

EXACT METHODS FOR DISCRETE Γ -ROBUST INTERDICTION PROBLEMS WITH AN APPLICATION TO THE BILEVEL KNAPSACK PROBLEM

YASMINE BECK, IVANA LJUBIĆ, AND MARTIN SCHMIDT

ABSTRACT. Developing solution methods for discrete bilevel problems is known to be a challenging task—even if all parameters of the problem are exactly known. Many real-world applications of bilevel optimization, however, involve data uncertainty. We study discrete min-max problems with a follower who faces uncertainties regarding the parameters of the lower-level problem. Adopting a Γ -robust approach, we present an extended formulation and a multi-follower formulation to model this type of problem. For both settings, we provide a generic branch-and-cut framework. Specifically, we investigate interdiction problems with a monotone Γ -robust follower and we derive problem-tailored cuts, which extend existing techniques that have been proposed for the deterministic case. For the Γ -robust knapsack interdiction problem, we computationally evaluate and compare the performance of the proposed algorithms for both modeling approaches.

1. INTRODUCTION

In the last years and decades, bilevel optimization problems have gained increasing attention due to their ability to model hierarchical decision making processes that occur in various applications such as transportation (Ben-Ayed et al. 1992; Migdalas 1995), energy markets (Arroyo 2010; Grimm et al. 2019), or pricing (Dempe and Zemkoho 2012; Labbé et al. 1998). In bilevel problems, the decision maker on the upper level (the leader) makes a decision anticipating the reaction of the lower-level player (the follower). In this paper, we consider discrete linear bilevel problems of the form

$$\min_x \quad c^\top x + d^\top y \quad (1a)$$

$$\text{s.t.} \quad Ax \geq a, \quad (1b)$$

$$x \in X \subseteq \mathbb{Z}^{n_x}, \quad (1c)$$

$$y \in \arg \max_{y'} \{d^\top y' : y' \in Y(x) \subseteq \mathbb{Z}_+^{n_y}\}, \quad (1d)$$

where $Y(x)$ denotes the lower-level feasible set that is parameterized by the leader's variables x . The set X is used to denote integrality constraints. Moreover, we have $c \in \mathbb{R}^{n_x}$, $d \in \mathbb{R}^{n_y}$, $A \in \mathbb{R}^{k \times n_x}$, and $a \in \mathbb{R}^k$. We refer to (1a)–(1c) as the upper-level and to (1d) as the lower-level problem. Note that Problem (1) is a min-max problem. Hence, the follower's response yields the worst-possible outcome for the leader, which is why there is no need to distinguish between the optimistic and the pessimistic approach; see, e.g., Dempe (2002). Let us further point out that we do not consider coupling constraints in the upper level, which is crucial for the validity of the branch-and-cut methods we propose in the following sections. In particular, this type of problem covers the important classes of interdiction (G. Brown et al. 2006; Cormican et al. 1998; DeNegre 2011; Fischetti, Ljubić, et al. 2019; Furini, Ljubić, Malaguti, et al. 2021; Israeli and Wood 2002; Wood 2011) and blocking problems (Bazgan et al. 2013; Furini, Ljubić, Malaguti, et al. 2020; Golden 1978; Pajouh 2020; Pajouh, Boginski, et al. 2014; Pajouh, Walteros, et al. 2015; Zenklusen et al. 2009) that arise in various real-world applications such as in critical infrastructure defense, network

Date: April 21, 2023.

2010 Mathematics Subject Classification. 90C11, 90C27, 90C57, 90C70.

Key words and phrases. Bilevel optimization, Robust optimization, Knapsack interdiction, Mixed-integer programming, Branch-and-Cut.

disruption, or marketing. A recent survey on network interdiction models and algorithms can be found in Smith and Song (2020).

Due to their nested structure, even linear bilevel problems are strongly NP-hard in general; see, e.g., Hansen et al. (1992). Moreover, merely checking feasibility for mixed-integer bilevel problems is an NP-hard problem. Thus, it is a difficult task to develop solution methods—especially for bilevel problems that involve discrete variables. In the seminal work by Moore and Bard (1990), the first branch-and-bound method for solving mixed-integer linear bilevel problems is discussed. The idea is extended by DeNegre and Ralphs (2009) who provide a branch-and-cut approach that is based on techniques of standard integer linear programming. In particular, this work can be considered as a turning point regarding computational mixed-integer bilevel optimization that is followed by many influential works on solution methods for bilevel problems; see, e.g., Fischetti, Ljubić, et al. (2017, 2018), Tahernejad, Ralphs, and DeNegre (2020), and Xu and Wang (2014). For a detailed discussion on further techniques for mixed-integer bilevel optimization, we refer to the recent survey in Kleinert et al. (2021).

Throughout this paper, we say that an upper-level decision x is feasible if $x \in X$ and $Ax \geq a$ are satisfied. Moreover, we will hold on to the following.

Assumption 1. *For every feasible decision x of the leader, the lower-level feasible set $Y(x)$ is non-empty.*

Assumption 2. *The shared constraint set $\{(x, y) : Ax \geq a, x \in X, y \in Y(x)\}$ is non-empty and compact.*

Assumption 3. *All linking variables, i.e., all variables of the leader that appear in the lower-level constraints, are bounded integers.*

Assumptions 1–3 are necessary to ensure that Problem (1) has an optimal solution. For a feasible upper-level decision x , we further define the lower-level optimal-value function

$$\Phi(x) = \max_y \{d^\top y : y \in Y(x)\}$$

to re-write Problem (1) as the single-level problem

$$\min_{x \in X, \eta \in \mathbb{R}} c^\top x + \eta \tag{2a}$$

$$\text{s.t. } Ax \geq a, \tag{2b}$$

$$\eta \geq \Phi(x). \tag{2c}$$

Up to this point, we implicitly made the assumption that the input data of both players is certain. However, in many practical situations, this is not the case and the players are forced to make decisions under uncertainty. In mathematical optimization, there are two main approaches to deal with data uncertainty—stochastic optimization (Birge and Louveaux 2011) and robust optimization (Ben-Tal et al. 2009; Bertsimas, D. Brown, et al. 2010; Soyster 1973). In stochastic optimization, it is assumed that the uncertainties can be described by probability distributions that are known in advance. In this setting, the decision maker hedges against uncertainties in a probabilistic sense, e.g., by optimizing over expected values, by considering chance constraints, or some risk-averse models. In between stochastic and robust optimization there is the further approach of distributional robustness; see, e.g., Goh and Sim (2010). In this paper, however, we focus on a robust approach. In robust optimization, the decision maker is interested in a solution that is feasible for all possible realizations of the uncertain data that are assumed to take values in a given uncertainty set. Thus, one pursues a worst-case-oriented philosophy. However, a major point of criticism regarding this approach is the possible over-conservatism of solutions in the sense that ensuring robustness can be very expensive. Addressing this matter, Bertsimas and Sim (2003, 2004) and Sim (2004) propose a more flexible robust approach—the so-called Γ -robust approach—which allows to control the level of conservatism of the solution. In this setting, it is assumed

that the decision maker hedges against the cases in which only a subset of the uncertain parameters will change as to adversely affect the solution of the problem at hand.

In the context of bilevel optimization, problems involving data uncertainties have been investigated using both stochastic as well as robust approaches. Cormican et al. (1998) and Israeli (1999) address stochastic network interdiction problems with uncertainties regarding the interdiction success and uncertain arc capacities, respectively. A stochastic approach for interdiction problems under uncertainty is also considered in the survey by Smith and Song (2020). Further works that pursue a stochastic approach for more general bilevel problems under uncertainty can be found, e.g., in Burtscheidt and Claus (2020), Burtscheidt, Claus, and Dempe (2020), Dempe, Ivanov, et al. (2017), Ivanov (2018), and Yanikoğlu and Kuhn (2018). To the best of our knowledge, robust approaches to address data uncertainty in bilevel optimization have been much less investigated. In the context of power markets, a Γ -robust approach to deal with uncertain lower-level data is considered in Haghghat (2014). Chuong and Jeyakumar (2017) consider problems with uncertain upper- as well as lower-level constraints and solve the robust counterpart via a sequence of semidefinite programming relaxations. In Zeng et al. (2020), a worst-case oriented approach for bilevel problems with lower-level data uncertainty is addressed. Buchheim and Henke (2020) and Buchheim, Henke, and Hommelsheim (2021) present complexity results for robust bilevel problems with uncertainties regarding the lower-level objective function coefficients. For a brief introduction to robust bilevel optimization, we refer to Beck, Ljubić, et al. (2022) and for a general discussion on bilevel optimization under uncertainty, we refer to the recent survey by Beck, Ljubić, et al. (2023).

Let us further mention that, in bilevel optimization, the sources for uncertainty are much richer compared to classic, i.e., single-level, optimization. In bilevel optimization, uncertainties may not only arise in the problem's data but there may also be uncertainty regarding the (observation of the) decisions of the two players. Despite being still in its infancy, there are a few works that consider robust approaches to deal with decision uncertainty. Besançon et al. (2019) propose a robust approach to hedge against near-optimal lower-level decisions. In this context, complexity results are discussed in Besançon et al. (2021). A similar setting in which the leader anticipates sub-optimal follower's decisions due to lower-level algorithmic uncertainty is considered in Zare et al. (2020). The authors consider the setting in which the leader hedges against the Γ th least damaging choices of the solution algorithm for the lower-level problem, which is to some extent related to the notion of Γ -robustness proposed in Bertsimas and Sim (2003). Lastly, robust optimization techniques are used in Beck and Schmidt (2021) to model follower's response uncertainty due to limited observability regarding the leader's decision.

The contributions of this paper are the following. We study discrete linear bilevel problems involving a follower who faces uncertainties regarding the parameters of the lower-level problem. In this context, we pursue the same idea as in Bertsimas and Sim (2003, 2004) and Sim (2004) so that the follower aims to only hedge against a subset of deviations of uncertain parameters. In contrast to the aforementioned literature, we consider bilevel problems that involve discrete variables on the lower level. Therefore, standard reformulation techniques like replacing the lower-level problem by its Karush–Kuhn–Tucker conditions (see, e.g., Fortuny-Amat and McCarl (1981)) cannot be applied anymore.

With regard to the uncertainties, we distinguish the following two cases. On the one hand, we assume that the lower-level's objective function coefficients are uncertain. Instead of d_i , we consider the uncertain coefficients \bar{d}_i , where $\bar{d}_i \in [d_i - \Delta d_i, d_i]$ for all $i \in [n_y] := \{1, \dots, n_y\}$. We denote d_i as the nominal value of the i th coefficient of the lower-level's objective function and Δd_i as its maximum deviation from the nominal value. For a feasible upper-level decision x , the robust counterpart of the lower-level problem (1d) in which the follower hedges against at most $\Gamma_d \in \{0, \dots, n_y\}$ deviations in the objective function coefficients is

given by

$$\Phi_d(x) = \max_{y \in Y(x)} \left\{ d^\top y - \max_{\{S \subseteq [n_y]: |S| \leq \Gamma_d\}} \sum_{i \in S} \Delta d_i y_i \right\}. \quad (3)$$

The overall bilevel problem with a Γ_d -robust follower facing uncertain objective function coefficients can thus be written as

$$\min_{x \in X, \eta \in \mathbb{R}} c^\top x + \eta \quad \text{s.t.} \quad Ax \geq a, \eta \geq \Phi_d(x). \quad (4)$$

On the other hand, we deal with uncertainties regarding the lower-level constraints. Here, we focus on the specific case of a single packing-type constraint of the form $\bar{w}^\top y + v^\top x \leq C$ with $v \in \mathbb{R}^{n_x}$, $C \in \mathbb{R}$, and $\bar{w} \in \mathbb{R}^{n_y}$, where $\bar{w}_i \in [w_i, w_i + \Delta w_i]$ for all $i \in [n_y]$ with $w, \Delta w \in \mathbb{R}^{n_y}$. For a feasible upper-level decision x , the robust counterpart in which the follower hedges against at most $\Gamma_w \in \{0, \dots, n_y\}$ deviations in the constraint coefficients is given by

$$\Phi_w(x) = \max_{y \in Y(x)} \left\{ d^\top y: w^\top y + v^\top x + \max_{\{S \subseteq [n_y]: |S| \leq \Gamma_w\}} \sum_{i \in S} \Delta w_i y_i \leq C \right\} \quad (5)$$

such that the Γ_w -robust counterpart of Problem (2) can be written as

$$\min_{x \in X, \eta \in \mathbb{R}} c^\top x + \eta \quad \text{s.t.} \quad Ax \geq a, \eta \geq \Phi_w(x). \quad (6)$$

Let us mention that we focus on this specific case to ensure that both of the two approaches we present in this paper to model discrete bilevel problems with a robust follower can deal with uncertain lower-level constraints. Nevertheless, possible extensions are sketched in Section 5.

Without loss of generality, we further impose the following.

Assumption 4. *The deviations are non-negative, i.e., $\Delta d_i, \Delta w_i \geq 0$ for all $i \in [n_y]$.*

To model the two types of situations, we consider an approach using an *extended formulation*. Additionally, we present an approach using a *multi-follower formulation* for the special case in which all lower-level variables are binary, i.e., $Y(x) \subseteq \{0, 1\}^{n_y}$. However, the multi-follower-based approach can be extended naturally to allow for additional non-binary lower-level variables as long as the coefficients of the objective function or the constraints corresponding to the non-binary variables are not subject to uncertainty. We present a generic branch-and-cut framework to solve the value-function reformulation of the robustified bilevel problem. Moreover, we derive problem-tailored cuts for interdiction problems that can be used in the proposed branch-and-cut procedure. These cuts assume that Γ -robust follower sub-problems satisfy a downward monotonicity property, which arises in many packing-type applications. In this context, it is our aim to provide a natural extension of the results that have been proposed in Fischetti, Ljubić, et al. (2019) for the deterministic case. The main results of this paper are stated in Theorems 2–5 and form the core for the implementation of the proposed solution methods.

The remainder of the paper is organized as follows. In Section 2, we provide an extended formulation and a multi-follower formulation to model discrete linear bilevel problems with a Γ -robust lower-level problem. We present a generic branch-and-cut method to solve these problems. In Section 3, we focus on interdiction problems with a follower problem that satisfies a downward monotonicity property. In Section 4, we evaluate the effectiveness of the proposed approaches in a numerical study using the bilevel knapsack interdiction problem, which is a prominent example of an interdiction problem that satisfies the monotonicity property. Finally, we conclude in Section 5.

2. GENERIC BRANCH-AND-CUT FRAMEWORKS

The aim of this section is to present generic branch-and-cut frameworks that can be used to solve the Γ -robustification of Problem (2) as stated around (4) and (6). The methods are

similar to a procedure proposed by Wood (2011), which resembles (generalized) Benders decomposition (Benders 1962; Geoffrion 1972). To initialize the methods, we start by solving the problem in which the integrality constraints on the variables x as well as $\eta \geq \Phi(x)$ are omitted. This means that we first consider the linear problem

$$\begin{aligned} \min_{x, \eta} \quad & c^\top x + \eta \\ \text{s.t.} \quad & (x, \eta) \in \Omega_0 := \{(x, \eta) \in \mathbb{R}^{n_x} \times \mathbb{R} : Ax \geq a, x \in \overset{\circ}{X}, \eta \geq \eta^-\}. \end{aligned} \quad (\text{P}_0)$$

Here, $\overset{\circ}{X}$ is a continuous relaxation of X , i.e., the integer points contained in $\overset{\circ}{X}$ coincide with X . Furthermore, we use an a priori lower bound $\eta^- \in \mathbb{R}$ on $\Phi(x)$ for all feasible leader's decisions x . In the following sections, we elaborate on how to obtain a valid lower bound in our setting. After considering Problem (P₀), we iteratively add valid inequalities or branch to cut off integer-infeasible points and also add valid inequalities to cut off bilevel infeasible points. Let

$$\begin{aligned} \min_{x, \eta} \quad & c^\top x + \eta \\ \text{s.t.} \quad & (x, \eta) \in \Omega_j \subseteq \mathbb{R}^{n_x} \times \mathbb{R} \end{aligned} \quad (\text{P}_j)$$

be the problem of node j of the branch-and-cut search tree. Here, the set Ω_j contains all valid inequalities that have been added previously to cut off integer-infeasible and bilevel-infeasible points as well as all branching decisions. If either Problem (P_j) is infeasible or, if the objective function value corresponding to an optimal solution (x^j, η^j) exceeds the current upper bound U , we can fathom node j . Otherwise, we do the following. First, we check if the upper-level variables x^j satisfy the integrality constraints, i.e., we check if $x^j \in X$ holds. If this is not the case, we separate a fractional solution by either exploiting standard cutting planes from mixed-integer linear optimization as elaborated in, e.g., Cornuéjols (2008), or by branching. Otherwise, we proceed by checking for bilevel feasibility, i.e., we check if $\eta^j \geq \Phi(x^j)$ is satisfied. For this purpose, we solve a reformulation of the robust counterpart of the lower-level problem that is parameterized by the current leader's decision $x^j \in X$. In the following sections, we will elaborate on how to obtain these reformulations. In particular, we present two approaches—an extended formulation and a multi-follower formulation—that are derived from Theorem 1 and 3, respectively, in Bertsimas and Sim (2003). Based on the latter two types of formulations, valid cuts to separate bilevel-infeasible points can be obtained. Nevertheless, the development of such cuts strongly depends on the specific problem considered at the lower level. Hence, the branch-and-cut frameworks presented in the remainder of this section remain fairly general and need to be adapted accordingly to capture the application problem at hand. We will show such adaptations for the bilevel knapsack interdiction problem in the following sections.

2.1. Extended Formulation. One possibility to reformulate the robust counterparts (3) and (5) is to allow for an extended variable space of the follower that involves additional continuous variables.

Lemma 1. *For a feasible upper-level decision x , the robust counterpart of the lower-level problem (3) can be solved as the mixed-integer linear problem*

$$\Phi_d(x) = \max_{y, z, \theta} \sum_{i=1}^{n_y} d_i y_i - \Gamma_d \theta - \sum_{i=1}^{n_y} z_i \quad (7a)$$

$$\text{s.t.} \quad z_i + \theta \geq \Delta d_i y_i, \quad i \in [n_y], \quad (7b)$$

$$(z, \theta) \geq 0, \quad (7c)$$

$$y \in Y(x). \quad (7d)$$

The extended formulation for the case of uncertainties in the lower-level constraint is stated in the following lemma.

Lemma 2. *For a feasible upper-level decision x , the robust counterpart of the lower-level problem (5) can be solved as the mixed-integer linear problem*

$$\Phi_w(x) = \max_{y,z,\theta} \sum_{i=1}^{n_y} d_i y_i \quad (8a)$$

$$\text{s.t.} \quad \sum_{i=1}^{n_y} w_i y_i + \Gamma_w \theta + \sum_{i=1}^{n_y} z_i \leq C - \sum_{i=1}^{n_x} v_i x_i, \quad (8b)$$

$$z_i + \theta \geq \Delta w_i y_i, \quad i \in [n_y], \quad (8c)$$

$$(z, \theta) \geq 0, \quad (8d)$$

$$y \in Y(x). \quad (8e)$$

Both lemmas can be shown in analogy to the proof of Theorem 1 in Bertsimas and Sim (2003). Before we comment on how the previous reformulations can be embedded in a branch-and-cut framework to separate bilevel-infeasible points, let us now provide valid lower bounds that can be used in Problem (P₀) to initialize the method.

Proposition 1. *Let x be a feasible upper-level decision. Then, by Assumption 2, there is a vector of finite upper bounds $u \in \mathbb{R}_+^{n_y}$ for the follower's variables such that $Y(x) \subseteq [0, u_1] \times \dots \times [0, u_{n_y}]$ and such that*

$$\eta^- := \sum_{i=1}^{n_y} \min\{d_i, 0\} u_i - \Gamma_d \max_{i \in [n_y]} \{\Delta d_i u_i\} - \sum_{i=1}^{n_y} \Delta d_i u_i$$

is a valid lower bound for the optimal objective function value of the x -parameterized problem (7).

All the proofs that we omit here can be found in Appendix A. In the following, we provide a lower bound for the variant of the problem with uncertainties regarding the follower's packing-type constraint.

Proposition 2. *Let x be a feasible upper-level decision. Then, by Assumption 2, there is a vector of finite upper bounds $u \in \mathbb{R}_+^{n_y}$ for the follower's variables such that $Y(x) \subseteq [0, u_1] \times \dots \times [0, u_{n_y}]$ and such that*

$$\eta^- := \sum_{i=1}^{n_y} \min\{d_i, 0\} u_i$$

is a valid lower bound for the optimal objective function value of the x -parameterized problem (8).

The method to process node j of the branch-and-cut search tree that exploits an extended formulation (as stated in the last two lemmas) is formally stated in Algorithm 1. To determine $\Phi(x^j)$ for the current leader's decision x^j in Step 9, we need to solve the x^j -parameterized robust lower-level problem. Depending on the considered uncertainty model, this can either be Problem (7) or Problem (8) for which we set $\Phi(x^j) = \Phi_d(x^j)$ or $\Phi(x^j) = \Phi_w(x^j)$ accordingly. In Step 11, we generate a valid cut to exclude the bilevel-infeasible point (x^j, η^j) . To this end, one can use generic cuts like (generalized) no-good cuts; see, e.g., Tahernejad and Ralphs (2020).

2.2. Multi-Follower Approach. An alternative reformulation of the robust counterparts (3) and (5) can be obtained under the following additional assumptions.

Assumption 5. *All lower-level variables y are binary, i.e., $Y(x) \subseteq \{0, 1\}^{n_y}$.*

Assumption 6. *The indices are ordered such that the deviations are given in non-increasing order, i.e., $\Delta d_i \geq \Delta d_{i+1}$ or $\Delta w_i \geq \Delta w_{i+1}$ for all $i \in [n_y]$ with $\Delta d_{n_y+1} = 0$ and $\Delta w_{n_y+1} = 0$.*

Algorithm 1 Processing node j using the extended formulation

-
- 1: Solve Problem (\mathbf{P}_j) .
 - 2: **if** Problem (\mathbf{P}_j) is infeasible **then**
 - 3: Fathom the current node.
 - 4: Let (x^j, η^j) denote the solution of Problem (\mathbf{P}_j) .
 - 5: **if** $c^\top x^j + \eta^j \geq U$ **then**
 - 6: Fathom the current node.
 - 7: **if** $x^j \notin X$ **then**
 - 8: Either generate cuts valid for $\Omega_j \cap (X \times \mathbb{R})$, augment Ω_j , and go to Step 1 or branch.
 - 9: Determine $\Phi(x^j)$ and set $U \leftarrow \min\{U, c^\top x^j + \Phi(x^j)\}$.
 - 10: **if** $\eta^j < \Phi(x^j)$ **then**
 - 11: Generate a valid cut that excludes (x^j, η^j) from Ω_j , augment Ω_j , and go to Step 1.
-

Note that Assumption 6 is w.l.o.g. as long as we do not have uncertainties in both the objective function coefficients and the constraint coefficients on the lower level. Assumptions 5 and 6 are necessary to exploit Theorem 3 in Bertsimas and Sim (2003), which is what we do in the following.

Lemma 3. *Let x be a feasible upper-level decision. Under Assumptions 5 and 6, solving the robust counterpart of the lower-level problem (3) is equivalent to solving $n_y + 1$ problems of the nominal type, i.e.,*

$$\Phi_d(x) = \max_{\ell \in \{1, \dots, n_y + 1\}} \Phi_d^\ell(x) \quad (9)$$

holds, where for all $\ell \in \{1, \dots, n_y + 1\}$, we have

$$\Phi_d^\ell(x) = -\Gamma_d \Delta d_\ell + \max_{y \in Y(x)} \left\{ \tilde{d}(\ell)^\top y \right\} \quad (10)$$

with

$$\tilde{d}(\ell)_i = \begin{cases} d_i - (\Delta d_i - \Delta d_\ell), & 1 \leq i \leq \ell, \\ d_i, & \ell + 1 \leq i \leq n_y. \end{cases}$$

Note that the lower-level optimal-value function (9) is defined as the maximum of $n_y + 1$ value functions. Thus, the robustification of Problem (1) can be interpreted as a single-leader-multi-follower problem with $n_y + 1$ many followers, which is why we refer to (9) as *multi-follower formulation*. Moreover, we refer to (10) as the ℓ th *follower sub-problem* throughout this paper.

In the following, we provide a valid lower bound for the optimal objective function value of (9) that can be used in Problem (\mathbf{P}_0) .

Proposition 3. *Let x be a feasible upper-level decision and let $\ell \in \{1, \dots, n_y + 1\}$ be arbitrary but fixed. Under Assumption 5, a valid lower bound for the optimal objective function value of the x -parameterized problem (9) is given by*

$$\eta^- := -\Gamma_d \Delta d_\ell + \sum_{i=1}^{n_y} \min \left\{ \tilde{d}(\ell)_i, 0 \right\}.$$

The multi-follower reformulation for the case of uncertainties in the lower-level constraint is stated in the following lemma.

Lemma 4. *Let x be a feasible upper-level decision. Under Assumptions 5 and 6, solving the robust counterpart of the lower-level problem (5) is equivalent to solving $n_y + 1$ problems of the nominal type, i.e.,*

$$\Phi_w(x) = \max_{\ell \in \{1, \dots, n_y + 1\}} \Phi_w^\ell(x)$$

holds, where for all $\ell \in \{1, \dots, n_y + 1\}$, we have

$$\Phi_w^\ell(x) = \max_{y \in Y(x)} \left\{ \sum_{i=1}^{n_y} d_i y_i : \Gamma_w \Delta w_\ell + \sum_{i=1}^{n_y} \tilde{w}(\ell)_i y_i + \sum_{i=1}^{n_x} v_i x_i \leq C \right\} \quad (11)$$

with

$$\tilde{w}(\ell)_i = \begin{cases} w_i + (\Delta w_i - \Delta w_\ell), & 1 \leq i \leq \ell, \\ w_i, & \ell + 1 \leq i \leq n_y. \end{cases}$$

Lemmas 3 and 4 can be shown in analogy to the proof of Theorem 3 in Bertsimas and Sim (2003).

Note that, in the case of uncertainties regarding the follower's inequality constraint, we have $\Delta d_\ell = 0$ such that $\tilde{d}(\ell) = d$ holds for all $\ell \in \{1, \dots, n_y + 1\}$. Using the latter, Proposition 3 also provides a valid lower bound for the setting considered in Lemma 4. Let us further note that we consider the same deterministic lower-level objective function in the extended formulation as well as in the multi-follower formulation. Moreover, we would like to point out that the fact that there are only *binary* variables corresponding to uncertain coefficients on the lower level is a crucial point for the validity of the multi-follower formulation.

In what follows, we omit the subscripts d and w that are used to denote the considered uncertainty modeling for notational convenience. Further, we will hold on to an improvement of the previous results that has been established in Álvarez-Miranda, Ljubić, et al. (2013) by reducing the number of nominal problems to be considered to $n_y - \Gamma + 2$, i.e.,

$$\Phi(x) = \max_{\ell \in \{\Gamma, \dots, n_y + 1\}} \Phi^\ell(x). \quad (12)$$

Note that a further reduction of the number of nominal problems to be solved has been published by Lee and Kwon (2014). Since this is, however, not the core focus of this paper, we stay with the version as published by Álvarez-Miranda, Ljubić, et al. (2013).

The method to process node j that exploits the multi-follower formulation for the robust counterpart of the lower-level problem is formally stated in Algorithm 2. In contrast to the approach using the extended formulation, in which a single cut is added at each node of the branch-and-cut search tree in case of bilevel-infeasibility, a cut for each follower sub-problem $\ell \in \{\Gamma, \dots, n_y + 1\}$ that satisfies $\eta^j < \Phi^\ell(x^j)$ is added in Step 12 of Algorithm 2. This means that up to $n_y - \Gamma + 2$ cuts could be added at each node. However, it would also be valid to consider, e.g., adding only the most violated cut for the given leader's decision x^j . We will address this aspect in detail when we discuss various cut separation strategies in Section 4.

Algorithm 2 Processing node j using the multi-follower approach

- 1: Solve Problem (P_j) .
 - 2: **if** Problem (P_j) is infeasible **then**
 - 3: Fathom the current node.
 - 4: Let (x^j, η^j) denote the solution of Problem (P_j) .
 - 5: **if** $c^\top x^j + \eta^j \geq U$ **then**
 - 6: Fathom the current node.
 - 7: **if** $x^j \notin X$ **then**
 - 8: Either generate cuts valid for $\Omega_j \cap (X \times \mathbb{R})$, augment Ω_j , and go to Step 1 or branch.
 - 9: **for all** $\ell \in \{\Gamma, \dots, n_y + 1\}$ **do**
 - 10: Solve the ℓ th lower-level sub-problem to obtain $\Phi^\ell(x^j)$.
 - 11: **if** $\eta^j < \Phi^\ell(x^j)$ **then**
 - 12: Generate a valid cut that excludes (x^j, η^j) from Ω_j , augment Ω_j .
 - 13: Set $\Phi(x^j) \leftarrow \max_{\ell \in \{\Gamma, \dots, n_y + 1\}} \Phi^\ell(x^j)$ and $U \leftarrow \min\{U, c^\top x^j + \Phi(x^j)\}$.
 - 14: **If** at least one cut was added in Step 12, **then** go to Step 1.
-

Theorem 1. *If we embed either Algorithm 1 or 2 into a usual branch-and-bound framework, we obtain a correct method that terminates with an optimal solution (x^*, η^*) after a finite number of nodes and after adding an overall finite number of cuts.*

Proof. We first recall that all linking variables are bounded integers. We then observe that, w.l.o.g., non-linking variables can be moved to the lower level; see also Bolusani et al. (2020). Hence, the finite termination is due to the finiteness of the number of feasible upper-level decisions, the finiteness of branch-and-cut methods to solve the lower-level sub-problems, and from the fact that a leader's decision cannot be selected twice. If (x, η) is a non-optimal leader's decision, i.e., $\eta < \Phi(x)$, adding a globally valid inequality excludes (x, η) from being feasible for all subsequent considerations. If an upper-level decision were ever to occur again, i.e., if there would exist solutions $(x^k, \eta^k) = (x^j, \eta^j)$ of Problem (P_j) with $j < k$, then $\eta^j = \eta^k \geq \Phi(x^j)$ would have to hold and the termination criterion would be satisfied. Thus, an optimal solution cannot be overlooked. In particular, the number of cuts possibly added to the problem formulation is finite and in $O(2^{n_x})$. \square

Note that the sub-problems that are solved in Step 10 of Algorithm 2 are independent. This means that the objective function and the constraints of a sub-problem only include the upper-level decision x and the lower-level variables corresponding to the ℓ th sub-problem with $\ell \in \{\Gamma, \dots, n_y + 1\}$. Thus, the sub-problems can be solved in parallel. Further note that we have not specified how the cuts that are added in Algorithm 1 and 2 are generated. For an overview of various cutting planes that can be used for general classes of mixed-integer linear bilevel problems, we refer to Tahernejad and Ralphs (2020). Nevertheless, stronger formulations can be obtained for certain problems; see, e.g., Fischetti, Ljubić, et al. (2019) and Furini, Ljubić, Segundo, et al. (2021). Thus, it is often essential to exploit specific properties of the application problem at hand to derive valid cuts, which is what we do in the remainder of the paper.

3. INTERDICTION CUTS FOR MONOTONE Γ -ROBUST FOLLOWERS

In the following, we focus on interdiction problems with a follower problem that satisfies a downward monotonicity property as in Fischetti, Ljubić, et al. (2019). We refer to this type of problem as an interdiction problem with a monotone follower. In this setting, both players share a common set of items indexed by $i \in [n]$. The leader has the ability to influence the follower's decision by prohibiting the usage of certain items by the follower. This is established by either setting the leader's variable $x_i = 1$ to interdict item $i \in [n]$ for the follower or $x_i = 0$, otherwise. For the ease of presentation, we assume the following.

Assumption 7. *The number of variables on the upper and the lower level is the same, i.e., $n_x = n_y = n$ holds.*

In particular, this means that all variables of the leader need to be binary, i.e., $x \in X = \{0, 1\}^n$. However, the following results can as well be adapted to account for non-interdicting (and thus possibly non-binary) variables of the leader. The case in which the lower-level problem also includes variables that are not subject to interdiction can be handled by partitioning the follower's variable set into interdicted and non-interdicted variables as it is done in Fischetti, Ljubić, et al. (2019).

Assumption 8. *For a feasible upper-level decision x , the x -parameterized lower-level feasible set is of the form*

$$Y(x) = \{y \in \mathcal{Y}: By \leq b, y_i \leq u_i(1 - x_i), i \in [n]\} \quad (13)$$

with $\mathcal{Y} \subseteq \mathbb{Z}_+^n$, $B \in \mathbb{R}_+^{m \times n}$, $b \in \mathbb{R}_+^m$, and a vector of finite upper bounds $u \in \mathbb{R}_+^n$.

Assumption 8 ensures that the nominal lower-level problem (1d) satisfies a downward monotonicity property, which we formally define in the following. Further note that the

leader's variables x are linked to the lower-level problem only via the interdiction constraints $y_i \leq u_i(1 - x_i)$. Both aspects are essential for the validity of the cuts we propose in the remainder of this section.

Proposition 4 (Monotonicity Property). *Let x be a feasible decision of the leader. Further, let $y \in Y(x)$ and let $y' \in \mathcal{Y}$ be such that $y' \leq y$ holds. Then, y' is a feasible follower's decision for the given leader's decision x , i.e., $y' \in Y(x)$.*

It is noteworthy that, under Assumption 8, all items with non-positive objective function coefficients are not chosen by the follower. Consequently, the leader does not need to spend interdiction resources on these items and we could thus omit all items with non-positive objective function coefficients in the problem formulation. This leads us to, w.l.o.g., making the following assumption.

Assumption 9. *All lower-level objective function coefficients are positive, i.e., $d_i > 0$ for all $i \in [n]$.*

Finally, we will hold on to the following, which is reasonable in the interdiction setting.

Assumption 10. *There are no terms depending on the leader's decision in the upper-level objective function, i.e., $c = 0$.*

In the following sections, we show that the downward monotonicity property remains satisfied when Γ -robust followers are considered, which we exploit to derive valid cuts. In Section 3.1, we focus on the case of uncertainties in the follower's objective function coefficients. We derive two variants of *interdiction cuts* based on the two approaches discussed in Section 2. We devote Section 3.2 to strengthened formulations for the proposed interdiction cuts. Finally, the case of uncertainties in the lower-level constraint is addressed in Section 3.3.

3.1. Problems with Uncertain Objective Function Coefficients. According to the notation considered in the previous sections, we assume that all lower-level objective function coefficients may be subject to uncertainty and that the follower hedges against at most Γ_d deviations in the objective function coefficients. The robust counterpart of the lower-level problem and of the overall bilevel problem is given in (3) and (4), respectively. The corresponding extended formulation and the multi-follower formulation have already been stated in Section 2. For the extended formulation, we only need to replace (7d) with (13), i.e., we consider the problem

$$\Phi(x) = \max_{y, z, \theta} \sum_{i=1}^n d_i y_i - \Gamma_d \theta - \sum_{i=1}^n z_i \quad (14a)$$

$$\text{s.t. } z_i + \theta \geq \Delta d_i y_i, \quad i \in [n], \quad (14b)$$

$$By \leq b, \quad (14c)$$

$$y_i \leq u_i(1 - x_i), \quad i \in [n], \quad (14d)$$

$$(z, \theta) \geq 0, y \in \mathcal{Y}. \quad (14e)$$

For the multi-follower approach, we need to replace the feasible set in (10) with (13). Due to Assumption 5, $u_i = 1$ is a trivially valid upper bound for all $i \in [n]$. For $\ell \in \{\Gamma_d, \dots, n_y + 1\}$, we thus consider the sub-problem

$$\Phi^\ell(x) = -\Gamma_d \Delta d_\ell + \max_y \left\{ \tilde{d}(\ell)^\top y : By \leq b, y_i \leq 1 - x_i, y_i \in \{0, 1\}, i \in [n] \right\}. \quad (15)$$

Proposition 5. *Let x be a feasible decision of the leader. Then, the x -parameterized problem (15) satisfies the downward monotonicity property. Moreover, let (y, z, θ) be feasible for the x -parameterized problem (14), and let $y' \in \mathcal{Y}$ be such that $y' \leq y$ holds. Then, (y', z, θ) is feasible for the x -parameterized problem (14) as well.*

Due to Proposition 5 and for the ease of presentation, we also say that the extended formulation (14) satisfies the downward monotonicity property. In what follows, we exploit the

previous result to introduce penalized formulations for Problems (14) and (15) that are used to derive valid interdiction cuts. To this end, we omit the interdiction constraints $y_i \leq u_i(1 - x_i)$ in the problem formulation and instead add the penalty terms $-d_i y_i x_i$ to the objective function for all $i \in [n]$.

Proposition 6. *Let x be a feasible upper-level decision. Then, Problem (14) and the mixed-integer linear problem*

$$\Phi(x) = \max_{y, z, \theta} \sum_{i=1}^n d_i y_i (1 - x_i) - \Gamma_d \theta - \sum_{i=1}^n z_i \quad (16a)$$

$$s.t. \quad z_i + \theta \geq \Delta d_i y_i, \quad i \in [n], \quad (16b)$$

$$By \leq b, \quad (16c)$$

$$y_i \leq u_i, \quad i \in [n], \quad (16d)$$

$$(z, \theta) \geq 0, y \in \mathcal{Y} \quad (16e)$$

admit the same optimal value.

For the multi-follower-based approach, we obtain the following similar result.

Proposition 7. *Let x be a feasible upper-level decision and let $\ell \in \{\Gamma_d, \dots, n+1\}$ be arbitrary but fixed. Then, Problem (15) and the problem*

$$\Phi^\ell(x) = -\Gamma_d \Delta d_\ell + \max_{y \in Y} \left\{ \sum_{i=1}^n \tilde{d}(\ell)_i y_i (1 - x_i) \right\} \quad (17)$$

with

$$Y = \{y \in \{0, 1\}^n : By \leq b\}$$

admit the same optimal value.

The feasible set of Problem (16) and of each sub-problem (17) is independent from the leader's decision. Moreover, the objective functions are linear for fixed x . Thus, an optimal solution is attained at a vertex of the convex hull of the respective feasible set. We set

$$\Psi = \{(y, z, \theta) \in \mathcal{Y} \times \mathbb{R}^n \times \mathbb{R} : (y, z, \theta) \text{ satisfy (16b)–(16e)}\}.$$

In what follows, we use $\hat{\Psi}$ and \hat{Y} to denote the set containing the finite number of vertices of the convex hull of Ψ and Y , respectively. Further, let (x, η) be feasible for Problem (2) with the lower-level optimal-value function (3). Due to Proposition 6, we have

$$\eta \geq \Phi(x) \geq \sum_{i=1}^n d_i \hat{y}_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i=1}^n \hat{z}_i$$

for arbitrary but fixed $(\hat{y}, \hat{z}, \hat{\theta}) \in \hat{\Psi}$. Consequently, the interdiction cuts

$$\eta \geq \sum_{i=1}^n d_i \hat{y}_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i=1}^n \hat{z}_i \quad \text{for all } (\hat{y}, \hat{z}, \hat{\theta}) \in \hat{\Psi} \quad (18)$$

are valid for Problem (2). To derive valid cuts for the multi-follower-based approach, we do the following. Under Assumptions 5 and 6, we have

$$\begin{aligned} \eta &\geq \Phi(x) = \max_{\ell \in \{\Gamma_d, \dots, n+1\}} \Phi^\ell(x) \\ &\geq \Phi^\ell(x) = -\Gamma_d \Delta d_\ell + \max_{y \in Y} \left\{ \sum_{i=1}^n \tilde{d}(\ell)_i y_i (1 - x_i) \right\} \\ &\geq -\Gamma_d \Delta d_\ell + \sum_{i=1}^n \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) \end{aligned}$$

for arbitrary but fixed $\hat{y} \in \hat{Y}$ and for all $\ell \in \{\Gamma_d, \dots, n+1\}$. The first equality follows from (12) and the second one holds due to Proposition 7. As a result, the cuts

$$\eta \geq -\Gamma_d \Delta d_\ell + \sum_{i=1}^n \tilde{d}(\ell)_i y_i (1 - x_i) \quad \text{for all } \hat{y} \in \hat{Y}, \ell \in \{\Gamma_d, \dots, n+1\} \quad (19)$$

are valid for Problem (2). In particular, we may obtain different valid cuts for each $\ell \in \{\Gamma_d, \dots, n+1\}$. Finally, we exploit the previous results to equivalently reformulate the interdiction problem with a Γ_d -robust follower facing uncertain objective function coefficients.

Theorem 2. *Problem (2) with the lower-level optimal-value function (3) can be equivalently reformulated by replacing Constraint (2c) with (18). Under Assumptions 5 and 6, an equivalent reformulation can be obtained by replacing Constraint (2c) with (19).*

3.2. Cut Strengthening and Enhanced Formulations. In this section, we provide enhancements and techniques to strengthen the cuts proposed in the previous section. First, we introduce the notion of maximal packings.

Definition 1. *A follower's decision $(\hat{y}, \hat{z}, \hat{\theta}) \in \hat{\Psi}$ is a maximal packing w.r.t. the extended formulation (16) if there is no $(y', \hat{z}, \hat{\theta}) \in \hat{\Psi} \setminus \{\hat{y}\}$ such that $\hat{y} \leq y'$ holds.*

We exploit this notion to avoid the generation of unnecessary cuts.

Proposition 8. *Let $(\hat{y}, \hat{z}, \hat{\theta}) \in \hat{\Psi}$ be a non-maximal packing for Problem (16) and let $(y', \hat{z}, \hat{\theta}) \in \hat{\Psi} \setminus \{\hat{y}\}$ be chosen such that $\hat{y} \leq y'$ holds. Then, the interdiction cut (18) associated with $(\hat{y}, \hat{z}, \hat{\theta})$ is dominated by the interdiction cut associated with $(y', \hat{z}, \hat{\theta})$.*

Here and in what follows, *domination* between two cuts is understood in the sense that the feasible set induced by the dominating cut is contained in the feasible set induced by the dominated one. Due to the previous result, it is sufficient to only consider the interdiction cuts that correspond to maximal packings of the follower.

Definition 2. *A follower's decision $\hat{y} \in \hat{Y}$ is a maximal packing w.r.t. the multi-follower formulation (17) if there is no $y' \in \hat{Y} \setminus \{\hat{y}\}$ such that $\hat{y} \leq y'$ holds.*

Note that there is no need to specify $\ell \in \{\Gamma_d, \dots, n+1\}$ in the previous definition since, in each sub-problem, we consider the same set \hat{Y} , which is independent from ℓ . To obtain a dominance result for interdiction cuts associated with maximal packings in the multi-follower setting, we need to further study the properties of the modified objective function coefficients $\tilde{d}(\ell)$ for each $\ell \in \{\Gamma_d, \dots, n+1\}$. Note that the modified objective function coefficients can be non-positive for certain items in some follower sub-problems. If $\tilde{d}(\ell)_i \leq 0$ holds for all $\ell \in \{\Gamma_d, \dots, n+1\}$, the i th item will not be chosen in any of the follower sub-problems. Thus, the leader does not need to spend interdiction resources on the i th item and $x_i = y_i = 0$ can be fixed. This is equivalent to completely omitting the i th item in the problem formulation. However, if there is an item $i \in [n]$ with non-positive modified objective function coefficients only for some follower sub-problems, i.e., $\tilde{d}(k)_i \leq 0$ for all $k \in \mathcal{S} \subset \{\Gamma_d, \dots, n+1\}$ and $\tilde{d}(l)_i > 0$ for all $l \in \{\Gamma_d, \dots, n+1\} \setminus \mathcal{S}$, the i th item might be part of an optimal solution. Therefore, we introduce the following notation. For each $\ell \in \{\Gamma_d, \dots, n+1\}$, we define the set

$$D_+^\ell := \{i \in [n]: \tilde{d}(\ell)_i > 0\}.$$

Proposition 9. *The interdiction cuts (19) can be replaced with*

$$\eta \geq -\Gamma_d \Delta d_\ell + \sum_{i \in D_+^\ell} \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) \quad \text{for all } \hat{y} \in \hat{Y}, \ell \in \{\Gamma_d, \dots, n+1\}. \quad (20)$$

In particular, the cuts (20) dominate the basic interdiction cuts (19).

With the previous considerations, we can finally state a dominance result for interdiction cuts associated with maximal packings in the multi-follower setting.

Proposition 10. *Let $\hat{y} \in \hat{Y}$ be a non-maximal packing for Problem (17) and let $y' \in \hat{Y} \setminus \{\hat{y}\}$ be such that $\hat{y} \leq y'$ holds. Then, the interdiction cuts (20) associated with \hat{y} are dominated by the interdiction cuts associated with y' .*

Note that the previous result is not valid for the basic interdiction cuts as stated in (19) since, in general, $\tilde{d}(\ell)_i \hat{y}_i (1 - x_i) \leq \tilde{d}(\ell)_i y'_i (1 - x_i)$ does not hold for all $i \notin D_+^\ell$. Moreover, we would like to mention that we also considered maximal packings for the leader. However, preliminary computational tests revealed that this does not improve the performance of the overall solution method. Thus, we decided to refrain from using this ingredient in our computational study in Section 4. However, we exploit dominance properties among items to obtain further enhancements. First, we introduce additional inequalities regarding the leader's decision x . In what follows, A_i denotes the i th column of A .

Theorem 3. *For all $s, t \in [n]$, $s \neq t$, chosen such that $A_{.s} > A_{.t}$, $B_{.s} \leq B_{.t}$, $u_s \geq u_t$, $d_s \geq d_t$, and $d_s - \Delta d_s \geq d_t - \Delta d_t$, the dominance inequality*

$$x_t \leq x_s \quad (21)$$

is satisfied in at least one optimal solution of Problem (2) with the lower-level optimal-value function (3).

Proof. Let (x^*, η^*) be an optimal solution of Problem (2) with the lower-level optimal-value function (3). Without loss of generality, suppose that there are exactly two distinct items $s, t \in [n]$ that satisfy the requirements of the theorem but for which the dominance inequality (21) does not hold, i.e., $x_s^* = 0$ and $x_t^* = 1$. Otherwise, we repeat the following procedure as long as there are still items left that satisfy the requirements but violate the corresponding dominance inequality. The idea is to construct an optimal leader's decision that satisfies the dominance inequality. To this end, we set

$$x'_i = \begin{cases} x_i^*, & i \in [n] \setminus \{s, t\}, \\ 1, & i = s, \\ 0, & i = t. \end{cases}$$

By construction, x' is feasible for Problem (2) with the lower-level optimal-value function (3) and satisfies the dominance inequality associated with s and t . Without loss of generality, we show that x' is also an optimal solution of Problem (2) using the extended formulation. Let (y', z', θ') be an optimal solution of Problem (14) for x' , i.e., we have $y'_s = 0$. Moreover, $z'_i = \max\{\Delta d_i y'_i - \theta', 0\}$ holds for all $i \in [n]$ due to Constraints (14b), (14e), and the objective function. If $y'_t = 0$ holds, (y', z', θ') is also a feasible follower's decision for x^* and we obtain

$$\Phi(x') = \sum_{i=1}^n d_i y'_i - \Gamma_d \theta' - \sum_{i=1}^n z'_i \leq \Phi(x^*).$$

If $y'_t \geq 1$ holds, we consider the alternative follower's decision $(\hat{y}, \hat{z}, \theta')$ with

$$\hat{y}_i = \begin{cases} y'_i, & i \in [n] \setminus \{s, t\}, \\ y'_t, & i = s, \\ 0, & i = t, \end{cases}$$

and

$$\hat{z}_i = \begin{cases} z'_i, & i \in [n] \setminus \{s, t\}, \\ \max\{\Delta d_s y'_t - \theta', 0\}, & i = s, \\ 0, & i = t. \end{cases}$$

By construction, $(\hat{y}, \hat{z}, \theta')$ is feasible for Problem (14) given the leader's decision x^* and we obtain

$$\begin{aligned}
\Phi(x') &= \sum_{i=1}^n d_i y'_i - \Gamma_d \theta' - \sum_{i=1}^n z'_i \\
&= d_s \underbrace{y'_s}_{=0} + d_t y'_t + \sum_{i \in [n] \setminus \{s,t\}} d_i \underbrace{y'_i}_{=\hat{y}_i} - \Gamma_d \theta' - \underbrace{z'_s}_{=0} - z'_t - \sum_{i \in [n] \setminus \{s,t\}} \underbrace{z'_i}_{=\hat{z}_i} \\
&= d_t y'_t + \sum_{i \in [n] \setminus \{s,t\}} d_i \hat{y}_i - \Gamma_d \theta' - \max\{\Delta d_t y'_t - \theta', 0\} - \sum_{i \in [n] \setminus \{s,t\}} \hat{z}_i \\
&= \sum_{i \in [n] \setminus \{s,t\}} d_i \hat{y}_i - \Gamma_d \theta' + \min\{(d_t - \Delta d_t) y'_t + \theta', d_t y'_t\} - \sum_{i \in [n] \setminus \{s,t\}} \hat{z}_i \\
&\leq \sum_{i \in [n] \setminus \{s,t\}} d_i \hat{y}_i - \Gamma_d \theta' + \min\{(d_s - \Delta d_s) \hat{y}_s + \theta', d_s \hat{y}_s\} - \sum_{i \in [n] \setminus \{s,t\}} \hat{z}_i \\
&= d_s \hat{y}_s + d_t \underbrace{\hat{y}_t}_{=0} + \sum_{i \in [n] \setminus \{s,t\}} d_i \hat{y}_i - \Gamma_d \theta - \max\{\Delta d_s \hat{y}_s - \theta', 0\} - \underbrace{\hat{z}_t}_{=0} - \sum_{i \in [n] \setminus \{s,t\}} \hat{z}_i \\
&= \sum_{i=1}^n d_i \hat{y}_i - \Gamma_d \theta' - \sum_{i=1}^n \hat{z}_i \\
&\leq \Phi(x^*).
\end{aligned}$$

Hence, the alternative leader's decision x' is optimal for Problem (2) with the lower-level optimal-value function (3) and satisfies the dominance inequality associated with s and t . This concludes the proof. \square

Note that, in the case of only binary follower's variables, the requirement $u_s \geq u_t$ in the previous theorem is trivially satisfied since $u_i = 1$ for all $i \in [n]$ is a valid upper bound. Further note that the requirement $A_{.s} > A_{.t}$ can be relaxed to $A_{.s} \geq A_{.t}$ if the matrix A has only non-negative entries.

In the remainder of this section, we provide lifted cuts that dominate their respective basic counterparts stated in (18) and (20). We start by lifting the basic interdiction cuts corresponding to the extended formulation.

Theorem 4. *For any arbitrary but fixed $(\hat{y}, \hat{z}, \hat{\theta}) \in \hat{\Psi}$, let $K \in [n]$, $S_a = \{a_1, \dots, a_K\} \subset [n]$, and $S_b = \{b_1, \dots, b_K\} \subset [n]$ be such that $S_a \cap S_b = \emptyset$, $\hat{y}_{a_k} \geq 1$, $\hat{y}_{b_k} = 0$, $B_{.a_k} \geq B_{.b_k}$, $u_{a_k} \leq u_{b_k}$, $\Delta d_{a_k} < \Delta d_{b_k}$, and $d_{a_k} - \Delta d_{a_k} < d_{b_k} - \Delta d_{b_k}$ for all $k \in [K]$. Then, the following lifted interdiction cut is valid for Problem (2) with the lower-level optimal-value function (3):*

$$\eta \geq \sum_{i=1}^n d_i \hat{y}_i (1 - x_i) + \sum_{k=1}^K ((d_{b_k} - \Delta d_{b_k}) - (d_{a_k} - \Delta d_{a_k})) \hat{y}_{a_k} (1 - x_{b_k}) - \Gamma_d \hat{\theta} - \sum_{i=1}^n \hat{z}_i. \quad (22)$$

Proof. Let (x, η) be a feasible leader's decision for Problem (2) with the lower-level optimal-value function (3). If $x_{b_k} = 1$ holds for all $k \in [K]$, we obtain the basic interdiction cut as stated in (18), which is satisfied by x . Otherwise, we define $\mathcal{K} := \{k \in [K] : x_{b_k} = 0\}$ and consider an alternative follower's decision $(y', z', \hat{\theta})$, which is obtained as follows. For all $k \in \mathcal{K}$, we set $y'_{a_k} = 0$, $y'_{b_k} = \hat{y}_{a_k}$ as well as $z'_{a_k} = 0$ and $z'_{b_k} = \max\{\Delta d_{b_k} \hat{y}_{a_k} - \hat{\theta}, 0\}$. We further define $\bar{\mathcal{K}} := [n] \setminus \{a_k, b_k : k \in \mathcal{K}\}$ and set $y'_i = \hat{y}_i$ as well as $z'_i = \hat{z}_i$ for all $i \in \bar{\mathcal{K}}$. By construction, we have $(y', z', \hat{\theta}) \in \Psi$. Since the interdiction cuts (18) are not only valid for vertices contained in the set $\hat{\Psi}$ but also for any lower-level feasible point, the leader's

decision x also satisfies the basic interdiction cut associated with $(y', z', \hat{\theta})$, i.e.,

$$\begin{aligned}
\eta &\geq \sum_{i=1}^n d_i y'_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i=1}^n z'_i \\
&= \sum_{i \in \mathcal{K}} d_i \underbrace{y'_i}_{=\hat{y}_i} (1 - x_i) + \sum_{k \in \mathcal{K}} (d_{a_k} \underbrace{y'_{a_k}}_{=0} (1 - x_{a_k}) + d_{b_k} \underbrace{y'_{b_k}}_{=\hat{y}_{a_k}} \underbrace{(1 - x_{b_k})}_{=1}) \\
&\quad - \Gamma_d \hat{\theta} - \sum_{i \in \mathcal{K}} \underbrace{z'_i}_{=\hat{z}_i} - \sum_{k \in \mathcal{K}} (\underbrace{z'_{a_k}}_{=0} + z'_{b_k}) \\
&= \sum_{i \in \mathcal{K}} d_i \hat{y}_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i \in \mathcal{K}} \hat{z}_i + \sum_{k \in \mathcal{K}} (d_{b_k} \hat{y}_{a_k} - \max\{\Delta d_{b_k} \hat{y}_{a_k} - \hat{\theta}, 0\}).
\end{aligned} \tag{23}$$

In particular, we have $\hat{z}_{a_k} = \max\{\Delta d_{a_k} \hat{y}_{a_k} - \hat{\theta}, 0\}$ and $\hat{z}_{b_k} = 0$ for all $k \in [K]$. Hence, we can re-write Inequality (22) as

$$\begin{aligned}
\eta &\geq \sum_{i=1}^n d_i \hat{y}_i (1 - x_i) + \sum_{k=1}^K ((d_{b_k} - \Delta d_{b_k}) - (d_{a_k} - \Delta d_{a_k})) \hat{y}_{a_k} (1 - x_{b_k}) - \Gamma_d \hat{\theta} - \sum_{i=1}^n \hat{z}_i \\
&= \sum_{i \in \mathcal{K}} d_i \hat{y}_i (1 - x_i) + \sum_{k \in \mathcal{K}} (d_{a_k} \hat{y}_{a_k} (1 - x_{a_k}) + d_{b_k} \underbrace{\hat{y}_{b_k}}_{=0} (1 - x_{b_k})) \\
&\quad + \sum_{k \in \mathcal{K}} ((d_{b_k} - \Delta d_{b_k}) - (d_{a_k} - \Delta d_{a_k})) \hat{y}_{a_k} (1 - \underbrace{x_{b_k}}_{=0}) \\
&\quad + \sum_{k \in [K] \setminus \mathcal{K}} ((d_{b_k} - \Delta d_{b_k}) - (d_{a_k} - \Delta d_{a_k})) \hat{y}_{a_k} (1 - \underbrace{x_{b_k}}_{=1}) - \Gamma_d \hat{\theta} \\
&\quad - \sum_{i \in \mathcal{K}} \hat{z}_i - \sum_{k \in \mathcal{K}} (\hat{z}_{a_k} + \underbrace{\hat{z}_{b_k}}_{=0}) \\
&= \sum_{i \in \mathcal{K}} d_i \hat{y}_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i \in \mathcal{K}} \hat{z}_i + \sum_{k \in \mathcal{K}} (d_{a_k} \hat{y}_{a_k} (1 - x_{a_k}) \\
&\quad + ((d_{b_k} - \Delta d_{b_k}) - (d_{a_k} - \Delta d_{a_k})) \hat{y}_{a_k} - \max\{\Delta d_{a_k} \hat{y}_{a_k} - \hat{\theta}, 0\}) \\
&= \sum_{i \in \mathcal{K}} d_i \hat{y}_i (1 - x_i) - \Gamma_d \hat{\theta} + \sum_{k \in \mathcal{K}} (d_{b_k} \hat{y}_{a_k} - \max\{\Delta d_{b_k} \hat{y}_{a_k} - \hat{\theta}, 0\}) \\
&\quad - \sum_{k \in \mathcal{K}} (d_{a_k} \hat{y}_{a_k} x_{a_k} + (\Delta d_{b_k} - \Delta d_{a_k}) \hat{y}_{a_k} \\
&\quad \quad + \max\{\Delta d_{a_k} \hat{y}_{a_k} - \hat{\theta}, 0\} - \max\{\Delta d_{b_k} \hat{y}_{a_k} - \hat{\theta}, 0\})
\end{aligned} \tag{24}$$

Note that the right-hand side of (24) corresponds to the right-hand side of (23) with the additional term

$$\begin{aligned}
&\sum_{k \in \mathcal{K}} (d_{a_k} \hat{y}_{a_k} x_{a_k} + (\Delta d_{b_k} - \Delta d_{a_k}) \hat{y}_{a_k} \\
&\quad + \max\{\Delta d_{a_k} \hat{y}_{a_k} - \hat{\theta}, 0\} - \max\{\Delta d_{b_k} \hat{y}_{a_k} - \hat{\theta}, 0\})
\end{aligned} \tag{25}$$

subtracted. We now show that the latter term is non-negative. For all $k \in \mathcal{K}$, we have

$$\begin{aligned}
&\max\{\Delta d_{a_k} \hat{y}_{a_k} - \hat{\theta}, 0\} - \max\{\Delta d_{b_k} \hat{y}_{a_k} - \hat{\theta}, 0\} \\
&= \begin{cases} (\Delta d_{a_k} - \Delta d_{b_k}) \hat{y}_{a_k}, & \Delta d_{b_k} \hat{y}_{a_k} \geq \Delta d_{a_k} \hat{y}_{a_k} \geq \hat{\theta}, \\ -\Delta d_{b_k} \hat{y}_{a_k} + \hat{\theta}, & \Delta d_{b_k} \hat{y}_{a_k} \geq \hat{\theta} \geq \Delta d_{a_k} \hat{y}_{a_k}, \\ 0, & \hat{\theta} \geq \Delta d_{b_k} \hat{y}_{a_k} \geq \Delta d_{a_k} \hat{y}_{a_k}, \end{cases}
\end{aligned}$$

and we consequently obtain

$$d_{a_k} \hat{y}_{a_k} x_{a_k} + (\Delta d_{b_k} - \Delta d_{a_k}) \hat{y}_{a_k} + \max\{\Delta d_{a_k} \hat{y}_{a_k} - \hat{\theta}, 0\} - \max\{\Delta d_{b_k} \hat{y}_{a_k} - \hat{\theta}, 0\}$$

$$= \begin{cases} d_{a_k} \hat{y}_{a_k} x_{a_k}, & \Delta d_{b_k} \hat{y}_{a_k} \geq \Delta d_{a_k} \hat{y}_{a_k} \geq \hat{\theta}, \\ d_{a_k} \hat{y}_{a_k} x_{a_k} - \Delta d_{a_k} \hat{y}_{a_k} + \hat{\theta}, & \Delta d_{b_k} \hat{y}_{a_k} \geq \hat{\theta} \geq \Delta d_{a_k} \hat{y}_{a_k}, \\ d_{a_k} \hat{y}_{a_k} x_{a_k} + (\Delta d_{b_k} - \Delta d_{a_k}) \hat{y}_{a_k}, & \hat{\theta} \geq \Delta d_{b_k} \hat{y}_{a_k} \geq \Delta d_{a_k} \hat{y}_{a_k}. \end{cases}$$

The latter is greater or equal to $d_{a_k} \hat{y}_{a_k} x_{a_k}$ for all $k \in \mathcal{K}$. In particular, we have $d_{a_k} \hat{y}_{a_k} x_{a_k} = 0$ if $x_{a_k} = 0$ and $d_{a_k} \hat{y}_{a_k} > 0$, otherwise. To sum up, the term in (25) is non-negative. This means that feasibility w.r.t. Inequality (23) implies feasibility w.r.t. Inequality (24). Since (23) is a valid inequality, this concludes the proof. \square

In the next theorem, we consider lifted versions for the enhanced interdiction cuts stated in (20).

Theorem 5. *For any arbitrary but fixed $\hat{y} \in \hat{Y}$ and $\ell \in \{\Gamma_d, \dots, n+1\}$, let $K \in [n]$, $S_a^\ell = \{a_1, \dots, a_K\} \subset D_+^\ell$, and $S_b^\ell = \{b_1, \dots, b_K\} \subset D_+^\ell$ be chosen such that $S_a^\ell \cap S_b^\ell = \emptyset$, $\hat{y}_{a_k} = 1$, $\hat{y}_{b_k} = 0$, $B_{a_k} \geq B_{b_k}$, and $\tilde{d}(\ell)_{a_k} < \tilde{d}(\ell)_{b_k}$ hold for all $k \in [K]$. Under Assumptions 5 and 6, the lifted interdiction cut*

$$\eta \geq -\Gamma_d \Delta d_\ell + \sum_{i \in D_+^\ell} \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) + \sum_{k=1}^K (\tilde{d}(\ell)_{b_k} - \tilde{d}(\ell)_{a_k}) (1 - x_{b_k}) \quad (26)$$

is valid for Problem (2) with the lower-level optimal-value function (3).

Proof. Let (x, η) be a feasible leader's decision for Problem (2) with the lower-level optimal-value function (3). If $x_{b_k} = 1$ holds for all $k \in [K]$, we obtain the enhanced interdiction cut as stated in (20), which is satisfied by x . Otherwise, let us define $\mathcal{K} := \{k \in [K] : x_{b_k} = 0\}$ as well as $\bar{\mathcal{K}} := D_+^\ell \setminus \{a_k, b_k : k \in \mathcal{K}\}$ and consider the alternative decision of the follower given by

$$y'_i = \begin{cases} \hat{y}_i, & i \in \bar{\mathcal{K}}, \\ 1, & i \in \{b_k : k \in \mathcal{K}\}, \\ 0, & i \in \{a_k : k \in \mathcal{K}\}. \end{cases}$$

By construction, we have $y' \in Y$. Since the interdiction cuts (20) are not only valid for vertices contained in the set \hat{Y} but also for any lower-level feasible point, the leader's decision x satisfies the basic interdiction cut associated with y' and ℓ , i.e.,

$$\begin{aligned} \eta &\geq -\Gamma_d \Delta d_\ell + \sum_{i \in D_+^\ell} \tilde{d}(\ell)_i y'_i (1 - x_i) \\ &= -\Gamma_d \Delta d_\ell + \sum_{i \in \bar{\mathcal{K}}} \tilde{d}(\ell)_i \underbrace{y'_i}_{=\hat{y}_i} (1 - x_i) \\ &\quad + \sum_{k \in \mathcal{K}} \left(\tilde{d}(\ell)_{a_k} \underbrace{y'_{a_k}}_{=0} (1 - x_{a_k}) + \tilde{d}(\ell)_{b_k} \underbrace{y'_{b_k} (1 - x_{b_k})}_{=1} \right) \\ &= -\Gamma_d \Delta d_\ell + \sum_{i \in \bar{\mathcal{K}}} \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) + \sum_{k \in \mathcal{K}} \tilde{d}(\ell)_{b_k}. \end{aligned} \quad (27)$$

Further, we can re-write Inequality (26) as

$$\begin{aligned}
\eta &\geq -\Gamma_d \Delta d_\ell + \sum_{i \in D_+^\ell} \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) + \sum_{k=1}^K (\tilde{d}(\ell)_{b_k} - \tilde{d}(\ell)_{a_k}) (1 - x_{b_k}) \\
&= -\Gamma_d \Delta d_\ell + \sum_{i \in \bar{\mathcal{K}}} \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) + \sum_{k \in \mathcal{K}} \tilde{d}(\ell)_{a_k} \underbrace{\hat{y}_{a_k}}_{=1} (1 - x_{a_k}) \\
&\quad + \sum_{k \in \mathcal{K}} \tilde{d}(\ell)_{b_k} \underbrace{\hat{y}_{b_k}}_{=0} (1 - x_{b_k}) + \sum_{k=1}^K (\tilde{d}(\ell)_{b_k} - \tilde{d}(\ell)_{a_k}) (1 - x_{b_k}) \\
&= -\Gamma_d \Delta d_\ell + \sum_{i \in \bar{\mathcal{K}}} \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) + \sum_{k \in \mathcal{K}} \tilde{d}(\ell)_{a_k} (1 - x_{a_k}) \\
&\quad + \sum_{k \in \mathcal{K}} (\tilde{d}(\ell)_{b_k} - \tilde{d}(\ell)_{a_k}) \underbrace{(1 - x_{b_k})}_{=1} + \sum_{k \in [K] \setminus \mathcal{K}} (\tilde{d}(\ell)_{b_k} - \tilde{d}(\ell)_{a_k}) \underbrace{(1 - x_{b_k})}_{=0} \\
&= -\Gamma_d \Delta d_\ell + \sum_{i \in \bar{\mathcal{K}}} \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) + \sum_{k \in \mathcal{K}} \tilde{d}(\ell)_{a_k} (1 - x_{a_k}) \\
&\quad + \sum_{k \in \mathcal{K}} (\tilde{d}(\ell)_{b_k} - \tilde{d}(\ell)_{a_k}) \\
&= -\Gamma_d \Delta d_\ell + \sum_{i \in \bar{\mathcal{K}}} \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) + \sum_{k \in \mathcal{K}} \tilde{d}(\ell)_{b_k} - \sum_{k \in \mathcal{K}} \tilde{d}(\ell)_{a_k} x_{a_k}.
\end{aligned} \tag{28}$$

Note that the right-hand side of (28) corresponds to the right-hand side of (27) with the additional term

$$\sum_{k \in \mathcal{K}} \tilde{d}(\ell)_{a_k} x_{a_k} \geq 0$$

subtracted. Here, we have $\tilde{d}(\ell)_{a_k} x_{a_k} = 0$ if $x_{a_k} = 0$ and $\tilde{d}(\ell)_{a_k} x_{a_k} > 0$, otherwise. Hence, feasibility w.r.t. Inequality (27) implies feasibility w.r.t. Inequality (28). Since (27) is a valid inequality, this concludes the proof. \square

Let us point out that the items in the sets S_a and S_b (or S_a^ℓ and S_b^ℓ in the multi-follower case) can be paired in different ways. This might yield different lifted cuts. To this end, we consider the following separation procedure. For an item a that is a candidate to enter the set S_a or S_a^ℓ , i.e., $\hat{y}_a \geq 1$ holds, we select its counterpart b among all items that satisfy the requirements with maximum value of $((d_b - \Delta d_b) - (d_a - \Delta d_a)) \hat{y}_{a_k} (1 - x_b)$ and $(\tilde{d}(\ell)_b - \tilde{d}(\ell)_a) (1 - x_b)$ for the extended formulation and the multi-follower-based approach, respectively. If such a pair (a, b) exists, items a and b are inserted into the sets S_a and S_b (or S_a^ℓ and S_b^ℓ in the multi-follower case) and then removed from any further consideration.

3.3. Problems with an Uncertain Lower-Level Constraint. To conclude this section, we briefly address uncertainties that only arise in a single packing-type constraint of the follower as stated around (5) and (6). For the validity of the proposed cuts, we need to impose the following.

Assumption 11. *The uncertain lower-level constraint does not contain terms depending on the leader's decision, i.e., $v = 0$.*

Assumption 12. *All constraint coefficients of the uncertain lower-level constraint are non-negative, i.e., $w_i \geq 0$ holds for all $i \in [n]$.*

Assumption 11 is an analog of Assumption 8 stating that the leader's variables x are linked to the lower-level problem only via the interdiction constraints $y_i \leq u_i(1 - x_i)$. Assumption 12 ensures that the Γ_w -robust lower-level problems (8) and (11) satisfy the downward monotonicity property. Both assumptions are necessary to exploit a penalized formulation of the Γ_w -robust follower's problem to derive valid cuts as it is done in the case

of uncertain objective function coefficients. Again, we remove interdiction constraints and add penalty terms $-d_i y_i x_i$ to the objective function for all $i \in [n]$. Hence, we consider the same deterministic objective function

$$\sum_{i=1}^n d_i y_i (1 - x_i)$$

in the extended formulation as well as in the multi-follower formulation. In particular, the objective function is linear for a fixed leader’s decision x . However, the description of the resulting feasible set differs for both approaches. For the extended formulation, we maximize over the feasible set of the penalized follower’s problem projected onto the y -space, which is given by

$$\Theta = \{y \in \mathcal{Y} : \exists z, \theta \geq 0 \text{ such that (8b) and (8c) are satisfied}\}.$$

The feasible set Θ is independent from the leader’s decision. Hence, an optimal solution of the Γ_w -robust follower’s problem is attained at a vertex of the convex hull of Θ . We denote $\hat{\Theta}$ as the set containing the finite number of vertices of the convex hull of Θ . Then, the interdiction cuts

$$\eta \geq \sum_{i=1}^n d_i \hat{y}_i (1 - x_i) \quad \text{for all } \hat{y} \in \hat{\Theta} \quad (29)$$

are valid for Problem (2) and can equivalently replace Constraint (2c). Under Assumptions 5 and 6, when considering the multi-follower approach, we maximize over the ℓ -dependent feasible sets

$$\left\{ y \in \{0, 1\}^n : \Gamma_w \Delta w_\ell + \sum_{i=1}^n \tilde{w}(\ell)_i y_i \leq C \right\},$$

which may differ in each follower sub-problem. However, it is easy to see that an optimal solution needs to be contained in the union of all ℓ -dependent sets. Let Y' be the set containing the finite number of vertices of the convex hull of the union of all ℓ -dependent sets. Then, we obtain the interdiction cuts for the multi-follower case by replacing $\hat{y} \in \hat{\Theta}$ with $\hat{y} \in Y'$ in (29), which can equivalently replace Constraint (2c) in Problem (2).

4. COMPUTATIONAL RESULTS

We now provide detailed numerical results for the proposed methods to solve interdiction problems with a monotone Γ -robust follower. Our solution approaches are implemented in Python 3.6.9 and Gurobi 9.1.2 is used to solve all arising optimization problems.¹ To add the interdiction cuts described in the previous sections, we use Gurobi’s lazy constraint callbacks, which requires to set the parameter `LazyConstraints` to 1. All other parameters have been left at their default settings. The tests have been realized on an Intel XEON SP 6126 at 2.6 GHz (6 cores) with 32 GB RAM, which is part of the high performance cluster “Elwetritsch” at TU Kaiserslautern within the “Alliance of High Performance Computing Rheinland-Pfalz” (AHRP).² For each test run, we set the time limit to 1 h. We refer to the branch-and-cut method that exploits the multi-follower approach as `MF` and to the one that is based on the extended formulation as `Ext`. In Section 3.2, we have discussed several enhancements to improve the performance of `MF` and `Ext`, which we assess in the following. The aim is to determine a “winner setting” for `MF` and `Ext` and to compare both approaches. For the ease of presentation, we focus on the bilevel knapsack interdiction problem as a typical example of an interdiction problem with a monotone follower. Moreover, we only discuss the results for problems with uncertainties regarding the objective function coefficients.

Before we start, let us also comment on that there is no other tailored method in the literature that solves bilevel knapsack interdiction problems using a Γ -robust treatment of uncertainties. Hence, there are no alternative methods that we could compare with. However,

¹The code for the approaches presented in this paper is publicly available at <https://github.com/YasmineBeck/gamma-robust-knapsack-interdiction-solver>.

²We kindly acknowledge the support of RHRK (<https://rz.rptu.de/en/>).

using the extended formulation, one could transform the Γ -robust interdiction problem into a standard mixed-integer linear bilevel problem. The latter can, in general, be solved with the MibS solver (Tahernejad, Ralphs, and DeNegre 2020) and the general branch-and-cut solver presented in Fischetti, Ljubić, et al. (2017). We tested both solvers for 40 robustified knapsack instances with 35 items; see below for more details on our test set. MibS is not able to solve any of these instances—although 35 items belong to the smallest class of instances in our test set. The smallest optimality gap we get is 105%. The other solver solved 2 out of 40 instances. The runtimes are 1797.27s and 3188.13s, which is more than a factor of 360 or 635 longer than our tailored methods take. Consequently, we omit a more detailed comparison with these general-purpose bilevel solvers.

4.1. Generation of Knapsack Test Instances. To test our solution approaches, we consider the bilevel knapsack interdiction problem that has been considered in Caprara et al. (2016) and which is formally stated as

$$\begin{aligned} \min_{x \in \{0,1\}^n} \quad & p^\top y \\ \text{s.t.} \quad & v^\top x \leq B, \\ & y \in \arg \max_{y' \in \{0,1\}^n} \{p^\top y' : w^\top y' \leq C, y'_i \leq 1 - x_i, i \in [n]\} \end{aligned}$$

with $B, C \in \mathbb{Z}_+$, and $p, v, w \in \mathbb{Z}_+^n$. In particular, the bilevel knapsack interdiction problem is a prominent example for an interdiction problem with a monotone follower. The motivation for us to focus on this type of problem is the following. Classic knapsack problems belong to the most intensively studied discrete optimization problems, which is due to their relevance in many real-world applications, e.g., in the field of economics. In particular, the bilevel knapsack interdiction problem naturally extends the classic knapsack problem such as to capture competitive situations; see, e.g., DeNegre (2011) for a specific application in corporate strategy. Moreover, the knapsack interdiction problem is commonly used as a benchmark for testing bilevel optimization solvers; see, e.g., DeNegre and Ralphs (2009) and Tang et al. (2016). In its deterministic variant, the knapsack interdiction problem has been studied, e.g., in Caprara et al. (2013, 2016), Della Croce and Scatamacchia (2020), DeNegre (2011), Fischetti, Ljubić, et al. (2019), Fischetti, Monaci, et al. (2018), Shi et al. (2020), and Tang et al. (2016). For our computational study, we generate deterministic knapsack interdiction instances according to Martello et al. (1999), which we adapt to account for a Γ -robust follower. Before we comment on the uncertainty parameterizations that we consider, let us briefly describe the generation of the knapsack instances. The profits p_i and the follower's weights w_i take uncorrelated integer values from the interval $[0, 100]$. For each instance size $n \in \{35, 40, 45, 50, 55, \dots, 100\}$, 10 instances have been generated. The follower's knapsack capacity C is set to $\lceil (N/11) \sum_{i=1}^n w_i \rceil$, where $N \in \{1, \dots, 10\}$ is used to identify the instance number. The leader's weights v_i and the interdiction budget B are uniformly random integers from the intervals $[0, 100]$ and $[C - 10, C + 10]$. Let us point out that the deterministic knapsack instances with size $n \in \{35, 40, \dots, 55\}$ are taken from Caprara et al. (2016). All other instances have not yet been studied in the literature and are newly generated. To study the effects of a Γ -robust follower, we consider four different uncertainty parameterizations. We assume that the deviations take either 10% or 25% of the nominal value. The parameter Γ is set to either 10% or 50% of the instance size n . In the case of a fractional value for Γ , we then consider the closest integer. Hence, our test set contains 560 robustified knapsack instances. Finally, let us mention that we do not consider any instances with more than 100 items because even the most advanced variants of the presented approaches cannot solve all of the 560 robustified knapsack instances described above within a reasonable amount of time.

4.2. Lifted Cuts and Dominance Inequalities. We now assess the influence of lifted cuts and dominance inequalities on the overall performance of the solution method. To this end, we consider the following four settings.

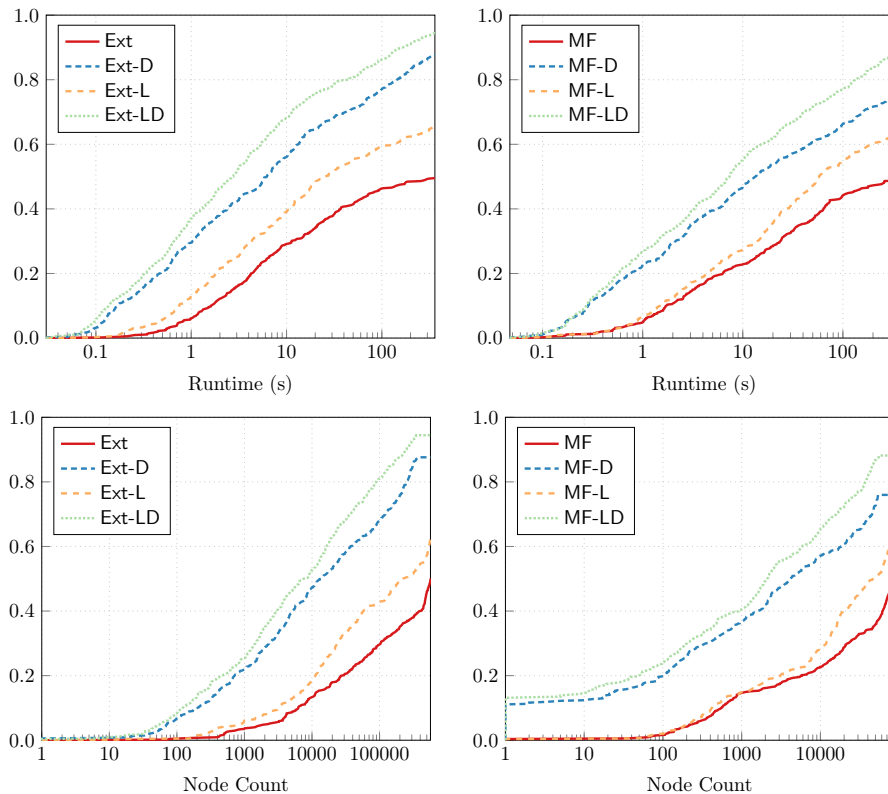


FIGURE 1. Log-scaled ECDF plots of the runtimes (in s) (top) and the number of branch-and-bound nodes (bottom) for Ext (left) and MF (right) using lifted cuts and/or dominance inequalities.

MF/Ext: The basic setting in which only basic interdiction cuts are added without any further enhancements.

MF-D/Ext-D: The basic setting with the addition of dominance inequalities (21) regarding the leader’s decision.

MF-L/Ext-L: The basic setting but instead of considering basic interdiction cuts, we add lifted interdiction cuts.

MF-LD/Ext-LD: Like MF-L or Ext-L but with the addition of dominance inequalities (21).

As a default for MF, we consider the cut separation strategy in which all violated (lifted) interdiction cuts are added to the problem formulation. Figure 1 shows the empirical cumulative distribution functions (ECDFs) w.r.t. the runtimes and the number of branch-and-bound nodes of the four settings. The ECDFs can be interpreted as the percentage of instances (y -axis) that can be solved within a certain amount of time and/or after investigating a certain number of branch-and-bound nodes (log-scaled x -axis). Note that, to have a fair comparison, we exclude 29 instances that none of the considered variants can solve within the time limit. Moreover, we exclude 19 instances that every variant can solve within 5 s to avoid drawing erroneous conclusions because of low runtimes. Hence, we consider a total of 512 instances at this point. While lifted interdiction cuts only slightly improve the performance of MF and Ext, the use of dominance inequalities significantly enhances the performance of both approaches. The combination of lifted cuts and dominance inequalities yields only minor further improvement compared to MF-D and Ext-D. Nevertheless, MF-LD and Ext-LD dominate all other settings of the respective solution approach. This observation is also underlined by the results in Table 1. In what follows, we thus hold on to the variants

TABLE 1. Mean and median runtimes (in s), mean and median number of branch-and-bound nodes as well as the number of solved instances for the variants with lifted cuts and/or dominance inequalities.

	runtimes		node count		solved
	mean	median	mean	median	
Ext	1863.94	1507.90	689196.96	574953.50	268
Ext-D	583.12	53.52	108588.91	10254.50	473
Ext-L	1337.09	188.62	435544.75	168709.50	353
Ext-LD	260.25	21.86	46412.61	5593.50	510
MF	1933.78	2147.53	106190.09	87702.00	265
MF-D	1018.79	100.61	26374.73	3197.50	410
MF-L	1482.54	510.86	69689.89	36435.00	344
MF-LD	573.54	62.09	15212.53	1771.00	476

with additional dominance inequalities and lifted interdiction cuts as our “winner setting” for both approaches.

4.3. The Benefits of Maximal Packings. We consider maximal packings of the follower to avoid the generation of unnecessary interdiction cuts. This is achieved in the following manner. We determine a feasible follower’s decision y for the current x by either solving the follower’s extended formulation (14) or the sub-problems (15). In the next step, we complete the follower’s decision to a maximal packing in a greedy-like fashion. To this end, we order the indices of the follower’s variables according to non-increasing profit-to-weight ratio and then gradually add items that still fit into the follower’s knapsack. When considering the extended formulation, however, we need to further check if the follower’s decision (y, z, θ) satisfies $\theta \geq \Delta p_i$. This requirement is necessary to ensure the feasibility of the maximal packing w.r.t. the constraints $z_i + \theta \geq \Delta p_i y_i$ for all $i \in [n]$. Finally, we generate and add only the interdiction cut that corresponds to the follower’s maximal packing.

To assess the influence of maximal packings, we adopt the parameterizations of our previous “winner settings” MF-LD and Ext-LD with the difference that we now only add cuts corresponding to maximal packings of the follower. We refer to these settings as MF-LD-Max and Ext-LD-Max. Again, for a fair comparison, we exclude 21 instances that none of the variants can solve within the time limit and 90 instances that every variant can solve within 5s so that we consider a total of 449 instances here. Based on Figure 2, it can be seen that adding only the interdiction cuts corresponding to maximal packings of the follower improves the performance of the overall solution method for both approaches. In particular, this holds true for the easier instances, which is also underlined by the number of solved problems presented in Table 2. Table 2 further shows considerably smaller mean and median runtimes for MF-LD-Max and Ext-LD-Max compared to their counterparts without maximal packings. Also the mean and median number of branch-and-bound nodes visited is significantly smaller in both settings when maximal packings of the follower are considered. To sum up, the observations drawn from Figure 2 and Table 2 clearly suggest that MF-LD-Max and Ext-LD-Max are the “winner settings” among the considered variants.

As mentioned in Section 3.2, we also considered leader’s maximal packings. However, preliminary computational results revealed that maximal packings for the leader interfere with Gurobi’s integrated branching rules and node selection, which is why we decided to refrain from this ingredient.

4.4. The Impact of Warmstarting. In our computational study, we also investigated how warmstarting the proposed methods may affect their performance. To this end, we considered two options—a heuristic similar to the one presented in Fischetti, Ljubić, et al. (2019) and solving the nominal knapsack interdiction problem—to determine an initial feasible (and

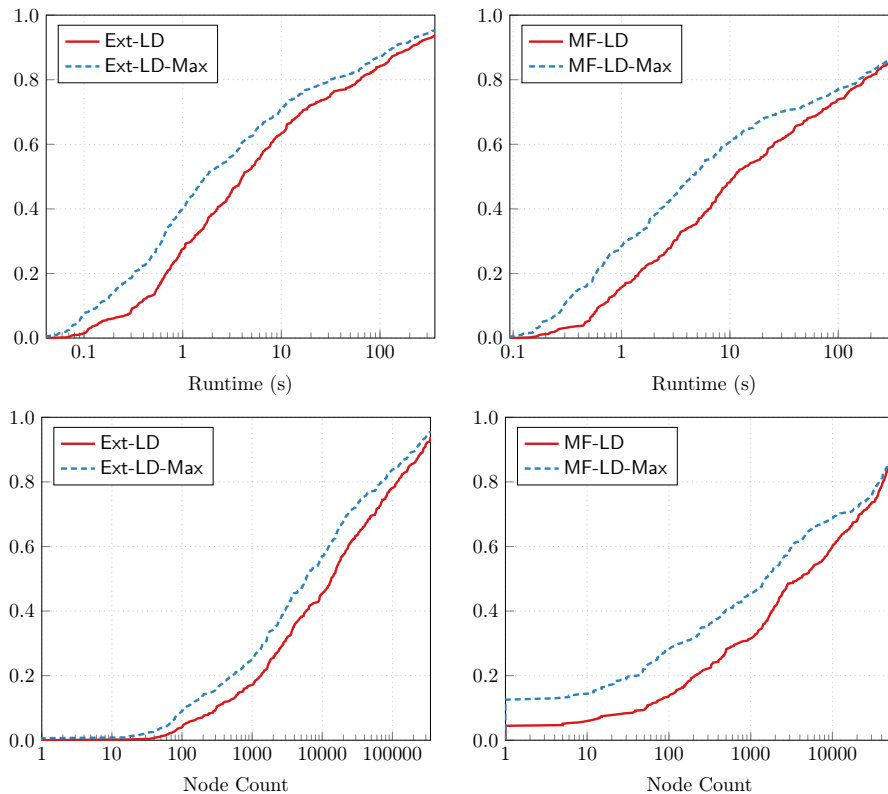


FIGURE 2. Log-scaled ECDF plots of the runtimes (in s) (top) and the number of branch-and-bound nodes (bottom) for Ext (left) and MF (right) considering maximal packings of the follower.

TABLE 2. Mean and median runtimes (in s), mean and median number of branch-and-bound nodes as well as the number of solved instances for the variants with and without maximal packings of the follower.

	runtimes		node count		solved
	mean	median	mean	median	
Ext-LD	360.70	37.71	61901.68	12293.00	439
Ext-LD-Max	252.71	15.32	41166.66	5253.00	448
MF-LD	718.13	94.62	18089.67	2789.00	405
MF-LD-Max	616.47	37.65	15079.47	1373.00	406

potentially good) decision of the leader for Problem (P_0). The pair (x, η) that we obtain using any of the two options is then provided to Gurobi as MIP start.

The computational results obtained for both settings revealed that warmstarting the methods has no significant impact on their performance, neither for MF nor for Ext. This is why we omit the details on the runtimes and the number of branch-and-bound nodes here. As per default, both methods are warmstarted using the heuristic in all subsequent considerations.

4.5. Comparison of Different Cut Separation Strategies and the Potential of Parallelization. We further evaluate enhancement techniques that exploit the special structure of the multi-follower formulation. Due to the fact that the overall problem

can be considered as a single-leader-multi-follower problem with independent followers, we can make use of parallelization as briefly mentioned in Section 2. In this context, we can further consider different cut separation strategies instead of adding all violated interdiction cuts in each iteration of the algorithm. This way, the number of cuts added to the problem formulation can be reduced, which might speed up the overall solution method. We adopt the parameterizations of the previous “winner setting”, i.e., we consider the multi-follower formulation with lifted cuts corresponding to maximal packings of the follower, dominance inequalities regarding the leader’s decision, and heuristic warmstarts. For notational convenience, we omit MF-LD-Max as a prefix for the considered variants when we focus on the following cut separation strategies.

All-In: The default setting in which all interdiction cuts that are violated by the current leader’s decision are added to the problem formulation.

Most-Violated: A single cut is added corresponding to the interdiction inequality (20) that is maximally violated by the current leader’s decision.

Sorting: In Álvarez-Miranda, Fernández, et al. (2015), the authors propose a learning mechanism to identify the sub-problems that produce potentially good cuts. This is done by taking the information of previous iterations into account. We adapt the proposed strategy to our setting such that a single potentially good cut is added in each iteration.

First-In: We iterate over $\ell \in \{\Gamma, \dots, n+1\}$, add the first interdiction cut that is violated by the current leader’s decision, and then break the loop.

Random: Among the violated interdiction cuts, we randomly choose a single cut and add it to the problem formulation.

To assess the potential of parallelization, we consider so-called *idealized runtimes*, which reflect the overall runtime of the solution method provided that there are sufficient capacities available to solve all sub-problems in parallel. For each instance, the idealized runtimes are computed after solving the problems sequentially by taking the maximum over the runtime of each sub-problem.

In Figure 3, we compare the aforementioned cut separation strategies w.r.t. sequential and idealized runtimes. Note that we exclude 21 instances that none of the considered variants can solve within the time limit and 43 instances that every variant can solve within 5 s. Hence, we consider a total of 496 instances here. It can be seen that the first insertion strategy First-In harms the performance of the solution method for both sequential and idealized runtimes as a benchmark. The last observation is also reflected by the results on the number of branch-and-bound nodes presented in Figure 4 (left). A possible reason for this might be the following. By adding the first violated interdiction cut, later follower sub-problems that might produce stronger cuts are neglected. In particular, it is possible that the cut corresponding to the same follower sub-problem is added in each iteration of the algorithm. To overcome this situation, we consider Random as a modification of First-In. Based on Figure 3, it can be observed that this variant performs significantly better. In particular, Random outperforms all other cut separation strategies w.r.t. sequential runtimes. The potential reasons for the effectiveness of the randomized first insertion strategy are twofold. On the one hand, it seems beneficial to add a single cut in each iteration of the algorithm; otherwise, as for the strategy All-In, the leader’s problem can get extremely large w.r.t. the number of constraints. On the other hand, Random has comparatively low computational costs. For Sorting, it can be seen that having information from previous iterations does not lead to a better choice than randomly adding a violated cut. When considering Most-Violated, we need to solve all sub-problems to determine the most violated interdiction cut, which seems to be rather expensive. The previous observations are underlined by the results in Table 3, since the sequential mean and median runtimes for Random are considerably smaller compared to all other cut separation strategies. Focusing on idealized runtimes, however, it is noteworthy that Random is no longer the “winning” cut separation strategy. Based on Figure 3 and Table 3, Most-Violated dominates the randomized first insertion strategy w.r.t. idealized runtimes.

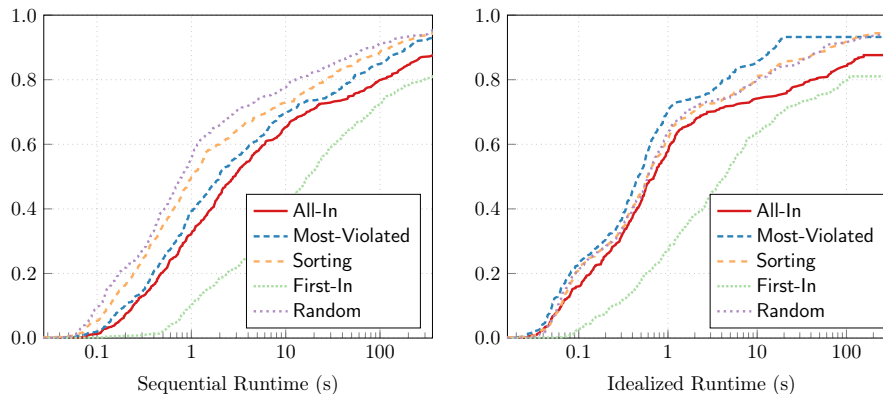


FIGURE 3. Log-scaled ECDF plots of the sequential runtimes (left) and the idealized runtimes (right) (in s) for different cut separation strategies.

TABLE 3. Sequential and idealized mean and median runtimes (in s) as well as the number of solved instances for the variants with different cut separation strategies.

	sequential runtimes		idealized runtimes		
	mean	median	mean	median	solved
All-In	543.17	24.85	267.35	6.03	452
Most-Violated	346.88	17.98	37.53	4.35	481
Sorting	245.16	9.19	138.89	5.53	487
First-In	855.10	157.92	466.59	37.68	418
Random	183.52	7.23	156.57	5.61	492

This is to be expected since we can benefit the most from parallelization for **Most-Violated**, where we indeed have to solve all of the sub-problems. In particular, the previous observations suggest that the higher computational costs for solving all sub-problems—especially if this is done in parallel—can be compensated to some extent by the strength of the added cuts. This strength is further visualized in Figure 4 (right) and Table 4. Here, to assess the quality of the added cuts, we consider the runtimes independently of the times spent for the cut generation. However, it is noteworthy that, apart from the strength of **Most-Violated**, the results also suggest that **Sorting** and **Random** yield cuts of good quality for the easier instances. Finally, we further highlight the benefit of adding the most-violated cut by considering the mean and median number of branch-and-bound nodes presented in Table 4. It can be observed that the mean and median number of nodes for **Most-Violated** are significantly smaller than the ones for almost all of the other cut separation strategies. Only **All-In** has a considerably smaller mean number of investigated nodes compared to **Most-Violated**. However, as stated earlier, it seems beneficial to add a single cut in each iteration of the algorithm instead of adding all violated cuts, which is underlined by both the runtime results as well as the number of solved instances.

To sum up, **Most-Violated** yields the strongest cuts among the considered variants. Moreover, provided that the necessary capacities are available to solve all sub-problems in parallel, **Most-Violated** particularly outperforms all other variants w.r.t. idealized runtimes. Hence, we consider **Most-Violated** as our “winning” strategy in the idealized setting. When considering sequential runtimes, however, the cost of solving all sub-problems cannot be completely compensated by the strength of the added cuts. Since the overall runtime is our decisive criterion, we prefer **Random** in the sequential setting.

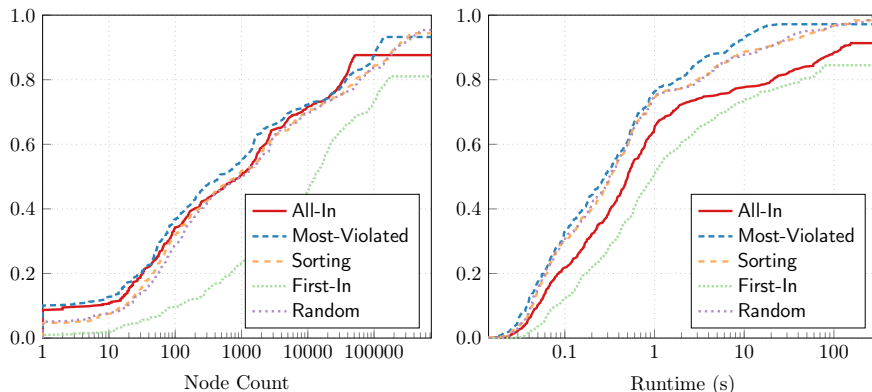


FIGURE 4. Log-scaled ECDF plots of the number of branch-and-bound nodes (left) and the runtimes (in s) independent of the times spent for the cut generation (right) for different cut separation strategies.

TABLE 4. Mean and median runtimes (in s) independent of the times spent for the cut generation as well as the mean and median number of branch-and-bound nodes for the variants with different cut separation strategies.

	runtimes		node count	
	mean	median	mean	median
All-In	258.36	5.05	14468.14	670.00
Most-Violated	27.67	2.83	24159.59	351.00
Sorting	104.84	3.26	39044.19	587.50
First-In	342.75	9.72	72248.80	11379.00
Random	113.29	3.16	45086.26	690.50

4.6. Comparison of the Solution Approaches. We now compare the “winning” parameterizations of the extended formulation and the multi-follower formulation. For the multi-follower-based approach, we particularly distinguish between the sequential and the idealized setting. Hence, we consider MF-LD-Max-Random, MF-LD-Max-Most-Violated, and Ext-LD-Max, which we abbreviate with MF-seq, MF-ideal, and Ext in the following. Again, we exclude 16 instances that none of the considered variants can solve within the time limit and 137 instances that every variant can solve within 5 s. In total, we thus consider 407 instances. Figure 5 (left) shows the ECDF plots w.r.t. the runtimes of the three considered variants. Note that we consider sequential and idealized runtimes for MF-seq and MF-ideal, respectively. We observe that MF-ideal clearly outperforms the remaining two approaches. This particularly affirms that the strength of the multi-follower-based approach lies in the possibility to parallelize the solution of the follower sub-problems. The previous observation is also underlined by the mean and median runtimes in Table 5. It can be seen that MF-ideal has significantly smaller mean and median runtimes compared to MF-seq and Ext. Despite being not as significant as when considering the runtimes, the same qualitative behavior can also be observed for the results on the number of nodes in Figure 5 (right) and Table 5. If, however, the capacity is not available to have an idealized parallelization, the multi-follower-based approach MF-seq still performs slightly better than Ext. Based on Figure 5, MF-seq seems to have an advantage over Ext on the easier instances. The last observation is also supported by the results on the mean and median runtimes presented in Table 5. Yet the number of solved instances suggests that Ext performs slightly better than MF-seq on the harder instances. Let us emphasize, however, that the amount of instances that can be solved by Ext but not

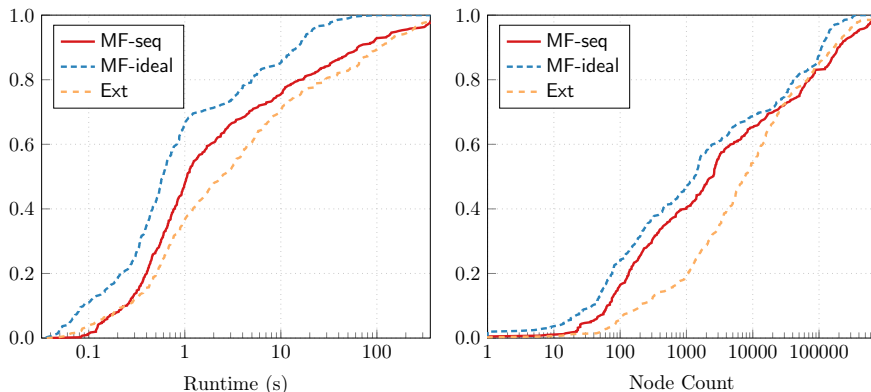


FIGURE 5. Log-scaled ECDF plots of the runtimes (in s) (left) and the number of branch-and-bound nodes (right) for the “winner settings” of Ext and MF.

TABLE 5. Mean and median runtimes (in s), mean and median number of branch-and-bound nodes as well as the number of solved instances for the “winner settings” of Ext and MF.

	runtimes		node count		
	mean	median	mean	median	solved
Ext	324.42	24.16	49818.23	7680.00	401
MF-seq	267.57	10.59	62861.15	2546.00	398
MF-ideal	43.64	5.99	29502.58	1333.00	387

by MF-seq is comparatively small, which is why we consider MF-seq as the overall better method in the sequential setting.

4.7. The Computational Cost of Robustness. To conclude our computational study, we address the computational cost of robustness. This expression captures the effect of robustification, e.g., on the overall runtimes of the solution methods. To evaluate the computational cost of robustness, we compare the three “winning” approaches MF-seq, MF-ideal, and Ext for the considered uncertainty parameterizations, which are referred to as (Δ, Γ) . Here, $\Delta \in \{10, 25\}$ is used to specify the considered percentage deviations in the objective function coefficients and $\Gamma \in \{10, 50\}$ denotes the percentage that the parameter Γ takes of the instance size. Based on Table 6, it can be seen that the mean and median runtimes to solve the Γ -robust knapsack interdiction problem increase with increasing values of Δ and Γ for Ext. For MF-seq and MF-ideal, however, this does not seem to be the case in principle. Detailed runtime results for each knapsack instance can further be found in Table 8. For both the multi-follower approach and the extended formulation, we compare the nominal runtimes to the mean runtimes obtained from all considered uncertainty parameterizations. Note that we label the sequential and idealized mean runtimes with the superscripts *seq* and *ideal*, respectively.

To further assess the cost of robustness w.r.t. the runtimes, we measure the relative performance of the method in the Γ -robust and in the nominal case. Note that we restrict ourselves to the instances that have been considered in Section 4.6 for a fair comparison, i.e., we exclude 16 instances that none of the “winner” settings can solve within the time limit and 137 instances that every variant can solve within 5 s. To measure the relative performance, we determine the coefficient of runtimes $q_i = t_{i,\text{rob}}/t_{i,\text{nom}}$ for each knapsack instance i . Here, $t_{i,\text{rob}}$ and $t_{i,\text{nom}}$ denote the runtimes of the considered solution method for instance i in

TABLE 6. Mean and median runtimes (in s) for the “winner” settings of Ext and MF for the uncertainty parameterizations given by (Δ, Γ) .

		uncertainty parameterization				
		$(\Delta, \Gamma) =$	(10,10)	(10,50)	(25,10)	(25,50)
mean	Ext		201.93	331.11	234.78	540.96
	MF-seq		301.37	234.55	334.64	192.01
	MF-ideal		43.63	52.38	43.93	34.94
median	Ext		6.34	41.50	10.70	84.54
	MF-seq		10.35	10.63	10.44	11.06
	MF-ideal		7.66	5.40	7.14	4.86

the Γ -robust and in the nominal case, respectively. In Figure 6, we show box-plots for the coefficients of runtimes corresponding to MF-seq, MF-ideal, and Ext. Each box in Figure 6 represents the distribution of the determined coefficients q over the 407 considered instances for the three “winning” solution methods. It can be seen that the box for Ext is considerably larger compared to the other two approaches. This shows that the coefficients for MF-seq and MF-ideal are less dispersed, which suggests that the performance of these methods is more stable compared to the one of Ext. In particular, the boxes for MF-seq and MF-ideal are of similar size, which reflects the similarity between the multi-follower-based solution approaches. Note that MF-seq and MF-ideal rely on the same solution procedure. The only difference between these approaches is that MF-seq is a sequential method and MF-ideal exploits parallelization. Let us point out that, in Figure 6, we use a logarithmic scaling of the y -axis for runtime coefficients greater than 40 such as to capture the spread of the outliers in a detailed and comprehensive way. For Ext, the outliers are widely scattered, which shows that the performance of the method is rather volatile. In contrast to that, smaller ranges of the outliers can be observed for MF-seq and MF-ideal. All previous observations are also affirmed by the results in Table 7. Based on Figure 6 and Table 7, it thus seems reasonable to prefer the multi-follower approach over the extended formulation since it seems to be the more stable method. Interestingly, we also observe coefficients of runtimes strictly smaller than 1 for some of the instances. This means that the robust counterpart can be solved faster than the nominal problem, i.e., robustification does not necessarily always lead to increased computational costs. However, the mean and median coefficients of runtimes shown in Table 7 emphasize that the robustification is not “for free”. However, the computational cost of robustness for the multi-follower-based approaches—MF-ideal in particular—is comparatively small.

TABLE 7. Minimum, mean, median, and maximum values of the coefficients of runtimes for the “winner” settings of Ext and MF.

	min	mean	median	max
Ext	0.79	128.59	14.68	4365.61
MF-seq	0.76	7.71	4.95	74.66
MF-ideal	0.04	4.60	1.66	65.37

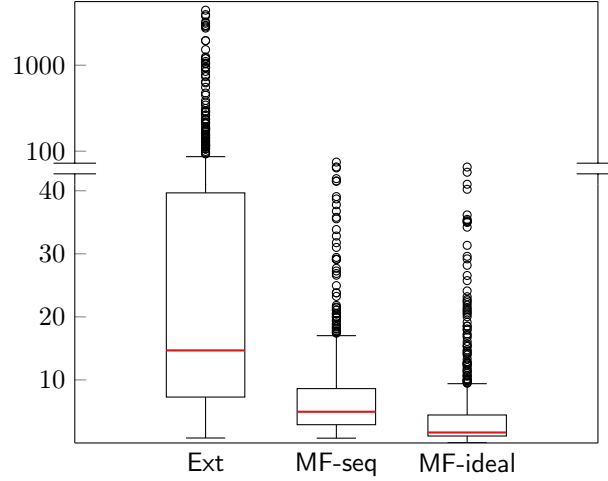


FIGURE 6. Box-plots of the coefficients of runtimes $q_i = t_{i,rob}/t_{i,nom}$ for the “winner” settings of Ext and MF.

To sum up, Γ -robust solutions are obtained at the expense of increased computational difficulty of the problem but, provided that there are sufficient capacities available to solve all of the follower sub-problems in parallel, the price of robustness w.r.t. runtimes is comparatively small. Thus, we prefer MF-ideal over MF-seq and Ext. In the sequential setting, the results in Table 7 clearly suggests that MF-seq outperforms Ext w.r.t. the stability of the method. Based on the results in Table 5, however, the advantage of MF-seq over Ext is not as significant as when the relative performance measure is considered. In particular, as elaborated in Section 4.6, Ext seems to have an advantage on harder instances, which also justifies the use of the extended formulation.

Table 8: Runtime results (in s) for the nominal knapsack instances and mean runtimes for their robust counterparts.

size	instance	Ext		MF		
		nominal	robust	nominal	robust ^{seq}	robust ^{ideal}
35	1	0.22	1.71	0.37	0.74	0.27
	2	0.49	6.00	0.49	1.13	0.45
	3	0.56	13.70	0.77	4.68	1.92
	4	0.22	1.75	1.34	6.08	3.23
	5	0.28	2.74	2.01	5.80	4.16
	6	0.14	0.71	0.38	0.69	0.49
	7	0.18	1.20	0.37	3.15	2.83
	8	0.16	0.53	0.14	0.39	0.18
	9	0.04	0.65	0.13	0.37	0.28

	10	0.05	0.31	0.04	0.32	0.14
40	1	0.33	2.28	0.36	1.43	0.50
	2	0.44	10.22	0.56	2.44	0.93
	3	1.47	42.99	1.25	11.75	3.41
	4	0.29	2.90	1.18	6.85	2.64
	5	0.20	2.51	2.09	6.11	4.53
	6	0.25	0.94	0.34	0.80	0.70
	7	0.15	1.18	0.35	0.86	0.52
	8	0.18	0.56	0.22	0.57	0.27
	9	0.18	4.85	0.31	3.81	3.45
	10	0.02	0.07	0.02	0.19	0.13
45	1	0.38	4.70	0.38	1.51	0.51
	2	0.38	10.91	0.46	1.61	0.61
	3	0.77	33.53	3.44	14.58	4.25
	4	0.43	34.35	7.34	22.32	6.71
	5	0.35	16.26	4.69	10.93	5.99
	6	0.21	1.39	0.46	3.41	2.98
	7	0.06	1.70	0.38	0.67	0.39
	8	0.13	0.86	0.27	1.02	0.67
	9	0.14	1.08	0.15	0.71	0.47
	10	0.20	1.01	0.22	0.70	0.45
50	1	0.36	5.11	0.40	1.57	0.77
	2	4.70	121.56	7.57	26.17	6.96
	3	5.30	76.00	10.08	19.37	7.45
	4	0.21	13.50	3.93	12.48	2.86
	5	0.42	45.12	13.53	34.86	9.56
	6	0.22	1.29	0.43	0.98	0.58
	7	0.17	2.46	0.40	1.46	0.97
	8	0.05	0.75	0.35	0.67	0.39
	9	0.16	9.69	0.30	4.36	4.08
	10	0.06	0.39	0.06	0.57	0.21
55	1	1.08	14.10	0.79	5.53	1.27
	2	2.55	86.23	4.31	19.58	4.45
	3	299.82	2191.23	261.03	870.73	196.44
	4	30.68	338.33	40.64	192.79	48.64
	5	0.67	38.71	15.82	20.45	12.10
	6	0.44	21.98	0.91	3.65	3.41
	7	0.23	8.73	0.57	4.25	3.59
	8	0.24	32.00	0.44	5.18	4.72
	9	0.14	3.21	0.16	1.18	0.74
	10	0.21	2.47	0.21	2.73	2.53
60	1	0.49	3.73	0.32	1.88	0.44
	2	3.28	58.43	1.49	14.08	3.30
	3	28.27	424.31	22.57	114.68	31.02
	4	32.84	358.24	35.87	182.28	54.59
	5	1.06	25.08	3.70	17.98	10.79
	6	0.66	6.44	0.43	5.37	5.16
	7	0.22	5.77	0.41	1.15	0.56
	8	0.24	4.99	0.44	3.47	2.88
	9	0.23	3.74	0.31	3.84	3.54
	10	0.20	1.53	0.27	0.74	0.46

65	1	0.39	4.97	0.51	1.93	0.45
	2	5.71	99.99	5.05	28.94	6.22
	3	66.36	705.40	61.75	172.25	56.87
	4	119.97	773.86	106.79	548.47	128.77
	5	0.59	78.25	9.81	60.65	28.67
	6	0.34	24.47	0.79	7.99	6.72
	7	0.28	19.71	0.79	5.67	5.27
	8	0.28	13.19	0.27	1.29	0.67
	9	0.28	7.37	0.18	5.67	5.28
	10	0.23	1.32	0.29	0.86	0.47
70	1	0.37	5.77	0.55	3.20	0.79
	2	6.95	173.91	10.91	48.16	10.97
	3	171.28	1477.50	197.35	594.18	166.50
	4	52.55	1168.15	143.76	566.51	160.45
	5	0.82	200.63	35.73	148.16	53.30
	6	0.42	32.51	0.74	6.44	4.02
	7	0.31	32.62	0.70	1.38	0.70
	8	0.18	20.74	0.21	2.16	1.42
	9	0.26	5.06	0.23	4.85	3.92
	10	0.23	2.67	0.15	1.56	1.01
75	1	0.65	10.65	0.76	4.53	0.97
	2	30.46	302.07	32.78	82.94	18.50
	3	185.97	1669.87	148.04	594.80	159.71
	4	2.37	1420.08	282.86	1476.37	349.61
	5	0.56	165.93	27.51	87.01	34.18
	6	0.40	40.84	1.29	4.12	3.31
	7	0.21	21.62	0.41	1.08	0.91
	8	0.21	15.62	0.30	6.81	6.63
	9	0.24	3.41	0.31	1.43	0.82
	10	0.11	0.91	0.28	0.85	0.48
80	1	0.66	15.50	0.65	9.21	1.94
	2	58.70	584.49	31.46	194.11	41.11
	3	513.98	2410.95	302.54	1554.35	399.95
	4	0.93	–	496.44	–	–
	5	0.52	187.66	27.88	109.24	38.73
	6	0.30	75.02	1.28	11.47	8.50
	7	0.35	42.64	0.65	7.76	6.42
	8	0.30	16.52	0.31	5.23	4.66
	9	0.10	4.79	0.08	1.85	1.13
	10	0.25	0.90	0.30	1.35	0.58
85	1	1.07	43.68	2.47	26.83	3.19
	2	84.63	875.17	89.83	353.61	73.44
	3	425.13	3112.44	333.26	2305.22	376.08
	4	0.97	–	635.59	–	–
	5	0.77	218.35	30.64	91.05	34.17
	6	0.36	55.90	1.31	4.23	2.17
	7	0.28	53.94	1.25	8.13	7.25
	8	0.21	22.15	0.24	7.73	6.85
	9	0.17	6.58	0.15	9.70	7.37
	10	0.11	0.71	0.09	1.37	0.52
90	1	2.60	43.73	2.73	19.40	3.87

	2	130.91	1184.05	110.48	506.53	106.59
	3	1324.50	–	1102.64	–	–
	4	1.25	–	1788.24	–	–
	5	0.85	298.27	29.03	74.70	28.46
	6	0.48	90.69	1.22	10.88	6.64
	7	0.29	53.34	0.92	9.32	7.33
	8	0.24	19.87	0.36	8.75	7.54
	9	0.21	7.75	0.21	8.54	6.76
	10	0.18	2.63	0.22	1.82	0.75
<hr/>						
95	1	1.54	64.46	2.55	33.70	4.15
	2	140.75	1203.87	101.87	702.22	104.00
	3	1935.09	–	1266.83	–	–
	4	1.87	–	–	–	–
	5	1.87	683.32	54.15	206.94	47.86
	6	0.54	155.16	2.57	12.58	3.53
	7	0.35	143.62	1.08	11.34	8.30
	8	0.31	39.91	0.46	8.81	7.65
	9	0.63	10.83	0.79	4.93	2.32
	10	0.14	1.79	0.14	1.56	0.53
<hr/>						
100	1	3.09	69.32	1.61	37.42	3.71
	2	203.20	1597.40	136.96	694.00	175.65
	3	–	–	–	–	–
	4	3.47	–	–	–	–
	5	0.82	1122.00	71.78	391.87	105.00
	6	0.63	491.01	3.35	23.15	10.84
	7	0.45	461.73	1.21	10.86	8.76
	8	0.71	58.45	0.70	2.37	1.62
	9	0.13	15.88	0.33	2.92	1.01
	10	0.32	3.56	0.33	9.13	8.12

5. CONCLUSION

In this paper, we consider discrete min-max problems with a follower facing uncertain lower-level data. We exploit a Γ -robust approach so that the follower only hedges against a subset of deviations in the uncertain parameters as to adversely affect the solution of the problem. We present two approaches—an extended formulation and a multi-follower formulation—to model this type of situation. For both frameworks, we present a fairly generic branch-and-cut method. Nevertheless, we can obtain stronger formulations for certain types of problems. As an example, we consider interdiction problems with a monotone Γ -robust follower to derive problem-tailored cuts that generalize existing interdiction cuts from the literature. Finally, we conduct a computational study to assess the performance of the two proposed solution approaches. To this end, we focus on the bilevel knapsack interdiction problem, which is one of the most prominent examples of monotone interdiction problems.

The computational results suggest that the extended formulation (Ext) performs slightly better on harder knapsack instances. However, smaller overall mean and median runtimes and a more stable performance of the method compared to Ext can be observed for the multi-follower formulation (MF). In particular, we can exploit parallelization for the multi-follower formulation, which is a major strength of this solution approach. Nevertheless, the study justifies the use of both the extended formulation as well as the multi-follower approach.

Despite the contribution of this paper, there are still several interesting research questions that require further investigation. We briefly sketch three of them.

- (1) Throughout this paper, we assume that there are no coupling constraints, i.e., there are no upper-level constraints explicitly depending on the variables of the follower. This is a crucial assumption for the validity of the proposed methods. Otherwise, we would not be able to project the follower’s variables out of the problem using the optimal-value function as it is done in Section 1. Nevertheless, developing solution methods for Γ -robust bilevel problems with coupling constraints is a reasonable aspect of future work.
- (2) In this paper, we focus on interdiction problems with a monotone Γ -robust follower to obtain problem-tailored cuts. An interesting direction for future research would be to investigate if generic cuts such as, e.g., intersection cuts (Fischetti, Ljubić, et al. 2016, 2018) can be adapted to the setting described in Section 2.
- (3) Finally, we would like to emphasize that we only consider uncertainties in a single packing-type constraint in the lower level. The extended formulation can easily be adapted to allow for deviations in multiple lower-level constraints. For the multi-follower formulation, however, this situation significantly increases the difficulty of the problem. This is due to the assumption regarding the ordering of the indices, which is required to exploit the results in Bertsimas and Sim (2003). In general, it is not possible to order the indices such that the deviations in the constraint coefficients are non-increasing if there are multiple uncertain constraints. Thus, this aspect is left for future research.

ACKNOWLEDGEMENTS

The third author thanks the DFG for their support within projects A05 and B08 in CRC TRR 154.

REFERENCES

- Álvarez-Miranda, E., E. Fernández, and I. Ljubić (2015). “The Recoverable Robust Facility Location Problem.” In: *Transportation Research Part B: Methodological* 79, pp. 93–120. DOI: [10.1016/j.trb.2015.06.001](https://doi.org/10.1016/j.trb.2015.06.001).
- Álvarez-Miranda, E., I. Ljubić, and P. Toth (2013). “A note on the Bertsimas & Sim algorithm for robust combinatorial optimization problems.” In: *4OR—A Quarterly Journal of Operations Research* 11, pp. 349–360. DOI: [10.1007/s10288-013-0231-6](https://doi.org/10.1007/s10288-013-0231-6).
- Arroyo, J. M. (2010). “Bilevel programming applied to power system vulnerability analysis under multiple contingencies.” In: *IET Generation, Transmission & Distribution* 4.2, pp. 178–190. DOI: [10.1049/iet-gtd.2009.0098](https://doi.org/10.1049/iet-gtd.2009.0098).
- Bazgan, C., S. Toubaline, and D. Vanderpooten (2013). “Critical edges/nodes for the minimum spanning tree problem: complexity and approximation.” In: *Journal of Combinatorial Optimization* 26, pp. 178–189. DOI: [10.1007/s10878-011-9449-4](https://doi.org/10.1007/s10878-011-9449-4).
- Beck, Y., I. Ljubić, and M. Schmidt (2022). “A Brief Introduction to Robust Bilevel Optimization.” In: *SIAG on Optimization Views and News* 30.2, pp. 1–10. URL: <https://siagoptimization.github.io/assets/views/ViewsAndNews-30-2.pdf>.
- (2023). “A survey on bilevel optimization under uncertainty.” In: *European Journal of Operational Research*. DOI: [10.1016/j.ejor.2023.01.008](https://doi.org/10.1016/j.ejor.2023.01.008).
- Beck, Y. and M. Schmidt (2021). “A robust approach for modeling limited observability in bilevel optimization.” In: *Operations Research Letters* 49.5, pp. 752–758. DOI: [10.1016/j.orl.2021.07.010](https://doi.org/10.1016/j.orl.2021.07.010).
- Ben-Ayed, O., C. Blair, D. Boyce, and L. LeBlanc (1992). “Construction of a real-world bilevel linear programming model of the highway network design problem.” In: *Annals of Operations Research* 34, pp. 219–254. DOI: [10.1007/BF02098181](https://doi.org/10.1007/BF02098181).
- Ben-Tal, A., L. Ghaoui, and A. Nemirovski (2009). *Robust Optimization*. DOI: [10.1515/9781400831050](https://doi.org/10.1515/9781400831050).
- Benders, J. F. (1962). “Partitioning procedures for solving mixed-variables programming problems.” In: *Numerische Mathematik* 4, pp. 238–252. DOI: [10.1007/BF01386316](https://doi.org/10.1007/BF01386316).

- Bertsimas, D., D. Brown, and C. Caramanis (2010). “Theory and Applications of Robust Optimization.” In: *SIAM Review* 53. DOI: [10.1137/080734510](https://doi.org/10.1137/080734510).
- Bertsimas, D. and M. Sim (2003). “Robust discrete optimization and network flows.” In: *Mathematical Programming* 98, pp. 49–71. DOI: [10.1007/s10107-003-0396-4](https://doi.org/10.1007/s10107-003-0396-4).
- (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).
- Besaçon, M., M. F. Anjos, and L. Brotcorne (2019). *Near-optimal Robust Bilevel Optimization*. URL: <https://arxiv.org/pdf/1908.04040.pdf>.
- Besaçon, M., M. F. Anjos, and L. Brotcorne (2021). “Complexity of near-optimal robust versions of multilevel optimization problems.” In: *Optimization Letters* 15, pp. 2597–2610. DOI: [10.1007/s11590-021-01754-9](https://doi.org/10.1007/s11590-021-01754-9).
- Birge, J. R. and F. Louveaux (2011). *Introduction to Stochastic Programming*. Springer-Verlag New York. DOI: [10.1007/978-1-4614-0237-4](https://doi.org/10.1007/978-1-4614-0237-4).
- Bolusani, S., S. Coniglio, T. K. Ralphs, and S. Tahernejad (2020). “A Unified Framework for Multistage Mixed Integer Linear Optimization.” In: *Bilevel Optimization: Advances and Next Challenges*. Ed. by S. Dempe and A. Zemkoho. Cham: Springer International Publishing, pp. 513–560. DOI: [10.1007/978-3-030-52119-6_18](https://doi.org/10.1007/978-3-030-52119-6_18).
- Brown, G., M. Carlyle, J. Salmerón, and R. Wood (2006). “Defending Critical Infrastructure.” In: *Interfaces* 36, pp. 530–544. DOI: [10.1287/inte.1060.0252](https://doi.org/10.1287/inte.1060.0252).
- Buchheim, C. and D. Henke (2020). *The bilevel continuous knapsack problem with uncertain follower’s objective*. URL: <https://arxiv.org/abs/1903.02810>.
- Buchheim, C., D. Henke, and F. Hommelsheim (2021). *On the Complexity of Robust Bilevel Optimization With Uncertain Follower’s Objective*. URL: <https://arxiv.org/abs/2105.08378>.
- Burtscheidt, J. and M. Claus (2020). “Bilevel Linear Optimization Under Uncertainty.” In: *Bilevel Optimization: Advances and Next Challenges*. Ed. by S. Dempe and A. Zemkoho. Springer International Publishing, pp. 485–511. DOI: [10.1007/978-3-030-52119-6_17](https://doi.org/10.1007/978-3-030-52119-6_17).
- Burtscheidt, J., M. Claus, and S. Dempe (2020). “Risk-Averse Models in Bilevel Stochastic Linear Programming.” In: *SIAM Journal on Optimization* 30.1, pp. 377–406. DOI: [10.1137/19M1242240](https://doi.org/10.1137/19M1242240).
- Caprara, A., M. Carvalho, A. Lodi, and G. J. Woeginger (2013). “A Complexity and Approximability Study of the Bilevel Knapsack Problem.” In: *Integer Programming and Combinatorial Optimization*. Ed. by M. Goemans and J. Correa. Vol. 7801. IPCO 2013. Springer, Berlin, Heidelberg, pp. 98–109. DOI: [10.1007/978-3-642-36694-9_9](https://doi.org/10.1007/978-3-642-36694-9_9).
- (2016). “Bilevel Knapsack with Interdiction Constraints.” In: *INFORMS Journal on Computing* 28.2, pp. 319–333. DOI: [10.1287/ijoc.2015.0676](https://doi.org/10.1287/ijoc.2015.0676).
- Chuong, T. D. and V. Jeyakumar (2017). “Finding Robust Global Optimal Values of Bilevel Polynomial Programs with Uncertain Linear Constraints.” In: *Journal of Optimization Theory and Applications* 173.2, pp. 683–703. DOI: [10.1007/s10957-017-1069-4](https://doi.org/10.1007/s10957-017-1069-4).
- Cormican, K. J., D. P. Morton, and R. K. Wood (1998). “Stochastic Network Interdiction.” In: *Operations Research* 46.2, pp. 184–197. DOI: [10.1287/opre.46.2.184](https://doi.org/10.1287/opre.46.2.184).
- Cornuéjols, G. (2008). “Valid inequalities for mixed integer linear programs.” In: *Mathematical Programming* 112, pp. 3–44. DOI: [10.1007/s10107-006-0086-0](https://doi.org/10.1007/s10107-006-0086-0).
- Della Croce, F. and R. Scatamacchia (2020). “An exact approach for the bilevel knapsack problem with interdiction constraints and extensions.” In: *Mathematical Programming* 183, pp. 249–281. DOI: [10.1007/s10107-020-01482-5](https://doi.org/10.1007/s10107-020-01482-5).
- Dempe, S. (2002). *Foundations of Bilevel Programming*. Springer US. DOI: [10.1007/b101970](https://doi.org/10.1007/b101970).
- Dempe, S. and A. B. Zemkoho (2012). “Bilevel road pricing: theoretical analysis and optimality conditions.” In: *Annals of Operations Research* 196, pp. 223–240. DOI: [10.1007/s10479-011-1023-z](https://doi.org/10.1007/s10479-011-1023-z).
- Dempe, S., S. Ivanov, and A. Naumov (2017). “Reduction of the bilevel stochastic optimization problem with quantile objective function to a mixed-integer problem.” In: *Applied Stochastic Models in Business and Industry* 33.5, pp. 544–554. DOI: [10.1002/asmb.2254](https://doi.org/10.1002/asmb.2254).

- DeNegre, S. T. (2011). “Interdiction and Discrete Bilevel Linear Programming.” PhD thesis. Lehigh University. URL: <https://coral.ise.lehigh.edu/~ted/files/papers/ScottDeNegreDissertation11.pdf>.
- DeNegre, S. T. and T. K. Ralphs (2009). “A Branch-and-cut Algorithm for Integer Bilevel Linear Programs.” In: *Operations research and cyber-infrastructure*. Springer, pp. 65–78. DOI: [10.1007/978-0-387-88843-9_4](https://doi.org/10.1007/978-0-387-88843-9_4).
- Fischetti, M., I. Ljubić, M. Monaci, and M. Sinnl (2016). “Intersection Cuts for Bilevel Optimization.” In: *Integer Programming and Combinatorial Optimization*. Ed. by Q. Louveaux and M. Skutella. IPCO 2016. Springer, pp. 77–88. DOI: [10.1007/978-3-319-33461-5_7](https://doi.org/10.1007/978-3-319-33461-5_7).
- (2017). “A New General-Purpose Algorithm for Mixed-Integer Bilevel Linear Programs.” In: *Operations Research* 65.6, pp. 1615–1637. DOI: [10.1287/opre.2017.1650](https://doi.org/10.1287/opre.2017.1650).
- (2018). “On the use of intersection cuts for bilevel optimization.” In: *Mathematical Programming* 172, pp. 77–103. DOI: [10.1007/s10107-017-1189-5](https://doi.org/10.1007/s10107-017-1189-5).
- (2019). “Interdiction Games and Monotonicity, with Application to Knapsack Problems.” In: *INFORMS Journal on Computing* 31.2, pp. 390–410. DOI: [10.1287/ijoc.2018.0831](https://doi.org/10.1287/ijoc.2018.0831).
- Fischetti, M., M. Monaci, and M. Sinnl (2018). “A dynamic reformulation heuristic for Generalized Interdiction Problems.” In: *European Journal of Operational Research* 267.16, pp. 40–51. DOI: [10.1016/j.ejor.2017.11.043](https://doi.org/10.1016/j.ejor.2017.11.043).
- Fortuny-Amat, J. and B. McCarl (1981). “A Representation and Economic Interpretation of a Two-Level Programming Problem.” In: *The Journal of the Operational Research Society* 32.9, pp. 783–792. DOI: [10.1057/jors.1981.156](https://doi.org/10.1057/jors.1981.156).
- Furini, F., I. Ljubić, E. Malaguti, and P. Paronuzzi (2020). “On Integer and Bilevel Formulations for the k -Vertex Cut Problem.” In: *Mathematical Programming Computation* 12, pp. 133–164. DOI: [10.1007/s12532-019-00167-1](https://doi.org/10.1007/s12532-019-00167-1).
- (2021). “Casting light on the hidden bilevel combinatorial structure of the capacitated vertex separator problem.” In: *Operations Research*. DOI: [10.1287/opre.2021.2110](https://doi.org/10.1287/opre.2021.2110).
- Furini, F., I. Ljubić, P. S. Segundo, and Y. Zhao (2021). “A branch-and-cut algorithm for the Edge Interdiction Clique Problem.” In: *European Journal of Operational Research* 294.1, pp. 54–69. DOI: [10.1016/j.ejor.2021.01.030](https://doi.org/10.1016/j.ejor.2021.01.030).
- Geoffrion, A. M. (1972). “Generalized Benders decomposition.” In: *Journal of Optimization Theory and Applications* 10, pp. 237–260. DOI: [10.1007/BF00934810](https://doi.org/10.1007/BF00934810).
- Goh, J. and M. Sim (2010). “Distributionally Robust Optimization and Its Tractable Approximations.” In: *Operations Research* 58.4-part-1, pp. 902–917. DOI: [10.1287/opre.1090.0795](https://doi.org/10.1287/opre.1090.0795).
- Golden, B. (1978). “A problem in network interdiction.” In: *Naval Research Logistics Quarterly* 25.4. DOI: [10.1002/nav.3800250412](https://doi.org/10.1002/nav.3800250412).
- Grimm, V., L. Schewe, M. Schmidt, and G. Zöttl (2019). “A Multilevel Model of the European Entry-Exit Gas Market.” In: *Mathematical Methods of Operations Research* 89.2, pp. 223–255. DOI: [10.1007/s00186-018-0647-z](https://doi.org/10.1007/s00186-018-0647-z).
- Haghighat, H. (2014). “Strategic offering under uncertainty in power markets.” In: *International Journal of Electrical Power & Energy Systems* 63, pp. 1070–1077. DOI: [10.1016/j.ijepes.2014.05.049](https://doi.org/10.1016/j.ijepes.2014.05.049).
- Hansen, P., B. Jaumard, and G. Savard (1992). “New branch-and-bound rules for linear bilevel programming.” In: *SIAM Journal on Scientific and Statistical Computing* 13.5, pp. 1194–1217. DOI: [10.1137/0913069](https://doi.org/10.1137/0913069).
- Israeli, E. (1999). “System Interdiction and Defense.” PhD thesis. Naval Postgraduate School Monterey, California. URL: <https://apps.dtic.mil/sti/pdfs/ADA361997.pdf>.
- Israeli, E. and R. K. Wood (2002). “Shortest-Path Network Interdiction.” In: *Networks* 40.2, pp. 97–111. DOI: [10.1002/net.10039](https://doi.org/10.1002/net.10039).
- Ivanov, S. (2018). “A Bilevel Stochastic Programming Problem with Random Parameters in the Follower’s Objective Function.” In: *Journal of Applied and Industrial Mathematics* 12.4, pp. 658–667. DOI: [10.1134/S1990478918040063](https://doi.org/10.1134/S1990478918040063).

- Kleinert, T., M. Labbé, I. Ljubić, and M. Schmidt (2021). “A Survey on Mixed-Integer Programming Techniques in Bilevel Optimization.” In: *EURO Journal on Computational Optimization*. DOI: [10.1016/j.ejco.2021.100007](https://doi.org/10.1016/j.ejco.2021.100007). Forthcoming.
- Labbé, M., P. Marcotte, and G. Savard (1998). “A Bilevel Model of Taxation and Its Application to Optimal Highway Pricing.” In: *Management Science* 44.12-part 1. DOI: [10.1287/mnsc.44.12.1608](https://doi.org/10.1287/mnsc.44.12.1608).
- Lee, T. and C. Kwon (2014). “A short note on the robust combinatorial optimization problems with cardinality constrained uncertainty.” In: *4OR* 12, pp. 373–378. DOI: [10.1007/s10288-014-0270-7](https://doi.org/10.1007/s10288-014-0270-7).
- Martello, S., D. Pisinger, and P. Toth (1999). “Dynamic Programming and Strong Bounds for the 0-1 Knapsack Problem.” In: *Management Science* 45.3. DOI: [10.1287/mnsc.45.3.414](https://doi.org/10.1287/mnsc.45.3.414).
- Migdalas, A. (1995). “Bilevel programming in traffic planning: Models, methods and challenge.” In: *Journal of Global Optimization* 7, pp. 381–405. DOI: [10.1007/BF01099649](https://doi.org/10.1007/BF01099649).
- Moore, J. T. and J. F. Bard (1990). “The mixed integer linear bilevel programming problem.” In: *Operations Research* 38.5, pp. 911–921. DOI: [10.1287/opre.38.5.911](https://doi.org/10.1287/opre.38.5.911).
- Pajouh, F. M. (2020). “Minimum cost edge blocker clique problem.” In: *Annals of Operations Research* 294, pp. 345–376. DOI: [10.1007/s10479-019-03315-x](https://doi.org/10.1007/s10479-019-03315-x).
- Pajouh, F. M., V. Boginski, and E. L. Pasiliao (2014). “Minimum vertex blocker clique problem.” In: *Networks* 64.1. DOI: [10.1002/net.21556](https://doi.org/10.1002/net.21556).
- Pajouh, F. M., J. L. Walteros, V. Boginski, and E. L. Pasiliao (2015). “Minimum edge blocker dominating set problem.” In: *European Journal of Operational Research* 247.1, pp. 16–26. DOI: [10.1016/j.ejor.2015.05.037](https://doi.org/10.1016/j.ejor.2015.05.037).
- Shi, X., O. Prokopyev, and T. K. Ralphs (2020). *Mixed Integer Bilevel Optimization with k-optimal Follower: A Hierarchy of Bounds*. Tech. rep. URL: http://www.optimization-online.org/DB_FILE/2020/06/7874.pdf.
- Sim, M. (2004). “Robust Optimization.” PhD thesis. Massachusetts Institute of Technology, Sloan School of Management. URL: <https://dspace.mit.edu/handle/1721.1/17725>.
- Smith, J. C. and Y. Song (2020). “A survey of network interdiction models and algorithms.” In: *European Journal of Operational Research* 283.3, pp. 797–811. DOI: [10.1016/j.ejor.2019.06.024](https://doi.org/10.1016/j.ejor.2019.06.024).
- Soyster, A. L. (1973). “Technical Note—Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming.” In: *Operations Research* 21.5, pp. 1154–1157. DOI: [10.1287/opre.21.5.1154](https://doi.org/10.1287/opre.21.5.1154).
- Tahernejad, S. and T. K. Ralphs (2020). *Valid Inequalities for Mixed Integer Bilevel Linear Optimization Problems*. Tech. rep. URL: https://engineering.lehigh.edu/sites/engineering.lehigh.edu/files/_DEPARTMENTS/ise/pdf/tech-papers/20/20T_013.pdf.
- Tahernejad, S., T. K. Ralphs, and S. T. DeNegre (2020). “A branch-and-cut algorithm for mixed integer bilevel linear optimization problems and its implementation.” In: *Mathematical Programming Computation* 12, pp. 529–568. DOI: [10.1007/s12532-020-00183-6](https://doi.org/10.1007/s12532-020-00183-6).
- Tang, Y., J. P. Richard, and J. C. Smith (2016). “A class of algorithms for mixed-integer bilevel min–max optimization.” In: *Journal of Global Optimization* 66, pp. 225–262. DOI: [10.1007/s10898-015-0274-7](https://doi.org/10.1007/s10898-015-0274-7).
- Wood, R. K. (2011). “Bilevel Network Interdiction Models: Formulations and Solutions.” In: *Wiley Encyclopedia of Operations Research and Management Science*. DOI: [10.1002/9780470400531.eorms0932](https://doi.org/10.1002/9780470400531.eorms0932).
- Xu, P. and L. Wang (2014). “An exact algorithm for the bilevel mixed integer linear programming problem under three simplifying assumptions.” In: *Computers & Operations Research* 41, pp. 309–318. DOI: [10.1016/j.cor.2013.07.016](https://doi.org/10.1016/j.cor.2013.07.016).
- Yamkoğlu, İ. and D. Kuhn (2018). “Decision Rule Bounds for Two-Stage Stochastic Bilevel Programs.” In: *SIAM Journal on Optimization* 28.1, pp. 198–222. DOI: [10.1137/16M1098486](https://doi.org/10.1137/16M1098486).
- Zare, M. H., O. A. Prokopyev, and D. Sauré (2020). “On Bilevel Optimization with Inexact Follower.” In: *Decision Analysis* 17.1. DOI: [10.1287/deca.2019.0392](https://doi.org/10.1287/deca.2019.0392).

- Zeng, B., H. Dong, R. Sioshansi, F. Xu, and M. Zeng (2020). “Bilevel Robust Optimization of Electric Vehicle Charging Stations With Distributed Energy Resources.” In: *IEEE Transactions on Industry Applications* 56.5, pp. 5836–5847. DOI: [10.1109/TIA.2020.2984741](https://doi.org/10.1109/TIA.2020.2984741).
- Zenklusen, R., B. Ries, C. Picouveau, D. de Werra, M.-C. Costa, and C. Bentz (2009). “Blockers and transversals.” In: *Discrete Mathematics* 309.13, pp. 4306–4314. DOI: [10.1016/j.disc.2009.01.006](https://doi.org/10.1016/j.disc.2009.01.006).

APPENDIX A. OMITTED PROOFS

Proof of Proposition 1. Let x be a feasible upper-level decision and let (y^*, z^*, θ^*) be an optimal solution of the x -parameterized lower-level problem (7). By Assumption 2, there is a vector $u \in \mathbb{R}_+^{n_y}$ such that $Y(x) \subseteq [0, u_1] \times \cdots \times [0, u_{n_y}]$ and, in particular, $0 \leq y_i^* \leq u_i$ holds for all $i \in [n_y]$. Consequently, we obtain

$$d^\top y^* \geq \sum_{i=1}^{n_y} \min\{d_i, 0\} u_i.$$

By (7b), $\theta^* \geq 0$, and $y^* \leq u$, an optimal solution of the follower further satisfies

$$z_i^* = \max\{\Delta d_i y_i^* - \theta^*, 0\} \leq \Delta d_i y_i^* \leq \Delta d_i u_i$$

for all $i \in [n_y]$. Finally, we have $\theta^* \leq \max\{\Delta d_i u_i : i \in [n_y]\}$ due to (7b) and the optimality of θ^* , which concludes the proof. \square

Proof of Proposition 2. This follows immediately from the first part of the proof of Proposition 1. \square

Proof of Proposition 3. Let x be a feasible upper-level decision and let $\ell \in \{1, \dots, n_y + 1\}$ be arbitrary but fixed. Further, let y^ℓ be an optimal solution of the ℓ th sub-problem (10) that is parameterized in x . By (9) and since $y^\ell \in \{0, 1\}^{n_y}$, we obtain

$$\Phi_d(x) \geq -\Gamma_d \Delta d_\ell + \sum_{i=1}^{n_y} \tilde{d}(\ell)_i y_i^\ell \geq -\Gamma_d \Delta d_\ell + \sum_{i=1}^{n_y} \min\{\tilde{d}(\ell)_i, 0\},$$

which concludes the proof. \square

Proof of Proposition 4. Let x be a feasible upper-level decision. Further, let $y \in Y(x)$ and let $y' \in \mathcal{Y}$ be such that $y' \leq y$ holds. Due to $B \in \mathbb{R}_+^{m \times n}$, we obtain

$$\begin{aligned} B y' &\leq B y \leq b, \\ y'_i &\leq y_i \leq u_i(1 - x_i), \quad i \in [n], \end{aligned}$$

Thus, y' is a feasible follower’s decision for the given x , i.e., $y' \in Y(x)$. \square

Proof of Proposition 5. Let x be a feasible upper-level decision. First, we show that the extended formulation (14) satisfies the monotonicity property. To this end, let (y, z, θ) be a feasible follower’s decision for Problem (14) for the given x . Further, let $y' \in \mathcal{Y}$ be such that $y' \leq y$ holds. Due to $B \in \mathbb{R}_+^{m \times n}$ and $\Delta d_i \geq 0$ for all $i \in [n]$, we obtain

$$\begin{aligned} z_i + \theta &\geq \Delta d_i y_i \geq \Delta d_i y'_i, \quad i \in [n], \\ B y' &\leq B y \leq b, \\ y'_i &\leq y_i \leq u_i(1 - x_i), \quad i \in [n], \end{aligned}$$

i.e., the follower’s decision (y', z, θ) is feasible for Problem (14) for the given x . Second, we show that each sub-problem (15) satisfies the monotonicity property. Note that there is no need to specify $\ell \in \{\Gamma_d, \dots, n + 1\}$ since the feasible set of (15) does not depend on ℓ . For the given x , let y be a feasible follower’s decision for sub-problem (15). Further, let $y' \in \{0, 1\}^n$ be such that $y' \leq y$. Since we restrict ourselves to binary follower’s variables in the multi-follower case, we have valid upper bounds $u_i = 1$ for all $i \in [n]$. Applying the

same arguments as before, y' is feasible for sub-problem (15) for the given x . Consequently, the Γ_d -robust lower-level problems (14) and (15) satisfy the monotonicity property. \square

Proof of Proposition 6. Let x be a feasible upper-level decision. For notational convenience, let $\Psi(x)$ and Ψ denote the feasible set of Problem (14) and (16), respectively. Further, let (y^*, z^*, θ^*) be an optimal solution of Problem (14) for the given leader's decision x . Then, (y^*, z^*, θ^*) is also feasible for Problem (16), i.e., $(y^*, z^*, \theta^*) \in \Psi$. In particular, $y_i^* x_i = 0$ holds for all $i \in [n]$, i.e., both problems have the same objective function value for (y^*, z^*, θ^*) . Thus, we obtain

$$\begin{aligned} & \max_{y, \theta, z} \left\{ \sum_{i=1}^n d_i y_i - \Gamma_d \theta - \sum_{i=1}^n z_i : (y, z, \theta) \in \Psi(x) \right\} \\ & \leq \max_{y, \theta, z} \left\{ \sum_{i=1}^n d_i y_i (1 - x_i) - \Gamma_d \theta - \sum_{i=1}^n z_i : (y, z, \theta) \in \Psi \right\}. \end{aligned} \quad (30)$$

Let $(\hat{y}, \hat{z}, \hat{\theta})$ be an optimal solution of Problem (16) for the given leader's decision x . Without loss of generality, we assume that there is exactly one item $k \in [n]$ for which the interdiction constraint $\hat{y}_k \leq u_k(1 - x_k)$ is not satisfied, i.e., $\hat{y}_k \geq 1$ and $x_k = 1$. Otherwise, we repeat the following as long as there are no more items left that violate the interdiction constraint. We consider the alternative follower's decision $(y', z', \hat{\theta})$ with

$$y'_i = \begin{cases} \hat{y}_i, & i \in [n] \setminus \{k\}, \\ 0, & i = k, \end{cases}$$

and

$$z'_i = \begin{cases} \hat{z}_i, & i \in [n] \setminus \{k\}, \\ 0, & i = k. \end{cases}$$

By construction, $(y', z', \hat{\theta})$ is feasible for Problem (16) and satisfies all interdiction constraints. Moreover, we obtain

$$\begin{aligned} & \max_{y, \theta, z} \left\{ \sum_{i=1}^n d_i y_i (1 - x_i) - \Gamma_d \theta - \sum_{i=1}^n z_i : (y, z, \theta) \in \Psi \right\} \\ & = d_k \hat{y}_k \underbrace{(1 - x_k)}_{=0} + \sum_{i \in [n] \setminus \{k\}} d_i \underbrace{\hat{y}_i}_{=y'_i} (1 - x_i) - \Gamma_d \hat{\theta} - \hat{z}_k - \sum_{i \in [n] \setminus \{k\}} \underbrace{\hat{z}_i}_{=z'_i} \\ & = d_k \underbrace{y'_k}_{=0} (1 - x_k) + \sum_{i \in [n] \setminus \{k\}} d_i y'_i (1 - x_i) - \Gamma_d \hat{\theta} - \underbrace{z'_k}_{=0} - \hat{z}_k - \sum_{i \in [n] \setminus \{k\}} z'_i \\ & \leq \sum_{i=1}^n d_i y'_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i=1}^n z'_i, \end{aligned}$$

i.e., the alternative follower's decision is optimal for Problem (16). In particular, it is also feasible for Problem (14), i.e., $(y', z', \hat{\theta}) \in \Psi(x)$, and we have $y'_i x_i = 0$ for all $i \in [n]$. Hence,

we obtain

$$\begin{aligned}
& \max_{y, \theta, z} \left\{ \sum_{i=1}^n d_i y_i (1 - x_i) - \Gamma_d \theta - \sum_{i=1}^n z_i : (y, z, \theta) \in \Psi \right\} \\
&= \sum_{i=1}^n d_i y'_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i=1}^n z'_i \\
&= \sum_{i=1}^n d_i y'_i - \Gamma_d \hat{\theta} - \sum_{i=1}^n z'_i \\
&\leq \max_{y, \theta, z} \left\{ \sum_{i=1}^n d_i y_i - \Gamma_d \theta - \sum_{i=1}^n z_i : (y, z, \theta) \in \Psi(x) \right\}.
\end{aligned} \tag{31}$$

Due to (30) and (31), Problem (14) and (16) admit the same optimal value. \square

Proof of Proposition 7. Let x be a feasible upper-level decision and let $\ell \in \{\Gamma_d, \dots, n+1\}$ be arbitrary but fixed. Further, let y^* be an optimal solution of the ℓ th sub-problem (15) for the given leader's decision x . Then, y^* is also feasible for the ℓ th sub-problem (17), i.e., $y^* \in Y$. In particular, $y_i^* x_i = 0$ holds for all $i \in [n]$, i.e., both sub-problems have the same objective function value for y^* . Thus, we obtain

$$\max_y \left\{ \sum_{i=1}^n \tilde{d}(\ell)_i y_i : y \in Y(x) \right\} \leq \max_y \left\{ \sum_{i=1}^n \tilde{d}(\ell)_i y_i (1 - x_i) : y \in Y \right\}. \tag{32}$$

Let \hat{y} be an optimal solution of the ℓ th sub-problem (17) for the given leader's decision x . Without loss of generality, suppose there is exactly one item $k \in [n]$ for which the interdiction constraint $\hat{y}_k \leq 1 - x_k$ is not satisfied, i.e., $\hat{y}_k = 1 = x_k$. Then, we consider the alternative follower's decision

$$y'_i = \begin{cases} \hat{y}_i, & i \in [n] \setminus \{k\}, \\ 0, & i = k. \end{cases}$$

By construction, y' is feasible for the ℓ th sub-problem (17) and satisfies all interdiction constraints. Moreover, we obtain

$$\begin{aligned}
\max_y \left\{ \sum_{i=1}^n \tilde{d}(\ell)_i y_i (1 - x_i) : y \in Y \right\} &= \tilde{d}(\ell)_k \underbrace{\hat{y}_k (1 - x_k)}_{=0} + \sum_{i \in [n] \setminus \{k\}} \tilde{d}(\ell)_i \underbrace{\hat{y}_i (1 - x_i)}_{=y'_i} \\
&= \tilde{d}(\ell)_k \underbrace{y'_k (1 - x_k)}_{=0} + \sum_{i \in [n] \setminus \{k\}} \tilde{d}(\ell)_i y'_i (1 - x_i) \\
&= \sum_{i=1}^n \tilde{d}(\ell)_i y'_i (1 - x_i),
\end{aligned}$$

i.e., the alternative follower's decision is optimal for Problem (17). In particular, it is also feasible for Problem (15), i.e., $y' \in Y(x)$, and we have $y'_i x_i = 0$ for all $i \in [n]$. Hence, we obtain

$$\begin{aligned}
\max_y \left\{ \sum_{i=1}^n \tilde{d}(\ell)_i y_i (1 - x_i) : y \in Y \right\} &= \sum_{i=1}^n \tilde{d}(\ell)_i y'_i (1 - x_i) \\
&= \sum_{i=1}^n \tilde{d}(\ell)_i y'_i \\
&\leq \max_y \left\{ \sum_{i=1}^n \tilde{d}(\ell)_i y_i : y \in Y(x) \right\}.
\end{aligned} \tag{33}$$

Due to (32) and (33), Problem (15) and (17) admit the same optimal value. \square

Proof of Theorem 2. Let $(x, \eta) \in X \times \mathbb{R}$ be a given leader's decision. Due to the validity of the proposed cuts, it suffices to show that the feasibility of (x, η) w.r.t. either the interdiction cuts (18) or (19) implies $\eta \geq \Phi(x)$. To this end, suppose that (x, η) satisfies $Ax \geq a$ and either the interdiction cuts

$$\eta \geq \sum_{i=1}^n d_i \hat{y}_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i=1}^n \hat{z}_i \quad \text{for all } (\hat{y}, \hat{z}, \hat{\theta}) \in \hat{\Psi}$$

or

$$\eta \geq -\Gamma_d \Delta d_\ell + \sum_{i=1}^n \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) \quad \text{for all } \hat{y} \in \hat{Y}, \ell \in \{\Gamma_d, \dots, n+1\}.$$

By Propositions 6 and 7, this is equivalent to

$$\eta \geq \max_{\hat{y}, \hat{\theta}, \hat{z}} \left\{ \sum_{i=1}^n d_i \hat{y}_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i=1}^n \hat{z}_i : (\hat{y}, \hat{z}, \hat{\theta}) \in \hat{\Psi} \right\} = \Phi(x)$$

and

$$\eta \geq \max_{\ell \in \{\Gamma_d, \dots, n+1\}} \left\{ -\Gamma_d \Delta d_\ell + \max_{\hat{y} \in \hat{Y}} \left\{ \sum_{i=1}^n \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) \right\} \right\} = \Phi(x),$$

which concludes the proof. \square

Proof of Proposition 8. This follows immediately from $d_i > 0$ and from the fact that $x_i \in \{0, 1\}$ implies $\hat{y}_i (1 - x_i) \leq y'_i (1 - x_i)$ for all $i \in [n]$. \square

Proof of Proposition 9. Let (x, η) be feasible for Problem (2) with the lower-level optimal-value function (3). Further, let $\ell \in \{\Gamma_d, \dots, n+1\}$ be arbitrary but fixed. Due to $x_i \in \{0, 1\}$ for all $i \in [n]$, we have $\tilde{d}(\ell)_i (1 - x_i) \leq 0$ for all $i \notin D_+^\ell$. Hence, all follower's variables y_i with $i \notin D_+^\ell$ could be omitted in this sub-problem, i.e., we obtain

$$\Phi^\ell(x) = \max_y \left\{ \sum_{i \in D_+^\ell} \tilde{d}(\ell)_i y_i (1 - x_i) : y \in \hat{Y} \right\}.$$

The validity of the new interdiction cuts (20) can be shown in a similar way as it is done in Section 3. In particular,

$$\begin{aligned} & -\Gamma_d \Delta d_\ell + \sum_{i=1}^n \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) \\ &= -\Gamma_d \Delta d_\ell + \sum_{i \in D_+^\ell} \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) + \sum_{i \notin D_+^\ell} \underbrace{\tilde{d}(\ell)_i \hat{y}_i (1 - x_i)}_{\leq 0} \\ &\leq -\Gamma_d \Delta d_\ell + \sum_{i \in D_+^\ell} \tilde{d}(\ell)_i \hat{y}_i (1 - x_i) \end{aligned}$$

holds for all $\hat{y} \in \hat{Y}$ and $\ell \in \{\Gamma_d, \dots, n+1\}$. Thus, the cuts (20) dominate the basic interdiction cuts (19). \square

Proof of Proposition 10. This can be shown in analogy to the proof of Proposition 8.

(Y. Beck, M. Schmidt) TRIER UNIVERSITY, DEPARTMENT OF MATHEMATICS, UNIVERSITÄTSRING 15, 54296 TRIER, GERMANY

Email address: yasmine.beck@uni-trier.de

Email address: martin.schmidt@uni-trier.de

(I. Ljubić) ESSEC BUSINESS SCHOOL OF PARIS, CERGY-PONTOISE, FRANCE

Email address: ljubic@essec.edu