Online Appendices for : "Linearized Robust Counterparts of Two-stage Robust Optimization Problems with Applications in Operations Management"

Amir Ardestani-Jaafari and Erick Delage

Appendices

A. Additional details regarding Example 1

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

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

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

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

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

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

Table 10 Detailed optimal solution of problem (3)

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

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

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

B. Multi-product assembly problem

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

$$\underset{\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{\zeta})}{\text{maximize}} \quad \min_{\boldsymbol{\zeta} \in \mathcal{U}} \; -\boldsymbol{c}^\top \boldsymbol{x} + (\boldsymbol{q} - \boldsymbol{l})^\top \boldsymbol{y}(\boldsymbol{\zeta}) + \boldsymbol{s}^\top (\boldsymbol{x} - A\boldsymbol{y}(\boldsymbol{\zeta})) \tag{45a}$$

$$\text{subject to} \quad \boldsymbol{y}(\boldsymbol{\zeta}) \leq \boldsymbol{\zeta}, \forall \boldsymbol{\zeta} \in \mathcal{U} \tag{45b}$$

$$A\boldsymbol{y}(\boldsymbol{\zeta}) \leq \boldsymbol{x}, \forall \boldsymbol{\zeta} \in \mathcal{U} \tag{45c}$$

$$\boldsymbol{y}(\boldsymbol{\zeta}) \geq 0, \forall \boldsymbol{\zeta} \in \mathcal{U} \tag{45d}$$

$$0 \leq \boldsymbol{x} \leq M, \tag{45e}$$

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

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

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

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

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

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

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

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

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

This indicates to us that

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

and that

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

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

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

We conclude this example with a description of the specific context in which exploiting the information about the bound $\boldsymbol{u}$ on $\boldsymbol{\lambda}^*$ leads to a strictly tighter conservative approximation. In particular, consider a multi-product assembly problem with three products and two different types of parts. The pre-order variable $\boldsymbol{x}$ is bounded by 100,000, the cost of parts A and B are, respectively, \$25 per unit and \$3 per unit, while the salvage value is \$4 per unit and \$1 per unit. Furthermore, the difference between the selling price and the unit production cost of each product is: \$380/unit, \$800/unit, and \$1200/unit respectively for products #1 to #3. Next, we have that product #1 requires 9 units of both parts, product #2 requires 5 units of part B, and #3 requires 9 units of A and 4 units of B. Finally, for products #1 to #3, the nominal demand is respectively of 9000, 10,000, and 8000 units while the worst-case demand for each is 1000, 2000, and 0 units respectively. In this specific context, one can exploit the above closed-form bounds $\boldsymbol{u}^1 := [\,425\ 805\ 1240\,]^\top$, $\boldsymbol{u}^2 := [\,138\ 310\,]^\top$, and $\boldsymbol{u}^3 := [\,4032\ 1550\ 2482\,]^\top$. However, using problem (17), with $M := 4050$, allows us to tighten these bounding vectors even more:

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

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

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

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

C. Relation to AARC for General Uncertainty Sets

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

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

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

where

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

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

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

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

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

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

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

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

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

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

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

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

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

where the last constraint can be reformulated as

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

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

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

and finally

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

These steps allow us to reach our conclusion:

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

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

D. Improved tractable approximations for surgery block allocation problems

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

$$\underset{\boldsymbol{x},Z,\boldsymbol{y}(\boldsymbol{\zeta})}{\text{minimize}} \quad \max_{\boldsymbol{\zeta}\in\mathcal{U}} \; c\sum_{i\in\mathcal{I}} x_i + d\sum_{i\in\mathcal{I}} y_i(\boldsymbol{\zeta}) \tag{49a}$$

$$\text{subject to} \quad y_i(\boldsymbol{\zeta}) \geq \sum_{j\in\mathcal{J}} \zeta_j Z_{ij} - w x_i \,,\; \forall i \in \mathcal{I} \,,\; \forall \boldsymbol{\zeta} \in \mathcal{U} \tag{49b}$$

$$\boldsymbol{y}(\boldsymbol{\zeta}) \geq 0 \,,\; \forall \boldsymbol{\zeta} \in \mathcal{U} \tag{49c}$$

$$\sum_{i\in\mathcal{I}} Z_{ij} = 1 \,,\; \forall j \in \mathcal{J} \tag{49d}$$

$$Z_{ij} \leq x_i \,,\; \forall i \in \mathcal{I} \,,\; \forall j \in \mathcal{J} \tag{49e}$$

$$\boldsymbol{x} \in \{0,1\}^m \,,\; Z \in \{0,1\}^{m\times n} \,, \tag{49f}$$

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

In particular, we consider a particular problem instance in which there are three surgery blocks and 2 operating rooms that can run for 8 hours. The cost of opening a room is \$390,000, and the overtime cost is \$1,000 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. The session length is 8 hours. We finally set the budget to $\Gamma = 2$. In this context, one can show that the model proposed by Denton will suggest opening only one OR, where all blocks will be scheduled for an estimated worst-case total cost of \$822,000. On the other hand, one can verify that opening both ORs and scheduling the biggest block in one OR and the two smaller ones in the second OR leads to a worst-case total cost of \$812,000. Note that the worst-case total cost of this solution is estimated at \$828,000 by the Denton model. One can further confirm that the exact optimal solution is the one that is returned by the AARC (and LRC) model. Table 13 summarizes the optimal bounds on worst-case cost obtained for the two types of solutions (*i.e.,* open one or two ORs) using RORA, AARC, and an exact approach.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Putting all these together we get

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

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

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

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

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

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

where

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

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

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

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

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

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

where

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

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

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

Hence, we have that

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

$$\sum_{j\in\mathcal{J}}\delta_j \le \Gamma$$
$$0\le \Delta_{ij}\le \delta_j\,,\,\forall i\in\mathcal{I}\,,\,\forall j\in\mathcal{J}$$
$$\Delta_{ij}\le z_i\,,\,\forall i\,,\,\forall j\in\mathcal{J}$$
$$\sum_{j\in\mathcal{J}}\Delta_{ij}\le \Gamma z_i\,,\,\forall i\in\mathcal{I}$$
$$1-\delta_j-\lambda_i+\Delta_{ij}\ge 0\,,\,\forall i\in\mathcal{I}\,,\,\forall j\in\mathcal{J}$$
$$\sum_{j\in\mathcal{J}}\delta_j-\sum_{j\in\mathcal{J}}\Delta_{ij}\le \Gamma(1-\lambda_i)\,,\,\forall i\in\mathcal{I}$$
$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\Delta_{ij}\le\Gamma\,,$$

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

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

E. Extension of the Decomposition Algorithm proposed in Ardestani-Jaafari and Delage (2018)

We extend the decomposition algorithm of Ardestani-Jaafari and Delage (2018) such that it can be applied for all proposed models in this paper: *i.e.,* MLRC, SDP-LRC, SDP-LRC2, GLRC, and GSDP-LRC2. We first define the following sets:

$$\mathcal{S}_{\text{MLRC}} := \{(\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta)\in\mathbb{R}^{n_\zeta}\times\mathbb{R}^{\bar{m}}\times\mathbb{R}^{n_\zeta\times\bar{m}}\,|\,(6\text{b})-(8\text{c}),(19),(21)\}$$
$$\mathcal{S}_{\text{SDP-LRC}} := \left\{(\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta)\in\mathbb{R}^{n_\zeta}\times\mathbb{R}^{\bar{m}}\times\mathbb{R}^{n_\zeta\times\bar{m}}\;\middle|\;\begin{array}{l}\exists(\Xi,\Lambda)\in\mathbb{R}^{n_\zeta\times n_\zeta}\times\mathbb{R}^{n_{\bar{m}}\times n_{\bar{m}}},\\(6\text{b})-(6\text{d}),(8\text{b})-(8\text{f}),(19),(21),\\(34\text{b}),(34\text{c}),(34\text{d})\end{array}\right\}$$
$$\mathcal{S}_{\text{SDP-LRC2}} := \left\{(\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta)\in\mathbb{R}^{n_\zeta}\times\mathbb{R}^{\bar{m}}\times\mathbb{R}^{n_\zeta\times\bar{m}}\;\middle|\;\begin{array}{l}\exists(\Xi,\Lambda)\in\mathbb{R}^{n_\zeta\times n_\zeta}\times\mathbb{R}^{n_{\bar{m}}\times n_{\bar{m}}},\\(6\text{b})-(6\text{d}),(8\text{b})-(8\text{f}),(19),(21),\\(34\text{b}),(34\text{c}),(35\text{b})\end{array}\right\}$$
$$\mathcal{S}_{\text{GLRC}} := \left\{(\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta)\in\mathbb{R}^{n_\zeta}\times\mathbb{R}^{\bar{m}}\times\mathbb{R}^{n_\zeta\times\bar{m}}\;\middle|\;\begin{array}{l}\exists(\Xi,\Lambda)\in\mathbb{R}^{n_\zeta\times n_\zeta}\times\mathbb{R}^{n_{\bar{m}}\times n_{\bar{m}}},\\(6\text{b})-(6\text{d}),(8\text{b}),(8\text{c}),(19),(21),\\(38\text{b}),(38\text{c}),(38\text{d}),(38\text{e})\end{array}\right\}$$
$$\mathcal{S}_{\text{GSDP-LRC2}} := \left\{(\boldsymbol{\zeta},\boldsymbol{\lambda},\Delta)\in\mathbb{R}^{n_\zeta}\times\mathbb{R}^{\bar{m}}\times\mathbb{R}^{n_\zeta\times\bar{m}}\;\middle|\;\begin{array}{l}\exists(\Xi,\Lambda)\in\mathbb{R}^{n_\zeta\times n_\zeta}\times\mathbb{R}^{n_{\bar{m}}\times n_{\bar{m}}},\\(6\text{b})-(6\text{d}),(8\text{b})-(8\text{f}),(19),(21),\\(34\text{b}),(34\text{c}),(38\text{b})-(38\text{e}),(39\text{b})\end{array}\right\}.$$

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

Row generation algorithm

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

Step #2: Solve the following subproblem

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

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

Step #3: Let $\kappa := \kappa + 1$ and solve the following master problem:

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

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

Let $\dot{\boldsymbol{x}}^{(\kappa)}$ and UB take on the values of any optimal solution and optimal value, respectively, of the MP.

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

Similar to Ardestani-Jaafari and Delage (2018), one can propose a set of valid inequalities to the MP model in the row generation algorithm. In particular, for all models it is possible to replace the MP with an enhanced version that exploits a scenario-based relaxation:

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

where $\mathcal{U}^\kappa \subset \mathcal{U}$ is any finite set of scenarios for $\boldsymbol{\zeta}$, e.g. the worst-case scenarios identified when solving the SP in the $\kappa - 1$ first iterations. Alternatively, or in addition, one can employ similar cuts as were used in Ardestani-Jaafari and Delage (2018) that exploits the specific structure of the approximation model. For example, in the case of MLRC we get

$$\begin{aligned}
\text{(MLRC-MP')} \quad & \underset{\boldsymbol{x}\in\mathcal{X},\rho,\boldsymbol{y},Y,\boldsymbol{z},Z}{\text{maximize}} \quad \rho \\
& \text{subject to} \quad (51b)
\end{aligned}$$

$$\rho \leq \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{d}^\top (\boldsymbol{y} + Y\boldsymbol{\zeta}) - \boldsymbol{u}^\top (\boldsymbol{z} + Z\boldsymbol{\zeta}), \forall \boldsymbol{\zeta} \in \mathcal{U}^\kappa$$
$$A\boldsymbol{x} + B(\boldsymbol{y} + Y\boldsymbol{\zeta}) \leq \Psi(\boldsymbol{x})\boldsymbol{\zeta} + \boldsymbol{z} + Z\boldsymbol{\zeta}, \forall \boldsymbol{\zeta} \in \mathcal{U}^\kappa$$
$$\boldsymbol{z} + Z\boldsymbol{\zeta} \geq 0, \forall \boldsymbol{\zeta} \in \mathcal{U}^\kappa,$$

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

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

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

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

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

with

$$P := \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \end{bmatrix}, \qquad \bar{\boldsymbol{\zeta}} := \begin{bmatrix} 80 \\ 80 \\ 60 \end{bmatrix}, \qquad \hat{\boldsymbol{\zeta}} := \begin{bmatrix} 60 \\ 60 \\ 40 \end{bmatrix},$$

where P models the property that the correlation between the demand of any two items is 0.5. We compare, as it is shown in Table 14, the optimal bound on worst-case profit and the achieved worst-case profit of solutions obtained from the LRC model, the SDP-LRC model, and the semi-definite programming model (denoted by SDP-A&D) proposed in Ardestani-Jaafari and Delage (2016). In this example, LRC is not exact and can actually be improved upon using models such that SDP-A&D and SDP-LRC. In particular, the bound on best achievable worst-case profit is increased by a factor of about 3 and 10 using SDP-A&D and SDP-LRC respectively. This translates directly in some improvement in performance of solutions of SDP-A&D and SDP-LRC which achieve a worst-case profit that are respectively near 4 and 16 times better than what is achieved by the solution of AARC. It is also clear that the SDP-LRC model is responsible for most of the improvement.

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

	AARC	SDP-A&D†	SDP-LRC	Exact model
Optimal bound on worst-case profit	41.83	113.01	411.08	825.83
Worst-case profit of solution	41.83	150.94	664.82	825.83

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

G. Implications for Copositive Programming Reformulations

In this section, we reuse the ideas of both Hanasusanto and Kuhn (2018) and Xu and Burer (2018) to strengthen the connections between the ARO model with relatively complete recourse, copositive programming, and both MLRC and SDP-LRC derived in Section 3. In order to improve readability, some of the proofs of the theorems that are presented are delayed to Appendix H.

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

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

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

$$
\begin{aligned}
g(\boldsymbol{x}) = \quad & \min_{\boldsymbol{\zeta},\boldsymbol{\lambda}} \; \boldsymbol{c}^\top \boldsymbol{x} + (\Psi(\boldsymbol{x})\boldsymbol{\zeta})^\top \boldsymbol{\lambda} - (A\boldsymbol{x})^\top \boldsymbol{\lambda} && \text{(52a)}\\
& \text{subject to} \;\; B^\top \boldsymbol{\lambda} = \boldsymbol{d} && \text{(52b)}\\
& \qquad P\boldsymbol{\zeta} \leq \boldsymbol{q} && \text{(52c)}\\
& \qquad \mathbf{0} \leq \boldsymbol{\lambda} \leq \boldsymbol{u}\,. && \text{(52d)}
\end{aligned}
$$

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

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

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

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

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

$$
\begin{aligned}
g(\boldsymbol{x}) = \quad \min_{Y,\boldsymbol{y}} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \mathrm{tr}(\tilde{Q}(\boldsymbol{x})^\top Y) + 2\tilde{\boldsymbol{c}}(\boldsymbol{x})^\top \boldsymbol{y} && (53a)\\
\text{subject to} \quad & \tilde{A}\boldsymbol{y} = \tilde{\boldsymbol{b}} && (53b)\\
& \tilde{A}Y = \tilde{\boldsymbol{b}}\boldsymbol{y}^\top && (53c)\\
& \begin{bmatrix} Y & \boldsymbol{y} \\ \boldsymbol{y}^\top & 1 \end{bmatrix} \in \mathcal{K}_{\mathrm{CP}}\,, && (53d)
\end{aligned}
$$

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.

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

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

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

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

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

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

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

At this point, we have assembled all the ingredient to present yet a second equivalent copositive programming formulation of ARO for relatively complete recourse problems (see $\overline{\mathrm{RLP}}$ model in Xu and Burer (2018) for the original equivalent model).

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

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

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

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

1. *if $\mathcal{K}=\mathcal{K}_1:=\mathbb{R}_+^{\tilde{n}+1\times\tilde{n}+1}$, then the conic program reduces to a linear program that is equivalent to MLRC* (18)*;*

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

3. *if $\mathcal{K}_1\subseteq\mathcal{K}\subseteq\mathcal{K}_{Cop}$ then the conic program provides a tighter approximation than MLRC*

4. *if $\mathcal{K}_2\subseteq\mathcal{K}\subseteq\mathcal{K}_{Cop}$ then the conic program provides a tighter approximation than $\max_{\boldsymbol{x}\in\mathcal{X}} g_{SDP\text{-}LRC}(\boldsymbol{x})$.*

Compared to Hanasusanto and Kuhn (2018), the particularity of this non-convex quadratic program (52) is that it has a bounded feasible space which can be exploited to establish strong duality even thought the ARO does not satisfy the complete recourse assumption. Alternatively, Xu and Burer (2018) did propose imposing a bound on $\|\boldsymbol{\lambda}\|_2$ in order to help with duality yet did not attempt to further connect the resulting model to the AARC approach. Furthermore, the reformulation that is obtained using $\|\boldsymbol{\lambda}\|_2$ cannot readily be approximated using linear programming.

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

H. Proofs of Appendix G

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

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

for some

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

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

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

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

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

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

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

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

Second, we demonstrate that

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

for the following assignment:

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

Studying each term separately we get:

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

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

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

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

This completes our proof. □

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

$\mathcal{K}_{\text{PSD}}^{\tilde{n}+1\times\tilde{n}+1}$ which is $\mathcal{K}_2 := \mathbb{R}_+^{\tilde{n}+1\times\tilde{n}+1}\cap\mathcal{K}_{\text{PSD}}^{\tilde{n}+1\times\tilde{n}+1}$ respectively. In what follows, we simply refer to the value of each of these two bounds as $g_1(\boldsymbol{x})$ and $g_2(\boldsymbol{x})$ respectively.

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

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

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

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

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

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

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

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

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

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

We can therefore derive the following implications:

$$(6\text{c}) \Rightarrow P\Delta \leq \boldsymbol{q}\boldsymbol{\lambda}^\top \Rightarrow QP\Delta \leq Q\boldsymbol{q}\boldsymbol{\lambda}^\top \Rightarrow \Delta \geq 0$$

$$(21) \Rightarrow P\Delta \geq \boldsymbol{q}\boldsymbol{\lambda}^\top - (\boldsymbol{q} - P\boldsymbol{\zeta})\boldsymbol{u}^\top \Rightarrow QP\Delta \geq Q\boldsymbol{q}\boldsymbol{\lambda}^\top - Q(\boldsymbol{q} - P\boldsymbol{\zeta})\boldsymbol{u}^\top \Rightarrow \Delta \leq \boldsymbol{\zeta}\boldsymbol{u}^\top$$

$$(8\text{f}) \Rightarrow P\Xi P^\top + \boldsymbol{q}\boldsymbol{q}^\top \geq P\boldsymbol{\zeta}\boldsymbol{q}^\top + \boldsymbol{q}\boldsymbol{\zeta}^\top P^\top \Rightarrow QP\Xi P^\top \geq Q\boldsymbol{q}(P\boldsymbol{\zeta} - \boldsymbol{q})^\top + QP\boldsymbol{\zeta}\boldsymbol{q}^\top \Rightarrow \Xi P^\top \leq \boldsymbol{\zeta}\boldsymbol{q}^\top$$

$$\Rightarrow \Xi P^\top \leq \boldsymbol{\zeta}\boldsymbol{q}^\top \Rightarrow Q\Xi P^\top \leq Q\boldsymbol{\zeta}\boldsymbol{q}^\top \Rightarrow \Xi \geq 0\,.$$

Hence, one can establish that

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

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

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

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

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

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

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

References

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

Ardestani-Jaafari A, Delage E (2018) The value of flexibility in robust locationtransportation problems. *Transportation Science* 52(1):189–209.

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

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

Denton BT, Miller AJ, Balasubramanian HJ, Huschka TR (2010) Optimal allocation of surgery blocks to operating rooms under uncertainty. *Oper. Res.* 58(4-part-1):802–816.

Hanasusanto G, Kuhn D (2018) Conic programming reformulations of two-stage distributionally robust linear programs over Wasserstein balls. *Operations Research* 66(3):849–869.

Parrilo P (2000) *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. Ph.D. thesis, California Institute of Technology.

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

Shapiro A, Dentcheva D, Ruszczyński A (2009) *Lectures on stochastic programming: modeling and theory* (SIAM).

Xu G, Burer S (2018) A copositive approach for two-stage adjustable robust optimization with uncertain right-hand sides. *Computational Optimization and Applications* 70(1):33–59.