Optimization under Decision-Dependent Uncertainty

Omid Nohadani, Kartikey Sharma
Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208,
nohadani@northwestern.edu, kartikeysharma2014@u.northwestern.edu

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 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 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 proactive uncertainty control also mitigates over conservatism of current robust optimization approaches.

Key words: robust optimization, endogenous uncertainty, decision-dependent uncertainty

History: October 10, 2016

1. Introduction

Optimization solutions become suboptimal, or even infeasible, when they are exposed to uncertainty. The two well-established approaches of stochastic optimization and robust optimization address this issue by taking uncertainties into account. Stochastic optimization (SO) was introduced by Dantzig (1955) and can be used when the distribution of the uncertainty is available. For a review of SO, refer to Shapiro et al. (2009).

When uncertainties can be regarded as residing in a set, robust optimization (RO) is a computationally attractive alternative (Ben-Tal et al. 2009, Bertsimas et al. 2011). The method of RO was first mentioned by Soyster (1973) and has been extended considerably in the past two decades and applied to problems ranging from portfolio management (Ghaoui et al. 2003), to healthcare (Chu et al. 2005), to electricity systems (Lorca et al. 2016), and to engineering design (Bertsimas et al. 2010). RO has also been extended to multistage problems. In order to allow for less conservative and more realistic solutions, Ben-Tal et al. (2004) introduced the notion of adjustable robust optimization (ARO). Since this approach is generally intractable, various approximations have been proposed, e.g., by allowing second-stage decisions to be functions of uncertainty with affine or quadratic structures (Iancu 2010, Georghiou et al. 2015). Bertsimas et al. (2015) provide bounds on the sub-optimality of static solutions for ARO problems.

RO and ARO both employ uncertainty sets that are predetermined and, hence, *exogenous*. However in many real-world problems, the uncertainty can be affected by decisions. In such decision-dependent cases, the uncertainty set is *endogenous*. Optimization over such endogenous uncertainties is the focus of this work. We introduce the general RO problem with endogenous uncertainties and provide a class of uncertainty sets, whose reformulations improve over standard techniques. 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 et al. (1998) 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 (2004) 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 schedule, that minimizes cost while satisfying the demand for different goods (Goel and Grossmann 2006). For package sorting centers, Novoa et al. (2016) 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. (2012) address a software partitioning problem, where code segments are assigned to different 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 et al. (2015) use a finite adaptability approximation to ARO, as introduced by Bertsimas and Caramanis (2010), and consider optimization problems with binary recourse decisions. For problems with uncertain objective and constraints, they formulate the RO problem with a decision-dependent uncertainty set before finally reformulating it as a MILP. Poss (2013) considers a combinatorial optimization problem with budgeted uncertainty set. This extends the work of Bertsimas and Sim (2004) to decision-dependent budgets and is confined to budget uncertainty with limited discussion on general sets. On the other hand, for a dynamic pricing problem with learning, Bertsimas and Vayanos (2015) consider 1 or ∞-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 schedule problems, Vujanic et al. (2016) 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. Note that in all approaches to date, the decision dependence is modeled in a specific context, often driven by an application.

In this work, we study single-stage RO problems with decision-dependent uncertainty sets

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

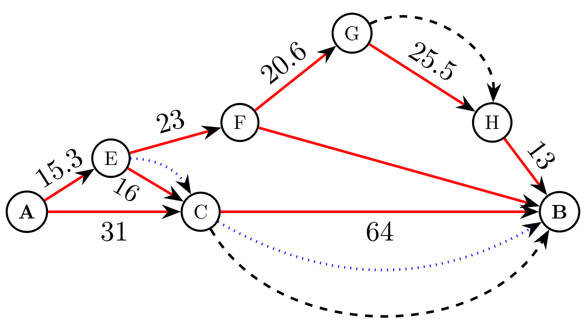
where $\mathbf{x} \in \Re^n$ and $\mathbf{y} \in \Re^n$ represent decision variables, which need to satisfy each constraint $i = 1, \ldots, m$ for every realization from the set $\mathcal{U}_i(\mathbf{x})$ that bounds the uncertain parameter $\boldsymbol{\xi}_i$.

To show the range of applicability of this framework, we illustrate two practical 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 for many real-world problems. In order to improve solutions, decisions on uncertainty reduction have to be included into the optimization problem. In this new setting, the uncertainty set is a function of decisions on acquiring additional information.

Example 2: Shortest Path on a Network. This example is intended to provide intuition for decision-dependent RO problems. Consider the graph in Figure 1 with the arcset $\mathcal{A}$ and let the uncertain length for arc e be $d_e = \bar{d}_e(1+0.5\xi_e)$, where $\bar{d}_e$ denotes the nominal value. The uncertain component ξ_e lies in $\mathcal{U}(\mathbf{x}) = \left\{\boldsymbol{\xi} \mid 0 \le \xi_e \le 1 - 0.8x_e,\ \sum_{e\in\mathcal{A}} \xi_e \le 1\right\}$. When $x_e = 0$, the maximum possible ξ_e is 1, whereas when $x_e = 1$, the maximum possible ξ_e reduces to 0.2. For simplicity, we assume the reduction to be possible for at most one of the arcs.



Shortest Path	**Path**	**Nominal**	**Worstcase**
Nominal	A–C–B	95	$31 + 1.5 \times 64 =$ **127**
Robust	A–E–F–G–H–B	97.4	$15.3 + 23 + 20.6 + 1.5 \times 25.5 + 13$ $=$ **110.15**
Endogenous Robust	A–E–C–B	95.3	$15.3 + 1.4 \times 16 + 1.1 \times 64 =$ **108.1**

Figure 1 Shortest path problem on a network. Nominal arc 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. The nominally shortest path A–C–B lengthens in the worst-case to 127 units. Standard RO optimizes against this case, resulting in A–E–F–G–H–B with an increased nominal length (price of robustness) but reduced worst-case of 110.15. If strengthening an arc is permitted, it is possible to reduce its uncertainty. When $x_{C-B} = 1$, the robust optimal path becomes A–E–C–B, reducing the worst-case cost to 108.5. This example demonstrates that decision-dependent sets can be leveraged to model decisions which mitigate the worst-case scenario.

The journey of RO has also included measures to reduce conservatism. The original RO formulation by Soyster (1973) produced over conservative solutions for many applications due to the use of box uncertainties. Later, Ben-Tal and Nemirovski (1999) provided less conservative solutions by using general polyhedral and ellipsoid uncertainty sets. ARO models and decision rule approximations took another step in this direction by allowing decisions to depend on the realizations. In this vein, decision-dependent uncertainty sets offer a new avenue to reduce the level of conservatism. For example, Poss (2013) decreases it for cardinality constrained sets. In this work, we introduce the concept of *proactive uncertainty control* by using decision-dependent sets to enable deliberate uncertainty reduction.

The contributions of this paper can be summarized as follows:

1. We study robust linear optimization problems with polyhedral decision-dependent uncertainty sets. In these sets, the parameter uncertainty depends on decisions. 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. 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 a modification to the Big-M linearization with a reduced number of constraints. This improves the performance over standard formulations.

This paper provides a first step towards RO solutions for endogenous uncertainties. Despite the wide prevalence of such uncertainties in many real-world settings, these problems have not received much attention in the literature, due largely to computational intractabilities. This work also offers a deeper understanding of decision dependence and 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. The case of continuous decisions influencing the uncertainty is not discussed. 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 2 provides the decision-dependent robust linear optimization problem and proves it to be NP-complete. Section 3 introduces the class of uncertainty sets discussed in this paper and provides reformulations. For comparison, Section 4 provides the corresponding Big-M formulation which can model more general sets. It also provides methods to improve these standard techniques. A numerical experiment is discussed in Section 5 to illustrate the advantages of the decision-dependent setting and to computationally compare the three formulations. Any proofs not presented immediately are relegated to the electronic companion.

Notation. Throughout this paper, we use bold lower and uppercase letters to denote vectors and matrices. Scalars are marked in regular font. All vectors are column vectors and the vector of ones is denoted by $\mathbf{e}$. Furthermore, $diag(\bullet)$ 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 $\mathbf{A}_{i,\bullet}$ and the j^{th} column is denoted by $\mathbf{A}_{\bullet,j}$. The problems have m constraints indexed by i. LHS denotes left-hand-side and RHS denotes 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. General Decision Dependence

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

Definition 1. The set with constraint matrix $\mathbf{D}$, constant vector $\mathbf{d}$, and decision coefficient matrix $\mathbf{\Delta}$ given by

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

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

Note that $\mathbf{\Delta}$ 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 5, 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 1. *The robust linear problem* (RO-DDU) *with uncertainty set* $\mathcal{U}^P$ *is NP-complete.*

The complete proof is given in the appendix, but it can be sketched by the following steps:

1. Consider an instance of the 3-Satisfiability problem (3-SAT) for a set $N = \{1, 2, \ldots, n\}$ of literals and m clauses, which seeks to find a solution $\mathbf{x} \in \{0,1\}^n$ that satisfies
$$x_{i_1} + x_{i_2} + (1 - x_{i_3}) \geq 1 \ \forall i = 1, \ldots, m.$$
2. Consider the following special case of (RO-DDU) with $\mathbf{x} \in \Re^n$, $\mathbf{y} \in \Re^m$, $z \in \Re$
$$\min_{\mathbf{x},\mathbf{y},z\geq 0} \left\{-z \mid z - \mathbf{a}^\top \mathbf{y} \leq 0, \ \forall \mathbf{a} \in \mathcal{U}(\mathbf{x}), \ \mathbf{x}, \mathbf{y} \leq \mathbf{e}, \ -\mathbf{y} \leq -\mathbf{e}\right\}, \qquad \text{(RO-SAT)}$$
where $\mathcal{U}(\mathbf{x}) = \{(a_1, \ldots, a_m) \mid a_i \geq x_{i_1} , a_i \geq x_{i_2} , a_i \geq 1 - x_{i_3} , a_i \leq 1\}$. Note that the 3-SAT problem is embedded in this set.
3. 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 $\mathcal{U}^P$.
5. Since the 3-SAT problem is NP-complete (Cook 1971), problem (RO-DDU) is NP-complete.

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 (e.g., Kolodziej et al. 2013). For binary decision variables $\mathbf{x}$ influencing $\mathcal{U}(\mathbf{x})$, the problem (RO-DDU) can be reformulated as an MILP, using the Big-M method (see Section 4). However, they suffer from weak relaxations.

3. 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 is expressed in the set:

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

Here, $\mathbf{D} \in \Re^{m\times n}$ is a coefficient matrix, $\mathbf{d} \in \Re^m$ is the RHS vector, $\mathbf{v} \in \Re^n$ are the minimum upper bounds, and $\mathbf{W} = diag(\mathbf{w}) \in \Re^{n\times n}$ (a diagonal matrix) are the incremental upper bounds, which apply when reduction is not applied. For $\mathcal{U}^{\overline{\mathbf{\Pi}}}$, the influence variable is $\mathbf{x} \in \{0,1\}^n$. The decision dependence in $\mathcal{U}^{\overline{\mathbf{\Pi}}}$ affects the upper bounds on each uncertain component ξ_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 the introductory example and will discuss it further in the numerical application.

We now discuss how such structures 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.

3.1. $\overline{\mathbf{\Pi}}$-Uncertainty

For succinctness, this section provides a reformulation of the following linear constraint

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

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

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

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

where in problem (P), $\boldsymbol{\pi}(\mathbf{x},\mathbf{y})$ denotes the corresponding dual variable. Problem (P) maximizes the LHS directly over $\mathcal{U}^{\overline{\mathbf{\Pi}}}(\mathbf{x})$. 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'), as suggested by Cormican et al. (1998) in the context of stochastic network interdiction. Proposition 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 matrix $\overline{\mathbf{\Pi}} = diag(\overline{\boldsymbol{\pi}})$. Here, $\overline{\boldsymbol{\pi}}$ is a component-wise 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 $\overline{\mathbf{\Pi}}$ 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 $\overline{\boldsymbol{\pi}}$ in Proposition 2. The dual problems of (P) and (P') are given by:

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

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

PROPOSITION 1. *If the set $\mathcal{U}^{\overline{\mathbf{\Pi}}}(\mathbf{x})$ is nonempty and $\mathbf{v},\mathbf{W} \geq 0$, then for all $\mathbf{x},\mathbf{y}$ with a binary $\mathbf{x}$:*

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

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 $(\mathbf{r} = \boldsymbol{\pi} - \mathbf{\Pi}\mathbf{x}, \mathbf{s} = \boldsymbol{\pi}, \mathbf{t} = \mathbf{q})$ with $\mathbf{\Pi} =$

$diag(\boldsymbol{\pi})$ be a potential feasible solution to (D'). It follows that $\mathbf{s}^\top + \mathbf{t}^\top\mathbf{D} = \boldsymbol{\pi}^\top + \mathbf{q}^\top\mathbf{D} \geq \mathbf{y}^\top$, and $\mathbf{r}^\top + \mathbf{t}^\top\mathbf{D} = \boldsymbol{\pi}^\top - \mathbf{x}^\top\boldsymbol{\Pi} + \mathbf{q}^\top\mathbf{D} \geq \mathbf{y}^\top - \mathbf{x}^\top\boldsymbol{\Pi} \geq \mathbf{y}^\top - \mathbf{x}^\top\overline{\boldsymbol{\Pi}}$. 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 $\mathbf{s}^\top + \mathbf{t}^\top\mathbf{D} \geq \mathbf{y}^\top$. Therefore, $\boldsymbol{\pi}^\top + \mathbf{q}^\top\mathbf{D} = \mathbf{s}^\top + \mathbf{t}^\top\mathbf{D} \geq \mathbf{y}^\top$, 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 $\sum_i w_i(s_i - r_i - s_i x_i)$. 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 $i \in X_0$:

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

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

$$\begin{aligned}l(\mathbf{t}) = \min_{r_i, s_i, \, i \in X_0} \; & \sum_{i \in X_0} r_i w_i + \sum_{i \in X_0} s_i v_i \\ \text{s.t.} \; & \left.\begin{aligned} & s_i + \mathbf{t}^\top\mathbf{D}_{\cdot,i} \geq y_i \\ & r_i + \mathbf{t}^\top\mathbf{D}_{\cdot,i} \geq y_i \\ & r_i, s_i \geq \mathbf{0} \end{aligned}\right\} \forall \, i \in X_0.\end{aligned}$$

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, the values of s_i and r_i at optimality are equal and set to their lower bounds $s_i = r_i = \max\{y_i - \mathbf{t}^\top \mathbf{D}_{\bullet,i}, 0\}$. Hence $\sum_{i \in X_0} s_i w_i - r_i w_i = 0$, which means $h(\mathbf{x},\mathbf{y}) \leq \bar{h}(\mathbf{x},\mathbf{y})$. □

Using Proposition 1 and problem (D'), the constraint (LC) can be reformulated as

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

Note that this reformulation does not contain any bilinear terms and includes fewer constraints than the standard Big-M formulations. Additionally, Proposition 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 $\overline{\mathbf{\Pi}}$.

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

Proposition 2. *If $\mathbf{D}$ and $\mathbf{y}$ are component-wise nonnegative, then $\pi_i(\mathbf{x},\mathbf{y}) \leq y_i \ \forall (\mathbf{x},\mathbf{y})$ for constraint* (LC) *under the uncertainty set $\mathcal{U}^{\overline{\mathbf{\Pi}}}$.*

This proposition allows us to estimate $\overline{\pi}_i$ by setting it equal to the maximum value taken by y_i in the overall problem. In some cases, such as shortest path or facility location problems, this is easily estimated from the underlying model. With this, all components of the decision-dependent problem with the polyhedral uncertainty set $\mathcal{U}^{\overline{\mathbf{\Pi}}}$ can be computed efficiently for practical size problems. We now extend Proposition 1 to more general uncertainty sets.

3.2. Extension to conic sets

Given a cone $\mathcal{K}$, the decision-dependent uncertainty set $\mathcal{U}^{\overline{\mathbf{\Pi}}}(\mathbf{x})$ 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 $\mathbf{W} = diag(\mathbf{w})$ denote upper bounds to the uncertain component $\boldsymbol{\xi}$. The objective is to reformulate the constraint $\mathbf{y}^\top \boldsymbol{\xi} \le b, \ \forall \boldsymbol{\xi} \in \mathcal{U}^{\mathcal{K}}(\mathbf{x})$. In order to satisfy this constraint for all $\boldsymbol{\xi} \in \mathcal{U}^{\mathcal{K}}(\mathbf{x})$, its LHS can be expressed with the following two problems:

$$\begin{aligned} h(\mathbf{x},\mathbf{y}) = \max_{\boldsymbol{\xi}} \ & \mathbf{y}^\top \boldsymbol{\xi} \\ \text{s.t.} \ & \mathbf{d} - \mathbf{D}\boldsymbol{\xi} \in \mathcal{K} \\ & \boldsymbol{\xi} \le \mathbf{v} + \mathbf{W}(\mathbf{e} - \mathbf{x}) \ : \boldsymbol{\pi}(\mathbf{x},\mathbf{y}) \\ & \boldsymbol{\xi} \ge \mathbf{0}, \end{aligned} \tag{KP}$$

$$\begin{aligned} \bar{h}(\mathbf{x},\mathbf{y}) = \max_{\boldsymbol{\xi},\boldsymbol{\zeta}} \ & (\mathbf{y} - \overline{\boldsymbol{\Pi}}\mathbf{x})^\top \boldsymbol{\xi} + \mathbf{y}^\top \boldsymbol{\zeta} \\ \text{s.t.} \ & \mathbf{d} - \mathbf{D}\boldsymbol{\xi} \in \mathcal{K} \\ & \boldsymbol{\xi} \le \mathbf{W}\mathbf{e} \\ & \boldsymbol{\zeta} \le \mathbf{v} \\ & \boldsymbol{\xi}, \boldsymbol{\zeta} \ge \mathbf{0}. \end{aligned} \tag{KP'}$$

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

PROPOSITION 3. *If* $\forall \mathbf{x} \in \{0,1\}^n$ *there exists a point in the relative interior of* $\mathcal{U}^{\mathcal{K}}(\mathbf{x})$ *(Slater point) and* $\mathbf{v}, \mathbf{W} \ge 0$*, then for all* $\mathbf{x}, \mathbf{y}$*:*

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

Using Proposition 3 and the dual problem of (KP'), the constraint (LC) can be reformulated as

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

with the dual cone $\mathcal{K}^*$. Note that this reformulation has only linear terms and, as we will see in Section 4, fewer constraints than the standard Big-M formulation, hence it is more suitable for larger sized problems. The proof of this formulation proceeds parallel to that of Proposition 1.

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 $\mathcal{U}^{\mathbf{\Pi}}$. The following section provides a reformulation of problem (RO-DDU) under general polyhedral uncertainty sets.

4. 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 4. *If the uncertainty set $\mathcal{U}_i(\mathbf{x})$ is a polyhedron as in $\mathcal{U}^P(\mathbf{x})$ with $\boldsymbol{D}_i \in \Re^{m_i \times p}$, $\mathbf{d}_i \in \Re^{m_i}$, and $\boldsymbol{\Delta}_i \in \Re^{m_i \times n}$ and if $\mathbf{x}$ is binary, then the robust counterpart of problem* (RO-DDU) *is*

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

where M is a sufficiently large number.

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 (Ben-Tal et al. 2015), which includes conic and budgeted structures. Their proofs (not shown) proceed parallel to that of Proposition 4.

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 4 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.

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	$(n+1)s$	$p+mK+nmK(s+1)$	$m+mp+2n+nmK(3s+1)$	$mK(ns+1)$

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) $\mathbf{x}\in\Re^n$ 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.

4.1. Modified Big-M Reformulation

Consider the uncertainty set $\mathcal{U}^{\mathrm{P}}(\mathbf{x})$ to be expressed as

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

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 $\boldsymbol{\Delta}$ are non-negative. Proposition 5 reformulates constraint (LC) for $\mathcal{U}^{\mathrm{P}}(\mathbf{x})$ under this assumption.

Proposition 5. *If $\Delta_{ij} \ge 0\ \forall i,j$, then the constraint* (LC) *with the uncertainty set $\mathcal{U}^P(\mathbf{x})$ and a large constant M can be reformulated as*

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

This proposition 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. 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 5. For a comparison, we reformulate the constraint (LC) over the uncertainty set $\mathcal{U}^{\overline{\mathbf{\Pi}}}(\mathbf{x})$ using all three presented techniques, namely (i) $\overline{\mathbf{\Pi}}$, (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

Formulations	Problem	Variables	Constraints
$\overline{\mathbf{\Pi}}$	$\mathbf{t}^\top\mathbf{d} + \mathbf{r}^\top\mathbf{W}\mathbf{e} + \mathbf{s}^\top\mathbf{v} \leq b$ $\mathbf{s}^\top + \mathbf{t}^\top\mathbf{D} \geq \mathbf{y}^\top$ $\mathbf{r}^\top + \mathbf{t}^\top\mathbf{D} \geq \mathbf{y}^\top - \mathbf{x}^\top\overline{\mathbf{\Pi}}$ $\mathbf{r}, \mathbf{s}, \mathbf{t} \geq \mathbf{0}.$	C: $m+2n$	A: $1+2n$ S: $m+2n$
Big-M	$\mathbf{t}^\top\mathbf{d} + \mathbf{s}^\top\mathbf{v} + \boldsymbol{s}^\top\mathbf{W}\mathbf{e} - \sum_i r_i \leq b$ $\mathbf{s}^\top + \mathbf{t}^\top\mathbf{D} \geq \mathbf{y}^\top$ $w_i s_i - M(1-x_i) \leq r_i \leq M x_i$ $r_i \leq w_i s_i$ $\mathbf{r}, \mathbf{s}, \mathbf{t} \geq \mathbf{0}.$	C: $m+2n$	A: $1+4n$ S: $m+2n$
Modified Big-M	$\mathbf{t}^\top\mathbf{d} + \mathbf{s}^\top\mathbf{v} + \mathbf{r}^\top\mathbf{e} \leq b$ $\mathbf{s}^\top + \mathbf{t}^\top\mathbf{D} \geq \mathbf{y}^\top$ $r_i \geq w_i s_i - M x_i$ $\mathbf{r}, \mathbf{s}, \mathbf{t} \geq \mathbf{0}.$	C: $m+2n$	A: $1+2n$ S: $m+2n$

Table 2 Comparison of reformulations of (LC) for $\mathcal{U}^{\overline{\mathbf{\Pi}}}(\mathbf{x})$ and their sizes (C: Continuous, A: Affine, S: Sign)

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.

5. 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 routing problems can be solved with polynomial time algorithms (Dijkstra 1959). However, this does not hold for uncertain arc lengths. Past research on robust shortest path problems focused on scenario-based (Yu and Yang 1998), cardinality (Bertsimas and Sim 2003), and interval uncertainty (Averbakh and Lebedev 2004, Zieliński 2004). 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:

1. *Comparing the numerical performance of different robust formulations,*
2. *Measuring the benefit of proactive reduction as a function of size, budget, or cost of reduction,*
3. *Measuring the number of arcs in the shortest path as a function of size, budget, or cost,*
4. *Evaluating the price of robustness and the benefit of interacting with uncertainties, and*
5. *Comparing 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 sections of the network.

To illustrate this setting, we discuss a problem on a graph $G = (\mathcal{V}, \mathcal{A}, d(\cdot))$ for the set of nodes $\mathcal{V}$, arcs $\mathcal{A}$, and the distance function $d(\cdot)$. The objective is to find the shortest path from the source to the target node ($s \to t$) when the actual realized distances from node i to j are uncertain and a function $d_{ij}(\boldsymbol{\xi}) = \left(1 + \frac{1}{2}\,\xi_{ij}\right)\bar{d}_{ij}$ 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}. The parameter $\boldsymbol{\xi}$ resides in a cardinality constrained uncertainty set with reducible upper bounds. The complete problem is given by

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

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

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

We solve problem (SP) using the three different formulations: (i) $\overline{\boldsymbol{\Pi}}-$formulation from Proposition 1, (ii) standard Big-M formulation, and (iii) Modified Big-M formulation from Proposition 5.

Formulation	Problem	Variables	Constraints
$\overline{\boldsymbol{\Pi}}$	$\min_{\mathbf{x},\mathbf{y},\mathbf{q},\mathbf{r},p} f(\mathbf{x},\mathbf{y}) + p\Gamma + \sum_{(i,j)\in\mathcal{A}} q_{ij}(1-\gamma_{ij}) + \sum_{(i,j)\in\mathcal{A}} r_{ij}\gamma_{ij}$ s.t. $p + q_{ij} \ge \frac{y_{ij}d_{ij} - \overline{\pi}_{ij}d_{ij}x_{ij}}{2}$, $p + r_{ij} \ge \frac{y_{ij}d_{ij}}{2}$, $p, q_{ij}, r_{ij} \ge 0,\ \mathbf{x},\mathbf{y} \in X \times Y.$	B: $2\|\mathcal{A}\|$ C: $2\|\mathcal{A}\|+1$	A: $\|\mathcal{V}\|+2\|\mathcal{A}\|$ S: $2\|\mathcal{A}\|+1$
Big-M	$\min_{\mathbf{x},\mathbf{y},\mathbf{q},\mathbf{r},p} f(\mathbf{x},\mathbf{y}) + p\Gamma + \sum_{(i,j)\in\mathcal{A}} q_{ij} - \sum_{(i,j)\in\mathcal{A}} \gamma_{ij}r_{ij}$ s.t. $p + q_{ij} \ge \frac{d_{ij}y_{ij}}{2}$, $0 \le r_{ij} \le Mx_{ij}$, $q_{ij} - M(1-x_{ij}) \le r_{ij} \le q_{ij}$, $p, q_{ij}, r_{ij} \ge 0,\ \mathbf{x},\mathbf{y} \in X \times Y.$	B: $2\|\mathcal{A}\|$ C: $2\|\mathcal{A}\|+1$	A: $\|\mathcal{V}\|+4\|\mathcal{A}\|$ S: $2\|\mathcal{A}\|+1$
Modified Big-M	$\min_{\mathbf{x},\mathbf{y},\mathbf{q},\mathbf{r},p} f(\mathbf{x},\mathbf{y}) + p\Gamma + \sum_{(i,j)\in\mathcal{A}} r_{ij} + \sum_{(i,j)\in\mathcal{A}} q_{ij}(1-\gamma_{ij})$ s.t. $p + q_{ij} \ge \frac{d_{ij}y_{ij}}{2}$, $r_{ij} \ge \gamma_{ij} - Mx_{ij}$, $p, q_{ij}, r_{ij} \ge 0,\ \mathbf{x},\mathbf{y} \in X \times Y.$	B: $2\|\mathcal{A}\|$ C: $2\|\mathcal{A}\|+1$	A: $\|\mathcal{V}\|+2\|\mathcal{A}\|$ S: $2\|\mathcal{A}\|+1$

Table 3 Shortest path formulations for $\mathcal{U}^{SP}(\mathbf{x})$ and their sizes (B: Binary, C: Continuous, A: Affine, S: Sign).

In Table 3, $X \times Y$ denote 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 in Table 2.

Experiment 1: Performance Comparison The numerical setup is as follows. We randomly generate points on a 100×100 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 Γ 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 $|\mathcal{V}| = \{50, 75, \ldots, 250\}$, 100 random graphs are generated. 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

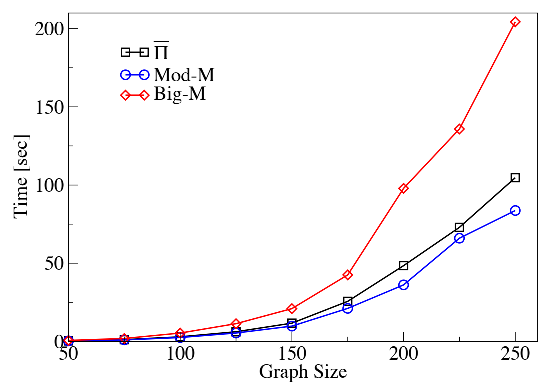


Figure 2 Comparison of median solution times of reformulations from Propositions 1, 5, and the standard Big-M.

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 $\overline{\mathbf{\Pi}}$ and the Modified Big-M formulation than for the Big-M formulation.
- The difference between the Big-M and the proposed formulations increases with growing $|\mathcal{V}|$. This highlights the advantage of the $\overline{\mathbf{\Pi}}$ and Modified Big-M formulation for larger graphs.
- The median time of the Modified Big-M formulation is less than that of the $\overline{\mathbf{\Pi}}$-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 $\overline{\mathbf{\Pi}}$ formulations are comparable over a broad range of network sizes, the subproblem in the $\overline{\mathbf{\Pi}}$ reformulation is convex, which can be exploited by cut-generating methods, which may be computationally advantageous. We also solved the $\overline{\mathbf{\Pi}}$ 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 Γ and the level of uncertainty reduction γ. These parameters create three scenarios:

(i) the *nominal* case, where no uncertainty is present, $n^*(\Gamma = 0, \gamma = 0)$;

(ii) the *standard robust* case with no decision dependence, $n^*(\Gamma > 0, \gamma = 0)$; and

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

There are four parameters that govern the effect of interactions with uncertainty: $\gamma, |\mathcal{V}|, c$, and Γ. 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, the problem (SP) is implemented on randomly generated graphs of $[20-50]$ nodes. This size is comparable to moderately sized transportation networks (Montemanni and Gambardella 2005). 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 EC.1a shows that $\gamma > 0$ reduces z^* (shorter paths), which is independent of $|\mathcal{V}|$. The inset of Figure EC.1a 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 non-zero c. We also observe that z^* is independent of $|\mathcal{V}|$, which can be explained by the fact that $|\mathcal{V}|$ only increases from $20 - 50$ and $n^*(\Gamma > 0, \gamma > 0)$ does not change sizably over this range 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. Further discussion and results are available in the electronic companion.

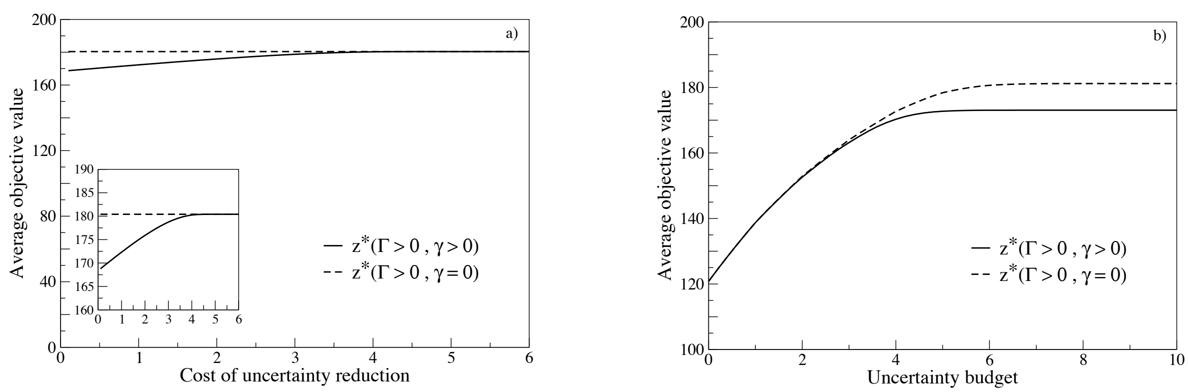


Figure 3 Average objective value as a function of: a) cost of uncertainty reduction c and b) maximum uncertainty Γ. The graph consists of $|\mathcal{V}| = 30$ nodes.

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. Figure 3a ($|\mathcal{V}| = 30$ 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 (difference between the dotted line in Figure 3a and $z^*(\Gamma = 0, \gamma = 0)$ in Figure 3b) is constant w.r.t. γ but changes with Γ. On the other hand, the benefit of interaction decreases with increase in c, as can be observed in Figure 4a.

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 ($c = 0$).

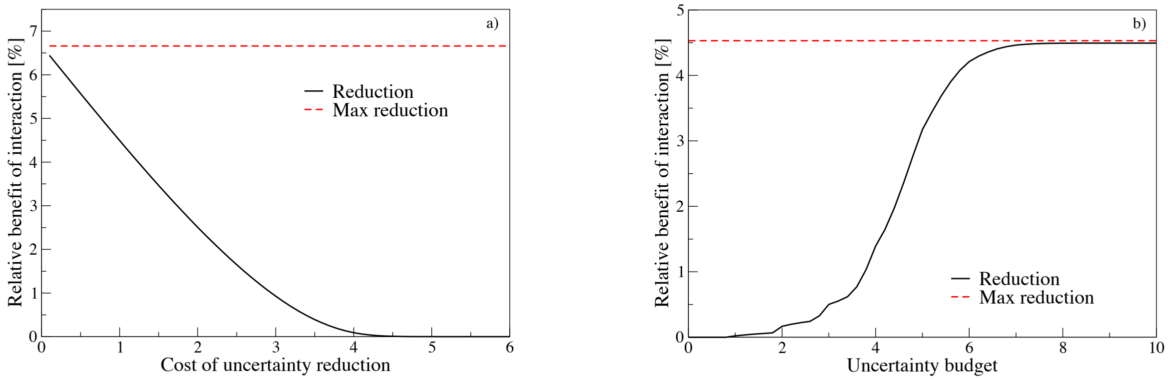


Figure 4 Average relative benefit of interaction as a function of: a) cost of uncertainty reduction c and b) maximum uncertainty Γ. The graph consists of $|\mathcal{V}| = 30$ nodes.

Experiment 5: Uncertainty Budget. Γ governs the number of arcs that can be affected by uncertainty. Figure 3b shows that z^* increases gradually with Γ 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 Γ 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 Γ and plateaus beyond a certain Γ (not shown). An analogous behavior can be observed for the benefit of interaction, as shown in Figure 4b. The maximum benefit is achieved at $c = 0$.

Figure 5a displays how the average n^* changes with Γ 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 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 Γ increases, it turns 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 Γ 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 $\gamma > 0$,

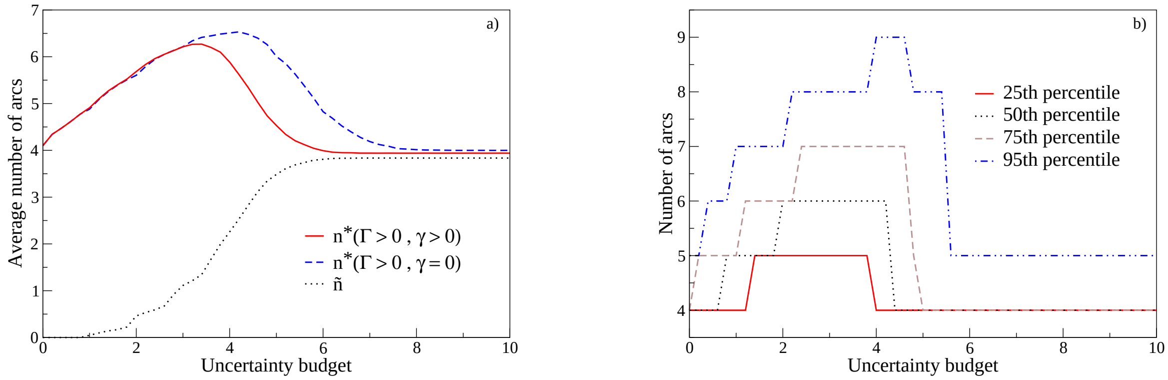


Figure 5 The dependence on the budget of uncertainty Γ for: a) average number of arcs and b) their distribution. The graph consists of $|\mathcal{V}| = 30$ nodes and uncertainty reduction is permitted.

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 5a. Note that for small Γ, 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 5b displays the distribution of the number of arcs using different percentiles of $n^*(\Gamma > 0, \gamma > 0)$ (corresponding to Figure 5a). Here, we observe that as Γ increases, the distribution of $n^*(\Gamma > 0, \gamma > 0)$ skews towards 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 DDU solutions and compares them to a similar SO problem. The details of the SO formulation are provided in the electronic companion. The average performance is evaluated by randomly generating the uncertain component ξ_{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 "-W."

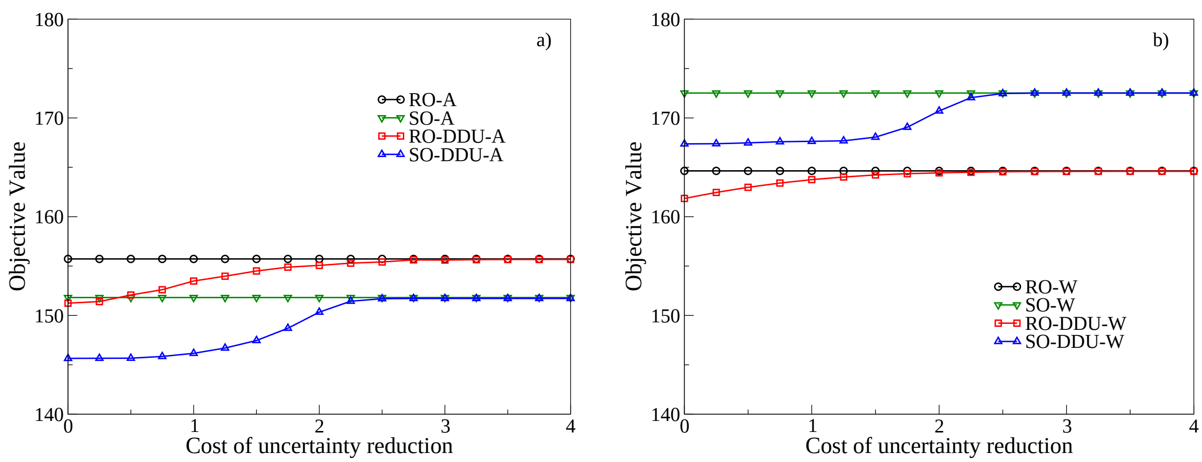


Figure 6 Comparison of RO and SO formulations : a) average and b) worst-case objective value.

Figure 6a 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 6b, RO-W is significantly less 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 $\overline{\mathbf{\Pi}}$-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.

6. 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 reformulations 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 sub-problem 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 envision decision-dependent sets to solidify the tradeoff between computation and optimality by inducing beneficial cuts.

Acknowledgments

We are grateful to David Morton for insightful comments.

References

Averbakh, Igor, Vasilij Lebedev. 2004. Interval data minmax regret network optimization problems. *Discrete Applied Mathematics* **138**(3) 289–301.

Ben-Tal, Aharon, Dick Den Hertog, Jean-Philippe Vial. 2015. Deriving robust counterparts of nonlinear uncertain inequalities. *Mathematical Programming* **149**(1-2) 265–299.

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

Ben-Tal, Aharon, Alexander Goryashko, Elana Guslitzer, Arkadi Nemirovski. 2004. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* **99**(2) 351–376.

Ben-Tal, Aharon, Arkadi Nemirovski. 1999. Robust solutions of uncertain linear programs. *Operations Research Letters* **25**(1) 1–13.

Bertsimas, Dimitris, David B Brown, Constantine Caramanis. 2011. Theory and applications of robust optimization. *SIAM Review* **53**(3) 464–501.

Bertsimas, Dimitris, Constantine Caramanis. 2010. Finite adaptability in multistage linear optimization. *IEEE Transactions on Automatic Control* **55**(12) 2751–2766.

Bertsimas, Dimitris, Vineet Goyal, Brian Y Lu. 2015. A tight characterization of the performance of static solutions in two-stage adjustable robust linear optimization. *Mathematical Programming* **150**(2) 281–319.

Bertsimas, Dimitris, Omid Nohadani, Kwong Meng Teo. 2010. Robust optimization for unconstrained simulation-based problems. *Operations Research* **58**(1) 161–178.

Bertsimas, Dimitris, Melvyn Sim. 2003. Robust discrete optimization and network flows. *Mathematical Programming* **98**(1) 49–71.

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

Bertsimas, Dimitris, John N Tsitsiklis. 1997. *Introduction to linear optimization*, vol. 6. Athena Scientific Belmont, MA.

Bertsimas, Dimitris, Phebe Vayanos. 2015. Data-driven learning in dynamic pricing using adaptive optimization URL `http://www.optimization-online.org/DB_HTML/2014/10/4595.html`.

Chu, Millie, Yuriy Zinchenko, Shane G Henderson, Michael B Sharpe. 2005. Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty. *Physics in Medicine and Biology* **50**(23) 5463.

Cook, Stephen A. 1971. The complexity of theorem-proving procedures. *Proceedings of the Third Annual ACM Symposium on Theory of Computing*. ACM, 151–158.

Cormican, Kelly J, David P Morton, R Kevin Wood. 1998. Stochastic network interdiction. *Operations Research* **46**(2) 184–197.

Dantzig, George B. 1955. Linear programming under uncertainty. *Management Science* **1**(3-4) 197–206.

Dijkstra, Edsger W. 1959. A note on two problems in connexion with graphs. *Numerische Mathematik* **1**(1) 269–271.

Georghiou, Angelos, Wolfram Wiesemann, Daniel Kuhn. 2015. Generalized decision rule approximations for stochastic programming via liftings. *Mathematical Programming* **152**(1-2) 301–338.

Ghaoui, Laurent El, Maksim Oks, Francois Oustry. 2003. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research* **51**(4) 543–556.

Goel, Vikas, Ignacio E Grossmann. 2004. A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. *Computers & Chemical Engineering* **28**(8) 1409–1429.

Goel, Vikas, Ignacio E Grossmann. 2006. A class of stochastic programs with decision dependent uncertainty. *Mathematical Programming* **108**(2-3) 355–394.

Hanasusanto, Grani A, Daniel Kuhn, Wolfram Wiesemann. 2015. K-adaptability in two-stage robust binary programming. *Operations Research* **63**(4) 877–891.

Iancu, Dan Andrei. 2010. Adaptive robust optimization with applications in inventory and revenue management. Ph.D. thesis, Massachusetts Institute of Technology.

Jonsbråten, Tore W, Roger JB Wets, David L Woodruff. 1998. A class of stochastic programs with decision dependent random elements. *Annals of Operations Research* **82** 83–106.

Kolodziej, Scott, Pedro M Castro, Ignacio E Grossmann. 2013. Global optimization of bilinear programs with a multiparametric disaggregation technique. *Journal of Global Optimization* **57**(4) 1039–1063.

Lorca, Alvaro, X Andy Sun, Eugene Litvinov, Tongxin Zheng. 2016. Multistage adaptive robust optimization for the unit commitment problem. *Operations Research* **64**(1) 32–51.

Montemanni, Roberto, Luca Maria Gambardella. 2005. The robust shortest path problem with interval data via benders decomposition. *4OR* **3**(4) 315–328.

Novoa, Luis J, Ahmad I Jarrah, David P Morton. 2016. Flow balancing with uncertain demand for automated package sorting centers. *Transportation Science* doi:10.1287/trsc.2015.0662.

Poss, Michael. 2013. Robust combinatorial optimization with variable budgeted uncertainty. *4OR* **11**(1) 75–92.

Shapiro, Alexander, Darinka Dentcheva, Andrzej Ruszczynski. 2009. *Lectures on Stochastic Programming*. Society for Industrial and Applied Mathematics.

Soyster, Allen L. 1973. Technical non-convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research* **21**(5) 1154–1157.

Spacey, Simon A, Wolfram Wiesemann, Daniel Kuhn, Wayne Luk. 2012. Robust software partitioning with multiple instantiation. *INFORMS Journal on Computing* **24**(3) 500–515.

Vujanic, Robin, Paul Goulart, Manfred Morari. 2016. Robust optimization of schedules affected by uncertain events. *Journal of Optimization Theory and Applications* 1–22.

Yu, Gang, Jian Yang. 1998. On the robust shortest path problem. *Computers & Operations Research* **25**(6) 457–468.

Zieliński, Paweł. 2004. The computational complexity of the relative robust shortest path problem with interval data. *European Journal of Operational Research* **158**(3) 570–576.

E-Companion

Proof to Theorem 1

It is sufficient to prove the following Lemma.

LEMMA EC.1. *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: ($\Longrightarrow$) Suppose the 3-SAT problem has a feasible solution $\mathbf{x}$. This means, $\mathbf{x}$ has to satisfy

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

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 $\mathcal{U}(\mathbf{x}) = \{(a_1, \ldots a_m) \mid a_i \geq x_{i_1}\, ,\, a_i \geq x_{i_2}\, ,\, a_i \geq 1 - x_{i_3}\, ,\, a_i \leq 1 \;\; \forall i = 1, \ldots, m\}$. Since at least one of x_{i_1} or x_{i_2} or $1 - x_{i_3}$ equals 1, a_i satsifies $a_i \geq 1$. This implies that $a_i = 1\, \forall i$ 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 $z \geq m$.

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

We now show that this result implies that for each i at least one of x_{i_1} or 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 $a'_i = \max\{x_{i_1}, x_{i_2}, (1 - x_{i_3})\}$ which is an element of the uncertainty set and $a'_i < 1$. However, this contradicts the result of $a_i = 1 \;\; \forall i$. Therefore, if $z^* = m$, then we can find a feasible solution for the 3-SAT problem. □

Proof to Proposition 2

Proof: Consider the following problem for some i

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

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 (EC.1), i.e., $\mathbf{b} = [\mathbf{d}^\top, \mathbf{v}^\top + (\mathbf{e} - \mathbf{x})^\top \mathbf{W}]^\top$. Assume that the solution is non-degenerate. This means $\mathbf{B}^{-1}\mathbf{b} > 0$. 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 ϵ to obtain a non-degenerate solution, which only marginally changes the objective (see, e.g., Bertsimas and Tsitsiklis 1997).

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

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

which represents the change in the objective value. Let $\boldsymbol{\xi}_0$ be the optimal solution of the problem with $\theta = 0$ and $\boldsymbol{\xi}_\theta$ be the optimal solution of 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 EC.2, we can state that

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

This implies that $\pi_{0,i} \le y_i \ \forall i$. As a consequence, the dual variables can be bound using $\mathbf{y}$, as

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

□

Lemma EC.2. *If the matrix* $\mathbf{D}$ *is element-wise greater than* 0*, then* $\boldsymbol{\xi}_\theta \le \boldsymbol{\xi}_0 + \theta \mathbf{e}_i$.

Proof: Suppose this is not true, i.e., there exists at least one index k such that $\boldsymbol{\xi}_{\theta,k} > \boldsymbol{\xi}_{0,k} + \theta \mathbf{e}_{i,k}$. In addition, it holds that for $\theta \ge 0$, $\mathbf{y}^\top \boldsymbol{\xi}_\theta > \mathbf{y}^\top \boldsymbol{\xi}_0$.

If $k \ne i$, then $\boldsymbol{\xi}_\theta \le \mathbf{v} + \mathbf{W}(\mathbf{e} - \mathbf{x})$, which suggests $\boldsymbol{\xi}_\theta$ to be feasible for the problem with $\theta = 0$. This would contradict the optimality of $\boldsymbol{\xi}_0$.

If $k = i$, then $\boldsymbol{\xi}_{\theta,i} > \boldsymbol{\xi}_{0,i} + \theta$. However this results in $\boldsymbol{\xi}_0 < \boldsymbol{\xi}_\theta - \theta \mathbf{e}_i \le \mathbf{v} + \mathbf{W}(\mathbf{e} - \mathbf{x})$. Since $D(\boldsymbol{\xi}_\theta - \theta \mathbf{e}_i) = \mathbf{D}\boldsymbol{\xi}_\theta - \theta \mathbf{D}\mathbf{e}_i \le \mathbf{d} - \theta \mathbf{D}\mathbf{e}_i \le \mathbf{d}$, $\boldsymbol{\xi}_\theta - \theta \mathbf{e}_i$ 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 $\boldsymbol{\xi}_\theta \le \boldsymbol{\xi}_0 + \theta \mathbf{e}_i$. □

Proof to Proposition 3

Proof: Consider the duals of the problems (KP) and (KP').

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

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

By assumption of existence of a Slater point, the strong duality holds and hence the optimal objectives of the primal and dual problems are equal. Though we work with the dual problems, we will be referring to the objective values as $h(\mathbf{x},\mathbf{y})$ and $\bar{h}(\mathbf{x},\mathbf{y})$. The proof uses the solution of each problem to create a feasible solution to the other problem and show the inequalities between $h(\bullet,\bullet) \le \bar{h}(\bullet,\bullet)$ and $\bar{h}(\bullet,\bullet) \le h(\bullet,\bullet)$. Let $(\boldsymbol{\pi}, \mathbf{q})$ be a solution to problem (KD). Then $(\mathbf{r} = \boldsymbol{\pi} - \boldsymbol{\Pi}\mathbf{x}, \mathbf{s} = \boldsymbol{\pi}, \mathbf{t} = \mathbf{q})$ is a potential feasible solution to (KD'), where $\boldsymbol{\Pi} = diag(\boldsymbol{\pi})$. The first constraint is satisfied since $\boldsymbol{\pi}^\top - \mathbf{x}^\top\boldsymbol{\Pi} + \mathbf{q}^\top\mathbf{D} \ge \mathbf{y}^\top - \mathbf{x}^\top\boldsymbol{\Pi} \ge \mathbf{y}^\top - \mathbf{x}^\top\overline{\boldsymbol{\Pi}}$. The second and third constraints are also satisfied because $\mathbf{q} \in \mathcal{K}^*$ and $\boldsymbol{\pi} \ge 0 \implies \boldsymbol{\pi} - \boldsymbol{\Pi}\mathbf{x} \ge \mathbf{0}$ for a binary $\mathbf{x}$. Hence, $(\mathbf{r} = \boldsymbol{\pi} - \boldsymbol{\Pi}\mathbf{x}, \mathbf{s} = \boldsymbol{\pi}, \mathbf{t} = \mathbf{q})$ is a feasible solution for the problem (KD'). By feasibility of $(\mathbf{r} = \boldsymbol{\pi} - \boldsymbol{\Pi}\mathbf{x}, \mathbf{s} = \boldsymbol{\pi}, \mathbf{t} = \mathbf{q})$ and since $\bar{h}(\mathbf{x}, \mathbf{y})$ is the optimal objective value, we can write

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

Consider a solution $(\mathbf{r}, \mathbf{s}, \mathbf{t})$ of problem (KD'). Then $(\mathbf{s}, \mathbf{t})$ is a solution to problem (KD). The feasibility requirements are immediately satisfied since $\mathbf{t} \in \mathcal{K}^*, \mathbf{s} \ge \mathbf{0}$ and $\mathbf{s}^\top + \mathbf{t}^\top\mathbf{D} \ge \mathbf{y}^\top$. Using this, the upper bound on $h(\mathbf{x}, \mathbf{y})$ can be expressed as

$$\begin{aligned}
h(\mathbf{x}, \mathbf{y}) &\le \mathbf{t}^\top\mathbf{d} + \mathbf{s}^\top\mathbf{v} + \mathbf{s}^\top\mathbf{W}(\mathbf{e} - \mathbf{x}) \\
&\le \bar{h}(\mathbf{x}, \mathbf{y}) - \mathbf{r}^\top\mathbf{W}\mathbf{e} + \mathbf{s}^\top\mathbf{W}\mathbf{e} - \mathbf{s}^\top\mathbf{W}\mathbf{x} \\
&\le \bar{h}(\mathbf{x}, \mathbf{y}) + \sum_{i=1}^{n} w_i(-r_i + s_i - s_i x_i).
\end{aligned}$$

For the values of x_i, the two cases can be distinguished as follows. If $x_i = 1$, then $w_i(-r_i + s_i - s_i x_i) = -w_i r_i \le 0$. If $x_i = 0$, then RHS $= -w_i r_i + w_i s_i$. For case $x_i = 0$, the optimization problem (KD') can be written as two nested minimization problems

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

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

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

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

From this inner minimization problem, we can conclude $\mathbf{r} = \mathbf{s}$. Since $\mathbf{W}$ and $\mathbf{v}$ are positive, at optimality, the values of r_i and s_i are set to their lower bounds of $r_i = s_i = \max\{y_i - \mathbf{t}^\top \mathbf{D}_{\boldsymbol{\cdot},i}, 0\}$. Therefore, $\sum_{i\in X_0} r_i w_i - s_i w_i = 0$, which yields,

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

Combining the two inequalities, we obtain $h(\mathbf{x}, \mathbf{y}) = \bar{h}(\mathbf{x}, \mathbf{y})$. □

Proof to Proposition 4

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

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

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 (EC.3) can now be expressed as

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

where $\boldsymbol{\pi}_i \in \Re^{m_i}$ 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 (EC.4) are satisfied. Therefore, the original problem (RO-DDU) can be written as

$$\left.\begin{aligned}
&\min_{\boldsymbol{\pi}_i,\mathbf{x},\mathbf{y}} \mathbf{c}^\top\mathbf{x}+\mathbf{f}^\top\mathbf{y} \\
&\quad \text{s.t. } \mathbf{a}_i^\top\mathbf{x}+\boldsymbol{\pi}_i^\top\mathbf{d}_i+\boldsymbol{\pi}_i^\top\boldsymbol{\Delta}_i\mathbf{x}\le b_i \\
&\qquad \boldsymbol{\pi}_i^\top\mathbf{D}_i=\mathbf{y}^\top \\
&\qquad \boldsymbol{\pi}_i\ge\mathbf{0}
\end{aligned}\right\}\forall i. \tag{EC.5}$$

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

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

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

$$w_{ijk}\le\pi_{ij},\ \ 0\le w_{ijk}\le Mx_k,\ \ w_{ijk}\ge\pi_{ij}-M(1-x_k),$$

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

$$\begin{aligned}
&\min_{\mathbf{x},\mathbf{y}}\ \mathbf{c}^\top\mathbf{x}+\mathbf{f}^\top\mathbf{y} \\
&\left.\begin{aligned}
&\text{s.t. } \mathbf{a}_i^\top\mathbf{x}+\boldsymbol{\pi}_i^\top\mathbf{d}_i+\sum_{j=1}^{m_i}\sum_{k=1}^{n}\Delta_{ijk}w_{ijk}\le b_i \\
&\quad \boldsymbol{\pi}_i^\top\mathbf{D}_i=\mathbf{y}^\top
\end{aligned}\right\}\ \forall i \\
&\left.\begin{aligned}
&\quad w_{ijk}\le Mx_k,\ \ w_{ijk}\le\pi_{ij} \\
&\quad w_{ijk}\ge\pi_{ij}-M(1-x_k) \\
&\quad \boldsymbol{\pi}_i\ge\mathbf{0},\ w_{ijk}\ge 0
\end{aligned}\right\}\ \forall i,j,k \\
&\quad \mathbf{x}\in\{0,1\}^n.
\end{aligned} \tag{EC.6}$$

Case 2: Problem (RO-DDU) is infeasible. Then the reformulation in (EC.6) is infeasible. To show this, consider the original problem (RO-DDU).

Suppose this problem is infeasible under the assumptions of Proposition 4. This means that

$\forall \mathbf{x} : \exists \boldsymbol{\xi} \in \mathcal{U}(\mathbf{x})$ such that $\mathbf{a}_i^\top \mathbf{x} + \boldsymbol{\xi}_i^\top \mathbf{y} > b_i$. This means that $\mathbf{a}_i^\top \mathbf{x} + \max_{\boldsymbol{\xi}_i \in \mathcal{U}_i(\mathbf{x})} \boldsymbol{\xi}_i^\top \mathbf{y} > b_i$ holds for at least one i. Using the dual of the inner problem, the constraints can be written $\forall \boldsymbol{\pi}_i$ as

$$\begin{aligned} &\mathbf{a}_i^\top \mathbf{x} + \boldsymbol{\pi}_i^\top (\mathbf{d}_i + \boldsymbol{\Delta}_i \mathbf{x}) > b_i \\ &\boldsymbol{\pi}_i^\top \mathbf{D}_i = \mathbf{y}^\top \\ &\boldsymbol{\pi}_i \geq \mathbf{0}. \end{aligned} \tag{EC.7}$$

Now, assume that the reformulation in (EC.6) is feasible. Given the constraints in (EC.6), 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 $\boldsymbol{\pi}_i = (\pi_{i1}, \pi_{i2}, \ldots, \pi_{ik}, \ldots, \pi_{m_i})$ 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 (EC.7) that there exist no such $\boldsymbol{\pi}_i$. □

Proof to Proposition 5

Proof: The LHS maximization problem can be written as

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

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

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

The above constraints can be rewritten by expanding the variable space as

$$\begin{aligned} &\sum_{i=1}^{m} \pi_i d_i + \sum_{i=1}^{m} \sum_{j=1}^{n} t_{ij} \leq b \\ &\pi_i \Delta_{ij} x_j \leq t_{ij} && \forall i, j \\ &\sum_{i=1}^{m} \pi_i D_{ij} = y_j && \forall j \\ &\boldsymbol{\pi} \geq \mathbf{0}. \end{aligned} \tag{EC.9}$$

If there is a variable $\boldsymbol{\pi}$ feasible for the set of equations given by (EC.8), then we can find a feasible variable for (EC.9) by $t_{ij} = \pi_i \Delta_{ij} x_j$. On the other hand, if there exists a feasible solution to (EC.9), then it is also feasible for (EC.8). 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 following set of constraints

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

Hence (EC.9) can be rewritten as

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

□

Numerical Experiments: Dependence on Graph Size

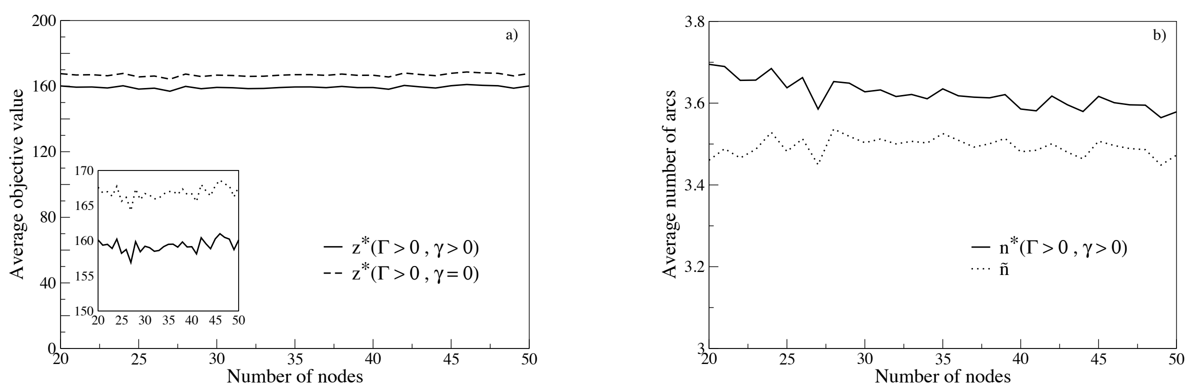


Figure EC.1 Dependence on graph size $|\mathcal{V}|$ for: a) average objective function and b) average number of arcs. The inset is a magnification.

Figure EC.1b 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 $|\mathcal{V}|$ as the number of arcs grows faster than the number of nodes, because in the experimental setup, only a fixed fraction of arcs are removed.

Stochastic Optimization formulation for DDU problem

Consider the arc lengths to be uncertain and given by the function $d:\mathcal{A}\times\mathcal{U}^{SSP}\rightarrow\Re_+$ of the form $d_e=\bar{d}_e\left(1+\frac{\xi_e}{2}\right)$, where $\boldsymbol{\xi}\in\mathcal{U}^{SSP}(\mathbf{x})$ with $\{x_e\}_{e\in\mathcal{A}}\in\{0,1\}^{|\mathcal{A}|}$. The stochastic uncertainty set is

$$\mathcal{U}^{SSP}(\mathbf{x})=\times_{e\in\mathcal{A}}[0,1-\gamma_e x_e].$$

The set $\mathcal{U}^{SSP}(\mathbf{x})$ is a Cartesian produce of intervals. We assume that the uncertainty $\boldsymbol{\xi}$ has a uniform distribution with support in $\mathcal{U}^{SSP}(\mathbf{x})$. Since this support depends on the parameter $\mathbf{x}$, we denote this decision-dependent uniform distribution by $\mathbb{P}(\mathbf{x})$. The corresponding stochastic shortest path problem can be expressed as

$$\begin{aligned}\min_{\mathbf{x}}\ &\sum_{e\in\mathcal{A}}c_e x_e+\mathbb{E}_{\mathbb{P}(\mathbf{x})}\left[\sum_{e\in\mathcal{A}}d_e(\boldsymbol{\xi})y_e\right]\\ \text{s.t. }&\mathbf{y}\in Y\\ &\mathbf{x}\in\{0,1\}^{|\mathcal{A}|},\end{aligned}$$

where c_e denotes the cost of reducing the uncertainty in arc e ($x_e=1$), and Y the set of combinatorial constraints (e.g. shortest path). Using the linearity of the expectation term, we obtain

$$\mathbb{E}_{\mathbb{P}(\mathbf{x})}\left[\sum_{e\in\mathcal{A}}d_e(\xi)y_e\right]=\sum_{e\in\mathcal{A}}\mathbb{E}_{\mathbb{P}(\mathbf{x})}[d_e]y_e=\sum_{e\in\mathcal{A}}\bar{d}_e y_e+\frac{\bar{d}_e}{2}\sum_{e\in\mathcal{A}}\mathbb{E}_{\mathbb{P}(\mathbf{x})}[\xi_e]y_e,$$

using the uncertain d_e. However, since ξ_e has a uniform distribution over $[0,1-\gamma_e x_e]$, the expectation can be expressed as $\frac{1}{2}(1-\gamma_e x_e)$. Therefore, we can rewrite the expectation as

$$\mathbb{E}_{\mathbb{P}(\mathbf{x})}\left[\sum_{e\in\mathcal{A}}d_e(\xi)y_e\right]=\sum_{e\in\mathcal{A}}\bar{d}_e y_e+\frac{\bar{d}_e}{2}\sum_{e\in\mathcal{A}}\frac{y_e-\gamma_e x_e y_e}{2}=\frac{5}{4}\sum_{e\in\mathcal{A}}\bar{d}_e y_e-\frac{\bar{d}_e}{4}\sum_{e\in\mathcal{A}}\gamma_e x_e y_e.$$

Consequently, the stochastic shortest path problem can be given as

$$\begin{aligned}\min_{x}\ &\sum_{e\in\mathcal{A}}c_e x_e+\frac{5}{4}\sum_{e\in\mathcal{A}}\bar{d}_e y_e-\sum_{e\in\mathcal{A}}\gamma_e\frac{\bar{d}_e x_e y_e}{4}\\ \text{s.t. }&\mathbf{y}\in Y,\ \mathbf{x}\in\{0,1\}^{|\mathcal{A}|}.\end{aligned}$$

Since this is a bilinear problem, its linearized form is

$$\begin{aligned}
\min_{\mathbf{x}} & \sum_{e\in\mathcal{A}} c_e x_e + \frac{5}{4}\sum_{e\in\mathcal{A}} \bar{d}_e y_e - \sum_{e\in\mathcal{A}} \gamma_e \frac{\bar{d}_e w_e}{4} \\
\text{s.t.} & \left.\begin{aligned} & w_e \geq x_e + y_e - 1 \\ & w_e \leq x_e,\ w_e \leq y_e \end{aligned}\right\} \quad \forall e\in\mathcal{A} \\
& \mathbf{x}\in\{0,1\}^{|\mathcal{A}|},\ \mathbf{y}\in Y.
\end{aligned}$$