

AFFINE DECISION RULES FOR ROBUST OPTIMIZATION WITH DECISION-DEPENDENT INFORMATION DISCOVERY

UGUR KARABAS, HENRI LEFEBVRE, MICHAEL POSS

ABSTRACT. We study robust optimization problems with decision-dependent information discovery under polyhedral uncertainty and continuous recourse. We first establish structural properties of the resulting min–max–min problem, including existence of solutions and a condition under which information discovery has no value. Then, we develop a heuristic approach that combines affine decision rules with scenario generation and show that its adversarial separation problem is NP-hard. Additionally, we introduce a sampling procedure to estimate solution quality without solving the overall problem to global optimality. Numerical experiments on knapsack and facility-location instances show that the method outperforms a K -adaptability approach recently proposed in the literature in terms of computation time while achieving comparable solution quality. Our approach is applicable to polyhedral uncertainty sets that can be expressed as $\Xi = \{\sigma^+ - \sigma^- : (\sigma^+, \sigma^-) \in \Theta\}$ with a downward monotone polytope Θ .

1. INTRODUCTION

Robust two-stage optimization has settled as a well-established paradigm to handle optimization problems facing uncertainty on their parameters. In essence, the paradigm considers optimization problems where some variables (first-stage) are to be decided before knowing the realizations of the uncertain parameters, while the remaining variables (second-stage) can be decided upon full knowledge of these parameters. The extension we study herein considers that some of the uncertain parameters can be queried in the first stage, and that only these queried parameters are known to the decision maker at the second stage. Hence, the resulting problem still contains three levels: (i) the set of queried parameters is decided together with the first-stage decisions, (ii) the adversary reveals the exact values for the queried parameters, (iii) the decision maker commits to her remaining decisions, with the constraint that they must remain feasible for all possible uncertain parameters that have not been queried. The resulting problem, called decision-dependent information discovery (DDID), is relevant in many applications, among which urban planning, project management, resource allocations, scheduling, among many others; see, e.g., Vayanos et al. (2026) and the references therein.

In this paper we consider more specifically robust optimization problems with DDID that can be cast as

$$\inf_{x \in \mathcal{X}, w \in \mathcal{W}} \sup_{\mu \in \Xi_r} \inf_{y \in \mathcal{Y}(w, x, \mu)} c^\top y, \quad (\text{DDID})$$

Date: October 7, 2026.

2020 Mathematics Subject Classification. 90-XX, 90C17, 90C47, 90C59, 90C60 .

Key words and phrases. robust optimization, decision-dependent information discovery, adjustable robust optimization, affine decision rules, scenario generation, K -adaptability.

in which $\mathcal{X} \subseteq \mathbb{R}^{n_x} \times \mathbb{Z}^{n_z}$ denotes the feasibility set for the first-stage decisions, together with the set of possible information discovery $\mathcal{W} \subseteq \{0, 1\}^{n_\xi}$,

$$\Xi_r := \left\{ \mu \in \mathbb{R}_{\geq 0}^{n_\xi} : A\mu \leq r \right\}$$

is the downward monotone polyhedral uncertainty set for some matrix $A \geq 0$ (see, e.g., Poss 2018),

$$\Xi_r(w, \mu) := \{ \xi \in \Xi_r : w \circ \xi = w \circ \mu \}$$

is the remaining uncertainty assuming that parameters corresponding to “ $w_i = 1$ ” have been queried and revealed the values corresponding to vector μ , and

$$\mathcal{Y}(w, x, \mu) := \{ y \in \mathbb{R}^{n_y} : T(\xi)x + W(\xi)y \leq h \text{ for each } \xi \in \Xi_r(w, \mu) \}$$

denotes the feasibility set for the adjustable variables, which are all assumed to be continuous. Observe that costs on variables x could be incorporated into the constraints of $\mathcal{Y}(w, x, \mu)$. Also, we assume, for the sake of simpler notations, that functions $T(\cdot)$ and $W(\cdot)$ are linear in ξ ; one readily extends our results to affine functions.

We observe that set $\mathcal{Y}(w, x, \mu)$ involves robust constraints in its definition, since some parameters are still uncertain when y is to be decided. In fact, choosing $w_i = 1$ for each $i \in \{1, \dots, n_\xi\}$ leads to the two-stage robust problem

$$\inf_{x \in \mathcal{X}} \sup_{\mu \in \Xi_r} \inf_{y \in \hat{\mathcal{Y}}(x, \mu)} c^\top y, \quad (\text{TS})$$

where the robust constraints are no longer present in the feasibility set of the second-stage variables

$$\hat{\mathcal{Y}}(x, \mu) := \{ y \in \mathbb{R}^{n_y} : T(\mu)x + W(\mu)y \leq h \}.$$

Similarly, taking $w_i = 0$ for each $i \in \{1, \dots, n_\xi\}$, problem (DDID) falls down to the static robust problem with constraint uncertainty

$$\inf_{x \in \mathcal{X}} \inf_{y \in \tilde{\mathcal{Y}}(x)} c^\top y, \quad (\text{static})$$

where y are now first-stage variables that belong to the set

$$\tilde{\mathcal{Y}}(x) := \{ y \in \mathbb{R}^{n_y} : T(\xi)x + W(\xi)y \leq h \text{ for each } \xi \in \Xi_r \}$$

featuring robust constraints characterized by Ξ_r . Therefore, the optimal solution cost of (DDID) lies between the optimal solution costs of (static) and (TS).

Problem (DDID) was originally introduced by Vayanos et al. (2026), wherein the authors provide numerous applications and suggest solving approximately the problem through a K -adaptability reformulation, extending the idea initially proposed in the context of two-stage robust optimization by Hanasusanto et al. (2015). Given the difficulty of solving large instances of the problem, subsequent works have focused on the special case where only cost is uncertain. On the one hand, Paradiso et al. (2025) provided a logic-based Benders decomposition algorithm, able to solve to optimality instances of medium size, as illustrated on a DDID variant of the orienteering problem. They also improved the previous K -adaptability reformulation. On the other hand, assuming A does not contain too many rows (such as the budget uncertainty set from Bertsimas and Sim 2004), Omer et al. (2024) proposed a branch-and-price algorithm for the problem, able to solve larger problems than the previous works for such uncertainty polytopes. In addition to these numerically-oriented papers, one could also mention the recent line of research seeking to characterize the hardness of very simple variants of (DDID) under cost uncertainty (Chen et al. 2026) and constraint uncertainty (Chen et al. 2026; Chen and Poss 2026). An extension of DDID to the distributionally robust setting was proposed in Jin et al. (2026).

One important assumption we make in this paper that is restrictive is the requirement that all recourse variables be continuous. If the assumption were relaxed, one would enter the realm of two-stage robust optimization with mixed-integer recourse which, despite having witnessed substantial advances in past years such as those presented, e.g., in Lefebvre and Subramanyam (2026) and Lefebvre et al. (2024), remains an extraordinarily challenging class of problems. In particular, we are not aware of any existing work that treats convex uncertainty sets such as the polytopes considered herein. We therefore focus on the case of continuous recourse, which, despite being Σ_2^P -hard in general (Lefebvre et al. 2026; Poss et al. 2026), is known to be significantly more tractable from a computational perspective. Exact approaches for two-stage robust problems based on scenario generation have been presented in Ayoub and Poss (2016), Flambard et al. (2026), and Zeng and Zhao (2013), while approximate approaches based on linear decision rules have been developed by Ben-Tal et al. (2004).

Contributions. Our contributions are both theoretical and algorithmic. On the theoretical side, we provide in Section 2 three important properties of (DDID). First, we show that (DDID) encompasses the more general setting where the polytope is given by $\Xi'_r := \left\{ \sigma^+ - \sigma^- \in \mathbb{R}^{n_\varepsilon} : \check{A} \begin{pmatrix} \sigma^+ \\ \sigma^- \end{pmatrix} \leq r, \sigma^+, \sigma^- \geq 0 \right\}$, for some $\check{A} \geq 0$. Second, we show that the problem is either infeasible or admits a solution, which is not immediate from the outermost infimum operator present in its definition, and might be somewhat surprising given that the supremum operator is not guaranteed to reach its optimal value in Ξ_r , as we further illustrate by an example. Third, we show that (DDID) is equivalent to the robust problem without information discovery (static) whenever $\mathcal{Y}(w, x, \mu)$ contains a single uncertain constraint. This generalizes an observation already made in Omer et al. (2024) in the setting of DDID with cost uncertainty only, since the unique constraint of $\mathcal{Y}(w, x, \mu)$ could represent the (uncertain) costs through an epigraphic reformulation. We further show numerically that whenever $\mathcal{Y}(w, x, \mu)$ contains two or more constraints, the optimal solution of (DDID) may be strictly better than the one of (static).

On the algorithmic side, we detail in Section 3 a heuristic algorithm for (DDID) combining the use of linear decision rules with scenario generation. We prove, in particular, that the adversarial separation problem is NP-hard. Computing exactly the objective value of a given solution being difficult, we further propose a sampling algorithm that computes a lower bound for that value, converging to it almost surely as the number of samples grows. We illustrate the algorithm numerically in Section 4, and compare it with the approximation provided by the K -adaptability from Vayanos et al. (2026). Our experiments, carried out on small knapsack instances, show how our heuristic scales much better than the K -adaptability heuristic, while providing solutions of the same quality as the best solutions returned by the latter. Our experiments further underline the reasonable scalability of our heuristic on a DDID variant of the robust facility location, illustrating its ability to be applicable to small instances.

Remark 1. *We have assumed throughout that no constraints link x and w only to ensure that the cost of (TS) is not higher than the one of (static), thereby clarifying the picture of how (DDID) stands between these two models. However, our technical developments readily extend to the case where this assumption is relaxed.*

Notation. For any vector $z \in \mathbb{R}^n$, we let $[z]^+$ denote the vector such that $[z]_i^+ = \max\{z_i, 0\}$ for all $i = 1, \dots, n$. Let f be a given function. We let $\text{dom}(f) := \{x : f(x) < \infty\}$ and use the convention $\min\{f(x) : x \in \emptyset\} = \infty$. For a set-valued map S , its domain is $\text{dom}(S) := \{x : \exists y \in S(x)\}$.

2. BASIC PROPERTIES

We begin by establishing basic structural properties of Problem (DDID). First, we show that the downward monotonicity of Ξ_r can be relaxed to a larger class of polytopes. Then, we study the existence of solutions and show with an example that the inner supremum may not be attained. Last, we turn to the value of information discovery and identify a class of problems for which information discovery cannot improve upon the corresponding robust counterpart.

2.1. Relaxing downward monotonicity. We show next that the results derived in this paper apply to the larger class of polytopes that can be defined as

$$\left\{ \sigma^+ - \sigma^- \in \mathbb{R}^{n_\xi} : \check{A} \begin{pmatrix} \sigma^+ \\ \sigma^- \end{pmatrix} \leq r, \sigma^+, \sigma^- \geq 0 \right\}$$

for some $\check{A} \geq 0$. A well-known example of the above polytope is the budget uncertainty polytope that arises when the uncertain parameters appear multiplied by variables that are not necessarily nonnegative. In this case, the downward monotone budget uncertainty polytope $\{\xi \in [0, 1]^{n_\xi} : \|\xi\|_1 \leq \Gamma\}$ is usually replaced by the larger polytope

$$\{\xi \in [-1, 1]^{n_\xi} : \|\xi\|_1 \leq \Gamma\} = \{\sigma^+ - \sigma^- : \|\sigma^+\|_1 + \|\sigma^-\|_1 \leq \Gamma, 0 \leq \sigma^+, \sigma^- \leq 1\},$$

which is not downward monotone anymore. We present this generalization in the next result, by describing the uncertainty polytope Ξ'_r as the image of a downward monotone polytope through a specific linear mapping.

Theorem 1. *Let $\Xi'_r \subseteq \mathbb{R}^{n_\xi}$ and $\Theta := \{(\sigma^+, \sigma^-) \in \mathbb{R}_{\geq 0}^{2n_\xi} : \check{A}\sigma \leq r\}$ be polytopes such that $\check{A} \geq 0$, and $f(\Theta) = \Xi'_r$ for the linear mapping f defined by $f(\sigma) := \sigma^+ - \sigma^-$. Then, $\Xi'_r = \left\{ \sigma^+ - \sigma^- \in \mathbb{R}^{n_\xi} : \check{A} \begin{pmatrix} \sigma^+ \\ \sigma^- \end{pmatrix} \leq r, \sigma^+, \sigma^- \geq 0 \right\}$, and the optimal solution cost of the variant of (DDID) considering Ξ'_r instead of Ξ_r is equal to the one of*

$$\inf_{x \in \mathcal{X}, v \in \check{\mathcal{W}}} \sup_{\sigma \in \Theta} \inf_{y \in \check{\mathcal{Y}}(v, x, \sigma)} c^\top y, \quad (1)$$

where $\check{\mathcal{W}} := \{(v^+, v^-) \in \mathcal{W} \times \mathcal{W} : v^+ = v^-\}$,

$$\check{\mathcal{Y}}(v, x, \sigma) := \{y \in \mathbb{R}^{n_y} : \check{T}(\theta)x + \check{W}(\theta)y \leq h \text{ for each } \theta \in \Theta(v, \sigma)\},$$

$\check{T}(\theta) = T(f(\theta))$, $\check{W}(\theta) = W(f(\theta))$, and $\Theta(v, \sigma) := \{\theta \in \Theta : v \circ \theta = v \circ \sigma\}$.

Proof. Observe first that $\Xi'_r = \left\{ \sigma^+ - \sigma^- \in \mathbb{R}^{n_\xi} : \check{A} \begin{pmatrix} \sigma^+ \\ \sigma^- \end{pmatrix} \leq r, \sigma^+, \sigma^- \geq 0 \right\}$ follows immediately from the definition of $f(\Theta) = \{f(\theta) : \theta \in \Theta\}$.

To prove the claimed equality, let us consider a pair $w \in \mathcal{W}$ and $\begin{pmatrix} w \\ w \end{pmatrix} \in \check{\mathcal{W}}$, together with a pair $\mu \in \Xi'_r$ and $\sigma(\mu) := ([\mu]^+, [-\mu]^+)$. Observe that $\sigma(\mu) \in \arg \min \{\check{A}_k^\top \sigma : \sigma \in f^{-1}(\mu), \sigma \geq 0\}$ for each row $\check{A}_k$ of $\check{A}$, so that $\Xi'_r = f(\Theta)$ implies that $\sigma(\mu) \in \Theta$. We first prove that $f(\Theta(v, \sigma(\mu))) = \Xi'_r(w, \mu)$, which implies

$$\check{\mathcal{Y}}(v, x, \sigma(\mu)) = \mathcal{Y}(w, x, \mu). \quad (2)$$

We start with the direct inclusion. For any $\theta = (\theta^+, \theta^-) \in \Theta(v, \sigma(\mu))$, we define $\xi = \theta^+ - \theta^-$. We have that $\xi \in \Xi'_r$ and $w \circ \xi = w \circ \theta^+ - w \circ \theta^- = w \circ \sigma(\mu)^+ - w \circ \sigma(\mu)^- = w \circ \mu$ so that $\xi \in \Xi'_r(w, \mu)$. To prove the reciprocal inclusion, consider $\xi \in \Xi'_r(w, \mu)$, and define $\theta = ([\xi]^+, [-\xi]^+)$, so $f(\theta) = \xi$. We see that $w \circ \xi = w \circ \mu$ implies $\begin{pmatrix} w \\ w \end{pmatrix} \circ \theta = \begin{pmatrix} w \\ w \end{pmatrix} \circ \sigma(\mu)$. Furthermore, $\theta \in \arg \min \{\check{A}_k^\top \tilde{\theta} : \tilde{\theta} \in f^{-1}(\xi), \tilde{\theta} \geq 0\}$ for each row $\check{A}_k$ of $\check{A}$, so $\theta \in \Theta$, implying also that $\theta \in \Theta(v, \sigma(\mu))$.

We turn then to proving

$$\sup_{\sigma \in \Theta} \inf_{y \in \check{\mathcal{Y}}(v, x, \sigma)} c^\top y = \sup_{\mu \in \Xi'_r} \inf_{y \in \mathcal{Y}(w, x, \mu)} c^\top y.$$

To do so, observe that for each $\sigma \in f^{-1}(\mu) \cap \Theta$, repeating the reasoning below (2) for the direct inclusion to σ yields $f(\Theta(v, \sigma)) \subseteq \Xi'_r(w, \mu) = f(\Theta(v, \sigma(\mu)))$, so that $\check{\mathcal{Y}}(v, x, \sigma(\mu)) \subseteq \check{\mathcal{Y}}(v, x, \sigma)$, which implies that

$$\inf_{y \in \check{\mathcal{Y}}(v, x, \sigma)} c^\top y \leq \inf_{y \in \check{\mathcal{Y}}(v, x, \sigma(\mu))} c^\top y,$$

so that

$$\sup_{\sigma \in f^{-1}(\mu) \cap \Theta} \inf_{y \in \check{\mathcal{Y}}(v, x, \sigma)} c^\top y = \inf_{y \in \check{\mathcal{Y}}(v, x, \sigma(\mu))} c^\top y. \quad (3)$$

Decomposing $\Theta = \bigcup_{\mu \in \Xi'_r} (f^{-1}(\mu) \cap \Theta)$, we obtain

$$\sup_{\sigma \in \Theta} \inf_{y \in \check{\mathcal{Y}}(v, x, \sigma)} c^\top y = \sup_{\mu \in \Xi'_r} \sup_{\sigma \in f^{-1}(\mu) \cap \Theta} \inf_{y \in \check{\mathcal{Y}}(v, x, \sigma)} c^\top y = \sup_{\mu \in \Xi'_r} \inf_{y \in \check{\mathcal{Y}}(w, x, \mu)} c^\top y,$$

where the last equality follows from combining (2) and (3).

The equality finally follows from the trivial bijection between $\mathcal{W}$ and $\check{\mathcal{W}}$ that maps any $w \in \mathcal{W}$ to $\begin{pmatrix} w \\ w \end{pmatrix} \in \check{\mathcal{W}}$, and reciprocally. $\square$

2.2. Existence of solutions. The results described in this subsection and in the subsequent one rely on the following assumption, which is naturally satisfied in most applications.

Assumption 1. *The sets $\mathcal{X}$ and Ξ_r are non-empty and compact. Moreover, for all $(w, x, \mu) \in \mathcal{W} \times \mathcal{X} \times \Xi_r$, $\mathcal{Y}(w, x, \mu) \subseteq \Omega$ holds for some compact set Ω .*

We note that Assumption 1 does not necessarily imply that Problem (DDID) is feasible, i.e., it is possible that $\{(w, x) \in \mathcal{W} \times \mathcal{X} : \forall \mu \in \Xi_r, \exists y \in \mathcal{Y}(w, x, \mu)\} = \emptyset$ holds.

To ease readability, we let ϕ denote the optimal-value function of the inner infimum in Problem (DDID), i.e., we let

$$\phi(w, x, \mu) := \inf_y \{c^\top y : y \in \mathcal{Y}(w, x, \mu)\}.$$

We start with a lemma showing some lower semicontinuity property of ϕ .

Lemma 1. *Let Assumption 1 hold. Then, for any fixed $w \in \mathcal{W}$, the extended-real-valued function $\phi(w, \cdot, \cdot)$ is lower semicontinuous on $\mathcal{X} \times \Xi_r$ and $\text{dom}(\phi(w, \cdot, \cdot))$ is closed.*

Proof. Let $w \in \mathcal{W}$ be fixed. By definition, $\Xi_r(w, \mu) \neq \emptyset$ holds for any $\mu \in \Xi_r$. Moreover, the set-valued map $\Xi_r(w, \cdot)$ is polyhedral, hence continuous, and compact-valued; see Assumption 1 as well as Example 9.35 from Rockafellar and Wets (1998). Now, consider the function

$$\psi(w, x, \mu, y) := \max_{\xi \in \Xi_r(w, \mu)} \| [T(\xi)x + W(\xi)y - h]^+ \|_\infty.$$

Clearly, the maximum is well-posed by compactness of $\Xi_r(w, \mu)$ and continuity of the objective function. In addition, $\mathcal{X} \times \Xi_r \times \Omega \subseteq \text{dom}(\psi(w, \cdot, \cdot, \cdot))$ holds. By Berge's maximum theorem, since the set-valued map $\Xi_r(w, \cdot)$ is continuous, we have that ψ is continuous on $\mathcal{X} \times \Xi_r \times \Omega$.

Now, we may write $\phi(w, x, \mu) = \min_y \{c^\top y : \psi(w, x, \mu, y) \leq 0, y \in \Omega\}$. By compactness of Ω and continuity of $\psi(w, \cdot, \cdot, \cdot)$, we obtain that

$$(x, \mu) \mapsto \{y \in \Omega : \psi(w, x, \mu, y) \leq 0\}$$

is outer semicontinuous. Moreover, since $\mathcal{Y}(w, x, \mu)$ is a closed subset of the compact set Ω , it is compact. Hence, by the Weierstrass theorem, the innermost infimum is attained and we have

$$\text{dom}(\phi(w, \cdot, \cdot)) = \text{dom}(\mathcal{Y}(w, \cdot, \cdot)).$$

Furthermore, since $\text{graph}(\mathcal{Y}(w, \cdot, \cdot))$ is closed and $\mathcal{Y}(w, x, \mu) \subseteq \Omega$ with Ω compact, its projection

$$\text{dom}(\mathcal{Y}(w, \cdot, \cdot)) = \text{proj}_{x, \mu}(\text{graph}(\mathcal{Y}(w, \cdot, \cdot)))$$

is closed in $\mathcal{X} \times \Xi_r$. Hence, $\text{dom}(\phi(w, \cdot, \cdot))$ is closed in $\mathcal{X} \times \Xi_r$.

Finally, by continuity of $y \mapsto c^\top y$ and outer semicontinuity, Lemma 17.30 from Aliprantis and Border (2006) gives that ϕ is lower semicontinuous on its domain. Since its domain is closed in $\mathcal{X} \times \Xi_r$ and $\phi = +\infty$ outside its domain, $\phi(w, \cdot, \cdot)$ is lower semicontinuous on $\mathcal{X} \times \Xi_r$. $\square$

We are now ready to state the theorem on the existence of solutions to Problem (DDID).

Theorem 2. *Let Assumption 1 hold. Then, Problem (DDID) is either infeasible or has a solution with finite value.*

Proof. Since $\mathcal{W}$ is a finite set, Problem (DDID) can be solved (or proved to be infeasible) by enumerating all $w \in \mathcal{W}$. Thus, let $w \in \mathcal{W}$ be fixed. We show that the remaining problem is either infeasible or has a solution.

By Lemma 1, $\phi(w, \cdot, \cdot)$ is lower semicontinuous on $\mathcal{X} \times \Xi_r$. Moreover, it is bounded from above by $\max_y \{c^\top y : y \in \Omega\} < \infty$ on its domain. Since pointwise supremum preserves lower semicontinuity, we obtain that the function $x \mapsto \sup_\mu \{\phi(w, x, \mu) : \mu \in \Xi_r\}$ is again lower semicontinuous. The result is achieved by compactness of $\mathcal{X}$. $\square$

Corollary 1. *Let Assumption 1 hold. Then, Problem (DDID) can be written as*

$$\min_{x \in \mathcal{X}, w \in \mathcal{W}} \sup_{\mu \in \Xi_r} \min_{y \in \mathcal{Y}(w, x, \mu)} c^\top y.$$

In the last corollary, we note that the supremum is not turned into a maximum. Indeed, one would need, e.g., to show that $\phi(w, x, \cdot)$ is upper semicontinuous for this to be the case. Yet, the next example shows that this is not true in general.

Example 1. *Consider Problem (DDID) with $n_x := 0$, $\mathcal{W} := \{1\}$, $\Xi_r := [0, 1]$ and the inner minimization problem given by*

$$\begin{aligned} \min_y \quad & -y - \xi \\ \text{s.t.} \quad & \xi y \leq \xi \quad \text{for each } \xi \in \Xi_r(w, \mu), \\ & y \in [0, 3]. \end{aligned}$$

Because w is fixed to 1, this problem is equivalent to

$$\min_y \{-y - \mu : \mu y \leq \mu, y \in [0, 3]\} =: \phi(1, \mu),$$

where the uncertain objective and right-hand side (formally absent from (DDID)) can be handled through an epigraph variable, and a constant coordinate for y . We now show that $\phi(1, \cdot)$ is not upper semicontinuous. First, assume that $\mu = 0$, then, the constraint “ $\mu y \leq \mu$ ” vanishes and the optimal objective function value equals -3 . Otherwise, assume that $\mu \in (0, 1]$, we have that the unique solution is $y = 1$ with value $-1 - \mu$. We obtain

$$\limsup_{\mu \rightarrow 0^+} \phi(1, \mu) = -1 > \phi(1, 0) = -3.$$

Hence, the function is not upper semicontinuous and we can readily verify that the supremum of -1 is not attained by any point in $[0, 1]$.

2.3. Value of information discovery. In the previous section, we gave sufficient conditions for Problem (DDID) to be well-posed. We now turn our attention to the question of whether there is an advantage in solving Problem (DDID) compared to its static robust version. To that end, we consider the so-called value of information discovery, i.e., the relative difference between $z(\text{DDID})$ and $z(\text{static})$, which denote the optimal solution costs of (DDID) and (static), respectively. More formally, the value of information discovery can be expressed as

$$\text{VoD} := 100 \times \frac{z(\text{static}) - z(\text{DDID})}{z(\text{static})}.$$

Hence, VoD represents the relative amount of resources the decision maker is willing to spend to set-up the information discovery represented by set $\mathcal{W}$.

We show next that $\text{VoD} = 0$ whenever $\mathcal{Y}(w, x, \mu)$ is defined by a single uncertain constraint. To show this result, we reformulate $\mathcal{Y}(w, x, \mu)$ by (i) removing the decision-dependence from the robust constraints and (ii) dualizing the resulting constraints. The next lemma achieves step (i); it is also leveraged in the next section to provide a solution algorithm for solving (DDID) approximately through linear decision rules.

Lemma 2. *Let $w \in \mathcal{W}$ and $x \in \mathcal{X}$ be given. Then, we have*

$$\sup_{\mu \in \Xi_r} \inf_{y \in \mathcal{Y}(w, x, \mu)} c^\top y = \sup_{\mu \in \Xi_r} \inf_{y \in \mathcal{Y}^\circ(w, x, \mu)} c^\top y,$$

where $\mathcal{Y}^\circ(w, x, \mu)$ denotes the set of all $y \in \mathbb{R}^{n_y}$ satisfying

$$T(\zeta(w, \mu, \xi))x + W(\zeta(w, \mu, \xi))y \leq h, \quad \text{for each } \xi \in \Xi_{r(\mu)}, \quad (4)$$

where $\zeta(w, \mu, \xi) := w \circ \mu + (1 - w) \circ \xi$ and $r(\mu) := r - A\mu$.

Proof. Let $(x, w) \in \mathcal{X} \times \mathcal{W}$ be given and define $\Xi_r^\circ(w) := \{\mu \in \Xi_r : (1 - w) \circ \mu = 0\}$. We first show that

$$\sup_{\mu \in \Xi_r} \inf_{y \in \mathcal{Y}(w, x, \mu)} c^\top y = \sup_{\mu \in \Xi_r^\circ(w)} \inf_{y \in \mathcal{Y}(w, x, \mu)} c^\top y$$

Consider any $\mu \in \Xi_r$ and let $\bar{\mu} := w \circ \mu$. Since $A \geq 0$ and $0 \leq \bar{\mu} \leq \mu$ holds, we have that $\bar{\mu} \in \Xi_r(w)$. Moreover, $w \circ \bar{\mu} = w \circ \mu$ implies that $\Xi_r(w, \bar{\mu}) = \Xi_r(w, \mu)$ and, therefore, $\mathcal{Y}(w, x, \bar{\mu}) = \mathcal{Y}(w, x, \mu)$ holds. Hence, for any value attained by a vector $\mu \in \Xi_r$, there is a vector $\bar{\mu} \in \Xi_r^\circ(w)$ that attains the same value, which proves the equality.

Second, we show that

$$\sup_{\mu \in \Xi_r} \inf_{y \in \mathcal{Y}^\circ(w, x, \mu)} c^\top y = \sup_{\mu \in \Xi_r^\circ(w)} \inf_{y \in \mathcal{Y}^\circ(w, x, \mu)} c^\top y$$

also holds. Let again $\mu \in \Xi_r$ and $\bar{\mu} := w \circ \mu$. Using $A \geq 0$, we easily show that that $r(\bar{\mu}) \geq r(\mu)$ holds, and thus $\Xi_{r(\mu)} \subseteq \Xi_{r(\bar{\mu})}$. Moreover, note that $\zeta(w, \bar{\mu}, \xi) = \zeta(w, \mu, \xi)$ also holds, implying that $\mathcal{Y}^\circ(w, x, \bar{\mu}) \subseteq \mathcal{Y}^\circ(w, x, \mu)$ and we get

$$\begin{aligned} \inf_{y \in \mathcal{Y}^\circ(w, x, \bar{\mu})} c^\top y &\geq \inf_{y \in \mathcal{Y}^\circ(w, x, \mu)} c^\top y \\ \implies \sup_{\bar{\mu} \in \Xi_r^\circ(w)} \inf_{y \in \mathcal{Y}^\circ(w, x, \bar{\mu})} c^\top y &\geq \sup_{\mu \in \Xi_r^\circ(w)} \inf_{y \in \mathcal{Y}^\circ(w, x, \mu)} c^\top y. \end{aligned}$$

The reverse inequality follows from $\Xi_r^\circ(w) \subseteq \Xi_r$, which shows the claimed equality.

It remains to show that

$$\sup_{\mu \in \Xi_r^\circ(w)} \inf_{y \in \mathcal{Y}(w, x, \mu)} c^\top y = \sup_{\mu \in \Xi_r^\circ(w)} \inf_{y \in \mathcal{Y}^\circ(w, x, \mu)} c^\top y$$

is satisfied. To this end, it is enough to show that, for any $\mu \in \Xi_r^\circ(w)$, we have

$$\Xi_r(w, \mu) = \{\zeta(w, \mu, \xi) : \xi \in \Xi_{r(\mu)}\}.$$

Thus, let $\mu \in \Xi_r^\circ(w)$ be given and let us show the “ $\supseteq$ ” direction first. Thus, take any $\xi \in \Xi_{r(\mu)}$. Since $A \geq 0$ and $\xi, \mu \geq 0$, we have

$$A\zeta(w, \mu, \xi) = A(w \circ \mu) + A((1-w) \circ \xi) \leq A\mu + A\xi \leq r.$$

Moreover, we have $w \circ \zeta(w, \mu, \xi) = w \circ \mu$. Hence, $\zeta(w, \mu, \xi) \in \Xi_r(w, \mu)$. Next, we show the “ $\subseteq$ ” direction. Let $\bar{\xi} \in \Xi_r(w, \mu)$, and define $\xi := (1-w) \circ \bar{\xi}$, so $\bar{\xi} = \zeta(w, \mu, \xi)$. Since $\mu \in \Xi_r^\circ(w)$, we have $\bar{\xi} = \mu + \xi$, and hence

$$A\xi = A\bar{\xi} - A\mu \leq r - A\mu = r(\mu),$$

so that $\xi \in \Xi_{r(\mu)}$. $\square$

The next lemma reformulates Constraint (4) by dualization, effectively achieving (ii).

Lemma 3. *The i -th robust constraint in (4), i.e.,*

$$T_i(\zeta(w, \mu, \xi))x + W_i(\zeta(w, \mu, \xi))y \leq h_i, \quad \text{for each } \xi \in \Xi_{r(\mu)},$$

holds if and only if there exists $\lambda_i \geq 0$ such that

$$\begin{aligned} (A^\top \lambda_i)_k &\geq (1-w_k)(T_i(e_k)x + W_i(e_k)y), \quad k \in \{1, \dots, n_\xi\}, \\ T_i(w \circ \mu)x + W_i(w \circ \mu)y + r(\mu)^\top \lambda_i &\leq h_i, \end{aligned} \quad (5)$$

where e_k denotes the k -th unit vector.

Proof. Write $c_{ik} := (1-w_k)(T_i(e_k)x + W_i(e_k)y)$ for the coefficient of ξ_k . By linearity of T_i and W_i , and writing $\xi = \sum_k \xi_k e_k$, the constraint is equivalent to

$$\max_{\xi \geq 0} \left\{ \sum_k c_{ik} \xi_k : A\xi \leq r(\mu) \right\} \leq h_i - T_i(w \circ \mu)x - W_i(w \circ \mu)y.$$

The left-hand side is a linear program in ξ . Associating multipliers $\lambda_i \geq 0$ with the constraints $A\xi \leq r$, LP duality gives

$$\max_{\xi \geq 0} \left\{ \sum_k c_{ik} \xi_k : A\xi \leq r(\mu) \right\} = \min_{\lambda_i \geq 0} \left\{ r(\mu)^\top \lambda_i : (A^\top \lambda_i)_k \geq c_{ik} \text{ for each } k \right\}.$$

A minimum is at most a given value if and only if some feasible point attains that value; dropping the minimum therefore yields (5). $\square$

The next theorem follows from a well-known result in the two-stage robust optimization literature due to Marandi and Hertog (2018), and shows that there is no value of information discovery if only one constraint is subject to uncertainty.

Theorem 3. *Let Assumption 1 hold and assume that $\mathcal{Y}(w, x, \mu)$ is defined by a single uncertain constraint. Then, $\text{VoD} = 0$ holds.*

Proof. Consider that $\mathcal{Y}(w, x, \mu)$ is defined by a single uncertain constraint and observe that the uncertainty appears only in one of the constraints of (5). Hence, the uncertainty appears in only one of the constraints of the resulting problem and it is evidently constraint-wise in the sense defined by Marandi and Hertog (2018). By Assumption 1, the sets Ξ_r , $\mathcal{X}$, $\mathcal{W}$ and Ω are compact, and after linearization of the products involving w and (x, y) , the constraints are defined by functions that are bilinear in (w, x, y, λ) and μ . Furthermore, by the boundedness of x and w and Cramer’s rule for basic solutions, we can also define $\Lambda > 0$ such that the constraints $\|\lambda\|_\infty \leq \Lambda$ can be added to (5) without affecting the solution; see, e.g., Buchheim (2023). Hence, all conditions of Marandi and Hertog (2018, Theorem 1) are met (constraint-wise, convexity and compactness of the uncertainty set and recourse set, and convexity/concavity of the involved functions) so the theorem immediately implies $z(\text{static}) = z(\text{DDID})$. $\square$

3. NUMERICAL SOLUTION

3.1. Linear decision rules and scenario generation. Problem (DDID) has a two-stage structure in terms of decisions variables, with x and w being the first-stage decisions. As such, it can be addressed approximately through linear decision rules, assuming that y is a linear function of μ , which we denote by $Y(w \circ \mu)$. Plugging the decision rules in (DDID) and considering the reformulation in (4) leads to

$$\min_{\substack{x \in \mathcal{X}, Y \\ w \in \mathcal{W}, \theta}} \left\{ \theta \left| \begin{array}{l} \theta \geq c^\top Y(w \circ \mu), \quad \text{for each } \mu \in \Xi_r, \\ T(\zeta(w, \mu, \xi))x + W(\zeta(w, \mu, \xi))Y(w \circ \mu) \leq h, \\ \text{for each } \mu \in \Xi_r, \xi \in \Xi_r(\mu) \end{array} \right. \right\}. \quad (\text{DDID-LDR})$$

The resulting problem (DDID-LDR) is a robust linear optimization problem with quadratic dependency on the uncertain parameters. It can be solved “efficiently” as long as the separation problems can be solved “efficiently”. Specifically, for the objective uncertainty, we solve the following separation problem:

$$\max_{\mu \in \Xi_r} c^\top Y(w \circ \mu).$$

Furthermore, since T and W are matrices, there is an additional separation problem for each row i of the constraints

$$\max_{\mu \in \Xi_r, \xi \in \Xi_r(\mu)} T_i(\zeta(w, \mu, \xi))x + W_i(\zeta(w, \mu, \xi))Y(w \circ \mu). \quad (6)$$

Within a standard cutting-plane algorithm, if the optimal value of the objective separation problem strictly exceeds θ , a new constraint is added to the master problem to update the objective bound. Similarly, if the optimal value of the i -th row separation problem strictly exceeds h_i , a new constraint is generated to enforce that row. Note that while the objective separation is a linear problem, the row separation problems are quadratic problems subject to linear constraints. Despite their non-convexity, solvers like, e.g., Gurobi, are able to solve them to global optimality in a short amount of time given the size of the problems at hand. Finally, the products between w and (x, Y, μ) can be linearized, thanks to w being binary.

3.2. NP-hardness of the row separation problem. In this section, we show that the row separation problem (6) is NP-hard by a reduction from the maximization of a quadratic form over a box. To that end, we formally introduce the latter problem.

Definition 1 (Box-QP). Given a symmetric matrix $Q \in \mathbb{Q}^{n \times n}$, solve

$$\max_z \{z^\top Q z : z \in [-1, 1]^n\}.$$

The next result is well known from the literature.

Proposition 1 (Sahni 1974). *Box-QP is NP-hard.*

We now give the complexity result.

Theorem 4. *The row separation problem (6) is NP-hard.*

Proof. Fix $w = 1$ and $x = 0$. With $w = 1$, $\zeta(w, \mu, \xi) = \mu$, and the coupling constraint in $\Xi_r(w, \mu)$ impose $\xi = \mu$. The term $T_i(\xi)x$ vanishes, so the row separation objective simplifies to $\max_{\mu \in \Xi_r} W_i(\mu)Y(\mu)$. Let n be the dimension of Q in Definition 1, and set the uncertainty dimension to $n_\xi = 2n$, writing $\mu = (\mu^+, \mu^-)$. Take

$$\Xi_r = \{\mu \in \mathbb{R}_+^{2n} : \mu \leq 1\},$$

which fits the general form $\Xi_r = \{\mu \in \mathbb{R}_+^{n_\xi} : A\mu \leq r\}$. Both $W_i(\cdot)$ and $Y(\cdot)$ are linear in μ and free to choose, so we take $Y(\mu) := \mu^+ - \mu^-$ and $W_i(\mu) := (\mu^+ - \mu^-)^\top Q$.

With these choices, and using the mapping $z := \mu^+ - \mu^-$, the row separation problem becomes

$$\max_{\mu \in [0,1]^{2n}} (\mu^+ - \mu^-)^\top Q (\mu^+ - \mu^-) = \max_{z \in [-1,1]^n} z^\top Q z,$$

so its optimal value equals the Box-QP value. By the proposition above, Box-QP is NP-hard; therefore the row separation problem is NP-hard. $\square$

3.3. Assessing the objective value. Let $(w, x) \in \mathcal{W} \times \mathcal{X}$ be a feasible first-stage solution for (DDID), such as, e.g., the optimal solution of (DDID-LDR). We typically do not know the value of (w, x) for the true objective function of the problem, namely

$$\Phi(w, x) := \sup_{\mu \in \Xi_r} \inf_{y \in \mathcal{Y}(w, x, \mu)} c^\top y.$$

Moreover, observe that computing $\Phi(w, x)$ is a difficult problem in general as (i) it is closely related to the coNP-hard adversarial problem in two-stage robust optimization (Chekuri et al. 2007), and (ii) it involves computing the supremum of a function that may not be upper semicontinuous as illustrated in Example 1. For these reasons, we take a further heuristic step herein and compute a naive lower bound of $\Phi(w, x)$ by sampling the uncertainty set. Specifically, we construct a set $\hat{\Xi}_r^k \subset \Xi_r$ by sampling k vectors in Ξ_r with uniform distribution and define

$$\hat{\Phi}^k(w, x) := \max_{\mu \in \hat{\Xi}_r^k} \inf_{y \in \mathcal{Y}(w, x, \mu)} c^\top y.$$

Note that the supremum has been replaced by a maximum because $\hat{\Xi}_r^k$ is finite. Moreover, $\hat{\Phi}^k(w, x)$ can be computed in polynomial time for each $\mu \in \hat{\Xi}_r^k$ by solving the dualized linear program resulting from Lemma 3.

We show next that the lower semicontinuity of $\phi(w, x, \cdot)$ proved in Lemma 1 implies that $\hat{\Phi}^k(w, x)$ converges almost surely to $\Phi(w, x)$ as k goes to infinity.

Proposition 2. *Let Assumption 1 hold and let $(w, x) \in \mathcal{W} \times \mathcal{X}$. Then, $\hat{\Phi}^k(w, x)$ is nondecreasing in k , bounded above by $\Phi(w, x)$, and*

$$\hat{\Phi}^k(w, x) \uparrow \Phi(w, x) \quad \text{almost surely.}$$

Proof. Monotonicity follows from $\hat{\Xi}_r^k \subseteq \hat{\Xi}_r^{k+1}$ and the upper bound from $\hat{\Xi}_r^k \subset \Xi_r$. Let $(\mu^j)_{j \geq 1} \subset \Xi_r$ be such that $\phi(w, x, \mu^j) \rightarrow \Phi(w, x)$ as j goes to infinity, which exists by definition of the supremum, and fix j . Every ball (relatively to the affine hull of Ξ_r) centered at μ^j having positive measure in Ξ_r , with probability 1 the sets $\hat{\Xi}_r^k$ contain a sequence $\hat{\xi}^k \in \hat{\Xi}_r^k$ that converges to μ^j as k goes to infinity. By Lemma 1, $\phi(w, \cdot, \cdot)$ is lower semicontinuous, so $\phi(w, x, \cdot)$ is, and we get $\liminf_{k \rightarrow \infty} \phi(w, x, \hat{\xi}^k) \geq \phi(w, x, \mu^j)$, so that $\liminf_{k \rightarrow \infty} \hat{\Phi}^k(w, x) \geq \phi(w, x, \mu^j)$ because $\hat{\Phi}^k(w, x) \geq \phi(w, x, \hat{\xi}^k)$. This holds with probability 1 for each j , hence, these events being countably many, with probability 1 for all j simultaneously, and letting j go to infinity yields $\liminf_{k \rightarrow \infty} \hat{\Phi}^k(w, x) \geq \Phi(w, x)$. The result follows since $\hat{\Phi}^k(w, x) \leq \Phi(w, x)$ for every k . $\square$

Following the above sampling, we can also define the approximate value of discovery of a given heuristic solution (w, x) as

$$\widehat{\text{VoD}}^k(w, x) := 100 \times \frac{\hat{\Phi}^k(x^{\text{static}}, w^{\text{static}}) - \hat{\Phi}^k(w, x)}{\hat{\Phi}^k(x^{\text{static}}, w^{\text{static}})},$$

where $(x^{\text{static}}, w^{\text{static}})$ is any optimal solution to (static); the superscript k will be removed from $\widehat{\text{VoD}}^k(w, x)$ when clear from the context. It is important to note here that the error of $\widehat{\text{VoD}}(w, x)$ arises from two sources: (i) $\hat{\Phi}^k(w, x) \leq \Phi(w, x)$, and (ii) (w, x) may not be a global solution to (DDID), so in general, $\Phi(w, x) \geq z(\text{DDID})$.

Overall, this means that it is not possible to relate $\widehat{\text{VoD}}^k(w, x)$ to VoD for an arbitrary point (w, x) . However, if (x^*, w^*) is guaranteed to be the global solution, then $\hat{\Phi}^k(x^*, w^*) \leq z(\text{DDID})$, and therefore, $\widehat{\text{VoD}}^k(x^*, w^*) \geq \text{VoD}$.

4. NUMERICAL RESULTS

We now assess the proposed linear decision rule approach on two classes of problems. We first consider a fractional knapsack problem, where we compare Problem (DDID-LDR) with the K -adaptability formulation of Vayanos et al. (2026) and evaluate the resulting discovery decisions using the sampling procedure introduced in Section 3.3. We then turn to a capacitated facility location problem to investigate the effect of the uncertainty budget on both the value of information discovery and the computational performance of Problem (DDID-LDR).

Models are implemented in Julia using JuMP and solved with Gurobi 13.0.2. All experiments were run on an Apple M2 with 8 cores. The code needed to reproduce all experiments reported in this paper is available at <https://github.com/hlefebvr/adr-for-robust-ddid>.

4.1. Fractional Knapsack Problem.

4.1.1. *Problem Definition.* We consider an uncertain variant of the fractional knapsack problem with n items. The deterministic version can be modeled as

$$\max_{y \in [0,1]^n} \left\{ \sum_{i=1}^n p_i y_i : \sum_{i=1}^n s_i y_i \leq C \right\},$$

where p_i and s_i denote the profit and the weight of item i , respectively, for all $i = 1, \dots, n$. The robust counterpart to this problem regards the profits and the weights as uncertain. To that end, let $\bar{p}_i$ and $\hat{p}_i$ denote the nominal profit and its maximum deviation, respectively. We assume that the profit of each item i , $i \in \{1, \dots, n\}$, can be expressed as $p_i(\xi) := \bar{p}_i - \hat{p}_i \xi_i$ where ξ_i is an uncertain parameter. Similarly, we denote by $\bar{s}_i$ and $\hat{s}_i$ the nominal value and maximum deviation of the weight, and write $s_i(\xi) := \bar{s}_i + \hat{s}_i \xi_i$. The uncertain parameters are assumed to belong to a budgeted uncertainty set with parameter Γ , i.e., $\Xi_\Gamma = \{\xi \in [0, 1]^n : \|\xi\|_1 \leq \Gamma\}$, while at most q items may be queried, i.e., $\mathcal{W} := \{w \in \{0, 1\}^n : \|w\|_1 \leq q\}$.

Given an information discovery decision $w \in \mathcal{W}$ with its corresponding parameter vector μ , the remaining robust fractional knapsack problem reads

$$\begin{aligned} \min_{y_0, y} \quad & y_0 \\ \text{s.t.} \quad & - \sum_i p_i(\xi) y_i \leq y_0 \quad \text{for each } \xi \in \Xi_\gamma(w, \mu), \\ & \sum_i s_i(\xi) y_i \leq C \quad \text{for each } \xi \in \Xi_\gamma(w, \mu), \\ & 0 \leq y \leq 1. \end{aligned}$$

Hence, denoting by $\mathcal{Y}(w, \mu)$ the feasible region of the above problem, the overall DDID problem reads

$$\min_{w \in \mathcal{W}} \sup_{\mu \in \Xi_\Gamma} \min_{y \in \mathcal{Y}(w, \mu)} y_0.$$

Since y_0 bounds the negative of the total profit, we report $-y_0$ throughout this section, so that larger reported values correspond to better guaranteed profits.

4.1.2. *Experimental setup.* We compare Problem (DDID-LDR) with the K -adaptability approximation of Problem (DDID) proposed by Vayanos et al. (2026). In the latter, K candidate second-stage policies are selected before the uncertainty is revealed, and the best feasible policy is chosen after the discovered coordinates become known. The corresponding reformulation for the fractional knapsack problem, together with the data mapping to Vayanos et al. (2026, Theorem 4), is given in Appendix A.

4.1.3. *A Toy Example.* We consider a two-item toy instance of the fractional knapsack problem for which the optimal discovery decision can be identified directly from the data. Item 1 has deterministic weight but uncertain profit, whereas item 2 is subject to substantial weight uncertainty. The instance parameters are

$$\bar{p} = (3, 3), \quad \hat{p} = (4, 3), \quad \bar{s} = (2, 2), \quad \hat{s} = (0, 3.5), \quad C = 4, \quad \Gamma = 1, \quad q = 1.$$

Both items have uncertain profit and item 2 additionally has uncertain weight, so the adversary can spend its budget on either coordinate.

Problem (DDID-LDR) is solved by cut generation in 0.68 seconds and terminates after 11 iterations. The resulting solution queries item 2, i.e., $w = (0, 1)$, and attains an objective value of 1.7507. This instance is small enough to be solved by hand: the static robust value is $z(\text{static}) = 9/7 \approx 1.2857$, while under $w = (0, 1)$ the adversary solves $\max_{\mu_2 \in [0, 1]} -7\mu_2^2 + 2.75\mu_2 + 2$, attained at $\mu_2 \approx 0.4688$, so that $z(\text{DDID}) \approx 1.75066$. The affine decision rule therefore matches the optimal value of Problem (DDID) on this instance to four decimals, while remaining a restriction of it. We then solve the K -adaptability formulation of Vayanos et al. (2026) for $K \in \{1, \dots, 5\}$, imposing a time limit of 300 seconds for each run. Table 1 reports, for each value of K , the objective value of the best feasible point found at termination, the corresponding relative optimality gap, and the total computation time.

TABLE 1. Empirical results for the K -adaptability approach on the fractional knapsack toy example.

K	Obj. Val.	Opt. Gap (%)	Time (s)
1	1.284	0.00	0.45
2	1.667	0.00	0.64
3	≥ 1.727	9.60	300
4	≥ 1.743	59.04	300
5	≥ 1.748	71.61	300

Because selecting $K = 1$ reduces the problem to its static robust counterpart, no benefit from information discovery can be realized: the objective value does not depend on w , so every discovery decision is optimal and we report $w = (0, 0)$. For every $K > 1$, however, the computed solution is $w = (0, 1)$; querying item 1 instead gives the same worst-case value, so on this instance either query is optimal. Increasing K improves the quality of the K -adaptability approximation, but the computational burden grows rapidly. No run with $K \geq 3$ is solved to optimality within the 300-second time limit; at $K = 3$ the remaining gap is still 9.60%, and it reaches 59.04% and 71.61% for $K = 4$ and $K = 5$, respectively. Note also that none of the K -adaptability runs reaches the value 1.7507 returned by Problem (DDID-LDR) in 0.68 seconds: the best bound, at $K = 5$, is ≥ 1.748 after 300 seconds.

We finally stress that the objective values produced by Problem (DDID-LDR) and by the K -adaptability formulation are not directly comparable. The two approaches restrict the recourse policy in fundamentally different ways: Problem (DDID-LDR)

restricts recourse decisions to affine functions of the revealed uncertainty, whereas K -adaptability restricts them to a finite set of K candidate policies. Consequently, the values reported by the two formulations correspond to different approximations of the original DDID problem. Hence, differences between their objective values reflect differences between the two approximations and do not, by themselves, indicate that one method has identified a better discovery decision. In the next section, we evaluate the discovery decisions returned by both methods on random instances via sampling, following Section 3.3.

4.1.4. *Random instances.* We now extend the comparison to four randomly generated instances of increasing size, with $n = 2, 4, 6, 8$. For each instance, we draw $\bar{p}_i \sim U[2, 5]$, $\bar{s}_i \sim U[1, 3]$, $\hat{p}_i \sim U[1, 5]$ and $\hat{s}_i \sim U[1, 4]$, and set $C = \sum_i \bar{s}_i$. The uncertainty budget is fixed to $\Gamma = 1$ and the discovery budget is $q = 1$.

We compare the discovery decisions returned by Problem (DDID-LDR) and the K -adaptability approach using the sampling procedure introduced in Section 3.3, with $k = 2000$ scenarios drawn from Ξ_Γ . Table 2 reports, for each instance size and method, from left to right, the objective value returned by the respective method, the sampled ex-post value $\hat{\Phi}$, the resulting approximate value of discovery $\widehat{\text{VoD}}$, and the computation time. Each run was executed with a time limit of 120 seconds.

TABLE 2. Empirical results for Problem (DDID-LDR) and the K -adaptability approach on four randomly generated instances of the fractional knapsack problem. With $K = 1$ a single recourse plan is chosen before anything is revealed, so the K -adaptability objective does not depend on w ; we therefore report the static robust value in those rows.

n	Method	Obj. Val.	$\hat{\Phi}$	$\widehat{\text{VoD}}$ (%)	Time (s)
2	(DDID-LDR)	4.9984	4.9984	4.45	0.08
	$K=1$	4.7808	4.7854	0.00	0.03
	$K=2$	≥ 4.9971	4.9987	4.46	120
	$K=3$	≥ 4.9986	4.9984	4.45	120
4	(DDID-LDR)	7.7087	7.7086	4.82	0.11
	$K=1$	7.3473	7.3543	0.00	0.17
	$K=2$	≥ 7.6900	7.7442	5.30	120
	$K=3$	≥ 7.7299	7.7442	5.30	120
6	(DDID-LDR)	15.6269	15.6272	2.71	2.18
	$K=1$	15.2003	15.2154	0.00	0.86
	$K=2$	≥ 15.7097	15.7184	3.31	120
	$K=3$	≥ 15.7166	15.7184	3.31	120
8	(DDID-LDR)	22.5930	22.5935	0.74	0.82
	$K=1$	22.4149	22.4270	0.00	0.33
	$K=2$	≥ 22.7299	22.7857	1.60	120
	$K=3$	≥ 22.7299	22.7857	1.60	120

Both approaches capture most of the value of information discovery relative to the static solution, and the discovery decisions they return are of comparable quality. For $n = 2$ the two approaches select the same item and attain the same sampled ex-post value. For $n = 4, 6$ and 8 they select different items, and the decisions returned by K -adaptability with $K \geq 2$ attain a slightly higher sampled value than those returned by Problem (DDID-LDR): 7.7442 against 7.7086 for $n = 4$, 15.7184

against 15.6272 for $n = 6$, and 22.7857 against 22.5935 for $n = 8$, i.e., differences of 0.46%, 0.58% and 0.85%, respectively. These gaps are small in absolute terms, but they are systematic and grow mildly with the instance size, so the richer recourse structure of K -adaptability does translate into marginally better discovery decisions here. The comparison must, however, be read together with the computation times reported in the same table.

The main difference between the two approaches is computational. Problem (DDID-LDR) is solved to optimality for all four instances in at most 2.18 seconds, and in less than one second for three of them. By contrast, none of the K -adaptability runs with $K \geq 2$ is solved to optimality: all of them reach the 120-second limit without closing the optimality gap. The marginally better discovery decisions of K -adaptability are therefore obtained at a cost that is two to three orders of magnitude higher, and they come without any optimality guarantee, since the corresponding runs are interrupted by the time limit.

Overall, the experiments indicate a clear trade-off. Increasing K enriches the recourse flexibility of the K -adaptability formulation and can yield slightly better discovery decisions, but exploiting this flexibility becomes increasingly difficult as the corresponding formulations become harder to solve, and none of the runs with $K \geq 2$ terminates within the time limit. Problem (DDID-LDR), in contrast, is solved to optimality in seconds and recovers most of the value of information discovery, losing under one percent of the ex-post value on the considered instances. This supports the use of the affine-decision-rule formulation as a computationally attractive approximation for robust optimization problems with decision-dependent information discovery, in particular when the number of items makes the K -adaptability reformulation impractical.

4.2. Facility Location Problem.

4.2.1. Problem Definition. We consider a capacitated facility location problem with n facilities and m customers, in which customer demands are uncertain and may be partially revealed before the allocation decisions are made. Opening facility i incurs a fixed cost f_i and provides capacity C_i , while t_{ij} denotes the unit transportation cost from facility i to customer j . For each customer j , let $\bar{d}_j$ and $\hat{d}_j$ denote the nominal demand and its maximum deviation, respectively, and define $d_j(\xi) := \bar{d}_j + \hat{d}_j \xi_j$. The uncertain parameters belong to the budgeted uncertainty set $\Xi_\Gamma := \{\xi \in [0, 1]^m : \|\xi\|_1 \leq \Gamma\}$, while at most q customer demands may be queried, i.e., $\mathcal{W} := \{w \in \{0, 1\}^m : \|w\|_1 \leq q\}$.

Given first-stage decisions $x \in \{0, 1\}^n$ and $w \in \mathcal{W}$, together with the revealed parameter vector μ , the remaining robust allocation problem reads

$$\begin{aligned}
& \min_{y_0, y} && y_0 \\
& \text{s.t.} && \sum_{i=1}^n f_i x_i + \sum_{i=1}^n \sum_{j=1}^m t_{ij} d_j(\xi) y_{ij} \leq y_0, \quad \text{for each } \xi \in \Xi_\gamma(w, \mu), \\
& && \sum_{j=1}^m d_j(\xi) y_{ij} \leq C_i x_i, \quad i = 1, \dots, n, \quad \text{for each } \xi \in \Xi_\gamma(w, \mu), \\
& && \sum_{i=1}^n y_{ij} \geq 1, \quad j = 1, \dots, m, \\
& && 0 \leq y \leq 1.
\end{aligned}$$

Here, x_i indicates whether facility i is opened and y_{ij} denotes the fraction of the demand of customer j served by facility i . Hence, denoting by $\mathcal{Y}(w, x, \mu)$ the feasible

region of the above problem, the overall DDID problem reads

$$\min_{x \in \{0,1\}^n, w \in \mathcal{W}} \sup_{\mu \in \Xi_\Gamma} \min_{(y_0, y) \in \mathcal{Y}(w, x, \mu)} y_0.$$

Thus, the first-stage decisions determine both which facilities are opened and which customer demands are queried, while the allocation decisions are adjusted after the selected demands have been revealed.

4.2.2. Instance Generation. We generate random instances using parameter ranges taken from Zeng and Zhao (2013) in the context of two-stage robust optimization. Nominal demands are sampled as $\bar{d}_j \sim U[10, 500]$, while the corresponding deviations are defined as $\hat{d}_j = \alpha_j \bar{d}_j$ with $\alpha_j \sim U[0.1, 0.5]$. Facility opening costs, capacities, and transportation costs are sampled as $f_i \sim U[100, 1000]$, $C_i \sim U[200, 700]$, and $t_{ij} \sim U[1, 1000]$, respectively. Since this procedure may generate instances that are infeasible under the worst-case demand realization, we rescale all capacities by a common factor whenever necessary so that the total available capacity exceeds the total demand at $\xi = 1$ by 5%. We consider different instance sizes $(n, m) \in \{(4, 4), (5, 5), (6, 6)\}$ and generate 10 instances for each size. For each instance, the uncertainty budget Γ ranges over all integer values in $\{1, \dots, m\}$. Note that the case $\Gamma = m$ corresponds to the full box uncertainty set, beyond which increasing Γ has no effect. Finally, since the discovery budget has no analogue in the reference model of Zeng and Zhao (2013), we set $q = \lceil m/2 \rceil$, allowing approximately half of the customer demands to be investigated.

4.2.3. Numerical Results. In this section, we evaluate the solutions returned by Problem (DDID-LDR) using the sampling estimator introduced in Section 3.3, with $k = 500$ scenarios drawn uniformly from Ξ_Γ . For each instance size and uncertainty budget, the same pool is used for every instance, so that the reported values are directly comparable.

Table 3 contains the main outcome of our experiments. For each instance size and uncertainty budget Γ , from left to right, we report the number of instances solved to optimality within the time limit, the approximate value of discovery $\widehat{\text{VoD}}$, the average computation time, the shifted geometric means of the time spent on the master problem and on the separation problems, and the average number of cutting-plane iterations. Each run is executed with a time limit of 180 seconds. The reported averages are computed over the instances solved to optimality.

First, we see that the approximate value of discovery decreases consistently as the uncertainty budget Γ increases. This pattern is visible across all instance sizes. For $(4, 4)$, $\widehat{\text{VoD}}$ falls from 7.21% at $\Gamma = 1$ to 0.32% at $\Gamma = 4$, while for $(5, 5)$ it decreases from 9.12% to 1.11% as Γ increases from 1 to 5: as Γ grows, a larger number of customer demands may deviate simultaneously, so the information gained by observing a subset of them becomes less valuable.

Second, the uncertainty budget also has a substantial impact on computational time. This effect is particularly visible for the $(6, 6)$ instances. Increasing Γ from 2 to 5 raises the average computation time from 5.97 to 128.38 seconds and the average number of cutting-plane iterations from 37.2 to 160.0. At $\Gamma = 5$, only four out of ten instances are solved within the 180-second time limit, while at $\Gamma = 6$ none of the instances is solved within the prescribed time.

The last two columns show where this effort is spent. The separation problems remain cheap throughout: on the $(6, 6)$ instances their shifted geometric mean grows only from 0.68 to 2.24 seconds as Γ goes from 2 to 5, because each subproblem keeps the same small size regardless of Γ . The time spent solving the master problem, by contrast, grows from 3.45 to 119.71 seconds over the same range. Its share of the total time increases steadily with Γ : master solves remain a minority of the

TABLE 3. Empirical results for Problem (DDID-LDR) on the facility location instances. Columns “Master” and “Sep.” report the shifted geometric mean (shift 0.01 s) of the time spent by the solver on the master problem and on the separation problems, respectively; note that these two columns use a different averaging scheme than the “Time” column, so they are not directly additive.

(n, m)	Γ	# Solved	$\widehat{\text{VoD}}$ (%)	Time (s)	Master (s)	Sep. (s)	# Iter.
(4,4), $q = 2$	1	10/10	7.21	0.27	0.03	0.09	12.1
	2	10/10	3.86	0.53	0.15	0.12	18.5
	3	10/10	1.51	1.11	0.60	0.16	26.6
	4	10/10	0.32	1.91	1.17	0.18	38.8
(5,5), $q = 3$	1	10/10	9.12	0.79	0.11	0.22	15.9
	2	10/10	7.60	2.56	1.35	0.48	29.6
	3	10/10	4.20	4.40	2.76	0.44	36.7
	4	10/10	2.03	14.84	11.81	0.78	83.0
	5	10/10	1.11	27.73	24.34	0.97	121.6
(6,6), $q = 3$	1	10/10	6.92	0.90	0.13	0.24	16.1
	2	10/10	5.51	5.97	3.45	0.68	37.2
	3	10/10	3.29	15.47	11.44	0.79	49.3
	4	10/10	1.85	48.94	41.92	1.40	106.0
	5	4/10	1.39	128.38	119.71	2.24	160.0
	6	0/10	–	–	–	–	–

total time at $\Gamma = 1$ for all three instance sizes, and additionally at $\Gamma = 2$ for the (4,4) instances, but become the majority of the total time from $\Gamma = 2$ onward for (5,5) and (6,6), and from $\Gamma = 3$ onward for (4,4). The bottleneck is therefore the master problem: as cuts accumulate, the mixed-integer program solved at each iteration becomes progressively harder, and larger uncertainty budgets require more cuts before termination. We also note that the columns “Master (s)” and “Sep. (s)” report shifted geometric means while “Time (s)” reports an arithmetic mean, so their sum need not equal the total even before accounting for solver overhead; the remaining gap is the cost of building the separation models in JuMP. At $\Gamma = 4$ on the (6,6) instances, for instance, each iteration solves 85 separation subproblems (one objective, six capacity, six demand, and 36 each of upper- and lower-bound separations), so an average run of 106 iterations builds around 9,000 such models.

5. CONCLUSION

In this paper, we studied robust optimization problems with decision-dependent information discovery under polyhedral uncertainty and continuous recourse. Our results give structural properties of this problem class and show that affine decision rules provide a practical approximation despite the difficulty of the resulting separation problems. The computational experiments further indicate that the proposed approach can produce discovery decisions of comparable quality to those obtained with K -adaptability, while requiring substantially less computational effort on the considered instances.

A natural direction for future research is the development of exact solution methods for Problem (DDID). In particular, it would be interesting to investigate whether decomposition or scenario-generation techniques can be designed to solve the original DDID problem without restricting the recourse policy to linear decision rules.

ACKNOWLEDGEMENTS

We are grateful to Yasmine Beck for fruitful discussions on the topic of this paper. The authors acknowledge the use of OpenAI’s ChatGPT and Anthropic’s Claude for assistance with editing and polishing the text, including grammar and stylistic improvements, and helping the correction of mathematical mistakes. ChatGPT and Claude were also used to support basic coding tasks.

REFERENCES

- Aliprantis, C. D. and K. C. Border (2006). *Infinite Dimensional Analysis: A Hitchhiker’s Guide*. 3rd ed. Berlin, Heidelberg: Springer. DOI: [10.1007/3-540-29587-9](https://doi.org/10.1007/3-540-29587-9).
- Ayoub, J. and M. Poss (2016). “Decomposition for adjustable robust linear optimization subject to uncertainty polytope.” In: *Computational Management Science* 13.2, pp. 219–239. DOI: [10.1007/s10287-016-0249-2](https://doi.org/10.1007/s10287-016-0249-2).
- Ben-Tal, A., A. Goryashko, E. Guslitzer, and A. Nemirovski (2004). “Adjustable robust solutions of uncertain linear programs.” In: *Mathematical Programming* 99.2, pp. 351–376. DOI: [10.1007/s10107-003-0454-y](https://doi.org/10.1007/s10107-003-0454-y).
- Bertsimas, D. and M. Sim (2004). “The price of robustness.” In: *Operations Research* 52.1, pp. 35–53. DOI: [10.1287/opre.1030.0065](https://doi.org/10.1287/opre.1030.0065).
- Buchheim, C. (2023). “Bilevel linear optimization belongs to NP and admits polynomial-size KKT-based reformulations.” In: *Operations Research Letters* 51.6, pp. 618–622. DOI: [10.1016/j.orl.2023.10.006](https://doi.org/10.1016/j.orl.2023.10.006).
- Chekuri, C., F. B. Shepherd, G. Oriolo, and M. G. Scutellà (2007). “Hardness of robust network design.” In: *Networks* 50.1, pp. 50–54. DOI: [10.1002/net.20165](https://doi.org/10.1002/net.20165).
- Chen, X., M. Goerigk, and M. Poss (2026). “The robust selection problem with information discovery.” In: *Discrete Applied Mathematics* 382, pp. 277–292. DOI: [10.1016/j.dam.2025.12.012](https://doi.org/10.1016/j.dam.2025.12.012).
- Chen, X. and M. Poss (2026). “Decision-dependent information discovery for robust cardinality-constrained combinatorial optimization problems.” Preprint, HAL hal-05667307. URL: <https://hal.science/hal-05667307/>.
- Flambard, P., A. N. Arslan, and B. Detienne (2026). “A semi-infinite constraint generation algorithm for two-stage robust optimization problems.” In: *INFORMS Journal on Computing*. In press. URL: <https://hal.science/hal-05184998/>.
- Hanasusanto, G. A., D. Kuhn, and W. Wiesemann (2015). “K-adaptability in two-stage robust binary programming.” In: *Operations Research* 63.4, pp. 877–891. DOI: [10.1287/opre.2015.1392](https://doi.org/10.1287/opre.2015.1392).
- Jin, Q., A. Georghiou, P. Vayanos, and G. A. Hanasusanto (2026). “Distributionally robust optimization with decision-dependent information discovery.” In: *Mathematical Programming*. Published online. DOI: [10.1007/s10107-026-02346-0](https://doi.org/10.1007/s10107-026-02346-0).
- Lefebvre, H., E. Malaguti, and M. Monaci (2024). “Adjustable robust optimization with discrete uncertainty.” In: *INFORMS Journal on Computing* 36.1, pp. 78–96. DOI: [10.1287/ijoc.2022.0086](https://doi.org/10.1287/ijoc.2022.0086).
- (2026). “Exact approaches for convex adjustable robust optimization.” In: *Mathematical Programming* 218.1–2, pp. 439–476. DOI: [10.1007/s10107-025-02311-3](https://doi.org/10.1007/s10107-025-02311-3).
- Lefebvre, H. and A. Subramanyam (2026). “Remarks on: a Lagrangian dual method for two-stage robust optimization with binary uncertainties.” In: *Optimization and Engineering* 27.1, pp. 531–553. DOI: [10.1007/s11081-025-10015-y](https://doi.org/10.1007/s11081-025-10015-y).
- Marandi, A. and D. den Hertog (2018). “When are static and adjustable robust optimization problems with constraint-wise uncertainty equivalent?” In: *Mathematical Programming* 170.2, pp. 555–568. DOI: [10.1007/s10107-017-1166-z](https://doi.org/10.1007/s10107-017-1166-z).

- Omer, J., M. Poss, and M. Rougier (2024). “Combinatorial robust optimization with decision-dependent information discovery and polyhedral uncertainty.” In: *Open Journal of Mathematical Optimization* 5, 5, pp. 1–25. DOI: [10.5802/ojmo.33](https://doi.org/10.5802/ojmo.33).
- Paradiso, R., A. Georghiou, S. Dabia, and D. Tönissen (2025). “Exact and approximate schemes for robust optimization problems with decision-dependent information discovery.” In: *INFORMS Journal on Computing* 37.6, pp. 1457–1477. DOI: [10.1287/ijoc.2023.0290](https://doi.org/10.1287/ijoc.2023.0290).
- Poss, M. (2018). “Robust combinatorial optimization with knapsack uncertainty.” In: *Discret. Optim.* 27, pp. 88–102. DOI: [10.1016/J.DISOPT.2017.09.004](https://doi.org/10.1016/J.DISOPT.2017.09.004). URL: <https://doi.org/10.1016/j.disopt.2017.09.004>.
- Poss, M., J. Kurtz, M. Goerigk, and D. Henke (2026). “The complexity landscape of robust (integer) linear programming.” DOI: [10.48550/arXiv.2608.21574](https://arxiv.org/abs/10.48550/arXiv.2608.21574).
- Rockafellar, R. T. and R. J. B. Wets (1998). *Variational Analysis*. Vol. 317. Grundlehren der mathematischen Wissenschaften. Berlin, Heidelberg: Springer. DOI: [10.1007/978-3-642-02431-3](https://doi.org/10.1007/978-3-642-02431-3).
- Sahni, S. (1974). “Computationally related problems.” In: *SIAM Journal on Computing* 3.4, pp. 262–279. DOI: [10.1137/0203021](https://doi.org/10.1137/0203021).
- Vayanos, P., A. Georghiou, and H. Yu (2026). “Robust optimization with decision-dependent information discovery.” In: *Management Science* 72.2, pp. 1509–1528. DOI: [10.1287/mnsc.2021.00160](https://doi.org/10.1287/mnsc.2021.00160).
- Zeng, B. and L. Zhao (2013). “Solving two-stage robust optimization problems using a column-and-constraint generation method.” In: *Operations Research Letters* 41.5, pp. 457–461. DOI: [10.1016/j.orl.2013.05.003](https://doi.org/10.1016/j.orl.2013.05.003).

APPENDIX A. K -ADAPTABILITY REFORMULATION OF THE FRACTIONAL KNAPSACK

We instantiate the K -adaptability model of Vayanos et al. (2026, Theorem 4) on the knapsack problem of Section 4.1. In that model, each policy $k \in [K] := \{1, \dots, K\}$ comes with its own copy ξ^k of the uncertain vector. We add a constant coordinate to that vector, writing $\xi = (\xi_0, \dots, \xi_n)$ with $\xi_0 = 1$, so that $n_\xi = n + 1$, and we let $\tilde{w} := (0, w_1, \dots, w_n)$, since the constant coordinate is never queried. The augmented counterpart of Ξ_Γ is then written as $\{\xi \in \mathbb{R}^{n+1} : A\xi \leq r\}$, where we keep the nonnegativity constraints as rows of A , which therefore has $R := 2n + 3$ rows.

Our instance has no first-stage variable x , and w appears neither in the objective nor in the capacity constraint. What is left of the data is

$$P := \begin{pmatrix} -\bar{p}_1 & \cdots & -\bar{p}_n \\ \hat{p}_1 & & \\ & \ddots & \\ & & \hat{p}_n \end{pmatrix}, \quad H(y^k) := (C \quad -\hat{s}_1 y_1^k \quad \cdots \quad -\hat{s}_n y_n^k),$$

plus the recourse row $\bar{s}^\top$. Indeed, $(\xi^k)^\top P y^k = -\sum_i (\bar{p}_i - \hat{p}_i \xi_i^k) y_i^k$ holds; note that the minus sign turns the profit maximization into a minimization as used by Vayanos et al. (2026). Further, we have that $\bar{s}^\top y^k \leq H(y^k) \xi^k$ models the capacity constraint “ $\sum_i (\bar{s}_i + \hat{s}_i \xi_i^k) y_i^k \leq C$ ”. Here the uncertain weights have been moved into H , as detailed in Remark 4 of Vayanos et al. (2026) which makes H depend on y^k .

Each vector ℓ tells us, for every plan k , whether that plan is feasible ($\ell_k = 0$) or which constraint it violates ($\ell_k \neq 0$). We have a single second-stage constraint, so $\mathcal{L} := \{0, 1\}^K$, with $\mathcal{L}_+ = \{1\}$ and $\partial\mathcal{L} = \mathcal{L} \setminus \{1\}$. For $\ell \in \partial\mathcal{L}$, we let $\Delta_K(\ell) := \{\mu \in \mathbb{R}_+^K : 1^\top \mu = 1, \mu_k = 0 \ \forall k : \ell_k \neq 0\}$. Finally, we bound all dual variables by a

constant $M > 0$. Given a scalar $\epsilon > 0$, the program reads

$$\begin{aligned}
& \min \quad \tau \\
& \text{s.t.} \quad w \in \mathcal{W}, \tau \in \mathbb{R}, y^k \in [0, 1]^n, \quad k \in [K], \\
& \quad \alpha(\ell) \in [0, M]^R, \gamma(\ell) \in [0, M]^K, \quad \ell \in \mathcal{L}, \\
& \quad \alpha^k(\ell) \in [0, M]^R, \eta^k(\ell) \in [-M, M]^{n+1}, \quad k \in [K], \ell \in \mathcal{L}, \\
& \quad \beta^k(\ell) \in [0, M], \quad k \in [K], \ell \in \partial\mathcal{L}, \\
& \quad \mu(\ell) \in \Delta_K(\ell), \quad \ell \in \partial\mathcal{L}, \\
& \quad A^\top \alpha(\ell) = \sum_{k \in [K]} \tilde{w} \circ \eta^k(\ell), \quad \ell \in \mathcal{L}, \\
& \quad A^\top \alpha^k(\ell) - H(y^k)^\top \beta^k(\ell) + \tilde{w} \circ \eta^k(\ell) = \mu_k(\ell) P y^k, \quad \ell \in \partial\mathcal{L}, k : \ell_k = 0, \\
& \quad A^\top \alpha^k(\ell) + H(y^k)^\top \gamma_k(\ell) + \tilde{w} \circ \eta^k(\ell) = \mu_k(\ell) P y^k, \quad \ell \in \partial\mathcal{L}, k : \ell_k \neq 0, \\
& \quad \tau \geq r^\top \left(\alpha(\ell) + \sum_{k \in [K]} \alpha^k(\ell) \right) - \sum_{k : \ell_k = 0} (\bar{s}^\top y^k) \beta^k(\ell) \\
& \quad \quad + \sum_{k : \ell_k \neq 0} (\bar{s}^\top y^k - \epsilon) \gamma_k(\ell), \quad \ell \in \partial\mathcal{L}, \\
& \quad A^\top \alpha^k(1) + H(y^k)^\top \gamma_k(1) + \tilde{w} \circ \eta^k(1) = 0, \quad k \in [K], \\
& \quad r^\top \left(\alpha(1) + \sum_{k \in [K]} \alpha^k(1) \right) + \sum_{k \in [K]} (\bar{s}^\top y^k - \epsilon) \gamma_k(1) \leq -1.
\end{aligned}$$

The products between $\tilde{w}$ and $\eta^k(\ell)$ have a binary factor and are therefore handled exactly by McCormick inequalities. However, the remaining products involve two continuous factors. Since our recourse variables are continuous, the linearization into a mixed-binary linear program discussed by Vayanos et al. (2026) after their Theorem 4 does not apply, and the model is solved directly as a nonconvex quadratically constrained program. We use $M = 10^3$ and $\epsilon = 10^{-4}$ in our experiments. The worst-case profit reported in Section 4.1 is $-\tau^*$.

(Henri Lefebvre, Michael Poss) LIRMM, UNIVERSITY OF MONTPELLIER, CNRS, MONTPELLIER, FRANCE

(Ugur Karabas) DEPARTMENT OF INDUSTRIAL ENGINEERING, BILKENT UNIVERSITY, ANKARA, TURKEY

Email address: `henri.lefebvre@cnrs.fr`

Email address: `michael.poss@lirmm.fr`

Email address: `ugur.karabas@ug.bilkent.edu.tr`