Linearized Robust Counterparts of Two-stage Robust Optimization Problems with Applications in Operations Management

Amir Ardestani-Jaafari, Erick Delage

HEC Montreal, Montreal, H3T 2A7, Canada

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

November 24, 2017

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 optimization problems, 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 inequality in the linearization process for bilinear models. We finally demonstrate how to employ this new scheme in a set of four operations management problems to improve the performance and guarantees of robust optimization.

Index terms— Two-stage adjustable robust optimization, affinely adjustable robust counterpart, linear programming relaxation, bilinear optimization.

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 realworld 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:

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

$$\text{subject to} \quad Ax + By(\zeta) \leq \Psi(x)\zeta\,,\ \forall\,\zeta \in \mathcal{U}\,, \tag{1b}$$

where $x \in \mathbb{R}^{n_x}$ identifies decisions that must be made "here and now" while $y : \mathbb{R}^{n_\zeta} \to \mathbb{R}^{n_y}$ identifies decisions that can adapt to the realization of the uncertain vector of parameters $\zeta \subset \mathbb{R}^{n_\zeta}$, with $\mathcal{U}$ as the "uncertainty set". Furthermore, the ARO model will have $A \in \mathbb{R}^{m\times n_x}$, $B \in \mathbb{R}^{m\times n_y}$, $c \in \mathbb{R}^{n_x}$, $d \in \mathbb{R}^{n_y}$, and $\Psi : \mathbb{R}^{n_x} \to \mathbb{R}^{m\times n_\zeta}$ such that $\Psi(x)$ is an affine function of variable 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 ζ, *i.e.,* the so called affine decision rule $y(\zeta) := Y\zeta + y$, for some $Y \in \mathbb{R}^{n_y\times n_\zeta}$ and $y \in \mathbb{R}^{n_y}$; therefore, problem (1) is conservatively approximated with

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

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

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). Although this literature seems to indicate that affine decision rules perform well in many applications, there are still a number of cases where AARC leads to 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:*

$$\begin{aligned}
\underset{x,v,y(\zeta)}{\text{maximize}} \quad & \min_{\zeta\in\mathcal{U}} \; -\sum_i c_i x_i + k_i v_i + \sum_i \sum_j \eta_{ij} y_{ij}(\zeta) && \text{(3a)}\\
\text{subject to} \quad & \sum_i y_{ij}(\zeta) \leq \zeta_j \,, \;\forall\, j \in \mathcal{J} \,, \;\forall\, \zeta \in \mathcal{U} && \text{(3b)}\\
& \sum_j y_{ij}(\zeta) \leq x_i \,, \;\forall\, i \in \mathcal{I} \,, \;\forall\, \zeta \in \mathcal{U} && \text{(3c)}\\
& y(\zeta) \geq 0 \,, \;\forall\, \zeta \in \mathcal{U} && \text{(3d)}\\
& 0 \leq x_i \leq M v_i \,, \;\forall\, i \in \mathcal{I} && \text{(3e)}\\
& v_i \in \{0,1\} \,, \;\forall\, i \in \mathcal{I} \,. && \text{(3f)}
\end{aligned}$$

In this problem, variable v_i indicates whether one opens a facility at location i for each $i \in \mathcal{I}$, 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 $j \in \mathcal{J}$. 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_i and k_i denote variable and fixed capacity cost for facility i 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 $\mathcal{U}$:

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

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 with a capacity of 24,000 units at the second location profit in order to achieve in the worst-case profit of 6,600 while under 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.

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

[1] To the best of our knowledge, the work of Chang et al. (2017) was developed independently from the results presented in

the first time a comprehensive investigation of how it can be employed on a general ARO problem. In doing so, we offer the following contributions:

- Our proposed scheme allows us to 1) re-interpret AARC as the model that is obtained when applying the linearization scheme found in Sherali and Alameddine (1992) to a certain bilinear optimization problem derived from the worst-case analysis (a.k.a. adversarial problem); 2) exploit valid linear and conic inequalities that are used in bilinear optimization problems to identify tractable reformulations that improve upon AARC; 3) demonstrate how such reformulations can identify solutions that are strict improvements over AARC on four types of operations management problems.
- 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:

$$\text{(Penalized ARO)} \quad \begin{array}{ll} \underset{x\in\mathcal{X},y(\zeta),z(\zeta))}{\text{maximize}} & \min_{\zeta\in\mathcal{U}} c^T x + d^T y(\zeta) - u^T z(\zeta) \\ \text{subject to} & Ax + By(\zeta) \leq \Psi(x)\zeta + z(\zeta)\,,\ \forall\,\zeta\in\mathcal{U}\,, \end{array}$$

  where $z : \mathbb{R}^{n_\zeta} \to \mathbb{R}^m$ is an adjustable variable that measures the amount of violation while $u \in \mathbb{R}^m$ is a vector of "large enough" (to be clarified later) marginal penalties. This procedure leads to a conservative approximation model (named MLRC) which is guaranteed to be feasible even when AARC fails to be feasible. An extensive set of numerical experiments performed on a facility location-transportation problem provides further evidence that this new scheme can significantly improve the quality of first stage decisions and bounds on worst-case profit for problem instances that suffer from large potential perturbations.
- Finally, we discuss 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 satisfied the relatively complete recourse property.

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. Next, in Section 3,

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.

we establish the equivalence between the LRC and AARC models. Section 4 describes the two methods that can be used to tighten the LRC/AARC approximation. Section 5 briefly describes how one might extend our results to general convex uncertainty sets. We then present in Section 6 four instances of operations management applications where the use of our new scheme leads to strict improvements compared to what is achieved by previous methods proposed in the literature. An exhaustive set of numerical experiments involving the location-transportation problem are presented in Section 7. Finally, we discuss the connection to recent copositive programming reformulations of ARO in Section 8 and conclude in Section 9.

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 $\mathcal{U} := \{\zeta | P\zeta \leq q\}$ where $P \in \mathbb{R}^{n_{\mathcal{U}} \times n_\zeta}, q \in \mathbb{R}^{n_{\mathcal{U}}}$.*

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

$$\forall\, x \in \mathcal{X}\,,\, \exists\, y(\zeta) :\, Ax + By(\zeta) \leq \Psi(x)\zeta \quad \forall\, \zeta \in \mathcal{U}\,.$$

Assumption 3. *For all $x \in \mathcal{X}$, there exists a feasible ζ, 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, a situation that is typically associated with an infinite loss. 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

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

where $g(x)$ is defined as

$$g(x) := \min_{\zeta \in \mathcal{U}} \quad \max_{y} \quad c^T x + d^T y \tag{5a}$$

$$\text{subject to} \quad Ax + By \leq \Psi(x)\zeta\,. \tag{5b}$$

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

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

where $\lambda \in \mathbb{R}^m$ is the dual variable associated with constraint (5b).

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

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

In Sherali and Alameddine (1992), the authors employ a linearization scheme that exploits a set of valid inequalities for a bilinear optimization problem similar to problem (6). In the context that we study here, this scheme leads us to consider that

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

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 $A = B$. Note that, in this model, constraints (7b) to (7e) are a set of redundant constraints that were added to problem (6). In particular, constraint (7b) is implied from

$$
\text{(6b)} \Rightarrow B^T \lambda \zeta^T = d\zeta^T \Rightarrow \zeta\lambda^T B = \zeta d^T\,.
$$

Moreover, constraints (7c)-(7e) can be similarly derived:

$$(6c)\&(6d) \Rightarrow P\zeta\lambda^T \leq q\lambda^T\,,$$

$$(6b) \Rightarrow B^T\lambda\lambda^T = d\lambda^T\,,$$

$$(6d) \Rightarrow \lambda\lambda^T \geq 0\,.$$

We next linearize problem (7) by introducing the variables $\Delta \in \mathbb{R}^{n_\zeta \times m}$ and $\Lambda \in \mathbb{R}^{m\times m}$, respectively defined as $\Delta := \zeta\lambda^T$ and $\Lambda := \lambda\lambda^T$, such that

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

A simple relaxation of problem (8) will lead to the linearized robust counterpart model for problem (1).

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

$$
\begin{aligned}
\textit{(LRC)} \quad & \underset{x\in\mathcal{X},Y,y,S,s}{\text{maximize}} && c^T x + d^T y - q^T s && \text{(9a)}\\
& \text{subject to} && P^T S = Y^T B^T - \Psi(x)^T && \text{(9b)}\\
& && Ax + By + S^T q \leq 0 && \text{(9c)}\\
& && P^T s = -Y^T d && \text{(9d)}\\
& && s \geq 0\,,\; S \geq 0\,, && \text{(9e)}
\end{aligned}
$$

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

Proof. First, we relax problem (8) by removing constraint (8g) to get a lower bound for $g(x)$. Next, we consider that since, when constraints (6b)-(6d), and (8c) are satisfied, one can simply let $\hat{\Lambda} := \hat{\lambda}\hat{\lambda}^T$

in order to satisfy constraints (8d)-(8f), the problem stays the same when disregarding Λ and the three constraints (8d)-(8f). Hence, we obtain a lower bound for $g(x)$ in the form

$$
\begin{aligned}
g(x) \geq g_{LRC}(x) := \quad & \min_{\zeta,\lambda,\Delta} && c^T x + \operatorname{tr}(\Psi(x)\Delta) - (Ax)^T \lambda && \text{(10a)}\\
& \text{subject to} && B^T \lambda = d && \text{(10b)}\\
& && P\zeta \leq q && \text{(10c)}\\
& && \lambda \geq 0 && \text{(10d)}\\
& && \Delta B = \zeta d^T && \text{(10e)}\\
& && P\Delta \leq q\lambda^T . && \text{(10f)}
\end{aligned}
$$

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

$$
\begin{aligned}
g_{LRC}(x) = \quad & \max_{Y,y,S,s} && c^T x + d^T y - q^T s && \text{(11a)}\\
& \text{subject to} && P^T S = Y^T B^T - \Psi(x)^T && \text{(11b)}\\
& && Ax + By + S^T q \leq 0 && \text{(11c)}\\
& && -P^T s = Y^T d && \text{(11d)}\\
& && s \geq 0 \, , \; S \geq 0 \, , && \text{(11e)}
\end{aligned}
$$

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

While it is known that evaluating the worst-case value of a given first-stage solution, *i.e.*, evaluating $g(x)$, is computationally intractable, we can show that it is possible to efficiently evaluate a bound on the worst-case value of the solution of the LRC model.

Lemma 2. *Given that $\hat{x}$ is the optimal solution of LRC* (9) *and $(\hat{\lambda},\, \hat{\zeta},\, \hat{\Delta})$ is the optimal solution of problem* (10)*, when x is fixed to $\hat{x}$, we have that*

$$0 \leq g(\hat{x}) - g_{LRC}(\hat{x}) \leq (\Psi(x)\hat{\zeta})^T \hat{\lambda} - \operatorname{tr}(\Psi(x)\hat{\Delta}) \, .$$

Proof. Given that $(\hat{\lambda},\, \hat{\zeta})$ is a feasible solution of problem (6) when x is fixed to $\hat{x}$, and since $g_{LRC}(\hat{x}) = c^T \hat{x} + \sum_{ij} \Psi(x)_{ij} \hat{\Delta}_{ji} - (A\hat{x})^T \hat{\lambda}$, this indicates that

$$0 \leq g(\hat{x}) - g_{LRC}(\hat{x}) \leq (\Psi(x)\hat{\zeta})^T \hat{\lambda} - \operatorname{tr}(\Psi(x)\hat{\Delta}) \, . \quad \square$$

$\square$

3 Relation to AARC

In this section, we explain how the LRC model can be considered equivalent to the conservative approximation model obtained with AARC.

Proposition 2. *LRC* (9) *is equivalent to AARC* (2)*.*

Proof. Proof. As a first step, we reformulate problem (10) in terms of an inner and an outer minimization operations

$$
\begin{aligned}
g_{LRC}(x) = \min_{\zeta \in \mathcal{U}} \quad & \min_{\lambda, \Delta} && c^T x + \mathrm{tr}(\Psi(x)\Delta) - (Ax)^T \lambda && \text{(12a)} \\
& \text{subject to} && B^T \lambda = d && \text{(12b)} \\
& && P\zeta \le q && \text{(12c)} \\
& && \lambda \ge 0 && \text{(12d)} \\
& && \Delta B = \zeta d^T && \text{(12e)} \\
& && P\Delta \le q\lambda^T \,. && \text{(12f)}
\end{aligned}
$$

We next derive the dual formulation of the inner minimization over λ and Δ as

$$
\begin{aligned}
\max_{y, Y, S} \quad & c^T x + d^T(y + Y\zeta) && \text{(13a)} \\
\text{subject to} \quad & Ax + By + S^T q \le 0 && \text{(13b)} \\
& P^T S = Y^T B^T - \Psi(x)^T && \text{(13c)} \\
& S \ge 0 \,, && \text{(13d)}
\end{aligned}
$$

where $y \in \mathbb{R}^{n_y}, Y \in \mathbb{R}^{n_y \times n_\zeta}$ and $S \in \mathbb{R}^{n_{\mathcal{U}} \times m}$ are the dual variables associated with constraints (12b), (12e), and (12f) respectively. Based on Sion's minimax theorem, since $\mathcal{U}$ is bounded, the same value for $g_{LRC}(x)$ can be obtained by reversing the order of minimization over ζ and maximization over y, Y, and S. Equivalently, we have that

$$
\begin{aligned}
g_{LRC}(x) = \quad & \max_{y, Y, S} && \min_{\zeta \in \mathcal{U}} c^T x + d^T(Y\zeta + y) && \text{(14a)} \\
& \text{subject to} && Ax + By + S^T q \le 0 && \text{(14b)} \\
& && P^T S = Y^T B^T - \Psi(x)^T && \text{(14c)} \\
& && S \ge 0 \,. && \text{(14d)}
\end{aligned}
$$

We next consider the i^{th} row of constraint (14b) and the i^{th} column of constraint (14c):

$$A_{i:}x + B_{i:}w + (S_{:i})^T q \leq 0, \tag{15a}$$

$$P^T S_{:i} = Y^T (B_{i:})^T - (\Psi(x)_{i:})^T \tag{15b}$$

where $A_{i:}$, $B_{i:}$, and $\Psi(x)_{i:}$ denote the i^{th} row of matrices A, B, and $\Psi(x)$ respectively, and $S_{:i}$ denotes the i^{th} column of matrix S. We show that constraints (15a) and (15b) are equivalent to

$$A_{i:}x + B_{i:}(Y\zeta + y) \leq \Psi(x)_{i:}\zeta,\ \forall\, \zeta \in \mathcal{U}\,. \tag{16}$$

We do so by considering that S is not in the objective function so that we can remove S from the set of decision variables and instead replace constraints (15a) and (15b) with

$$\begin{aligned} \min_{S_{:i}} \quad & A_{i:}x + B_{i:}w + (S_{:i})^T q && \leq 0\,, && (17a)\\ \text{subject to} \quad & P^T S_{:i} = Y^T (B_{i:})^T - (\Psi(x)_{i:})^T\,, && && (17b) \end{aligned}$$

where the embedded minimization problem can be replaced by a maximization problem, using duality theory. This leads us to considering constraint (17) as equivalent to

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

with ζ the dual variable of (17b). We have thus confirmed that constraints (15a) and (15b) are equivalent to (16), and likewise that constraints (14b) and (14c) are equivalent to constraint (2b). Therefore, the LRC model (9) is equivalent to the AARC model (2). $\square$ $\square$

4 Improving LRC and AARC using Valid Inequalities

In this section, we identify two types of valid inequalities that can be employed to formulate improved versions of LRC 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 λ in problem (6). This process will lead to a modified LRC model that preserves the computational complexity of LRC while guaranteeing the feasibility of the resulting approximation model. Secondly, we will identify a set of valid conic inequalities that will lead to a semi-definite programming formulation for LRC.

4.1 Exploiting an implicit bound on λ^*

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 $u \in \mathbb{R}^m$ such that, for all $x \in \mathcal{X}$ and for all $\zeta \in \mathcal{U}$, there exists an optimal solution $\lambda^* \leq u$ for the problem*

$$\begin{aligned} \underset{\lambda}{\text{minimize}} \quad & (\Psi(x)\zeta - Ax)^T \lambda && \text{(18a)} \\ \text{subject to} \quad & B^T \lambda = d && \text{(18b)} \\ & \lambda \geq 0\,. && \text{(18c)} \end{aligned}$$

At first glance, one might consider that Assumption 4 is rather limiting given that λ is an abstract object. One should, however, be aware that such a bound u is guaranteed to exists. Indeed, based on Lemma 1 there always exist a finite optimal solution for problem (18) 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 constraint is finite, we must conclude that such a u exists. A challenge remains in identifying the tightest values possible for 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 will later in subsections 6.3 and 6.4 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 $u := M$ satisfies Assumption 4, for any fixed $k = 1, 2, \ldots, m$, one can identify a tighter bound u_k^* that satisfies Assumption 4 by solving the following mixed-integer linear program*

$$\begin{aligned} u_k^* := \quad \max_{\lambda, v} \quad & \lambda_k && \text{(19a)} \\ \text{subject to} \quad & B^T \lambda = d && \text{(19b)} \\ & \lambda_i \leq M(1 - v_i)\,,\, \forall\, i && \text{(19c)} \\ & \sum_{i=1}^m v_i + n_y = n_\lambda && \text{(19d)} \\ & \lambda \geq 0\,,\, v \in \{0,1\}^m\,. && \text{(19e)} \end{aligned}$$

Proof. Proof. For any fixed $x \in \mathcal{X}$ and $\zeta \in \mathcal{U}$, given that problem (18) is known to be feasible and finite (see Lemma 1), it must be the case that an optimal solution λ^* for the problem is located at one of the vertices of the polyhedra defined by $B^T\lambda = d$ and $\lambda \geq 0$. By definition, for any vertex λ^ν of this polyhedra, there must exists a set of indices $\mathcal{I} \subseteq \{1, \ldots, m\}$ such that:

1. $|\mathcal{I}| = n_\lambda - n_y$

2. the set $\{B_{:1}, \ldots, B_{:n_y}\} \cup \{e_i\}_{i\in\mathcal{I}}$ contains linearly independent vectors

3. λ^ν is the unique solution to the following system of equations

$$B^T\lambda^\nu = d \qquad\qquad e_i^T\lambda^\nu = 0\,,\ \forall\, i \in \mathcal{I}\,.$$

It therefore follows that if $\|\lambda^\nu\|_\infty \leq M$ for all the vertex v of this polyhedron, then problem (19) necessarily returns an optimal value $\bar{\lambda}_k^*$ that satisfies $\lambda_k^* \leq \max_\nu \lambda_k^\nu \leq u_k^*$. □ □

It is worth mentioning that the bound obtained using Proposition 3 can be further tightened by also imposing that

2'. the set $\{B_{:1}, \ldots, B_{:n_y}\} \cup \{e_i\}_{i\in\mathcal{I}:v_i=1}$ contains linearly independent vectors.

Although it is not clear how to solve the resulting problem in the form of a MILP, since verifying whether such a condition is satisfied is rather straightforward (using the rank of the matrix composed of the different vectors), in practice one can employ a constraint-generation scheme that progressively adds linear constraints to problem (19) in order to reject the current optimal solutions for v, until problem (19) returns an optimal solution that satisfies this condition.

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

$$\begin{aligned}
\text{(MLRC)} \quad \underset{x\in\mathcal{X},Y,y,S,s,W,w}{\text{maximize}} \quad & c^Tx + d^Ty - q^Ts - u^Tw - u^TWq && (20a)\\
\text{subject to} \quad & Y^TB^T - P^T(S-W) = \Psi(x)^T && (20b)\\
& Ax + By + (S-W)^Tq - w \leq 0 && (20c)\\
& -P^T(s + Wu) = Y^Td && (20d)\\
& s \geq 0\,,\ S \geq 0\,,\ w \geq 0\,,\ W \geq 0\,, && (20e)
\end{aligned}$$

where $Y \in \mathbb{R}^{n_y\times n_\zeta}$, $y \in \mathbb{R}^{n_y}$, $S \in \mathbb{R}^{n_\mathcal{U}\times m}$, $s \in \mathbb{R}^{n_\mathcal{U}}$, $W \in \mathbb{R}^{n_\mathcal{U}\times m}$, *and* $w \in \mathbb{R}^m$. *Furthermore, the optimal value of problem* (20) *is necessarily larger than or equal to the optimal value of LRC* (9)*.*

Proof. 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:

$$\lambda \leq u \tag{21a}$$

$$(q - P\zeta)(u - \lambda)^T \geq 0 \tag{21b}$$

where constraint (21b) can be linearized by replacing $\Delta := \zeta\lambda^T$ as

$$P\Delta \geq q\lambda^T - (q - P\zeta)u^T \,. \tag{22}$$

Adding constraints (21a) and (22) to problem (10) leads to MLRC after applying duality theory. □ □

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* (20) *is equivalent to applying affine decision rules to the following two-stage problem:*

$$\begin{aligned}
&\underset{x\in\mathcal{X},y(\zeta),z(\zeta)}{\text{maximize}} && \min_{\zeta\in\mathcal{U}} c^T x + d^T y(\zeta) - u^T z(\zeta) && \text{(23a)}\\
&\text{subject to} && Ax + By(\zeta) \leq \Psi(x)\zeta + z(\zeta),\ \forall\, \zeta \in \mathcal{U} && \text{(23b)}\\
& && z(\zeta) \geq 0\,,\ \forall\, \zeta \in \mathcal{U}\,. && \text{(23c)}
\end{aligned}$$

When observing problem (23) closely, one can readily recognize that it replaces the ARO problem with

$$\underset{x\in\mathcal{X},y(\zeta),z(\zeta)}{\text{maximize}} \quad \min_{\zeta\in\mathcal{U}} c^T x + d^T y(\zeta) - \sum_i u_i \max(0\,,\, e_i^T(\Psi(x)\zeta - Ax - By(\zeta)))\,,$$

where $e_i \in \mathbb{R}^m$ 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 $y(\cdot)$ 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 (23) is equivalent to problem (1), i.e., that there is an optimal solution of (23) with $z^*(\zeta) := 0$. The penalties described in u should, however, be selected as small as possible in order for the estimate of the worst-case profit 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 rules obtained from solving penalized models such as (23) might not be implementable for all possible realizations of ζ 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 x^*, it implies that once x^* is implemented and $\bar{\zeta}$ is observed, one needs to seek the recourse decision that minimizes the specific recourse problem that is being experienced,[2] *i.e.,*

$$\begin{aligned} \underset{y}{\text{maximize}} \quad & d^T y \\ \text{subject to} \quad & Ax^* + By \leq \Psi(x^*)\bar{\zeta} \end{aligned}$$

When doing so, the incurred cost is guaranteed to be smaller than what had been anticipated by the conservative approximation scheme applied to (23).

Proof. Proof. Adding constraints (21a) and (22) to problem (10) leads to the following formulation:

$$\begin{aligned} g_{MLRC}(x) := \quad \min_{\zeta,\lambda,\Delta} \quad & c^T x + \text{tr}(\Psi(x)\Delta) - (Ax)^T \lambda && \text{(24a)} \\ \text{subject to} \quad & \text{(10b)} - \text{(10f)} \\ & \lambda \leq u && \text{(24b)} \\ & P\Delta \geq q\lambda^T - (q - P\zeta)u^T \,. && \text{(24c)} \end{aligned}$$

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

$$\begin{aligned} g_{MLRC}(x) = \quad \max_{Y,y,W,w,S} \quad & \min_{\zeta\in\mathcal{U}} c^T x + d^T(Y\zeta + y) - u^T(w + W^T(q - P\zeta)) && \text{(25a)} \\ \text{subject to} \quad & P^T S = Y^T B^T - \Psi(x)^T + P^T W && \text{(25b)} \\ & Ax + Bw + S^T q \leq W^T q + w && \text{(25c)} \\ & S \geq 0 && \text{(25d)} \\ & w \geq 0 \,,\, W \geq 0 \,, && \text{(25e)} \end{aligned}$$

where $w \in \mathbb{R}^m$ and $W \in \mathbb{R}^{n_{\mathcal{U}} \times m}$ are respectively the dual variables associated with constraints (24b) and (24c). Again, the constraints (25b)-(25d) can be replaced with

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

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

$$\begin{aligned} g_{MLRC}(x) = \quad \max_{Y,y,W,w} \quad & \min_{\zeta\in\mathcal{U}} c^T x + d^T(Y\zeta + y) - u^T(w + W^T(q - P\zeta)) && \text{(26a)} \\ \text{subject to} \quad & Ax + B(Y\zeta + y) \leq \Psi(x)\zeta + w + W^T(q - P\zeta), \,\forall\, \zeta \in \mathcal{U} && \text{(26b)} \\ & w \geq 0 \,,\, W \geq 0 \,. && \text{(26c)} \end{aligned}$$

[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.

We next introduce new variables z and Z as

$$z := w + W^T q\,, \qquad Z := -W^T P\,,$$

where $z \in \mathbb{R}^m$ and $Z \in \mathbb{R}^{m \times n_\zeta}$. Therefore, problem (25) can be reformulated as

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

We finally show that constraints (27c)-(27e) are equivalent to the following constraint:

$$z + Z\zeta \geq 0\,, \forall\, \zeta \in \mathcal{U}\,. \tag{28}$$

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

$$\begin{aligned}
& \max_{w_{:j}} && z_j - W_{:j}^T q && \geq 0 && \text{(29a)}\\
& \text{subject to} && Z_{j:}^T = -P^T W_{:j} && && \text{(29b)}\\
& && W_{:j} \geq 0 && && \text{(29c)}
\end{aligned}$$

as the constraint that

$$\begin{aligned}
& \min_{\zeta} && z_j + Z_{j:}\zeta && \geq 0\,. && \text{(30a)}\\
& \text{subject to} && P\zeta \leq q && && \text{(30b)}
\end{aligned}$$

This completes our proof. $\square$ $\square$

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* (20) *and the optimization problem obtained by employing static or affine decision rules in problem* (23) *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 (23) which can be shown to have a static feasible solution. In particular, let $\bar{x}$ be any member of $\mathcal{X}$, while letting $y(\zeta) := 0$ and for each $z_i(\zeta) := \max(0; \max_{\zeta\in\mathcal{U}} e_i^T(A\bar{x} - \Psi(\bar{x})\zeta))$, where $e_i \in \mathbb{R}^m$ 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 (20) otherwise it would be in contradiction with Proposition 5.

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

$$\begin{aligned}
\underset{x\in\mathcal{X},y(\zeta),z(\zeta)}{\text{maximize}} \quad & \min_{\zeta\in\mathcal{U}} c^T x + d^T y(\zeta) - q_\lambda^T z(\zeta) \\
\text{subject to} \quad & Ax + By(\zeta) \leq \Psi(x)\zeta + P_\lambda^T z(\zeta),\, \forall\, \zeta \in \mathcal{U} \\
& z(\zeta) \geq 0\,,\, \forall\, \zeta \in \mathcal{U}\,.
\end{aligned}$$

Furthermore, if P_λ is such that there exists some z such that $P_\lambda^T z > 0$ (as is the case when $P_\lambda = I$ and $q_\lambda = u$) 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 $\{\lambda \,|\, \lambda \geq 0,\, B^T\lambda = d\}$.

4.2 Exploiting Valid Conic Inequalities

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

$$\begin{bmatrix} \Lambda & \Delta^T \\ \Delta & \Xi \end{bmatrix} = \begin{bmatrix} \lambda \\ \zeta \end{bmatrix} \begin{bmatrix} \lambda^T & \zeta^T \end{bmatrix}, \tag{31}$$

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

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

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

$$\begin{bmatrix} \Lambda & \Delta^T & \lambda \\ \Delta & \Xi & \zeta \\ \lambda^T & \zeta^T & 1 \end{bmatrix} \succeq 0\,. \tag{32}$$

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)*:*

$$
\begin{aligned}
g_{\text{SDP-LRC}}(x) = \quad & \min_{\zeta,\lambda,\Delta,\Lambda,\Xi} && c^T x + \mathrm{tr}(\Psi(x)\Delta) - (Ax)^T\lambda && \text{(33a)}\\
& \text{subject to} && B^T\lambda = d && \text{(33b)}\\
& && P\zeta \le q && \text{(33c)}\\
& && 0 \le \lambda \le u && \text{(33d)}\\
& && \Delta B = \zeta d^T && \text{(33e)}\\
& && P\Delta \le q\lambda^T && \text{(33f)}\\
& && P\Delta \ge q\lambda^T - (q - P\zeta)u^T && \text{(33g)}\\
& && \begin{bmatrix} \Lambda & \Delta^T & \lambda \\ \Delta & \Xi & \zeta \\ \lambda^T & \zeta^T & 1 \end{bmatrix} \succeq 0 && \text{(33h)}\\
& && \Lambda B = \lambda d^T && \text{(33i)}\\
& && P\Xi P^T + qq^T \ge P\zeta q^T + q\zeta^T P^T && \text{(33j)}\\
& && 0 \le \Lambda \le u\lambda^T && \text{(33k)}\\
& && \Lambda + uu^T \ge \lambda u^T + u\lambda^T\,. && \text{(33l)}
\end{aligned}
$$

Moreover, the optimal value of $\max_{x\in\mathcal{X}} g_{\text{SDP-LRC}}(x)$ *is necessarily larger than or equal to the optimal value of LRC* (9) *and MLRC* (20)*.*

Proof. Proof. We start by including the new variables Λ and Ξ and constraint (32) in problem (10). One can then realize that constraints (8d) and (8e) can now help tighten the feasible region. Finally, a final tightening step can be achieved by exploiting the fact that $P\zeta \le q$ implies the following:

$$q - P\zeta \ge 0 \Rightarrow (q - P\zeta)(q - P\zeta)^T \ge 0 \Rightarrow P\zeta\zeta^T P^T + qq^T \ge P\zeta q^T + q\zeta^T P^T$$

and that $\lambda \leq u$ implies that $\lambda\lambda^T \leq u\lambda^T$ and

$$(u-\lambda)(u-\lambda)^T \geq 0 \;\Rightarrow uu^T - u\lambda^T - \lambda u^T + \lambda\lambda^T \geq 0\,,$$

which together lead to constraints (33j), (33k), and (33l) after replacing $\Xi := \zeta\zeta^T$ and $\Lambda := \lambda\lambda^T$. $\square$ $\square$

Note that, although we presented $g_{\text{SDP-LRC}}(x)$ as a minimization problem, a semi-definite programming duality can be employed to obtain a maximization representation of this function that can be integrated with the maximization in x as was done with other LRC models. We, however, omit the details of this reformulation 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 model. In fact, the authors of Ardestani-Jaafari and Delage (2016) were able to establish such a connection for a special case of the SDP-LRC model. 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.

5 LRC for General Uncertainty Sets

In this section, we extend our LRC model so that it can accomodate general convex uncertainty sets, *i.e.*, $\mathcal{U}$ is not polyhedral. In this regard, we will instead consider uncertainty sets that can be represented as

$$\mathcal{U}_{general} := \{\zeta \in \mathbb{R}^{n_\zeta} \mid f_l(\zeta) \leq q_l, \forall\, l = 1, \ldots, L\} \tag{34}$$

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 $\bar{\zeta}$ such that $f_l(\bar{\zeta}) < q_l$ for all $l = 1, \ldots, L$. Furthermore, for each $l = 1, \ldots, L$, 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 $f_l(x) = \sup_y x^T y - f_*(y)$, where $f_{l*}(y)$ is the convex conjugate of $f_l(\cdot)$, i.e., $f_{l*}(y) := \sup_z y^T z - f_l(z)$.*

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

Assumption 5, we have that

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

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

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

$$g(x) = \min_{\zeta,\lambda} \quad c^T x + (\Psi(x)\zeta)^T\lambda - (Ax)^T\lambda \tag{35a}$$
$$\text{subject to} \quad B^T\lambda = d \tag{35b}$$
$$f_l(\zeta) \le q_l, \forall\, l \tag{35c}$$
$$0 \le \lambda \le u \tag{35d}$$
$$\zeta\lambda^T B = \zeta d^T \tag{35e}$$
$$\lambda_i f_l(\zeta) \le \lambda_i q_l\,, \forall\, i\,, \forall\, l \tag{35f}$$
$$(q_l - f_l(\zeta))(u_i - \lambda_i) \ge 0\,, \forall\, i\,, \forall\, l\,. \tag{35g}$$

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

$$g(x) \ge\ g_{GLRC}(x) = \min_{\zeta,\lambda,\Delta} \quad c^T x + \mathrm{tr}(\Psi(x)\Delta) - (Ax)^T\lambda \tag{36a}$$
$$\text{subject to} \quad B^T\lambda = d \tag{36b}$$
$$f_l(\zeta) \le q_l, \forall\, l \tag{36c}$$
$$0 \le \lambda \le u \tag{36d}$$
$$\Delta B = \zeta d^T \tag{36e}$$
$$h_l(\Delta_{:i}, \lambda_i) \le q_l\lambda_i\,, \forall\, i\,, \forall\, l \tag{36f}$$
$$h_l(\zeta u_i - \Delta_{:i}, u_i - \lambda_i) \le q_l(u_i - \lambda_i)\,, \tag{36g}$$

where constraint (36f) is derived from

$$f_l(\zeta) \le q_l \Rightarrow \lambda_i f_l(\zeta) \le q_l\lambda_i \Rightarrow \lambda_i(\sup_y \zeta^T y - f_{l*}(y)) \le q_l\lambda_i \Rightarrow \sup_y (\lambda_i\zeta)^T y - \lambda_i f_{l*}(y) = h(\Delta_{:i}, \lambda_i) \le q_l\lambda_i$$

and constraint (36g) is derived from

$$\begin{aligned}\big(q_l - f_l(\zeta)\big)(u_i - \lambda_i) \geq 0 &\Rightarrow (u_i - \lambda_i) f_l(\zeta) \leq q_l(u_i - \lambda_i)\\
&\Rightarrow \sup_y (u_i - \lambda_i)\zeta^T y - (u_i - \lambda_i) f_{l*}(y) \leq q_l(u_i - \lambda_i)\\
&\Rightarrow h_l(\zeta(u_i - \lambda_i), u_i - \lambda_i) \leq q_l(u_i - \lambda_i)\\
&\Rightarrow h_l(\zeta u_i - \Delta_{:i}, u_i - \lambda_i) \leq q_l(u_i - \lambda_i)\,,\end{aligned}$$

for all i and for all l, and where the term $\zeta\lambda_i$ is linearized through $\Delta_{:i}$. One might apply duality theory to problem (36) 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 A will demonstrate that the problem maximize$_{x\in\mathcal{X}}\, g_{GLRC}(x)$ always produces a tighter conservative approximation than employing affine decision rules in problem (23). It is worth noting however that, based on the proof presented in Appendix A, it appears legitimate to believe that the two schemes are equivalent in most practical situations.

6 Examples

In this section, we provide examples of four applications for which the schemes presented in Section 4 (namely MLRC and SDP-LRC) can provide strict improvement compared to approaches that are commonly used in practice. The first and second applications, namely a multi-item newsvendor problem and a surgery block allocation problem, will involve special cases of ARO (known as complete recourse problem) for which MLRC trivially reduces to AARC so that only SDP-LRC can improve performance. Surprisingly, for the surgery allocation problem we will show that AARC already provides strictly better solutions than the model recently presented in Denton et al. (2010), although the latter model was believed by the authors to provide an optimal robust solution. For the third and fourth application, namely the facility location problem introduced in Section 1 and a multi-product assembly problem, we will show how one obtains a u that satisfies Assumption 4 in order to formulate the MLRC model and demonstrate that using this MLRC can generate strict improvement. We provide further evidence of the extent of this improvement in Section 7 where extensive numerical experiments for the facility location problem are presented.

We note that, in this section, in order to be more concise in our descriptions, we let $\mathcal{I}$ and $\mathcal{J}$ represent the sets $\{1,\ldots,m\}$ and $\{1,\ldots,n\}$ respectively, and consider that $i\in\mathcal{I}$ and $j\in\mathcal{J}$ when the membership for i and j is left unspecified.

6.1 Multi-item newsvendor problem

Consider the following robust multi-item newsvendor problem:

$$\max_{x\in\mathcal{X}} \min_{\zeta\in\mathcal{U}} \sum_j r_j \min(x_j,\zeta_j) - c_j x_j + s_j \max_j(x_j-\zeta_j,0) - p_j \max(\zeta_j - x_j, 0)\,, \tag{37}$$

where r_j, c_j, $s_j \leq r_j$, and p_j denote sale price, ordering cost, salvage price, and shortage cost of a unit of the j-th item, $j \in \mathcal{J}$, respectively, and ζ_j denotes the demand for item j for each j. Problem (37) is a special case of ARO, as

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

In Ardestani-Jaafari and Delage (2016), the authors show that affine decision rules are actually optimal in this problem when $\mathcal{U}$ is defined as

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

where affine decision rules are made with respect to (δ^+, δ^-), and when Γ is an integer value. One should however explore for other uncertainty set whether it is possible to get a conservative approximation that is tighter than AARC using the MLRC or SDP-LRC models. Regarding the MLRC model, one can actually show that the dual variables of the recourse problem are already bounded for this multi-item newsvendor problem. Specifically, the dual problem takes the shape of

$$\begin{aligned}
\min_{\lambda^1,\lambda^2} \quad & \sum_j \lambda_j^1((r_j - c_j)x_j - (r_j - s_j)(x_j - \zeta_j)) + \sum_j \lambda_j^2((r_j - c_j)x_j - p_j(\zeta_j - x_j)) \\
\text{subject to} \quad & \lambda_j^1 + \lambda_j^2 = 1\,,\ \forall\, j \in \mathcal{J}\,, \\
& \lambda^1 \geq 0\,,\ \lambda^2 \geq 0\,,
\end{aligned}$$

which already implies that $\lambda^1 \leq 1$ and $\lambda^2 \leq 1$ for any feasible solution. Hence, when formulating MLRC model (20) with $u = 1$, one obtains a model that is exactly equivalent to LRC model (9) and AARC (following Proposition 2). We are therefore left with the question of whether SDP-LRC can provide strict performance improvements. We will achieve this with the following example.

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

$$\tilde{\mathcal{U}}(\Gamma) = \left\{ \zeta \;\middle|\; \exists\, \delta^+, \delta^- \in \mathbb{R}^n, \quad \begin{array}{l} \delta_j^- \geq 0, \delta_j^+ \geq 0 \\ \delta_j^+ + \delta_j^- \leq 1\,, \forall j \\ \sum_j \delta_j^+ + \delta_j^- = \Gamma \end{array} \quad \& \quad \begin{array}{l} \zeta_1 = \bar{\zeta}_j + \hat{\zeta}_j(\delta_1^+ + \delta_2^+ - \delta_1^- - \delta_2^-)/2 \\ \zeta_2 = \bar{\zeta}_j + \hat{\zeta}_j(\delta_2^+ + \delta_3^+ - \delta_2^- - \delta_3^-)/2 \\ \zeta_3 = \bar{\zeta}_j + \hat{\zeta}_j(\delta_3^+ + \delta_1^+ - \delta_3^- - \delta_1^-)/2 \end{array} \right\}.$$

We compare, as it is shown in Table 1, 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. We next study what is the price that is paid in terms of computations.

In order to evaluate how each model trades-off between solution computation time and quality, we generated 10 random instances of multi-item newsvendor problems with 5, 10, 20, 30, and 40 items and uncertainty budgets Γ varying between 10% and 90% of the total number of items. For each case, a random instance was generated by letting $r = 100$ and selecting each c_j uniformly on the interval [1, 60], each $p_j \sim [1, c_j]$, and $s_j \sim [1, c_j]$. The uncertainty set was randomly constructed using

$$\tilde{\mathcal{U}}_P(\Gamma) = \left\{ \zeta \;\middle|\; \exists\, \delta^+, \delta^- \in \mathbb{R}^n, \quad \delta_j^- \geq 0,\ \delta_j^+ \geq 0,\ \delta^+ + \delta^- \leq 1,\ \sum_j \delta_j^+ + \delta_j^- = \Gamma,\ \zeta = \bar{\zeta} + \mathrm{diag}(\hat{\zeta})P(\delta^+ - \delta^-) \right\}.$$

where $\hat{\zeta} \sim [1, \bar{\zeta}]$ with $\bar{\zeta} = 100$, and where $P \in \mathbb{R}^{n \times n}$ is a matrix where each term was uniformly drawn from $[-1,\ 1]$ and each row normalized so that it sums to one.

Table 2 presents the average computation time take by CPLEX 12.7.1 to solve AARC, and by DSDP 5.8 (Benson et al., 2000) to solve SDP-LRC, and SDP-A&D. It also presents the average relative improvement achieved by SDP-LRC and SDP-A&D in terms of optimal bound on best achievable worst-case profit. Based on these results, it appears that SDP-LRC does give rise to a heavier computation burden than SDP-A&D but also to solutions of noticeably better quality (19% improvement on average).

Table 1: Comparison of optimal bound and worst-case profit associated to solutions obtained from conservative approximation models and exact models in a newsvendor problem

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

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

6.2 Surgery block allocation problem

Consider the following surgery block allocation problem:

$$\begin{aligned}
\underset{x,Z,y(\zeta)}{\text{minimize}} \quad & \max_{\zeta\in\mathcal{U}} \; c\sum_i x_i + d\sum_i y_i(\zeta) && \text{(38a)}\\
\text{subject to} \quad & y_i(\zeta) \geq \sum_j \zeta_j Z_{ij} - wx_i \,,\; \forall\, i \in \mathcal{I} \,,\; \forall\, \zeta \in \mathcal{U} && \text{(38b)}\\
& y(\zeta) \geq 0 \,,\; \forall\, \zeta \in \mathcal{U} && \text{(38c)}\\
& \sum_i Z_{ij} = 1 \,,\; \forall\, j \in \mathcal{J} && \text{(38d)}\\
& Z_{ij} \leq x_i \,,\; \forall\, i \in \mathcal{I} && \text{(38e)}\\
& x \in \{0,1\}^m \;,\; Z \in \{0,1\}^{m\times n} \,, && \text{(38f)}
\end{aligned}$$

where for each $i = 1, 2, \ldots, m$, variable x_i denotes whether we will open Operating Room (OR) i or not, while, for each $j = 1, 2, \ldots, n$, 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 (38d) captures the fact that a surgery block needs to be assigned to exactly one OR, while constraint (38e) captures the fact that surgery blocks can be assigned to an OR only if it is opened.

Proposition 7. *When* $\mathcal{U} := \{\zeta \in \mathbb{R}^n \,|\, \exists\, \delta \in [0,1]^n, \zeta = \bar{\zeta} + \hat{\zeta}\delta,\; \sum_j \delta_j \leq \Gamma\}$*, employing affine decision rules in the surgery block allocation problem provides a conservative approximation that is at least as tight as the reformulation proposed in Denton et al. (2010) (see model (40) in that paper).*

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

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

where

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

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

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

where $\lambda \in \mathbb{R}^m$ 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 (38) is equivalent to optimizing

$$\begin{aligned}
\underset{x,Z}{\text{minimize}} \quad & g_{LRC}(x, Z) \\
\text{subject to} \quad & \text{(38d)}, \text{(38e)}, \text{(38f)} \,,
\end{aligned}$$

where

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

We will now exploit the fact that we can add the constraint $\sum_{ij} \Delta_{ij} \leq \Gamma$ to the problem associated to $g_{LRC}(x,Z)$ without affecting the optimal value that it will return. This is because, for any optimal solution $(\lambda^*, \delta^*, \Delta^*)$, 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

$$
\begin{aligned}
& 0 \leq \Delta_{ij} \leq \delta_j \Rightarrow 0 \leq \Delta_{ij} Z_{ij} \leq \delta_j \Rightarrow 0 \leq \Delta'_{ij} \leq \delta_j \\
& \Delta_{ij} \leq z_i \Rightarrow \Delta_{ij} Z_{ij} \leq z_i \Rightarrow \Delta'_{ij} \leq z_i \\
& \sum_j \Delta_{ij} \leq \Gamma z_i \Rightarrow \sum_j \Delta_{ij} Z_{ij} \leq \Gamma z_i \Rightarrow \sum_j \Delta'_{ij} \leq \Gamma z_i \\
& \sum_{ij} \Delta'_{ij} = \sum_{ij} \Delta_{ij} Z_{ij} \leq \sum_{ij} \delta_j Z_{ij} =\leq \sum_j \delta_j \sum_i Z_{ij} = \sum_j \delta_j \leq \Gamma\,.
\end{aligned}
$$

Hence, we have that

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

meaning that, for any feasible x and Z, it must be that $g_{LRC}(x, Z) \leq g_{Denton}(x, Z)$, 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. □ □

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. Hence, this exact solution will also be identified by MLRC, which is actually equivalent to AARC following a siilar argument as presented in Section 6.1, and by SDP-LRC. Table 3 summarizes the optimal bounds on worst-case cost obtained for the two types of solutions (*i.e.,* open one or two ORs) using AARC, MLRC, SDP-LRC, and an exact solution of ARO.

The fact that employing affine decision rules can provide better solutions than the model in Denton et al. (2010) is surprising, as it is stated in Denton et al. (2010) that their proposed reformulation provides an exact solution to the robust surgery block allocation problem. The issue in the argument presented by the authors of that paper is found in their Proposition 6, which states that a certain polyhedron only has integer extreme points. Appendix B provides arguments that disprove this proposition.

6.3 Location-transportation problem

We now revisit the location-transportation problem presented in (3) and describe how to formulate the associated MLRC. We will then establish whether MLRC is able to provide a less conservative solution than the one obtained by AARC and the model proposed in Ardestani-Jaafari and Delage (2017) for the instance of this problem described in Example 1.

In order to formulate the MLRC model, one first needs to identify a u that satisfies Assumption 4. To do so, we start by reminding the reader of the recourse problem that arises in this application:

$$\begin{aligned}
\underset{y}{\text{maximize}} \quad & -\sum_i c_i x_i + k_i v_i + \sum_i \sum_j \eta_{ij} y_{ij} && \text{(39a)}\\
\text{subject to} \quad & \sum_i y_{ij} \leq \zeta_j \,,\ \forall\, j \in \mathcal{J} && \text{(39b)}\\
& \sum_j y_{ij} \leq x_i \,,\ \forall\, i \in \mathcal{I} && \text{(39c)}\\
& y \geq 0\,. && \text{(39d)}
\end{aligned}$$

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

$$\begin{aligned}
\underset{\lambda^1,\lambda^2,\lambda^3}{\text{minimize}} \quad & -\sum_i c_i x_i + k_i v_i + \sum_j \zeta_j \lambda_j^1 + \sum_i x_i \lambda_i^2 && \text{(40a)}\\
\text{subject to} \quad & \lambda_j^1 + \lambda_i^2 - \lambda_{ij}^3 = \eta_{ij} \,,\ \forall\, i \in \mathcal{I} \,,\ \forall\, j \in \mathcal{J} && \text{(40b)}\\
& \lambda^1 \geq 0\,,\ \lambda^2 \geq 0\,,\ \lambda^3 \geq 0\,, && \text{(40c)}
\end{aligned}$$

where $\lambda^1 \in \mathbb{R}^n$, $\lambda^2 \in \mathbb{R}^m$, and $\lambda^3 \in \mathbb{R}^{m\times n}$ are dual variables associated with constraints (39b), (39c), and (39d) respectively. Our goal is therefore to identify tight bounds for λ^{1*}, λ^{2*}, and λ^{3*}.

Since $x \geq 0$ and $\zeta \geq 0$, the objective function of problem (40) is non-decreasing in λ^1 and λ^2. One can therefore conclude that, at optimum, each term of λ^{1*} will be such that it will either be equal to 0 or involved in a one 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

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

Similarly, in the case of λ^{2*}, we have

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

Finally, since $\lambda_{ij}^{3*} = \lambda_j^{1*} + \lambda_i^{2*} - \eta_{ij} := u_{ij}^3$, one could conclude that $\lambda_{ij}^3 \leq \max(0\,,\, \max_i \eta_{ij} + \max_j \eta_{ij} - \eta_{ij})$.

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:

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

where $z^1 : \mathbb{R}^n \to \mathbb{R}^n$, $z^2 : \mathbb{R}^n \to \mathbb{R}^m$, and $z^3 : \mathbb{R}^n \to \mathbb{R}^{m\times n}$ 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 where $z_j^1(\zeta) := z_j^1 \zeta_j$, $z^2(\zeta) := 0$, and $z^3(\zeta) := 0$, 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 λ^{1*}, λ^{2*}, and λ^{3*} can be improved. This can be done by solving problem (19) with $M := 2\|\eta\|_\infty = 11.8$ since for $k = 1, 2, 3$ we have that $\lambda^{k*} \leq \max_{ij} |\eta_{ij}|$, and exploiting the proposed constraint generation scheme. This leads to establishing the following bounds:

$$u^1 := \begin{bmatrix} 5.9 \\ 5.9 \\ 4.9 \end{bmatrix} \qquad u^2 := \begin{bmatrix} 5.9 \\ 5.9 \end{bmatrix} \qquad u^3 := \begin{bmatrix} 0 & 0.6 & 1 \\ 7.2 & 0 & 1 \end{bmatrix} .$$

Table 4 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 remind the reader that AARC (and implicitely 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 Section 7.

Finally, we compare the computational time of LRC, MLRC, and ELAARC as it is illustrated in Table 5. From this table, one can observe that the computational time of MLRC is not significantly heavier than that of the other models, and it can be solved in less than 23 minutes using CPLEX solver for medium-sized instances. We also remark that larger sized instances of MLRC might be more effectively addressed using a similar decomposition scheme as presented in Ardestani-Jaafari and Delage (2017).

6.4 Multi-product assembly problem

In this problem, 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 $i \in \mathcal{I}$ 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 $j \in \mathcal{J}$. The robust multi-product assembly problem can be formulated as follows:

$$
\begin{aligned}
\underset{x,y(\zeta)}{\text{maximize}} \quad & \min_{\zeta\in\mathcal{U}} \; -c^T x + (q-l)^T y(\zeta) + s^T(x - Ay(\zeta)) && \text{(41a)}\\
\text{subject to} \quad & y(\zeta) \le \zeta \,, \forall \zeta \in \mathcal{U} && \text{(41b)}\\
& Ay(\zeta) \le x \,, \forall \zeta \in \mathcal{U} && \text{(41c)}\\
& y(\zeta) \ge 0 \,, \forall \zeta \in \mathcal{U} && \text{(41d)}\\
& 0 \le x \le M \,, && \text{(41e)}
\end{aligned}
$$

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

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

where $z^1 : \mathbb{R}^n \to \mathbb{R}^n$, $z^2 : \mathbb{R}^n \to \mathbb{R}^m$, and $z^3 : \mathbb{R}^n \to \mathbb{R}^n$ can be interpreted as violation adjustments for constraints (41b), (41c), and (41d). Yet, in this case, the u bounds are obtained from the dual problem:

$$\begin{aligned}
\underset{\lambda^1,\lambda^2,\lambda^3}{\text{minimize}} \quad & \zeta^T\lambda^1 + x^T\lambda^2 & \text{(42a)} \\
\text{subject to} \quad & \lambda^1_j + \sum_i A_{ij}\lambda^2_i - \lambda^3_j = q_j - l_j + A^T_{:j}s\,,\ \forall\, j \in \mathcal{J} & \text{(42b)} \\
& \lambda^1 \geq 0\,,\ \lambda^2 \geq 0\,,\ \lambda^3 \geq 0\,, & \text{(42c)}
\end{aligned}$$

where $\lambda^1 \in \mathbb{R}^n$, $\lambda^2 \in \mathbb{R}^m$, and $\lambda^3 \in \mathbb{R}^n$ are the dual variables associated to constraints (41b), (41c), and (41d). Here again, the objective function is non-decreasing in λ^1 and λ^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:

$$\lambda^1_j + \sum_i A_{ij}\lambda^2_i \geq q_j - l_j + A^T_{:j}s\,,\ \forall\, j \in \mathcal{J}\ .$$

This indicates to us that

$$\lambda^{1*}_j \leq \max(0\,,\ q_j - l_j + A^T_{:j}s - \sum_i A_{ij}\lambda^{2*}_i) \leq \max(0\,,\ q_j - l_j + A^T_{:j}s) := u^1_j\,,$$

and that

$$\lambda^{2*}_i \leq \max(0\,,\ \max_{j\in\mathcal{J}_i}\frac{1}{A_{ij}}(q_j - l_j + A^T_{:j}s - \lambda^1_j - \sum_{i'\neq i} A_{ij}\lambda^{2*}_i)) \leq \max(0\,,\ \max_{j\in\mathcal{J}_i}\frac{1}{A_{ij}}(q_j - l_j + A^T_{:j}s)) := u^2_i\,,$$

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

$$\lambda_j^3 = \lambda_j^1 + \sum_i A_{ij}\lambda_i^2 - q_j + l_j - A_{:j}^T s \leq u_j^1 + \sum_i A_{ij}u_i^2 - q_j + l_j - A_{:j}^T s := u_j^3 \,.$$

We conclude this example with a description of the specific context in which exploiting the information about the bound u on λ^* 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 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 $u^1 := [\ 425 \quad 805 \quad 1240\ ]^T$, $u^2 := [\ 140 \quad 310\ ]^T$, and $u^3 := [\ 4050 \quad 1550 \quad 2500\ ]^T$. However, using problem (19), with $M := 4050$, allows us to tighten these bounding vectors even more:

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

As it is shown in Table 6, when the budget of uncertainty is set to $\Gamma = 2$, a direct application of affine decision rules in problem (41) 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.

7 Extensive Numerical Study of Location-transportation Problem

In this section, we perform an extensive comparison of the quality of the solutions that are provided for the location-transportation problem of Section 6.3 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 particular, we generated 1000 instances of a problem containing 5 facility and 10 customer locations, where the nominal demand and maximum deviation are respectively equal to $\bar{\zeta}_j = 20,000$ and $\hat{\zeta}_j = 0.9\bar{\zeta}_j = 18,000$ for all j while the variable and fixed capacity costs are respectively $c = 0.6$ and $K = 50,000$. 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 $[4, 7]$.

In discussing our finding, we will make use the following expressions which are borrowed from Ardestani-Jaafari and Delage (2017):

- The "optimized worst-case bound" of a conservative approximation model refers to the best lower bound on worst-case profit that can be achieved according to this model (i.e., optimal value of the conservative approximation model).
- The "optimal worst-case profit" of a problem instance refers to the best worst-case profit that can be achieved for this instance.[3]
- The "relative optimized bound gap" of a conservative approximation model refers to the relative difference between the optimal worst-case profit of the problem instance and the optimized worst-case bound of this model.
- The "relative suboptimality gap" of the solution of a conservative approximation model refers to the relative difference between the optimal worst-case profit 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.

Regarding the quality of the optimized worst-case bound, one might first observe in Table 7 that the average optimized bounds always strictly improves 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 ELAARC and MLRC models are able to improve the average relative optimized bound gap obtained with AARC by 0.19% and 10.10% respectively. Furthermore, it appears that a significant gain is achieved with the introduction of MLRC, such that the proposed relative optimized bound gap are on average less than 1.67% from being exact. Finally, one might notice that, this reduction becomes more significant as one increases the required

[3]Implementation detail: The optimal value of the ARO model is obtained using the column-and-constraint generation algorithm presented in Zeng and Zhao (2013).

level of robustness, Γ, as long as $\Gamma < 10$. A similar observation can be made for the ELAARC model for which the difference with AARC is barely noticeable for small values of Γ. While Ardestani-Jaafari and Delage (2017) had noticed that marginal improvement in performance could be obtained by employing affine decision rules in a penalized form of the recourse problem (leading to the ELAARC model), this new study clearly illustrates how a more thorough application of this scheme can be beneficial when the amount of uncertainty is large.

Table 8 provides additional statistics about the relative suboptimality gap of the different solutions proposed by each approximation model in the 10,000 problem instances surveyed in Table 7 (i.e., 1000 randomly generated instances evaluated for all $\Gamma = 1, \ldots, 10$). 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 could be repeated here, the main one in this table might be how the added flexibility employed in the MLRC truly gives this approximation scheme significantly better chances of a high-quality solution in terms of relative suboptimality gap. Numerically speaking, this did not appear to come at a price in terms of solution time (resolution time of all three models were comparable) although this would need to be verified in the context of larger scale problems.

8 Implications for Copositive Programming Reformulations

Since the publication of the thesis of Ardestani-Jaafari (2016) on which this article is based, a few studies have shed some additional light on 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 sections 6.3 and 6.4), 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 λ 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 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 4. 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 ζ 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 $\zeta := \zeta^+ - \zeta^-$ with $\zeta^+ \geq 0$ and $\zeta^- \geq 0$.*

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

$$\begin{aligned}
g(x) = \quad & \min_{\zeta,\lambda} && c^T x + (\Psi(x)\zeta)^T\lambda - (Ax)^T\lambda && \text{(43a)}\\
& \text{subject to} && B^T\lambda = d && \text{(43b)}\\
& && P\zeta \leq q && \text{(43c)}\\
& && 0 \leq \lambda \leq u\,. && \text{(43d)}
\end{aligned}$$

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

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

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

$$\tilde{Q}(x) := \begin{bmatrix} 0 & (1/2)\Psi(x) & 0 \\ (1/2)\Psi(x)^T & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \tilde{c}(x) := \begin{bmatrix} -(1/2)Ax \\ 0 \\ 0 \end{bmatrix} \quad \tilde{A} := \begin{bmatrix} B^T & 0 & 0 & 0 \\ 0 & P & I & 0 \\ I & 0 & 0 & I \end{bmatrix} \quad \tilde{b} := \begin{bmatrix} d \\ q \\ u \end{bmatrix},$$

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

$$
\begin{aligned}
g(x) = \quad & \min_{Y,y} && c^T x + \mathrm{tr}(\tilde{Q}(x)^T Y) + 2\tilde{c}(x)^T y && (44a)\\
& \text{subject to} && \tilde{A}y = \tilde{b} && (44b)\\
& && \tilde{A}Y = \tilde{b}y^T && (44c)\\
& && \begin{bmatrix} Y & y \\ y^T & 1 \end{bmatrix} \in \mathcal{K}_{\mathrm{CP}}\,, && (44d)
\end{aligned}
$$

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

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

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

$$
\begin{aligned}
g(x) \geq \quad & \max_{W,w,t} && c^T x + \tilde{b}^T w - t && (45a)\\
& \text{subject to} && \begin{bmatrix} \tilde{Q}(x) - (1/2)(W^T\tilde{A} + \tilde{A}^T W) & \tilde{c}(x) - (1/2)(\tilde{A}^T w - W^T\tilde{b}) \\ \tilde{c}(x)^T - (1/2)(\tilde{A}^T w - W^T\tilde{b})^T & t \end{bmatrix} \in \mathcal{K}_{\mathrm{Cop}}\,, && (45b)
\end{aligned}
$$

where $t \in \mathbb{R}$, and where $w \in \mathbb{R}^{n_y+n_{\mathcal{U}}+m}$ and $W \in \mathbb{R}^{n_y+n_{\mathcal{U}}+m\times\tilde{n}}$ contain the dual variables associated to constraints (44b) and (44c) respectively, while $\mathcal{K}_{\mathrm{Cop}}$ refers to the dual cone of $\mathcal{K}_{\mathrm{CP}}$ also known as the cone of copositive matrices, i.e.

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

When attempting to prove that strong duality holds, a sufficient step consists in verifying whether problem (45) is strictly feasible. We refer the reader to Appendix C for a complete proof of this claim.

Lemma 3. *Given assumptions 4 and 6, problem* (45) *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 2. *Given assumptions 4 and 6, the following copositive program is equivalent to problem* (1)*:*

$$\underset{x\in\mathcal{X},W,w,t}{\text{maximize}} \quad c^T x + \tilde{b}^T w - t \tag{46a}$$

$$\text{subject to} \quad \begin{bmatrix} \tilde{Q}(x) - (1/2)(W^T\tilde{A} + \tilde{A}^T W) & \tilde{c}(x) - (1/2)(\tilde{A}^T w - W^T\tilde{b}) \\ \tilde{c}(x)^T - (1/2)(\tilde{A}^T w - W^T\tilde{b})^T & t \end{bmatrix} \in \mathcal{K}_{Cop}. \tag{46b}$$

Proof. Proof. Strong duality follows from the fact that the dual problem (45) is strictly feasible, following Lemma 3, and bounded, which follows easily from Assumption 3 since it states that $\max_{x\in\mathcal{X}} g(x)$ is bounded. □ □

Corollary 3. *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* (46) *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* (20)*;*
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* $\max_{x\in\mathcal{X}} g_{\textit{SDP-LRC}}(x)$.
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* $\max_{x\in\mathcal{X}} g_{\textit{SDP-LRC}}(x)$.

See Appendix C for proof.

Compared to Hanasusanto and Kuhn (2016), the particularity of this non-convex quadratic program (43) 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 $\|\lambda\|_2$ 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 $\|\lambda\|_2$ 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.

9 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 optimization problem that needs to 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 surveyed the types of improvement that our proposed models might offer in four different operations management applications and shed some light on how this scheme relates to approaches that are based on copositive programming reformulations as were presented in Hanasusanto and Kuhn (2016) and Xu and Burer (2016).

References

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 locationtransportation 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.

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

Benson, Steven J, Yinyu Ye, Xiong Zhang. 2000. Solving large-scale sparse semidefinite programs for combinatorial optimization. *SIAM Journal on Optimization* **10**(2) 443–461.

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.

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. *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.

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.

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. Ph.D. thesis, California Institute of Technology.

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* .

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.

A Relation to AARC for General Uncertainty Sets

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

Proposition 8. *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 $\mathcal{U}_{general}$:*

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

where

$$
\begin{aligned}
g_{GLRC}(x) := \quad & \min_{\zeta,\lambda,\Delta} && c^T x + \mathrm{tr}(\Psi(x)\Delta) - (Ax)^T \lambda && (47a)\\
& \text{subject to} && B^T\lambda = d && (47b)\\
& && f(\zeta) \le q && (47c)\\
& && \lambda \ge 0 && (47d)\\
& && \Delta B = \zeta d^T && (47e)\\
& && h(\Delta_{:i}, \lambda_i) \le q\lambda_i \,,\ \forall\, i\,. && (47f)
\end{aligned}
$$

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

$$\sup_z \Delta_{:i}^T z - \lambda_i f_*(z) \le q\lambda_i \,,\ \forall\, i\,.$$

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

$$
\begin{aligned}
\mathcal{L}(\zeta,\lambda,\Delta,y,Y,s) &:= c^T x + \mathrm{tr}(\Psi(x)\Delta) - (Ax)^T\lambda + y^T(d - B^T\lambda) + \mathrm{tr}(Y(\zeta d^T - \Delta B)\\
&\qquad + \sum_i s_i(\sup_z \Delta_{:i}^T z - \lambda_i f_*(z) - q\lambda_i)\\
&= \sup_{z_1,\dots,z_m} c^T x + \mathrm{tr}(\Psi(x)\Delta) - (Ax)^T\lambda + y^T(d - B^T\lambda) + \mathrm{tr}(Y(\zeta d^T - \Delta B)\\
&\qquad + \sum_i s_i \Delta_{:i}^T z_i - \lambda_i s_i f_*(z_i) - q s_i \lambda_i)\,,
\end{aligned}
$$

where $y \in \mathbb{R}^{n_y}$, $Y \in \mathbb{R}^{n_y \times n_\zeta}$, and $s \in \mathbb{R}^m$ are respectively the dual variables associated to constraints (47b), (47e), and (47f). Now letting $\mathcal{L}(\zeta,\lambda,\Delta,y,Y,s,\{z_i\}_{i=1}^m)$ denote the expression on the right of the $\sup_{z_1,\dots,z_m}$ operator, we necessarily have that

$$
\begin{aligned}
g_{GLRC}(x) \quad &= \inf_{\substack{\zeta\,:\,f(\zeta)\le q\\ \Delta,\lambda\,:\,\lambda\ge 0}} \ \sup_{y,Y,s\ge 0,\{z_i\}_{i=1}^m} \mathcal{L}(\zeta,\lambda,\Delta,y,Y,s,\{z_i\}_{i=1}^m) && (48a)\\
&\ge \inf_{\zeta: f(\zeta)\le q} \ \sup_{y,Y,s\ge 0,\{z_i\}_{i=1}^m} \ \inf_{\Delta,\lambda\ge 0} \mathcal{L}(\zeta,\lambda,\Delta,y,Y,s,\{z_i\}_{i=1}^m)\,. && (48b)
\end{aligned}
$$

One can then analytically resolve the optimum in terms of λ and Δ as

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

The equality constraint can further be used in conjunction with the fact that $s_i f_*(z_i) = h_*(s_i z_i, s_i) := \sup_y s_i z_i^T y - s_i f(y)$, to obtain

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

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

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

where the last constraint can be reformulated as

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

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

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

and finally

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

These steps allow us to reach our conclusion:

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

which is the conservative approximation of the worst-case performance for 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. $\square$ $\square$

B Counterargument to Proposition 6 in Denton et al. (2010)

We first recall the proposition that can be found in Denton et al. (2010).

Proposition 9. (Proposition 6 from Denton et al. (2010)) *The polyhedron defined by the following constraints has integer extreme points when τ is an integer*

$$\sum_{ij} \Delta_{ij} \leq \tau \tag{49a}$$

$$0 \leq \Delta_{ij} \leq Y_{ij} z_j \,,\ \forall\, i \in \mathcal{I} \,,\ \forall\, j \in \mathcal{J} \tag{49b}$$

$$0 \leq z_j \leq 1 \,,\ \forall\, j \in \mathcal{J} \,, \tag{49c}$$

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

Here is a counter-proposition.

Proposition 10. *Let $n = 3$, $m = 2$, $Y_{12} = Y_{21} = Y_{31} = 1$, and $\Gamma = 2$, then the polyhedron defined by equations* (49a), (49b), *and* (49c) *has the following extreme point:*

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

Proof. 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 $(\bar{\Delta}, \bar{z})$ satisfies this system of equations, we have confirmed that this assignment describes an extreme point of the polyhedron. $\square$

C Proofs of Section 8

Proof. Proof of Lemma 3. We will exploit Farkas lemma to identify a ray $(\gamma\bar{t}, \gamma\bar{w}, \gamma\bar{W})$, parametrized by $\gamma > 0$, such that

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

for some

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

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

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

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

$$s \geq 0 \qquad s_0 \geq 0 \qquad P^T s \geq s_0 \qquad q^T s < M s_0\,.$$

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 $\hat{\zeta} \geq 0$ leads to a contradiction:

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

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

Second, we demonstrate that

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

for the following assignment:

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

Studying each term separately we get:

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

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

$$\geq\frac{1}{2}((2+\max(\bar{s}^Tq+1^Tu\ ,\ 0))-q^T\bar{s}-1^Tu)\geq 1$$

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

This completes our proof. □ □

Proof. Proof of Corollary 3. Given that we have established that strong duality applies for problem (45) 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 $g_{\text{MLRC}}(x)$ and $g_{\text{SDP-LRC}}(x)$ to the optimal value of problem (44) 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 $g_1(x)$ and $g_2(x)$ respectively.

First, whether the cone is replaced with $\mathcal{K}_1$ or $\mathcal{K}_2$, we can exploit the equalities of problem (44) to reformulate $g_1(x)$ and $g_2(x)$ as

$$g_i(x)=\min_{Y,y,\Lambda,\Xi,\lambda,\xi}\quad c^Tx+\text{tr}(\tilde{Q}(x)^TY)+2\tilde{c}(x)^Ty \tag{50a}$$

$$\text{subject to}\quad (44\text{b})-(44\text{c})$$

$$\begin{bmatrix} Y & y \\ y^T & 1 \end{bmatrix}=\Phi\begin{bmatrix} \Lambda & \Delta^T & \lambda \\ \Delta & \Xi & \zeta \\ \lambda^T & \zeta^T & 1 \end{bmatrix}\Phi^T \tag{50b}$$

$$\begin{bmatrix} Y & y \\ y^T & 1 \end{bmatrix}\in\mathcal{K}_i\,, \tag{50c}$$

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

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

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 (50b) to see that each one of them simply repeats an equality constraint present in constraints (44b) and (44c).

Next, in problem (24), one can certainly consider additional decision variables $\Lambda \in \mathbb{R}^{m\times m}$ and $\Xi \in \mathbb{R}^{n_\zeta \times n_\zeta}$ that need to satisfy constraints (33j), (33k), and (33l), without affecting the feasible set in terms of (ζ, λ, Δ) given that the assignment $\Lambda := \lambda\lambda^T$ and $\Xi = \zeta\zeta^T$ always satisfies these constraints. Furthermore, in both problems (24) and (33) 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 $Q \in \mathbb{R}^{m\times m}$ such that:

$$Q \geq 0\,,\; QP = -I\,,\; Qq \leq 0\,.$$

We can therefore derive the following implications:

$$\begin{aligned}
&(10c) \;\Rightarrow\; P\Delta \leq q\lambda^T \;\Rightarrow\; QP\Delta \leq Qq\lambda^T \;\Rightarrow\; \Delta \geq 0 \\
&(24c) \;\Rightarrow\; P\Delta \geq q\lambda^T - (q - P\zeta)u^T \;\Rightarrow\; QP\Delta \geq Qq\lambda^T - Q(q - P\zeta)u^T \;\Rightarrow\; \Delta \leq \zeta u^T \\
&(33j) \;\Rightarrow\; P\Xi P^T + qq^T \geq P\zeta q^T + q\zeta^T P^T \;\Rightarrow\; QP\Xi P^T \geq Qq(P\zeta - q)^T + QP\zeta q^T \;\Rightarrow\; \Xi P^T \leq \zeta q^T \\
&\qquad\;\Rightarrow\; \Xi P^T \leq \zeta q^T \;\Rightarrow\; Q\Xi P^T \leq Q\zeta q^T \;\Rightarrow\; \Xi \geq 0\,.
\end{aligned}$$

Hence, one can establish that

$$\begin{aligned}
g_{\mathrm{MLRC}}(x) = \quad & \min_{Y,y,\Lambda,\Xi,\lambda,\xi} && c^T x + \mathrm{tr}(\tilde{Q}(x)^T Y) + 2\tilde{c}(x)^T y \\
& \text{subject to} && (44b) - (44c), (50b) \\
& && \begin{bmatrix} Y & y \\ y^T & 1 \end{bmatrix} \succeq 0\,.
\end{aligned}$$

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

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

$$\begin{aligned} g_{\text{SDP-LRC}}(x) = \quad & \min_{Y,y,\Lambda,\Xi,\lambda,\xi} && c^T x + \mathrm{tr}(\tilde{Q}(x)^T Y) + 2\tilde{c}(x)^T y \\ & \text{subject to} && (44b) - (44c), (50b) \\ & && \begin{bmatrix} Y & y \\ y^T & 1 \end{bmatrix} \geq 0 \,, \; \begin{bmatrix} \Lambda & \Delta^T & \lambda \\ \Delta & \Xi & \zeta \\ \lambda^T & \zeta^T & 1 \end{bmatrix} \succeq 0 \,. \end{aligned}$$

Although this does not exactly give rise to the optimization problem associated to $g_2(x)$ because of the different linear matrix inequalities, they are both equivalent because of constraint (50b) and the fact that Φ is full rank. Namely, it is clear that

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

This implies that $g_2(x) = g_{\text{SDP-LRC}}(x)$. 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 (46) 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. $\square$ $\square$

Table 2: Computation time and quality of bounds obtained from AARC/(M)LRC and SDP-LRC compared to SDP-A&D†

		Computational time (in sec)			Improvement (in %)	
	Γ%	AARC/(M)LRC	SDP-LRC	SDP-A&D†	SDP-LRC	SDP-A&D
5	10	<1	1	1	0.00	0.00
	30	<1	1	1	0.07	0.02
	50	<1	1	1	0.19	0.07
	70	<1	1	1	0.14	0.01
	90	<1	1	1	0.00	0.00
10	10	<1	5	1	1.79	0.00
	30	<1	5	1	14.18	0.04
	50	<1	6	1	18.65	0.03
	70	<1	6	1	43.01	0.01
	90	<1	6	1	47.12	0.00
20	10	<1	61	6	3.52	0.01
	30	<1	61	7	0.82	0.02
	50	<1	60	7	21.57	0.02
	70	1	66	6	20.91	0.01
	90	1	70	7	35.10	0.00
30	10	3	360	27	4.63	0.03
	30	3	361	30	18.40	0.03
	50	2	356	27	10.63	0.04
	70	2	408	26	15.72	0.02
	90	3	368	31	71.84	0.01
40	10	3	1209	66	3.71	0.19
	30	2	1260	76	8.02	0.64
	50	3	1214	78	15.28	0.04
	70	2	1244	69	44.24	0.03
	90	2	1192	85	81.78	0.01

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

Table 3: Comparison of the worst-case cost for different solution methods to the surgery block allocation problem

Alternative	AARC's/(M)LRC's	SDP-LRC's	Exact model	RORA[†]
Open one OR	\$822,000	\$822,000	\$822,000	\$822,000
Open two ORs	\$812,000	\$812,000	\$812,000	\$828,000

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

Table 4: Comparison of optimal bound on worst-case profit and worst-case profit associated to solutions of each model in an instance of RLTP

	AARC/LRC	MLRC	SDP-LRC	Exact model	ELAARC[†]
Optimal bound on worst-case profit	0	6600	6600	6600	4024
Worst-case profit of solution	0	6600	6600	6600	4622

[†]ELAARC refers to the approximation model proposed in Ardestani-Jaafari and Delage (2017).

Table 5: Computational time of LRC, MLRC and the available approximation in the literature

$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{J}\rvert$	Γ	AARC/LRC	MLRC	ELAARC†
10	20	10%	3	7	5
		30%	3	11	5
		50%	4	10	7
		70%	9	15	10
		90%	6	14	8
20	20	10%	18	67	31
		30%	31	78	25
		50%	44	110	41
		70%	70	215	70
		90%	61	274	102
20	40	10%	489	891	775
		30%	488	1074	391
		50%	690	452	750
		70%	1027	1339	616
		90%	457	556	356

†ELAARC refers to the approximation model proposed in Ardestani-Jaafari and Delage (2017) solved directly using CPLEX.Is this the constraint generation algorithm or one-shot reformulation. This should be specified.

Table 6: Optimal solution of AARC and MLRC in the instance of multi-product assembly problem

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

Table 7: Average relative optimized bound gap (Bound gap) and average relative suboptimality gap (Subopt. 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 $\Gamma = 1$ and $\Gamma = 10$ in these problem instances.

Γ	AARC		MLRC		ELAARC†	
	Bound gap	Subopt. gap	Bound gap	Subopt. gap	Bound gap	Subopt. gap
1	0	0	0	0	0	0
2	0.88	0.66	0.25	0.20	0.88	0.66
3	2.26	2.03	0.75	0.59	2.25	2.02
4	4.18	3.99	1.15	0.93	4.15	3.96
5	7.03	6.87	1.48	1.19	6.98	6.82
6	11.12	10.99	1.45	1.19	11.01	10.86
7	17.63	17.54	1.65	1.48	17.36	17.26
8	28.31	28.29	2.97	2.92	27.73	27.45
9	46.25	34.18	7.01	6.99	45.45	21.17
10	0	0	0	0	0	0
Avg.	11.77	10.46	1.67	1.54	11.58	9.02

†ELAARC refers to the approximation model proposed in Ardestani-Jaafari and Delage (2017).

Table 8: Proportion of the 10,000 problem instances analysed in Table 7 for which the relative sub-optimality gap of each approximation model was within a certain range. Average gap and maximum gap are also reported.

Gap range	AARC/LRC	MLRC	ELAARC
=0	20.6	23.8	20.6
≤0.1	20.9	27.4	20.9
≤1	28.4	56.3	28.4
≤10	61.4	99.3	62.3
Avg.	10.5	1.6	9.0
Max.	50.0	13.3	46.3

†ELAARC refers to the approximation model proposed in Ardestani-Jaafari and Delage (2017).