{\boldsymbol{x},\boldsymbol{y},Y} h(\boldsymbol{x},\boldsymbol{y},Y)\,,$$

where

$$\begin{aligned} h(\boldsymbol{x},\boldsymbol{y},Y) := \max_{S,\boldsymbol{s}} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top \boldsymbol{y} - \boldsymbol{q}^\top \boldsymbol{s} && \text{(12a)}\\ \text{subject to} \quad & P^\top S = Y^\top B^\top - \Psi(\boldsymbol{x})^\top && \text{(12b)}\\ & A\boldsymbol{x} + B\boldsymbol{y} + S^\top \boldsymbol{q} \leq 0 && \text{(12c)}\\ & -P^\top \boldsymbol{s} = Y^\top \boldsymbol{d} && \text{(12d)}\\ & \boldsymbol{s} \geq 0\,,\; S \geq 0\,. && \text{(12e)} \end{aligned}$$

Noting that, we interpret the value of $h(\boldsymbol{x},\boldsymbol{y},Y)$ as returning minus infinity when problem (12) is infeasible. Applying LP duality to the maximization in $\boldsymbol{s}$ we obtain that:

$$\begin{aligned} h(\boldsymbol{x},\boldsymbol{y},Y) = \max_{S} \quad & \min_{\boldsymbol{\zeta}\in\mathcal{U}} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top(\boldsymbol{y} + Y\boldsymbol{\zeta}) && \text{(13a)}\\ \text{subject to} \quad & A\boldsymbol{x} + B\boldsymbol{y} + S^\top \boldsymbol{q} \leq 0 && \text{(13b)}\\ & P^\top S = Y^\top B^\top - \Psi(\boldsymbol{x})^\top && \text{(13c)}\\ & S \geq 0\,. && \text{(13d)} \end{aligned}$$

One can confirm that strong duality applies given that the minimization in $\boldsymbol{\zeta}$ is feasible based on Assumption 1.

Considering that S is not in the objective function of problem (13), we can conclude that $h(\boldsymbol{x},\boldsymbol{y},Y) = \min_{\boldsymbol{\zeta}\in\mathcal{U}} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top(\boldsymbol{y} + Y\boldsymbol{\zeta})$ as long as, for all $i = 1,\ldots,\bar{m}$, there exists a vector $S_{:i}$ (capturing the i-th column of S) that satisfies the constraints:

$$\begin{aligned} & A_{i:}\boldsymbol{x} + B_{i:}\boldsymbol{w} + (S_{:i})^\top \boldsymbol{q} \leq 0 && \text{(14a)}\\ & P^\top S_{:i} = Y^\top (B_{i:})^\top - (\Psi(\boldsymbol{x})_{i:})^\top && \text{(14b)}\\ & S_{:i} \geq 0\,, && \text{(14c)} \end{aligned}$$

where $A_{i:}$, $B_{i:}$, and $\Psi(\boldsymbol{x})_{i:}$ denote the i^{th} row of matrices A, B, and $\Psi(\boldsymbol{x})$ respectively. This is equivalent to verifying that:

$$\begin{aligned} \min_{S_{:i}\geq 0} \quad & A_{i:}\boldsymbol{x} + B_{i:}\boldsymbol{w} + (S_{:i})^\top \boldsymbol{q} \qquad \leq 0 && \text{(15a)}\\ \text{subject to} \quad & P^\top S_{:i} = Y^\top (B_{i:})^\top - (\Psi(\boldsymbol{x})_{i:})^\top\,, && \text{(15b)} \end{aligned}$$

which can be replaced by a maximization problem, using duality theory. This leads us to considering constraint (15) as equivalent to

$$\begin{aligned} \max_{\boldsymbol{\zeta}} \quad & A_{i:}\boldsymbol{x} + B_{i:}\boldsymbol{w} + (Y^\top (B_{i:})^\top - (\Psi(\boldsymbol{x})_{i:})^\top)^\top \boldsymbol{\zeta} \qquad \leq 0\\ \text{subject to} \quad & P\zeta \leq \boldsymbol{q}\,, \end{aligned}$$

with ζ the dual variable of (15b) and where strong duality follows from the same argument as before. We have thus confirmed that

$$h(\boldsymbol{x},\boldsymbol{y},Y) = \begin{cases} \min_{\boldsymbol{\zeta}\in\mathcal{U}} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top(\boldsymbol{y} + Y\boldsymbol{\zeta}) & \text{if } A\boldsymbol{x} + B(Y\boldsymbol{\zeta} + \boldsymbol{y}) \leq \Psi(\boldsymbol{x})\boldsymbol{\zeta},\ \forall\,\boldsymbol{\zeta}\in\mathcal{U} \\ -\infty & \text{otherwise} \end{cases}\,,$$

which is the function that is optimized in AARC model (2). Therefore, the LRC model (9) produces the same optimal value as the optimal value of AARC, and additionally has the same set of of optimal solutions for $(\boldsymbol{x},\boldsymbol{y},Y)$.

□

3 Improving LRC and AARC using valid Inequalities

In this section, we identify two types of valid inequalities that can be employed to formulate new versions of LRC models that provide tighter conservative approximations. First, we will make use of valid linear inequalities that can be derived from an implicit upper bound on the optimal solution for $\boldsymbol{\lambda}$ in problem (6). This process will lead to a modified LRC model that also takes the form of a linear program when $\mathcal{X}$ is polyhedral while guaranteeing the feasibility of the resulting approximation model. Secondly, we will identify a set of valid conic inequalities that will lead to semi-definite programming formulations for LRC.

3.1 Exploiting an implicit bound on $\boldsymbol{\lambda}^*$

We start with a reasonable assumption that can be used to generate helpful valid inequalities for problem (6) and obtain our Modified LRC (MLRC) model.

Assumption 4 *One can identify a bounding vector $\boldsymbol{u} \in \mathbb{R}^{\bar{m}}$ such that, for all $\boldsymbol{x} \in \mathcal{X}$ and for all $\boldsymbol{\zeta} \in \mathcal{U}$, there exists an optimal solution $\boldsymbol{\lambda}^* \leq \boldsymbol{u}$ for the problem*

$$\begin{aligned} \underset{\boldsymbol{\lambda}}{\text{minimize}} \quad & (\Psi(\boldsymbol{x})\boldsymbol{\zeta} - A\boldsymbol{x})^\top \boldsymbol{\lambda} && \text{(16a)} \\ \text{subject to} \quad & B^\top \boldsymbol{\lambda} = \boldsymbol{d} && \text{(16b)} \\ & \boldsymbol{\lambda} \geq 0\,. && \text{(16c)} \end{aligned}$$

At first glance, one might consider that Assumption 4 is rather limiting given that $\boldsymbol{\lambda}$ is an abstract object. One should, however, be aware that such a bound $\boldsymbol{u}$ is guaranteed to exist. Indeed, based on Lemma 1 there always exists a finite optimal solution for problem (16) and, in particular, one that lies at one of the vertices of the feasible polyhedron. Since the number of vertices of a polyhedra defined by a finite number of constraints is finite, we must conclude that such a $\boldsymbol{u}$ exists. A challenge remains in identifying the tightest values possible for $\boldsymbol{u}$. For this reason, we propose a numerical procedure that can be used to identify tighter bounds once an initial gross estimate has been found. We later in Section 5 and in Appendix B investigate two practical examples where such initial estimates can be found analytically.

Proposition 3 *Given that the columns of B are linearly independent and an initial bound M such that $\boldsymbol{u} := M$ satisfies Assumption 4, for any fixed $k = 1, 2, \ldots, \bar{m}$, one can identify a tighter bound u_k^* that satisfies Assumption 4 by solving the following mixed-integer linear program*

$$\begin{aligned} u_k^* := \max_{\boldsymbol{\lambda}, \boldsymbol{v}} \quad & \boldsymbol{e}_k^\top \boldsymbol{\lambda} && \text{(17a)} \\ \text{subject to} \quad & B^\top \boldsymbol{\lambda} = \boldsymbol{d} && \text{(17b)} \\ & \lambda_i \leq M(1 - v_i)\,,\ \forall\, i && \text{(17c)} \\ & \sum_{i=1}^{\bar{m}} v_i + n_y = \bar{m} && \text{(17d)} \\ & \sum_{i \in \mathcal{I}} v_i \leq |\mathcal{I}| - 1\,,\ \forall \mathcal{I} \in \mathfrak{I}(B) && \text{(17e)} \\ & \boldsymbol{\lambda} \geq 0\,,\ \boldsymbol{v} \in \{0, 1\}^{\bar{m}}\,, && \text{(17f)} \end{aligned}$$

where $\boldsymbol{e}_k$ is the k-th column of the identity matrix while $\mathfrak{I}(B) := \{\mathcal{I} \subseteq \{1, \ldots, \bar{m}\} \,|\, \{B_{:1}, \ldots, B_{:n_y}\} \cup \{\boldsymbol{e}_i\}_{i \in \mathcal{I}}$ contains linearly dependent vectors$\}$.

Proof. For any fixed $\boldsymbol{x} \in \mathcal{X}$ and $\boldsymbol{\zeta} \in \mathcal{U}$, given that problem (16) is known to be feasible and finite (see Lemma 1), it must be the case that an optimal solution $\boldsymbol{\lambda}^*$ for the problem is located at one of the vertices of the polyhedra defined by $B^\top \boldsymbol{\lambda} = \boldsymbol{d}$ and $\boldsymbol{\lambda} \geq 0$. By definition, for any vertex $\boldsymbol{\lambda}^\nu$ of this polyhedra, there must exists a set of indices $\mathcal{I} \subseteq \{1, \ldots, \bar{m}\}$ such that:

1. $|\mathcal{I}| = \bar{m} - n_y$
2. the set $\{B_{:1}, \ldots, B_{:n_y}\} \cup \{\boldsymbol{e}_i\}_{i\in\mathcal{I}}$ contains linearly independent vectors
3. $\boldsymbol{\lambda}^\nu$ is the unique solution to the following system of equations

$$B^\top \boldsymbol{\lambda}^\nu = \boldsymbol{d} \qquad\qquad \boldsymbol{e}_i^\top \boldsymbol{\lambda}^\nu = 0\,,\, \forall\, i \in \mathcal{I}\,.$$

It therefore follows that if $\|\boldsymbol{\lambda}^\nu\|_\infty \leq M$ for all the vertex v of this polyhedron, then problem (17) necessarily returns an optimal value u_k^* that satisfies $\boldsymbol{\lambda}_k^* \leq \max_\nu \boldsymbol{\lambda}_k^\nu = u_k^*$. □

Note that the MILP presented in (17) has a potentially large number of constraints. It is however possible to identify an initial bound for u_k^* by relaxing constraint (17e). One can then further improve on this bound by progressively generating the constraints in (17e) that are violated by solutions of the relaxed problem. In particular, verifying whether a particular assignment for $\boldsymbol{v}$ satisfies constraint (17e) reduces to verifying whether the matrix composed by the vectors $\{B_{:1}, \ldots, B_{:n_y}\} \cup \{\boldsymbol{e}_i\}_{i:v_i=1}$ is full rank.

Proposition 4 *Given Assumption 4, the following modified linearized robust counterpart model is a conservative approximation to problem (1):*

$$\begin{aligned}
\text{(MLRC)} \quad & \underset{\boldsymbol{x}\in\mathcal{X},Y,\boldsymbol{y},S,\boldsymbol{s},W,\boldsymbol{w}}{\text{maximize}} && \boldsymbol{c}^\top\boldsymbol{x} + \boldsymbol{d}^\top\boldsymbol{y} - \boldsymbol{q}^\top\boldsymbol{s} - \boldsymbol{u}^\top\boldsymbol{w} - \boldsymbol{u}^\top W\boldsymbol{q} && \text{(18a)}\\
& \text{subject to} && Y^\top B^\top - P^\top(S - W) = \Psi(\boldsymbol{x})^\top && \text{(18b)}\\
&&& A\boldsymbol{x} + B\boldsymbol{y} + (S-W)^\top\boldsymbol{q} - \boldsymbol{w} \leq 0 && \text{(18c)}\\
&&& -P^\top(\boldsymbol{s} + W\boldsymbol{u}) = Y^\top\boldsymbol{d} && \text{(18d)}\\
&&& \boldsymbol{s} \geq 0\,,\, S \geq 0\,,\, \boldsymbol{w} \geq 0\,,\, W \geq 0\,, && \text{(18e)}
\end{aligned}$$

where $Y \in \mathbb{R}^{n_y \times n_\zeta}$, $\boldsymbol{y} \in \mathbb{R}^{n_y}$, $S \in \mathbb{R}^{n_\mathcal{U} \times \bar{m}}$, $\boldsymbol{s} \in \mathbb{R}^{n_\mathcal{U}}$, $W \in \mathbb{R}^{n_\mathcal{U} \times \bar{m}}$, and $\boldsymbol{w} \in \mathbb{R}^{\bar{m}}$. Furthermore, the optimal value of problem (18) is necessarily larger or equal than the optimal value of LRC (9), can in some case be strictly larger, and is actually equal when the constraint $\boldsymbol{\lambda} \leq \boldsymbol{u}$ is redundant with respect to the set $\{\boldsymbol{\lambda} \in \mathbb{R}^{\bar{m}} \,|\, \boldsymbol{\lambda} \geq 0,\, B^\top\boldsymbol{\lambda} = \boldsymbol{d}\}$.

Note that Proposition 4 implies that MLRC provides a tighter conservative approximation than AARC while reducing to AARC under the special condition that $\boldsymbol{\lambda} \leq \boldsymbol{u}$ is already implied from the constraints $\boldsymbol{\lambda} \geq 0$ and $B^\top\boldsymbol{\lambda} = \boldsymbol{d}$.

Proof. Given that Assumption 4 is satisfied, the following constraints are valid inequalities for problem (6) in the sense that they can be added to this problem without affecting its optimal value:

$$\boldsymbol{\lambda} \leq \boldsymbol{u} \tag{19}$$

$$(\boldsymbol{q} - P\boldsymbol{\zeta})(\boldsymbol{u} - \boldsymbol{\lambda})^\top \geq 0 \tag{20}$$

where constraint (20) can be linearized by replacing $\Delta := \boldsymbol{\zeta}\boldsymbol{\lambda}^\top$ as

$$P\Delta \geq \boldsymbol{q}\boldsymbol{\lambda}^\top - (\boldsymbol{q} - P\boldsymbol{\zeta})\boldsymbol{u}^\top\,. \tag{21}$$

Adding constraints (19) and (21) to problem (10) leads to the following formulation:

$$\begin{aligned}
g_{MLRC}(\boldsymbol{x}) := \min_{\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta} \quad & \boldsymbol{c}^\top\boldsymbol{x} + \mathrm{tr}(\Psi(\boldsymbol{x})\Delta) - (A\boldsymbol{x})^\top\boldsymbol{\lambda} && \text{(22a)}\\
\text{subject to} \quad & \text{(10b)} - \text{(10f)} &&\\
& \boldsymbol{\lambda} \leq \boldsymbol{u} && \text{(22b)}\\
& P\Delta \geq \boldsymbol{q}\boldsymbol{\lambda}^\top - (\boldsymbol{q} - P\boldsymbol{\zeta})\boldsymbol{u}^\top\,, && \text{(22c)}
\end{aligned}$$

that will lead to MLRC after applying duality theory.

In order to support the claim that MLRC can achieve strictly better bounds than LRC and AARC, we refer the reader to the example of a multi-product assembly problem presented in Appendix B. The numerical evidence presented in Section 5 also further confirms this fact.

Finally, in the case that $\boldsymbol{\lambda} \leq \boldsymbol{u}$ is redundant with respect to the set $\{\boldsymbol{\lambda} \in \mathbb{R}^{\bar{m}} \,|\, \boldsymbol{\lambda} \geq 0,\, B^\top \boldsymbol{\lambda} = \boldsymbol{d}\}$, we can show that $g_{LRC}(\boldsymbol{x}) = g_{MLRC}(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathcal{X}$. Comparing problems (10) and (22), we see that this result can be obtained by showing that constraints (22b) and (22c) are redundant in (22). While the fact that (22b) is redundant follows directly from the condition we are investigating, one needs to pay a closer look at the case of (22c). In particular, for each $i \in \{1, \ldots, \bar{m}\}$, we start by formulating the fact that $\lambda_i \leq u_i$ is redundant as the infeasibility of the following set :

$$\{\boldsymbol{\lambda} \in \mathbb{R}^{\bar{m}} \,|\, B^\top \boldsymbol{\lambda} = \boldsymbol{d}\,,\ \boldsymbol{\lambda} \geq 0\,,\ \lambda_i > u_i\}\,. \tag{23}$$

Given that Lemma 1 tells us that $B^\top \boldsymbol{\lambda} = \boldsymbol{d}\,,\ \boldsymbol{\lambda} \geq 0$ is feasible for some $\bar{\boldsymbol{\lambda}}$, we can show that infeasibility of (23) is equivalent to the infeasibility of the set:

$$\{(\boldsymbol{\lambda}, \gamma) \in \mathbb{R}^{\bar{m}} \times \mathbb{R} \,|\, \boldsymbol{\lambda} \geq 0\,,\ \gamma \geq 0\,,\ B^\top \boldsymbol{\lambda} = \boldsymbol{d}\gamma\,,\ \lambda_i - u_i\gamma > 0\}\,. \tag{24}$$

This is first due to the fact that any member of the set in (24) can be used to construct a member of (23) by setting $\gamma = 1$. Conversely, given any member of the set in (24), either $\gamma > 0$ and one can construct a member of (23) using $\boldsymbol{\lambda}' := \boldsymbol{\lambda}/\gamma$, otherwise $\gamma = 0$ and then there exists an $\alpha > (u_i - \bar{\lambda}_i)/\lambda_i$ such that $\boldsymbol{\lambda}' := \bar{\boldsymbol{\lambda}} + \alpha\boldsymbol{\lambda}$ is in the former set since $B^\top \boldsymbol{\lambda}' = B^\top(\bar{\boldsymbol{\lambda}} + \alpha\boldsymbol{\lambda}) = B^\top \bar{\boldsymbol{\lambda}} = \boldsymbol{d}$ and $\lambda_i' = \bar{\lambda}_i + \alpha\lambda_i > u_i$. We follow this argument by using, for all i, Farkas lemma on the constraints in (24). Hence, if (23) is infeasible for all i then there must exist a matrix $\boldsymbol{V} \in \mathbb{R}^{n_y \times \bar{m}}$ that satisfies $BV \geq I$ and $\boldsymbol{d}^\top V \leq \boldsymbol{u}^\top$. This in turn can be used to establish the following:

$$\begin{aligned}
\boldsymbol{q}\boldsymbol{\lambda}^\top - (\boldsymbol{q} - P\boldsymbol{\zeta})\boldsymbol{u}^\top - P\Delta \ &\leq\ \boldsymbol{q}\boldsymbol{\lambda}^\top - (\boldsymbol{q} - P\boldsymbol{\zeta})\boldsymbol{d}^\top V - P\Delta \\
&=\ \boldsymbol{q}\boldsymbol{\lambda}^\top - \boldsymbol{q}\boldsymbol{d}^\top V + P\boldsymbol{\zeta}\boldsymbol{d}^\top V - P\Delta \\
&=\ \boldsymbol{q}\boldsymbol{\lambda}^\top - \boldsymbol{q}\boldsymbol{\lambda}^\top BV + P\Delta BV - P\Delta \\
&=\ (\boldsymbol{q}\boldsymbol{\lambda}^\top - P\Delta)(I - BV) \leq 0
\end{aligned}$$

where we first exploited $\boldsymbol{d}^\top V \leq \boldsymbol{u}^\top$, then constraints (10b) and (10e), and finally the fact that $BV \geq I$. □

Similarly as was the case for the original LRC model, one can uncover an intimate connection between MLRC and conservative approximations that are obtained using AARC. This connection is made explicit in the following proposition.

Proposition 5 *The MLRC (18) is equivalent to applying affine decision rules to the following two-stage problem:*

$$\begin{aligned}
\underset{\boldsymbol{x} \in \mathcal{X}, \boldsymbol{y}(\boldsymbol{\zeta}), \boldsymbol{z}(\boldsymbol{\zeta})}{\text{maximize}} \quad & \min_{\boldsymbol{\zeta} \in \mathcal{U}} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top \boldsymbol{y}(\boldsymbol{\zeta}) - \boldsymbol{u}^\top \boldsymbol{z}(\boldsymbol{\zeta}) && (25a)\\
\text{subject to} \quad & A\boldsymbol{x} + B\boldsymbol{y}(\boldsymbol{\zeta}) \leq \Psi(\boldsymbol{x})\boldsymbol{\zeta} + \boldsymbol{z}(\boldsymbol{\zeta}),\ \forall\, \boldsymbol{\zeta} \in \mathcal{U} && (25b)\\
& \boldsymbol{z}(\boldsymbol{\zeta}) \geq 0\,,\ \forall\, \boldsymbol{\zeta} \in \mathcal{U}\,. && (25c)
\end{aligned}$$

Namely, the optimal value of problems (18) and (25) are the same when $\boldsymbol{y}(\boldsymbol{\zeta})$ and $\boldsymbol{z}(\boldsymbol{\zeta})$ are limited to the space of affine functions, and so is the set of optimal solutions for $\boldsymbol{x}$.

Proof. Similarly to what was described in the proof of Proposition 2, the function $g_{MLRC}(\boldsymbol{x})$ can be reformulated as

$$\begin{aligned}
g_{MLRC}(\boldsymbol{x}) = \max_{Y, \boldsymbol{y}, W, \boldsymbol{w}, S} \quad & \min_{\boldsymbol{\zeta} \in \mathcal{U}} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top (Y\boldsymbol{\zeta} + \boldsymbol{y}) - \boldsymbol{u}^\top (\boldsymbol{w} + W^\top(\boldsymbol{q} - P\boldsymbol{\zeta})) && (26a)\\
\text{subject to} \quad & P^\top S = Y^\top B^\top - \Psi(\boldsymbol{x})^\top + P^\top W && (26b)
\end{aligned}$$

$$
\begin{aligned}
& A\boldsymbol{x} + B\boldsymbol{w} + S^\top \boldsymbol{q} \leq W^\top \boldsymbol{q} + \boldsymbol{w} && \text{(26c)}\\
& S \geq 0 && \text{(26d)}\\
& \boldsymbol{w} \geq 0\,,\ W \geq 0\,, && \text{(26e)}
\end{aligned}
$$

where $\boldsymbol{w} \in \mathbb{R}^{\bar{m}}$ and $W \in \mathbb{R}^{n_{\mathcal{U}} \times \bar{m}}$ are respectively the dual variables associated with constraints (22b) and (22c). Again, the constraints (26b)–(26d) can be replaced with

$$
A\boldsymbol{x} + B(Y\boldsymbol{\zeta} + \boldsymbol{y}) \leq \Psi(\boldsymbol{x})\boldsymbol{\zeta} + \boldsymbol{w} + W^\top(\boldsymbol{q} - P\boldsymbol{\zeta}),\ \forall\, \boldsymbol{\zeta} \in \mathcal{U}\,,
$$

and decision variable S removed from the optimization problem. In this way, we obtain

$$
\begin{aligned}
g_{MLRC}(\boldsymbol{x}) = \max_{Y,\boldsymbol{y},W,\boldsymbol{w}} \quad & \min_{\boldsymbol{\zeta}\in\mathcal{U}} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top(Y\boldsymbol{\zeta} + \boldsymbol{y}) - \boldsymbol{u}^\top(\boldsymbol{w} + W^\top(\boldsymbol{q} - P\boldsymbol{\zeta})) && \text{(27a)}\\
\text{subject to} \quad & A\boldsymbol{x} + B(Y\boldsymbol{\zeta} + \boldsymbol{y}) \leq \Psi(\boldsymbol{x})\boldsymbol{\zeta} + \boldsymbol{w} + W^\top(\boldsymbol{q} - P\boldsymbol{\zeta}),\ \forall\, \boldsymbol{\zeta} \in \mathcal{U} && \text{(27b)}\\
& \boldsymbol{w} \geq 0\,,\ W \geq 0\,. && \text{(27c)}
\end{aligned}
$$

We next introduce new variables $\boldsymbol{z}$ and Z as

$$
\boldsymbol{z} := \boldsymbol{w} + W^\top \boldsymbol{q}\,, \qquad Z := -W^\top P\,,
$$

where $\boldsymbol{z} \in \mathbb{R}^{\bar{m}}$ and $Z \in \mathbb{R}^{\bar{m} \times n_\zeta}$. Therefore, problem (26) can be reformulated as

$$
\begin{aligned}
g_{MLRC}(\boldsymbol{x}) = \max_{Y,\boldsymbol{y},W,\boldsymbol{z},Z} \quad & \min_{\boldsymbol{\zeta}\in\mathcal{U}} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top(Y\boldsymbol{\zeta} + \boldsymbol{y}) - \boldsymbol{u}^\top(\boldsymbol{z} + Z\boldsymbol{\zeta}) && \text{(28a)}\\
\text{subject to} \quad & A\boldsymbol{x} + B(Y\boldsymbol{\zeta} + \boldsymbol{y}) \leq \Psi(\boldsymbol{x})\boldsymbol{\zeta} + \boldsymbol{z} + Z\boldsymbol{\zeta},\ \forall\, \boldsymbol{\zeta} \in \mathcal{U} && \text{(28b)}\\
& Z = -W^\top P && \text{(28c)}\\
& \boldsymbol{z} - W^\top \boldsymbol{q} \geq 0 && \text{(28d)}\\
& W \geq 0\,. && \text{(28e)}
\end{aligned}
$$

We finally show that constraints (28c)–(28e) are equivalent to the following constraint:

$$
\boldsymbol{z} + Z\boldsymbol{\zeta} \geq 0\,,\ \forall\, \boldsymbol{\zeta} \in \mathcal{U}\,. \tag{29}
$$

This is done by considering that, for some fixed j, duality can once again be used to reformulate the constraint that

$$
\begin{aligned}
\max_{MW_{:j}} \quad & z_j - W_{:j}^\top \boldsymbol{q} \geq 0 && \text{(30a)}\\
\text{subject to} \quad & Z_{j:}^\top = -P^\top W_{:j} && \text{(30b)}\\
& W_{:j} \geq 0 && \text{(30c)}
\end{aligned}
$$

as the constraint that

$$
\begin{aligned}
\min_{\zeta} \quad & z_j + Z_{j:}\zeta \geq 0\,. && \text{(31a)}\\
\text{subject to} \quad & P\boldsymbol{\zeta} \leq \boldsymbol{q} && \text{(31b)}
\end{aligned}
$$

This completes our proof. □

When observing problem (25) closely, one can readily recognize that it replaces the ARO problem with

$$
\underset{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y}(\boldsymbol{\zeta}),\boldsymbol{z}(\boldsymbol{\zeta})}{\text{maximize}} \quad \min_{\boldsymbol{\zeta}\in\mathcal{U}} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top \boldsymbol{y}(\boldsymbol{\zeta}) - \sum_i u_i \max(0\,,\, \boldsymbol{e}_i^\top(\Psi(\boldsymbol{x})\boldsymbol{\zeta} - A\boldsymbol{x} - B\boldsymbol{y}(\boldsymbol{\zeta})))\,,
$$

where $\boldsymbol{e}_i \in \mathbb{R}^{\bar{m}}$ is the i-th column of the identify matrix. In other words, this is a two-stage robust optimization where one does not impose that the recourse policy $\boldsymbol{y}(\cdot)$ be feasible but instead imposes a penalty to the profit that is proportional to the magnitude of the violation of each constraint. In fact, one can show that the marginal penalties u defined in Assumption 4 are large enough to ensure that problem (25) is equivalent to problem (1), i.e., that there is an optimal solution of (25) with $\boldsymbol{z}^*(\boldsymbol{\zeta}) := 0$. The penalties described in $\boldsymbol{u}$ should, however, be selected as small as possible in order for the approximation obtained using affine decision rules to be more accurate.

We believe this interpretation of the MLRC model provides valuable new insights on how to apply affine decision rules (or more general ones) to multi-stage decision problems. Note however, that the optimal affine decision rule for $\boldsymbol{y}(\boldsymbol{\zeta})$ obtained from solving penalized models such as (25) might not be implementable for all possible realizations of $\boldsymbol{\zeta}$ and are therefore more difficult to interpret as optimal recourse policies. While this does not prevent the approximation model from identifying good first-stage decisions $\boldsymbol{x}^*$, it implies that once $\boldsymbol{x}^*$ is implemented and $\bar{\boldsymbol{\zeta}}$ is observed, one needs to seek the recourse decision that minimizes the specific recourse problem that is being experienced,[2] i.e.,

$$\begin{aligned} \underset{\boldsymbol{y}}{\text{maximize}} \quad & \boldsymbol{d}^\top \boldsymbol{y} \\ \text{subject to} \quad & A\boldsymbol{x}^* + B\boldsymbol{y} \le \Psi(\boldsymbol{x}^*)\bar{\boldsymbol{\zeta}} \end{aligned}$$

When doing so, the incurred cost is guaranteed to be smaller than what had been anticipated by the conservative approximation scheme applied to (25).

The relation between MLRC and schemes that employ affine decision rules can be exploited to demonstrate that MLRC is always a feasible approximation model. Namely, it provides a feasible first-stage solution even in situations where AARC model (2) is an infeasible problem.

Corollary 1 *Both MLRC (18) and the optimization problem obtained by employing static or affine decision rules in problem (25) are conservative approximations of problem (1) which necessarily admit a feasible solution.*

This result follows from the fact that MLRC is equivalent to employing affine decision rules in problem (25) which can be shown to have a static feasible solution. In particular, let $\bar{\boldsymbol{x}}$ be any member of $\mathcal{X}$, while letting $\boldsymbol{y}(\boldsymbol{\zeta}) := 0$ and for each $z_i(\zeta) := \max(0; \max_{\boldsymbol{\zeta}\in\mathcal{U}} \boldsymbol{e}_i^\top (A\bar{\boldsymbol{x}} - \Psi(\bar{\boldsymbol{x}})\boldsymbol{\zeta}))$, where $\boldsymbol{e}_i \in \mathbb{R}^{\bar{m}}$ is the i-th column of the identify matrix, and which is finite since $\mathcal{U}$ was assumed to be bounded. There must therefore exist a feasible assignment for MLRC (18) otherwise it would be in contradiction with Proposition 5.

Remark 1 *It is worth mentioning that any information about $\boldsymbol{\lambda}^*$ that takes the shape of linear inequalities $P_\lambda \boldsymbol{\lambda}^* \le \boldsymbol{q}_\lambda$ can be exploited in a very similar way as was done in this section with the information $\boldsymbol{\lambda}^* \le \boldsymbol{u}$. Indeed, in doing so, one obtains a conservative approximation model that is equivalent to employing affine decision rules in the following penalized ARO:*

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y}(\boldsymbol{\zeta}),\boldsymbol{z}(\boldsymbol{\zeta})}{\text{maximize}} \quad & \min_{\boldsymbol{\zeta}\in\mathcal{U}} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top \boldsymbol{y}(\boldsymbol{\zeta}) - \boldsymbol{q}_\lambda^\top \boldsymbol{z}(\boldsymbol{\zeta}) \\ \text{subject to} \quad & A\boldsymbol{x} + B\boldsymbol{y}(\boldsymbol{\zeta}) \le \Psi(\boldsymbol{x})\boldsymbol{\zeta} + P_\lambda^\top \boldsymbol{z}(\boldsymbol{\zeta}), \, \forall\, \boldsymbol{\zeta} \in \mathcal{U} \\ & \boldsymbol{z}(\boldsymbol{\zeta}) \ge 0\,,\, \forall\, \boldsymbol{\zeta} \in \mathcal{U}\,. \end{aligned}$$

Furthermore, if P_λ is such that there exists some $\boldsymbol{z}$ such that $P_\lambda^\top \boldsymbol{z} > 0$ (as is the case when $P_\lambda = I$ and $\boldsymbol{q}_\lambda = \boldsymbol{u}$) then the resulting conservative approximation is guaranteed to be feasible. The question remains however of how to efficiently identify a compact set of inequalities, through P_λ and q_λ, that describes a tight outer approximation of the convex hull of the vertices of $\{\boldsymbol{\lambda} \,|\, \boldsymbol{\lambda} \ge 0,\, B^\top \boldsymbol{\lambda} = \boldsymbol{d}\}$.

[2] As discussed in Delage and Iancu (2015), such a procedure needs in principle to be applied for any solution scheme that does not guarantee "Bellman-optimality" and in particular for the solution of models where affine decision rules are employed.

3.2 Exploiting valid conic inequalities

Our second source of improvement for the LRC model comes from considering the following set of quadratic equalities:

$$\begin{bmatrix} \Lambda & \Delta^\top \\ \Delta & \Xi \end{bmatrix} = \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\zeta} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}^\top & \boldsymbol{\zeta}^\top \end{bmatrix}, \tag{32}$$

where $\Lambda \in \mathbb{R}^{\bar{m}\times\bar{m}}$ and $\Xi \in \mathbb{R}^{n_\zeta \times n_\zeta}$, such that $\Lambda := \boldsymbol{\lambda}\boldsymbol{\lambda}^\top$ and $\Xi := \boldsymbol{\zeta}\boldsymbol{\zeta}^\top$. It is well known that this system of equations can be relaxed using the following matrix inequality

$$\begin{bmatrix} \Lambda & \Delta^\top \\ \Delta & \Xi \end{bmatrix} \succeq \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\zeta} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}^\top & \boldsymbol{\zeta}^\top \end{bmatrix},$$

where $A \succeq B$ indicates that $A - B$ is in the cone of positive semi-definite matrices. This non-linear matrix inequality reduces to a linear matrix inequality after applying Schur's complement

$$\begin{bmatrix} \Lambda & \Delta^\top & \boldsymbol{\lambda} \\ \Delta & \Xi & \boldsymbol{\zeta} \\ \boldsymbol{\lambda}^\top & \boldsymbol{\zeta}^\top & 1 \end{bmatrix} \succeq 0\,. \tag{33}$$

This constraint can be added to problem (10) with additional valid inequalities involving Λ and Ξ to obtain the tighter SDP-LRC model.

Proposition 6 *Given Assumption 4, the following semi-definite programming linearized robust counterpart is a conservative approximation of problem* (1)*:*

$$\begin{aligned} g_{\text{SDP-LRC}}(\boldsymbol{x}) = \min_{\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta,\Lambda,\Xi} \quad & \boldsymbol{c}^\top\boldsymbol{x} + \operatorname{tr}(\Psi(\boldsymbol{x})\Delta) - (A\boldsymbol{x})^\top\boldsymbol{\lambda} && \text{(34a)} \\ \text{subject to} \quad & (6b) - (6d), (8b) - (8f), (19), (21) \\ & \Lambda \le u\boldsymbol{\lambda}^\top && \text{(34b)} \\ & \Lambda + \boldsymbol{u}\boldsymbol{u}^\top \ge \boldsymbol{\lambda}\boldsymbol{u}^\top + \boldsymbol{u}\boldsymbol{\lambda}^\top && \text{(34c)} \\ & \begin{bmatrix} \Lambda & \Delta^\top & \boldsymbol{\lambda} \\ \Delta & \Xi & \boldsymbol{\zeta} \\ \boldsymbol{\lambda}^\top & \boldsymbol{\zeta}^\top & 1 \end{bmatrix} \succeq 0\,. && \text{(34d)} \end{aligned}$$

Moreover, the optimal value of $\max_{x\in\mathcal{X}} g_{\text{SDP-LRC}}(\boldsymbol{x})$ *is necessarily larger than or equal to the optimal value of LRC* (9) *and MLRC* (18) *and can in some cases be strictly larger.*

Proof. We start by including the new variables Λ and Ξ and constraint (33) in problem (8). One can also exploit $\boldsymbol{\lambda} \le \boldsymbol{u}$ which implies that $\boldsymbol{\lambda}\boldsymbol{\lambda}^\top \le \boldsymbol{u}\boldsymbol{\lambda}^\top$ and that

$$(\boldsymbol{u}-\boldsymbol{\lambda})(\boldsymbol{u}-\boldsymbol{\lambda})^\top \ge 0 \;\Rightarrow \boldsymbol{u}\boldsymbol{u}^\top - \boldsymbol{u}\boldsymbol{\lambda}^\top - \boldsymbol{\lambda}\boldsymbol{u}^\top + \boldsymbol{\lambda}\boldsymbol{\lambda}^\top \ge 0\,.$$

This leads to problem (34) after replacing $\Xi := \boldsymbol{\zeta}\boldsymbol{\zeta}^\top$ and $\Lambda := \boldsymbol{\lambda}\boldsymbol{\lambda}^\top$, and relaxing the quadratic equality constraint (8g)–(8i). We further refer the reader to Appendix F for a simple multi-item newsvendor problem where SDP-LRC strictly improves the bound of AARC. □

Given that the computational difficulties of solving the SDP-LRC model arise from the use of large linear matrix inequality constraint (34d), it might be beneficial to approximate this model conservatively using a set of smaller linear matrix inequalities as proposed in the following corollary.

Corollary 2 *Given Assumption 4 and a set of matrices* $\{\mathbb{M}_k\}_{i=k,\dots,K}$*, with each* $\mathbb{M}_k \in \mathbb{R}^{(\bar{m}+n_\zeta+1)\times n_k}$ *for some* $n_k \le \bar{m}+n_\zeta+1$*, the following semi-definite programming linearized robust counterpart is a conservative approximation of problem (1):*

$$\max_{\boldsymbol{x}\in\mathcal{X}} g_{\textit{SDP-LRC2}}(\boldsymbol{x}) = \min_{\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta,\Lambda,\Xi} \quad \boldsymbol{c}^\top\boldsymbol{x} + \mathrm{tr}(\Psi(\boldsymbol{x})\Delta) - (A\boldsymbol{x})^\top\boldsymbol{\lambda} \tag{35a}$$

$$\text{subject to} \quad (6b) - (6d), (8b) - (8f), (19), (21), (34b), (34c)$$

$$\mathbb{M}_k^\top \begin{bmatrix} \Lambda & \Delta^\top & \boldsymbol{\lambda} \\ \Delta & \Xi & \boldsymbol{\zeta} \\ \boldsymbol{\lambda}^\top & \boldsymbol{\zeta}^\top & 1 \end{bmatrix} \mathbb{M}_k \succeq 0\,, \forall\, k = 1,\dots,K\,. \tag{35b}$$

Moreover, the optimal value of $\max_{x\in\mathcal{X}} g_{\textit{SDP-LRC2}}(\boldsymbol{x})$ *is necessarily larger than or equal to the optimal value of LRC (9) and MLRC (18) and necessarily smaller or equal to* $\max_{x\in\mathcal{X}} g_{\textit{SDP-LRC}}(\boldsymbol{x})$*.*

Note that, although we presented $g_{\text{SDP-LRC}}(\boldsymbol{x})$ and $g_{\text{SDP-LRC2}}(\boldsymbol{x})$ as minimization problems, semi-definite programming duality can be employed to obtain a maximization representation of these functions that can be integrated with the maximization in x as was done with other LRC models. We, however, omit the details of these reformulations for aesthetics reasons. Given the connections to AARC that were established regarding the LRC and MLRC models, we suspect that a similar connection could be obtained for the SDP-LRC and SDP-LRC2 models. In fact, Ardestani-Jaafari and Delage (2016) establishes such a connection for a special case of the SDP-LRC2 model in the context where $g(\boldsymbol{x})$ captures the worst-case of a sum of piecewise linear convex functions. A quick look at their result suggests that the connection that could be established here is highly technical and would provide rather limited new insights.

Remark 2 *It is worth indicating to the reader that in the process of identifying practical applications that might benefit from the application of models such as SDP-LRC and SDP-LRC2 (and in fact AARC and MLRC), we discovered that in the surgery allocation problem considered in Denton et al. (2010), these models could in some cases improve the quality of the solution obtained using RORA, the "exact" tractable reformulation proposed by the authors. We refer to Appendix D for a complete demonstration of how and why one can improve the quality of the solutions identified by the reformulation in Denton et al. (2010). This result is especially interesting given that it re-opens the question of how to identify exact or good quality solutions for this important class of problems in a tractable way.*

4 LRC for general uncertainty sets

In this section, we extend our LRC model so that it can accommodate general convex uncertainty sets, i.e., non-polyhedral $\mathcal{U}$, while preserving the strong connection to AARC. This will in turn allow us to generalize the MLRC and SDP-LRC models to the non-polyhedral setting while ensuring that the quality of the bounds that are obtained improve on the bounds that can be obtained using AARC. Specifically, in this section we consider uncertainty sets that can be represented as

$$\mathcal{U}_{general} := \{\boldsymbol{\zeta} \in \mathbb{R}^{n_\zeta} \mid P\boldsymbol{\zeta} \le \boldsymbol{q},\ f_l(\boldsymbol{\zeta}) \le 0, \forall\, l = 1,\dots,L\} \tag{36}$$

using a set of convex $f_l(\cdot)$ functions. To establish an extension of LRC, we will need to make use of the notion of perspective functions for a special class of convex functions.

Assumption 5 *The uncertainty set* $\mathcal{U}_{general}$ *defines a bounded convex set containing a strictly feasible solution* $\bar{\boldsymbol{\zeta}}$ *such that* $f_l(\bar{\boldsymbol{\zeta}}) < 0$ *for all* $l = 1,\dots,L$*. Furthermore, for each* $l = 1,\dots,L$*, the function* $f_l(\cdot)$ *is a lower semi-continuous convex function. This implies that, according to the Fenchel-Moreau Theorem, it must be that* $f_l(\boldsymbol{x}) = \sup_{\boldsymbol{y}} \boldsymbol{x}^\top\boldsymbol{y} - f_*(\boldsymbol{y})$*, where* $f_{l*}(\boldsymbol{y})$ *is the convex conjugate of* $f_l(\cdot)$*, i.e.,* $f_{l*}(\boldsymbol{y}) := \sup_{\boldsymbol{z}} \boldsymbol{y}^\top\boldsymbol{z} - f_l(\boldsymbol{z})$*.*

Definition 1 *(Hiriart-Urruty and Lemaréchal 2001) For each* $l = 1,\dots,L$*, let* $h_l : \mathbb{R}^{n_\zeta} \times \mathbb{R}_+ \to \mathbb{R}$ *be the closure, denoted by cl, of the perspective function* $h_l(\boldsymbol{z},t) := cl\ t f_l(\boldsymbol{z}/t)$*. In particular, given that* $f_l(\cdot)$ *satisfies*

Assumption 5, we have that

$$h_l(\boldsymbol{z},t) := \sup_{\boldsymbol{y}} \boldsymbol{z}^\top \boldsymbol{y} - t f_{l*}(\boldsymbol{y}) = \begin{cases} t f_l(\boldsymbol{z}/t) & \text{if } t > 0 \\ \lim_{t\to 0^+} t f_l(\boldsymbol{z}/t) & \text{if } t = 0 \\ \infty & \text{otherwise} \end{cases} .$$

Based on this definition, it is clear that $h_l(\boldsymbol{z},t)$ is jointly convex in $\boldsymbol{z}$ and t.

Under the uncertainty set $\mathcal{U}_{general}$, when Assumption 4 is satisfied, the value of $g(\boldsymbol{x})$ becomes

$$\begin{aligned} g(\boldsymbol{x}) = \min_{\boldsymbol{\zeta},\boldsymbol{\lambda}} \quad & \boldsymbol{c}^\top \boldsymbol{x} + (\Psi(\boldsymbol{x})\boldsymbol{\zeta})^\top \boldsymbol{\lambda} - (A\boldsymbol{x})^\top \boldsymbol{\lambda} && \text{(37a)} \\ \text{subject to} \quad & (6b) - (6d), (7b), (7c), (19), (20) \\ & f_l(\boldsymbol{\zeta}) \leq 0, \forall\, l = 1,\ldots,L && \text{(37b)} \\ & \lambda_i f_l(\boldsymbol{\zeta}) \leq 0\,, \forall\, i = 1,\ldots,\bar{m}\,, \forall\, l = 1,\ldots,L && \text{(37c)} \\ & f_l(\boldsymbol{\zeta})(u_i - \lambda_i) \leq 0\,, \forall\, i = 1,\ldots,\bar{m}\,, \forall\, l = 1,\ldots,L && \text{(37d)} \\ & f_l(\boldsymbol{\zeta})\boldsymbol{e}_i^\top(\boldsymbol{q} - P\boldsymbol{\zeta}) \leq 0\,, \forall\, i = 1,\ldots,n_{\mathcal{U}}\,, \forall\, l = 1,\ldots,L\,. && \text{(37e)} \end{aligned}$$

As was done for polyhedral sets, under Assumption 5 this optimization model can be linearized using perspective functions:

$$\begin{aligned} g(\boldsymbol{x}) \geq g_{GLRC}(\boldsymbol{x}) = \min_{\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \mathrm{tr}(\Psi(\boldsymbol{x})\Delta) - (A\boldsymbol{x})^\top \boldsymbol{\lambda} && \text{(38a)} \\ \text{subject to} \quad & (6b) - (6d), (8b), (8c), (19), (21) \\ & f_l(\boldsymbol{\zeta}) \leq 0, \forall\, l = 1,\ldots,L && \text{(38b)} \\ & h(\Delta \boldsymbol{e}_i, \boldsymbol{e}_i^\top \boldsymbol{\lambda}) \leq 0\,, \forall\, i = 1,\ldots,\bar{m}\,, \forall\, l = 1,\ldots,L && \text{(38c)} \\ & h_l(\boldsymbol{\zeta}\boldsymbol{e}_i^\top \boldsymbol{u} - \Delta \boldsymbol{e}_i, \boldsymbol{e}_i^\top(\boldsymbol{u} - \boldsymbol{\lambda})) \leq 0\,, \forall\, i = 1,\ldots,\bar{m}, \forall\, l = 1,\ldots,L && \text{(38d)} \\ & h_l(\boldsymbol{\zeta}\boldsymbol{e}_i^\top \boldsymbol{q} - \Xi P^\top \boldsymbol{e}_i,\, \boldsymbol{e}_i^\top(\boldsymbol{q} - P\boldsymbol{\zeta})) \leq 0\,, \forall\, i = 1,\ldots,n_{\mathcal{U}}, \forall\, l = 1,\ldots,L\,, && \text{(38e)} \end{aligned}$$

where constraint (38c) is derived from

$$\boldsymbol{\lambda}_i f_l(\boldsymbol{\zeta}) \leq 0 \Rightarrow \boldsymbol{e}_i^\top \boldsymbol{\lambda}(\sup_{\boldsymbol{y}} \boldsymbol{\zeta}^\top \boldsymbol{y} - f_{l*}(\boldsymbol{y})) \leq 0 \Rightarrow \sup_{\boldsymbol{y}} (\boldsymbol{e}_i^\top \boldsymbol{\lambda}\boldsymbol{\zeta})^\top \boldsymbol{y} - \boldsymbol{e}_i^\top \boldsymbol{\lambda} f_{l*}(\boldsymbol{y}) = h(\Delta \boldsymbol{e}_i, \boldsymbol{e}_i^\top \boldsymbol{\lambda}) \leq 0$$

and constraint (38d) is derived from

$$\begin{aligned} f_l(\boldsymbol{\zeta})(u_i - \boldsymbol{\lambda}_i) \leq 0 &\Rightarrow \sup_{\boldsymbol{y}} \boldsymbol{e}_i^\top(\boldsymbol{u} - \boldsymbol{\lambda})\boldsymbol{\zeta}^\top \boldsymbol{y} - \boldsymbol{e}_i^\top(\boldsymbol{u} - \boldsymbol{\lambda}) f_{l*}(\boldsymbol{y}) \leq 0 \\ &\Rightarrow h_l(\boldsymbol{\zeta}\boldsymbol{e}_i^\top(\boldsymbol{u} - \boldsymbol{\lambda}), \boldsymbol{e}_i^\top(\boldsymbol{u} - \boldsymbol{\lambda})) \leq 0 \\ &\Rightarrow h_l(\boldsymbol{\zeta}\boldsymbol{e}_i^\top \boldsymbol{u} - \Delta \boldsymbol{e}_i, \boldsymbol{e}_i^\top(\boldsymbol{u} - \boldsymbol{\lambda})) \leq 0\,, \end{aligned}$$

for all i and for all l, and where the term $\boldsymbol{\zeta}\lambda_i$ is linearized through $\Delta_{:i}$. Finally, constraint (38e) is obtained from:

$$\begin{aligned} f_l(\boldsymbol{\zeta})\boldsymbol{e}_i^\top(\boldsymbol{q} - P\boldsymbol{\zeta}) \leq 0 &\Rightarrow h(\boldsymbol{\zeta}\boldsymbol{e}_i^\top(\boldsymbol{q} - P\boldsymbol{\zeta}),\, \boldsymbol{e}_i^\top(\boldsymbol{q} - P\boldsymbol{\zeta})) \leq 0 \\ &\Rightarrow h(\boldsymbol{\zeta}\boldsymbol{e}_i^\top \boldsymbol{q} - \boldsymbol{\zeta}\boldsymbol{\zeta}^\top P^\top \boldsymbol{e}_i,\, \boldsymbol{e}_i^\top(\boldsymbol{q} - P\boldsymbol{\zeta})) \leq 0 \\ &\Rightarrow h(\boldsymbol{\zeta}\boldsymbol{e}_i^\top \boldsymbol{q} - \Xi P^\top \boldsymbol{e}_i,\, \boldsymbol{e}_i^\top(\boldsymbol{q} - P\boldsymbol{\zeta})) \leq 0\,. \end{aligned}$$

One might then apply duality theory to problem (38) to derive a compact mathematical programming representation of the LRC model under $\mathcal{U}_{general}$. Regarding the relation between this more general LRC model and AARC, Appendix C will demonstrate that the problem maximize$_{\boldsymbol{x}\in\mathcal{X}}\, g_{GLRC}(\boldsymbol{x})$ always produces a tighter conservative approximation than employing affine decision rules in problem (25). It is worth noting however that, based on the proof presented in Appendix C, it appears legitimate to believe that the two schemes are equivalent in most practical situations.

Finally, for completeness we can present the extension of SDP-LRC and SDP-LRC2 to the context of general uncertainty sets as follows:

$$g(\boldsymbol{x}) \geq g_{GSDP-LRC2}(\boldsymbol{x}) = \min_{\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta,\Lambda,\Xi} \quad \boldsymbol{c}^\top \boldsymbol{x} + \mathrm{tr}(\Psi(\boldsymbol{x})\Delta) - (A\boldsymbol{x})^\top \boldsymbol{\lambda} \tag{39a}$$

$$\text{subject to} \quad (6b) - (6d), (8b) - (8f), (19), (21), (34b), (34c), (38b) - (38e)$$

$$\mathbb{M}_k^\top \begin{bmatrix} \Lambda & \Delta^\top & \boldsymbol{\lambda} \\ \Delta & \Xi & \boldsymbol{\zeta} \\ \boldsymbol{\lambda}^\top & \boldsymbol{\zeta}^\top & 1 \end{bmatrix} \mathbb{M}_k \succeq 0\,,\ \forall\, k = 1, \ldots, K\,. \tag{39b}$$

for some arbitrary set of $\{\mathbb{M}_k\}_{k=1,\ldots,K}$. When using the conservative approximation of the form $\max_{\boldsymbol{x}\in\mathcal{X}} g_{GSDP-LRC2}(\boldsymbol{x})$, based on Proposition 7 presented in Appendix C, one has the guarantee that the optimal value is larger or equal to the optimal value of AARC and MLRC while different level of numerical efficiency can be achieved by choosing the size of K and each $\mathbb{M}_k$ appropriately.

5 Numerical study with the robust location-transportation problem

In this section, we first revisit the robust location-transportation problem (RLTP) presented in (3). As a short background, we could say that RLTP is one of the most popular two-stage robust optimization problems. It was first studied by Atamtürk and Zhang (2007) and since then several researchers, e.g., Baron et al. (2011), Ardestani-Jaafari and Delage (2017), Bertsimas and de Ruiter (2016) have studied this problem. In the remainder of this section, we describe how to formulate the associated MLRC in Section 5.1 and show that it identifies the optimal solution for the instance presented in Example 1. We will then, in Section 5.2, evaluate the quality of solutions obtained from MLRC, and its numerical efficiency, to AARC, the conservative approximation model proposed in Ardestani-Jaafari and Delage (2017), and an exact solution scheme presented in Zeng and Zhao (2013), on an extensive set of problem instances.

5.1 Implementation of MLRC for the RLTP

In order to formulate the MLRC model, one first needs to identify a $\boldsymbol{u}$ that satisfies Assumption 4. To do so, we start by reminding the reader of the recourse problem associated problem 3:

$$\underset{\boldsymbol{y}}{\text{maximize}} \quad \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \eta_{ij} y_{ij} \tag{40a}$$

$$\text{subject to} \quad \sum_{i\in\mathcal{I}} y_{ij} \leq \zeta_j\,,\ \forall\, j \in \mathcal{J} \tag{40b}$$

$$\sum_{j\in\mathcal{J}} y_{ij} \leq x_i\,,\ \forall\, i \in \mathcal{I} \tag{40c}$$

$$\boldsymbol{y} \geq 0\,. \tag{40d}$$

Hence, the dual formulation of the recourse function takes the form

$$\underset{\boldsymbol{\lambda}^1,\boldsymbol{\lambda}^2,\boldsymbol{\lambda}^3}{\text{minimize}} \quad \sum_{j\in\mathcal{J}} \zeta_j \lambda_j^1 + \sum_{i\in\mathcal{I}} x_i \lambda_i^2 \tag{41a}$$

$$\text{subject to} \quad \lambda_j^1 + \lambda_i^2 - \lambda_{ij}^3 = \eta_{ij}\,,\ \forall\, i \in \mathcal{I}\,, \forall\, j \in \mathcal{J} \tag{41b}$$

$$\boldsymbol{\lambda}^1 \geq 0\,,\ \boldsymbol{\lambda}^2 \geq 0\,,\ \boldsymbol{\lambda}^3 \geq 0\,, \tag{41c}$$

where $\boldsymbol{\lambda}^1 \in \mathbb{R}^n$, $\boldsymbol{\lambda}^2 \in \mathbb{R}^m$, and $\boldsymbol{\lambda}^3 \in \mathbb{R}^{m\times n}$ are dual variables associated with constraints (40b), (40c), and (40d) respectively. Our goal is therefore to identify tight bounds for $\boldsymbol{\lambda}^{1*}$, $\boldsymbol{\lambda}^{2*}$, and $\boldsymbol{\lambda}^{3*}$.

Since $\boldsymbol{x} \geq 0$ and $\boldsymbol{\zeta} \geq 0$, the objective function of problem (41) is non-decreasing in $\boldsymbol{\lambda}^1$ and $\boldsymbol{\lambda}^2$. One can therefore conclude that, at optimum, each term of $\boldsymbol{\lambda}^{1*}$ will be such that it will either be equal to 0 or involved in an active constraint among the set of constraints

$$\lambda_j^1 + \lambda_i^2 \geq \eta_{ij}\,,\ \forall\, i \in \mathcal{I}\,, \forall\, j \in \mathcal{J}\,.$$

It must therefore be that

$$\lambda_j^{1*} \le \max(0\,,\, \max_i \eta_{ij} - \lambda_i^{2*}) \le \max(0\,,\, \max_i \eta_{ij}) := u_j^1\,,\, \forall j \in \mathcal{J}\,.$$

Similarly, in the case of $\boldsymbol{\lambda}^{2*}$, we have

$$\lambda_i^2 \le \max(0\,,\, \max_j \eta_{ij} - \lambda_j^{1*}) \le \max(0\,,\, \max_j \eta_{ij}) := u_i^2\,,\, \forall i \in \mathcal{I}\,.$$

Finally, since $\lambda_{ij}^{3*} = \lambda_j^{1*} + \lambda_i^{2*} - \eta_{ij}$, one could conclude that

$$\lambda_{ij}^3 \le \max(0\,,\, \max_i \eta_{ij} + \max_j \eta_{ij} - \eta_{ij}) := u_{ij}^3\,. \tag{42}$$

Based on propositions 4 and 5, we know that we obtain conservative approximation to problem (3) when employing affine decision rules in the following augmented model:

$$\begin{aligned}
\underset{\boldsymbol{x},\boldsymbol{y}(\boldsymbol{\zeta}),\boldsymbol{z}^1(\boldsymbol{\zeta}),\boldsymbol{z}^2(\boldsymbol{\zeta}),\boldsymbol{z}^3(\boldsymbol{\zeta}),v}{\text{maximize}} \; \min_{\boldsymbol{\zeta}\in\mathcal{U}} & - \sum_{i\in\mathcal{I}} \left(c_i x_i + k_i v_i\right) + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \eta_{ij} y_{ij}(\boldsymbol{\zeta}) - \sum_{i\in\mathcal{I}} u_i^1 z_i^1(\boldsymbol{\zeta}) - \sum_{j\in\mathcal{J}} u_j^2 z_j^2(\boldsymbol{\zeta}) - \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} u_{ij}^3 z_{ij}^3 \\
\text{subject to} & \sum_i y_{ij}(\boldsymbol{\zeta}) \le \zeta_j + z_j^1(\boldsymbol{\zeta}) \,\forall j \in \mathcal{J}\,\forall \boldsymbol{\zeta} \in \mathcal{U} \\
& \sum_{j\in\mathcal{J}} y_{ij}(\boldsymbol{\zeta}) \le x_i + z_i^2(\boldsymbol{\zeta}) \,\forall i \in \mathcal{I}\,\forall \boldsymbol{\zeta} \in \mathcal{U} \\
& \boldsymbol{y}(\boldsymbol{\zeta}) + \boldsymbol{z}^3(\boldsymbol{\zeta}) \ge 0 \,\forall \boldsymbol{\zeta} \in \mathcal{U} \\
& \boldsymbol{z}^1(\boldsymbol{\zeta}) \ge 0 \; \boldsymbol{z}^2(\boldsymbol{\zeta}) \ge 0 \; \boldsymbol{z}^3(\boldsymbol{\zeta}) \ge 0 \,\forall \boldsymbol{\zeta} \in \mathcal{U} \\
& 0 \le x_i \le M v_i \,\forall i \in \mathcal{I} \\
& v_i \in \{0,1\} \,\forall i \in \mathcal{I}\,.
\end{aligned}$$

where $\boldsymbol{z}^1 : \mathbb{R}^n \to \mathbb{R}^n$, $\boldsymbol{z}^2 : \mathbb{R}^n \to \mathbb{R}^m$, and $\boldsymbol{z}^3 : \mathbb{R}^n \to \mathbb{R}^{m\times n}$ can be interpreted as violation adjustments for constraints (3b), (3c), and (3d) respectively. In Ardestani-Jaafari and Delage (2017), the authors employed a special case of such a conservative approximation, that is called as the Extended Lifted Affinely Adjustable Robust Counterpart (ELAARC), where $z_j^1(\boldsymbol{\zeta}) := z_j^1 \zeta_j$, $\boldsymbol{z}^2(\boldsymbol{\zeta}) := 0$, and $\boldsymbol{z}^3(\boldsymbol{\zeta}) := 0$, and showed that there exist instances of the location-transportation problem for which this conservative approximation is strictly tighter than employing affine decision rules directly in model (3).

When addressing the specific problem instance presented in Example 1, one can verify whether the bounds for $\boldsymbol{\lambda}^{1*}$, $\boldsymbol{\lambda}^{2*}$, and $\boldsymbol{\lambda}^{3*}$ can be improved. This can be done by solving problem (17) with $M := 2\|\eta\|_\infty = 11.8$ since for $k = 1, 2, 3$ we have that $\boldsymbol{\lambda}^{k*} \le \max_{ij} |\eta_{ij}|$, and exploiting the proposed constraint generation scheme. This leads to establishing the following bounds:

$$\boldsymbol{u}^{1*} = \begin{bmatrix} 5.9 \\ 5.9 \\ 4.9 \end{bmatrix} = \boldsymbol{u}^1 \qquad \boldsymbol{u}^{2*} = \begin{bmatrix} 5.9 \\ 5.9 \end{bmatrix} = \boldsymbol{u}^2 \qquad \boldsymbol{u}^{3*} = \begin{bmatrix} 0 & 0.6 & 1 \\ 0.6 & 0 & 1 \end{bmatrix} < \boldsymbol{u}^3 := \begin{bmatrix} 5.9 & 6.2 & 5.9 \\ 6.2 & 5.9 & 5.9 \end{bmatrix}.$$

Note that for this problem instance our analytical estimates of $\boldsymbol{u}^1$ and $\boldsymbol{u}^2$ were "optimal" while the optimal penalties $\boldsymbol{u}^{3*}$ were significantly smaller than the analytical ones.

Table 1 compares the performance of AARC proposed by Ben-Tal et al. (2004), ELAARC proposed by Ardestani-Jaafari and Delage (2017) and our proposed MLRC in this problem instance. One can observe from the table that while AARC exhibits "over-conservatism" by refusing to open any of the facilities, ELAARC instead provides facility location plan that achieves 70% of the best worst-case profit possible. The optimality gap is actually reduced to zero when using MLRC. It is worth reminding the reader that AARC (and implicitly ELAARC and MLRC) was shown in Ardestani-Jaafari and Delage (2017) to provide an exact solution to this model when $\Gamma = 1$ or $\Gamma = 3$, however, this example confirms that some improvement is possible for other sizes of budgets and that MLRC is a promising alternative to consider. This point of view will be further reinforced in the extensive numerical experiments presented in the next section.

Table 1: Comparison of optimal bound on worst-case profit and worst-case profit associated with solutions of each model in an instance of RLTP.

	AARC/LRC	ELAARC	MLRC	SDP-LRC	Exact model
Optimal bound on worst-case profit	0	4,024	6,600	6,600	6,600
Worst-case profit of solution	0	4,622	6,600	6,600	6,600

5.2 Numerical results of location-transportation problem

In this section, we perform an extensive comparison of the quality of the solutions that are provided for the RLTP by AARC, MLRC, and the ELAARC model proposed by Ardestani-Jaafari and Delage (2017) on a set of randomly generated location-transportation problem instances. In the case of MLRC, we choose to focus on an implementation that fixes $\boldsymbol{z}^3 = 0$ given that a preliminary study revealed that no improvements where achieved using $\boldsymbol{z}^3$ with $\boldsymbol{u}^3$ set as in (42).[3] Subsection 5.2.1 will study the impact of varying the amount of uncertainty on the quality of the robust strategy proposed by each approximation model in a large set of small artificially generated instances. Next, Subsection 5.2.2 will focus on comparing the computational requirements for solving these approaches on the large scale instances used in Ardestani-Jaafari and Delage (2017).

5.2.1 Optimality gap analysis

In this subsection, we attempt to empirically compare the quality of the approximate robust solutions obtained from the different conservative approximation models. We artificially generated 100 instances of an uncapacitated problem with 10 facility and 20 customer locations, where the variable and fixed capacity costs are respectively fixed to $c = 0.6$ and $K = 10,000$. In each instance, the marginal revenue η_{ij} associated to goods shipped from each location $i \in \mathcal{I}$ to each customer $j \in \mathcal{J}$ is independently and identically drawn from a uniform distribution over the interval [4, 7]. Finally, for all j, the nominal demand is fixed to $\bar{\zeta}_j = 20,000$ while the maximum demand deviation is drawn using $\hat{\zeta}_j = \varepsilon_j \bar{\zeta}_j$, for some maximum deviation ratio ε_j generated uniformly over the $[0.3, 0.9]$ interval unless specified otherwise. The quality of the different approaches will be measured in terms of "relative suboptimality gap" which refers to the relative difference between the optimal worst-case profit, that is obtained by C&CG algorithm of Zeng and Zhao (2013) for a problem instance and the worst-case profit achieved when implementing the approximate facility location decisions while the transportation plan is re-optimized when the demand is observed.

In the first set of experiments we focus on the quality of solutions obtained under different values of Γ. Table 2 presents average, 90-percentile, and maximum relative suboptimality gap of each approximation over 100 instances. Regarding the quality of the optimized worst-case bound, one might first observe in Table 2 that the average suboptimality gap is always strictly improved when using ELAARC and MLRC models. One might further notice that the most significant improvements appear to occur exactly when passing from the ELAARC model to the MLRC model. Specifically, the MLRC model is able to improve 90-percentile and maximum suboptimality gap obtained with ELAARC by factor 2.6 and 2.0 respectively. Furthermore, one might notice that this reduction becomes more significant as one increases the required level of robustness, Γ, as long as $\Gamma < |\mathcal{J}|$. A similar observation can be made for the ELAARC model for which the difference with AARC is barely noticeable for small values of Γ.

Table 3 provides additional statistics about the relative suboptimality gap of the different solutions proposed by each approximation model in the 1,000 problem instances surveyed in Table 2 (i.e., 100 randomly generated instances evaluated for all $\Gamma = 2, 4, \ldots, 20$). Specifically, the table indicates, the proportion of instances for which an approximation model was able to identify a solution whose relative suboptimality gap was within a given range. The table also presents the average and maximum relative suboptimality gap for each conservative approximation model. While similar observations as before can be made here, this table

[3]Note that one could potentially improve the quality of the solution obtained from MLRC and SDP-LRC by setting $\boldsymbol{u}^3$ using problem (17). Yet, one should then account for the time spent doing so in comparing the numerical efficiency of the different methods. Comparatively, the estimates for $\boldsymbol{u}^1$ and $\boldsymbol{u}^2$ appeared to be optimal in all instances that were considered.

highlights how the added flexibility employed in the MLRC gives this approximation scheme significantly better chances of identifying high-quality solutions.

Table 2: Relative suboptimality gap for the solutions obtained from each approximation model under different values of budget. Note that Ardestani-Jaafari and Delage (2017) showed that AARC provides an exact solution when $\Gamma = |\mathcal{J}|$ in these problem instances.

$\Gamma\%^*$	Avg. suboptimality gap			90-percentile suboptimality gap			Max. suboptimality gap		
	AARC	ELAARC	MLRC†	AARC	ELAARC	MLRC†	AARC	ELAARC	MLRC†
10	0.06	0.05	0.03	0.12	0.11	0.07	0.19	0.17	0.17
20	0.19	0.17	0.10	0.34	0.30	0.21	0.58	0.50	0.28
30	0.34	0.32	0.15	0.53	0.51	0.32	0.78	0.70	0.68
40	0.48	0.46	0.19	0.78	0.72	0.39	1.18	1.05	0.93
50	0.63	0.61	0.15	1.05	0.98	0.38	1.41	1.33	1.00
60	0.73	0.71	0.17	1.36	1.28	0.44	1.66	1.58	0.94
70	0.80	0.77	0.20	1.51	1.46	0.46	2.15	1.94	0.74
80	0.79	0.76	0.17	1.46	1.36	0.43	2.48	2.28	0.77
90	0.71	0.69	0.16	1.23	1.20	0.34	3.01	2.54	0.59
100	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Avg.	0.47	0.45	0.13	0.84	0.79	0.30	1.34	1.21	0.61

†This implementation of MLRC enforces $\boldsymbol{z}^3 = 0$. * $\Gamma\%$ denotes that the units are in % of $|\mathcal{J}|$.

Table 3: Proportion of the 1,000 problem instances analyzed in Table 2 for which the relative suboptimality gap of each approximation model was within a certain range. Average gap and maximum gap are also reported.

Gap range	AARC	ELAARC	MLRC†
=0	10.1	10.1	13.6
≤0.1	20.6	21.8	55.9
≤1	89.4	90.4	100
≤10	100	100	100
Avg.	0.47	0.45	0.13
Max.	3.01	2.54	0.99

†This implementation of MLRC enforces $\boldsymbol{z}^3 = 0$.

In the second set of our results, Tables 4 and 5, we consider five different Γ, i.e., $\Gamma = 10\%, 30\%, 50\%, 70\%, 90\%$ of $|\mathcal{J}|$, and seven different ε from 0.3 to 0.9. We report in Table 4 the average, 90-percentile, and maximum suboptimality gap achieved under each conditions. One can first observe in Table 4 that the suboptimality gap increases for all three models as ε and Γ increase. However, the increase is more significant in AARC and ELAARC such that their average suboptimality gaps reaches 35.0% and 25.7% respectively when $\Gamma = 90\%$ of $|\mathcal{J}|$ and $\varepsilon = 0.9$. On the other hand, MLRC's performance is less dramatically affected by such changes reaching an average relative suboptimality of 3.46% and maximum of 7.8% under the same conditions.

Table 5 provides additional information about the results reported in Table 4 in a similar way as Table 3. The main new result of this table are that the average suboptimality gap of MLRC is less than 1% and while the maximum suboptimality gap of AARC and ELAARC can exceed 30%, it is always less than 10% with maximum of 7.75% for MLRC. In sum, while our results show the improvement in performance of AARC when we move to MLRC, as shown in Table 2, they clearly illustrates how MLRC can be beneficial especially when the amount of uncertainty is large (Table 4).

Finally, we compare SDP-LRC in the presence of other approximations in the above instances when both budget and size of uncertainty are high. As shown in Table 6, AARC and ELAARC are somewhat overly conservative compared to MLRC. In particular, the latter can improve reduce the average suboptimality gap from 10.8%, with AARC, to 1.9%. Furthermore, SDP-LRC can additionally reduce it down to 0.2%.

Numerically speaking, this did not appear to come at a price in terms of solution time (resolution time of all three models were comparable) as verified in the context of larger scale problems in Subsection 5.2.2.

From the computational perspective, the improvement in quality of MLRC did not appear to come at a price in terms of solution time. Specifically, all three linear programming based reformulation (AARC, ELAARC, and MLRC) led to similar solution time (6.1, 6.7, and 8.8 seconds on average respectively). On the other hand, SDP-LRC was much more demanding (i.e., 109 seconds) yet its efficiency could perhaps be improved using SDP-LRC2 model or the decomposition scheme presented in the next subsection.

Table 4: Statistics of the suboptimality gap for the solutions obtained from each approximation model under different values of demand interval (ε) and average over different Γ.

Γ%	ε	Avg. suboptimality gap			90-percentile suboptimality gap			Max. suboptimality gap		
		AARC	ELAARC	MLRC†	AARC	ELAARC	MLRC†	AARC	ELAARC	MLRC†
10	0.3	0.00	0.00	0.00	0.01	0.01	0.01	0.03	0.03	0.03
	0.4	0.01	0.01	0.01	0.02	0.02	0.02	0.04	0.04	0.04
	0.5	0.01	0.01	0.01	0.02	0.02	0.02	0.05	0.05	0.05
	0.6	0.03	0.02	0.02	0.04	0.04	0.03	0.07	0.06	0.06
	0.7	0.07	0.05	0.04	0.11	0.09	0.07	0.13	0.12	0.10
	0.8	0.18	0.11	0.08	0.27	0.14	0.11	0.33	0.17	0.16
	0.9	0.42	0.18	0.13	0.52	0.22	0.19	0.61	0.24	0.22
30	0.3	0.05	0.05	0.05	0.08	0.08	0.08	0.16	0.16	0.16
	0.4	0.11	0.11	0.11	0.19	0.19	0.18	0.26	0.26	0.26
	0.5	0.27	0.26	0.23	0.39	0.38	0.33	0.49	0.47	0.45
	0.6	0.58	0.50	0.41	0.75	0.62	0.52	0.88	0.72	0.64
	0.7	1.05	0.78	0.61	1.18	0.90	0.77	1.38	1.03	0.85
	0.8	1.58	1.06	0.81	1.74	1.20	1.04	1.92	1.35	1.14
	0.9	2.21	1.35	1.03	2.46	1.54	1.33	2.65	1.69	1.49
50	0.3	0.07	0.07	0.07	0.13	0.13	0.12	0.23	0.23	0.23
	0.4	0.18	0.18	0.16	0.33	0.33	0.25	0.46	0.46	0.42
	0.5	0.51	0.50	0.36	0.76	0.74	0.55	1.03	0.91	0.78
	0.6	1.31	1.14	0.68	1.60	1.38	1.01	1.76	1.51	1.22
	0.7	2.45	1.91	1.13	2.69	2.19	1.55	2.91	2.38	1.84
	0.8	3.82	2.78	1.65	4.11	3.17	2.25	4.25	3.34	2.55
	0.9	5.59	3.77	2.23	6.07	4.33	3.17	6.43	4.61	3.46
70	0.3	0.05	0.05	0.04	0.11	0.11	0.09	0.18	0.18	0.18
	0.4	0.17	0.17	0.12	0.32	0.32	0.19	0.45	0.45	0.32
	0.5	0.64	0.63	0.28	0.92	0.89	0.44	1.14	1.10	0.55
	0.6	2.26	1.97	0.57	2.58	2.24	0.84	2.74	2.55	1.09
	0.7	4.79	3.84	1.04	5.03	4.26	1.49	5.25	4.67	1.99
	0.8	8.23	6.27	1.73	8.60	7.01	2.52	8.82	7.32	3.16
	0.9	13.3	9.65	2.73	14.2	10.9	4.35	14.4	11.6	5.20
90	0.3	0.03	0.03	0.03	0.07	0.06	0.06	0.13	0.13	0.12
	0.4	0.12	0.12	0.10	0.21	0.21	0.17	0.29	0.29	0.26
	0.5	0.58	0.57	0.24	0.82	0.81	0.39	0.93	0.91	0.57
	0.6	2.56	2.14	0.49	2.94	2.53	0.77	3.22	2.72	1.05
	0.7	7.64	5.78	0.95	8.20	6.53	1.47	8.54	6.87	2.08
	0.8	16.3	12.07	1.78	17.2	13.4	2.99	17.7	14.5	3.90
	0.9	35.0	25.7	3.46	36.7	28.7	5.91	37.7	30.8	7.8

†This implementation of MLRC enforces $\boldsymbol{z}^3 = 0$. * Γ% denotes that the units are in % of $|\mathcal{J}|$.

Table 5: Proportion of the 3,500 problem instances analysed in Table 4 for which the relative suboptimality gap of each approximation model was within a certain range. Average gap and maximum gap are also reported.

Gap range	AARC	ELAARC	MLRC†
=0	1.46	1.51	1.54
≤0.1	26.6	27.9	33.1
≤1	57.6	67.1	77.3
≤10	91.4	93.3	100
Avg.	3.21	2.39	0.67
Max.	37.7	30.8	7.75

†This implementation of MLRC enforces $\boldsymbol{z}^3 = 0$.

Table 6: The improvement of suboptimality gap using SDP-LRC in comparison with other approximations.

ε	$\Gamma\%$	Avg. suboptimality gap				Max. suboptimality gap			
		AARC	ELAARC	MLRC†	SDP-LRC	AARC	ELAARC	MLRC†	SDP-LRC
0.7	50	2.45	1.91	1.13	0.07	2.91	2.91	1.84	0.27
	70	4.79	3.84	1.04	0.09	5.25	4.67	1.99	0.55
	90	7.64	5.78	0.95	0.16	8.54	6.87	2.08	0.71
0.8	50	3.82	2.78	1.65	0.11	4.25	3.34	2.55	0.64
	70	8.23	6.27	1.73	0.14	8.82	7.32	3.16	1.16
	90	16.3	12.1	1.78	0.22	17.7	14.5	3.90	0.92
0.9	50	5.59	3.77	2.23	0.16	6.43	4.61	3.46	0.93
	70	13.3	9.65	2.73	0.37	14.3	11.6	5.20	1.43
	90	35.0	25.7	3.46	0.51	37.7	30.8	7.75	2.52

†This implementation of MLRC enforces $\boldsymbol{z}^3 = 0$. * $\Gamma\%$ denotes that the units are in % of $|\mathcal{J}|$.

5.2.2 Computational analysis

In this subsection, we compare the computational time needed to solve AARC, ELAARC, and MLRC on the problem instances presented in Section 6 of Ardestani-Jaafari and Delage (2017). In order to produce a fair comparison, we extended the decomposition algorithm proposed in Ardestani-Jaafari and Delage (2017) to MLRC (see Appendix E for the detailed description). Specifically, we compare the computation time of each model (with and without decomposition) to the computational requirements associated to the exact column-and-constraint (C&CG) algorithm presented in Zeng and Zhao (2013) for varying sizes of problem instances, the budget of uncertainty Γ, and maximum demand perturbation ε that is defines as $\varepsilon_j = \hat{\zeta}_j/\bar{\zeta}_j\,, \forall j$. All approaches are implemented using Optimization Programming Language (OPL) within IBM ILOG CPLEX Optimization Studio 12.8.0 and a computing platform equipped with Intel Xeon E5-2667 processors, 2.9 GHz, and 32GB RAM.

Table 7 presents the computation time for the three problem instances. Note that in these instances the maximum demand perturbation ratio $\varepsilon_j := \hat{\zeta}_j/\bar{\zeta}_j$ is drawn randomly between 0.15 and 1 for all j. For each instance, we measured the impact of varying the budget of uncertainty between different proportions of the total number of locations.

Our first observation is that the computation time of MLRC is comparable with that of AARC and ELAARC when they are solved directly as MILP. We also observe that the C&CG algorithm requires much less effort than any of these direct implementations. It leaves us with the question of whether our conservative approximation models could also benefit from a well-designed decomposition scheme. Indeed, looking at the "Decomposition" columns, we notice that the time needed to solve all approximations can be significantly improved using our proposed decomposition algorithm. More precisely, we estimate that this algorithm is responsible for reducing the computation requirements for solving MLRC by a factor of 9 for the medium instance, and allowed us to solve in two hours on average, large scale problems that could not be solver in less than 2 days.The table also provides evidence that the C&CG algorithm is unable to converge in less than 48 hours for the instance when Γ equals 30% and 50% of the number of locations while all instances were solved in less than three hours using the decomposed MLRC approach. This seems to supports the idea that MLRC is more tractable than C&CG.

Table 8 presents the computation time for the same problem instances but with $\Gamma = 20\%|\mathcal{J}|$. For each instance, we measured the impact of varying the demand perturbation ε on computation time. The result of this table reconfirm the efficiency of our proposed decomposition scheme for MLRC. One can also observe that the MLRC model becomes more difficulty to solve as ε is increased.

Table 7: Computational time of robust approximations (in seconds).

$\|\mathcal{I}\|$	$\|\mathcal{J}\|$	$\Gamma\%^*$	MILP			Decomposition			C&CG
			AARC/LRC	ELAARC	MLRC†	AARC/LRC	ELAARC	MLRC†	Exact
10	20	10	3	5	7	5	5	9	1
		30	3	5	11	7	7	15	1
		50	4	7	10	7	7	15	1
		70	9	10	15	5	5	9	1
		90	6	8	14	2	2	2	1
		Avg.	5	7	11.4	5.2	5.2	10	1
20	40	10	489	775	891	67	81	81	8
		30	488	391	1074	89	81	121	16
		50	690	750	452	92	103	97	49
		70	1027	616	1339	92	94	147	30
		90	457	356	556	28	29	44	15
		Avg.	630.2	577.6	862.4	73.6	77.6	98	23.6
50	100	10	-	-	-	3,108	3,072	4,256	7,989
		30	-	-	-	4,263	4,146	6,593	-
		50	-	-	-	6,253	6,121	9,028	-
		70	-	-	-	5,453	5,882	9,271	7,364
		90	-	-	-	1,049	966	1,631	18
		Avg.	-	-	-	4,025.2	4,037.4	6,155.8	-

* $\Gamma\%$ denotes that the units are in % of $|\mathcal{J}|$. †This implementation of MLRC enforces $\boldsymbol{z}^3 = 0$.
$-$ denotes that resolution took longer than 48 hours.

Table 8: Computational time of robust approximations (in seconds).

$\|\mathcal{I}\|$	$\|\mathcal{J}\|$	ε	MILP			Decomposition			C&CG
			AARC/LRC	ELAARC	MLRC†	AARC/LRC	ELAARC	MLRC†	Exact
10	20	0.1	3	5	8	5	5	6	1
		0.3	4	6	9	5	8	7	1
		0.5	4	6	8	8	11	8	1
		0.7	4	6	9	7	8	5	1
		0.9	5	6	9	8	5	5	1
		Avg.	4	5.8	8.6	6.6	7.4	6.2	1
20	40	0.1	450	672	290	33	25	60	13
		0.3	735	301	481	30	36	56	14
		0.5	434	320	508	73	56	61	10
		0.7	247	263	467	105	73	158	8
		0.9	179	35	290	128	104	193	10
		Avg.	409	318.2	407.2	73.8	58.8	105.6	11
50	100	0.1	-	-	-	472	577	601	-
		0.3	-	-	-	3108	2958	3368	-
		0.5	-	-	-	5648	5553	5967	-
		0.7	-	-	-	6032	5840	6390	-
		0.9	-	-	-	6548	6349	6840	-
		Avg.	-	-	-	5334	5175	5641.25	-

†This implementation of MLRC enforces $\boldsymbol{z}^3 = 0$.

6 Numerical study with the multi-item newsvendor problem

The robust newsvendor problem has been studied since the early work of Scarf (1958). Since then, many researchers have contributed to this problem. In particular, Vairaktarakis (2000) studied robust multi-item newsvendor problem with interval uncertainty set and proposed tractable algorithm to obtain the optimal worst-case profit. Reccently, Ardestani-Jaafari and Delage (2017) studied robust multi-item newsvendor problem with polyhedral uncertainty set. They showed that affine policy can achieve an optimal worst-case profit when the uncertainty set is a budgeted uncertainty set. They also proposed a SDP based conservative approximation that we refer to as SDP-A&D hereafter. In this section, we compare the efficiency of our proposed approaches to the bounds obtained using AARC and SDP-A&D. In doing so, we will pay special attention to uncertainty sets that do not take the polyhedral form.

Consider the robust multi-item newsvendor problem described in Ardestani-Jaafari and Delage (2016):

$$\max_{\boldsymbol{x}\in\mathcal{X}} \min_{\boldsymbol{\zeta}\in\mathcal{U}} \sum_{j\in\mathcal{J}} r_j \min(x_j,\zeta_j) - c_j x_j + s_j \max(x_j-\zeta_j,0) - p_j \max(\zeta_j - x_j, 0)\,, \tag{43}$$

where, for each $j \in \mathcal{J} := 1,\ldots,n$, r_j, c_j, $s_j \le r_j$, and p_j denote the sale price, ordering cost, salvage price, and shortage cost of a unit of the j-th item respectively. Furthermore, for each j, the decision variable x_j denotes the number of units of item j that will be ordered, within a feasible set of joint order $\mathcal{X} \subseteq \mathbb{R}^n_+$, and each ζ_j denotes the realized demand for item j that is only known to be within an uncertainty set $\mathcal{U} \subset \mathbb{R}^n_+$ at the time of choosing $\boldsymbol{x}$. Note that the condition $\boldsymbol{s} \le \boldsymbol{r}$ states that salvage price is lower or equal to sales price, which is a natural assumption, and ensures that the problem is convex with respect to $\boldsymbol{x}$. It is well known that problem (43) is a special case of ARO, which is made explicit by the following equivalent form:

$$\begin{aligned}
\max_{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y}(\boldsymbol{\zeta})} \quad & \min_{\boldsymbol{\zeta}\in\mathcal{U}} \sum_{j\in\mathcal{J}} y_j(\boldsymbol{\zeta}) \\
\text{subject to} \quad & y_j(\boldsymbol{\zeta}) \le (r_j - c_j)x_j - (r_j - s_j)(x_j - \zeta_j)\,, \forall\, j\in\mathcal{J}\,, \forall\, \boldsymbol{\zeta}\in\mathcal{U} \\
& y_j(\boldsymbol{\zeta}) \le (r_j - c_j)x_j - p_j(\zeta_j - x_j)\,, \forall\, j\in\mathcal{J}\,, \forall\, \boldsymbol{\zeta}\in\mathcal{U}\,,
\end{aligned}$$

where each $y_j(\cdot)$ is a second-stage auxiliary variable that allows to account either for overstock and backlog of item j. Interestingly, in Ardestani-Jaafari and Delage (2016), the authors show that affine decision rules can actually be optimal for this problem when $\mathcal{U}$ is defined as the "budgeted uncertainty set":

$$\mathcal{U}(\Gamma) = \left\{ \boldsymbol{\zeta} \;\middle|\; \exists\, \boldsymbol{\delta}^+, \boldsymbol{\delta}^- \in \mathbb{R}^n,\; \begin{array}{l} \boldsymbol{\delta}^- \ge 0, \boldsymbol{\delta}^+ \ge 0 \\ \delta_j^+ + \delta_j^- \le 1\,, \forall\, j\in\mathcal{J} \\ \sum_{j\in\mathcal{J}} \delta_j^+ + \delta_j^- = \Gamma \\ \zeta_j = \bar{\zeta}_j + \hat{\zeta}_j(\delta_j^+ - \delta_j^-) \end{array} \right\},$$

where each $\bar{\zeta}_j$ and $\hat{\zeta}_j$ represent respectively the nominal demand and maximum deviation for item j, and Γ is a budget on the maximum number of items that might be perturbed[4]. Optimality of affine decision rules occurs under the condition that Γ is integer, and that the decision rules are defined in terms of both positive $\boldsymbol{\delta}^+$ and negative $\boldsymbol{\delta}^-$ deviations from the nominal amount.

A second interesting fact about this problem has to do with the structure of the feasible set of the dual recourse problem which takes the following form: $\boldsymbol{\lambda}^1 \ge 0$, $\boldsymbol{\lambda}^2 \ge 0$, and $\boldsymbol{\lambda}^1 + \boldsymbol{\lambda}^2 = 1$ for some $(\boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2) \in \mathbb{R}^n \times \mathbb{R}^n$. When attempting to employ MLRC to this problem, one quickly realizes that the optimal $\boldsymbol{u}^*$ obtained from problem (17) indicates that $\boldsymbol{\lambda}^i \le 1$, for $i \in \{1,2\}$, which is already implied by the constraints of the dual recourse problem. Based on Proposition 4, we can conclude that MLRC will not provide any improvement compared to the application of AARC.

The two properties of multi-item newsvendor problems described above indicate that only the valid conic inequalities, among the two types presented in Section 3, can offer improvement in accuracy and that this improvement can only occur in context that employ different uncertainty sets than $\mathcal{U}(\Gamma)$, with integer Γ. In particular, in what follows we will consider the following uncertainty set:

$$\tilde{\mathcal{U}}_P := \left\{ \boldsymbol{\zeta} \;\middle|\; \exists\, (\boldsymbol{\delta}^+, \boldsymbol{\delta}^-) \in \mathcal{U}_\delta,\; \boldsymbol{\zeta} = \bar{\boldsymbol{\zeta}} + \mathrm{diag}(\hat{\boldsymbol{\zeta}}) P (\boldsymbol{\delta}^+ - \boldsymbol{\delta}^-) \right\},$$

where $\boldsymbol{\delta}^+ \in \mathbb{R}^n_+$ and $\boldsymbol{\delta}^- \in \mathbb{R}^n_+$ capture respectively the positive and negative deviations from the nominal value and where $\mathcal{U}_\delta$ is either the lifted representation of a budgeted set, also known as the intersection of the ℓ_1-norm ball of radius Γ with the ℓ_∞-ball of radius one:

$$\mathcal{U}_\delta(\Gamma) := \left\{ (\boldsymbol{\delta}^+, \boldsymbol{\delta}^-) \in \mathbb{R}^n \times \mathbb{R}^n \;\middle|\; \begin{array}{l} \boldsymbol{\delta}^- \ge 0, \boldsymbol{\delta}^+ \ge 0 \\ \delta_j^+ + \delta_j^- \le 1\,, \forall\, j\in\mathcal{J} \\ \sum_{j\in\mathcal{J}} \delta_j^+ + \delta_j^- = \Gamma \end{array} \right\},$$

[4] Note that in Ardestani-Jaafari and Delage (2016) it is shown that $\le \Gamma$ and $= \Gamma$ are equivalent for this budgeted set yet the latter is the one needed for AARC to be proven exact.

where, for each j, the constraints $\delta_j^+ \geq 0$, $\delta_j^- \geq 0$, and $\delta_j^+ + \delta_j^- \leq 1$ capture the convex hull of $\left\{(\delta_j^+, \delta_j^-) \in \mathbb{R} \times \mathbb{R} \,\middle|\, 0 \leq \delta_j^- \leq 1,\ 0 \leq \delta_j^+ \leq 1,\ \delta_j^+ \delta_j^- = 0\right\}$, or the lifted representation of the intersection of the ℓ_2-norm ball of radius γ with the ℓ_∞-ball of radius one:

$$\mathcal{U}_\delta(\gamma) := \left\{ (\boldsymbol{\delta}^+, \boldsymbol{\delta}^-) \in \mathbb{R}^n \times \mathbb{R}^n \,\middle|\, \begin{array}{l} \boldsymbol{\delta}^- \geq 0, \boldsymbol{\delta}^+ \geq 0 \\ \delta_j^+ + \delta_j^- \leq 1,\, \forall j \in \mathcal{J} \\ \sqrt{\sum_{j \in \mathcal{J}} {\delta_j^+}^2 + {\delta_j^-}^2} \leq \gamma \end{array} \right\},$$

while $P \in \mathbb{R}^{n \times n}$ is a matrix that models the amount of correlation between each ζ_j (as discussed in Section 5 of Bertsimas and Sim (2004)). In particular, both versions of $\tilde{\mathcal{U}}_P$ have strong statistical motivation when the triplet $(\bar{\boldsymbol{\zeta}}, \hat{\boldsymbol{\zeta}}, P)$ is chosen such that each perturbation $\delta_j := \delta_j^+ - \delta_j^-$ can be considered independent, unbiased (i.e. with an expected value of zero), and supported on $[-1, 1]$ (see Chapter 2.3 in Ben-Tal et al. (2009)). We refer the reader to Appendix F for a simple three items example where it is shown that SDP-LRC can strictly improve on AARC even under $\mathcal{U}_\delta(\Gamma)$ when $\Gamma = 2$ as long as $P \neq I$.

The result of Ardestani-Jaafari and Delage (2016) also motivates us to compare the performance of AARC and SDP-LRC on the so-called "lifted" reformulation of problem (43):

$$\begin{aligned}
\max_{\boldsymbol{x} \in \mathcal{X}, \boldsymbol{y}(\boldsymbol{\delta}^+, \boldsymbol{\delta}^-)} \quad & \min_{(\boldsymbol{\delta}^+, \boldsymbol{\delta}^-) \in \mathcal{U}_\delta} \sum_{j \in \mathcal{J}} y_j(\boldsymbol{\delta}^+, \boldsymbol{\delta}^-) && \text{(44a)} \\
\text{subject to} \quad & y_j(\boldsymbol{\delta}^+, \boldsymbol{\delta}^-) \leq (r_j - c_j)x_j - (r_j - s_j)(x_j - \zeta_j(\boldsymbol{\delta}^+, \boldsymbol{\delta}^-)),\, \forall j \in \mathcal{J}, \forall (\boldsymbol{\delta}^+, \boldsymbol{\delta}^-) \in \mathcal{U}_\delta && \text{(44b)} \\
& y_j(\boldsymbol{\delta}^+, \boldsymbol{\delta}^-) \leq (r_j - c_j)x_j - p_j(\zeta_j(\boldsymbol{\delta}^+, \boldsymbol{\delta}^-) - x_j),\, \forall j \in \mathcal{J}, \forall (\boldsymbol{\delta}^+, \boldsymbol{\delta}^-) \in \mathcal{U}_\delta, && \text{(44c)}
\end{aligned}$$

where $\boldsymbol{\zeta}(\boldsymbol{\delta}^+, \boldsymbol{\delta}^-) := \bar{\boldsymbol{\zeta}} + \mathrm{diag}(\hat{\boldsymbol{\zeta}})P(\boldsymbol{\delta}^+ - \boldsymbol{\delta}^-)$ since this is the version of the problem where AARC is known to return optimal solutions under some conditions.

In what follows, we compare the empirical performance (in both accuracy and numerical efficiency) of the following approaches for solving problem (44) when $\mathcal{U}_\delta$ is polyhedral and non-polyhedral, i.e. $\mathcal{U}_\delta(\Gamma)$ and $\mathcal{U}_\delta(\gamma)$ respectively:

- AARC employs affine decision rules to solve problem (44)
- (G)SDP-LRC employs the reformulation obtained using $\max_{\boldsymbol{x} \in \mathcal{X}} g_{\text{(G)SDP-LRC}}(\boldsymbol{x})$ with $\mathbb{M}_1 = I$ and $K = 1$.
- (G)SDP-LRC2 employs the reformulation obtained using $\max_{\boldsymbol{x} \in \mathcal{X}} g_{\text{(G)SDP-LRC}}(\boldsymbol{x})$ with

  $$\mathbb{M}_k := \begin{bmatrix} \boldsymbol{e}_k \mathbf{1}^\top & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 1 \end{bmatrix},\ \forall k = 1, \ldots, \bar{m},$$

  such that each of the constraint (39b) reduces to:

  $$\begin{bmatrix} \Lambda_{k,k} & \Delta_{:,k}^\top & \lambda_k \\ \Delta_{:,k} & \Xi & \boldsymbol{\xi} \\ \lambda_k & \boldsymbol{\xi}^\top & 1 \end{bmatrix} \succeq 0,\ \forall k = 1, \ldots, \bar{m},$$

  where $\boldsymbol{\xi} := [{\boldsymbol{\delta}^+}^\top {\boldsymbol{\delta}^-}^\top]^\top$.
- SDP-A&D refers to the semi-definite programming model proposed in Ardestani-Jaafari and Delage (2016). In the context of $\mathcal{U}_\delta(\gamma)$, we employ SDP-A&D from problem (7) from Ardestani-Jaafari and Delage (2016) with $A := [I - I]^\top$ and $\boldsymbol{b} := 1$, and add the non linear constraint that $\sqrt{\sum_{j \in \mathcal{J}} {\zeta_j^+}^2 + {\zeta_j^-}^2} \leq \gamma$, following the notation of Ardestani-Jaafari and Delage (2016).

Note that in the case of the polyhedral set $\mathcal{U}_\delta(\Gamma)$, the two methods SDP-LRC2 and SDP-A&D are equivalent while SDP-LRC2 can be shown to be tighter than SDP-A&D for the non-polyhedral set $\mathcal{U}_\delta(\gamma)$. This is because our extension of SDP-A&D to general uncertainty sets does not impose constraints (38c), (38d), and (38e). All experiments performed in this section were done on a Mac Pro 2.7 GHz 12-Core Intel Xeon E5 with 64 Go of memory. Furthermore, while AARC and the C&CG method were implemented using CPLEX 12.8, the semi-definite programs associated to (G)SDP-LRC, (G)SDP-LRC2, and SDP-A&D, were solved using Mosek 7 (see ApS (2017)) through the YALMIP interface (see Löfberg (2004)). All algorithms were implemented using a single-thread.

Finally, the accuracy of each approach will be compared using their respective "relative bound gap", i.e., the relative difference between the optimal worst-case profit of the problem instance and the optimal value of the conservative approximation model.

6.1 Experimental set-up

In our experiments, each problem instance is randomly generated by drawing each sale price r_j uniformly from [0, 100], each ordering cost c_j from $[0, r_j]$, each salvage price s_j from $[0, c_j]$, each shortage cost p_j from $[0, c_j]$. The uncertainty in the demand is characterized by drawing each $\hat{\bar{\zeta}}_j$ from [0, 100] and each maximum deviation $\hat{\zeta}_j$ from $[0, \bar{\zeta}_j]$. Finally, the matrix P is constructed based on a randomly generated correlation matrix $\mathfrak{C}(\nu)$ using a variant of the vine method presented in Lewandowski et al. (2009) which samples partial correlations from a beta distribution with parameters $\alpha = \beta = \nu$ and rescaled to the $[-1, 1]$ interval. Note that ν controls the amount of correlation and is itself uniformly chosen from the range $\{0.01, 0.03, 0.1, 0.32, 1, 3.16, 10, 31.62, 100, \infty\}$. While, for brevity, we do not provide further details about the generation of $\mathfrak{C}(\nu)$, one can observe in Figure 1 a distribution of the average absolute correlation found in the non-diagonal element of $\mathfrak{C}(\nu)$ in our experiments. For completeness, we also mention that $\mathcal{X} = \mathbb{R}^n_+$ throughout.

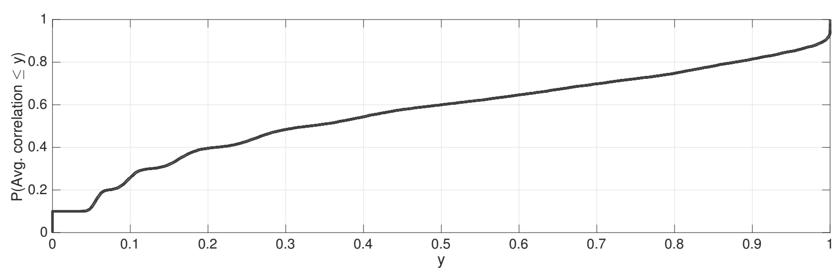


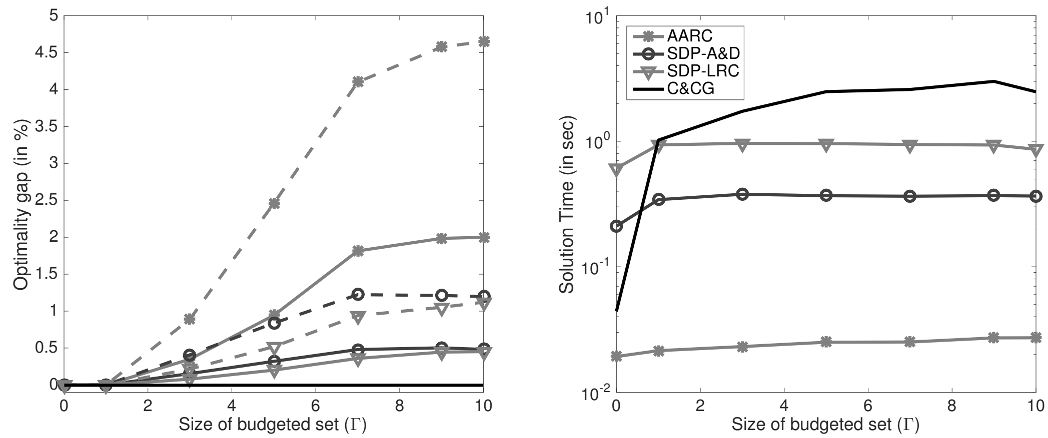
Figure 1: Distribution of average absolute correlation captured by P in the randomly generated multi-item newsvendor problem instances that used in the experiments of Section 6.

6.2 Performance comparison for polyhedral uncertainty sets $\mathcal{U}_\delta(\Gamma)$

In a first set of experiments, we evaluate the performance of each approach on a set of 1000 randomly generated problem instances with $n = 10$ items. For each of these instances, we also studied the sensitivity of performances with respect to the budget $\Gamma \in \{0, 1, 3, 5, 7, 9, 1\}$. Figure 2(a) presents the average relative bound gap achieved by AARC, SDP-A&D, SDP-LRC in the presence of optimal worst-case profit obtained by C&CG, together with the 90-th percentile of this gap in the set of experiments. Figure 2(b) presents the average computation time needed by each method to produce an optimal solution. Looking at Figure 2(a), one might observe that the quality of the solution proposed by AARC degrades as the amount of uncertainty increases. In particular, the average gap decreases from about 2.2% to nearly 0.5% by employing an approach that is based on a semi-definite programming formulation (either SDP-A&D or SDP-LRC). One also notes that the improvement from SDP-A&D/SDP-LRC2 to SDP-LRC is only marginal. On the other hand, when looking at Figure 2(b), we notice that when $\Gamma \geq 1$ all approximation methods provide on average their solution in a time that is reasonably faster than C&CG, with AARC performing on average 10 to 100 times faster than the other schemes.

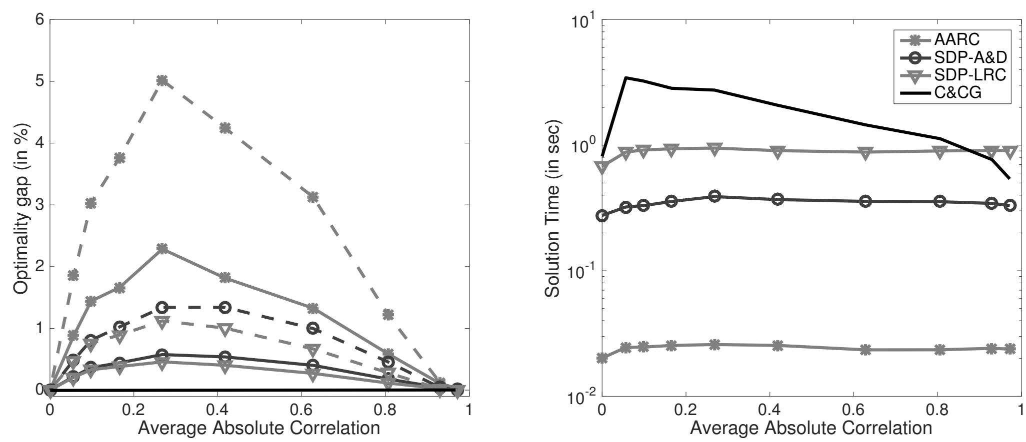
We finally present in Figure 3 the quality of solutions and efficiency of all methods as a function of average absolute correlation between items. One can confirm from Figure 3(a) that the multi-item newsvendor problem becomes harder to approximate when the absolute correlation between items is of medium size.

Indeed, for small size of correlation, the theory states that AARC becomes optimal for this problem. On the other hand of the spectrum, when absolute correlation is large then the uncertainty about $\boldsymbol{\zeta}$ can asymptotically be captured using a single factor, i.e., $P = \boldsymbol{f}\boldsymbol{g}^\top$, which in turns allows AARC to become exact once again.



(a) (b)

Figure 2: Comparison of performances achieved by all methods for different values of budgets when using the polyhedral uncertainty set $\mathcal{U}_\delta(\Gamma)$. (a) presents the average (solid line) and 90-th percentile (dashed line) of the relative bound gap that is achieved a set of 1000 experiments for different values of Γ. (b) presents the average computation time needed by each method to produce an optimal solution for different values of Γ.

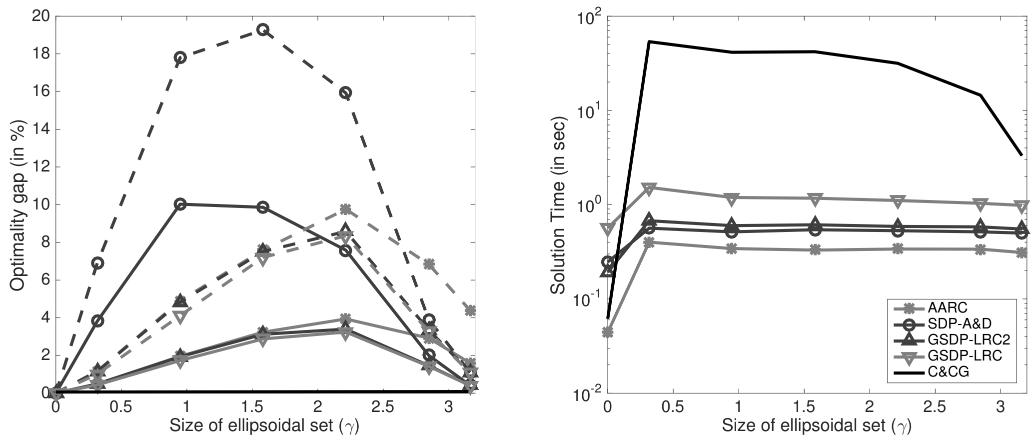


(a) (b)

Figure 3: Comparison of performances achieved by all methods as a function of average absolute correlation between items when using the polyhedral uncertainty set $\mathcal{U}_\delta(\Gamma)$. (a) presents the average (solid line) and 90-th percentile (dashed line) of the relative bound gap that is achieved a set of 1000 experiments for different values of Γ. (b) presents the average computation time needed by each method to produce an optimal solution for different values of Γ.

6.3 Performance comparison for non-polyhedral uncertainty sets $\mathcal{U}_\delta(\gamma)$

In this second set of experiments, we start by evaluating the performance of each approach on a set of 100 randomly generated problem instances with $n = 10$ items. For each of these instances we also studied the sensitivity of performances with respect to the budget $\gamma \in \{0,\ 0.1\sqrt{n},\ 0.3\sqrt{n},\ 0.5\sqrt{n},\ 0.7\sqrt{n},\ 0.9\sqrt{n},\ \sqrt{n}\}$. Figure 4(a) presents the average relative bound gap achieved by AARC, SDP-A&D, SDP-LRC2, SDP-LRC, and C&CG, together with the 90-th percentile of this gap in the set of experiments. Figure 4(b) presents the average computation time needed by each method to produce an optimal solution. Looking at Figure 4(a), one might now be surprised by the fact that SDP-A&D achieves a worst performance than AARC for midrange sizes of γ. This is due to the fact that SDP-A&D was only designed in a way that ensures that it provide tighter bounds than AARC for polyhedral sets while this property is lost for non-polyhedral ones. We could conclude from this that SDP-LRC2, which brings into play constraints (38c), (38d), and (38e), is the better way of extending this approach to such uncertainty sets. We can also observe, as was the case for $\mathcal{U}_\delta(\Gamma)$, that the quality of solutions are only marginally improved by using SDP-LRC instead of SDP-LRC2 while SDP-LRC is much less attractive from a computational point of view. Finally, Figure 4(b) clearly reveals the inefficiency of C&CG in contexts of non-polyhedral uncertainty sets.



(a) (b)

Figure 4: Comparison of performances achieved by each method for different values of budgets when using the non-polyhedral uncertainty set $\mathcal{U}_\delta(\Gamma)$. (a) presents the average (solid line) and 90-th percentile (dashed line) of the relative bound gap that is achieved a set of 100 experiments for different values of Γ. (b) presents the average computation time needed by each method to produce an optimal solution for different values of Γ.

We conclude this section by presenting Table 9 which reports the average computation time and solution quality of AARC, SDP-A&D, and SDP-LRC2 on problem instances which sizes range from 10 to 100 items. Specifically, it presents the average relative improvement achieved by both SDP-A&D and SDP-LRC2 compared to AARC in terms of optimal bound on best achievable worst-case profit. Based on these results, it appears that SDP-LRC2 consistently provides a tighter bound than both AARC and SDP-A&D. In problems with a large number of items $n = 100$, SDP-LRC2 does give rise to a heavier computation burden than AARC but produces solutions of noticeably better quality, i.e. 2 to 5 % improvement on average when $\gamma \in [0.5\sqrt{n},\ 0.9\sqrt{n}]$. On the other hand, SDP-LRC2's numerical requirements appears to be only slightly more important than SDP-A&D.

7 Conclusions

In this paper, we proposed a linearization scheme that can be used to construct tractable conservative approximation models for two-stage adjustable robust optimization problems with right-hand side uncertainty. We showed that this scheme provides an alternate interpretation of models obtained through the use of AARC. Yet, by considering the adversarial problem as a "bilinear program" that can be linearized, it becomes very natural to identify modifications based on valid linear and conic inequalities that will improve LRC and consequently provide tightening procedures for AARC. Based on these results, it is clear that the LRC model can help us to better understand the quality of solutions obtained from AARC and offers a perspective that might help design better approximation methods for two-stage adjustable robust optimization models (e.g., the penalty based method presented in Proposition 5) . We finally extensively surveyed the types of improvement that our proposed models might offer in two different operations management applications.

Table 9: Average computation time and improvement of quality of bounds obtained from SDP-A&D and SDP-LRC2 compared to AARC taken over 10 randomly generated problem instances, for different problem sizes n and uncertainty sets γ.

	γ	Computational time (in sec)			Improvement (in %)	
	(in % of $\sqrt{n}$)	AARC	SDP-A&D	SDP-LRC2	SDP-A&D	SDP-LRC2
10	10	<1	<1	<1	-2.8	0.0
	30	<1	<1	<1	-5.5	0.0
	50	<1	<1	<1	-7.8	0.1
	70	<1	<1	<1	-1.6	0.2
	90	<1	<1	<1	9.8	10.2
25	10	7	8	10	-2.9	0.0
	30	7	7	10	-5.1	0.1
	50	8	7	11	-5.2	0.5
	70	8	8	10	-0.7	2.0
	90	11	8	11	1.6	1.8
50	10	104	174	218	-2.8	0.0
	30	124	169	220	-4.0	0.3
	50	129	180	223	-2.5	1.1
	70	158	187	241	0.5	2.5
	90	187	216	264	4.5	4.7
100	10	1605	8092	11135	-2.5	0.1
	30	1849	8251	11438	-3.3	0.6
	50	2100	9079	11819	-0.8	2.1
	70	2765	9941	12410	2.0	3.5
	90	3141	10119	11534	4.8	5.0

Since the publication of the thesis of Ardestani-Jaafari (2016) on which this article is based, a few studies have further investigated the connections between two-stage robust optimization problems and reformulation schemes for non-convex quadratic programs, and in particular with copositive programming. In Hanasusanto and Kuhn (2016), the authors consider a generalization of our ARO problem (1) which takes the shape of a distributionally robust problem that is based on Wasserstein balls where the uncertainty lies in both the objective and right-hand side of the recourse problem. Unlike our Assumption 2, they assume that the problem has complete recourse and present an equivalent copositive program. In the case where complete recourse does not hold (as in the examples of Section 5 and Appendix B), the authors propose a sequence of copositive programs known to provide solutions that converge to the optimal solution. In an article developed independently, Xu and Burer (2016) also present an equivalent copositive programming reformulation for two-stage robust optimization problems with right-hand side uncertainty. Unlike in Hanasusanto and Kuhn (2016), their approach can handle convex uncertainty set of arbitrary structure. The authors also establish a connection with AARC and a conic programming reformulation that is obtained by approximating the copositive cone from within. In a recent revision, the authors were inspired by our idea of bounding the dual recourse variable $\boldsymbol{\lambda}$ in order to provide an exact copositive programming reformulation for relatively complete recourse problem although no connections to AARC have yet been identified for the resulting reformulation.

In view of this recent literature, one might wonder how MLRC/AARC and SDP-LRC models derived in Section 3 relate to these recent copositive programming reformulations. In an attempt to answer this question, in Appendix G we reuse the ideas presented in of both Hanasusanto and Kuhn (2016) and Xu and Burer (2016) to identify a copositive programming that is exactly equivalent to the ARO problem (1) and which reduces to MLRC and SDP-LRC under an appropriate choice of inner approximations of the cone of copositive matrices. We hope this can help readers that plan on designing approximations of ARO through the copositive programming paradigm to establish useful connections with the models presented in this paper.

Appendix A Additional details regarding Example 1

Based on the numerical optimization of the ARO problem (3) described in this example, we identified that an optimal two-stage solution takes the configuration presented in Table 10. Note that in this table we only present the shipments made from facility at location #1 since the optimal solution does not open any facility at location #2. On the other hand, while it is clear that the optimal solution of the AARC model simply prescribes a production capacity and shipments of zero units under any circumstances, it can be interesting to look at the strategy that would be recommended by AARC under the condition that it was forced to open a facility at location #1. Table 11 presents the details of such a solution. In particular, the initial production capacity is set to 42699 in this context, while the affine decision rule recommends the following shipments to the three locations:

$$\bar{y}_{11}(\delta) := 12949 - 12949\delta_1 + 2000\delta_2 + 2000\delta_3$$
$$\bar{y}_{12}(\delta) := 13750 - 13750\delta_2 + 2000\delta_1 + 2000\delta_3$$
$$\bar{y}_{13}(\delta) := 16000 - 16000\delta_3 + 2000\delta_1 + 2000\delta_2$$

and let $\boldsymbol{y}(\boldsymbol{\zeta}) := \bar{\boldsymbol{y}}((\boldsymbol{\zeta} - 20000)/18000)$. Overall, this policy achieves a worst-case profit of -4619 which motivates the closure of the facility.

Based on the analysis of these two solutions, we observe that, once the cost of opening the facility is considered a sunk cost, it is actually possible for the affine policy to make profits by preparing some production capacity. The policy fails to make reasonable profits however because in order to generate profits in the worst-case scenarios where the demand is low, it also needs to plan even larger shipments (by linearity of the policy) for the less pessimistic scenarios. This in turns requires a larger production capacity which does not pay off given that we are only interested in the worst-case profit. The ARO model is instead capable of planning some shipments for the low demand scenarios without being forced to make bigger shipments in other situations thus can better control the production capacity in order to perform better in terms of the worst-case profit.

Table 10: Detailed optimal solution of problem (3).

Scenario			Prod. capacity	Shipment from facility #1			First stage cost	Second stage profit	Total profit
ζ_1	ζ_2	ζ_3		Loc. #1	Loc. #2	Loc. #3			
20000	20000	20000	24000	20000	4000	0	114400	140400	26000
20000	20000	2000	24000	20000	4000	0	114400	140400	26000
2000	20000	20000	24000	2000	20000	2000	114400	133600	19200
20000	2000	20000	24000	20000	2000	2000	114400	139000	24600
2000	2000	20000	24000	2000	2000	20000	114400	121000	6600
2000	20000	2000	24000	2000	20000	2000	114400	133600	19200
20000	2000	2000	24000	20000	2000	2000	114400	139000	24600

Table 11: Detailed optimal solution of AARC approximation of problem (3) when facility #1 is opened.

Scenario			Prod. capacity	Shipment from facility #1			First stage cost	Second stage profit	Total profit
ζ_1	ζ_2	ζ_3		Loc. #1	Loc. #2	Loc. #3			
20000	20000	20000	42699	12949	13750	16000	125619	231800	106181
20000	20000	2000	42699	14949	15750	0	125619	176400	50781
2000	20000	20000	42699	0	15750	18000	125619	176400	50781
20000	2000	20000	42699	14949	0	18000	125619	176400	50781
2000	2000	20000	42699	2000	2000	20000	125619	121000	-4619
2000	20000	2000	42699	2000	17750	2000	125619	121000	-4619
20000	2000	2000	42699	16949	2000	20003	125619	121000	-4619

Appendix B Multi-product assembly problem

In the multi-product assembly problem discussed in (Shapiro et al. 2009, Chapter 1), a manufacturer produces n products using m different types of parts. It is a two-stage problem wherein, the manufacturer pre-orders x_i units for part $i \in \mathcal{I} := \{1, \ldots, m\}$ with a cost of c_i per unit in the first stage; and when demand is realized, it must be determined how many products, y_j, to make for each type $j \in \mathcal{J} := \{1, \ldots, n\}$. The robust multi-product assembly problem can be formulated as follows:

$$\begin{aligned}
\underset{\boldsymbol{x},\boldsymbol{y}(\boldsymbol{\zeta})}{\text{maximize}} \quad & \min_{\boldsymbol{\zeta}\in\mathcal{U}} \; -\boldsymbol{c}^\top\boldsymbol{x} + (\boldsymbol{q}-\boldsymbol{l})^\top\boldsymbol{y}(\boldsymbol{\zeta}) + \boldsymbol{s}^\top(\boldsymbol{x} - A\boldsymbol{y}(\boldsymbol{\zeta})) && \text{(45a)}\\
\text{subject to} \quad & \boldsymbol{y}(\boldsymbol{\zeta}) \le \boldsymbol{\zeta}\,, \forall\,\boldsymbol{\zeta}\in\mathcal{U} && \text{(45b)}\\
& A\boldsymbol{y}(\boldsymbol{\zeta}) \le \boldsymbol{x}\,, \forall\,\boldsymbol{\zeta}\in\mathcal{U} && \text{(45c)}\\
& \boldsymbol{y}(\boldsymbol{\zeta}) \ge 0\,, \forall\,\boldsymbol{\zeta}\in\mathcal{U} && \text{(45d)}\\
& 0 \le \boldsymbol{x} \le M\,, && \text{(45e)}
\end{aligned}$$

where $\boldsymbol{\zeta} \in \mathbb{R}^n$ is the uncertain demand for each product and where parameters $\boldsymbol{q}$ and $\boldsymbol{l}$ denote, respectively, the selling price and production cost per unit of the products, while $\boldsymbol{s}$ denotes the salvage unit value of unused parts. Furthermore A_{ij} denotes the number of units of part i that is required to assemble product j. Finally, the uncertain demand $\boldsymbol{\zeta}$ is assumed to lie in the following budgeted uncertainty set $\mathcal{U}$:

$$\mathcal{U} = \{\boldsymbol{\zeta} | \exists\, \boldsymbol{\delta} \in [0,1]^n,\, \zeta_j = \bar{\zeta}_j - \hat{\zeta}_j\delta_j\,,\, \forall\, j\,,\, \sum_j \delta_j \le \Gamma\}\,,$$

where $\bar{\zeta}_j$ and $\hat{\zeta}_j$ denote the nominal demand and the interval demand, respectively, for all j.

As was done for the previous example, one can hope to identify a tighter conservative approximation than with AARC by employing affine decision rules in the following augmented model:

$$\begin{aligned}
\underset{\boldsymbol{x},\boldsymbol{y}(\boldsymbol{\zeta}),\boldsymbol{z}^1(\boldsymbol{\zeta}),\boldsymbol{z}^2(\boldsymbol{\zeta}),\boldsymbol{z}^3(\boldsymbol{\zeta})}{\text{maximize}} \quad & \min_{\boldsymbol{\zeta}\in\mathcal{U}} \; -\boldsymbol{c}^\top\boldsymbol{x} + (\boldsymbol{q}-\boldsymbol{l})^\top\boldsymbol{y}(\boldsymbol{\zeta}) + \boldsymbol{s}^\top(\boldsymbol{x} - A\boldsymbol{y}(\boldsymbol{\zeta})) - {\boldsymbol{u}^1}^\top\boldsymbol{z}^1(\boldsymbol{\zeta}) - {\boldsymbol{u}^2}^\top\boldsymbol{z}^2(\boldsymbol{\zeta}) - {\boldsymbol{u}^3}^\top\boldsymbol{z}^3(\boldsymbol{\zeta})\\
\text{subject to} \quad & \boldsymbol{y}(\boldsymbol{\zeta}) \le \boldsymbol{\zeta} + \boldsymbol{z}^1(\boldsymbol{\zeta})\,, \forall\,\boldsymbol{\zeta}\in\mathcal{U}\\
& A\boldsymbol{y}(\boldsymbol{\zeta}) \le \boldsymbol{x} + \boldsymbol{z}^2(\boldsymbol{\zeta})\,, \forall\,\boldsymbol{\zeta}\in\mathcal{U}\\
& \boldsymbol{y}(\boldsymbol{\zeta}) \ge 0 - \boldsymbol{z}^3(\boldsymbol{\zeta})\,, \forall\,\boldsymbol{\zeta}\in\mathcal{U}\\
& \boldsymbol{z}^1(\boldsymbol{\zeta}) \ge 0\,, \forall\,\boldsymbol{\zeta}\in\mathcal{U}\\
& \boldsymbol{z}^2(\boldsymbol{\zeta}) \ge 0\,, \forall\,\boldsymbol{\zeta}\in\mathcal{U}\\
& \boldsymbol{z}^3(\boldsymbol{\zeta}) \ge 0\,, \forall\,\boldsymbol{\zeta}\in\mathcal{U}\\
& 0 \le \boldsymbol{x} \le M\,,
\end{aligned}$$

where $\boldsymbol{z}^1 : \mathbb{R}^n \to \mathbb{R}^n$, $\boldsymbol{z}^2 : \mathbb{R}^n \to \mathbb{R}^m$, and $\boldsymbol{z}^3 : \mathbb{R}^n \to \mathbb{R}^n$ can be interpreted as violation adjustments for constraints (45b), (45c), and (45d). Yet, in this case, the $\boldsymbol{u}$ bounds are obtained from the dual problem:

$$\begin{aligned}
\underset{\boldsymbol{\lambda}^1,\boldsymbol{\lambda}^2,\boldsymbol{\lambda}^3}{\text{minimize}} \quad & \boldsymbol{\zeta}^\top\boldsymbol{\lambda}^1 + \boldsymbol{x}^\top\boldsymbol{\lambda}^2 && \text{(46a)}\\
\text{subject to} \quad & \lambda^1_j + \sum_{i\in\mathcal{I}} A_{ij}\lambda^2_i - \lambda^3_j = q_j - l_j + A_{:j}^\top\boldsymbol{s}\,, \forall\, j \in \mathcal{J} && \text{(46b)}\\
& \boldsymbol{\lambda}^1 \ge 0\,,\; \boldsymbol{\lambda}^2 \ge 0\,,\; \boldsymbol{\lambda}^3 \ge 0\,, && \text{(46c)}
\end{aligned}$$

where $\boldsymbol{\lambda}^1 \in \mathbb{R}^n$, $\boldsymbol{\lambda}^2 \in \mathbb{R}^m$, and $\boldsymbol{\lambda}^3 \in \mathbb{R}^n$ are the dual variables associated to constraints (45b), (45c), and (45d). Here again, the objective function is non-decreasing in $\boldsymbol{\lambda}^1$ and $\boldsymbol{\lambda}^2$ so that, at optimum, each term of these two vectors is either zero or is involved in at least one active constraint among the following set:

$$\lambda^1_j + \sum_{i\in\mathcal{I}} A_{ij}\lambda^2_i \ge q_j - l_j + A_{:j}^\top\boldsymbol{s}\,, \forall\, j \in \mathcal{J}\;.$$

This indicates to us that

$$\lambda_j^{1*} \leq \max\left(0\,,\, q_j - l_j + A_{:j}^\top \boldsymbol{s} - \sum_{i \in \mathcal{I}} A_{ij}\lambda_i^{2*}\right) \leq \max(0\,,\, q_j - l_j + A_{:j}^\top \boldsymbol{s}) := u_j^1\,,$$

and that

$$\lambda_i^{2*} \leq \max(0\,,\, \max_{j \in \mathcal{J}_i} \frac{1}{A_{ij}}(q_j - l_j + A_{:j}^\top \boldsymbol{s} - \lambda_j^1 - \sum_{i' \neq i} A_{ij}\lambda_i^{2*})) \leq \max(0\,,\, \max_{j \in \mathcal{J}_i} \frac{1}{A_{ij}}(q_j - l_j + A_{:j}^\top \boldsymbol{s})) := u_i^2\,,$$

where the set of indices $\mathcal{J}_i := \{j \,|\, A_{ij} \neq 0\}$. Finally, since $\boldsymbol{\lambda}^3$ is uniquely determined based on $\boldsymbol{\lambda}^1$ and $\boldsymbol{\lambda}^2$, we have that

$$\lambda_j^3 = \lambda_j^1 + \sum_{i \in \mathcal{I}} A_{ij}\lambda_i^2 - q_j + l_j - A_{:j}^\top \boldsymbol{s} \leq u_j^1 + \sum_{i \in \mathcal{I}} A_{ij} u_i^2 - q_j + l_j - A_{:j}^\top \boldsymbol{s} := u_j^3\,.$$

We conclude this example with a description of the specific context in which exploiting the information about the bound $\boldsymbol{u}$ on $\boldsymbol{\lambda}^*$ leads to a strictly tighter conservative approximation. In particular, consider a multi-product assembly problem with three products and two different types of parts. The pre-order variable $\boldsymbol{x}$ is bounded by 100,000, the cost of parts A and B are, respectively, \$25 per unit and \$3 per unit, while the salvage value is \$4 per unit and \$1 per unit. Furthermore, the difference between the selling price and the unit production cost of each product is: \$380/unit, \$800/unit, and \$1200/unit respectively for products #1 to #3. Next, we have that product #1 requires 9 units of both parts, product #2 requires 5 units of part B, and #3 requires 9 units of A and 4 units of B. Finally, for products #1 to #3, the nominal demand is respectively of 9000, 10,000, and 8000 units while the worst-case demand for each is 1000, 2000, and 0 units respectively. In this specific context, one can exploit the above closed-form bounds $\boldsymbol{u}^1 := [\ 425 \quad 805 \quad 1240\ ]^\top$, $\boldsymbol{u}^2 := [\ 140 \quad 310\ ]^\top$, and $\boldsymbol{u}^3 := [\ 4050 \quad 1550 \quad 2500\ ]^\top$. However, using problem (17), with $M := 4050$, allows us to tighten these bounding vectors even more:

$$\boldsymbol{u}^{1^*} := \begin{bmatrix} 335 \\ 795 \\ 1160 \end{bmatrix} \qquad \boldsymbol{u}^{2^*} := \begin{bmatrix} 129 \\ 290 \end{bmatrix} \qquad \boldsymbol{u}^{3^*} := \begin{bmatrix} 2275 \\ 655 \\ 0 \end{bmatrix}.$$

As it is shown in Table 12, when the budget of uncertainty is set to $\Gamma = 2$, a direct application of affine decision rules in problem (45) will lead to a worst-case profit estimated at 2.474 million dollars; meanwhile employing affine decision rules in the equivalent formulation that allows penalized violations achieves a worst-case profit estimated at 2.722 million dollars (namely a 10% increase in profit). This confirms that the MLRC model can provide a strictly tighter conservative approximation for this type of problem.

Table 12: Optimal solution of AARC and MLRC in the instance of multi-product assembly problem.

	AARC/LRC	MLRC	Exact model
# of parts type A	92,793	81,000	81,000
# of parts type B	91,000	91,000	91,000
Optimal bound on worst-case profit	\$2.474 million	\$2.722 million	\$2.722 million
Worst-case profit of solution	\$2.474 million	\$2.722 million	\$2.722 million

Appendix C Relation to AARC for general uncertainty sets

For simplicity, we present the connection between GLRC and AARC for a convex uncertainty set described as $\mathcal{U}_{general} := \{\boldsymbol{\zeta} \in \mathbb{R}^{n_\zeta} \,|\, f(\boldsymbol{\zeta}) \leq 0\}$ and when no bounds are known for the dual variables $\boldsymbol{\lambda} \in \mathbb{R}^{\bar{m}}$.

Proposition 7 *Given that $f(\cdot)$ satisfies Assumption 5, the GLRC model presented below provides a tighter conservative approximation than the AARC model presented in (2) when the uncertainty set is described as $\mathcal{U}_{general}$:*

$$\underset{\boldsymbol{x} \in \mathcal{X}}{\text{maximize}} \quad g_{GLRC}(\boldsymbol{x})\,,$$

where

$$\begin{aligned} g_{GLRC}(\boldsymbol{x}) := \min_{\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \mathrm{tr}(\Psi(\boldsymbol{x})\Delta) - (A\boldsymbol{x})^\top \boldsymbol{\lambda} && (47a)\\ \text{subject to} \quad & B^\top \boldsymbol{\lambda} = \boldsymbol{d} && (47b)\\ & f(\boldsymbol{\zeta}) \leq 0 && (47c)\\ & \boldsymbol{\lambda} \geq 0 && (47d)\\ & \Delta B = \boldsymbol{\zeta}\boldsymbol{d}^\top && (47e)\\ & h(\Delta_{:i}, \lambda_i) \leq 0\,,\ \forall\, i = 1,\ldots,\bar{m}\,. && (47f) \end{aligned}$$

Proof. Based on Definition 1, constraint (47f) can be explicitly described as

$$\sup_{\boldsymbol{z}} \Delta_{:i}^\top \boldsymbol{z} - \lambda_i f_*(\boldsymbol{z}) \leq 0\,,\ \forall\, i = 1,\ldots,\bar{m}\,.$$

One can then construct the Lagrangian function of problem (47) using the following form :

$$\begin{aligned} \mathcal{L}(\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta,\boldsymbol{y},Y,\boldsymbol{s}) :=\ & \boldsymbol{c}^\top \boldsymbol{x} + \mathrm{tr}(\Psi(\boldsymbol{x})\Delta) - (A\boldsymbol{x})^\top \boldsymbol{\lambda} + \boldsymbol{y}^\top(\boldsymbol{d} - B^\top \boldsymbol{\lambda}) + \mathrm{tr}(Y(\boldsymbol{\zeta}\boldsymbol{d}^\top - \Delta B)\\ & + \sum_{i=1}^{\bar{m}} s_i(\sup_{\boldsymbol{z}} \Delta_{:i}^\top \boldsymbol{z} - \lambda_i f_*(\boldsymbol{z}))\\ = \sup_{\boldsymbol{z}^1,\ldots,\boldsymbol{z}^m}\ & \boldsymbol{c}^\top \boldsymbol{x} + \mathrm{tr}(\Psi(\boldsymbol{x})\Delta) - (Ax)^\top \boldsymbol{\lambda} + \boldsymbol{y}^\top(\boldsymbol{d} - B^\top \boldsymbol{\lambda}) + \mathrm{tr}(Y(\boldsymbol{\zeta}\boldsymbol{d}^\top - \Delta B)\\ & + \sum_{i=1}^{\bar{m}} s_i \Delta_{:i}^\top \boldsymbol{z}^i - \lambda_i s_i f_*(\boldsymbol{z}^i))\,, \end{aligned}$$

where $\boldsymbol{y} \in \mathbb{R}^{n_y}$, $Y \in \mathbb{R}^{n_y \times n_\zeta}$, and $\boldsymbol{s} \in \mathbb{R}^{\bar{m}}$ are respectively the dual variables associated to constraints (47b), (47e), and (47f). Now letting $\mathcal{L}(\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta,\boldsymbol{y},Y,\boldsymbol{s},\{\boldsymbol{z}^i\}_{i=1}^{\bar{m}})$ denote the expression on the right of the $\sup_{\boldsymbol{z}^1,\ldots,\boldsymbol{z}^m}$ operator, we necessarily have that

$$\begin{aligned} g_{GLRC}(\boldsymbol{x}) &= \inf_{\substack{\boldsymbol{\zeta}\,:\, f(\boldsymbol{\zeta}) \leq 0\\ \Delta,\,\boldsymbol{\lambda}\,:\,\boldsymbol{\lambda} \geq 0}} \sup_{\boldsymbol{y},Y,\boldsymbol{s}\geq 0,\{\boldsymbol{z}^i\}_{i=1}^{\bar{m}}} \mathcal{L}(\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta,\boldsymbol{y},Y,\boldsymbol{s},\{\boldsymbol{z}^i\}_{i=1}^{\bar{m}}) && (48a)\\ &\geq \inf_{\boldsymbol{\zeta}:f(\boldsymbol{\zeta})\leq 0} \sup_{\boldsymbol{y},Y,\boldsymbol{s}\geq 0,\{\boldsymbol{z}^i\}_{i=1}^{\bar{m}}} \inf_{\Delta,\boldsymbol{\lambda}\geq 0} \mathcal{L}(\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta,\boldsymbol{y},Y,\boldsymbol{s},\{\boldsymbol{z}^i\}_{i=1}^{\bar{m}})\,. && (48b) \end{aligned}$$

One can then analytically resolve the optimum in terms of $\boldsymbol{\lambda}$ and Δ as

$$\begin{aligned} g_{GLRC}(\boldsymbol{x}) \geq \min_{\boldsymbol{\zeta}:f(\boldsymbol{\zeta})\leq 0} \max_{\boldsymbol{y},Y,\boldsymbol{s}\geq 0,\{\boldsymbol{z}^i\}_{i=1}^{\bar{m}}} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top(\boldsymbol{y} + Y\boldsymbol{\zeta})\\ \text{subject to} \quad & (\Psi(\boldsymbol{x}))_{i:}^\top - Y^\top B_{i:}^\top + s_i \boldsymbol{z}^i = 0\,,\ \forall\, i\\ & A_{i:}\boldsymbol{x} + B_{i:}\boldsymbol{y} + s_i f_*(\boldsymbol{z}^i) \leq 0\,,\ \forall\, i\,. \end{aligned}$$

The equality constraint can further be used in conjunction with the fact that $s_i f_*(\boldsymbol{z}) = h_*(s_i\boldsymbol{z}, s_i) := \sup_{\boldsymbol{y}} s_i \boldsymbol{z}^\top \boldsymbol{y} - s_i f(\boldsymbol{y})$, to obtain

$$\begin{aligned} g_{GLRC}(\boldsymbol{x}) \geq \min_{\boldsymbol{\zeta}:f(\boldsymbol{\zeta})\leq 0} \max_{\boldsymbol{y},Y,\boldsymbol{s}\geq 0} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top(\boldsymbol{y} + Y\boldsymbol{\zeta})\\ \text{subject to} \quad & A_{i:}\boldsymbol{x} + B_{i:}\boldsymbol{y} + h_*(Y^\top B_{i:}^\top - (\Psi(\boldsymbol{x}))_{i:}^\top, s_i) \leq 0\,,\ \forall\, i\,. \end{aligned}$$

After applying Sion's minimax theorem as was done in the proof of Proposition 2, one obtains

$$\begin{aligned} g_{GLRC}(\boldsymbol{x}) \geq \max_{\boldsymbol{y},Y,\boldsymbol{s}\geq 0} \quad & \min_{\boldsymbol{\zeta}:f(\boldsymbol{\zeta})\leq 0} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top(\boldsymbol{y} + Y\boldsymbol{\zeta})\\ \text{subject to} \quad & A_{i:}\boldsymbol{x} + B_{i:}\boldsymbol{y} + h_*(Y^\top B_{i:}^\top - (\Psi(\boldsymbol{x}))_{i:}^\top, s_i) \leq 0\,,\ \forall\, i\,, \end{aligned}$$

where the last constraint can be reformulated as

$$A_{i:}\boldsymbol{x} + B_{i:}\boldsymbol{y} + \inf_{s_i \geq 0} \sup_{\boldsymbol{\zeta}} B_{i:}Y\boldsymbol{\zeta} - \Psi(\boldsymbol{x})_{i:}\boldsymbol{\zeta} - s_i f(\boldsymbol{\zeta}) \leq 0$$

since s_i is not involved in the objective function. Given that there exists a point $\bar{\boldsymbol{\zeta}}$ such that $f(\bar{\boldsymbol{\zeta}}) < 0$, strong duality theory will apply here and allow one to reformulate this constraint as

$$A_{i:}\boldsymbol{x} + B_{i:}\boldsymbol{y} + \sup_{\boldsymbol{\zeta}} \inf_{s_i \geq 0} B_{i:}Y\boldsymbol{\zeta} - \Psi(\boldsymbol{x})_{i:}\boldsymbol{\zeta} - s_i f(\boldsymbol{\zeta}) \leq 0\,,$$

and finally

$$A_{i:}\boldsymbol{x} + B_{i:}\boldsymbol{y} + \sup_{\boldsymbol{\zeta}:f(\boldsymbol{\zeta})\leq 0} B_{i:}Y\boldsymbol{\zeta} - \Psi(\boldsymbol{x})_{i:}\boldsymbol{\zeta} \leq 0\,.$$

These steps allow us to reach our conclusion:

$$\begin{aligned} g_{GLRC}(\boldsymbol{x}) \geq \max_{\boldsymbol{y},Y} \quad & \min_{\boldsymbol{\zeta}:f(\boldsymbol{\zeta})\leq 0} \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top(\boldsymbol{y} + Y\boldsymbol{\zeta}) \\ \text{subject to} \quad & Ax + B(\boldsymbol{y} + Y\boldsymbol{\zeta}) \leq \Psi(\boldsymbol{x})\boldsymbol{\zeta} \leq 0\,,\ \forall \boldsymbol{\zeta} : f(\boldsymbol{\zeta}) \leq 0\,, \end{aligned}$$

which is the conservative approximation of the worst-case performance for $\boldsymbol{x}$ when employing affine decision rules. Note that equality is met in this expression if one is able to establish the right constraint qualification conditions for a minimax theorem to apply in (48). In this case, LRC becomes equivalent to AARC. □

Appendix D Improved tractable approximations for surgery block allocation problems

Consider the following surgery block allocation problem proposed in Denton et al. (2010):

$$\begin{aligned} \underset{\boldsymbol{x},Z,\boldsymbol{y}(\boldsymbol{\zeta})}{\text{minimize}} \quad & \max_{\boldsymbol{\zeta}\in\mathcal{U}} \ c\sum_{i\in\mathcal{I}} x_i + d\sum_{i\in\mathcal{I}} y_i(\boldsymbol{\zeta}) && \text{(49a)} \\ \text{subject to} \quad & y_i(\boldsymbol{\zeta}) \geq \sum_{j\in\mathcal{J}} \zeta_j Z_{ij} - w x_i\,,\ \forall\, i\in\mathcal{I}\,,\ \forall\,\boldsymbol{\zeta}\in\mathcal{U} && \text{(49b)} \\ & \boldsymbol{y}(\boldsymbol{\zeta}) \geq 0\,,\ \forall\,\boldsymbol{\zeta}\in\mathcal{U} && \text{(49c)} \\ & \sum_{i\in\mathcal{I}} Z_{ij} = 1\,,\ \forall\, j\in\mathcal{J} && \text{(49d)} \\ & Z_{ij} \leq x_i\,,\ \forall\, i\in\mathcal{I}\,,\ \forall\, j\in\mathcal{J} && \text{(49e)} \\ & \boldsymbol{x}\in\{0,1\}^m\,,\ Z\in\{0,1\}^{m\times n}\,, && \text{(49f)} \end{aligned}$$

where for each $i \in \mathcal{I} := \{1, 2, \ldots, m\}$, variable x_i denotes whether we will open Operating Room (OR) i or not, while, for each $j \in \mathcal{J} := \{1, 2, \ldots, n\}$, the variable $Z_{ij} \in \{0, 1\}$ decides whether surgery block j will be allocated to OR i. Each ζ_j captures the duration of surgery block j, which is a priori not known exactly. As the surgeries are performed, if the total amount of time needed in OR i exceeds the planned session length w, then one has to schedule some overtime y_i. The cost model includes a fixed cost c for opening any OR and a variable overtime cost d. Note that constraint (49d) captures the fact that a surgery block needs to be assigned to exactly one OR, while constraint (49e) captures the fact that surgery blocks can be assigned to an OR only if it is opened. In what follows we demonstrate that the RORA reformulation proposed in Denton et al. (2010) does not always provide an exact solution to problem (49).

In particular, we consider a particular problem instance in which there are three surgery blocks and 2 operating rooms that can run for 8 hours. The cost of opening a room is $39,000, and the overtime cost is $100 per minute. The duration of each of the three surgery blocks is planned to be equal to 0 min, 240 min, and 320 min, but could last up to 160 min, 352 min, and 512 min respectively. We set the budget to $\Gamma = 2$.

In this context, one can show that the model proposed by Denton will suggest opening only one OR, where all blocks will be scheduled for an estimated worst-case total cost of \$822,000. On the other hand, one can verify that opening both ORs and scheduling the biggest block in one OR and the two smaller ones in the second OR leads to a worst-case total cost of \$812,000. Note that the worst-case total cost of this solution is estimated at \$828,000 by the Denton model. One can further confirm that the exact optimal solution is the one that is returned by the AARC (and LRC) model. Table 13 summarizes the optimal bounds on worst-case cost obtained for the two types of solutions (i.e., open one or two ORs) using RORA, AARC, and an exact approach.

Table 13: Comparison of the worst-case cost for different solution methods to the surgery block allocation problem.

Alternative	RORA†	AARC	Exact
Open one OR	\$822,00	\$822,000	\$822,0000
Open two ORs	\$828,000	\$812,000	\$812,000

†RORA refers to the "exact" reformulation proposed in Denton et al. (2010).

The issue with the argument presented by the authors of Denton et al. (2010) to support the exactness of their reformulation is found in their Proposition 6, which states that a certain polyhedron only has integer extreme points.

Proposition 8 (Proposition 6 from Denton et al. (2010)) *The polyhedron defined by the following constraints has integer extreme points when τ is an integer*

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\Delta_{ij} \leq \tau \tag{50a}$$

$$0 \leq \Delta_{ij} \leq Y_{ij} z_j \,, \forall\, i \in \mathcal{I} \,, \forall\, j \in \mathcal{J} \tag{50b}$$

$$0 \leq z_j \leq 1 \,, \forall\, j \in \mathcal{J} \,, \tag{50c}$$

where $\Delta \in \mathbb{R}^{n\times m}$ and $\boldsymbol{z} \in \mathbb{R}^m$ and where $Y \in \{0,1\}^{n\times m}$ satisfies the property that $\sum_{j\in\mathcal{J}} Y_{ij} = 1$ for all $i \in \mathcal{I}$.

In fact, one can claim the following counter-proposition.

Proposition 9 *Let $n = 3$, $m = 2$, $Y_{12} = Y_{21} = Y_{31} = 1$, and $\Gamma = 2$, then the polyhedron defined by Equations (50a), (50b), and (50c) has the following extreme point:*

$$\bar{\Delta} = \begin{bmatrix} 0 & 1 \\ 0.5 & 0 \\ 0.5 & 0 \end{bmatrix} \quad \bar{\boldsymbol{z}} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} .$$

Proof. This can easily be shown by verifying that this solution is feasible and that it satisfies exactly a set of 8 linearly independent constraints. The eight constraints are

$$\sum_{ij}\Delta_{ij} \leq 2 \qquad -\Delta_{11} \leq 0 \qquad \Delta_{12} \leq z_2 \qquad \Delta_{21} \leq z_1$$

$$-\Delta_{22} \leq 0 \qquad \Delta_{31} \leq z_1 \qquad -\Delta_{32} \leq 0 \qquad z_2 \leq 1 \,.$$

Putting all these together we get

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta_{11} \\ \Delta_{12} \\ \Delta_{21} \\ \Delta_{22} \\ \Delta_{31} \\ \Delta_{32} \\ z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Since this matrix is invertible and the pair $(\bar{\Delta}, \bar{\boldsymbol{z}})$ satisfies this system of equations, we have confirmed that this assignment describes an extreme point of the polyhedron. □

Given that the reformulation proposed in Denton et al. (2010) is inexact, it is worth investigating how the conservative approximation obtained from applying AARC to problem (49) compares to the Denton et al.'s reformulation. We do so through the following proposition which confirms that AARC provides tighter approximations than the reformulation of Denton et al. (2010).

Proposition 10 *When $\mathcal{U} := \{\boldsymbol{\zeta} \in \mathbb{R}^n \,|\, \exists\, \boldsymbol{\delta} \in [0,1]^n, \boldsymbol{\zeta} = \bar{\boldsymbol{\zeta}} + diag(\hat{\boldsymbol{\zeta}})\boldsymbol{\delta},\ \sum_{j\in\mathcal{J}} \delta_j \leq \Gamma\}$, employing affine decision rules in the surgery block allocation problem provides a conservative approximation that is at least as tight and in some cases strictly tighter as the reformulation proposed in Denton et al. (2010) (see model (40) in that paper).*

Proof. Indeed, the model presented in Denton et al. (2010) can be rewritten as

$$\begin{aligned} \underset{\boldsymbol{x},Z}{\text{minimize}} \quad & g_{Denton}(\boldsymbol{x}, Z) \\ \text{subject to} \quad & \sum_{i\in\mathcal{I}} Z_{ij} = 1\,,\ \forall\, j \in \mathcal{J} \\ & Z_{ij} \leq x_i\,,\ \forall\, i \in \mathcal{I} \\ & \boldsymbol{x} \in \{0,1\}^m\ ,\ Z \in \{0,1\}^{m\times n}\,, \end{aligned}$$

where

$$\begin{aligned} g_{Denton}(\boldsymbol{x}, Z) := \min_{\alpha,\boldsymbol{\gamma},\kappa} \quad & \sum_{i\in\mathcal{I}} c x_i + \sum_{i\in\mathcal{I}} \gamma_i + \Gamma\alpha \\ \text{subject to} \quad & \alpha \geq d\hat{\zeta}_j Z_{ij} - \kappa_{ij}\,,\ \forall\, i \in \mathcal{I}\,,\ \forall\, j \in \mathcal{J} \\ & \gamma_i \geq \sum_{j\in\mathcal{J}} \kappa_{ij} - d(w x_i - \sum_{j\in\mathcal{J}} \bar{\zeta}_j Z_{ij})\,,\ \forall\, i \in \mathcal{I} \\ & \alpha \geq 0\,,\ \boldsymbol{\gamma} \geq 0\,,\ \kappa \geq 0\,, \end{aligned}$$

where $\alpha \in \mathbb{R}$, $\boldsymbol{\gamma} \in \mathbb{R}^m$, and $\kappa \in \mathbb{R}^{m\times n}$. By duality, we can also represent this function as

$$\begin{aligned} g_{Denton}(\boldsymbol{x}, Z) := \max_{\boldsymbol{\lambda},\Delta} \quad & \sum_{i\in\mathcal{I}} c x_i + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} d\hat{\zeta}_j Z_{ij}\Delta_{ij} - \sum_{i\in\mathcal{I}} d(w x_i - \sum_{j\in\mathcal{J}} \bar{\zeta}_j Z_{ij})\lambda_i \\ \text{subject to} \quad & 0 \leq \lambda_i \leq 1\,,\ \forall\, i \in \mathcal{I} \\ & 0 \leq \Delta_{ij} \leq \lambda_i\,,\ \forall\, i \in \mathcal{I}\,,\ \forall\, j \in \mathcal{J} \\ & \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \Delta_{ij} \leq \Gamma\,, \end{aligned}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^m$ and $\Delta \in \mathbb{R}^{m\times n}$.

Based on Proposition 2 and the details presented in the proof of Proposition 9, we now know that employing affine decision rules in problem (49) is equivalent to optimizing

$$\begin{aligned}\underset{\boldsymbol{x},Z}{\text{minimize}} \quad & g_{LRC}(\boldsymbol{x}, Z)\\ \text{subject to} \quad & (49d), (49e), (49f)\,,\end{aligned}$$

where

$$\begin{aligned}g_{LRC}(\boldsymbol{x}, Z) := \max_{\boldsymbol{\lambda},\boldsymbol{\delta},\Delta} \quad & \sum_{i\in\mathcal{I}} cx_i + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} d\hat{\zeta}_j Z_{ij}\Delta_{ij} - \sum_{i\in\mathcal{I}} d(wx_i - \sum_{j\in\mathcal{J}} \bar{\zeta}_j Z_{ij})\lambda_i\\ \text{subject to} \quad & 0 \le \lambda_i \le 1\,,\ \forall\, i\in\mathcal{I}\\ & 0 \le \delta_j \le 1\,,\ \forall\, j\in\mathcal{J}\\ & \sum_{j\in\mathcal{J}} \delta_j \le \Gamma\\ & 0 \le \Delta_{ij} \le \delta_j\,,\ \forall\, i\in\mathcal{I}\,,\ \forall\, j\in\mathcal{J}\\ & \Delta_{ij} \le \lambda_i\,,\ \forall\, i\in\mathcal{I}\,,\ \forall\, j\in\mathcal{J}\\ & \sum_{j\in\mathcal{J}} \Delta_{ij} \le \Gamma z_i\,,\ \forall\, i\in\mathcal{I}\\ & 1 - \delta_j - \lambda_i + \Delta_{ij} \ge 0\,,\ \forall\, i\in\mathcal{I}\,,\ \forall\, j\in\mathcal{J}\\ & \sum_{j\in\mathcal{J}} \delta_j - \sum_{j\in\mathcal{J}} \Delta_{ij} \le \Gamma(1-\lambda_i)\,,\ \forall\, i\in\mathcal{I}\,.\end{aligned}$$

We will now exploit the fact that we can add the constraint $\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \Delta_{ij} \le \Gamma$ to the problem associated to $g_{LRC}(\boldsymbol{x}, Z)$ without affecting the optimal value that it will return. This is because, for any optimal solution $(\boldsymbol{\lambda}^*, \boldsymbol{\delta}^*, \Delta^*)$, one can simply replace Δ^* with Δ' such that $\Delta'_{ij} := Z_{ij}\Delta_{ij}$ satisfies all constraints and achieves the same objective. Indeed we have that

$$\begin{aligned}& 0 \le \Delta_{ij} \le \delta_j \Rightarrow 0 \le \Delta_{ij}Z_{ij} \le \delta_j \Rightarrow 0 \le \Delta'_{ij} \le \delta_j\\ & \Delta_{ij} \le z_i \Rightarrow \Delta_{ij}Z_{ij} \le z_i \Rightarrow \Delta'_{ij} \le z_i\\ & \sum_{j\in\mathcal{J}} \Delta_{ij} \le \Gamma z_i \Rightarrow \sum_{j\in\mathcal{J}} \Delta_{ij}Z_{ij} \le \Gamma z_i \Rightarrow \sum_{j\in\mathcal{J}} \Delta'_{ij} \le \Gamma z_i\\ & \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \Delta'_{ij} = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \Delta_{ij}Z_{ij} \le \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \delta_j Z_{ij} \le \sum_{j\in\mathcal{J}} \delta_j \sum_{i\in\mathcal{I}} Z_{ij} = \sum_{j\in\mathcal{J}} \delta_j \le \Gamma\,.\end{aligned}$$

Hence, we have that

$$\begin{aligned}g_{LRC}(\boldsymbol{x}, Z) = \max_{\boldsymbol{\lambda},\boldsymbol{\delta},\Delta} \quad & \sum_{i\in\mathcal{I}} cx_i + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} d\hat{\zeta}_j Z_{ij}\Delta_{ij} - \sum_{i\in\mathcal{I}} d(wx_i - \sum_{j\in\mathcal{J}} \hat{\zeta}_j Z_{ij})z_i\\ \text{subject to} \quad & 0 \le z_i \le 1\,,\ \forall\, i\in\mathcal{I}\\ & 0 \le \delta_j \le 1\,,\ \forall\, j\in\mathcal{J}\\ & \sum_{j\in\mathcal{J}} \delta_j \le \Gamma\\ & 0 \le \Delta_{ij} \le \delta_j\,,\ \forall\, i\in\mathcal{I}\,,\ \forall\, j\in\mathcal{J}\\ & \Delta_{ij} \le z_i\,,\ \forall\, i\,,\ \forall\, j\in\mathcal{J}\\ & \sum_{j\in\mathcal{J}} \Delta_{ij} \le \Gamma z_i\,,\ \forall\, i\in\mathcal{I}\\ & 1 - \delta_j - \lambda_i + \Delta_{ij} \ge 0\,,\ \forall\, i\in\mathcal{I}\,,\ \forall\, j\in\mathcal{J}\\ & \sum_{j\in\mathcal{J}} \delta_j - \sum_{j\in\mathcal{J}} \Delta_{ij} \le \Gamma(1-\lambda_i)\,,\ \forall\, i\in\mathcal{I}\\ & \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \Delta_{ij} \le \Gamma\,,\end{aligned}$$

meaning that, for any feasible $\boldsymbol{x}$ and Z, it must be that $g_{LRC}(\boldsymbol{x}, Z) \leq g_{Denton}(\boldsymbol{x}, Z)$, since the latter involves an optimization model that is exactly the same as the former except that it imposes fewer constraints. We conclude that exploiting affine decision rules must lead to a tighter conservative approximation.

The fact that the use of affine decision rules can provide a strictly tighter approximation is verified in the example that led to the results presented in Table 13. □

Appendix E Extension of the Decomposition Algorithm proposed in Ardestani-Jaafari and Delage (2017)

We extend the decomposition algorithm of Ardestani-Jaafari and Delage (2017) such that it can be applied for all proposed models in this paper: i.e., MLRC, SDP-LRC, SDP-LRC2, GLRC, and GSDP-LRC2. We first define the following sets:

$$\mathcal{S}_{\text{MLRC}} := \left\{(\boldsymbol{\zeta}, \boldsymbol{\lambda}, \Delta) \in \mathbb{R}^{n_\zeta} \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{n_\zeta \times \bar{m}} \mid (10b) - (10f), (22b), (22c)\right\}$$

$$\mathcal{S}_{\text{SDP-LRC}} := \left\{ (\boldsymbol{\zeta}, \boldsymbol{\lambda}, \Delta) \in \mathbb{R}^{n_\zeta} \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{n_\zeta \times \bar{m}} \;\middle|\; \begin{array}{l} \exists(\Xi, \Lambda) \in \mathbb{R}^{n_\zeta \times n_\zeta} \times \mathbb{R}^{n_{\bar{m}} \times n_{\bar{m}}}, \\ (6b) - (6d), (8b) - (8f), (19), (21), \\ (34b), (34c), (34d) \end{array} \right\}$$

$$\mathcal{S}_{\text{SDP-LRC2}} := \left\{ (\boldsymbol{\zeta}, \boldsymbol{\lambda}, \Delta) \in \mathbb{R}^{n_\zeta} \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{n_\zeta \times \bar{m}} \;\middle|\; \begin{array}{l} \exists(\Xi, \Lambda) \in \mathbb{R}^{n_\zeta \times n_\zeta} \times \mathbb{R}^{n_{\bar{m}} \times n_{\bar{m}}}, \\ (6b) - (6d), (8b) - (8f), (19), (21), \\ (34b), (34c), (35b) \end{array} \right\}$$

$$\mathcal{S}_{\text{GLRC}} := \left\{ (\boldsymbol{\zeta}, \boldsymbol{\lambda}, \Delta) \in \mathbb{R}^{n_\zeta} \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{n_\zeta \times \bar{m}} \;\middle|\; \begin{array}{l} \exists(\Xi, \Lambda) \in \mathbb{R}^{n_\zeta \times n_\zeta} \times \mathbb{R}^{n_{\bar{m}} \times n_{\bar{m}}}, \\ (6b) - (6d), (8b), (8c), (19), (21), \\ (38b), (38c), (38d), (38e) \end{array} \right\}$$

$$\mathcal{S}_{\text{GSDP-LRC2}} := \left\{ (\boldsymbol{\zeta}, \boldsymbol{\lambda}, \Delta) \in \mathbb{R}^{n_\zeta} \times \mathbb{R}^{\bar{m}} \times \mathbb{R}^{n_\zeta \times \bar{m}} \;\middle|\; \begin{array}{l} \exists(\Xi, \Lambda) \in \mathbb{R}^{n_\zeta \times n_\zeta} \times \mathbb{R}^{n_{\bar{m}} \times n_{\bar{m}}}, \\ (6b) - (6d), (8b) - (8f), (19), (21), \\ (34b), (34c), (38b) - (38e), (39b) \end{array} \right\}.$$

For any *model* $\in$ {MLRC, SDP-LRC, SDP-LRC2, GLRC, and GSDP-LRC2}, one can propose the following decomposition algorithm.

Algorithm 1 Row generation algorithm

Step #1: Set $UB = \infty$ and $LB = -\infty$. Identify any feasible first-stage solution $\dot{\boldsymbol{x}}^{(1)} \in \mathcal{X}$ (e.g., a solution of the deterministic model using a nominal scenario for $\boldsymbol{\zeta}$). Let $\kappa = 1$.
Step #2: Solve the following subproblem

$$\text{(SP)} \quad \underset{(\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta)\in\mathcal{S}_{model}}{\text{minimize}} \quad \boldsymbol{c}^\top \dot{\boldsymbol{x}}^{(\kappa)} + \text{tr}(\Psi(\dot{\boldsymbol{x}}^{(\kappa)})\Delta) - (A\dot{\boldsymbol{x}}^{(\kappa)})^\top \boldsymbol{\lambda}$$

Note that following Corollary 1, SP is necessarily feasible. One can thus set $\dot{\boldsymbol{\lambda}}^{(\kappa)}$ and $\dot{\Delta}^{(\kappa)}$ as its optimal solution and let ρ^* be its optimal value. Set $LB := \max(LB, \rho^*)$ and $\boldsymbol{x}^* := \dot{\boldsymbol{x}}^{(\kappa)}$ if $\rho^* \geq LB$.
Step #3: Let $\kappa := \kappa + 1$ and solve the following master problem:

$$\text{(MP)} \quad \underset{\boldsymbol{x}\in\mathcal{X},\rho}{\text{maximize}} \quad \rho \tag{51a}$$

$$\text{subject to} \quad \rho \leq \boldsymbol{c}^\top \boldsymbol{x} + \text{tr}(\Psi(\boldsymbol{x})\dot{\Delta}^{(l)}) - (A\boldsymbol{x})^\top \dot{\boldsymbol{\lambda}}^{(l)}, \forall l \in \{1, 2, \ldots, \kappa - 1\}. \tag{51b}$$

Let $\dot{\boldsymbol{x}}^{(\kappa)}$ and UB take on the values of any optimal solution and optimal value, respectively, of the MP.
Step #4: If $UB - LB \leq \varepsilon$ then terminate and return $\boldsymbol{x}^*$ and LB as the optimal solution; otherwise, repeat from Step #2. (Note that the termination condition can also be verified at the end of Step #2.)

Similar to Ardestani-Jaafari and Delage (2017), one can propose a set of valid inequalities to the MP model in the row generation algorithm. In particular, for all models it is possible to replace the MP with an enhanced version that exploits a scenario-based relaxation:

$$\begin{array}{lll}
\text{(MP')} & \underset{\boldsymbol{x}\in\mathcal{X},\rho,\{\boldsymbol{y}^l\}_{l=1}^{|\mathcal{U}^\kappa|}}{\text{maximize}} & \rho \\
& \text{subject to} & (51b) \\
& & \rho \leq \boldsymbol{c}^\top\boldsymbol{x} + \boldsymbol{d}^\top\boldsymbol{y}^l\,,\ \forall l:\boldsymbol{\zeta}^l\in\mathcal{U}^\kappa \\
& & A\boldsymbol{x} + B\boldsymbol{y}^l \leq \Psi(\boldsymbol{x})\boldsymbol{\zeta}\,,\ \forall l:\boldsymbol{\zeta}^l\in\mathcal{U}^\kappa\ ,
\end{array}$$

where $\mathcal{U}^\kappa\subset\mathcal{U}$ is any finite set of scenarios for $\boldsymbol{\zeta}$, e.g. the worst-case scenarios identified when solving the SP in the $\kappa-1$ first iterations. Alternatively, or in addition, one can employ similar cuts as were used in Ardestani-Jaafari and Delage (2017) that exploits the specific structure of the approximation model. For example, in the case of MLRC we get

$$\begin{array}{lll}
\text{(MLRC-MP')} & \underset{\boldsymbol{x}\in\mathcal{X},\rho,\boldsymbol{y},Y,\boldsymbol{z},Z}{\text{maximize}} & \rho \\
& \text{subject to} & (51b) \\
& & \rho \leq \boldsymbol{c}^\top\boldsymbol{x} + \boldsymbol{d}^\top(\boldsymbol{y}+Y\boldsymbol{\zeta}) - \boldsymbol{u}^\top(\boldsymbol{z}+Z\boldsymbol{\zeta})\,,\ \forall\boldsymbol{\zeta}\in\mathcal{U}^\kappa \\
& & A\boldsymbol{x} + B(\boldsymbol{y}+Y\boldsymbol{\zeta}) \leq \Psi(\boldsymbol{x})\boldsymbol{\zeta} + \boldsymbol{z} + Z\boldsymbol{\zeta}\,,\ \forall\boldsymbol{\zeta}\in\mathcal{D}^\kappa \\
& & \boldsymbol{z} + Z\boldsymbol{\zeta} \geq 0\,,\ \forall\boldsymbol{\zeta}\in\mathcal{U}^\kappa\ ,
\end{array}$$

where $\boldsymbol{y}\in\mathbb{R}^{n_y}$, $Y\in\mathbb{R}^{n_y\times n_\zeta}$, $\boldsymbol{z}\in\mathbb{R}^{n_{\bar{m}}}$, and $Z\in\mathbb{R}^{n_{\bar{m}}\times n_\zeta}$. In the implementation used in Section 5.2 for the MLRC model, we employed $\mathcal{U}^\kappa := \{\dot{\boldsymbol{\zeta}}^{(\kappa-1)}\}$ as the worst-case scenario identified when solving the SP in the previous iteration. This effectively makes both MP' and MLRC-MP' equivalent approaches. We refer the reader to Rahmaniani et al. (2017) for a survey of methods that can be used to improve solution time of Benders decomposition schemes.

Finally, we note that in the context of MLRC, this row generation algorithm is guaranteed to converge as long as one makes sure that the SP always return an optimal vertex of $\mathcal{S}_{\text{MLRC}}$. This is due to the fact that 1) if the same vertex is returned twice, then necessarily $UB = LB$; and 2) the polyhedra $\mathcal{S}_{\text{MLRC}}$ only has a finite number of vertices.

Appendix F SDP-LRC strictly improves on AARC in small multi-item example

Let us consider an example of problem (43) with $n=3$, $r=[\ 80\ \ 80\ \ 80\ ]$, $\boldsymbol{c}=[\ 70\ \ 50\ \ 20\ ]$, $\boldsymbol{s}=[\ 20\ \ 15\ \ 10\ ]$, and $p=[\ 60\ \ 60\ \ 50\ ]$. Demand vector $\boldsymbol{\zeta}$ is defined in the following uncertainty set $\mathcal{U}$:

$$\tilde{\mathcal{U}}_P(\Gamma) := \left\{\boldsymbol{\zeta}\,\middle|\,\exists\,(\boldsymbol{\delta}^+,\boldsymbol{\delta}^-)\in\mathcal{U}_\delta(\Gamma),\ \boldsymbol{\zeta}=\bar{\boldsymbol{\zeta}}+\text{diag}(\hat{\boldsymbol{\zeta}})P(\boldsymbol{\delta}^+-\boldsymbol{\delta}^-)\right\}\,,$$

where

$$P := \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}\,,$$

which models the property that the correlation between the demand of any two items is 0.5. We compare, as it is shown in Table 14, the optimal bound on worst-case profit and the achieved worst-case profit of solutions obtained from the LRC model, the SDP-LRC model, and the semi-definite programming model (denoted by SDP-A&D) proposed in Ardestani-Jaafari and Delage (2016). In this example, LRC is not exact and can actually be improved upon using models such that SDP-A&D and SDP-LRC. In particular, the bound on best achievable worst-case profit is increased by a factor of about 3 and 10 using SDP-A&D and SDP-LRC respectively. This translates directly in some improvement in performance of solutions of SDP-A&D and

SDP-LRC which achieve a worst-case profit that are respectively near 4 and 16 times better than what is achieved by the solution of AARC. It is also clear that the SDP-LRC model is responsible for most of the improvement.

Table 14: Comparison of optimal bound and worst-case profit associated to solutions obtained from conservative approximation models and exact models in a newsvendor problem.

	AARC	SDP-A&D†	SDP-LRC	Exact model
Optimal bound on worst-case profit	41.83	113.01	411.08	825.83
Worst-case profit of solution	41.83	150.94	664.76	825.83

† SDP-A&D refers to the semi-definite programming model proposed in Ardestani-Jaafari and Delage (2016)

Appendix G Implications for copositive programming reformulations

In this section, we reuse the ideas of both Hanasusanto and Kuhn (2016) and Xu and Burer (2016) to strengthen the connections between the ARO model with relatively complete recourse, copositive programming, and both MLRC and SDP-LRC derived in Section 3. In order to improve readability, some of the proofs of the theorems that are presented are delayed to Appendix H.

We start with an essential assumption that can be made without loss of generality in order to apply the theory related to copositive programming.

Assumption 6 *The uncertain vector $\boldsymbol{\zeta}$ is known to lie in the non-negative orthant, i.e. $\mathcal{U} \subset \mathbb{R}_+^{n_\zeta}$. This assumption is made without loss of generality since one can always redefine $\boldsymbol{\zeta} := \boldsymbol{\zeta}^+ - \boldsymbol{\zeta}^-$ with $\boldsymbol{\zeta}^+ \geq 0$ and $\boldsymbol{\zeta}^- \geq 0$.*

We next repeat an important result of Section 3 which stated that under assumptions 1–4, the ARO model is equivalent to maximize$_{\boldsymbol{x}\in\mathcal{X}}$ $g(\boldsymbol{x})$ where $g(\boldsymbol{x})$ is evaluated using

$$g(\boldsymbol{x}) = \min_{\boldsymbol{\zeta},\boldsymbol{\lambda}} \quad \boldsymbol{c}^\top \boldsymbol{x} + (\Psi(\boldsymbol{x})\boldsymbol{\zeta})^\top \boldsymbol{\lambda} - (A\boldsymbol{x})^\top \boldsymbol{\lambda} \tag{52a}$$

$$\text{subject to} \quad B^\top \boldsymbol{\lambda} = \boldsymbol{d} \tag{52b}$$

$$P\boldsymbol{\zeta} \leq \boldsymbol{q} \tag{52c}$$

$$\boldsymbol{0} \leq \boldsymbol{\lambda} \leq \boldsymbol{u}\,. \tag{52d}$$

In particular, it can be reformulated in a form that is more standard for non-convex quadratic programs:

$$\begin{aligned} g(\boldsymbol{x}) = \min_{\boldsymbol{y}} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{y}^\top \tilde{Q}(\boldsymbol{x}) \boldsymbol{y} + 2\tilde{\boldsymbol{c}}(\boldsymbol{x})^\top \boldsymbol{y} \\ \text{subject to} \quad & \tilde{A}\boldsymbol{y} = \tilde{\boldsymbol{b}} \\ & \boldsymbol{y} \geq 0\,, \end{aligned}$$

where $\boldsymbol{y} \in \mathbb{R}^{\tilde{n}}$, with $\tilde{n} := 2m + n_\zeta + n_\mathcal{U}$ so that $\boldsymbol{y}$ captures $[\boldsymbol{\lambda}^\top \ \boldsymbol{\zeta}^\top \ (\boldsymbol{q} - P\boldsymbol{\zeta})^\top \ (\boldsymbol{u} - \boldsymbol{\lambda})^\top]$ and where

$$\tilde{Q}(\boldsymbol{x}) := \begin{bmatrix} 0 & (1/2)\Psi(\boldsymbol{x}) & 0 & 0 \\ (1/2)\Psi(\boldsymbol{x})^\top & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \tilde{\boldsymbol{c}}(\boldsymbol{x}) := \begin{bmatrix} -(1/2)A\boldsymbol{x} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\tilde{A} := \begin{bmatrix} B^\top & 0 & 0 & 0 \\ 0 & P & \boldsymbol{I} & 0 \\ \boldsymbol{I} & 0 & 0 & \boldsymbol{I} \end{bmatrix} \qquad \tilde{\boldsymbol{b}} := \begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{q} \\ \boldsymbol{u} \end{bmatrix},$$

Based on corollaries 8.1 and 8.3 of Burer (2012), one can directly establish the following completely positive reformulation for $g(\boldsymbol{x})$:

$$g(\boldsymbol{x}) = \min_{Y,\boldsymbol{y}} \quad \boldsymbol{c}^\top \boldsymbol{x} + \mathrm{tr}(\tilde{Q}(\boldsymbol{x})^\top Y) + 2\tilde{\boldsymbol{c}}(\boldsymbol{x})^\top \boldsymbol{y} \tag{53a}$$

$$\text{subject to} \quad \tilde{A}\boldsymbol{y} = \tilde{\boldsymbol{b}} \tag{53b}$$

$$\tilde{A}Y = \tilde{\boldsymbol{b}}\boldsymbol{y}^\top \tag{53c}$$

$$\begin{bmatrix} Y & \boldsymbol{y} \\ \boldsymbol{y}^\top & 1 \end{bmatrix} \in \mathcal{K}_{\text{CP}}\,, \tag{53d}$$

where $Y \in \mathbb{R}^{\tilde{n}\times\tilde{n}}$ and where $\mathcal{K}_{\text{CP}}$ is the cone of completely positive matrices, i.e.

$$\mathcal{K}_{\text{CP}} := \left\{ \mathbb{M} \in \mathbb{R}^{\tilde{n}+1\times\tilde{n}+1} \,\middle|\, \mathbb{M} = \sum_{k\in K} \boldsymbol{z}^k {\boldsymbol{z}^k}^\top \text{ for some finite } \{\boldsymbol{z}^k\}_{k\in K} \subset \mathbb{R}_+^{\tilde{n}+1+1} \setminus 0 \right\} \cup \{0\}\,.$$

Given that completely positive programs are convex optimization model, one can hope to obtain a tight bound using conic duality so that

$$g(\boldsymbol{x}) \geq \max_{\boldsymbol{W},\boldsymbol{w},t} \quad \boldsymbol{c}^\top \boldsymbol{x} + \tilde{\boldsymbol{b}}^\top \boldsymbol{w} - t \tag{54a}$$

$$\text{subject to} \quad \begin{bmatrix} \tilde{Q}(\boldsymbol{x}) - (1/2)(\boldsymbol{W}^\top \tilde{A} + \tilde{A}^\top \boldsymbol{W}) & \tilde{\boldsymbol{c}}(\boldsymbol{x}) - (1/2)(\tilde{A}^\top \boldsymbol{w} - \boldsymbol{W}^\top \tilde{\boldsymbol{b}}) \\ \tilde{\boldsymbol{c}}(\boldsymbol{x})^\top - (1/2)(\tilde{A}^\top \boldsymbol{w} - \boldsymbol{W}^\top \tilde{\boldsymbol{b}})^\top & t \end{bmatrix} \in \mathcal{K}_{\text{Cop}}\,, \tag{54b}$$

where $t \in \mathbb{R}$, and where $\boldsymbol{w} \in \mathbb{R}^{n_y+n_{\mathcal{U}}+\bar{m}}$ and $\boldsymbol{W} \in \mathbb{R}^{n_y+n_{\mathcal{U}}+\bar{m}\times\tilde{n}}$ contain the dual variables associated to constraints (53b) and (53c) respectively, while $\mathcal{K}_{\text{Cop}}$ refers to the dual cone of $\mathcal{K}_{\text{CP}}$ also known as the cone of copositive matrices, i.e.

$$\mathcal{K}_{\text{Cop}} := \left\{ \mathbb{M} \in \mathbb{R}^{\tilde{n}+1\times\tilde{n}+1} \,\middle|\, \mathbb{M} = \mathbb{M}^\top,\ \boldsymbol{z}^\top \mathbb{M} \boldsymbol{z} \geq 0\,, \forall\, \boldsymbol{z} \in \mathbb{R}_+^{\tilde{n}} \right\}\,.$$

When attempting to prove that strong duality holds, a sufficient step consists in verifying whether problem (54) is strictly feasible.

Lemma 2 *Given Assumptions 4 and 6, problem (54) is strictly feasible. In particular it is even strictly feasible when $\mathcal{K}_{Cop}$ is replaced with $\mathbb{R}_+^{\tilde{n}+1\times\tilde{n}+1} \subset \mathcal{K}_{Cop}$.*

At this point, we have assembled all the ingredient to present yet a second equivalent copositive programming formulation of ARO for relatively complete recourse problems (see $\overline{\text{RLP}}$ model in Xu and Burer (2016) for the original equivalent model).

Corollary 3 *Given assumptions 1–4 and 6, the following copositive program is equivalent to problem (1):*

$$\underset{\boldsymbol{x}\in\mathcal{X},\boldsymbol{W},\boldsymbol{w},t}{\text{maximize}} \quad \boldsymbol{c}^\top \boldsymbol{x} + \tilde{\boldsymbol{b}}^\top \boldsymbol{w} - t \tag{55a}$$

$$\text{subject to} \quad \begin{bmatrix} \tilde{Q}(\boldsymbol{x}) - (1/2)(\boldsymbol{W}^\top \tilde{A} + \tilde{A}^\top \boldsymbol{W}) & \tilde{\boldsymbol{c}}(\boldsymbol{x}) - (1/2)(\tilde{A}^\top \boldsymbol{w} - \boldsymbol{W}^\top \tilde{\boldsymbol{b}}) \\ \tilde{\boldsymbol{c}}(\boldsymbol{x})^\top - (1/2)(\tilde{A}^\top \boldsymbol{w} - \boldsymbol{W}^\top \tilde{\boldsymbol{b}})^\top & t \end{bmatrix} \in \mathcal{K}_{Cop}\,. \tag{55b}$$

Proof. Strong duality follows from the fact that the dual problem (54) is strictly feasible, following Lemma 2, and bounded, which follows easily from Assumption 3 since it states that $\max_{\boldsymbol{x}\in\mathcal{X}} g(\boldsymbol{x})$ is bounded. □

Corollary 4 *Given assumptions 4 and 6, and some cone $\mathcal{K} \subseteq \mathcal{K}_{Cop}$, the conic program obtained by replacing $\mathcal{K}_{Cop}$ by $\mathcal{K}$ in problem (55) provides a conservative approximation to problem (1). Furthermore,*

1. *if $\mathcal{K} = \mathcal{K}_1 := \mathbb{R}_+^{\tilde{n}+1\times\tilde{n}+1}$, then the conic program reduces to a linear program that is equivalent to MLRC (18);*
2. *if $\mathcal{K} = \mathcal{K}_2 := \mathbb{R}_+^{\tilde{n}+1\times\tilde{n}+1} + \mathcal{K}_{PSD}^{\tilde{n}+1\times\tilde{n}+1}$, i.e. the Minkowski sum of the non-negative orthant and the cone $\mathcal{K}_{PSD}^{\tilde{n}+1\times\tilde{n}+1}$ of positive semi-definite matrices, then the conic program reduces to a semi-definite program that is equivalent to $\max_{\boldsymbol{x}\in\mathcal{X}} g_{\text{SDP-LRC}}(\boldsymbol{x})$.*
3. *if $\mathcal{K}_1 \subseteq \mathcal{K} \subseteq \mathcal{K}_{Cop}$ then the conic program provides a tighter approximation than MLRC*
4. *if $\mathcal{K}_2 \subseteq \mathcal{K} \subseteq \mathcal{K}_{Cop}$ then the conic program provides a tighter approximation than $\max_{\boldsymbol{x}\in\mathcal{X}} g_{\text{SDP-LRC}}(\boldsymbol{x})$.*

Compared to Hanasusanto and Kuhn (2016), the particularity of this non-convex quadratic program (52) is that it has a bounded feasible space which can be exploited to establish strong duality even thought the ARO does not satisfy the complete recourse assumption. Alternatively, Xu and Burer (2016) did propose imposing a bound on $\|\boldsymbol{\lambda}\|_2$ in order to help with duality yet did not attempt to further connect the resulting model to the AARC approach. Furthermore, the reformulation that is obtained using $\|\boldsymbol{\lambda}\|_2$ cannot readily be approximated using linear programming.

One should also be aware that there exists hierarchies of both polyhedral and semi-definite cones that can be used to cover the range $\mathcal{K}_1 \subseteq \mathcal{K} \subseteq \mathcal{K}_{\text{Cop}}$ and $\mathcal{K}_2 \subseteq \mathcal{K} \subseteq \mathcal{K}_{\text{Cop}}$ respectively and produce tighter conservative approximations albeit at a higher computational price. We refer the reader to Parrilo (2000) and Bomze and de Klerk (2002) for some examples.

Appendix H Proofs of Appendix G

Proof. Proof of Lemma 2. We will exploit Farkas lemma to identify a ray $(\gamma\bar{t}, \gamma\bar{\boldsymbol{w}}, \gamma\bar{\boldsymbol{W}})$, parametrized by $\gamma > 0$, such that

$$\begin{bmatrix} -(1/2)((\gamma\bar{\boldsymbol{W}})^\top \tilde{A} + \tilde{A}^\top(\gamma\bar{\boldsymbol{W}})) & -(1/2)(\tilde{A}^\top(\gamma\bar{\boldsymbol{w}}) - (\gamma\bar{\boldsymbol{W}})^\top\tilde{\boldsymbol{b}}) \\ -(1/2)(\tilde{A}^\top(\gamma\bar{\boldsymbol{w}}) - (\gamma\bar{\boldsymbol{W}})^\top\tilde{\boldsymbol{b}})^\top & \gamma\bar{t} \end{bmatrix} = \gamma\bar{\mathbb{M}}\,,$$

for some

$$\bar{\mathbb{M}} := \begin{bmatrix} -(1/2)(\bar{\boldsymbol{W}}^\top \tilde{A} + \tilde{A}^\top\bar{\boldsymbol{W}}) & -(1/2)(\tilde{A}^\top\bar{\boldsymbol{w}} - \bar{\boldsymbol{W}}^\top\tilde{\boldsymbol{b}}) \\ -(1/2)(\tilde{A}^\top\bar{\boldsymbol{w}} - \bar{\boldsymbol{W}}^\top\tilde{\boldsymbol{b}})^\top & \bar{t} \end{bmatrix} \geq 1\,.$$

This will consequently imply that there exists a $\bar{\gamma} > 0$ for which

$$\begin{bmatrix} \tilde{Q}(\boldsymbol{x}) - (1/2)((\gamma\bar{\boldsymbol{W}})^\top \tilde{A} + \tilde{A}^\top(\gamma\bar{\boldsymbol{W}})) & \tilde{\boldsymbol{c}}(\boldsymbol{x})^\top - (1/2)(\tilde{A}^\top(\gamma\bar{\boldsymbol{w}}) - (\gamma\bar{\boldsymbol{W}})^\top\tilde{\boldsymbol{b}}) \\ \tilde{\boldsymbol{c}}(\boldsymbol{x})^\top - (1/2)(\tilde{A}^\top(\gamma\bar{\boldsymbol{w}}) - (\gamma\bar{\boldsymbol{W}})^\top\tilde{\boldsymbol{b}})^\top & \gamma\bar{t} \end{bmatrix} = \begin{bmatrix} \tilde{Q}(\boldsymbol{x}) & \tilde{\boldsymbol{c}}(\boldsymbol{x}) \\ \tilde{\boldsymbol{c}}(\boldsymbol{x})^\top & 0 \end{bmatrix} + \bar{\gamma}\bar{\mathbb{M}} > 0\,.$$

First, based on Assumption 1, the boundedness of $\mathcal{U}$ implies that there is an $M > 0$ such that $\mathbf{1}^\top\boldsymbol{\zeta} \geq M$ is inconsistent with $P\boldsymbol{\zeta} \leq \boldsymbol{q}$. By Farkas lemma, this implies that there necessarily exists some $\boldsymbol{s} \in \mathbb{R}^{n_\mathcal{U}}$ and $s_0 \in \mathbb{R}$ that satisfy the following linear inequalities:

$$\boldsymbol{s} \geq 0 \qquad s_0 \geq 0 \qquad P^\top\boldsymbol{s} \geq s_0 \qquad \boldsymbol{q}^\top\boldsymbol{s} < Ms_0\,.$$

Yet, since Assumptions 1 and 6 state that $\mathcal{U} \subset \mathbb{R}_+^{n_\zeta}$ is non-empty, this implies that $s_0 > 0$. Indeed, if $s_0 = 0$, the existence of a feasible $\hat{\boldsymbol{\zeta}} \geq 0$ leads to a contradiction:

$$0 \leq \boldsymbol{s}^\top P\hat{\boldsymbol{\zeta}} \leq \boldsymbol{s}^\top\boldsymbol{q} < 0 \;\Rightarrow\; 0 < 0\,.$$

We finally conclude from this exercise that there must exist a $\bar{\boldsymbol{s}} := 1 + \alpha\boldsymbol{s}$, with $\alpha \geq 0$ such that $\bar{\boldsymbol{s}} \geq \mathbf{1}$ and $P^\top\bar{\boldsymbol{s}} = P^\top\mathbf{1} + \alpha P^\top\boldsymbol{s} \geq P^\top\mathbf{1} + \alpha s_0 \geq 1$. This occurs in fact when choosing $\alpha := (1/s_0)(1 + \max(\max_{i=1,\ldots,n_\zeta} P_{:i}^\top\mathbf{1}\;,\;0))$.

Second, we demonstrate that

$$\bar{\mathbb{M}} := \begin{bmatrix} -(1/2)(\bar{\boldsymbol{W}}^\top \tilde{A} + \tilde{A}^\top\bar{\boldsymbol{W}}) & -(1/2)(\tilde{A}^\top\bar{\boldsymbol{w}} - \bar{\boldsymbol{W}}^\top\tilde{\boldsymbol{b}}) \\ -(1/2)(\tilde{A}^\top\bar{\boldsymbol{w}} - \bar{\boldsymbol{W}}^\top\tilde{\boldsymbol{b}})^\top & \bar{t} \end{bmatrix} \geq 1$$

for the following assignment:

$$\bar{t} := 1 \qquad \bar{\boldsymbol{w}} := -(2 + \max(\bar{\boldsymbol{s}}^\top\boldsymbol{q} + \mathbf{1}^\top\boldsymbol{u}\;,\;0))\begin{bmatrix} 0 \\ \bar{\boldsymbol{s}} \\ \mathbf{1} \end{bmatrix} \qquad \bar{\boldsymbol{W}} := -\begin{bmatrix} 0 \\ \bar{\boldsymbol{s}} \\ \mathbf{1} \end{bmatrix}\mathbf{1}^\top\,.$$

Studying each term separately we get:

$$-\frac{1}{2}\left(\bar{\boldsymbol{W}}^\top \tilde{A} + \tilde{A}^\top \bar{\boldsymbol{W}}\right) = \frac{1}{2}\left(\begin{bmatrix} P^\top \bar{\boldsymbol{s}} \\ \boldsymbol{1} \\ \bar{\boldsymbol{s}} \\ \boldsymbol{1} \end{bmatrix} \boldsymbol{1}^\top + \boldsymbol{1} \begin{bmatrix} P^\top \bar{\boldsymbol{s}} \boldsymbol{1} \bar{\boldsymbol{s}} \boldsymbol{1} \end{bmatrix}\right) \geq 1$$

$$-\frac{1}{2}\left(\tilde{A}^\top \bar{\boldsymbol{w}} - \bar{\boldsymbol{W}}^\top \tilde{\boldsymbol{b}}\right) = \frac{1}{2}\left((1 + \max(\bar{\boldsymbol{s}}^\top \boldsymbol{q} + \boldsymbol{1}^\top \boldsymbol{u}\ ,\ 0)) \begin{bmatrix} P^\top \bar{\boldsymbol{s}} \\ \boldsymbol{1} \\ \bar{\boldsymbol{s}} \\ \boldsymbol{1} \end{bmatrix} - \boldsymbol{q}^\top \bar{\boldsymbol{s}} - \boldsymbol{1}^\top \boldsymbol{u}\right)$$

$$\geq \frac{1}{2}((2 + \max(\bar{\boldsymbol{s}}^\top \boldsymbol{q} + \boldsymbol{1}^\top \boldsymbol{u}\ ,\ 0)) - \boldsymbol{q}^\top \bar{\boldsymbol{s}} - \boldsymbol{1}^\top \boldsymbol{u}) \geq 1$$

$$\bar{t} \geq 1$$

This completes our proof. □

Proof of Corollary 4. Given that we have established that strong duality applies for problem (54) whether the copositive cone is replaced with $\mathbb{R}_+^{\tilde{n}+1\times\tilde{n}+1}$ or $\mathbb{R}_+^{\tilde{n}+1\times\tilde{n}+1} + \mathcal{K}_{\text{PSD}}^{\tilde{n}+1\times\tilde{n}+1}$, our efforts can focus on comparing $g_{\text{MLRC}}(\boldsymbol{x})$ and $g_{\text{SDP-LRC}}(\boldsymbol{x})$ to the optimal value of problem (53) with $\mathcal{K}_{\text{CP}}$ replaced with the dual cone of $\mathbb{R}_+^{\tilde{n}+1\times\tilde{n}+1}$, which is $\mathcal{K}_1 := \mathbb{R}_+^{\tilde{n}+1\times\tilde{n}+1}$, and the dual cone of $\mathbb{R}_+^{\tilde{n}+1\times\tilde{n}+1} + \mathcal{K}_{\text{PSD}}^{\tilde{n}+1\times\tilde{n}+1}$ which is $\mathcal{K}_2 := \mathbb{R}_+^{\tilde{n}+1\times\tilde{n}+1} \cap \mathcal{K}_{\text{PSD}}^{\tilde{n}+1\times\tilde{n}+1}$ respectively. In what follows, we simply refer to the value of each of these two bounds as $g_1(\boldsymbol{x})$ and $g_2(\boldsymbol{x})$ respectively.

First, whether the cone is replaced with $\mathcal{K}_1$ or $\mathcal{K}_2$, we can exploit the equalities of problem (53) to reformulate $g_1(\boldsymbol{x})$ and $g_2(\boldsymbol{x})$ as

$$g_i(\boldsymbol{x}) = \min_{Y,\boldsymbol{y},\Lambda,\Xi,\boldsymbol{\lambda},\boldsymbol{\zeta}} \quad \boldsymbol{c}^\top \boldsymbol{x} + \text{tr}(\tilde{Q}(\boldsymbol{x})^\top Y) + 2\tilde{\boldsymbol{c}}(\boldsymbol{x})^\top \boldsymbol{y} \tag{56a}$$

$$\text{subject to} \quad (53\text{b}) - (53\text{c})$$

$$\begin{bmatrix} Y & \boldsymbol{y} \\ \boldsymbol{y}^\top & 1 \end{bmatrix} = \Phi \begin{bmatrix} \Lambda & \Delta^\top & \boldsymbol{\lambda} \\ \Delta & \Xi & \boldsymbol{\zeta} \\ \boldsymbol{\lambda}^\top & \boldsymbol{\zeta}^\top & 1 \end{bmatrix} \Phi^\top \tag{56b}$$

$$\begin{bmatrix} Y & \boldsymbol{y} \\ \boldsymbol{y}^\top & 1 \end{bmatrix} \in \mathcal{K}_i\,, \tag{56c}$$

with $i = 1, 2$ and where $\Phi \in \mathbb{R}^{\tilde{n}+1\times n_\lambda+n_z eta+1}$ is the matrix defined as

$$\Phi := \begin{bmatrix} \boldsymbol{I} & 0 & 0 \\ 0 & \boldsymbol{I} & 0 \\ 0 & -P & \boldsymbol{q} \\ -\boldsymbol{I} & 0 & \boldsymbol{u} \\ 0 & 0 & 1 \end{bmatrix}.$$

We omit to provide the details of this equivalence as they are purely algebraic. One can indeed expand the list of linear equalities expression in constraint (56b) to see that each one of them simply repeats an equality constraint present in constraints (53b) and (53c).

Next, in problem (22), one can certainly consider additional decision variables $\Lambda \in \mathbb{R}^{\bar{m}\times\bar{m}}$ and $\Xi \in \mathbb{R}^{n_\zeta\times n_\zeta}$ that need to satisfy constraints (8d), (8f), (34b), and (34c), without affecting the feasible set in terms of $(\boldsymbol{\zeta}, \boldsymbol{\lambda}, \Delta)$ given that the assignment $\Lambda := \boldsymbol{\lambda}\boldsymbol{\lambda}^\top$ and $\Xi = \boldsymbol{\zeta}\boldsymbol{\zeta}^\top$ always satisfies these constraints. Furthermore, in both problems (22) and (34) one can exploit Assumption 6 to identify a list of redundant constraints. Namely, based on Farkas lemma, this assumption implies that there must exist some matrix $Q \in \mathbb{R}^{n_\mathcal{U}\times n_\mathcal{U}}$ such that:

$$Q \geq 0\ ,\ QP = -\boldsymbol{I}\ ,\ Q\boldsymbol{q} \leq 0\,.$$

We can therefore derive the following implications:

$$\begin{aligned}
(10c) &\Rightarrow P\Delta \leq \boldsymbol{q}\boldsymbol{\lambda}^\top \Rightarrow QP\Delta \leq Q\boldsymbol{q}\boldsymbol{\lambda}^\top \Rightarrow \Delta \geq 0 \\
(22c) &\Rightarrow P\Delta \geq \boldsymbol{q}\boldsymbol{\lambda}^\top - (\boldsymbol{q} - P\boldsymbol{\zeta})\boldsymbol{u}^\top \Rightarrow QP\Delta \geq Q\boldsymbol{q}\boldsymbol{\lambda}^\top - Q(\boldsymbol{q} - P\boldsymbol{\zeta})\boldsymbol{u}^\top \Rightarrow \Delta \leq \boldsymbol{\zeta}\boldsymbol{u}^\top \\
(8f) &\Rightarrow P\Xi P^\top + \boldsymbol{q}\boldsymbol{q}^\top \geq P\boldsymbol{\zeta}\boldsymbol{q}^\top + \boldsymbol{q}\boldsymbol{\zeta}^\top P^\top \Rightarrow QP\Xi P^\top \geq Q\boldsymbol{q}(P\boldsymbol{\zeta} - \boldsymbol{q})^\top + QP\boldsymbol{\zeta}\boldsymbol{q}^\top \Rightarrow \Xi P^\top \leq \boldsymbol{\zeta}\boldsymbol{q}^\top \\
&\Rightarrow \Xi P^\top \leq \boldsymbol{\zeta}\boldsymbol{q}^\top \Rightarrow Q\Xi P^\top \leq Q\boldsymbol{\zeta}\boldsymbol{q}^\top \Rightarrow \Xi \geq 0\,.
\end{aligned}$$

Hence, one can establish that

$$\begin{aligned}
g_{\text{MLRC}}(\boldsymbol{x}) = \min_{Y,\boldsymbol{y},\Lambda,\Xi,\boldsymbol{\lambda},\xi} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \operatorname{tr}(\tilde{Q}(\boldsymbol{x})^\top Y) + 2\tilde{\boldsymbol{c}}(\boldsymbol{x})^\top \boldsymbol{y} \\
\text{subject to} \quad & (53b) - (53c), (56b) \\
& \begin{bmatrix} Y & \boldsymbol{y} \\ \boldsymbol{y}^\top & 1 \end{bmatrix} \geq 0\,.
\end{aligned}$$

This readily implies that $g_1(\boldsymbol{x}) = g_{\text{MLRC}}(\boldsymbol{x})$ which completes the first part of the proof.

In the case of $g_{\text{SDP-LRC}}(\boldsymbol{x})$, the same argument leads us to establish

$$\begin{aligned}
g_{\text{SDP-LRC}}(\boldsymbol{x}) = \min_{Y,\boldsymbol{y},\Lambda,\Xi,\boldsymbol{\lambda},\xi} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \operatorname{tr}(\tilde{Q}(\boldsymbol{x})^\top Y) + 2\tilde{\boldsymbol{c}}(\boldsymbol{x})^\top \boldsymbol{y} \\
\text{subject to} \quad & (53b) - (53c), (56b) \\
& \begin{bmatrix} Y & \boldsymbol{y} \\ \boldsymbol{y}^\top & 1 \end{bmatrix} \geq 0\,, \begin{bmatrix} \Lambda & \Delta^\top & \boldsymbol{\lambda} \\ \Delta & \Xi & \boldsymbol{\zeta} \\ \boldsymbol{\lambda}^\top & \boldsymbol{\zeta}^\top & 1 \end{bmatrix} \succeq 0\,.
\end{aligned}$$

Although this does not exactly give rise to the optimization problem associated to $g_2(\boldsymbol{x})$ because of the different linear matrix inequalities, they are both equivalent because of constraint (56b) and the fact that Φ is full rank. Namely, it is clear that

$$\begin{bmatrix} \boldsymbol{I} & 0 & 0 & 0 & 0 \\ 0 & \boldsymbol{I} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \Phi = \boldsymbol{I} \quad \Rightarrow \quad \begin{bmatrix} \Lambda & \Delta^\top & \boldsymbol{\lambda} \\ \Delta & \Xi & \boldsymbol{\zeta} \\ \boldsymbol{\lambda}^\top & \boldsymbol{\zeta}^\top & 1 \end{bmatrix} \succeq 0 \Leftrightarrow \Phi \begin{bmatrix} \Lambda & \Delta^\top & \boldsymbol{\lambda} \\ \Delta & \Xi & \boldsymbol{\zeta} \\ \boldsymbol{\lambda}^\top & \boldsymbol{\zeta}^\top & 1 \end{bmatrix} \Phi^\top \succeq 0 \Leftrightarrow \begin{bmatrix} Y & \boldsymbol{y} \\ \boldsymbol{y}^\top & 1 \end{bmatrix} \succeq 0\,.$$

This implies that $g_2(\boldsymbol{x}) = g_{\text{SDP-LRC}}(\boldsymbol{x})$. The last two conclusions 3 and 4 of the theorem follow naturally from the fact that employing $\mathcal{K} \supseteq \mathcal{K}_1$ in problem (55) would mean that the feasible set is relaxed and must achieve a larger optimal value than MLRC, and similarly in the case of SDP-LRC. This completes our proof. □

References

ApS, MOSEK. 2017. MOSEK Optimization Toolbox for MATLAB 8.1.0.80. https://docs.mosek.com/8.1/toolbox/index.html.

Ardestani-Jaafari, A. 2016. Linearized robust counterparts with applications in location and inventory management problems. Ph.D. thesis, HEC Montréal.

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

Ardestani-Jaafari, Amir, Erick Delage. 2017. The value of flexibility in robust location transportation problems. Transportation Science Article in Advance.

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

Baron, Opher, Joseph Milner, Hussein Naseraldin. 2011. Facility location: A robust optimization approach. Production and Operations Management 20(5) 772–785.

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

Ben-Tal, Aharon, Laurent El Ghaoui, Arkadi Nemirovski. 2009. Robust optimization. Princeton University Press.

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

Bertsimas, Dimitris, Melvyn Sim. 2004. The price of robustness. Operations Research 52(1) 35–53.

Bomze, I. M., E. de Klerk. 2002. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. J. Global Optim. 24(2) 163–185

Buhayenko, Viktoryia, Dick den Hertog. 2017. Adjustable robust optimisation approach to optimise discounts for multi-period supply chain coordination under demand uncertainty. Int. J. Prod. Res. Article in Advance.

Burer, S. 2012. Copositive programming. In Handbook on semidefinite, conic and polynomial optimization. Springer, 201–218.

Chang, Y., S. Rao, M. Tawarmalani. 2017. Robust validation of network designs under uncertain demands and failures. 14th USENIX Symposium on Networked Systems Design and Implementation (NSDI 17). USENIX Association, Boston, MA, 347–362.

Dehghan, Shahab, Nima Amjady, Antonio J Conejo. 2017. Adaptive robust transmission expansion planning using linear decision rules. IEEE Trans. Power Syst. 32(5) 4024–4034.

Delage, E., D. A. Iancu. 2015. Robust Multistage Decision Making, chap. 2. INFORMS, 20–46.

Denton, B. T., A. J. Miller, H. J . Balasubramanian, T. R. Huschka. 2010. Optimal allocation of surgery blocks to operating rooms under uncertainty. Oper. Res. 58(4–part–1) 802–816.

Gauvin, Charles, Erick Delage, Michel Gendreau. 2017. Decision rule approximations for the risk averse reservoir management problem. Eur. J. Oper. Res. 261(1) 317–336.

Hanasusanto, G. A., D. Kuhn. 2016. Conic programming reformulations of two-stage distributionally robust linear programs over wasserstein balls. working paper .

Hiriart-Urruty, Jean-Baptiste, Claude Lemaréchal. 2001. Fundamentals of convex analysis. Springer, Berlin.

Jabr, Rabih A, Sami Karaki, Joe Akl Korbane. 2015. Robust multi-period opf with storage and renewables. IEEE Trans. Power Syst. 30(5) 2790–2799.

Kim, Byung Soo, Byung Do Chung. 2017. Affinely adjustable robust model for multiperiod production planning under uncertainty. IEEE Trans. Eng. Manag. 64(4) 505–514.

Lewandowski, D., D. Kurowicka, H. Joe. 2009. Generating random correlation matrices based on vines and extended onion method. Journal of Multivariate Analysis 100(9) 1989–2001.

Li, Zhigang, Wenchuan Wu, Boming Zhang, Bin Wang. 2015. Adjustable robust real-time power dispatch with large-scale wind power integration. IEEE Trans. Sustain. Energy 6(2) 357–368.

Löfberg, J. 2004. Yalmip : A toolbox for modeling and optimization in matlab. In Proceedings of the CACSD Conference. Taipei, Taiwan.

Lorca, Alvaro, Xu Andy Sun. 2017. Multistage robust unit commitment with dynamic uncertainty sets and energy storage. IEEE Trans. Power Syst. 32(3) 1678–1688.

Melamed, Michal, Aharon Ben-Tal, Boaz Golany. 2016. On the average performance of the adjustable RO and its use as an offline tool for multi-period production planning under uncertainty. Comput. Manag. Sci. 13(2) 1–23.

Parrilo, P. A. 2000. Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization. California Institute of Technology.

R. Rahmaniani, T. G. Crainic, M. Gendreau, W. Rei. 2017. The Benders decomposition algorithm: A literature review. European Journal of Operational Research 259(3) 801–817.

Scarf, H. 1958. A min-max solution of an inventory problem. H. E. Scarf K. S. Arrow, S. Karlin, ed., Studies in the mathematical theory of inventory and production. Stanford University Press, Stanford, CA, 201–209.

Shapiro, Alexander, Darinka Dentcheva, Andrzej Ruszczyński. 2009. Lectures on stochastic programming: modeling and theory. SIAM

Sherali, H. D., A. Alameddine. 1992. A new reformulation-linearization technique for bilinear programming problems. J. Global Optim. 2(4) 379–410.

Simchi-Levi, David, Nikolaos Trichakis, Peter Yun Zhang. 2016. Designing response supply chain against bioattacks. working paper .

Vairaktarakis, G.L. 2000. Robust multi-item newsboy models with a budget constraint. International Journal of Production Economics 66(3) 213–226.

Xu, G., S. Burer. 2016. A copositive approach for two-stage adjustable robust optimization with uncertain right-hand sides. working paper .

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