

## OPTIMIZATION UNDER DECISION-DEPENDENT UNCERTAINTY\*

OMID NOHADANI<sup>†</sup> AND KARTIKEY SHARMA<sup>†</sup>

**Abstract.** The efficacy of robust optimization spans a variety of settings with uncertainties bounded in predetermined sets. In many applications, uncertainties are affected by decisions and cannot be modeled with current frameworks. This paper takes a step towards generalizing robust linear optimization to problems with decision-dependent uncertainties. In general settings, we show these problems to be NP-complete. To alleviate the computational inefficiencies, we introduce a class of uncertainty sets whose size depends on binary decisions. We propose reformulations that improve upon alternative standard linearization techniques. To illustrate the advantages of this framework, a shortest path problem is discussed, where the uncertain arc lengths are affected by decisions. Beyond the modeling and performance advantages, the proposed notion of proactive uncertainty control also mitigates over conservatism of current robust optimization approaches.

**Key words.** robust optimization, endogenous uncertainty, decision-dependent uncertainty

**AMS subject classification.** 90C05

**DOI.** 10.1137/17M1110560

**1. Introduction.** The two well-established approaches of optimization under uncertainty are stochastic and robust optimization. Stochastic optimization (SO) can be used when the distribution of the uncertainty is available [31]. When uncertainties can be regarded as residing in sets, robust optimization (RO) is a computationally attractive alternative [3, 6]. The method of RO has been extended considerably and applied to problems ranging from portfolio management [18], to healthcare [13], electricity systems [25], and engineering design [8].

RO employs uncertainty sets that are predetermined and, hence, *exogenous*. For instance, temporal changes to the uncertainty can be explicitly modeled via time-dependent uncertainty sets [27]. In many real-world problems, however, the uncertainty can be affected by decisions. In such decision-dependent cases, the uncertainty set is *endogenous*. Despite the wide prevalence of such uncertainties in real-world settings, these problems have not received much attention in the past, largely due to computational intractabilities. In this paper, we take a first step towards robust linear optimization problems with endogenous uncertainties and provide a class of uncertainty sets whose reformulations improve on standard techniques. Specifically, we study a single-stage RO problem with decision-dependent uncertainty sets

$$\begin{aligned}
 \text{(RO-DDU)} \quad & \min_{\mathbf{x}, \mathbf{y}} \mathbf{c}^\top \mathbf{x} + \mathbf{f}^\top \mathbf{y} \\
 & \text{s.t. } \mathbf{a}_i^\top \mathbf{x} + \boldsymbol{\xi}_i^\top \mathbf{y} \leq b_i, \quad \forall \boldsymbol{\xi}_i \in \mathcal{U}_i(\mathbf{x}) \subseteq \mathbb{R}^n \quad \forall i = 1, \dots, m,
 \end{aligned}$$

where “s.t.” indicates “subject to” and  $\mathbf{x} \in \mathbb{R}^n$  and  $\mathbf{y} \in \mathbb{R}^n$  represent decision variables, which need to satisfy each constraint  $i = 1, \dots, m$  for every realization from the set  $\mathcal{U}_i(\mathbf{x})$  that bounds the uncertain parameter  $\boldsymbol{\xi}_i$ . Further, the parameters defining  $\mathcal{U}_i(\mathbf{x})$  depend on decisions  $\mathbf{x}$ . We first study the complexity of (RO-DDU) for polyhedral  $\mathcal{U}_i(\mathbf{x})$ . We then assume  $\mathbf{x}$  is binary and provide reformulations for special classes of polyhedral and conic uncertainty sets and conclude with numerical experiments.

\*Received by the editors January 5, 2017; accepted for publication (in revised form) February 26, 2018; published electronically June 14, 2018.

<http://www.siam.org/journals/siopt/28-2/M111056.html>

<sup>†</sup>Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208 (nohadani@northwestern.edu, kartikeysharma2014@u.northwestern.edu).

To show the range of applicability of this model, we illustrate two examples.

*Example 1* (uncertainty reduction). In facility location or inventory management problems with uncertain demand, the uncertainty can be reduced by spending resources to acquire information. Similarly, in healthcare problems, additional medical tests can improve the diagnosis. This type of uncertainty reduction is characteristic of many real-world problems. In order to improve solutions, decisions on uncertainty reduction have to be included in the optimization problem, making the uncertainty a function of decisions on acquiring additional information.

*Example 2* (shortest path on a network). Consider the graph in Figure 1 with the arcset  $\mathcal{A}$  and let the uncertain length for any arc  $e$  be  $d_e = \bar{d}_e(1 + 0.5\xi_e)$ , where  $\bar{d}_e$  denotes the nominal value. The uncertain vector  $\xi$  lies in the uncertainty set  $\mathcal{U}(\mathbf{x}) = \{\xi \mid 0 \leq \xi_e \leq 1 - 0.8x_e \ \forall e, \sum_{e \in \mathcal{A}} \xi_e \leq 1\}$ . The binary decision  $x_e$  determines whether to reduce the maximum possible uncertainty  $\xi_e$  to 0.2 ( $x_e = 1$ ) or leave it at 1 ( $x_e = 0$ ). For simplicity, we assume the reduction to be possible for at most one of the arcs.

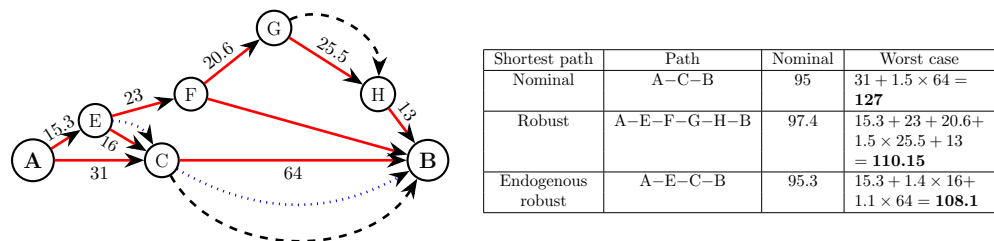


FIG. 1. Shortest path on a network. Nominal lengths are labeled. Worst-case and reduced-case lengths are displayed with dashed and dotted lines. The table shows the lengths in different settings.

Figure 1 displays a network with source node A and destination B. In the worst case, the nominally shortest path lengthens to 127 units. RO optimizes against this case, improving the worst-case length. If it is permitted to reduce the uncertainty of an arc, then A-E-C-B is selected with  $x_{C-B} = 1$  and the worst-case path becomes 108.5. This example demonstrates that decision-dependent sets can be leveraged to model decisions that mitigate the worst-case scenario.

The contributions of this paper can be summarized as follows.

1. We study robust linear optimization problems with a polyhedral decision-dependent set for the uncertain parameters. We prove such problems to be NP-complete. We also show that when decisions that influence the uncertainties are binary, the problem can be reformulated as a mixed integer optimization problem.
2. For binary  $\mathbf{x}$ , we provide a class of uncertainty sets for which a more efficient reformulation of the decision-dependent RO problem is possible. The set structure and the nature of decision dependence are leveraged to provide reformulations with fewer constraints.
3. We provide an improvement to big-M linearization for bilinear terms, which can reduce the number of constraints.

This work also showcases the advantages that can be gained in both stochastic and robust optimization by proactively controlling uncertainties.

We also emphasize what this paper fails to address. Reformulations for continuous decisions influencing the uncertainty are not provided. Furthermore, the primary problem in this paper is a static optimization problem, i.e., the decisions do not adapt to uncertainty realizations. In fact, it is the uncertainty set and the corresponding worst-case realization that are affected by decisions.

Section 3 discusses the complexity of the decision-dependent robust linear optimization problem. Section 4 introduces a class of uncertainty sets which allow improved reformulations. Section 5 provides a comparison to the corresponding big-M formulation. It also provides methods to improve these standard techniques. A numerical experiment is discussed in section 6 to illustrate the advantages of the decision-dependent setting and to computationally compare the three formulations.

*Notation.* Throughout this paper, we use bold lower and uppercase letters to denote vectors and matrices, respectively. Scalars are marked in regular font. All vectors are column vectors and the vector of ones is denoted by  $\mathbf{e}$ . Furthermore,  $\text{diag}(\bullet)$  denotes a diagonal matrix with  $\bullet$  on the diagonal and zeros elsewhere. For any given matrix  $\mathbf{A}$ , the  $i$ th row is denoted by  $\mathbf{A}_{i,\cdot}$  and the  $j$ th column is denoted by  $\mathbf{A}_{\cdot,j}$ . The problems have  $m$  constraints indexed by  $i$ . LHS denotes the left-hand side and RHS denotes the right-hand side. We use the phrases “decision-dependent” and “endogenous” interchangeably. Similarly, we refer to variables affecting an uncertainty set as influence variables.

**2. Background.** In the following, we first review endogenous settings in SO before discussing RO approaches.

The notion of endogenous uncertainty in SO generally corresponds to scenario trees, where decisions determine the probabilities. For example, Jonsbråten, Wets, and Woodruff [23] consider the cost of an item to remain uncertain until it is produced. The probability distribution depends upon which item is to be produced and when. Goel and Grossmann [19] address the problem of offshore oil and gas planning, with the objective of maximizing revenues and investments over a period of time, when the recovery and size of oil fields are not known in advance. They provide a disjunctive formulation that is solved by a decomposition algorithm. This approach is extended to a multistage SO problem for optimal production scheduling that minimizes cost while satisfying the demand for different goods [20]. For package sorting centers, Novoa, Jarrah, and Morton [28] seek to balance the flow across working stations. Capacities are modeled via budgeted uncertainties where the budget is a function of workstation allocation. These and other approaches address endogenous uncertainties probabilistically.

In RO, the endogenous nature of uncertainty is imposed directly on the uncertainty set itself. For example, Spacey et al. [33] address a software partitioning problem, where code segments are assigned to different computing nodes to reduce runtime with uncertain execution order and for unknown frequency of segment calls. They employ tailored decision-dependent uncertainty sets. Such sets also occur as a result of reformulations. For example, Hanasusanto, Kuhn, and Wiesemann [21] use a finite adaptability approximation to adjustable robust optimization (ARO), as introduced by Bertsimas and Caramanis [7], and consider optimization problems with binary recourse decisions. For problems with uncertain objective and constraints, they provide a formulation with decision-dependent uncertainty sets before finally reformulating it as a mixed integer linear programming (MILP) problem. Poss [29, 30] considers combinatorial optimization problems with budgeted uncertainty sets. This extends the work of Bertsimas and Sim [10] to decision-dependent budgets. These works focus on

budget uncertainty sets, with limited discussion on general sets. On the other hand, for a dynamic pricing problem with learning, Bertsimas and Vayanos [12] consider 1 or  $\infty$ -norm uncertainty sets for price-dependent demand. Specifically, the uncertain demand curve is driven by past realizations of price-demand pairs. Since the price is a decision variable, this leads to decision-dependent uncertainty sets. In the context of robust scheduling problems, Vujanic, Goulart, and Morari [34] consider a decision-dependent uncertainty set which is a vector sum of a collection of sets. The sets in the vector combination are selected by a decision which is a part of the original problem. They probe the performance of an affine policy for the problem. More recently, decision-dependent sets were studied in the context of control problems with *primitive* uncertainty sets [36]. Note that in all approaches to date, the decision dependence is modeled in a specific context, often driven by an application.

The journey of RO has also included measures to reduce conservatism. The original RO formulation by Soyster [32] produced over conservative solutions for many applications due to the use of box uncertainties. Later, Ben-Tal and Nemirovski [5] provided less conservative solutions by using general polyhedral and ellipsoid uncertainty sets. ARO models [4] and decision rule approximations took another step in this direction by allowing decisions to depend on the realizations [22, 17]. In this vein, decision-dependent uncertainty sets offer a new avenue to reduce the level of conservatism. For example, Poss [29] decreases it for cardinality constrained sets. This work also motivates the notion of *proactive uncertainty control* by using decision-dependent sets to enable deliberate uncertainty reduction.

**3. General decision dependence.** Robust linear optimization problems encompass a wide variety of applications, in portfolio optimization, healthcare, inventory management, and routing, amongst others. The tractability of robust linear programs provides a suitable starting point to analyze the complexity of RO problems with decision-dependent uncertainty. Here, we investigate the robust linear optimization problem (RO-DDU). The underlying uncertainty set is endogenous and defined as follows.

**DEFINITION 3.1.** *The set with constraint matrix  $\mathbf{D}$ , constant vector  $\mathbf{d}$ , and decision coefficient matrix  $\Delta$  given by*

$$\mathcal{U}^P(\mathbf{x}) = \{\boldsymbol{\xi} \mid \mathbf{D}\boldsymbol{\xi} \leq \mathbf{d} + \Delta\mathbf{x}\}$$

*is a polyhedral uncertainty set with affine decision dependence.*

Note that  $\Delta$  determines the influence of  $\mathbf{x}$  on the set and can be estimated from the data or from the context of an application. In section 6, we quantify it for a specific application.

The following theorem shows that RO problems with decision-dependent sets cannot be reformulated in a tractable fashion, a departure from standard RO problems. This occurs despite the fact that linear programs with polyhedral uncertainty sets have tractable robust counterparts.

**THEOREM 3.2.** *The robust linear problem (RO-DDU) with uncertainty set  $\mathcal{U}^P$  is NP-complete.*

*Proof.* The proof proceeds by way of the following steps:

1. Consider an instance of the 3-satisfiability problem (3-SAT) for a set of literals  $N = \{1, 2, \dots, n\}$  and  $m$  clauses, which seeks to find a solution  $\mathbf{x} \in \{0, 1\}^n$

that satisfies

$$x_{i_1} + x_{i_2} + (1 - x_{i_3}) \geq 1 \text{ for } m \text{ clauses and } i_1, i_2, i_3 \in \{1, \dots, n\}.$$

2. Consider the following special case of (RO-DDU) with  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{y} \in \mathbb{R}^m$ ,  $z \in \mathbb{R}$ :  
(RO-SAT)

$$\min_{\mathbf{x}, \mathbf{y}, z \geq 0} \left\{ -z \mid z - \boldsymbol{\xi}^\top \mathbf{y} \leq 0, \forall \boldsymbol{\xi} \in \mathcal{U}(\mathbf{x}), \mathbf{x}, \mathbf{y} \leq \mathbf{e}, -\mathbf{y} \leq -\mathbf{e} \right\},$$

where  $\mathcal{U}(\mathbf{x}) = \{(\xi_1, \dots, \xi_m) \mid \xi_i \geq x_{i_1}, \xi_i \geq x_{i_2}, \xi_i \geq 1 - x_{i_3}, \xi_i \leq 1\}$ . Note that the 3-SAT problem is embedded in this set.

3. By Lemma 3.3 (provided after these steps), the optimal value of (RO-SAT) is  $-m$  if and only if the 3-SAT problem has a solution.
4. Problem (RO-SAT) is a special case of (RO-DDU) with polyhedral set  $\mathcal{U}(\mathbf{x})$ .
5. Since the 3-SAT problem is NP-complete [14], problem (RO-DDU) is also NP-complete.  $\square$

LEMMA 3.3. *The 3-SAT problem has a feasible solution  $\mathbf{x}$  if and only if problem (RO-SAT) has an optimal value of at most  $-m$ .*

*Proof.* ( $\implies$ ) Suppose the 3-SAT problem has a feasible solution  $\mathbf{x}$ . This means  $\mathbf{x}$  has to satisfy

$$x_{i_1} + x_{i_2} + (1 - x_{i_3}) \geq 1 \quad \forall i = 1, \dots, m.$$

Since  $\mathbf{x} \in \{0, 1\}^n$ , for each  $i$  at least one of  $x_{i_1}$ ,  $x_{i_2}$ , or  $1 - x_{i_3}$  must be equal to 1. Now, consider the uncertainty set

$$\mathcal{U}(\mathbf{x}) = \{(\xi_1, \dots, \xi_m) \mid \xi_i \geq x_{i_1}, \xi_i \geq x_{i_2}, \xi_i \geq 1 - x_{i_3}, \xi_i \leq 1 \quad \forall i = 1, \dots, m\}.$$

Since at least one of  $x_{i_1}$ ,  $x_{i_2}$ , or  $1 - x_{i_3}$  equals 1,  $\xi_i$  satisfies  $\xi_i \geq 1$ . This implies that  $\xi_i = 1 \quad \forall i$  is the only point in  $\mathcal{U}(\mathbf{x})$ . For this uncertainty set, the feasible solution is  $\mathbf{x}, \mathbf{y} = \mathbf{1}$ ,  $z = m$ . This leads to the optimal solution  $-z \leq -m$  or  $z \geq m$ .

( $\impliedby$ ) Suppose (RO-SAT) has an optimal solution  $(\mathbf{x}^*, \mathbf{y}^*)$  with the objective value of  $-z^* \leq -m$ . We first show that strict inequality is not possible. Assume  $-z^* < -m$ . The constraints in (RO-SAT) imply  $z^* - \boldsymbol{\xi}^\top \mathbf{y}^* \leq 0$ , i.e.,  $\boldsymbol{\xi}^\top \mathbf{y}^* \geq z^* > m \quad \forall \boldsymbol{\xi} \in \mathcal{U}(\mathbf{x}^*)$ . The constraints also imply  $y_i^* = 1 \quad \forall i$ . This means that  $\sum_{i=1}^m \xi_i > m \quad \forall \boldsymbol{\xi} \in \mathcal{U}(\mathbf{x}^*)$ . However, the construction of the uncertainty set yields  $\xi_i \leq 1$ . This leads to a contradiction because  $\sum_{i=1}^m \xi_i \not> m$ , and hence  $-z^* = -m$ . Thus,  $\boldsymbol{\xi}^\top \mathbf{y}^* = m \quad \forall \boldsymbol{\xi} \in \mathcal{U}(\mathbf{x}^*)$ . Therefore, we can write  $\sum_{i=1}^m \xi_i = m \quad \forall \boldsymbol{\xi} \in \mathcal{U}(\mathbf{x}^*)$ , which implies  $\min_{\boldsymbol{\xi} \in \mathcal{U}(\mathbf{x}^*)} \sum_{i=1}^m \xi_i = m$ . However, since the uncertainty set implies  $\xi_i \leq 1 \quad \forall i$ , we can conclude that the sum can only be equal to  $m$  if  $\xi_i = 1 \quad \forall i$ .

We now show that this result implies for each  $i$  at least one of  $x_{i_1}^*$ ,  $x_{i_2}^*$ , or  $(1 - x_{i_3}^*)$  is equal to 1. Suppose this is not true. This implies  $\exists i$  for which  $x_{i_1}^* < 1$ ,  $x_{i_2}^* < 1$  and  $(1 - x_{i_3}^*) < 1$ . That means that we can construct a  $\xi'_i = \max\{x_{i_1}^*, x_{i_2}^*, (1 - x_{i_3}^*)\}$  which is an element of the uncertainty set and  $\xi'_i < 1$ . However, this contradicts the result of  $\xi_i = 1 \quad \forall i$ . Therefore, if  $z^* = m$ , then we can find a feasible solution for the 3-SAT problem.  $\square$

Although problem (RO-DDU) is NP-complete, it can be reformulated as a bilinear or biconvex program, which may be solved by global optimization techniques (see, e.g., [24]). For binary decision variables  $\mathbf{x}$  influencing  $\mathcal{U}(\mathbf{x})$ , problem (RO-DDU) can be reformulated as a MILP problem using the big-M method (see section 5). However, they suffer from weak relaxations.

**4. Structured uncertainty sets.** The weak numerical performance of big-M linearization can be overcome if the decision  $\mathbf{x}$  plays a decisive role in governing the elements of the uncertainty set. Specifically, if the effect of  $\mathbf{x}$  on the uncertainty set constraints can be modeled by penalizing the objective coefficients, then the number of constraints in the robust counterpart can be reduced. Here, we discuss the setting where  $\mathbf{x}$  controls the upper bounds of the uncertain variables. This mechanism can be expressed in the following set.

$$\bar{\Pi}\text{-uncertainty: } \mathcal{U}^{\bar{\Pi}}(\mathbf{x}) = \{\boldsymbol{\xi} \mid \mathbf{D}\boldsymbol{\xi} \leq \mathbf{d}, \boldsymbol{\xi} \leq \mathbf{v} + \mathbf{W}(\mathbf{e} - \mathbf{x}), \boldsymbol{\xi} \geq \mathbf{0}\}.$$

Here,  $\mathbf{D} \in \mathbb{R}^{m \times n}$  is a coefficient matrix,  $\mathbf{d} \in \mathbb{R}^m$  is the RHS vector,  $\mathbf{v} \in \mathbb{R}_+^n$  are the minimum upper bounds, and  $\mathbf{W} = \text{diag}(\mathbf{w}) \in \mathbb{R}_+^{n \times n}$  (a diagonal matrix) are the incremental upper bounds, which apply when reduction is not applied. For  $\mathcal{U}^{\bar{\Pi}}$ , the influence variable is  $\mathbf{x} \in \{0, 1\}^n$ . The decision dependence in  $\mathcal{U}^{\bar{\Pi}}$  affects the upper bounds on each uncertain component  $\xi_i$ . This means, if the problem allows influencing uncertainties, this set can model *proactive* uncertainty reduction. One possible example is disaster planning, where a decision to reduce the fragility of certain roads yields an improved worst-case outcome. Another example is measurement applications, where a decision for additional expenditure leads to increased accuracy. We employed such a set in Example 2 and discuss it further in the numerical application.

We now discuss how this structure can be leveraged to reformulate the original problem (RO-DDU). Note that the objective function remains unaffected by the definition of the uncertainty set, as does the first term of the constraint. Therefore, we focus our discussion on the parts of the constraint in problem (RO-DDU) that are affected by uncertainty.

**4.1.  $\bar{\Pi}$ -uncertainty.** For succinctness, this section provides a reformulation of the following linear constraint:

$$(LC) \quad \mathbf{y}^\top \boldsymbol{\xi} \leq b \quad \forall \boldsymbol{\xi} \in \mathcal{U}^{\bar{\Pi}}(\mathbf{x}).$$

To satisfy this constraint for all  $\boldsymbol{\xi} \in \mathcal{U}^{\bar{\Pi}}(\mathbf{x})$ , the uncertain LHS needs to be replaced by its maximum over the set. For this, consider the following two problems:

$$\begin{array}{ll} h(\mathbf{x}, \mathbf{y}) = & \bar{h}(\mathbf{x}, \mathbf{y}) = \\ \max_{\boldsymbol{\xi}} \mathbf{y}^\top \boldsymbol{\xi} & \max_{\boldsymbol{\xi}, \boldsymbol{\zeta}} (\mathbf{y} - \bar{\Pi}\mathbf{x})^\top \boldsymbol{\xi} + \mathbf{y}^\top \boldsymbol{\zeta} \\ (P) \quad \text{s.t. } \mathbf{D}\boldsymbol{\xi} \leq \mathbf{d}, & (P') \quad \text{s.t. } \mathbf{D}\boldsymbol{\xi} + \mathbf{D}\boldsymbol{\zeta} \leq \mathbf{d}, \\ & \boldsymbol{\xi} \leq \mathbf{W}\mathbf{e}, \\ & \boldsymbol{\zeta} \leq \mathbf{v}, \\ & \boldsymbol{\xi}, \boldsymbol{\zeta} \geq \mathbf{0}, \\ & \boldsymbol{\xi} \leq \mathbf{v} + \mathbf{W}(\mathbf{e} - \mathbf{x}) \quad : \boldsymbol{\pi}(\mathbf{x}, \mathbf{y}), \end{array}$$

where in problem (P),  $\boldsymbol{\pi}(\mathbf{x}, \mathbf{y})$  denotes the corresponding dual variable. Problem (P) maximizes the LHS directly over  $\mathcal{U}^{\bar{\Pi}}(\mathbf{x})$ . However, the standard reformulation of this problem leads to bilinear terms. To avoid them, we can leverage the structure of the uncertainty set and formulate problem (P) as problem (P'). Such a problem pair was also suggested in the context of stochastic network interdiction [15]. Proposition 4.1 uses the duals of (P) and (P') to prove that they have the same objective value at optimality. Formulating problem (P') requires the use of the matrix  $\bar{\Pi} = \text{diag}(\bar{\boldsymbol{\pi}})$ . Here,  $\bar{\boldsymbol{\pi}}$  is a componentwise upper bound of the optimal value of the dual variable  $\boldsymbol{\pi}(\mathbf{x}, \mathbf{y})$  for all  $\mathbf{x}, \mathbf{y}$ . Note that the matrix  $\bar{\Pi}$  is similar to  $M$  of the big-M method

in that it estimates an upper bound to the dual variables. We provide a method to estimate  $\bar{\pi}$  in Proposition 4.2. The dual problems of (P) and (P') are given by

$$\begin{aligned}
 g(\mathbf{x}, \mathbf{y}) = & \min_{\boldsymbol{\pi}, \mathbf{q}} \mathbf{q}^\top \mathbf{d} + \boldsymbol{\pi}^\top \mathbf{v} + \boldsymbol{\pi}^\top \mathbf{W}(\mathbf{e} - \mathbf{x}) \\
 \text{s.t. } & \boldsymbol{\pi}^\top + \mathbf{q}^\top \mathbf{D} \geq \mathbf{y}^\top, \\
 & \boldsymbol{\pi}, \mathbf{q} \geq \mathbf{0},
 \end{aligned}
 \tag{D}$$

$$\begin{aligned}
 \bar{g}(\mathbf{x}, \mathbf{y}) = & \min_{\mathbf{r}, \mathbf{s}, \mathbf{t}} \mathbf{t}^\top \mathbf{d} + \mathbf{r}^\top \mathbf{W}\mathbf{e} + \mathbf{s}^\top \mathbf{v} \\
 \text{s.t. } & \mathbf{s}^\top + \mathbf{t}^\top \mathbf{D} \geq \mathbf{y}^\top, \\
 & \mathbf{r}^\top + \mathbf{t}^\top \mathbf{D} \geq \mathbf{y}^\top - \mathbf{x}^\top \bar{\boldsymbol{\Pi}}, \\
 & \mathbf{r}, \mathbf{s}, \mathbf{t} \geq \mathbf{0}.
 \end{aligned}
 \tag{D'}$$

PROPOSITION 4.1. *Given a binary  $\mathbf{x}$ , if the set  $\mathcal{U}^{\bar{\boldsymbol{\Pi}}}(\mathbf{x})$  is nonempty and  $\mathbf{v}, \mathbf{W} \geq \mathbf{0}$ , then, for all  $\mathbf{y}$ ,*

$$h(\mathbf{x}, \mathbf{y}) = \bar{h}(\mathbf{x}, \mathbf{y}).$$

*Proof.* Strong duality warrants the equalities  $g(\mathbf{x}, \mathbf{y}) = h(\mathbf{x}, \mathbf{y})$  and  $\bar{g}(\mathbf{x}, \mathbf{y}) = \bar{h}(\mathbf{x}, \mathbf{y})$ . In the following, we also refer to the optimal objective values of the dual problems as  $h(\mathbf{x}, \mathbf{y})$  and  $\bar{h}(\mathbf{x}, \mathbf{y})$ . Let  $(\boldsymbol{\pi}, \mathbf{q})$  be an optimal solution to (D). Furthermore, let  $(\mathbf{r} = \boldsymbol{\pi} - \boldsymbol{\Pi}\mathbf{x}, \mathbf{s} = \boldsymbol{\pi}, \mathbf{t} = \mathbf{q})$  with  $\boldsymbol{\Pi} = \text{diag}(\boldsymbol{\pi})$  be a potential feasible solution to (D'). For these solutions, it follows that  $\mathbf{s}^\top + \mathbf{t}^\top \mathbf{D} = \boldsymbol{\pi}^\top + \mathbf{q}^\top \mathbf{D} \geq \mathbf{y}^\top$  and

$$\mathbf{r}^\top + \mathbf{t}^\top \mathbf{D} = \boldsymbol{\pi}^\top - \mathbf{x}^\top \boldsymbol{\Pi} + \mathbf{q}^\top \mathbf{D} \geq \mathbf{y}^\top - \mathbf{x}^\top \boldsymbol{\Pi} \geq \mathbf{y}^\top - \mathbf{x}^\top \bar{\boldsymbol{\Pi}}.$$

Since  $\boldsymbol{\pi}, \mathbf{q} \geq \mathbf{0}$ , and  $\mathbf{x}$  is binary, we obtain  $\mathbf{r}, \mathbf{s}, \mathbf{t} \geq \mathbf{0}$ . This means  $(\mathbf{r}, \mathbf{s}, \mathbf{t})$  is a feasible solution to problem (D'). This yields

$$\begin{aligned}
 \bar{h}(\mathbf{x}, \mathbf{y}) & \leq \mathbf{q}^\top \mathbf{d} + \boldsymbol{\pi}^\top \mathbf{v} + (\boldsymbol{\pi} - \boldsymbol{\Pi}\mathbf{x})^\top \mathbf{W}\mathbf{e} \\
 & = h(\mathbf{x}, \mathbf{y}).
 \end{aligned}$$

For the converse, let  $(\mathbf{r}, \mathbf{s}, \mathbf{t})$  be an optimal solution to (D'). Consider  $(\boldsymbol{\pi} = \mathbf{s}, \mathbf{q} = \mathbf{t})$  to be a solution to (D). The feasibility of  $(\mathbf{r}, \mathbf{s}, \mathbf{t})$  leads to  $\boldsymbol{\pi}^\top + \mathbf{q}^\top \mathbf{D} = \mathbf{s}^\top + \mathbf{t}^\top \mathbf{D} \geq \mathbf{y}^\top$  and  $\boldsymbol{\pi} = \mathbf{s} \geq \mathbf{0}, \mathbf{q} = \mathbf{t} \geq \mathbf{0}$ . Hence,  $(\boldsymbol{\pi}, \mathbf{q})$  is a feasible solution to (D), resulting in

$$\begin{aligned}
 h(\mathbf{x}, \mathbf{y}) & \leq \mathbf{t}^\top \mathbf{d} + \mathbf{s}^\top \mathbf{v} + \mathbf{s}^\top \mathbf{W}(\mathbf{e} - \mathbf{x}) \\
 & = \bar{h}(\mathbf{x}, \mathbf{y}) + (\mathbf{s} - \mathbf{r})^\top \mathbf{W}\mathbf{e} - \mathbf{s}^\top \mathbf{W}\mathbf{x}.
 \end{aligned}$$

In order to prove  $h(\mathbf{x}, \mathbf{y}) \leq \bar{h}(\mathbf{x}, \mathbf{y})$ , it is required to prove  $(\mathbf{s} - \mathbf{r})^\top \mathbf{W}\mathbf{e} - \mathbf{s}^\top \mathbf{W}\mathbf{x} \leq 0$ . This can be expressed as  $\sum_i w_i (s_i - r_i - s_i x_i) \leq 0$ . For all  $i$  with  $x_i = 1$ , it holds that  $w_i (s_i - r_i - s_i x_i) = -w_i r_i \leq 0$ .

Consider now the set of all  $i$  with  $x_i = 0$ , denoted by  $X_0$ . Problem (D') can be rewritten as two nested minimization problems, where the outer problem is over  $\mathbf{t}$  and  $r_j, s_j$  with  $j \notin X_0$  and the inner problem over  $r_i, s_i$  with  $i \in X_0$ :

$$\begin{aligned}
 \bar{h}(\mathbf{x}, \mathbf{y}) = & \min_{\mathbf{t}, r_j, s_j, j \notin X_0} \mathbf{t}^\top \mathbf{d} + \sum_{j \notin X_0} r_j w_j + \sum_{j \notin X_0} s_j v_j + l(\mathbf{t}) \\
 \text{s.t. } & \left. \begin{aligned}
 s_j + \mathbf{t}^\top \mathbf{D}_{\cdot, j} & \geq y_j, \\
 r_j + \mathbf{t}^\top \mathbf{D}_{\cdot, j} & \geq y_j - \bar{\pi}_j, \\
 r_j, s_j & \geq \mathbf{0}
 \end{aligned} \right\} \forall j \notin X_0.
 \end{aligned}$$

The inner minimization is captured by the function  $l(\mathbf{t})$ , which is given by

$$l(\mathbf{t}) = \min_{r_i, s_i, i \in X_0} \sum_{i \in X_0} r_i w_i + \sum_{i \in X_0} s_i v_i$$

$$\left. \begin{aligned} \text{s.t. } s_i + \mathbf{t}^\top \mathbf{D}_{\cdot, i} &\geq y_i, \\ r_i + \mathbf{t}^\top \mathbf{D}_{\cdot, i} &\geq y_i, \\ r_i, s_i &\geq \mathbf{0} \end{aligned} \right\} \forall i \in X_0.$$

Note that in this inner minimization problem the same constraints act on  $s_i$  and  $r_i$ . Since  $w_i$  and  $v_i$  are nonnegative, there exist optimal solutions  $s_i$  and  $r_i$  that are equal and set to their lower bounds  $s_i = r_i = \max\{y_i - \mathbf{t}^\top \mathbf{D}_{\cdot, i}, 0\}$ . Therefore,  $\sum_{i \in X_0} s_i w_i - r_i w_i = 0$ , which means  $h(\mathbf{x}, \mathbf{y}) \leq \bar{h}(\mathbf{x}, \mathbf{y})$ .  $\square$

Using Proposition 4.1 and problem (D'), the constraint (LC) can be reformulated as

$$\begin{aligned} \mathbf{t}^\top \mathbf{d} + \mathbf{r}^\top \mathbf{W} \mathbf{e} + \mathbf{s}^\top \mathbf{v} &\leq b, \\ \mathbf{s}^\top + \mathbf{t}^\top \mathbf{D} &\geq \mathbf{y}^\top, \\ \mathbf{r}^\top + \mathbf{t}^\top \mathbf{D} &\geq \mathbf{y}^\top - \mathbf{x}^\top \bar{\mathbf{\Pi}}, \\ \mathbf{r}, \mathbf{s}, \mathbf{t} &\geq \mathbf{0}. \end{aligned}$$

Note that this reformulation does not contain any bilinear terms and includes fewer constraints than the standard big-M formulations. Additionally, Proposition 4.1 allows us to replace  $h(\mathbf{x}, \mathbf{y})$  with  $\bar{h}(\mathbf{x}, \mathbf{y})$ . This is important because  $\bar{h}(\mathbf{x}, \mathbf{y})$  is convex in  $(\mathbf{x}, \mathbf{y})$ . Therefore, cut-generation algorithms can be used to solve this problem, which is not possible for the original problem with the constraint (LC). In the following, we discuss the matrix  $\bar{\mathbf{\Pi}}$ .

**Estimation of  $\bar{\mathbf{\Pi}}$ .** The following proposition sheds light on how to estimate  $\bar{\mathbf{\Pi}}$ .

PROPOSITION 4.2. *If  $\mathbf{D}$  and  $\mathbf{y}$  are elementwise nonnegative, then  $\pi_i(\mathbf{x}, \mathbf{y}) \leq y_i \forall (\mathbf{x}, \mathbf{y})$  for constraint (LC) under the uncertainty set  $\mathcal{U}^{\bar{\mathbf{\Pi}}}$ .*

*Proof.* Consider the following problem for some index  $i$ :

$$(1) \quad \begin{aligned} F(\theta) &= \max_{\boldsymbol{\xi}} \mathbf{y}^\top \boldsymbol{\xi} \\ \text{s.t. } \mathbf{D} \boldsymbol{\xi} &\leq \mathbf{d} && : \mathbf{q}, \\ \boldsymbol{\xi} &\leq \mathbf{v} + \mathbf{W}(\mathbf{e} - \mathbf{x}) + \theta \mathbf{e}_i && : \boldsymbol{\pi}, \\ \boldsymbol{\xi} &\geq \mathbf{0}. \end{aligned}$$

Let  $\boldsymbol{\xi}_0$  be the optimal solution at  $\theta = 0$  and the corresponding optimal dual variables are  $\mathbf{q}_0$  and  $\boldsymbol{\pi}_0$ . Let the optimal basis of the above problem be given by some matrix  $\mathbf{B}$ . Since  $\boldsymbol{\xi}_0$  is the optimal solution, the vector of basic variables is given by  $\boldsymbol{\xi}_0^B = \mathbf{B}^{-1} \mathbf{b}$ , where  $\mathbf{b}$  denotes the RHS vector of problem (1), i.e.,  $\mathbf{b} = [\mathbf{d}^\top, \mathbf{v}^\top + (\mathbf{e} - \mathbf{x})^\top \mathbf{W}]^\top$ . Assume that the solution is nondegenerate. This means  $\mathbf{B}^{-1} \mathbf{b} > \mathbf{0}$ . Then for a small enough change in  $\mathbf{b}$ , the optimal basis does not change. If it is degenerate, then  $\mathbf{b}$  can be perturbed by a small  $\epsilon$  to obtain a nondegenerate solution, which only marginally changes the objective (see, e.g., [11]).

When  $\theta > 0$  is small enough, the basis matrix does not change. This means that both solutions (corresponding to  $\theta = 0$  and  $\theta > 0$ ) have the same dual variables



because the dual variables do not depend on the RHS vector. This means

$$\begin{aligned} F(\theta) - F(0) &= \boldsymbol{\pi}_0^\top \mathbf{v} + \boldsymbol{\pi}_0^\top \mathbf{W}(\mathbf{e} - \mathbf{x}) + \theta \boldsymbol{\pi}_0^\top \mathbf{e}_i + \mathbf{q}_0^\top \mathbf{d} - \boldsymbol{\pi}_0^\top \mathbf{v} - \boldsymbol{\pi}_0^\top \mathbf{W}(\mathbf{e} - \mathbf{x}) - \mathbf{q}_0^\top \mathbf{d} \\ &= \theta \boldsymbol{\pi}_0^\top \mathbf{e}_i, \end{aligned}$$

which represents the change in the objective value. Let  $\boldsymbol{\xi}_0$  be the optimal solution of the problem with  $\theta = 0$  and let  $\boldsymbol{\xi}_\theta$  be the optimal solution of the problem with  $\theta > 0$ . Then the change in the objective value is

$$\theta \boldsymbol{\pi}_0^\top \mathbf{e}_i = \mathbf{y}^\top \boldsymbol{\xi}_\theta - \mathbf{y}^\top \boldsymbol{\xi}_0.$$

Using Lemma 4.4, we can state that

$$\begin{aligned} \theta \boldsymbol{\pi}_0^\top \mathbf{e}_i &= \mathbf{y}^\top \boldsymbol{\xi}_\theta - \mathbf{y}^\top \boldsymbol{\xi}_0 \\ &\leq \mathbf{y}^\top \boldsymbol{\xi}_0 + \theta \mathbf{y}^\top \mathbf{e}_i - \mathbf{y}^\top \boldsymbol{\xi}_0 \\ &= \theta \mathbf{y}^\top \mathbf{e}_i. \end{aligned}$$

This implies that  $\pi_{0,i} \leq y_i \quad \forall i$ . □

**COROLLARY 4.3.** *Proposition 4.2 allows the estimation of  $\bar{\boldsymbol{\Pi}}$  by*

$$(2) \quad \begin{aligned} \bar{\pi}_i &= \max_{\mathbf{y}} \mathbf{y}^\top \mathbf{e}_i \\ &\text{s.t. } (\mathbf{x}, \mathbf{y}) \in Y, \\ &\quad x_i \in \{0, 1\}, \end{aligned}$$

where set  $Y$  denotes the remaining constraints of the original full problem.

**LEMMA 4.4.** *If the matrix  $\mathbf{D}$  is elementwise greater than 0, then  $\boldsymbol{\xi}_\theta \leq \boldsymbol{\xi}_0 + \theta \mathbf{e}_i$ .*

*Proof.* Suppose this is not true, i.e., there exists at least one index  $k$  such that  $\xi_{\theta,k} > \xi_{0,k} + \theta e_{i,k}$ . In addition, it holds that for  $\theta \geq 0$ ,  $\mathbf{y}^\top \boldsymbol{\xi}_\theta > \mathbf{y}^\top \boldsymbol{\xi}_0$ . If  $k \neq i$ , then  $\boldsymbol{\xi}_\theta \leq \mathbf{v} + \mathbf{W}(\mathbf{e} - \mathbf{x})$ , which suggests that  $\boldsymbol{\xi}_\theta$  would be feasible for the problem with  $\theta = 0$ . This would contradict the optimality of  $\boldsymbol{\xi}_0$ . If  $k = i$ , then  $\xi_{\theta,i} > \xi_{0,i} + \theta$ . However this results in  $\boldsymbol{\xi}_0 < \boldsymbol{\xi}_\theta - \theta \mathbf{e}_i \leq \mathbf{v} + \mathbf{W}(\mathbf{e} - \mathbf{x})$ . Since

$$D(\boldsymbol{\xi}_\theta - \theta \mathbf{e}_i) = \mathbf{D}\boldsymbol{\xi}_\theta - \theta \mathbf{D}\mathbf{e}_i \leq \mathbf{d} - \theta \mathbf{D}\mathbf{e}_i \leq \mathbf{d},$$

$\boldsymbol{\xi}_\theta - \theta \mathbf{e}_i$  is a feasible solution to the problem with  $\theta = 0$ . However, this indicates that  $\mathbf{y}^\top (\boldsymbol{\xi}_\theta - \theta \mathbf{e}_i) > \mathbf{y}^\top \boldsymbol{\xi}_0$ , which also contradicts the optimality of  $\boldsymbol{\xi}_0$ . Therefore, we can conclude that  $\boldsymbol{\xi}_\theta \leq \boldsymbol{\xi}_0 + \theta \mathbf{e}_i$ . □

This proposition allows us to estimate  $\bar{\pi}_i$  by setting it equal to the maximum value that  $y_i$  can take in the overall problem. In some cases, such as shortest path or facility location problems, this is straightforwardly estimated from the underlying model. With this, all components of the decision-dependent problem with the polyhedral uncertainty set  $\mathcal{U}^{\bar{\boldsymbol{\Pi}}}$  can be computed efficiently for practical size problems. We now extend Proposition 4.1 to more general uncertainty sets.

**4.2. Extension to conic sets.** Given a cone  $\mathcal{K}$ , the decision-dependent uncertainty set  $\mathcal{U}^{\bar{\boldsymbol{\Pi}}}(\mathbf{x})$  can be extended to

$$\mathcal{U}^{\mathcal{K}}(\mathbf{x}) = \{\boldsymbol{\xi} \mid \mathbf{d} - \mathbf{D}\boldsymbol{\xi} \in \mathcal{K}, \boldsymbol{\xi} \leq \mathbf{v} + \mathbf{W}(\mathbf{e} - \mathbf{x}), \boldsymbol{\xi} \geq \mathbf{0}\}.$$

Here  $\mathbf{d}$  and  $\mathbf{D}$  are coefficients and  $\mathbf{v}$  and  $\mathbf{W} = \text{diag}(\mathbf{w})$  denote upper bounds to the uncertain component  $\boldsymbol{\xi}$ . The objective is to reformulate the constraint  $\mathbf{y}^\top \boldsymbol{\xi} \leq b$ ,

$\forall \xi \in \mathcal{U}^{\mathcal{K}}(\mathbf{x})$ . In order to satisfy this constraint for all  $\xi \in \mathcal{U}^{\mathcal{K}}(\mathbf{x})$ , its LHS can be expressed with the following two problems:

$$\begin{aligned}
 h(\mathbf{x}, \mathbf{y}) = & \max_{\xi} \mathbf{y}^\top \xi \\
 \text{(KP)} \quad \text{s.t. } & \mathbf{d} - \mathbf{D}\xi \in \mathcal{K}, \\
 & \xi \leq \mathbf{v} + \mathbf{W}(\mathbf{e} - \mathbf{x}) \quad : \quad \pi(\mathbf{x}, \mathbf{y}), \\
 & \xi \geq \mathbf{0},
 \end{aligned}
 \qquad
 \begin{aligned}
 \bar{h}(\mathbf{x}, \mathbf{y}) = & \max_{\xi, \zeta} (\mathbf{y} - \bar{\mathbf{\Pi}}\mathbf{x})^\top \xi + \mathbf{y}^\top \zeta \\
 \text{(KP')} \quad \text{s.t. } & \mathbf{d} - \mathbf{D}\xi \in \mathcal{K}, \\
 & \xi \leq \mathbf{W}\mathbf{e}, \\
 & \zeta \leq \mathbf{v}, \\
 & \xi, \zeta \geq \mathbf{0}.
 \end{aligned}$$

Here,  $\pi(\mathbf{x}, \mathbf{y})$  denotes the dual variable for the corresponding constraint. Let  $\bar{\mathbf{\Pi}}$  be an elementwise upper bound on the dual variables  $\pi(\mathbf{x}, \mathbf{y})$ . The following proposition shows that the problems (KP) and (KP') have the same optimal objective value.

PROPOSITION 4.5. *If  $\forall \mathbf{x} \in \{0, 1\}^n$  there exists a point in the relative interior of  $\mathcal{U}^{\mathcal{K}}(\mathbf{x})$  (a Slater point) and  $\mathbf{v}, \mathbf{W} \geq 0$ , then, for all  $\mathbf{x}, \mathbf{y}$ ,*

$$h(\mathbf{x}, \mathbf{y}) = \bar{h}(\mathbf{x}, \mathbf{y}).$$

The proof of this proposition proceeds similarly to that of Proposition 4.1. It uses strong duality, which holds due to Slater’s condition.

Using Proposition 4.5 and the dual problem of (KP'), the constraint (LC) can be reformulated as

$$\begin{aligned}
 \mathbf{t}^\top \mathbf{d} + \mathbf{r}^\top \mathbf{W}\mathbf{e} + \mathbf{s}^\top \mathbf{v} &\leq b, \\
 \mathbf{s}^\top + \mathbf{t}^\top \mathbf{D} &\geq \mathbf{y}^\top, \\
 \mathbf{r}^\top + \mathbf{t}^\top \mathbf{D} &\geq \mathbf{y}^\top - \mathbf{x}^\top \bar{\mathbf{\Pi}}, \\
 \mathbf{t} \in \mathcal{K}^*, \mathbf{r}, \mathbf{s} &\geq \mathbf{0},
 \end{aligned}$$

with the dual cone  $\mathcal{K}^*$ . Note that this reformulation has only linear terms and, as we will see in section 5, fewer constraints than the standard big-M formulation, hence it is more suitable for larger-size problems.

In summary, these results allow the modeling of uncertainty sets with reducible upper bounds. Such bounds motivate the notion of *proactive uncertainty control*. It mitigates conservatism and better actualizes the tradeoff between cost of control and disadvantage of uncertainty, both of which are instrumental parts of many real-world applications. Until now, we discussed the special polyhedral set  $\mathcal{U}^{\bar{\mathbf{\Pi}}}$ . The following section provides a reformulation of problem (RO-DDU) under general polyhedral uncertainty sets.

**5. Extensions to general polyhedral sets.** The previous section leveraged the specific structure of the uncertainty set to obtain smaller reformulations. The big-M reformulation, however, has the advantage of not requiring any special set structure. For completeness and a comparison of formulation sizes, the following proposition reformulates problem (RO-DDU) for the general polyhedral set  $\mathcal{U}^P(\mathbf{x})$ .

PROPOSITION 5.1. *If the uncertainty set  $\mathcal{U}_i(\mathbf{x})$  is a polyhedron as in  $\mathcal{U}^P(\mathbf{x})$  with  $\mathbf{D}_i \in \mathbb{R}^{m_i \times p}$ ,  $\mathbf{d}_i \in \mathbb{R}^{m_i}$ , and  $\Delta_i \in \mathbb{R}^{m_i \times n}$ , and if  $\mathbf{x}$  is binary, then the robust*

counterpart of problem (RO-DDU) is

$$\begin{aligned}
 & \min_{\mathbf{x}, \mathbf{y}, \mathbf{w}, \boldsymbol{\pi}} \mathbf{c}^\top \mathbf{x} + \mathbf{f}^\top \mathbf{y} \\
 & \text{s.t. } \left. \begin{aligned}
 & \mathbf{a}_i^\top \mathbf{x} + \boldsymbol{\pi}_i^\top \mathbf{d}_i + \sum_{j=1}^{m_i} \sum_{k=1}^n \Delta_{ijk} w_{ijk} \leq b_i, \\
 & \boldsymbol{\pi}_i^\top \mathbf{D}_i = \mathbf{y}^\top \\
 & w_{ijk} \leq Mx_k, \quad w_{ijk} \leq \pi_{ij}, \\
 & w_{ijk} \geq \pi_{ij} - M(1 - x_k), \\
 & \pi_{ij} \geq 0, \quad w_{ijk} \geq 0 \\
 & \mathbf{x} \in \{0, 1\}^n,
 \end{aligned} \right\} \begin{array}{l} \forall i, \\ \forall i, j, k, \end{array}
 \end{aligned}$$

where  $M$  is a sufficiently large number.

*Proof.* We consider two cases.

*Case 1.* There exists a feasible solution  $(\mathbf{x}, \mathbf{y})$  to (RO-DDU). Therefore,  $\mathbf{x}$  and  $\mathbf{y}$  must satisfy all constraints  $\mathbf{a}_i^\top \mathbf{x} + \boldsymbol{\xi}_i^\top \mathbf{y} \leq b_i \quad \forall \boldsymbol{\xi}_i \in \mathcal{U}_i(\mathbf{x})$  for all  $i$ . This is equivalent to

$$(3) \quad \mathbf{a}_i^\top \mathbf{x} + \max_{\boldsymbol{\xi}_i \in \mathcal{U}_i(\mathbf{x})} \boldsymbol{\xi}_i^\top \mathbf{y} \leq b_i \quad \forall i.$$

If this problem is feasible and has a finite optimal solution, then by strong duality the corresponding dual problem has the same objective value. Problem (3) can now be expressed as

$$(4) \quad \left. \begin{aligned}
 & \mathbf{a}_i^\top \mathbf{x} + \boldsymbol{\pi}_i^\top (\mathbf{d}_i + \boldsymbol{\Delta}_i \mathbf{x}) \leq b_i, \\
 & \boldsymbol{\pi}_i^\top \mathbf{D}_i = \mathbf{y}^\top, \\
 & \boldsymbol{\pi}_i \geq \mathbf{0}
 \end{aligned} \right\} \forall i,$$

where  $\boldsymbol{\pi}_i \in \mathbb{R}^{m_i}$  is the dual variable for constraints corresponding to the uncertainty set  $\mathcal{U}_i(\mathbf{x})$ . Here  $m_i$  refers to the number of constraints in the set  $\mathcal{U}_i(\mathbf{x})$ . Since the primal problem is feasible and finitely valued, there exists a  $\boldsymbol{\pi}_i$  for which the constraints (4) are satisfied. Therefore, the original problem (RO-DDU) can be written as

$$\begin{aligned}
 & \min_{\boldsymbol{\pi}_i, \mathbf{x}, \mathbf{y}} \mathbf{c}^\top \mathbf{x} + \mathbf{f}^\top \mathbf{y} \\
 & \text{s.t. } \left. \begin{aligned}
 & \mathbf{a}_i^\top \mathbf{x} + \boldsymbol{\pi}_i^\top \mathbf{d}_i + \boldsymbol{\pi}_i^\top \boldsymbol{\Delta}_i \mathbf{x} \leq b_i, \\
 & \boldsymbol{\pi}_i^\top \mathbf{D}_i = \mathbf{y}^\top, \\
 & \boldsymbol{\pi}_i \geq \mathbf{0}
 \end{aligned} \right\} \forall i.
 \end{aligned}$$

Note the bilinear term in the first constraint. By expanding the variable space, the  $i$ th constraint can be rewritten as

$$\mathbf{a}_i^\top \mathbf{x} + \sum_{j=1}^{m_i} \pi_{ij} d_{ij} + \sum_{j=1}^{m_i} \sum_{k=1}^n \Delta_{ijk} w_{ijk} \leq b_i, \quad \text{with } w_{ijk} = \pi_{ij} x_k.$$

In the bilinear term,  $w_{ijk} = \pi_{ij}x_k$ ,  $x_k$  is binary, allowing us to rewrite the term as

$$w_{ijk} \leq \pi_{ij}, \quad 0 \leq w_{ijk} \leq Mx_k, \quad w_{ijk} \geq \pi_{ij} - M(1 - x_k),$$

where  $M$  is a sufficiently large constant. Consequently, problem (RO-DDU) can be reformulated as

$$(5) \quad \begin{aligned} & \min_{\mathbf{x}, \mathbf{y}} \quad \mathbf{c}^\top \mathbf{x} + \mathbf{f}^\top \mathbf{y} \\ & \text{s.t.} \quad \mathbf{a}_i^\top \mathbf{x} + \boldsymbol{\pi}_i^\top \mathbf{d}_i + \sum_{j=1}^{m_i} \sum_{k=1}^n \Delta_{ijk} w_{ijk} \leq b_i, \\ & \quad \boldsymbol{\pi}_i^\top \mathbf{D}_i = \mathbf{y}^\top \\ & \quad w_{ijk} \leq Mx_k, \quad w_{ijk} \leq \pi_{ij}, \\ & \quad w_{ijk} \geq \pi_{ij} - M(1 - x_k), \\ & \quad \boldsymbol{\pi}_i \geq \mathbf{0}, \quad w_{ijk} \geq 0 \\ & \quad \mathbf{x} \in \{0, 1\}^n. \end{aligned} \quad \left. \begin{array}{l} \\ \\ \\ \\ \\ \\ \end{array} \right\} \begin{array}{l} \forall i, \\ \\ \\ \forall i, j, k, \end{array}$$

*Case 2.* Problem (RO-DDU) is infeasible. Then the reformulation in (5) is infeasible. To show this, consider the original problem (RO-DDU).

Suppose this problem is infeasible under the assumptions of Proposition 5.1. This means that for all  $\mathbf{x}$  there exists an  $i$  and  $\boldsymbol{\xi}_i \in \mathcal{U}_i(\mathbf{x})$  such that  $\mathbf{a}_i^\top \mathbf{x} + \boldsymbol{\xi}_i^\top \mathbf{y} > b_i$ . Consequently it holds that  $\mathbf{a}_i^\top \mathbf{x} + \max_{\boldsymbol{\xi}_i \in \mathcal{U}_i(\mathbf{x})} \boldsymbol{\xi}_i^\top \mathbf{y} > b_i$ . Using the dual of the inner problem, the constraint can be written as

$$(6) \quad \mathbf{a}_i^\top \mathbf{x} + \boldsymbol{\pi}_i^\top (\mathbf{d}_i + \boldsymbol{\Delta}_i \mathbf{x}) > b_i \quad \forall \boldsymbol{\pi}_i \in \{\boldsymbol{\pi}_i \mid \boldsymbol{\pi}_i^\top \mathbf{D}_i = \mathbf{y}^\top, \boldsymbol{\pi}_i \geq \mathbf{0}\}.$$

Now, assume that the reformulation in (5) is feasible. Given its constraints, there exists a binary vector  $\mathbf{x}$  and a vector  $\mathbf{w}$  such that  $w_{ijk} = \pi_{ij}x_k$ . However, this implies a variable  $\boldsymbol{\pi}_i = (\pi_{i1}, \pi_{i2}, \dots, \pi_{ik}, \dots, \pi_{im_i})$  that satisfies  $\boldsymbol{\pi}_i^\top \mathbf{D}_i = \mathbf{y}^\top$ ,  $\boldsymbol{\pi}_i \geq \mathbf{0}$ , and

$$\mathbf{a}_i^\top \mathbf{x} + \sum_{j=1}^{m_i} \pi_{ij} d_{ij} + \sum_{j=1}^{m_i} \sum_{k=1}^n \Delta_{ijk} \pi_{ij} x_k \leq b_i.$$

This contradicts the earlier assertion in (6) that there exists no such  $\boldsymbol{\pi}_i$ . □

This proposition allows us to reformulate the original decision-dependent RO problem as a mixed integer linear program which can be solved for many realistic size problems using off-the-shelf algorithms. Such mixed integer reformulations can also be provided for general convex uncertainty sets [2], which includes conic and budgeted structures. Their proofs (not shown) proceed parallel to that of Proposition 5.1.

Note that problem (RO-DDU) has  $n$  binary and  $p$  continuous variables, along with  $m$  constraints. The  $i$ th uncertain  $\boldsymbol{\xi}_i$  lies in an uncertainty set with  $m_i$  constraints. Table 1 presents the size of the reformulation under two settings: (i)  $\mathbf{x}$  is binary as in Proposition 5.1 and (ii)  $x_i$  can take  $s$  possible values. For the sake of clarity, we assume that  $m_i = K \forall i$ , where  $K$  is some constant. Table 1 shows that for (ii) the size of the reformulation increases rapidly with growing  $s$ . In certain cases, it is possible

to improve the big-M reformulation by imposing mild assumptions, as we will discuss next.

TABLE 1

Size of big-M formulation of (RO-DDU) for  $\mathcal{U}_i(\mathbf{x})$  with respect to (i)  $\mathbf{x} \in \{0, 1\}^n$  and (ii)  $\mathbf{x} \in \mathbb{R}^n$  with  $x_i$  taking  $s$  possible values:  $\dim(\mathbf{y}) = p$ ,  $K$  constraints in  $\mathcal{U}_i(\mathbf{x})$ , and  $m$  constraints in the complete problem.

Nature of $\mathbf{x}$	Binary var.	Continuous var.	Affine constr.	Sign constr.
Binary	$n$	$p + mK + nK$	$m + mp + 3nK$	$mK(n + 1)$
Finite valued	$(n + 1)s$	$p + mK + nmK(s + 1)$	$m + mp + 2n + nmK(3s + 1)$	$mK(ns + 1)$

**5.1. Modified big-M reformulation.** Consider the uncertainty set  $\mathcal{U}^P(\mathbf{x})$  to be expressed as

$$\mathcal{U}^P(\mathbf{x}) = \left\{ \boldsymbol{\xi} \mid \mathbf{D}_i^\top \boldsymbol{\xi} \leq d_i + \sum_{j=1}^n \Delta_{ij} x_j \quad \forall i = 1, \dots, m \right\}.$$

To overcome the poor numerical performance of standard big-M reformulation due to its weak relaxations, we impose the mild assumption that all elements of the coefficient matrix  $\Delta$  are nonnegative. Proposition 5.2 reformulates constraint (LC) for  $\mathcal{U}^P(\mathbf{x})$  under this assumption.

PROPOSITION 5.2. *If  $\Delta_{ij} \geq 0 \quad \forall i, j$ , then the constraint (LC) with the uncertainty set  $\mathcal{U}^P(\mathbf{x})$  and a large constant  $M$  can be reformulated as*

$$\left. \begin{aligned} \sum_{i=1}^m \pi_i d_i + \sum_{i=1}^m \sum_{j=1}^n t_{ij} &\leq b, \\ \sum_{i=1}^m \pi_i D_{ij} &= y_j \quad \forall j, \\ t_{ij} &\geq \pi_i \Delta_{ij} - M(1 - x_j), \\ \pi_i &\geq 0, \quad t_{ij} \geq 0 \end{aligned} \right\} \quad \forall i, j.$$

*Proof.* The LHS maximization problem for the constraint (LC) can be written as

$$\begin{aligned} \max_{\boldsymbol{\xi}} \quad & \mathbf{y}^\top \boldsymbol{\xi} \\ \text{s.t.} \quad & \mathbf{D}_i^\top \boldsymbol{\xi} \leq d_i + \sum_{j=1}^n \Delta_{ij} x_j \quad \forall i. \end{aligned}$$

Using the dual of this problem, the original constraint  $\mathbf{y}^\top \boldsymbol{\xi} \leq b \quad \forall \boldsymbol{\xi} \in \mathcal{U}^P(\mathbf{x})$  can be written as

$$(7) \quad \begin{aligned} \sum_{i=1}^m \pi_i (d_i + \sum_{j=1}^n \Delta_{ij} x_j) &\leq b, \\ \sum_{i=1}^m \pi_i D_{ij} &= y_j \quad \forall j, \\ \boldsymbol{\pi} &\geq \mathbf{0}. \end{aligned}$$

The constraints in (7) can be rewritten by expanding the variable space as

$$\begin{aligned}
 & \sum_{i=1}^m \pi_i d_i + \sum_{i=1}^m \sum_{j=1}^n t_{ij} \leq b, \\
 (8) \quad & \pi_i \Delta_{ij} x_j \leq t_{ij} \quad \forall i, j, \\
 & \sum_{i=1}^m \pi_i D_{ij} = y_j \quad \forall j, \\
 & \boldsymbol{\pi} \geq \mathbf{0}.
 \end{aligned}$$

If there is a variable  $\boldsymbol{\pi}$  feasible for the set of equations given by (7), then we can find a feasible variable for (8) by using  $t_{ij} = \pi_i \Delta_{ij} x_j$ . On the other hand, if there exists a feasible solution to (8), then it is also feasible for (7). If  $x_j = 0$ , then  $t_{ij} \geq 0$  and if  $x_j = 1$ , then  $t_{ij} \geq \pi_i \Delta_{ij}$ . This can be expressed as the set of constraints

$$0 \leq t_j \leq \pi_i \Delta_{ij} - M(1 - x_j),$$

which completes the proof. □

Proposition 5.2 leverages the fact that the variable  $t_{ij}$  remains at its lower bound, making the upper bounding constraints from the big-M linearization redundant. However, if  $t_{ij}$  can be negative, the two lower bounding constraints are not sufficient.

TABLE 2  
 Comparison of (LC) reformulations for the set  $\mathcal{U}^{\overline{\boldsymbol{\Pi}}}(\mathbf{x})$  (C: continuous, A: affine, S: sign).

Formulations	Problem	Variables
		Constraints
$\overline{\boldsymbol{\Pi}}$	$  \begin{aligned}  & \mathbf{t}^\top \mathbf{d} + \mathbf{r}^\top \mathbf{W} \mathbf{e} + \mathbf{s}^\top \mathbf{v} \leq b \\  & \mathbf{s}^\top + \mathbf{t}^\top \mathbf{D} \geq \mathbf{y}^\top \\  & \mathbf{r}^\top + \mathbf{t}^\top \mathbf{D} \geq \mathbf{y}^\top - \mathbf{x}^\top \overline{\boldsymbol{\Pi}} \\  & \mathbf{r}, \mathbf{s}, \mathbf{t} \geq \mathbf{0}  \end{aligned}  $	C: $m + 2n$ <hr style="width: 50%; margin: 5px auto;"/> A: $1 + 2n$ S: $m + 2n$
Big-M	$  \begin{aligned}  & \mathbf{t}^\top \mathbf{d} + \mathbf{s}^\top \mathbf{v} + \mathbf{s}^\top \mathbf{W} \mathbf{e} - \sum_i r_i \leq b \\  & \mathbf{s}^\top + \mathbf{t}^\top \mathbf{D} \geq \mathbf{y}^\top \\  & w_i s_i - M(1 - x_i) \leq r_i \leq M x_i \\  & r_i \leq w_i s_i \\  & \mathbf{r}, \mathbf{s}, \mathbf{t} \geq \mathbf{0}  \end{aligned}  $	C: $m + 2n$ <hr style="width: 50%; margin: 5px auto;"/> A: $1 + 4n$ S: $m + 2n$
Modified big-M	$  \begin{aligned}  & \mathbf{t}^\top \mathbf{d} + \mathbf{s}^\top \mathbf{v} + \mathbf{r}^\top \mathbf{e} \leq b \\  & \mathbf{s}^\top + \mathbf{t}^\top \mathbf{D} \geq \mathbf{y}^\top \\  & r_i \geq w_i s_i - M x_i \\  & \mathbf{r}, \mathbf{s}, \mathbf{t} \geq \mathbf{0}  \end{aligned}  $	C: $m + 2n$ <hr style="width: 50%; margin: 5px auto;"/> A: $1 + 2n$ S: $m + 2n$

In some cases, it is possible to reformulate the problem even if the RHS coefficients are negative. Consider the shortest path example presented in the introduction, which has constraints of the form  $\xi_e \leq 1 - \gamma_e x_e$ . Here, the coefficient  $\Delta_e = -\gamma_e$  is negative. However, we can rewrite the constraint as  $\xi_e \leq (1 - \gamma_e) + \gamma_e(1 - x_e)$  and apply the big-M linearization on the variable  $(1 - x_e)$  instead of on  $x_e$ . This substitution allows the use of the modified big-M reformulation in more general settings. We report the numerical performance of this approach in comparison with the earlier reformulations in section 6. For a comparison, we reformulate the constraint (LC) over the uncertainty set  $\mathcal{U}^{\overline{\boldsymbol{\Pi}}}(\mathbf{x})$  using all three presented techniques, namely, (i)  $\overline{\boldsymbol{\Pi}}$ , (ii) big-M, and (iii) modified big-M. Table 2 presents this comparison along with the

corresponding problem sizes. The sign constraints correspond to  $(\bullet \geq 0)$ , which are presented separately since they can be solved more efficiently. It displays that the primary difference between the big-M and the other two reformulations is the larger number of affine (linear) constraints. To gain intuition and provide computational comparison between the different formulations, we extend the introductory example of section 1 to a more detailed numerical experiment.

**6. Numerical experiments.** Shortest path problems on networks constitute a general class of models, describing the most efficient connection between a source and target. Deterministic shortest path routing problems can be solved with polynomial time algorithms [16]. However, this does not hold for uncertain arc lengths. Past research on robust shortest path problems focused on scenario [35], cardinality [9], and interval uncertainty [1, 37]. Despite a large body of literature, to the best of our knowledge, there is no work in the context of uncertainties that depend on decisions. To this end, our goals are the following:

1. To compare the numerical performance of different robust formulations.
2. To measure the benefit of proactive reduction as a function of size, budget, or cost of reduction.
3. To measure the number of arcs in the shortest path as a function of size, budget, or cost.
4. To evaluate the price of robustness and the benefit of interacting with uncertainties.
5. To compare the average and worst-case cost of decision dependence for RO and SO.

Here, we aim to model challenges that arise, e.g., in scenario planning of natural disasters. When sections of a transportation network are damaged, the actual travel times along arcs become uncertain. To plan for such a scenario, a decision-dependent RO solution can determine the arcs which should be strengthened (by reducing uncertainty) in order to improve the performance in an actual disaster. This strengthening incurs a fee. This means that it is possible to mitigate the impact of a disaster by managing the damage of a few particular arcs. Similarly, for transportation problems (e.g., air, ground), travel time can be improved by acquiring additional traffic or weather information on segments of the network.

To illustrate this setting, we discuss a problem on a graph  $G = (\mathcal{V}, \mathcal{A}, d(\bullet))$  for the set of nodes  $\mathcal{V}$ , arcs  $\mathcal{A}$ , and the distance function  $d(\bullet)$ . The objective is to find the shortest path from the source to the target node ( $s \rightarrow t$ ) when the actual realized distances from node  $i$  to  $j$  are uncertain and a function  $d_{ij}(\boldsymbol{\xi}) = (1 + \frac{1}{2} \xi_{ij})\bar{d}_{ij}$  of  $\boldsymbol{\xi}$ . The variable  $x_{ij}$  decides whether to reduce the maximum uncertainty in  $d_{ij}$ . This inquiry comes at a cost  $c_{ij}$ , which can be motivated as an investment in road improvement and is imposed on travelers via taxes or tolls. The parameter  $\boldsymbol{\xi}$  resides in a cardinality constrained uncertainty set with reducible upper bounds. The complete problem is given by

$$\begin{aligned}
 \text{(SP)} \quad & \min_{\mathbf{x}, \mathbf{y}} \max_{\boldsymbol{\xi} \in \mathcal{U}^{SP}(\mathbf{x})} \sum_{(i,j) \in \mathcal{A}} c_{ij}x_{ij} + \sum_{(i,j) \in \mathcal{A}} d_{ij}(\boldsymbol{\xi})y_{ij} \\
 & \text{s.t. } \mathbf{x} \in X \subseteq \{0, 1\}^{|\mathcal{A}|}, \mathbf{y} \in Y,
 \end{aligned}$$

where  $y_{ij}$  decides whether the arc  $(i, j)$  lies in the shortest path.  $X$  denotes any

constraints on  $\mathbf{x}$  and  $Y$  the set of routing constraints. The uncertainty set is given by

$$\mathcal{U}^{SP}(\mathbf{x}) = \left\{ \boldsymbol{\xi} \mid \sum_{(i,j) \in \mathcal{A}} \xi_{ij} \leq \Gamma, \xi_{ij} \leq 1 - \gamma_{ij}x_{ij}, \xi_{ij} \geq 0 \ \forall (i,j) \in \mathcal{A} \right\}.$$

We solve problem (SP) using three different formulations: (i) the  $\bar{\Pi}$  formulation from Proposition 4.1, (ii) the standard big-M formulation, and (iii) the modified big-M formulation from Proposition 5.2. In Table 3,  $X \times Y$  denotes the collection of both the shortest path and decision constraints. Furthermore,

$$f(\mathbf{x}, \mathbf{y}) = \sum_{(i,j) \in \mathcal{A}} c_{ij}x_{ij} + \sum_{(i,j) \in \mathcal{A}} \bar{d}_{ij}y_{ij}$$

denotes the total cost of reduction and nominal length. Table 3 shows that the difference between the big-M formulation and the other two formulations lies in the number of affine (linear) constraints, as shown in Table 2. We now discuss the numerical experiments.

TABLE 3  
Shortest path formulations for the set  $\mathcal{U}^{SP}(\mathbf{x})$  (B: binary, C: continuous, A: affine, S: sign).

Form.	Problem	Variables Constraints
$\bar{\Pi}$	$\min_{\mathbf{x}, \mathbf{y}, p} f(\mathbf{x}, \mathbf{y}) + p\Gamma + \sum_{(i,j) \in \mathcal{A}} q_{ij}(1 - \gamma_{ij}) + \sum_{(i,j) \in \mathcal{A}} r_{ij}\gamma_{ij}$ $\text{s.t. } p + q_{ij} \geq \frac{y_{ij}d_{ij} - \bar{\pi}_{ij}d_{ij}x_{ij}}{2}$ $p + r_{ij} \geq \frac{y_{ij}d_{ij}}{2}$ $p, q_{ij}, r_{ij} \geq 0, \mathbf{x}, \mathbf{y} \in X \times Y$	B: $2 \mathcal{A} $ C: $2 \mathcal{A}  + 1$ <hr/> A: $ \mathcal{V}  + 2 \mathcal{A} $ S: $2 \mathcal{A}  + 1$
Big-M	$\min_{\mathbf{x}, \mathbf{y}, p} f(\mathbf{x}, \mathbf{y}) + p\Gamma + \sum_{(i,j) \in \mathcal{A}} q_{ij} - \sum_{(i,j) \in \mathcal{A}} \gamma_{ij}r_{ij}$ $\text{s.t. } p + q_{ij} \geq \frac{d_{ij}y_{ij}}{2}$ $0 \leq r_{ij} \leq Mx_{ij}$ $q_{ij} - M(1 - x_{ij}) \leq r_{ij} \leq q_{ij}$ $p, q_{ij}, r_{ij} \geq 0, \mathbf{x}, \mathbf{y} \in X \times Y$	B: $2 \mathcal{A} $ C: $2 \mathcal{A}  + 1$ <hr/> A: $ \mathcal{V}  + 4 \mathcal{A} $ S: $2 \mathcal{A}  + 1$
Modified big-M	$\min_{\mathbf{x}, \mathbf{y}, p} f(\mathbf{x}, \mathbf{y}) + p\Gamma + \sum_{(i,j) \in \mathcal{A}} r_{ij} + \sum_{(i,j) \in \mathcal{A}} q_{ij}(1 - \gamma_{ij})$ $\text{s.t. } p + q_{ij} \geq \frac{d_{ij}y_{ij}}{2}$ $r_{ij} \geq \gamma_{ij} - Mx_{ij}$ $p, q_{ij}, r_{ij} \geq 0, \mathbf{x}, \mathbf{y} \in X \times Y$	B: $2 \mathcal{A} $ C: $2 \mathcal{A}  + 1$ <hr/> A: $ \mathcal{V}  + 2 \mathcal{A} $ S: $2 \mathcal{A}  + 1$

*Experiment 1* (performance comparison). The numerical setup is as follows. We randomly generate points on a  $100 \times 100$  area and connect them to create a complete graph. The two furthest nodes constitute the source and destination. The final graph is selected after removing 60% of the longest arcs in order to avoid direct connections between the source and destination. The uncertainty budget  $\Gamma$  is set to 2. The cost of reduction  $c_{ij} = c$  and the fraction of uncertainty reduced  $\gamma_{ij} = \gamma$  are 1.0 and 0.2, respectively. For each size  $|\mathcal{V}| = \{50, 75, \dots, 300\}$ , 100 random graphs are generated. These values serve as an illustration of the qualitative comparison of the formulations. In practical applications, they need to be estimated from the economical value of travel time ( $d_{ij}$ ) relative to the per-trip tax burden for road investments ( $c_{ij}$ ).



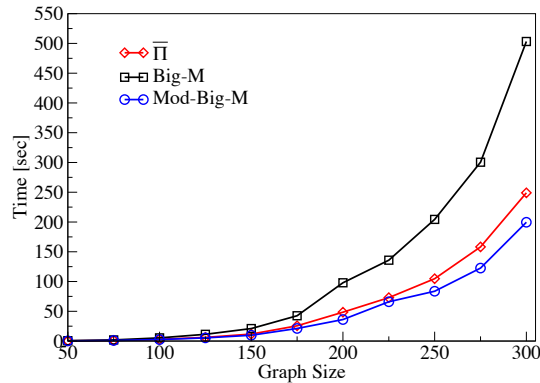


FIG. 2. Comparison of median solution times of reformulations from Propositions 4.1 and 5.2 and the standard big-M.

To solve these problems, we used the Gurobi 7.0 solver on a commercially available computing unit with Intel Core i7 at 3.6 GHz. The median computation times for different approaches and varying sizes are reported in Figure 2. Note that all three methods lead to the same solution. The observations from Figure 2 can be summarized as follows.

- The time increases with growing  $|\mathcal{V}|$  for all formulations. However, the increase is less steep for the  $\bar{\Pi}$  and modified big-M formulations than for the big-M formulation.
- The difference between big-M and the proposed formulations increases with growing  $|\mathcal{V}|$ . This highlights the advantage of the  $\bar{\Pi}$  and modified big-M formulations for larger graphs.
- The median time of the modified big-M formulation is less than that of the  $\bar{\Pi}$  formulation.

Figure 2 highlights the benefits of using the proposed formulations to solve such decision-dependent optimization problems. While the performance of the modified big-M and  $\bar{\Pi}$  formulations are comparable over a broad range of network sizes, the subproblem in the  $\bar{\Pi}$  formulation is convex, which can be exploited by cut-generating methods, which may be computationally advantageous. We also solved the  $\bar{\Pi}$  formulation using a cut-generation approach (not shown). However, for this application, it converged slowly and required a sizable number of cuts.

We now focus on analyzing how the solution changes as the parameters of the uncertainty set are varied. For this purpose, we introduce additional notation for observable quantities.

*Notation for observables.* The number of arcs in the shortest path is  $n^*$ , which is a function of the budget  $\Gamma$  and the level of uncertainty reduction  $\gamma$ . These parameters create three scenarios:

- the *nominal* case, where no uncertainty is present,  $n^*(\Gamma = 0, \gamma = 0)$ ;
- the *standard robust* case with no decision dependence,  $n^*(\Gamma > 0, \gamma = 0)$ ; and
- the *decision-dependent robust* case with uncertainty reduction  $n^*(\Gamma > 0, \gamma > 0)$ , in which case  $\tilde{n}$  is the number of arcs whose uncertainty was reduced.

We also follow this notation for the optimal objective value  $z^*$ . Consequently, the difference  $(z^*(\Gamma > 0, \gamma = 0) - z^*(\Gamma = 0, \gamma = 0))$  constitutes the *price of robustness*,

whereas the difference ( $z^*(\Gamma > 0, \gamma = 0) - z^*(\Gamma > 0, \gamma > 0)$ ) constitutes the *benefit of interaction*.

There are four parameters that govern the effect of interactions with uncertainty:  $\gamma$ ,  $|\mathcal{V}|$ ,  $c$ , and  $\Gamma$ . To evaluate their role and to infer the underlying mechanism, we devise four experiments by tuning across their range. Specifically, by adjusting one parameter while keeping the other three fixed, we explore four orthogonal settings.

In these experiments, problem (SP) is implemented on randomly generated graphs of 20 to 50 nodes. This size is comparable to moderately sized transportation networks [26]. For each size, 2000 graphs are generated in a manner similar to the previous experiment. We maintain these parameter values throughout the following experiments, except in those where their change is probed. In the following, we discuss the four experiments.

*Experiment 2* (uncertainty reduction). We compare  $z^*$  when reduction is permitted ( $\gamma > 0$ ) or not ( $\gamma = 0$ ). Figure 3(a) shows that  $\gamma > 0$  reduces  $z^*$  (shorter paths), which is independent of  $|\mathcal{V}|$ . The inset of Figure 3(a) is a magnification, displaying marginal fluctuations that stem from the random nature of graphs.

*Experiment 3* (graph size). We observe that not all arcs in the shortest path experience uncertainty reduction ( $\tilde{n} < n^*(\Gamma > 0, \gamma > 0)$ ) independent of  $|\mathcal{V}|$ . This is attributed to the nonzero  $c$ . We also observe that  $z^*$  is independent of  $|\mathcal{V}|$ , which can be explained by the fact that  $|\mathcal{V}|$  only increases from 20 to 50 and  $n^*(\Gamma > 0, \gamma > 0)$  does not change sizably over this range and, as such, the effect on  $z^*$  is undetectable. We expect  $n^*$  and  $z^*$  to increase measurably when  $|\mathcal{V}|$  varies by a few orders of magnitude. Larger experiments come at a significant computational burden and are outside the scope of this study.

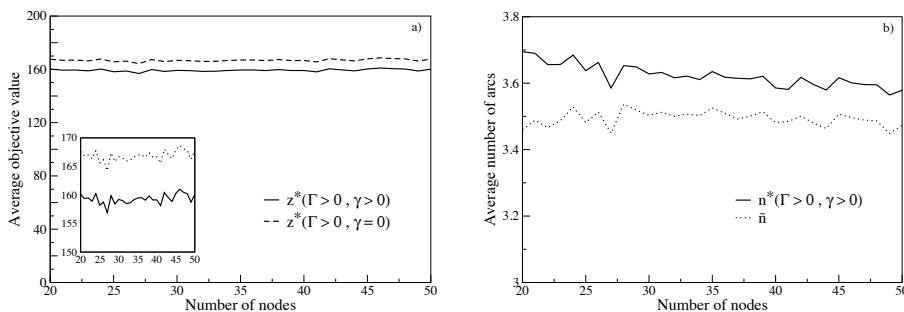


FIG. 3. Dependence on graph size  $|\mathcal{V}|$  for (a) average objective function and (b) average number of arcs. The inset is a magnification.

Figure 3(b) illustrates the average  $n^*(\Gamma > 0, \gamma > 0)$  and the average  $\tilde{n}$  for varying  $|\mathcal{V}|$ . We also observe a slight downward trend of  $n^*(\Gamma > 0, \gamma > 0)$  with increasing  $|\mathcal{V}|$ . This is because the connectivity within a graph increases with  $|\mathcal{V}|$  since the number of arcs grows faster than the number of nodes, because in the experimental setup only a fixed fraction of arcs are removed.

*Experiment 4* (cost of uncertainty reduction). The reduction cost  $c$  determines the trade-off between accepting the uncertainty level and its reduction. It can be expected that an increasing  $c$  marginalizes the benefits of reducing uncertainty. This means that for a sufficiently low  $c$ , uncertainty can be reduced in every arc in the shortest path. On the other hand, for high  $c$ , the opposite is true. Figure 4(a)

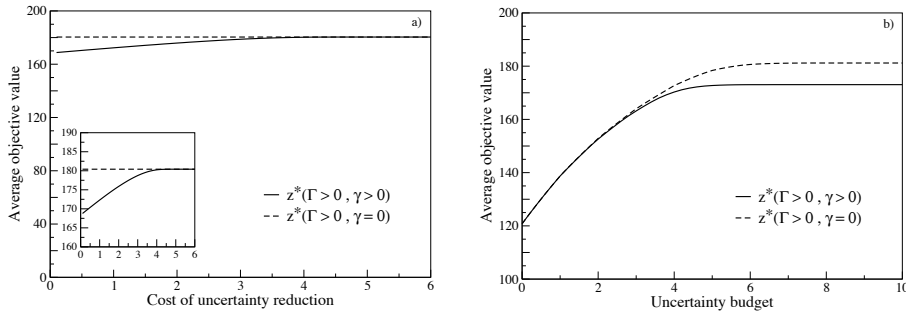


FIG. 4. Average objective value as a function of (a) cost of uncertainty reduction  $c$  and (b) maximum uncertainty  $\Gamma$ . The graph consists of  $|\mathcal{V}| = 30$  nodes.

( $|\mathcal{V}| = 30$  and  $\Gamma = 12$ ) shows that for  $c \leq 4$  the average  $z$  can be decreased. However, for large  $c$  the high cost of reduction makes it disadvantageous to reduce uncertainty. The price of robustness (the difference between the dotted line in Figure 4(a) and  $z^*(\Gamma = 0, \gamma = 0)$  in Figure 4(b)) is constant w.r.t.  $\gamma$  but changes with  $\Gamma$ . On the other hand, the benefit of interaction decreases with increase in  $c$ , as can be observed in Figure 5(a). Note that the maximum benefit of interaction is calculated by assuming uncertainty is reduced on all the arcs in the shortest path, at zero cost ( $c = 0$ ).

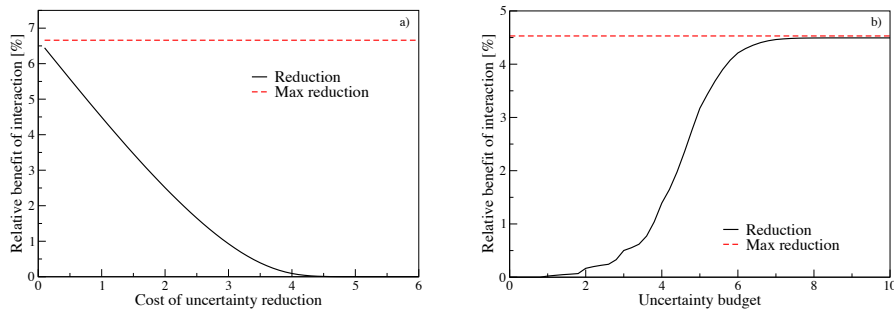


FIG. 5. Average relative benefit of interaction as a function of (a) cost of uncertainty reduction  $c$  and (b) maximum uncertainty  $\Gamma$ . The graph consists of  $|\mathcal{V}| = 30$  nodes.

*Experiment 5* (uncertainty budget).  $\Gamma$  governs the number of arcs that can be affected by uncertainty. Figure 4(b) shows that  $z^*$  increases gradually with  $\Gamma$  until it reaches the level of the corresponding shortest path length affected by the relative uncertainty  $(1 + \frac{1}{2})$  and plateaus thereafter. This is because increasing  $\Gamma$  beyond a certain point does not have any effect on  $n^*$ , since all the arcs in the path are already uncertain and additional budget remains untapped. Consequently, the price of robustness increases with  $\Gamma$  and plateaus beyond a certain  $\Gamma$  (not shown). An analogous behavior can be observed for the benefit of interaction, as shown in Figure 5(b). The maximum benefit is achieved at  $c = 0$ .

Figure 6(a) displays how the average  $n^*$  changes with  $\Gamma$  for the different settings. Note that the values of uncertainty are relative to the nominal arc length. This provides an upper bound on the maximum objective value, i.e., when every arc in the

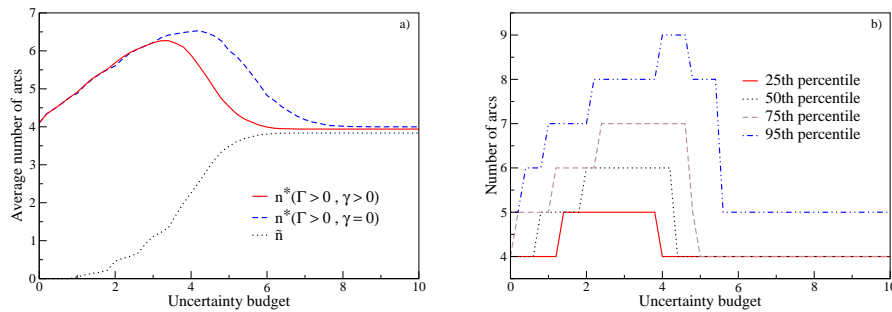


FIG. 6. The dependence on budget of uncertainty  $\Gamma$  for (a) average number of arcs and (b) their distribution. The graph consists of  $|\mathcal{V}| = 30$  nodes and uncertainty reduction is permitted.

shortest path (contributing to  $n^*$ ) is affected by the uncertainty. At  $\Gamma = 0$ , we observe  $n^*(\Gamma = 0, \gamma = 0)$ , and  $\tilde{n} = 0$ . As  $\Gamma$  increases, it becomes beneficial to choose more but shorter arcs, hence, the average  $n^*(\Gamma > 0, \gamma = 0)$  initially increases and reaches a maximum at  $\Gamma \approx n^*(\Gamma = 0, \gamma = 0)$ . As  $\Gamma$  grows even further, the standard robust solution  $n^*(\Gamma > 0, \gamma = 0)$  decreases and plateaus at the same level as  $n^*(\Gamma = 0, \gamma = 0)$ . When  $\gamma > 0$ , we observe that an increasing  $\Gamma \geq 0$  permits more uncertain arc lengths to be reduced ( $\tilde{n} \geq 0$ ) to a maximum of  $\tilde{n} \leq n^*(\Gamma = 0, \gamma = 0)$ . Since some of the arc uncertainty can be reduced, the peak of  $n^*(\Gamma > 0, \gamma > 0)$  occurs at a lower budget than when no reduction is allowed, as seen in Figure 6(a). Note that for small  $\Gamma$ , in order to cope with uncertainty, the optimal solution minimizes the length of each individual arc so that the impact of the uncertainty is minimized.

To further support this observation, Figure 6(b) displays the distribution of the number of arcs using different percentiles of  $n^*(\Gamma > 0, \gamma > 0)$  (corresponding to Figure 6(a)). Here, we observe that as  $\Gamma$  increases the distribution of  $n^*(\Gamma > 0, \gamma > 0)$  skews towards a larger number of arcs (the gaps between the percentiles increase). This means that the optimal solution becomes more diversified. Specifically, the model selects a path consisting of some certain and some uncertain arcs, with a subset of the latter experiencing uncertainty reduction. This continues until the saturation point (here  $\Gamma \approx 4$ ) because beyond a certain budget, diversification of paths becomes redundant. At this point, the shortest path is chosen exclusively amongst uncertain arcs, almost all of which experience uncertainty reduction (since  $\Gamma > n^*(\Gamma = 0, \gamma = 0)$ ).

*Experiment 6* (comparison to SO). This experiment evaluates the average and worst-case performance of the robust decision-dependent solutions and compares them to a similar SO problem. The SO formulation is given by

$$\begin{aligned} \min_{\mathbf{x}} \quad & \sum_{(i,j) \in \mathcal{A}} c_{ij} x_{ij} + \mathbb{E}_{\mathbb{P}(\mathbf{x})} \left[ \sum_{(i,j) \in \mathcal{A}} d_{ij}(\boldsymbol{\xi}) y_{ij} \right] \\ \text{s.t.} \quad & \mathbf{y} \in Y, \\ & \mathbf{x} \in \{0, 1\}^{|\mathcal{A}|}, \end{aligned}$$

with the uncertainty set

$$\boldsymbol{\xi} \in \mathcal{U}^{SSP}(\mathbf{x}) = \times_{i,j \in \mathcal{A}} [0, 1 - \gamma x_{ij}].$$

The distribution  $\mathbb{P}(\mathbf{x})$  is the uniform distribution over the support  $\mathcal{U}^{SSP}(\mathbf{x})$ . The average performance is evaluated by randomly generating the uncertain component  $\xi_{ij}$  (from  $[0, 1]$  for unreduced arcs and  $[0, 1 - \gamma x_{ij}]$  for reduced arcs) and implementing the existing robust and stochastic solutions for these randomly generated arc costs. The following solutions are evaluated. (i) RO: Robust solution for  $\gamma = 0$ . (ii) RO-DDU: Robust solution for  $\gamma > 0$ . (iii) SO: Stochastic solution for  $\gamma = 0$ . (iv) SO-DDU: Stochastic solution for  $\gamma > 0$ . The suffix of the average performances is “-A” and of the worst-case performances is “-W.”

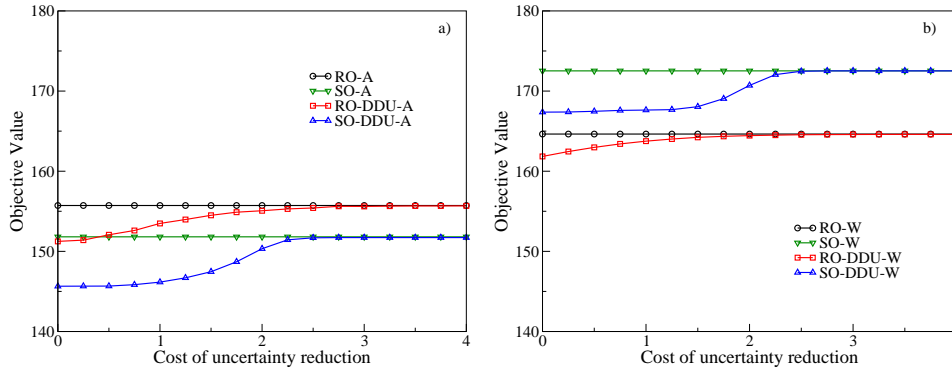


FIG. 7. Comparison of RO and SO formulations: (a) average and (b) worst-case objective value.

Figure 7(a) shows that the average objective of SO is less than the average RO objective. This is because RO optimizes the worst-case instead of the average performance as in SO. However, analogously in Figure 7(b), RO-W is significantly smaller than SO-W. The same applies to the decision-dependent counterparts for both cases. As can be expected, the objective values increase with  $c$  until it is no longer beneficial to reduce the uncertainty, i.e., the objective value of the RO-DDU solution increases until it matches that of the RO solution. The same holds true for the SO-DDU and SO solutions.

In summary, the  $\overline{\Pi}$  formulation and the modified big-M formulation perform considerably better than the standard big-M formulation and their benefits increase with graph size. The worst-case cost for the shortest path can be improved by proactively reducing the uncertainty on a subset of arcs. As the budget of uncertainty grows, these benefits improve but plateau beyond a certain level. At the same time, the cost of reduction curbs these benefits. The RO-DDU problem performs better than SO-DDU for the worst-case scenario. As expected, this benefit comes at the price of the average cost. This numerical study provides an overview of the impact of different formulations, probes various model parameters, and highlights the power of the proactive uncertainty control for both the worst-case and average performance.

**7. Concluding remarks.** In this paper, we present a novel optimization approach for solving problems with decision-dependent uncertainties. We show that for general polyhedral sets, such problems are, even in basic cases, NP-complete. To alleviate this, we introduce a class of uncertainty sets whose upper bounds are affected by decisions. They enable more realistic modeling of a broad range of applications and extend RO beyond the currently used exogenous sets. We provide reformula-

tions that have considerably fewer constraints compared to standard linearization techniques, allowing for faster computations. Our approach should be viewed as one option among many to model decision dependence while maintaining computational advantages. The induced convexity of the subproblem in the proposed reformulation reveals a path forward to use advanced cut-generating algorithms. We believe that finding new and appropriate conditions on sets will further improve the quality of the reformulations.

In addition, this work provides an alternative way of addressing one of the criticisms of RO approaches, namely overly conservative solutions. The description via decision-dependent sets enables mitigation of this issue by exercising proactive control on uncertainties. This setting offers an immediate way to manage the tradeoff between conservatism and optimality. Finally, novel cutting plane methods have instrumentally enhanced solution times and we envisage decision-dependent sets to solidify the tradeoff between computation and optimality by inducing beneficial cuts.

**Acknowledgment.** We are grateful to David Morton for insightful comments.

#### REFERENCES

- [1] I. AVERBAKH AND V. LEBEDEV, *Interval data minmax regret network optimization problems*, Discrete Appl. Math., 138 (2004), pp. 289–301.
- [2] A. BEN-TAL, D. DEN HERTOOG, AND J.-P. VIAL, *Deriving robust counterparts of nonlinear uncertain inequalities*, Math. Program., 149 (2015), pp. 265–299.
- [3] A. BEN-TAL, L. EL GHAOU, AND A. NEMIROVSKI, *Robust Optimization*, Princeton University Press, Princeton, NJ, 2009.
- [4] A. BEN-TAL, A. GORYASHKO, E. GUSLITZER, AND A. NEMIROVSKI, *Adjustable robust solutions of uncertain linear programs*, Math. Program., 99 (2004), pp. 351–376.
- [5] A. BEN-TAL AND A. NEMIROVSKI, *Robust solutions of uncertain linear programs*, Oper. Res. Lett., 25 (1999), pp. 1–13.
- [6] D. BERTSIMAS, D. B. BROWN, AND C. CARAMANIS, *Theory and applications of robust optimization*, SIAM Rev., 53 (2011), pp. 464–501.
- [7] D. BERTSIMAS AND C. CARAMANIS, *Finite adaptability in multistage linear optimization*, IEEE Trans. Automat. Control, 55 (2010), pp. 2751–2766.
- [8] D. BERTSIMAS, O. NOHADANI, AND K. M. TEO, *Robust optimization for unconstrained simulation-based problems*, Oper. Res., 58 (2010), pp. 161–178.
- [9] D. BERTSIMAS AND M. SIM, *Robust discrete optimization and network flows*, Math. Program., 98 (2003), pp. 49–71.
- [10] D. BERTSIMAS AND M. SIM, *The price of robustness*, Oper. Res., 52 (2004), pp. 35–53.
- [11] D. BERTSIMAS AND J. N. TSITSIKLIS, *Introduction to Linear Optimization*, Athena Sci. Ser. Optim. Neural Comput. 6, Athena Scientific, Belmont, MA, 1997.
- [12] D. BERTSIMAS AND P. VAYANOS, *Data-Driven Learning in Dynamic Pricing Using Adaptive Optimization*, [www.optimization-online.org/DB\\_HTML/2014/10/4595.html](http://www.optimization-online.org/DB_HTML/2014/10/4595.html), 2014.
- [13] M. CHU, Y. ZINCHENKO, S. G. HENDERSON, AND M. B. SHARPE, *Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty*, Phys. Med. Biol., 50 (2005), pp. 5463–5477.
- [14] S. A. COOK, *The complexity of theorem-proving procedures*, in STOC '71: Proceedings of the Third Annual ACM Symposium on Theory of Computing, ACM, New York, 1971, pp. 151–158.
- [15] K. J. CORMICAN, D. P. MORTON, AND R. K. WOOD, *Stochastic network interdiction*, Oper. Res., 46 (1998), pp. 184–197.
- [16] E. W. DIJKSTRA, *A note on two problems in connexion with graphs*, Numer. Math., 1 (1959), pp. 269–271.
- [17] A. GEORGHIOU, W. WIESEMANN, AND D. KUHN, *Generalized decision rule approximations for stochastic programming via liftings*, Math. Program., 152 (2015), pp. 301–338.
- [18] L. E. GHAOU, M. OKS, AND F. OUSTRY, *Worst-case value-at-risk and robust portfolio optimization: A conic programming approach*, Oper. Res., 51 (2003), pp. 543–556.
- [19] V. GOEL AND I. E. GROSSMANN, *A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves*, Comput. Chem. Eng., 28 (2004), pp. 1409–1429.

- [20] V. GOEL AND I. E. GROSSMANN, *A class of stochastic programs with decision dependent uncertainty*, Math. Program., 108 (2006), pp. 355–394.
- [21] G. A. HANASUSANTO, D. KUHN, AND W. WIESEMANN, *K-adaptability in two-stage robust binary programming*, Oper. Res., 63 (2015), pp. 877–891.
- [22] D. A. IANCU, *Adaptive Robust Optimization with Applications in Inventory and Revenue Management*, Ph.D. thesis, MIT, Cambridge, MA, 2010.
- [23] T. W. JONSBRÅTEN, R. J.-B. WETS, AND D. L. WOODRUFF, *A class of stochastic programs with decision dependent random elements*, Ann. Oper. Res., 82 (1998), pp. 83–106.
- [24] S. KOŁODZIEJ, P. M. CASTRO, AND I. E. GROSSMANN, *Global optimization of bilinear programs with a multiparametric disaggregation technique*, J. Global Optim., 57 (2013), pp. 1039–1063.
- [25] A. LORCA, X. A. SUN, E. LITVINOV, AND T. ZHENG, *Multistage adaptive robust optimization for the unit commitment problem*, Oper. Res., 64 (2016), pp. 32–51.
- [26] R. MONTEMANNI AND L. M. GAMBARDELLA, *The robust shortest path problem with interval data via Benders decomposition*, 4OR, 3 (2005), pp. 315–328.
- [27] O. NOHADANI AND A. ROY, *Robust optimization with time-dependent uncertainty in radiation therapy*, IISE Trans. Healthcare Syst. Eng., 7 (2017), pp. 81–92.
- [28] L. J. NOVOA, A. I. JARRAH, AND D. P. MORTON, *Flow balancing with uncertain demand for automated package sorting centers*, Transport. Sci., 52 (2016), <https://doi.org/10.1287/trsc.2015.0662>.
- [29] M. POSS, *Robust combinatorial optimization with variable budgeted uncertainty*, 4OR, 11 (2013), pp. 75–92.
- [30] M. POSS, *Robust combinatorial optimization with variable cost uncertainty*, Eur. J. Oper. Res., 237 (2014), pp. 836–845.
- [31] A. SHAPIRO, D. DENTCHEVA, AND A. RUSZCZYŃSKI, *Lectures on Stochastic Programming: Modeling and Theory*, MOS-SIAM Ser. Optim. 9, SIAM, Philadelphia, PA, 2009.
- [32] A. L. SOYSTER, *Technical non-convex programming with set-inclusive constraints and applications to inexact linear programming*, Oper. Res., 21 (1973), pp. 1154–1157.
- [33] S. A. SPACEY, W. WIESEMANN, D. KUHN, AND W. LUK, *Robust software partitioning with multiple instantiation*, INFORMS J. Comput., 24 (2012), pp. 500–515.
- [34] R. VUJANIC, P. GOULART, AND M. MORARI, *Robust optimization of schedules affected by uncertain events*, J. Optimiz. Theory App., 171 (2016), pp. 1033–1054.
- [35] G. YU AND J. YANG, *On the robust shortest path problem*, Comput. Oper. Res., 25 (1998), pp. 457–468.
- [36] X. ZHANG, M. KAMGARPOUR, A. GEORGHIOU, P. GOULART, AND J. LYGEROS, *Robust optimal control with adjustable uncertainty sets*, Autom., 75 (2017), pp. 249–259.
- [37] P. ZIELIŃSKI, *The computational complexity of the relative robust shortest path problem with interval data*, Eur. J. Oper. Res., 158 (2004), pp. 570–576.