A Copositive Approach for Two-Stage Adjustable Robust Optimization with Uncertain Right-Hand Sides

Guanglin Xu* Samuel Burer†

September 23, 2016

Abstract

We study two-stage adjustable robust linear programming in which the right-hand sides are uncertain and belong to a convex, compact uncertainty set. This problem is NP-hard, and the affine policy is a popular, tractable approximation. We prove that under standard and simple conditions, the two-stage problem can be reformulated as a copositive optimization problem, which in turn leads to a class of tractable, semidefinite-based approximations that are at least as strong as the affine policy. We investigate several examples from the literature demonstrating that our tractable approximations significantly improve the affine policy. In particular, our approach solves exactly in polynomial time a class of instances of increasing size for which the affine policy admits an arbitrarily large gap.

Keywords: Two-stage adjustable robust optimization, robust optimization, bilinear programming, non-convex quadratic programming, semidefinite programming, copositive programming.

1 Introduction

Ben-Tal et. al. [9] introduced two-stage *adjustable robust optimization (ARO)*, which considers both first-stage ("here-and-now") and second-stage ("wait-and-see") variables. ARO can be significantly less conservative than regular robust optimization, and real-world applications of ARO abound: unit commitment in renewable energy [14, 32, 35], facility location problems [4, 6, 24], emergency supply chain planning [7], and inventory management [1, 30];

*Department of Management Sciences, University of Iowa, Iowa City, IA, 52242-1994, USA. Email: guanglin-xu@uiowa.edu.

†Department of Management Sciences, University of Iowa, Iowa City, IA, 52242-1994, USA. Email: samuel-burer@uiowa.edu.

see also [8, 23, 28]. We refer the reader to the excellent, recent tutorial [20] for background on ARO.

Since ARO is intractable in general [9], multiple tractable approximations have been proposed for it. In certain situations, a static, robust-optimization-based solution can be used to approximate ARO, and sometimes this static solution is optimal [10, 12]. The *affine policy* [9], which forces the second-stage variables to be an affine function of the uncertainty parameters, is another common approximation for ARO, but it is generally suboptimal. Several nonlinear policies have also been used to approximate ARO. Chen and Zhang [19] proposed the *extended affine policy* in which the primitive uncertainty set is reparameterized by introducing auxiliary variables after which the regular affine policy is applied. Bertsimas et. al. [13] introduced a more accurate, yet more complicated, approximation which forces the second-stage variables to depend polynomially (with a user-specified, fixed degree) on the uncertain parameters. Their approach yields a hierarchy of Lasserre-type semidefinite approximations and can be extended to multi-stage robust optimization.

The approaches just described provide upper bounds when ARO is stated as a minimization. On the other hand, a single lower bound can be calculated, for example, by fixing a specific value in the uncertainty set and solving the resulting LP (linear program), and Monte Carlo simulation over the uncertainty set can then be used to compute a best lower bound. Finally, global approaches for solving ARO exactly include column and constraint generation [34] and Benders decomposition [14, 21].

In this paper, we consider the following two-stage adjustable robust linear minimization problem with uncertain right-hand side:

$$\begin{aligned} v^*_{\mathrm{RLP}} := \min_{x,y(\cdot)} \quad & c^T x + \max_{u \in \mathcal{U}} d^T y(u) \\ \text{s.t.} \quad & Ax + By(u) \geq Fu \quad \forall\, u \in \mathcal{U} \\ & x \in \mathcal{X}, \end{aligned} \tag{RLP}$$

where $A \in \mathbb{R}^{m\times n_1}, B \in \mathbb{R}^{m\times n_2}, c \in \mathbb{R}^{n_1}$, $d \in \mathbb{R}^{n_2}$, $F \in \mathbb{R}^{m\times k}$ and $\mathcal{X} \subseteq \mathbb{R}^{n_1}$ is a closed convex set containing the first-stage decision x. The uncertainty set $\mathcal{U} \subseteq \mathbb{R}^k$ is compact, convex, and nonempty and, in particular, we model it as a slice of a closed, convex, full-dimensional cone $\widehat{\mathcal{U}} \subseteq \mathbb{R}^k$:

$$\mathcal{U} := \{u \in \widehat{\mathcal{U}} : e_1^T u = u_1 = 1\}, \tag{1}$$

where e_1 is the first canonical basic vector in $\mathbb{R}^k$. In words, $\widehat{\mathcal{U}}$ is the homogenization of $\mathcal{U}$. We choose this homogenized version for notational convenience throughout the paper and note that it allows the modeling of affine effects of the uncertain parameters. The second-stage variable is $y(\cdot)$, formally defined as a mapping $y : \mathcal{U} \to \mathbb{R}^{n_2}$. It is well known that (RLP) is equivalent to

$$v^*_{\text{RLP}} = \min_{x \in \mathcal{X}} \quad c^T x + \max_{u \in \mathcal{U}} \min_{y(u) \in \mathbb{R}^{n_2}} \{d^T y(u) : By(u) \geq Fu - Ax\}, \tag{2}$$

where $y(u)$ is a vector variable specifying the value of $y(\cdot)$ at u.

Regarding (RLP), we make three standard assumptions.

Assumption 1. *The closed, convex set $\mathcal{X}$ and the closed, convex cone $\widehat{\mathcal{U}}$ are both full-dimensional and tractable.*

For example, $\mathcal{X}$ and $\widehat{\mathcal{U}}$ could be represented using linear, second-order-cone, and semidefinite inequalities.

Assumption 2. *Problem* (RLP) *is feasible, i.e., there exists a choice of $x \in \mathcal{X}$ and $y(\cdot)$ such that $Ax + By(u) \geq Fu$ for all $u \in \mathcal{U}$.*

The existence of an affine policy, which can be checked in polynomial time, is sufficient to establish that Assumption 2 holds.

Assumption 3. *Problem* (RLP) *is bounded, i.e., v^*_{RLP} is finite.*

Note that the negative directions of recession $\{r : d^T r < 0, Br \geq 0\}$ for the innermost LP in (2) do not depend on x and u. Hence, in light of Assumptions 2 and 3, there must exist no negative directions of recession; otherwise, v^*_{RLP} would clearly equal $-\infty$. So every innermost LP in (2) is either feasible and bounded or infeasible. In particular, Assumption 2 implies that at least one such LP is feasible and bounded. It follows that the associated dual LP $\max\{(Fu - Ax)^T w : B^T w = d, w \geq 0\}$ is also feasible with bounded value. In particular, the fixed set

$$\mathcal{W} := \{w \geq 0 : B^T w = d\}$$

is nonempty. For this paper, we also make one additional assumption:

Assumption 4. *The set $\mathcal{W}$ is bounded.*

Note that Assumption 4 holds if and only if there exists r such that $Br > 0$.

In Section 2, under Assumptions 1–4, we reformulate (RLP) as an equivalent copositive program, which first and foremost enables a new perspective on two-stage robust optimization. Compared to most existing copositive approaches for difficult problems, ours exploits copositive duality; indeed, Assumption 4 is sufficient for establishing strong duality between the copositive primal and dual. In Section 3, we then apply a similar approach to derive a new formulation of the affine policy, which is then, in Section 4, directly related to the copositive version of (RLP). This establishes two extremes: on the one side is the copositive representation of (RLP), while on the other is the affine policy. Section 4 also proposes semidefinite-based approximations of (RLP) that interpolate between the full copositive program and the affine policy. Finally, in Section 5, we investigate several examples from the literature that demonstrate our bounds can significantly improve the affine-policy value. In particular, we prove that our semidefinite approach solves a class of instances of increasing size for which the affine policy admits arbitrarily large gaps. We end the paper with a short discusson of future directions in Section 6.

During the writing of this paper, we became aware of a recent technical report by Ardestani-Jaafari and Delage [5], which introduces an approach for (RLP) that is very similar in spirit to ours in Section 2. Whereas we use copositive duality to reformulate (RLP) exactly and then approximate it using semidefinite programming, [5] uses semidefinite duality to approximate (RLP) in one step. We prefer our two-step approach because it clearly separates the use of conic duality from the choice of approximation. We also feel that our derivation is more compact. On the other hand, the "cost" of our more generic approach is the additional Assumption 4, which [5] does not make, although it is a mild cost because, even without Assumption 4, our approach still leads to strong approximations of (RLP). In addition, [5] focuses mainly on the case when $\mathcal{U}$ is polyhedral, whereas our approach builds semidefinite-based approximations for any $\mathcal{U}$ that can be represented, say, by linear, second-order-cone, and semidefinite inequalities.

1.1 Notation, terminology, and background

Let $\mathbb{R}^n$ denote n-dimensional Euclidean space represented as column vectors, and let $\mathbb{R}^n_+$ denote the nonnegative orthant in $\mathbb{R}^n$. For a scalar $p \ge 1$, the p-norm of $v \in \mathbb{R}^n$ is defined $\|v\|_p := (\sum_{i=1}^n |v_i|^p)^{1/p}$, e.g., $\|v\|_1 = \sum_{i=1}^n |v_i|$. We will drop the subscript for the 2-norm, i.e., $\|v\| := \|v\|_2$. For $v, w \in \mathbb{R}^n$, the inner product of v and w is $v^T w := \sum_{i=1}^n v_i w_i$. The symbol $\mathbb{1}_n$ denotes the all-ones vector in $\mathbb{R}^n$.

The space $\mathbb{R}^{m\times n}$ denotes the set of real $m \times n$ matrices, and the trace inner product of two matrices $A, B \in \mathbb{R}^{m\times n}$ is $A \bullet B := \text{trace}(A^T B)$. $\mathcal{S}^n$ denotes the space of $n \times n$ symmetric matrices, and for $X \in \mathcal{S}^n$, $X \succeq 0$ means that X is positive semidefinite. In addition, $\text{diag}(X)$ denotes the vector containing the diagonal entries of X, and $\text{Diag}(v)$ is the diagonal matrix with vector v along its diagonal. We denote the null space of a matrix A as $\text{Null}(A)$, i.e., $\text{Null}(A) := \{x : Ax = 0\}$. For $\mathcal{K} \subseteq \mathbb{R}^n$ a closed, convex cone, $\mathcal{K}^*$ denotes its dual cone. For a matrix A with n columns, the inclusion $\text{Rows}(A) \in \mathcal{K}$ indicates that the rows of A—considered as column vectors—are members of $\mathcal{K}$.

We next introduce some basics of *copositive programming* with respect to the cone $\mathcal{K} \subseteq \mathbb{R}^n$. The *copositive cone* is defined as

$$\mathrm{COP}(\mathcal{K}) := \{M \in \mathcal{S}^n : x^T M x \geq 0 \ \forall \ x \in \mathcal{K}\},$$

and its dual cone, the *completely positive cone*, is

$$\mathrm{CPP}(\mathcal{K}) := \{X \in \mathcal{S}^n : X = \textstyle\sum_i x^i (x^i)^T, \ x^i \in \mathcal{K}\},$$

where the summation over i is finite but its cardinality is unspecified. The term *copositive programming* refers to linear optimization over $\mathrm{COP}(\mathcal{K})$ or, via duality, linear optimization over $\mathrm{CPP}(\mathcal{K})$. In fact, these problems are sometimes called *generalized copositive programming* or *set-semidefinite optimization* [18, 22] in contrast with the standard case $\mathcal{K} = \mathbb{R}^n_+$. In this paper, we work with generalized copositive programming, although we use the shorter phrase for convenience.

Finally, for the specific dimensions k and m of problem (RLP), we let e_i denote the i-th coordinator vector in $\mathbb{R}^k$, and similarly, f_j denotes the j-th coordinate vector in $\mathbb{R}^m$. We will also use $g_1 := \binom{e_1}{0} \in \mathbb{R}^{k+m}$.

2 A Copositive Reformulation

In this section, we construct a copositive representation of (RLP) under Assumptions 1–4 by first reformulating the inner maximization of (2) as a copositive problem and then employing copositive duality.

Within (2), define

$$\pi(x) := \max_{u \in \mathcal{U}} \min_{y(u) \in \mathbb{R}^{n_2}} \{d^T y(u) : By(u) \geq Fu - Ax\}.$$

The dual of the inner minimization is $\max_{w \in \mathcal{W}} (Fu - Ax)^T w$, which is feasible as discussed in the Introduction. Hence, strong duality for LP implies

$$\pi(x) = \max_{u \in \mathcal{U}} \max_{w} \{(Fu - Ax)^T w : w \in \mathcal{W}\} = \max_{(u,w) \in \mathcal{U} \times \mathcal{W}} (Fu - Ax)^T w, \tag{3}$$

In words, $\pi(x)$ equals the optimal value of a bilinear program over convex constraints, which is NP-hard in general [26].

It holds also that $\pi(x)$ equals the optimal value of an associated copositive program

[15, 16], which we now describe. Define

$$z := \begin{pmatrix} u \\ w \end{pmatrix} \in \mathbb{R}^{k+m}, \quad E := \begin{pmatrix} -de_1^T & B^T \end{pmatrix} \in \mathbb{R}^{n_2 \times (k+m)}, \tag{4}$$

where $e_1 \in \mathbb{R}^k$ is the first coordinate vector, and homogenize via the relationship (1) and the definition of $\mathcal{W}$:

$$\begin{aligned} \pi(x) = \max \quad & (F - Axe_1^T) \bullet wu^T \\ \text{s.t.} \quad & Ez = 0 \\ & z \in \widehat{\mathcal{U}} \times \mathbb{R}^m_+, \quad g_1^T z = 1, \end{aligned}$$

where g_1 is the first coordinate vector in $\mathbb{R}^{k+m}$. The copositive representation is thus

$$\begin{aligned} \pi(x) = \max \; & (F - Axe_1^T) \bullet Z_{21} \\ \text{s.t.} \;\; & \operatorname{diag}(EZE^T) = 0 \\ & Z \in \text{CPP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+), \quad g_1 g_1^T \bullet Z = 1, \end{aligned} \tag{5}$$

where Z has the block structure

$$Z = \begin{pmatrix} Z_{11} & Z_{21}^T \\ Z_{21} & Z_{21} \end{pmatrix} \in \mathcal{S}^{k+m}.$$

Note that under positive semidefiniteness, which is implied by the completely positive constraint, the constraint $\operatorname{diag}(EZE^T) = 0$ is equivalent to $ZE^T = 0$; see proposition 1 of [16], for example. For the majority of this paper, we will focus on this second version:

$$\begin{aligned} \pi(x) = \max \; & (F - Axe_1^T) \bullet Z_{21} \\ \text{s.t.} \;\; & ZE^T = 0 \\ & Z \in \text{CPP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+), \quad g_1 g_1^T \bullet Z = 1. \end{aligned} \tag{6}$$

Because both $\mathcal{U}$ and $\mathcal{W}$ are bounded by assumption, there exists a scalar $r > 0$ such that the constraint $z^T z = u^T u + w^T w \leq r$ is redundant for (3). Hence, the lifted and linearized constraint $I \bullet Z \leq r$ can be added to (6) without changing its optimal value, although some feasible directions of recession may be cut off. We arrive at

$$\begin{aligned} \pi(x) = \max\ & (F - Axe_1^T) \bullet Z_{21} \\ \text{s.t.}\ & ZE^T = 0, \quad I \bullet Z \leq r \\ & Z \in \mathrm{CPP}(\widehat{\mathcal{U}} \times \mathbb{R}_+^m), \quad g_1 g_1^T \bullet Z = 1. \end{aligned} \tag{7}$$

Letting $\Lambda \in \mathbb{R}^{(k+m)\times n_2}$, $\lambda \in \mathbb{R}$, and $\rho \in \mathbb{R}$ be the respective dual multipliers of $ZE^T = 0$, $g_1 g_1^T \bullet Z = 1$, and $I \bullet Z \leq r$, standard conic duality theory implies the dual of (7) is

$$\begin{aligned} \min_{\lambda,\Lambda,\rho}\ & \lambda \\ \text{s.t.}\ & \lambda g_1 g_1^T - \tfrac{1}{2}G(x) + \tfrac{1}{2}(E^T\Lambda^T + \Lambda E) + \rho I \in \mathrm{COP}(\widehat{\mathcal{U}} \times \mathbb{R}_+^m) \\ & \rho \geq 0, \end{aligned} \tag{8}$$

where

$$G(x) := \begin{pmatrix} 0 & (F - Axe_1^T)^T \\ F - Axe_1^T & 0 \end{pmatrix} \in \mathcal{S}^{k+m}$$

is affine in x. Holding all other dual variables fixed, for $\rho > 0$ large, the matrix variable in (8) is strictly copositive—in fact, positive definite—which establishes that Slater's condition is satisfied, thus ensuring strong duality:

Proposition 1. *Under Assumption 4, suppose $r > 0$ is a constant such that $z^T z \leq r$ for all $z = (u, w) \in \mathcal{U} \times \mathcal{W}$. Then the optimal value of (8) equals $\pi(x)$.*

Now, with $\pi(x)$ expressed as a minimization that depends affinely on x, we can collapse (2) into a single minimization that is equivalent to (RLP):

$$\begin{aligned} \min_{x,\lambda,\Lambda,\rho}\ & c^T x + \lambda + r\rho \\ \text{s.t.}\ & x \in \mathcal{X},\ \lambda g_1 g_1^T - \tfrac{1}{2}G(x) + \tfrac{1}{2}(E^T\Lambda^T + \Lambda E) + \rho I \in \mathrm{COP}(\widehat{\mathcal{U}} \times \mathbb{R}_+^m) \\ & \rho \geq 0. \end{aligned} \tag{$\overline{RLP}$}$$

Theorem 1. *The optimal value of ($\overline{RLP}$) equals v^*_{RLP}.*

An equivalent version of ($\overline{RLP}$) can be derived based on the representation of $\pi(x)$ in (5):

$$\begin{aligned} \min_{x,\lambda,v,\rho}\ & c^T x + \lambda + r\rho \\ \text{s.t.}\ & x \in \mathcal{X},\ \lambda g_1 g_1^T - \tfrac{1}{2}G(x) + E^T \operatorname{Diag}(v) E + \rho I \in \mathrm{COP}(\widehat{\mathcal{U}} \times \mathbb{R}_+^m) \\ & \rho \geq 0. \end{aligned} \tag{9}$$

Our example in Section 5.1 will be based on this version.

We remark that, even if Assumption 4 fails and strong duality between (7) and (8) cannot be established, it still holds that the optimal value of ($\overline{RLP}$) is an upper bound on v^*_{RLP}. Note that, in this case, (7) should be modified to exclude $I \bullet Z \leq r$, and ρ should be set to 0 in (8).

3 The Affine Policy

Under the affine policy, the second-stage decision variable $y(\cdot)$ in (RLP) is modeled as a linear funciton of u via a free variable $Y \in \mathbb{R}^{n_2 \times k}$:

$$\begin{array}{rll} v^*_{\text{Aff}} := & \min\limits_{x, y(\cdot), Y} & c^T x + \max\limits_{u \in \mathcal{U}} d^T y(u) \\ & \text{s. t.} & Ax + By(u) \geq Fu \quad \forall\, u \in \mathcal{U} \\ & & y(u) = Yu \qquad\qquad \forall\, u \in \mathcal{U} \\ & & x \in \mathcal{X}. \end{array} \tag{Aff}$$

Here, Y acts as a "dummy" first-stage decision, and so (Aff) can be recast as a regular robust optimization problem over $\mathcal{U}$. Specifically, using standard techniques [9], (Aff) is equivalent to

$$\begin{array}{rl} \min\limits_{x, Y, \lambda} & c^T x + \lambda \\ \text{s. t.} & (\lambda e_1 - Y^T d) \in \widehat{\mathcal{U}}^* \\ & \text{Rows}(Axe_1^T - F + BY) \in \widehat{\mathcal{U}}^* \\ & x \in \mathcal{X}. \end{array} \tag{10}$$

Problem (10) is tractable, but in general, the affine policy is only an approximation of (RLP), i.e., $v^*_{\text{RLP}} < v^*_{\text{Aff}}$. In what follows, we provide a copositive representation ($\overline{\text{Aff}}$) of (Aff), which is then used to develop an alternative formulation (IA) of (10). Later, in Section 4, problem (IA) will be compared directly to ($\overline{\text{RLP}}$).

Following the approach of Section 2, we may express (Aff) as $\min\limits_{x \in \mathcal{X}, Y} c^T x + \Pi(x, Y)$ where

$$\Pi(x, Y) := \max_{u \in \mathcal{U}} \min_{y(u) \in \mathbb{R}^{n_2}} \{ d^T y(u) : By \geq Fu - Ax,\ y(u) = Yu \}.$$

The inner minimization has dual

$$\begin{aligned} & \max_{w \geq 0, v} \{ (Fu - Ax)^T w + (Yu)^T v : B^T w + v = d \} \\ & = \max_{w \geq 0} \left((Fu - Ax)^T w + (Yu)^T (d - B^T w) \right). \end{aligned}$$

After collecting terms, homogenizing, and converting to copostive optimization, we have

$$\begin{array}{rll} \Pi(x,Y) = & \max & (F - Axe_1^T - BY) \bullet Z_{21} + e_1 d^T Y \bullet Z_{11} \\ & \text{s.t.} & Z \in \text{CPP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+), \quad g_1 g_1^T \bullet Z = 1 \end{array} \tag{11}$$

with dual

$$\begin{array}{ll} \min\limits_{\lambda} & \lambda \\ \text{s.t.} & \lambda g_1 g_1^T - \frac{1}{2} G(x) + \frac{1}{2} H(Y) \in \text{COP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+), \end{array} \tag{12}$$

where $G(x)$ is defined as in Section 2 and

$$H(Y) := \begin{pmatrix} -e_1 d^T Y - Y^T d e_1^T & (BY)^T \\ BY & 0 \end{pmatrix} \in \mathcal{S}^{k+m}.$$

Since $\widehat{\mathcal{U}}$ has interior by Assumption 1, it follows that (11) also has interior, and so Slater's condition holds, implying strong duality between (11) and (12). Thus, repeating the logic of Section 2, (Aff) is equivalent to

$$\begin{array}{ll} \min\limits_{x,\lambda,Y} & c^T x + \lambda \\ \text{s.t.} & x \in \mathcal{X}, \quad \lambda g_1 g_1^T - \frac{1}{2} G(x) + \frac{1}{2} H(Y) \in \text{COP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+). \end{array} \tag{$\overline{Aff}$}$$

Proposition 2. *The optimal value of ($\overline{Aff}$) is v^*_{Aff}.*

We now show that $\text{COP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$ in ($\overline{\text{Aff}}$) can be replaced by a particular inner approximation without changing the optimal value. Moreover, this inner approximation is tractable, so that the resulting optimization problem serves as an alternative to the formulation (10) of (Aff).

Using the mnemonic "IA" for "inner approximation," we define

$$\text{IA}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+) := \left\{ S = \begin{pmatrix} S_{11} & S_{21}^T \\ S_{21} & S_{22} \end{pmatrix} : \begin{array}{l} S_{11} = e_1 \alpha^T + \alpha e_1^T, \ \alpha \in \widehat{\mathcal{U}}^*, \\ \text{Rows}(S_{21}) \in \widehat{\mathcal{U}}^*, \ S_{22} \geq 0 \end{array} \right\}.$$

This set is tractable because it is defined by affine constraints in $\widehat{\mathcal{U}}^*$ as well as nonnegativity constraints. Moreover, $\text{IA}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$ is indeed a subset of $\text{COP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$:

Lemma 1. $\text{IA}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+) \subseteq \text{COP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$.

Proof. We first note that (1) implies that the first coordinate of every element of $\widehat{\mathcal{U}}$ is nonnegative; hence, $e_1 \in \widehat{\mathcal{U}}^*$. Now, for arbitrary $\binom{p}{q} \in \widehat{\mathcal{U}} \times \mathbb{R}^m_+$ and $S \in \text{IA}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$, we prove $t := \binom{p}{q}^T S \binom{p}{q} \geq 0$. We have

$$t = \begin{pmatrix} p \\ q \end{pmatrix}^T \begin{pmatrix} S_{11} & S_{21}^T \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = p^T S_{11} p + 2\, q^T S_{21} p + q^T S_{22} q.$$

Analyzing each of the three summands separately, we first have

$$e_1, \alpha \in \widehat{\mathcal{U}}^* \quad \Longrightarrow \quad p^T S_{11} p = p^T (e_1 \alpha^T + \alpha e_1^T) p = 2(p^T e_1)(\alpha^T p) \geq 0.$$

Second, $p \in \widehat{\mathcal{U}}$ and $\text{Rows}(S_{21}) \in \widehat{\mathcal{U}}^*$ imply $S_{21} p \geq 0$, which in turn implies $q^T S_{21} p = q^T (S_{21} p) \geq 0$ because $q \geq 0$. Finally, it is clear that $q^T S_{22} q \geq 0$ as $S_{22} \geq 0$ and $q \geq 0$. Thus, $t \geq 0 + 0 + 0 = 0$, as desired. □

The following tightening of ($\overline{\text{Aff}}$) simply replaces $\text{COP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$ with its inner approximation $\text{IA}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$:

$$\begin{array}{llll} v^*_{\text{IA}} & := & \min\limits_{x,\lambda,Y} & c^T x + \lambda \\ & & \text{s.t.} & x \in \mathcal{X}, \quad \lambda g_1 g_1^T - \tfrac{1}{2} G(x) + \tfrac{1}{2} H(Y) \in \text{IA}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+). \end{array} \tag{IA}$$

By construction, $v^*_{\text{IA}} \geq v^*_{\text{Aff}}$, but in fact these values are equal.

Theorem 2. $v^*_{\text{IA}} = v^*_{\text{Aff}}$.

Proof. We show $v^*_{\text{IA}} \leq v^*_{\text{Aff}}$ by demonstrating that every feasible solution of (10) yields a feasible solution of (IA) with the same objective value. Let (x, Y, λ) be feasible for (10); we prove

$$S := \lambda g_1 g_1^T - \tfrac{1}{2} G(x) + \tfrac{1}{2} H(Y) \in \text{IA}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+),$$

which suffices. Note that the block form of S is

$$S = \begin{pmatrix} \lambda e_1 e_1^T - \frac{1}{2}(e_1 d^T Y + Y^T d e_1^T) & \frac{1}{2}(A x e_1^T - F + BY)^T \\ \frac{1}{2}(A x e_1^T - F + BY) & 0 \end{pmatrix}.$$

The argument decomposes into three pieces. First, we define $\alpha := \frac{1}{2}(\lambda e_1 - Y^T d)$, which satisfies $\alpha \in \widehat{\mathcal{U}}^*$ due to (10). Then

$$\begin{aligned} S_{11} &= \lambda e_1 e_1^T - \tfrac{1}{2}(e_1 d^T Y + Y^T d e_1^T) \\ &= \left(\tfrac{1}{2}\lambda e_1 e_1^T - \tfrac{1}{2} e_1 d^T Y\right) + \left(\tfrac{1}{2}\lambda e_1 e_1^T - \tfrac{1}{2} Y^T d e_1^T\right) \\ &= e_1 \alpha^T + \alpha e_1^T \end{aligned}$$

as desired. Second, we have $2\,\mathrm{Rows}(S_{21}) = \mathrm{Rows}(Axe_1^T - F + BY) \in \widehat{\mathcal{U}}^*$ by (10). Finally, $S_{22} = 0 \geq 0$. $\square$

4 Improving the Affine Policy

A direct relationship holds between ($\overline{\text{RLP}}$) and (IA):

Proposition 3. *In problem* ($\overline{\text{RLP}}$)*, write* $\Lambda = \binom{\Lambda_1}{\Lambda_2}$*, where* $\Lambda_1 \in \mathbb{R}^{k\times n_2}$ *and* $\Lambda_2 \in \mathbb{R}^{m\times n_2}$*. Problem* (IA) *is a restriction of* ($\overline{\text{RLP}}$) *in which* $\Lambda_2 = 0$*,* Y *is identified with* Λ_1^T*,* $\rho = 0$*, and* $\mathrm{COP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$ *is tightened to* $\mathrm{IA}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$*.*

Proof. Examining the similar structure of ($\overline{\text{RLP}}$) and (IA), it suffices to equate the terms $E^T\Lambda^T + \Lambda E$ and $H(Y)$ in the respective problems under the stated restrictions. From (4),

$$E^T\Lambda^T + \Lambda E = \begin{pmatrix} -e_1 d^T \Lambda_1^T - \Lambda_1 d e_1^T & \Lambda_1 B^T - e_1 d^T \Lambda_2^T \\ B\Lambda_1^T - \Lambda_2 d e_1^T & B\Lambda_2^T + \Lambda_2 B^T \end{pmatrix}.$$

Setting $\Lambda_2 = 0$ and identifying $Y = \Lambda_1^T$, we see

$$E^T\Lambda^T + \Lambda E = \begin{pmatrix} -e_1 d^T Y - Y^T d e_1^T & Y^T B^T \\ BY & 0 \end{pmatrix} = H(Y),$$

as desired. $\square$

Now let $\mathrm{IB}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$ be any closed convex cone satisfying

$$\mathrm{IA}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+) \subseteq \mathrm{IB}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+) \subseteq \mathrm{COP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+),$$

where the mnemonic "IB" stands for "in between", and consider the following problem gotten by replacing $\mathrm{COP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$ in ($\overline{\text{RLP}}$) with $\mathrm{IB}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$:

$$\begin{aligned} v^*_{\mathrm{IB}} \;:=\; \min_{x,\lambda,\Lambda} \quad & c^T x + \lambda \\ \text{s.t.} \quad & x \in \mathcal{X}, \quad \lambda g_1 g_1^T - \tfrac{1}{2}G(x) + \tfrac{1}{2}(E^T\Lambda^T + \Lambda E) \in \mathrm{IB}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+). \end{aligned} \tag{IB}$$

Problem (IB) is clearly a restriction of ($\overline{\text{RLP}}$), and by Proposition 3, it is simultaneously no tighter than (IA). Combining this with Theorems 1 and 2, we thus have:

Theorem 3. $v^*_{\mathrm{RLP}} \leq v^*_{\mathrm{IB}} \leq v^*_{\mathrm{Aff}}$.

We end this section with a short discussion of example approximations $\mathrm{IB}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$ for typical cases of $\widehat{\mathcal{U}}$. In fact, there are complete hierarchies of approximations of $\mathrm{COP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$

[36], but we present a relatively simple construction that starts from a given inner approximation $\mathrm{IB}(\widehat{\mathcal{U}})$ of $\mathrm{COP}(\widehat{\mathcal{U}})$:

Proposition 4. *Suppose* $\mathrm{IB}(\widehat{\mathcal{U}}) \subseteq \mathrm{COP}(\widehat{\mathcal{U}})$, *and define*

$$\mathrm{IB}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+) := \left\{ S + M + R \ : \ \begin{array}{c} S \in \mathrm{IA}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+), \ M \succeq 0 \\ R_{11} \in \mathrm{IB}(\widehat{\mathcal{U}}), \ R_{21} = 0, \ R_{22} = 0 \end{array} \right\}.$$

Then $\mathrm{IA}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+) \subseteq \mathrm{IB}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+) \subseteq \mathrm{COP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$.

Proof. For the first inclusion, simply take $M = 0$ and $R_{11} = 0$. For the second inclusion, let arbitrary $\binom{p}{q} \in \widehat{\mathcal{U}} \times \mathbb{R}^m_+$ be given. We need to show

$$\begin{pmatrix} p \\ q \end{pmatrix}^T (S + M + R) \begin{pmatrix} p \\ q \end{pmatrix} = \tbinom{p}{q}^T S \tbinom{p}{q} + \tbinom{p}{q}^T M \tbinom{p}{q} + p^T R_{11} p \geq 0.$$

The first term is nonnegative because $S \in \mathrm{IA}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$; the second term is nonnegative because $M \succeq 0$; and the third is nonnegative because $R_{11} \in \mathrm{COP}(\widehat{\mathcal{U}})$. □

When $\widehat{\mathcal{U}} = \{u \in \mathbb{R}^k : \|(u_2, \ldots, u_k)^T\| \leq u_1\}$ is the second-order cone, it is known [31] that

$$\mathrm{COP}(\widehat{\mathcal{U}}) = \{R_{11} = \tau J + M_{11} : \tau \geq 0, \ M_{11} \succeq 0\},$$

where $J = \mathrm{Diag}(1, -1, \ldots, -1)$. Because of this simple structure, it often makes sense to take $\mathrm{IB}(\widehat{\mathcal{U}}) = \mathrm{COP}(\widehat{\mathcal{U}})$ in practice. Note also that $M_{11} \succeq 0$ can be absorbed into $M \succeq 0$ in the definition of $\mathrm{IB}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$ above. When $\widehat{\mathcal{U}} = \{u \in \mathbb{R}^k : Pu \geq 0\}$ is a polyhedral cone based on some matrix P, a typical inner approximation of $\mathrm{COP}(\widehat{\mathcal{U}})$ is

$$\mathrm{IB}(\widehat{\mathcal{U}}) := \{R_{11} = P^T N P : N \geq 0\},$$

where N is a symmetric matrix variable of appropriate size. This corresponds to the RLT approach of [2, 17, 29].

5 Examples

In this section, we demonstrate our approximation v^*_{IB} satisfying $v^*_{\mathrm{RLP}} \leq v^*_{\mathrm{IB}} \leq v^*_{\mathrm{Aff}}$ on several examples from the literature. The first example is treated analytically, while the remaining examples are verified numerically. All computations are conducted with Mosek version 8.0.0.28 beta [3] on an Intel Core i3 2.93 GHz Windows computer with 4GB of RAM and implemented using the modeling language YALMIP [27] in MATLAB (R2014a).

5.1 A temporal network example

The paper [33] studies a so-called *temporal network* application, which for any integer $s \geq 2$ leads to the following problem, which is based on an uncertainty set $\Xi \subseteq \mathbb{R}^s$ and in which the first-stage decision x is fixed, say, at 0 and $y(\cdot)$ maps into $\mathbb{R}^s$:

$$\begin{array}{lll} \min\limits_{y(\cdot)} & \max\limits_{\xi\in\Xi} \; y(\xi)_s & \\ \text{s. t.} & y(\xi)_1 \geq \max\{\xi_1, 1-\xi_1\} & \forall\, \xi \in \Xi \\ & y(\xi)_2 \geq \max\{\xi_2, 1-\xi_2\} + y(\xi)_1 & \forall\, \xi \in \Xi \\ & \vdots & \\ & y(\xi)_s \geq \max\{\xi_s, 1-\xi_s\} + y(\xi)_{s-1} & \forall\, \xi \in \Xi. \end{array} \tag{13}$$

Note that each of the above linear constraints can be expressed as two separate linear constraints. The authors of [33] consider a polyhedral uncertainty set (based on the 1-norm). A related paper [25] considers a conic uncertainty set (based on the 2-norm) for $s = 2$; we will extend this to $s \geq 2$. In particular, we consider the following two uncertainty sets for general s:

$$\Xi_1 := \{\xi \in \mathbb{R}^s : \|\xi - \tfrac{1}{2}\mathbb{1}_s\|_1 \leq \tfrac{1}{2}\},$$
$$\Xi_2 := \{\xi \in \mathbb{R}^s : \|\xi - \tfrac{1}{2}\mathbb{1}_s\| \leq \tfrac{1}{2}\},$$

where $\mathbb{1}_s$ denotes the all-ones vector in $\mathbb{R}^s$. For $j = 1, 2$, let $v^*_{\mathrm{RLP},j}$ and $v^*_{\mathrm{Aff},j}$ be the robust and affine values associated with (13) for the uncertainty set Ξ_j. Note that $\Xi_1 \subseteq \Xi_2$, and hence $v^*_{\mathrm{RLP},1} \leq v^*_{\mathrm{RLP},2}$.

The papers [25, 33] show that $v^*_{\mathrm{Aff},1} = v^*_{\mathrm{Aff},2} = s$, and [33] establishes $v^*_{\mathrm{RLP},1} = \frac{1}{2}(s+1)$. Moreover, we can calculate $v^*_{\mathrm{RLP},2}$ in this paper by the following analysis. Any feasible $y(\xi)$ satisfies

$$\begin{aligned} y(\xi)_s &\geq \max\{\xi_s, 1-\xi_s\} + y(\xi)_{s-1} \\ &\geq \max\{\xi_s, 1-\xi_s\} + \max\{\xi_{s-1}, 1-\xi_{s-1}\} + y(\xi)_{s-2} \\ &\geq \cdots \geq \sum_{i=1}^{s} \max\{\xi_i, 1-\xi_i\} \end{aligned}$$

Hence, applying this inequality at an optimal $y(\cdot)$, it follows that

$$v^*_{\mathrm{RLP},2} \geq \max_{\xi\in\Xi_2} \sum_{i=1}^{s} \max\{\xi_i, 1-\xi_i\}.$$

Under the change of variables $\mu := 2\xi - \mathbb{1}_s$, we have

$$\begin{aligned} v^*_{\text{RLP},2} &\geq \max_{\xi \in \Xi_2} \sum_{i=1}^{s} \max\{\xi_i, 1-\xi_i\} = \max_{\|\mu\| \leq 1} \sum_{i=1}^{s} \tfrac{1}{2} \max\{1+\mu_i, 1-\mu_i\} \\ &= \frac{1}{2} \max_{\|\mu\| \leq 1} \sum_{i=1}^{s} (1 + |\mu_i|) = \frac{1}{2} \left(s + \max_{\|\mu\| \leq 1} \|\mu\|_1 \right) = \tfrac{1}{2}(\sqrt{s} + s), \end{aligned}$$

where the last equality follows from the fact that the largest 1-norm over the Euclidean unit ball is $\sqrt{s}$. Moreover, one can check that the specific, sequentially defined mapping

$$\begin{aligned} y(\xi)_1 &:= \max\{\xi_1, 1-\xi_1\} \\ y(\xi)_2 &:= \max\{\xi_2, 1-\xi_2\} + y(\xi)_1 \\ &\vdots \\ y(\xi)_s &:= \max\{\xi_s, 1-\xi_s\} + y(\xi)_{s-1} \end{aligned}$$

is feasible with objective value $\frac{1}{2}(\sqrt{s}+s)$. So $v^*_{\text{RLP},2} \leq \frac{1}{2}(\sqrt{s}+s)$, and this completes the argument that $v^*_{\text{RLP},2} = \frac{1}{2}(\sqrt{s}+s)$. Overall, we find that each $j = 1, 2$ yields a class of problems with arbitrarily large gaps between the true robust adjustable and affine-policy values.

Using the similar change of variables

$$u := (1, u_2, \ldots, u_{s+1})^T = (1, 2\xi_1 - 1, \ldots, 2\xi_s - 1)^T \in \mathbb{R}^{s+1},$$

for each Ξ_j, we may cast (13) in the form of (RLP) by setting $x = 0$, defining

$$m = 2s, \quad k = s+1, \quad n_2 = s,$$

and taking $\widehat{\mathcal{U}}_j$ to be the k-dimensional cone associated with the j-norm. For convenience, we continue to use s in the following discussion, but we will remind the reader of the relationships between s, m, k, and n_2 as necessary (e.g., $s = m/2$). We also set

$$d = (0, \ldots, 0, 1)^T \in \mathbb{R}^s,$$

$$B = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix} \in \mathbb{R}^{2s \times s}, \quad F = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & -1 \end{pmatrix} \in \mathbb{R}^{2s \times (s+1)}.$$

Furthermore,

$$\widehat{\mathcal{U}}_2 := \{u \in \mathbb{R}^{s+1} : \|(u_2, \ldots, u_{s+1})^T\| \le u_1\}$$

is the second-order cone, and

$$\widehat{\mathcal{U}}_1 := \{u \in \mathbb{R}^{s+1} : Pu \ge 0\},$$

where each row of $P \in \mathbb{R}^{2^s \times (s+1)}$ has the following form: $(1, \pm 1, \ldots, \pm 1)$. That is, each row is an $(s+1)$-length vector with a 1 in its first position and some combination of $+1$'s and -1's in the remaining s positions. Note that the size of P is exponential in s. Using extra nonnegative variables, we could also represent $\widehat{\mathcal{U}}_1$ as the projection of a cone with size polynomial in s, and all of the subsequent discussion would still apply. In other words, the exact representation of $\widehat{\mathcal{U}}_1$ is not so relevant to our discussion here; we choose the representation $Pu \ge 0$ in the original space of variables for convenience.

It is important to note that, besides $\widehat{\mathcal{U}}_1$ and $\widehat{\mathcal{U}}_2$, all other data required for representing (13) in the form of (RLP), such as the matrices B and F, do not depend on j. Assumptions 1–3 clearly hold, and the following proposition shows that (13) also satisfies Assumption 4:

Proposition 5. *For (13) and its formulation as an instance of (RLP), $\mathcal{W}$ is nonempty and bounded.*

Proof. The system $B^T w = d$ is equivalent to the $2s - 1$ equations $w_1 + w_2 = 1$, $w_2 + w_3 = 1, \cdots, w_{2s-1} + w_{2s} = 1$. It is thus straightforward to check that $\mathcal{W}$ is nonempty and bounded. □

5.1.1 The case $j = 2$

Let us focus on the case $j = 2$; we continue to make use of the subscript 2. Recall $v^*_{\mathrm{RLP},2} = \frac{1}{2}(\sqrt{s} + s)$, and consider problem ($\mathrm{IB}_2$) with $\mathrm{IB}(\widehat{\mathcal{U}}_2 \times \mathbb{R}^{2s}_+)$ built as described for the second-order cone at the end of Section 4. We employ the equivalent formulation (9) of ($\overline{\mathrm{RLP}}$), setting $x = 0$ and replacing $\mathrm{COP}(\widehat{\mathcal{U}}_2 \times \mathbb{R}^{2s}_+)$ by $\mathrm{IB}(\widehat{\mathcal{U}}_2 \times \mathbb{R}^{2s}_+)$:

$$\begin{array}{rll} v^*_{\mathrm{IB},2} = & \min & \lambda + r\rho \\ & \text{s.t.} & \lambda g_1 g_1^T - \frac{1}{2}G(0) + E^T \operatorname{Diag}(v) E + \rho I \in \mathrm{IB}(\widehat{\mathcal{U}}_2 \times \mathbb{R}^{2s}_+) \\ & & \rho \geq 0. \end{array} \tag{14}$$

Note that the dimension of g_1 is $k + m = (s + 1) + 2s = 3s + 1$.

We know $v^*_{\mathrm{RLP},2} \leq v^*_{\mathrm{IB},2}$ by Theorem 3. Substituting the definition of $\mathrm{IB}(\widehat{\mathcal{U}}_2 \times \mathbb{R}^{2s}_+)$ from Section 4, using the fact that $\widehat{\mathcal{U}}_2^* = \widehat{\mathcal{U}}_2$, and simplifying, we have

$$\begin{array}{rll} v^*_{\mathrm{IB},2} = & \min & \lambda + r\rho \\ & \text{s.t.} & \rho I + \lambda g_1 g_1^T - \frac{1}{2}G(0) + E^T \operatorname{Diag}(v) E - S - R \succeq 0 \\ & & \rho \geq 0,\ S_{11} = e_1 \alpha^T + \alpha e_1^T,\ \alpha \in \widehat{\mathcal{U}}_2,\ S_{22} \geq 0,\ \operatorname{Rows}(S_{21}) \in \widehat{\mathcal{U}}_2 \\ & & R_{11} = \tau J,\ \tau \geq 0,\ R_{21} = 0,\ R_{22} = 0. \end{array} \tag{15}$$

We will show that, for every $\rho > 0$, (15) has a feasible solution with objective value $v^*_{\mathrm{RLP},2} + r\rho$. Then, by letting $\rho \to 0$, we conclude that $v^*_{\mathrm{IB},2} \leq v^*_{\mathrm{RLP},2}$, which in turn establishes that $v^*_{\mathrm{IB},2} = v^*_{\mathrm{RLP},2}$, i.e., that our relaxation is in fact exact.

For fixed $\rho > 0$, let us construct the claimed feasible solution. Set

$$\lambda = v^*_{\mathrm{RLP},2} = \tfrac{1}{2}(\sqrt{s} + s), \quad \alpha = 0, \quad \tau = \tfrac{1}{4}\sqrt{s}, \quad S_{21} = 0,$$

and

$$S_{22} = \frac{1}{2\sqrt{s}} \sum_{i=1}^{s} \left(f_{2i} f_{2i-1}^T + f_{2i-1} f_{2i}^T \right) \geq 0,$$

where f_j denotes the j-th coordinate vector in $\mathbb{R}^m = \mathbb{R}^{2s}$. Note that clearly $\alpha \in \widehat{\mathcal{U}}_2$ and $\operatorname{Rows}(S_{21}) \in \widehat{\mathcal{U}}_2$. Also forcing $v = \mu \mathbb{1}_k$ for a single scalar variable μ, where $\mathbb{1}_k$ is the all-ones vector of size $k = s + 1$, the feasibility constraints of (15) simplify further to

$$\rho I + \begin{pmatrix} \frac{1}{2}(s + \sqrt{s}) e_1 e_1^T - \frac{1}{4}\sqrt{s} J & -\frac{1}{2}F^T \\ -\frac{1}{2}F & -S_{22} \end{pmatrix} + \mu E^T E \succeq 0, \tag{16}$$

where $e_1 \in \mathbb{R}^k = \mathbb{R}^{s+1}$ is the first coordinate vector. For compactness, we write

$$V := \begin{pmatrix} \frac{1}{2}(s+\sqrt{s})e_1e_1^T - \frac{1}{4}\sqrt{s}J & -\frac{1}{2}F^T \\ -\frac{1}{2}F & -S_{22} \end{pmatrix} \tag{17}$$

so that (16) reads $\rho I + V + \mu E^T E \succeq 0$. We next claim that, given ρ, V, and E, μ can be chosen so that (16) is indeed satisfied, which we prove in the Appendix. By the discussion above, it follows that indeed $v^*_{\mathrm{IB},2} = v^*_{\mathrm{RLP},2}$ for the instance (13) of (RLP) based on Ξ_2.

For completeness—and also to facilitate Section 5.1.2 next—we construct the corresponding optimal solution of the dual of (14), which can be derived from (5) by setting $x = 0$, adding the redundant constraint $I \bullet Z \le r$, and replacing $\mathrm{CPP}(\widehat{\mathcal{U}}_2 \times \mathbb{R}^{2s}_+)$ by its relaxation $\mathrm{IB}(\widehat{\mathcal{U}}_2 \times \mathbb{R}^{2s}_+)^*$, the dual cone of $\mathrm{IB}(\widehat{\mathcal{U}}_2 \times \mathbb{R}^{2s}_+)$. Specifically, the dual is

$$\begin{array}{rll} v^*_{\mathrm{IB},2} = & \max & F \bullet Z_{21} \\ & \text{s.t.} & \mathrm{diag}(EZE^T) = 0, \ I \bullet Z \le r \\ & & J \bullet Z_{11} \ge 0, \ Z_{11}e_1 \in \widehat{\mathcal{U}}_2, \ Z_{22} \ge 0, \ \mathrm{Rows}(Z_{21}) \in \widehat{\mathcal{U}}_2 \\ & & Z \succeq 0, \ g_1 g_1^T \bullet Z = 1. \end{array} \tag{18}$$

We construct the optimal solution of (18) to be

$$Z = \frac{1}{4}\left[\begin{pmatrix} 2e_1 \\ \mathbb{1}_m \end{pmatrix}\begin{pmatrix} 2e_1 \\ \mathbb{1}_m \end{pmatrix}^T + \sum_{i=1}^{s} \begin{pmatrix} \frac{2}{\sqrt{s}} e_{i+1} \\ f_{2i-1} - f_{2i} \end{pmatrix}\begin{pmatrix} \frac{2}{\sqrt{s}} e_{i+1} \\ f_{2i-1} - f_{2i} \end{pmatrix}^T \right],$$

where each $e_\bullet$ is a coordinate vector in $\mathbb{R}^k = \mathbb{R}^{s+1}$, each $f_\bullet$ is a coordinate vector in $\mathbb{R}^m = \mathbb{R}^{2s}$, and $\mathbb{1}_m \in \mathbb{R}^m$ is the all-ones vector. By construction, Z is positive semidefinite, and one can argue in a straightforward manner that

$$Z_{11} = \mathrm{Diag}(1, \tfrac{1}{s}, \ldots, \tfrac{1}{s}), \quad Z_{22} = \frac{1}{4}\left(I + \mathbb{1}_m \mathbb{1}_m^T - \sum_{i=1}^{s} (f_{2i} f_{2i-1}^T + f_{2i-1} f_{2i}^T) \right),$$

and

$$Z_{21} = \frac{1}{2}\begin{pmatrix} 1 & \frac{1}{\sqrt{s}} & 0 & \cdots & 0 \\ 1 & -\frac{1}{\sqrt{s}} & 0 & \cdots & 0 \\ 1 & 0 & \frac{1}{\sqrt{s}} & \cdots & 0 \\ 1 & 0 & -\frac{1}{\sqrt{s}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & \frac{1}{\sqrt{s}} \\ 1 & 0 & 0 & \cdots & -\frac{1}{\sqrt{s}} \end{pmatrix}.$$

Then Z clearly satisfies $g_1 g_1^T \bullet Z = 1$, $Z_{11} e_1 \in \widehat{\mathcal{U}}_2$, $J \bullet Z_{11} \geq 0$, $Z_{22} \geq 0$, and $\text{Rows}(Z_{21}) \in \widehat{\mathcal{U}}_2$. Furthermore, the constraint $I \bullet Z \leq r$ is easily satisfied for sufficiently large r. To check the constraint $\text{diag}(EZE^T) = 0$, it suffices to verify $EZ = 0$, which amounts to two equations. First,

$$0 = E \begin{pmatrix} 2e_1 \\ \mathbb{1}_m \end{pmatrix} = -2\, d e_1^T e_1 + B^T \mathbb{1}_m = -2d + 2d = 0,$$

and second, for each $i = 1, \ldots, s$,

$$0 = E \begin{pmatrix} \frac{2}{\sqrt{s}} e_{i+1} \\ f_{2i-1} - f_{2i} \end{pmatrix} = -\frac{2}{\sqrt{s}}\, d e_1^T e_{i+1} + B^T (f_{2i-1} - f_{2i}) = 0 + B^T f_{2i-1} - B^T f_{2i} = 0.$$

So the proposed Z is feasible. Finally, it is clear that the corresponding objective value is $F \bullet Z_{21} = \frac{1}{2}(\sqrt{s} + s)$. So Z is indeed optimal.

5.1.2 The case $j = 1$

Recall that Ξ_1 is properly contained in Ξ_2. So $v^*_{\text{RLP},1}$ cannot exceed $v^*_{\text{RLP},2}$ due to its smaller uncertainty set. In fact, as discussed above, we have $\frac{1}{2}(\sqrt{s}+1) = v^*_{\text{RLP},1} < v^*_{\text{RLP},2} = \frac{1}{2}(\sqrt{s}+s)$ and $v^*_{\text{Aff},1} = v^*_{\text{Aff},2} = s$. In this subsection, we further exploit the inclusion $\Xi_1 \subseteq \Xi_2$ and the results of the previous subsection (case $j = 2$) to prove that, for the particular tightening $\text{IB}(\widehat{\mathcal{U}}_1 \times \mathbb{R}^{2s}_+)$ proposed at the end of Section 4, we have $v^*_{\text{RLP},1} < v^*_{\text{IB},1} = \frac{1}{2}(\sqrt{s} + s) < v^*_{\text{Aff},1}$. In other words, the case $j = 1$ provides an example in which our approach improves the affine value but does not completely close the gap with the robust value.

The inclusion $\Xi_1 \subseteq \Xi_2$ implies $\widehat{\mathcal{U}}_1 \subseteq \widehat{\mathcal{U}}_2$ and $\text{CPP}(\widehat{\mathcal{U}}_1 \times \mathbb{R}^{2s}_+) \subseteq \text{CPP}(\widehat{\mathcal{U}}_2 \times \mathbb{R}^{2s}_+)$. Hence, $\text{COP}(\widehat{\mathcal{U}}_1 \times \mathbb{R}^{2s}_+) \supseteq \text{COP}(\widehat{\mathcal{U}}_2 \times \mathbb{R}^{2s}_+)$. Moreover, it is not difficult to see that the construction of $\text{IB}(\widehat{\mathcal{U}}_1 \times \mathbb{R}^{2s}_+)$ introduced at the end of Section 4 for the polyhedral cone $\widehat{\mathcal{U}}_1$ satisfies $\text{IB}(\widehat{\mathcal{U}}_1 \times \mathbb{R}^{2s}_+) \supseteq \text{IB}(\widehat{\mathcal{U}}_2 \times \mathbb{R}^{2s}_+)$. Thus, we conclude $v^*_{\text{IB},1} \leq v^*_{\text{IB},2} = \frac{1}{2}(\sqrt{s} + s)$.

We finally show $v^*_{\text{IB},1} \geq v^*_{\text{IB},2}$. Based on the definition of $\widehat{\mathcal{U}}_1$ using the matrix P, similar to (18) the corresponding dual problem is

$$\begin{aligned} v^*_{\text{IB},1} = \quad \max \quad & F \bullet Z_{21} \\ \text{s.t.} \quad & \text{diag}(EZE^T) = 0,\ I \bullet Z \leq r \\ & PZ_{11} e_1 \geq 0,\ PZ_{11} P^T \geq 0,\ Z_{22} \geq 0,\ PZ_{21}^T \geq 0 \\ & Z \succeq 0,\ g_1 g_1^T \bullet Z = 1. \end{aligned} \tag{19}$$

To complete the proof, we claim that the specific Z detailed in the previous subsection is also feasible for (19). It remains to show that $PZ_{11} e_1 \geq 0$, $PZ_{11} P^T \geq 0$, and $PZ_{21}^T \geq 0$.

Recall that $Z_{11} = \text{Diag}(1, \frac{1}{s}, \ldots, \frac{1}{s})$ and every row of P has the form $(1, \pm 1, \ldots, \pm 1)$.

Clearly, we have $PZ_{11}e_1 \geq 0$. Moreover, each entry of $PZ_{11}P^T$ can be expressed as $\binom{1}{\alpha}^T Z_{11}\binom{1}{\beta}$ for some $\alpha, \beta \in \mathbb{R}^s$ each of the form $(\pm 1, \ldots, \pm 1)$. We have

$$\binom{1}{\alpha}^T Z_{11}\binom{1}{\beta} = 1 + \tfrac{1}{s} \cdot \alpha^T \beta \geq 1 + \tfrac{1}{s}(-s) \geq 0.$$

So indeed $PZ_{11}P^T \geq 0$. To check $PZ_{21}^T \geq 0$, recall also that every column of Z_{21}^T has the form $\frac{1}{2}(e_1 \pm \frac{1}{\sqrt{s}} e_{i+1})$ for $i = 1, \ldots, s$, where $e_\bullet$ is a coordinate vector in $\mathbb{R}^k = \mathbb{R}^{s+1}$. Then each entry of $2PZ_{21}^T$ can be expressed as

$$\binom{1}{\alpha}^T e_1 \pm \tfrac{1}{\sqrt{s}} \binom{1}{\alpha}^T e_{i+1} \geq 1 - \tfrac{1}{\sqrt{s}} > 0.$$

So $PZ_{21}^T \geq 0$, as desired.

5.2 Lot-sizing problem on a network

We next consider a network lot-sizing problem derived from section 5 of [11] for which the mathematical formulation is:

$$\begin{array}{lll} \min\limits_{x, y(\cdot)} & c^T x + \max\limits_{\xi \in \Xi} \sum_{i=1}^N \sum_{j=1}^N t_{ij} y(\xi)_{ij} & \\ \text{s. t.} & x_i + \sum_{j=1}^N y(\xi)_{ji} - \sum_{j=1}^N y(\xi)_{ij} \geq \xi_i & \forall\, \xi \in \Xi,\ i = 1, \ldots, N \\ & y(\xi)_{ij} \geq 0 & \forall\, \xi \in \Xi,\ i, j = 1, \ldots, N \\ & 0 \leq x_i \leq V_i & \forall\, i = 1, \ldots, N, \end{array}$$

where N is the number of locations in the network, x denotes the first-stage stock allocations, $y(\xi)_{ij}$ denotes the second-stage shipping amounts from location i to location j, and the uncertainty set is the ball $\Xi := \{\xi : \|\xi\| \leq \Gamma\}$ for a given radius Γ. (The paper [11] uses a polyhedral uncertainty set, which we will also discuss below.) The vector c consists of the first-stage costs, the t_{ij} are the second-stage transportation costs for all location pairs, and V_i represents the capacity of store location i. We refer the reader to [11] for a full description.

Consistent with [11], we consider an instance with $N = 8$, $\Gamma = 10\sqrt{N}$, each $V_i = 20$, and each $c_i = 20$. We randomly generate the positions of the N locations from $[0, 10]^2$ in the plane. Then we set t_{ij} to be the (rounded) Euclidean distances between all pairs of locations; see Table 1.

Omitting the details, we reformulate this problem as an instance of (RLP), and we calculate $v^*_{\text{LB}} = 1573.8$ (using the Monte Carlo sampling procedure mentioned in the Introduction) and $v^*_{\text{Aff}} = 1950.8$. It is also easy to see that Assumption 1 holds, and the existence of an affine policy implies that Assumption 2 holds. Moreover, Assumption 3 holds because the

Location j	Location i 1	2	3	4	5	6	7	8
1	0	4	3	2	2	2	3	5
2	4	0	6	5	4	4	2	8
3	3	6	0	1	5	2	6	2
4	2	5	1	0	4	1	4	3
5	2	4	5	4	0	4	2	7
6	2	4	2	1	4	0	4	4
7	3	2	6	4	2	4	0	7
8	5	8	2	3	7	4	7	0

Table 1: Unit transportation costs t_{ij} associated with pairs of locations

original objective value above is clearly bounded below by 0. As mentioned at the end of Section 2, whether or not Assumption 4 holds, in practice we can still use our approach to calculate bounds. Indeed, we solve (IB) with the approximating cone $\text{IB}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$ defined in Section 4, where $\widehat{\mathcal{U}}$ is the second-order cone, and obtain $v^*_{\text{IB}} = 1794.0$, which closes the gap significantly. The first-stage allocations given by the affine policy and our approach, respectively, are

$$x^*_{\text{Aff}} \approx (9.097, 11.246,\ 9.516, 8.320, 10.384, 9.493, 10.211, 12.316),$$
$$x^*_{\text{IB}} \approx (0.269, 16.447, 15.328, 0.091, 18.124, 0.375,\ 9.951, 19.934).$$

Letting other data remain the same, we also ran tests on a budget uncertainty set $\Xi := \{\xi : 0 \le \xi \le \hat{\xi} e,\ e^T \xi \le \Gamma\}$, where $\hat{\xi} = 20$ and $\Gamma = 20\sqrt{N}$, which is consistent with [11]. We found that, in this case, our method did not perform better than the affine policy.

5.3 Randomly generated instances

Finally, we used the same method presented in [25] to generate random instances of (RLP) with $(k, m, n_1, n_2) = (17, 16, 3, 5)$, $\mathcal{X} = \mathbb{R}^{n_1}$, $\mathcal{U}$ equal to the unit ball, and $\widehat{\mathcal{U}}$ equal to the second-order cone. Specifically, the instances are generated as follows: (i) the elements of A and B are independently and uniformly sampled in $[-5, 5]$; (ii) the rows of F are uniformly sampled in $[-5, 5]$ such that each row is in $-\widehat{\mathcal{U}}^* = -\widehat{\mathcal{U}}$ guaranteeing $Fu \le 0$ for all $u \in \mathcal{U}$; and (iii) a random vector $\mu \in \mathbb{R}^m$ is repeatedly generated according to the uniform distribution on $[0, 1]^m$ until $c := A^T \mu \ge 0$ and $d := B^T \mu \ge 0$. Note that, by definition, $\mu \in \mathcal{W}$.

Clearly Assumption 1 is satisfied. In addition, we can see that Assumption 2 is true as follows. Consider $x = 0$ and set $y(\cdot)$ to be the zero map, i.e., $y(u) = 0$ for all $u \in \mathcal{U}$. Then $Ax + By(u) \ge Fu$ for all u if and only $0 \ge Fu$ for all u, which has been guaranteed by construction. Finally, Assumption 3 holds due to the following chain, where $\pi(x)$ is defined as at the beginning of Section 2:

$$\begin{aligned}
c^T x + \pi(x) &= c^T x + \max_{u \in \mathcal{U}} \max_{w \in \mathcal{W}} (Fu - Ax)^T w \\
&\geq c^T x + \max_{u \in \mathcal{U}} (Fu - Ax)^T \mu = c^T x - (Ax)^T \mu + \max_{u \in \mathcal{U}} (Fu)^T \mu \\
&= (c - A^T \mu)^T x + \max_{u \in \mathcal{U}} (Fu)^T \mu = 0^T x + \max_{u \in \mathcal{U}} (Fu)^T \mu \\
&> -\infty.
\end{aligned}$$

We do not know if Assumption 4 necessarily holds for this construction, but as mentioned at the end of Section 2, our approximations still hold even if Assumption 4 does not hold.

For 1,000 generated instances, we computed v^*_{Aff}, the lower bound v^*_{LB} from the sampling procedure of the Introduction, and our bound v^*_{IB} using the the approximating cone $\text{IB}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$ defined in Section 4, where $\widehat{\mathcal{U}}$ is the second-order cone. Of all 1,000 instances, 971 have $v^*_{\text{LB}} \leq v^*_{\text{IB}} = v^*_{\text{Aff}}$, while the remaining 29 have $v^*_{\text{LB}} < v^*_{\text{IB}} < v^*_{\text{Aff}}$. For those 29 instances with a positive gap, the average relative gap closed is 20.2%, where

$$\text{relative gap closed} := \frac{v^*_{\text{Aff}} - v^*_{\text{IB}}}{v^*_{\text{Aff}} - v^*_{\text{LB}}} \times 100\%.$$

6 Future Directions

In this paper, we have provided a new perspective on the two-stage problem (RLP). It would be interesting to study tighter inner approximations $\text{IB}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$ of $\text{COP}(\widehat{\mathcal{U}} \times \mathbb{R}^m_+)$ or to pursue other classes of problems, such as the one described in Section 5.1, for which our approach allows one to establish the tractability of (RLP). A significant open question for our approach—one which we have not been able to resolve—is whether the copositive approach corresponds to enforcing a particular class of policies $y(\cdot)$. For example, the paper [13] solves (RLP) by employing polynomial policies, but the form of our "copositive policies" is unclear even though we have proven they are rich enough to solve (RLP). A related question is how to extract a specific policy $y(\cdot)$ from the solution of the approximation (IB).

References

[1] B.-T. Aharon, G. Boaz, and S. Shimrit. Robust multi-echelon multi-period inventory control. *European Journal of Operational Research*, 199(3):922–935, 2009.

[2] K. M. Anstreicher. Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming. *Journal of Global Optimization*, 43(2-3):471–484, 2009.

[3] M. ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 8.0.*, 2016.

[4] A. Ardestani-Jaafari and E. Delage. The value of flexibility in robust location-transportation problem. *Les Cahiers du GERAD G–2014–83, GERAD, HEC Montréal*, 2014.

[5] A. Ardestani-Jaafari and E. Delage. Linearized robust counterparts of two-stage robust optimization problem with applications in operations management. Manuscript, HEC Montreal, 2016.

[6] A. Atamtürk and M. Zhang. Two-stage robust network flow and design under demand uncertainty. *Operations Research*, 55(4):662–673, 2007.

[7] A. Ben-Tal, B. Do Chung, S. R. Mandala, and T. Yao. Robust optimization for emergency logistics planning: Risk mitigation in humanitarian relief supply chains. *Transportation research part B: methodological*, 45(8):1177–1189, 2011.

[8] A. Ben-Tal, B. Golany, A. Nemirovski, and J.-P. Vial. Retailer-supplier flexible commitments contracts: a robust optimization approach. *Manufacturing & Service Operations Management*, 7(3):248–271, 2005.

[9] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Math. Program., Ser. A*, 99:351–376, 2004.

[10] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations Research Letters*, 25(1):1–14, 1999.

[11] D. Bertsimas and F. J. de Ruiter. Duality in two-stage adaptive linear optimization: Faster computation and stronger bounds. *INFORMS Journal on Computing*, 28(3):500–511, 2016.

[12] D. Bertsimas, V. Goyal, and P. Y. Lu. A tight characterization of the performance of static solutions in two-stage adjustable robust linear optimization. *Mathematical Programming Ser. A*, 150(2):281–319, 2014.

[13] D. Bertsimas, D. A. Iancu, and P. A. Parrilo. A hierarchy of near-optimal policies for multistage adaptive optimization. *IEEE Transactions on Automatic Control*, 56(12):2809–2824, 2011.

[14] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng. Adaptive robust optimization for the security constrained unit commitment problem. *IEEE transactions on power systems*, 28(1):52–63, 2013.

[15] S. Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. *Mathematical Programming Series A*, 120(2):479–495, September 2009.

[16] S. Burer. Copositive programming. In M. Anjos and J. Lasserre, editors, *Handbook of Semidefinite, Cone and Polynomial Optimization: Theory, Algorithms, Software and Applications*, International Series in Operational Research and Management Science, pages 201–218. Springer, 2011.

[17] S. Burer. A gentle, geometric introduction to copositive optimization. *Mathematical Programming*, 151(1):89–116, 2015.

[18] S. Burer and H. Dong. Representing quadratically constrained quadratic programs as generalized copositive programs. *Operations Research Letters*, 40:203–206, 2012.

[19] X. Chen and Y. Zhang. Uncertain linear programs: Extended affinely adjustable robust counterparts. *Operations Research*, 57(6):1469–1482, 2009.

[20] E. Delage and D. A. Iancu. *Robust Multistage Decision Making*, chapter 2, pages 20–46. INFORMS, 2015.

[21] S. H. H. Doulabi, P. Jaillet, G. Pesant, and L.-M. Rousseau. Exploiting the structure of two-stage robust optimization models with integer adversarial variables. Manuscript, MIT, 2016.

[22] G. Eichfelder and J. Jahn. Set-semidefinite optimization. *Journal of Convex Analysis*, 15:767–801, 2008.

[23] R. J. Fonseca and B. Rustem. International portfolio management with affine policies. *European Journal of Operational Research*, 223(1):177–187, 2012.

[24] V. Gabrel, M. Lacroix, C. Murat, and N. Remli. Robust location transportation problems under uncertain demands. *Discrete Applied Mathematics*, 164:100–111, 2014.

[25] M. J. Hadjiyiannis, P. J. Goulart, and D. Kuhn. A scenario approach for estimating the suboptimality of linear decision rules in two-stage robust optimization. In *50th IEEE conference on decision and control and European control conference (CDC-ECC)*, Orlando, FL, USA, December 12-15 2011.

[26] H. Konno. A cutting plane algorithm for solving bilinear programs. *Math. Program.*, 11:14–27, 1976.

[27] J. Lofberg. Yalmip: A toolbox for modeling and optimization in matlab. In *Computer Aided Control Systems Design, 2004 IEEE International Symposium on*, pages 284–289. IEEE, 2004.

[28] M. Poss and C. Raack. Affine recourse for the robust network design problem: Between static and dynamic routing. *Networks*, 61(2):180–198, 2013.

[29] H. D. Sherali and W. P. Adams. *A reformulation-linearization technique for solving discrete and continuous nonconvex problems*, volume 31. Springer Science & Business Media, 2013.

[30] O. Solyalı, J.-F. Cordeau, and G. Laporte. The impact of modeling on robust inventory management under demand uncertainty. *Management Science*, 62(4):1188–1201, 2016.

[31] J. F. Sturm and S. Zhang. On cones of nonnegative quadratic functions. *Mathematics of Operations Research*, 28(2):246–267, 2003.

[32] Q. Wang, J. P. Watson, and Y. Guan. Two-stage robust optimization for nk contingency-constrained unit commitment. *Power Systems, IEEE Transactions on*, 28:2366–2375, 2013.

[33] W. Wiesemann, D. Kuhn, and B. Rustem. Robust resource allocations in temporal networks. *Math. Programm., Ser. A*, 135(1):437–471, 2011.

[34] B. Zeng and L. Zhao. Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Research letters*, 41:457–461, 2013.

[35] L. Zhao and B. Zeng. Robust unit commitment problem with demand response and wind energy. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–8. IEEE, 2012.

[36] L. F. Zuluaga, J. Vera, and J. Pea. LMI approximations for cones of positive semidefinite forms. *SIAM Journal on Optimization*, 16(4):1076–1091, 2006.

Appendix

The existence of μ requires the following lemma:

Lemma 2. *If V is positive semidefinite on the null space of E (that is, $z \in \mathrm{Null}(E) \Rightarrow z^T V z \geq 0$), then there exists $\mu > 0$ such that $\rho I + V + \mu E^T E \succ 0$.*

Proof. We prove the contrapositive. Suppose $\rho I + V + \mu E^T E$ is not positive definite for all $\mu > 0$. In particular, there exists a sequence of vectors $\{z_\ell\}$ such that

$$z_\ell^T(\rho I + V + \ell E^T E) z_\ell \leq 0, \quad \|z_\ell\| = 1.$$

Since $\{z_\ell\}$ is bounded, there exists a limit point $\bar{z}$ such that

$$z_\ell^T(\tfrac{1}{\ell}(\rho I + V) + E^T E) z_\ell \leq 0 \quad \Rightarrow \quad \bar{z}^T E^T E \bar{z} = \|Ez\|^2 \leq 0 \quad \Leftrightarrow \quad \bar{z} \in \mathrm{Null}(E).$$

Furthermore,

$$\begin{aligned} z_\ell^T(\rho I + V) z_\ell \leq -\ell z_\ell^T E^T E z_\ell = -\ell \|E z_\ell\|^2 \leq 0 \quad &\Rightarrow \quad \bar{z}^T(\rho I + V)\bar{z} \leq 0 \\ &\Leftrightarrow \quad \bar{z}^T V \bar{z} \leq -\rho \|\bar{z}\|^2 < 0. \end{aligned}$$

Thus, V is not positive semidefinite on $\mathrm{Null}(E)$. $\square$

With the lemma in hand, it suffices to prove that V is positive semidefinite on $\mathrm{Null}(E)$, a fact which we establish directly.

Recall that $E \in \mathbb{R}^{n_2 \times (k+m)} = \mathbb{R}^{s \times (3s+1)}$. For notational convenience, we partition any $z \in \mathbb{R}^{k+m}$ into $z = \binom{u}{w}$ with $u \in \mathbb{R}^k = \mathbb{R}^{s+1}$ and $w \in \mathbb{R}^m = \mathbb{R}^{2s}$. Then, from the definition of E, we have

$$z = \begin{pmatrix} u \\ w \end{pmatrix} \in \mathrm{Null}(E) \iff \left\{ \begin{array}{l} w_1 + w_2 = w_3 + w_4 \\ w_3 + w_4 = w_5 + w_6 \\ \vdots \\ w_{2s-3} + w_{2s-2} = w_{2s-1} + w_{2s} \\ w_{2s-1} + w_{2s} = u_1 \end{array} \right\}$$
$$\iff w_{2i-1} = u_1 - w_{2i} \quad \forall\, i = 1, \ldots, s.$$

So, taking into account the definition (17) of V,

$$4\, z^T V z = 4 \tbinom{u}{w}^T V \tbinom{u}{w} = u^T \left(2(s + \sqrt{s}) e_1 e_1^T - \sqrt{s} J \right) u - 4\, w^T F u - 4\, w^T S_{22} w,$$

which breaks into the three summands, and we will simplify each one by one. First,

$$\begin{aligned} u^T \left(2(s + \sqrt{s}) e_1 e_1^T - \sqrt{s} J \right) u &= 2(s + \sqrt{s}) u_1^2 - \sqrt{s} u_1^2 + \sqrt{s} \sum_{j=2}^{s+1} u_j^2 \\ &= 2\, s\, u_1^2 + \sqrt{s}\, u^T u. \end{aligned}$$

Second,

$$
\begin{aligned}
-4w^T Fu &= -4\sum_{j=1}^{2s} w_j [Fu]_j = -4\sum_{i=1}^{s} \left(w_{2i-1}[Fu]_{2i-1} + w_{2i}[Fu]_{2i}\right) \\
&= -2\sum_{i=1}^{s} \left(w_{2i-1}(u_1 + u_{i+1}) + w_{2i}(u_1 - u_{i+1})\right) \\
&= -2\sum_{i=1}^{s} \left((w_{2i-1} + w_{2i})u_1 + u_{i+1}(w_{2i-1} - w_{2i})\right) \\
&= -2\sum_{i=1}^{s} \left(u_1^2 + u_{i+1}(w_{2i-1} - w_{2i})\right) \\
&= -2\, s\, u_1^2 - 2\sum_{i=1}^{s} u_{i+1}(w_{2i-1} - w_{2i}) \\
&= -2\, s\, u_1^2 + 2\sum_{i=1}^{s} u_{i+1}(w_{2i} - w_{2i-1}) = -2\, s\, u_1^2 + 2\sum_{i=1}^{s} u_{i+1}(2w_{2i} - u_1).
\end{aligned}
$$

Finally,

$$
\begin{aligned}
-4w^T S_{22} w &= -4w^T \left(\frac{1}{2\sqrt{s}} \sum_{i=1}^{s} \left(f_{2i} f_{2i-1}^T + f_{2i-1} f_{2i}^T\right)\right) w \\
&= -\frac{4}{\sqrt{s}} \sum_{i=1}^{s} w_{2i-1} w_{2i} = -\frac{4}{\sqrt{s}} \sum_{i=1}^{s} (u_1 - w_{2i}) w_{2i}.
\end{aligned}
$$

Combining the three summands, we have as desired

$$
\begin{aligned}
4z^T V z &= \left(2s\, u_1^2 + \sqrt{s}\, u^T u\right) + \left(-2s\, u_1^2 + 2\sum_{i=1}^{s} u_{i+1}(2w_{2i} - u_1)\right) + \left(-\frac{4}{\sqrt{s}} \sum_{i=1}^{s} (u_1 - w_{2i}) w_{2i}\right) \\
&= \sqrt{s}\, u^T u + 2\sum_{i=1}^{s} u_{i+1}(2w_{2i} - u_1) - \frac{4}{\sqrt{s}} \sum_{i=1}^{s} (u_1 - w_{2i}) w_{2i} \\
&= \sum_{i=1}^{s} \left(\frac{1}{\sqrt{s}}\, u_1^2 + \sqrt{s}\, u_{i+1}^2 + 2\, u_{i+1}(2w_{2i} - u_1) - \frac{4}{\sqrt{s}} (u_1 - w_{2i}) w_{2i}\right) \\
&= \sum_{i=1}^{s} \left(\frac{1}{\sqrt{s}}\, u_1^2 - 2\, u_1\, u_{i+1} - \frac{4}{\sqrt{s}}\, u_1\, w_{2i} + \sqrt{s}\, u_{i+1}^2 + 4\, u_{i+1}\, w_{2i} + \frac{4}{\sqrt{s}}\, w_{2i}^2\right) \\
&= \sum_{i=1}^{s} \left(-(s)^{-1/4}\, u_1 + (s)^{1/4}\, u_{i+1} + 2(s)^{-1/4}\, w_{2i}\right)^2 \\
&\geq 0.
\end{aligned}
$$