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 mixed-integer 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 due to some kind of bounded rationality. We study mixed-integer 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-scenario 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 mixed-integer linear bilevel problems of the form

\[
\begin{align*}
\min_x & \quad c^\top x + d^\top y \\
\text{s.t.} & \quad Ax \geq a, \\
& \quad x \in X \subseteq \mathbb{Z}^{n_x}, \\
& \quad y \in \arg \max_{y'} \left\{ d^\top y' : y' \in Y(x) \subseteq \mathbb{Z}^{n_y} \right\},
\end{align*}
\]

where \( Y(x) \) denotes the lower-level feasible set that is parameterized by the leader’s variables \( x \). Without loss of generality, we assume that all variables \( x \) are discrete (Bolusani et al. 2020). 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 a crucial assumption 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 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 the easiest instantiations of bilevel problems, namely linear bilevel problems, are strongly NP-hard; 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).

In what follows, we say that an upper-level decision \( x \) is feasible if \( x \in X \) and \( Ax \geq a \) are satisfied. We assume that \( Y(x) \neq \emptyset \) holds for all feasible leader’s decisions \( x \) and that the shared constraint set
\[
\{(x, y) : Ax \geq a, x \in X, y \in Y(x)\}
\]
is non-empty and compact. Moreover, we assume that all linking variables, i.e., all variables of the leader that appear in the lower-level constraints, are bounded integers. Thus, 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
\[
\begin{align*}
\min_{x, \eta} & \quad c^\top x + \eta \\
\text{s.t.} & \quad Ax \geq a, \quad (2a) \\
& \quad x \in X, \quad (2b) \\
& \quad \eta \geq \Phi(x), \quad (2c) \\
& \quad \eta \in \mathbb{R}. \quad (2d)
\end{align*}
\]
Up to this point, we implicitly made the assumption that both players act perfectly rational. In many real-world situations, however, this assumption is not valid since decision makers may face some kind of bounded rationality; see, e.g., Simon (1972). For applications that demonstrate the relevance of considering bounded rationality in practice, we refer the reader to Pita, Jain, Ordóñez, et al. (2009), Pita, Jain, Tambe, et al. (2010), and Pita, Portway, et al. (2008) and the references therein as well as to the survey by Smith and Song (2020). An elaboration on how decision makers are confronted with cognitive limitations preventing them from reaching a perfectly rational decision can be found in, e.g., Chariri (2017). In principle, both the leader and the follower may be affected by bounded rationality. In the literature, however, one typically focuses on the consideration of bounded rationality for the lower-level player since this is the more challenging variant from a mathematical point of view. A detailed survey on existing methods to account for bounded rationality in bilevel optimization can be found in the recent survey by Beck, Ljubić, et al. (2022). For a general discussion on bounded rationality, we refer to Rubinstein (1998). In this paper, we consider data uncertainty as one important aspect of bounded rationality.

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 or by considering chance constraints. 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 $\Gamma$-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 Yankïlo and Kuhn (2018). To the best of our knowledge, robust approaches to address bounded rationality in bilevel optimization have been much less investigated. In the context of power markets, a $\Gamma$-robust approach to deal with uncertain lower-level data is considered in Haghighat (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 Sariddichainunta and Inuiguchi (2017) and Zeng et al. (2020), worst-case oriented approaches for bilevel problems with lower-level data uncertainty are addressed. In Buchheim and Henke (2020) and Buchheim, Henke, and Hommelsheim (2021), complexity results for robust bilevel problems with uncertainties regarding the lower-level objective function coefficients are established. 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 $\Gamma$th least damaging choices of the solution algorithm for the lower-level problem, which is to some extent related to the notion of $\Gamma$-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. For a general discussion on bilevel optimization under uncertainty we refer to the recent survey by Beck, Ljubić, et al. (2022).

The contributions of this paper are the following. We study mixed-integer 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 $d_i^\Gamma$, where $d_i^\Gamma \in [d_i - \Delta d_i, d_i]$ for all $i \in \{n_y := \{1, \ldots, n_y\}\}$. We denote $d_i$ as the nominal value of the $i$th coefficient of the lower-level's objective function and $\Delta d_i$ is its maximum deviation from the nominal value. For a feasible upper-level
To model the two types of situations, we consider an approach using an extended formulation. Additionally, we present an approach using a multi-scenario formulation for the special case in which all lower-level variables are binary, i.e., \( Y(x) \subseteq \{0,1\}^{n_y} \). However, the scenario-based approach can be extended naturally to allow for additional non-binary follower’s 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 the follower’s 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 the \( \Gamma \)-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 1–4 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-scenario formulation to model mixed-integer linear bilevel problems with a \( \Gamma \)-robust follower 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 \( \Gamma \)-robustification of Problem (2). 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 Constraint (2d) are omitted. This means that we first consider the linear problem

\[
\min_{x,\eta} \ c^\top x + \eta \\
\text{s.t. } (x,\eta) \in \Omega_0 := \{(x,\eta) \in \mathbb{R}^{n_x} \times \mathbb{R} : Ax \geq a, x \in \bar{X}, \eta \geq \eta^-\}. 
\]  

(P_0)
Here, \( \hat{X} \) is the continuous relaxation of \( X \) and we also use an a priori lower bound on \( \Phi(x) \) for all feasible leader’s decisions \( x \). A trivial lower bound is, e.g., given by

\[
\eta^- := \sum_{i=1}^{n_y} \min\{d_i, 0\}.
\]

We then iteratively add valid inequalities or branch to cut off integer-infeasible points and also add valid inequalities to cut off bilevel infeasible points. Let

\[
\min_{x,\eta} c^T x + \eta \quad \text{s.t.} \quad (x, \eta) \in \Omega_j \subseteq \mathbb{R}^{n_x} \times \mathbb{R}
\]

be the problem of node \( j \) of the branch-and-cut search tree. Here, the set \( \Omega_j \) contains all valid inequalities that have been added previously to cut off integer-infeasible and bilevel-infeasible points. If either Problem \( (P_j) \) is infeasible or, if the objective function value corresponding to an optimal solution \( (x^j, \eta^j) \) 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 \( (2d) \) 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-scenario 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 \( (4) \) 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

\[
\Phi_d(x) = \max_{y,z,\theta} \sum_{i=1}^{n_x} d_i y_i - \Gamma d \theta - \sum_{i=1}^{n_x} z_i \quad (5a)
\]

\[
s.t. \quad z_i + \theta \geq \Delta d_i y_i, \quad i \in [n_y], \quad (5b)
\]

\[
(z, \theta) \geq 0, \quad y \in Y(x). \quad (5c)
\]

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 (4) can be solved as the mixed-integer linear problem

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

s.t. $\sum_{i=1}^{n_y} \nu_i y_i + \Gamma_w \theta + \sum_{i=1}^{n_y} z_i \leq C - \sum_{i=1}^{n_r} v_i x_i,$

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

$$(z, \theta) \geq 0,$$

$y \in Y(x).$

Both lemmas can be shown in analogy to the proof of Theorem 1 in Bertsimas and Sim (2003).

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 (5) or Problem (6) 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 $(x^j, \eta^i)$. To this end, one can use generic cuts like (generalized) no-good cuts; see, e.g., Tahernejad and Ralphs (2020).

Algorithm 1 Processing node $j$ using the extended formulation

1: Solve Problem ($P_j$).
2: if Problem ($P_j$) is infeasible then
3: Fathom the current node.
4: else Let $(x^j, \eta^i)$ denote the solution of Problem ($P_j$).
5: if $c^\top x^j + \eta^i \geq U$ 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 $\Omega_j$, 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 $\eta^i < \Phi(x^j)$ then
11: Generate a valid cut that excludes $(x^j, \eta^i)$ from $\Omega_j$, augment $\Omega_j$, and go to Step 1.

2.2. Multi-Scenario Approach. An alternative reformulation of the robust counterparts (3) and (4) can be obtained under the following additional assumptions.

Assumption 2. All lower-level variables $y$ are binary, i.e., $Y(x) \subseteq \{0, 1\}^{n_y}$.

Assumption 3. 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 $\Delta w_{n_y+1} = 0$.

Assumptions 2 and 3 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 2 and 3, solving the robust counterpart of the lower-level problem (3) is equivalent to solving $n_y + 1$ problems of the nominal type, i.e.,

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

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

$$\Phi_d^\ell(x) = -\Gamma_d \Delta d_\ell + \max_{y \in Y(x)} \{\hat{d}(\ell)^\top y\}$$
with
\[ \tilde{d}(\ell)_i = \begin{cases} d_i - (\Delta d_i - \Delta d_\ell), & 1 \leq i \leq \ell, \\ d_i, & \ell + 1 \leq i \leq n_y. \end{cases} \]

We refer to (7) as the multi-scenario formulation. Note, however, that this reformulation does not depend on scenarios in the traditional sense in which a finite number of scenarios is used for a scenario-based description of the uncertain problem data. Throughout this paper, a scenario \( \ell \in \{1, \ldots, n_y + 1\} \) instead refers to the corresponding sub-problem (8).

The multi-scenario 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 2 and 3, solving the robust counterpart of the lower-level problem (4) is equivalent to solving \( n_y + 1 \) problems of the nominal type, i.e.,

\[ \Phi_w(x) = \max_{\ell \in \{1, \ldots, n_y + 1\}} \Phi^\ell_w(x) \]

holds, where for all \( \ell \in \{1, \ldots, n_y + 1\} \), we have

\[ \Phi^\ell_w(x) = \max_{y \in Y(x)} \left\{ \sum_{i=1}^{n_y} d_i y_i : \Gamma_w \Delta w_\ell + \sum_{i=1}^{n_y} \tilde{w}(\ell)_i y_i + \sum_{i=1}^{n_y} v_i x_i \leq C \right\} \]

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

Lemma 3 can be shown in analogy to the proof of Theorem 3 in Bertsimas and Sim (2003), whereas a proof of Lemma 4 could follow the one of Lemma 2 in Álvarez-Miranda, Ljubić, et al. (2013).

Note that, in the case of uncertainties regarding the follower’s inequality constraint, we consider the same deterministic lower-level objective function in the extended formulation as well as in the multi-scenario formulation. Further, 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-scenario 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.,

\[ \Phi(x) = \max_{\ell \in \{\Gamma, \ldots, n_y + 1\}} \Phi^\ell(x). \]

Note that the lower-level optimal-value function (10) is defined as the maximum of \( n_y - \Gamma + 2 \) value functions. Thus, the robustification of Problem (1) can be interpreted as a single-leader-multi-follower problem with \( n_y - \Gamma + 2 \) many followers.

The method to process node \( j \) that exploits the multi-scenario 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 scenario cut for each lower-level sub-problem \( \ell \in \{\Gamma, \ldots, n_y + 1\} \) that satisfies \( \eta_j < \Phi^\ell(x^j) \) is added in Step 12 of Algorithm 2. This means that up to \( n_y - \Gamma + 2 \) scenario cuts could be added at each node. However, it would also be valid to consider, e.g., adding only the most violated scenario 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.

**Remark 1.** If we embed either Algorithm 1 or 2 into a usual branch-and-bound framework, we obtain a correct method that terminates after a finite number of iterations with an optimal solution \( (x^*, \eta^*) \). We first recall that all linking variables are bounded integers. We then
Thus, an optimal solution cannot be overlooked. In particular, the number of cuts possibly occur again, i.e., if there would exist solutions, w.l.o.g., non-linking variables can be moved to the lower level; see also Bolusani et al. (2020). Hence, the convergence is due to the finiteness of the number of feasible upper-level decisions and from the fact that a leader’s decision cannot be selected twice. If \( x, \eta \) is a non-optimal leader’s decision, i.e., \( \eta < \Phi(x) \), adding a globally valid inequality excludes \( (x, \eta) \) 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_\ell) 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 \( O(2^{n_y}) \).

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 \( \ell \)th sub-problem with \( \ell \in \{1, \ldots, n_y + 1\} \). 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 \( \Gamma \)-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 restrict ourselves to the case in which \( n_y = n_x = n \) 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). Moreover, we impose that the

---

**Algorithm 2** Processing node \( j \) using the multi-scenario approach

1. Solve Problem (P_\ell).
2. If Problem (P_\ell) is infeasible then
3. Fathom the current node.
4. Let \((x^j, \eta^j)\) denote the solution of Problem (P_\ell).
5. If \( c^j x^j + \eta^j \geq U \) 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 \( \Omega_j \), and go to Step 1 or branch.
9. For all \( \ell \in \{1, \ldots, n_y + 1\} \) do
10. Solve the \( \ell \)th lower-level sub-problem to obtain \( \Phi^\ell(x^j) \).
11. If \( \eta^j < \Phi^\ell(x^j) \) then
12. Generate a valid scenario cut that excludes \((x^j, \eta^j)\) from \( \Omega_j \), augment \( \Omega_j \).
13. Set \( \Phi(x^j) = \max_{x \in [\Gamma, \ldots, n_y + 1]} \Phi^\ell(x^j) \) and \( U = \min\{U, c^j x^j + \Phi(x^j)\} \).
14. If at least one cut was added in Step 12 then go to Step 1.
\( x \)-parameterized lower-level feasible set is of the form
\[
Y(x) = \{ y \in \mathcal{Y} : By \leq b, \ y_i \leq u_i(1-x_i), \ i \in [n] \}
\] (11)
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 \( u \in \mathbb{R}^n_+ \). Note that the leader’s variables \( x \) are linked to the lower-level problem only via the interdiction constraints \( y_i \leq u_i(1-x_i) \), which is crucial for the validity of the cuts we propose in the following. Finally, we impose that there are no terms depending on the leader’s decision in the upper-level objective function, i.e., \( c = 0 \), which is reasonable in the interdiction setting. We will hold on to the aforementioned assumptions for the remainder of the paper. In this setting, the nominal lower-level problem (1d) satisfies the following monotonicity property.

**Proposition 1** (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., \( y' \in Y(x) \).

All the proofs that we omit here can be found in Appendix A. Further, we assume that all lower-level objective function coefficients are positive, i.e., \( d_i > 0 \) for all \( i \in [n] \). This is w.l.o.g. since 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.

In the following sections, we show that the monotonicity property remains satisfied when \( \Gamma \)-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 valid cuts—interdiction cuts and scenario interdiction cuts—based on the two approaches discussed in Section 2. We devote Section 3.2 to obtain 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 \( \Gamma_d \) deviations in the objective function coefficients. The robust counterpart of the lower-level problem is given in (3). The corresponding extended formulation and the multi-scenario formulation have already been stated in Section 2. For the extended formulation, we only need to replace (5d) with (11), i.e., we consider the problem

\[
\Phi(x) = \max_{y,z,\theta} \sum_{i=1}^n d_i y_i - \Gamma_d \theta - \sum_{i=1}^n z_i
\]
\[
\text{s.t. } z_i + \theta \geq \Delta d_i y_i, \quad i \in [n],
\]
\[
By \leq b,
\]
\[
y_i \leq u_i(1-x_i), \quad i \in [n],
\]
\[
(z,\theta) \geq 0, \ y \in \mathcal{Y}.
\]

For the multi-scenario approach, we need to replace the feasible set in (8) with (11). Due to Assumption 2, \( u_j = 1 \) is a trivially valid upper bound for all \( i \in [n] \). Thus, we consider the scenario sub-problems

\[
\Phi'(x) = -\Gamma_d \Delta d_e + \max_{y} \left\{ \delta(\ell)^\top y : By \leq b, \ y_i \leq 1-x_i, \ y_i \in \{0,1\}, \ i \in [n] \right\}.
\]

**Proposition 2.** Suppose that Assumption 1 holds. Then, Problem (13) satisfies the monotonicity property. Moreover, let \( x \) be a feasible decision of the leader, let \( (y,z,\theta) \) be feasible for the \( x \)-parameterized problem (12), and let \( y' \in \mathcal{Y} \) be such that \( y' \leq y \) holds. Then, \( (y',z,\theta) \) is feasible for the \( x \)-parameterized problem (12) as well.
Due to Proposition 2 and for the ease of presentation, we also say that the extended formulation (12) satisfies the monotonicity property. In what follows, we exploit the previous result to introduce penalized formulations for Problems (12) and (13) 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 3.** Let \( x \) be a feasible upper-level decision. Under Assumption 1, Problem (12) and the mixed-integer linear problem

\[
\Phi(x) = \max_{y,z,\theta} \sum_{i=1}^{n} d_i y_i (1 - x_i) - \Gamma d \theta - \sum_{i=1}^{n} z_i \\
\text{s.t.} \quad z_i + \theta \geq \Delta d_i y_i, \quad i \in [n], \\
By \leq b, \\
y_i \leq u_i, \quad i \in [n], \\
(z,\theta) \geq 0, \quad y \in \mathcal{Y}
\]  

admit the same optimal value.

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

**Proposition 4.** Let \( x \) be a feasible upper-level decision and let \( \ell \in \{\Gamma d, \ldots, n + 1\} \) be arbitrary but fixed. Then, Problem (13) and the problem

\[
\Phi^\ell(x) = -\Gamma d \Delta d_\ell + \max_{y \in \hat{Y}} \left\{ \sum_{i=1}^{n} \tilde{d}(\ell)_i y_i (1 - x_i) \right\}
\]

with

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

admit the same optimal value.

The feasible set of Problem (14) and of each scenario sub-problem (15) 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

\[
\Psi = \{ (y, z, \theta) \in \mathcal{Y} \times \mathbb{R}^n \times \mathbb{R} : (y, z, \theta) \text{ satisfy (14b)} - (14e) \}.
\]

In what follows, we use \( \hat{\Psi} \) and \( \hat{Y} \) to denote the set containing all vertices of the convex hull of \( \Psi \) and \( Y \), respectively. Further, let \( (x, \eta) \) be feasible for Problem (2) with the lower-level optimal-value function (3). Under Assumption 1, Proposition 3 holds and we thus 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

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

are valid for Problem (2). To derive valid cuts for the scenario-based approach, we do the following. Under Assumptions 2 and 3, we have

\[
\eta \geq \Phi(x) = \max_{\ell \in \{\Gamma d, \ldots, n+1\}} \Phi^\ell(x)
\]

\[
\geq \Phi^\ell(x) = -\Gamma d \Delta d_\ell + \max_{y \in \hat{Y}} \left\{ \sum_{i=1}^{n} \tilde{d}(\ell)_i y_i (1 - x_i) \right\}
\]

\[
\geq -\Gamma d \Delta d_\ell + \sum_{i=1}^{n} \tilde{d}(\ell)_i \hat{y}_i (1 - x_i)
\]
for arbitrary but fixed $\hat{y} \in \hat{Y}$ and for all $\ell \in \{\Gamma_d, \ldots, n + 1\}$. The first equality follows from (10) and the second one holds due to Proposition 4. As a result, the cuts

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

are valid for Problem (2). In particular, we obtain different valid cuts for each scenario $\ell \in \{\Gamma_d, \ldots, n + 1\}$. Hence, we refer to (17) as scenario interdiction cuts. Finally, we exploit the previous results to equivalently reformulate the interdiction problem with a $\Gamma_d$-robust follower facing uncertain objective function coefficients.

**Theorem 1.** Under Assumption 1, Problem (2) with the lower-level optimal-value function (3) can be equivalently reformulated by replacing Constraint (2d) with (16). Under Assumptions 2 and 3, an equivalent reformulation can be obtained by replacing Constraint (2d) with (17).

### 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 (14) 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 5.** Let $(\hat{y}, \hat{z}, \hat{\theta}) \in \hat{\Psi}$ be a non-maximal packing for Problem (14) 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 (16) associated with $(\hat{y}, \hat{z}, \hat{\theta})$ is dominated by the interdiction cut associated with $(y', \hat{z}, \hat{\theta})$.

Due to the previous result, it is sufficient to consider only 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-scenario formulation (15) 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 the scenario $\ell \in \{\Gamma_d, \ldots, n + 1\}$ in the previous definition since we consider the same scenario-independent set $\hat{Y}$ in each sub-problem. To obtain a dominance result for scenario interdiction cuts associated with maximal packings, we need to further study the properties of the modified objective function coefficients $\bar{d}(\ell)_i$ for each scenario $\ell \in \{\Gamma_d, \ldots, n + 1\}$. Note that the modified objective function coefficients can be non-positive for certain items in some scenarios. If $\bar{d}(\ell)_i \leq 0$ holds for all $\ell \in \{\Gamma_d, \ldots, n + 1\}$, the $i$th item will not be chosen by the follower in any scenario. 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 of the scenarios, i.e., $\bar{d}(k)_i \leq 0$ for all $k \in \mathcal{S} \subset \{\Gamma_d, \ldots, n + 1\}$ and $\bar{d}(l)_i > 0$ for all $l \in \{\Gamma_d, \ldots, n + 1\} \setminus \mathcal{S}$, the $i$th item might be part of an optimal solution of the follower. Therefore, we introduce the following notation. For each scenario $\ell \in \{\Gamma_d, \ldots, n + 1\}$, we define the set

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

**Proposition 6.** The scenario interdiction cuts (17) can be replaced with

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

In particular, the cuts (18) dominate the basic scenario interdiction cuts (17).

With the previous considerations, we can finally state a dominance result for scenario interdiction cuts associated with maximal packings.
Proposition 7. Let \( y \in \hat{Y} \) be a non-maximal packing for Problem (15) and let \( y' \in \hat{Y} \setminus \{ y \} \) be such that \( y \leq y' \) holds. Then, the scenario interdiction cuts (18) associated with \( y \) are dominated by the scenario interdiction cuts associated with \( y' \).

Note that the previous result is not valid for the basic scenario interdiction cuts as stated in (17) since, in general, \( \Delta(\ell, y_i)(1 - x_i) \leq \Delta(\ell, y_{i}')(1 - x_i) \) does not hold for all \( i \notin D^t_{i} \). 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 2. Let \( s, t \in [n], s \neq t, \) be chosen such that \( A_s \geq A_t, B_s \leq B_t, u_s \geq u_t, \) \( d_s \geq d_t, \) and \( d_s - \Delta d_s \geq d_t - \Delta d_t \) hold. Then, the dominance inequality

\[
x_t \leq x_s
\]

(19)
is satisfied in at least one optimal solution of Problem (2) with the lower-level optimal-value function (3).

Proof. Let \((x^*, y^*)\) be an optimal solution of Problem (2) with the lower-level optimal-value function (3) such that the dominance inequality (19) is not satisfied. This means that there are two distinct items \( s, t \in [n] \) that satisfy the requirements of the theorem but for which \( x_s^* = 0 \) and \( x_t^* = 1 \) hold. The idea now 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. Without loss of generality, we show that \( x' \) is also an optimal solution of Problem (2) using the extended formulation. Let \((y', z', \theta')\) be an optimal solution of Problem (12) for \( x' \), i.e., we have \( y_i' = 0 \). Moreover, \( z_i' = \max\{\Delta d_i y_i' - \theta', 0\} \) holds due to Constraints (12b), (12e), and the objective function. If \( y_i' = 0 \) holds, \((y', z', \theta')\) is also a feasible follower’s decision for \( x^* \) and we obtain

\[
\Phi(x') = \sum_{i=1}^{n} d_i y_i' - \Gamma_i \theta' - \sum_{i=1}^{n} z_i' \leq \Phi(x^*).
\]

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

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

and

\[
\hat{z}_i = \begin{cases} z_i', & i \in [n] \setminus \{s, t\}, \\ \max\{\Delta d_i y_i' - \theta', 0\}, & i = s, \\ 0, & i = t. \end{cases}
\]

Then, \( \Phi(x') \leq \Phi(\hat{y}, \hat{z}, \theta') \) and we obtain a contradiction.
By construction, \((\hat{y}, \hat{z}, \hat{\theta})\) is feasible for Problem (12) given the leader’s decision \(x^*\) and we obtain

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

\[
= d_y y'_i + \sum_{i \in [n]\backslash\{s,t\}} d_i \hat{y}_i - \Gamma_d \theta' - \sum_{i \in [n]\backslash\{s,t\}} \hat{z}_i
\]

\[
= \sum_{i \in [n]\backslash\{s,t\}} d_i \hat{y}_i - \Gamma_d \theta' + \max\{\Delta d_i y'_i - \theta', 0\} - \sum_{i \in [n]\backslash\{s,t\}} \hat{z}_i
\]

\[
\leq \sum_{i \in [n]\backslash\{s,t\}} d_i \hat{y}_i - \Gamma_d \theta' + \min\{\Delta s_i y'_i + \theta', d_i \hat{y}_i\} - \sum_{i \in [n]\backslash\{s,t\}} \hat{z}_i
\]

\[
= \sum_{i \in [n]\backslash\{s,t\}} d_i \hat{y}_i - \Gamma_d \theta' - \sum_{i \in [n]\backslash\{s,t\}} \hat{z}_i
\]

Note that, in the case of only binary follower’s variables, the requirement \(u_a \geq u_t\) in the previous theorem is trivially satisfied since \(u_t = 1\) for all \(a \in [n]\) is a valid upper bound.

Further, we provide lifted cuts that dominate their respective basic counterparts stated in (16) and (18). We start by lifting the basic interdiction cuts corresponding to the extended formulation.

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

\[
\eta \geq \sum_{i=1}^{n} d_i \hat{y}_i (1 - x_i) + \sum_{k=1}^{K} (\Delta d_{ak} + \Delta d_{bk}) \hat{y}_{ak} (1 - x_{ak}) - \Gamma_d \hat{\theta} - \sum_{i=1}^{n} \hat{z}_i.
\]

**Proof.** Let \((x, \eta)\) be a feasible leader’s decision for Problem (2) with the lower-level optimal-value function (3). If \(x_{bk} = 1\) holds for all \(k \in [K]\), we obtain the basic interdiction cut as stated in (16), which is satisfied by \(x\). Otherwise, we define \(K := \{k \in [K]: x_{bk} = 0\}\). Further, we consider an alternative follower’s decision \((y', z', \hat{\theta})\), which is obtained as follows. We define \(\hat{K} := [n] \backslash \{a_k, b_k: k \in K\}\). Then, we consider \(y'_k = \hat{y}_i\) for all \(i \in \hat{K}\) and flipped entries for all \(k \in K\), i.e., \(y'_{ak} = 0\), \(y'_{bk} = \hat{\theta}_{ak}\). Further, we set \(z'_k = \hat{z}_i\) for all \(i \in \hat{K}\) as well as \(z'_{ak} = 0\) and \(z'_{bk} = \max\{\Delta d_{ak} \hat{y}_{ak} - \theta, 0\}\) for all \(k \in \hat{K}\). By construction, we have \((y', z', \hat{\theta}) \in \hat{\Psi}\). Hence, the leader’s decision \(x\) satisfies the basic interdiction cut associated with \((y', z', \hat{\theta})\), i.e.,

\[
\eta \geq \sum_{i=1}^{n} d_i y'_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i=1}^{n} z'_i
\]

\[
= \sum_{i \in \hat{K}} d_i y'_i (1 - x_i) + \sum_{k \in \hat{K}} (d_{ak} \hat{y}_{ak} (1 - x_{ak}) + d_{bk} \hat{\theta}_{ak} (1 - x_{bk}))
\]

\[
- \Gamma_d \hat{\theta} + \sum_{i \in \hat{K}} z'_i - \sum_{k \in \hat{K}} (z'_{ak} + z'_{bk})
\]

\[
= \sum_{i \in \hat{K}} d_i \hat{y}_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i \in \hat{K}} \hat{z}_i + \sum_{k \in \hat{K}} (d_{bk} \hat{y}_{ak} - \max\{\Delta d_{bk} \hat{y}_{ak} - \hat{\theta}, 0\})
\]

(21)
In particular, we have \( \bar{z}_{ak} = \max \{ \Delta d_{ak} \bar{y}_{ak} - \hat{\theta}, 0 \} \) and \( \bar{z}_{bh} = 0 \) for all \( k \in [K] \). Hence, we can re-write Inequality (20) as

\[
\eta \geq \sum_{i=1}^{n} d_i \hat{y}_i (1 - x_i) + \sum_{k=1}^{K} ((d_{bk} - \Delta d_{bk}) - (d_{ak} - \Delta d_{ak})) \hat{y}_{ak} (1 - x_{bh}) - \Gamma_d \hat{\theta} - \sum_{i=1}^{n} \bar{z}_i
\]

\[
= \sum_{i\in K} d_i \hat{y}_i (1 - x_i) + \sum_{k\in K} (d_{ak} \hat{y}_{ak} (1 - x_{ak}) + d_{bk} \hat{y}_{bk} (1 - x_{bh}))
\]

\[
+ \sum_{k\in K} ((d_{bk} - \Delta d_{bk}) - (d_{ak} - \Delta d_{ak})) \hat{y}_{ak} (1 - x_{bh})
\]

\[
+ \sum_{k\in [K]\setminus K} ((d_{bk} - \Delta d_{bk}) - (d_{ak} - \Delta d_{ak})) \hat{y}_{ak} (1 - x_{bh}) - \Gamma_d \hat{\theta}
\]

\[
- \sum_{i\in K} \bar{z}_i - \sum_{k\in K} (\bar{z}_{ak} + \bar{z}_{bh})
\]

\[
= \sum_{i\in K} d_i \hat{y}_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i\in K} \bar{z}_i + \sum_{k\in K} (d_{ak} \hat{y}_{ak} (1 - x_{ak})
\]

\[
+ ((d_{bk} - \Delta d_{bk}) - (d_{ak} - \Delta d_{ak})) \hat{y}_{ak} - \max \{ \Delta d_{ak} \hat{y}_{ak} - \hat{\theta}, 0 \} \).
\]

Subtracting the right-hand side of (22) from the right-hand side of (21) yields

\[
\sum_{k\in K} (d_{ak} \hat{y}_{ak} x_{ak} + (\Delta d_{bk} - \Delta d_{ak}) \hat{y}_{ak}
\]

\[
+ \max \{ \Delta d_{ak} \hat{y}_{ak} - \hat{\theta}, 0 \} - \max \{ \Delta d_{bk} \hat{y}_{ak} - \hat{\theta}, 0 \} \).
\]

For all \( k \in K \), we have

\[
\max \{ \Delta d_{ak} \hat{y}_{ak} - \hat{\theta}, 0 \} - \max \{ \Delta d_{bk} \hat{y}_{ak} - \hat{\theta}, 0 \}
\]

\[
= \begin{cases} 
(\Delta d_{ak} - \Delta d_{bk}) \hat{y}_{ak}, & \Delta d_{bk} \hat{y}_{ak} \geq \Delta d_{ak} \hat{y}_{ak} \geq \hat{\theta}, \\
-\Delta d_{bk} \hat{y}_{ak} + \hat{\theta}, & \Delta d_{bk} \hat{y}_{ak} \geq \hat{\theta} \geq \Delta d_{ak} \hat{y}_{ak}, \\
0, & \hat{\theta} \geq \Delta d_{bk} \hat{y}_{ak} \geq \Delta d_{ak} \hat{y}_{ak},
\end{cases}
\]

and we consequently obtain

\[
d_{ak} \hat{y}_{ak} x_{ak} + (\Delta d_{ak} - \Delta d_{bk}) \hat{y}_{ak} + \max \{ \Delta d_{ak} \hat{y}_{ak} - \hat{\theta}, 0 \} - \max \{ \Delta d_{bk} \hat{y}_{ak} - \hat{\theta}, 0 \}
\]

\[
= \begin{cases} 
d_{ak} \hat{y}_{ak} x_{ak}, & \Delta d_{bk} \hat{y}_{ak} \geq \Delta d_{ak} \hat{y}_{ak} \geq \hat{\theta}, \\
d_{ak} \hat{y}_{ak} x_{ak} - \Delta d_{ak} \hat{y}_{ak} + \hat{\theta}, & \Delta d_{bk} \hat{y}_{ak} \geq \hat{\theta} \geq \Delta d_{ak} \hat{y}_{ak}, \\
d_{ak} \hat{y}_{ak} x_{ak} + (\Delta d_{bk} - \Delta d_{ak}) \hat{y}_{ak}, & \hat{\theta} \geq \Delta d_{bk} \hat{y}_{ak} \geq \Delta d_{ak} \hat{y}_{ak}. \n\end{cases}
\]

Thus, each term in (23) is greater or equal to \( d_{ak} \hat{y}_{ak} x_{ak} \), which is equal to zero if \( x_{ak} = 0 \) and \( d_{ak} \hat{y}_{ak} > 0 \) otherwise. To sum up, Inequality (21) is valid for Problem (2), the left-hand side of both (21) and (22) are the same, and the sum (23) is non-negative. This concludes the proof. 

In the next theorem, we consider lifted versions for the enhanced scenario interdiction cuts stated in (18).

**Theorem 4.** For any arbitrary but fixed \( \hat{y} \in \hat{Y} \) and for fixed \( \ell \in \{ \Gamma_d, \ldots, n + 1 \} \), let \( K \in [n] \), \( S_{\ell}^t = \{ a_1, \ldots, a_K \} \subset D^t_{\ell} \), and \( S_{\ell}^b = \{ b_1, \ldots, b_K \} \subset D^b_{\ell} \) be chosen such that \( S_{\ell}^t \cap S_{\ell}^b = \emptyset \), \( \hat{y}_{ak} = 1 \), \( \hat{y}_{bh} = 0 \) and \( B_{ak} \geq B_{bk} \) and \( d(\ell)_{ak} < d(\ell)_{bk} \) hold for all \( k \in [K] \). Under Assumptions 2 and 3, the lifted scenario interdiction cut

\[
\eta \geq -\Gamma_d \Delta d_{\ell} + \sum_{i\in D^t_{\ell}} \hat{d}(\ell)_{i} x_i + \sum_{k=1}^{K} \sum_{i\in S_{\ell}^t} ((d(\ell)_{bk} - d(\ell)_{ak}) (1 - x_{bh}))
\]
is valid for Problem (2) with the lower-level optimal-value function (3).

Proof. Let \((x, \eta)\) 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 scenario interdiction cut as stated in (18), which is satisfied by \(x\). Otherwise, we define \(\mathcal{K} := \{k \in [K]: x_{b_k} = 0\}\) and \(\mathcal{K} := D^+ \setminus \{a_k, b_k: k \in \mathcal{K}\}\). Further, we consider the alternative follower’s decision
\[
y'_i = \begin{cases} y_i, & i \in \bar{\mathcal{K}}, \\ 1, & i \in \{b_k: k \in \mathcal{K}\}, \\ 0, & i \in \{a_k: k \in \mathcal{K}\}. \end{cases}
\]

By construction, we have \(y' \in \bar{Y}\). Hence, the leader’s decision \(x\) satisfies the basic scenario interdiction cut associated with \(y'\) and the scenario \(\ell\), i.e.,
\[
\eta \geq -\Gamma_d \Delta_d + \sum_{i \in D^+\ell} \tilde{d}(\ell)_i y'_i (1 - x_i)
= -\Gamma_d \Delta_d + \sum_{i \in \mathcal{K}} \tilde{d}(\ell)_i y'_i (1 - x_i)
+ \sum_{k \in \mathcal{K}} \left( \tilde{d}(\ell)_{a_k} y'_{a_k} (1 - x_{a_k}) + \tilde{d}(\ell)_{b_k} y'_{b_k} (1 - x_{b_k}) \right) = -\Gamma_d \Delta_d + \sum_{i \in \mathcal{K}} \tilde{d}(\ell)_i y_i (1 - x_i) + \sum_{k \in \mathcal{K}} \tilde{d}(\ell)_{b_k}.
\]

Further, we can re-write Inequality (24) as
\[
\eta \geq -\Gamma_d \Delta_d + \sum_{i \in \mathcal{K}} \tilde{d}(\ell)_i y'_{i} (1 - x_i) + \sum_{k = 1}^{K} \left( \tilde{d}(\ell)_{a_k} - \tilde{d}(\ell)_{a_k} \right)(1 - x_{b_k})
= -\Gamma_d \Delta_d + \sum_{i \in \mathcal{K}} \tilde{d}(\ell)_i y_{i} (1 - x_i) + \sum_{k = 1}^{K} \left( \tilde{d}(\ell)_{a_k} y_{a_k} (1 - x_{a_k}) + \sum_{k = 1}^{K} \left( \tilde{d}(\ell)_{b_k} - \tilde{d}(\ell)_{a_k} \right)(1 - x_{b_k}) \right)
= -\Gamma_d \Delta_d + \sum_{i \in \mathcal{K}} \tilde{d}(\ell)_i y_{i} (1 - x_i) + \sum_{k = 1}^{K} \left( \tilde{d}(\ell)_{a_k} y_{a_k} (1 - x_{a_k}) + \sum_{k \in \mathcal{K}\setminus\mathcal{K}} \tilde{d}(\ell)_{b_k} + \sum_{k = 1}^{K} \left( \tilde{d}(\ell)_{b_k} - \tilde{d}(\ell)_{a_k} \right)(1 - x_{b_k}) \right)
= -\Gamma_d \Delta_d + \sum_{i \in \mathcal{K}} \tilde{d}(\ell)_i y_{i} (1 - x_i) + \sum_{k = 1}^{K} \left( \tilde{d}(\ell)_{a_k} y_{a_k} (1 - x_{a_k}) + \sum_{k \in \mathcal{K}\setminus\mathcal{K}} \tilde{d}(\ell)_{b_k} \right)
\]

Subtracting the right-hand side of (26) from the right-hand side of (25) yields
\[
\sum_{k \in \mathcal{K}} \tilde{d}(\ell)_{a_k} x_{a_k} \geq 0
\]
as \(\tilde{d}(\ell)_{a_k} x_{a_k} = 0\) holds if \(x_{a_k} = 0\) and, otherwise, \(\tilde{d}(\ell)_{a_k} x_{a_k} > 0\) holds. Since Inequality (25) is valid for Problem (2) and the left-hand side of both (25) and (26) are the same, the lifted scenario interdiction cut (24) is valid for Problem (2).

Let us point out that the items in the sets \(S_a\) and \(S_b\) (or \(S^+_{\ell}\) and \(S^-_{\ell}\) in the multi-scenario 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 \( S_a^t \), i.e., \( \tilde{y}_a \geq 1 \) holds, we select its counterpart \( b \) among all items that satisfy the requirements with maximum value of \( ((d_b - \Delta d_b) - (d_a - \Delta d_a))\tilde{y}_a(1 - x_b) \) and \((\tilde{d}(\ell)_b - \bar{d}(\ell)_a)(1 - x_b)\) for the extended formulation and the scenario-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 \( S_a^t \) and \( S_b^t \) in the multi-scenario 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 (4). We restrict ourselves to the case in which the uncertain lower-level constraint does not contain terms depending on the leader’s decision, i.e., \( v = 0 \). This is an important assumption for the validity of the proposed cuts. Furthermore, we assume that \( w_i > 0 \) holds for all \( i \in [n] \) so that the \( \Gamma_w \)-robust lower-level problems (6) and (9) satisfy the monotonicity property. As in the case of uncertain objective function coefficients, we exploit a penalized formulation of the \( \Gamma_w \)-robust follower’s problem in which interdiction constraints are removed and penalty terms \(-d_i\tilde{y}_i x_i\) are added 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-scenario formulation, which 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

\[
\Theta = \{ y \in \mathcal{Y} : \exists z, \theta \geq 0 \text{ such that } (6b) \text{ and } (6c) \text{ are satisfied} \}.
\]

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

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

are valid for Problem (2) and can equivalently replace Constraint (2d). Under Assumption 2, i.e., \( \mathcal{Y} = \{0, 1\}^n \), and Assumption 3, we maximize over scenario-dependent feasible sets

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

when considering the multi-scenario approach. However, it is easy to see that an optimal solution needs to be contained in the union of all scenario-dependent sets. Let \( Y' \) be the set of all vertices of the convex hull of the union of all scenario-dependent sets. Then, we obtain the interdiction cuts for the multi-scenario case by replacing \( y \in \hat{\Theta} \) with \( y \in Y' \) in (27), which can equivalently replace Constraint (2d) in Problem (2).

### 4. Computational Results

We now provide detailed numerical results for the proposed methods to solve interdiction problems with a monotone \( \Gamma \)-robust follower. The solution approaches are implemented in Python 3.6.9 and Gurobi 9.1.2 is used to solve all arising optimization problems. 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 CPU 5-2699 V4 at 2.2 GHz (88 cores) with 756 GB RAM. For each test run, we set the time limit to 1h. We refer to the branch-and-cut method that exploits the multi-scenario approach as MS 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 MS and Ext, which we
assess in the following. The aim is to determine a “winner setting” for MS 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.

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

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

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 have adapted the knapsack instances from Caprara et al. (2016) to account for a $\Gamma$-robust follower. Before we comment on the uncertainty parameterizations that we consider, let us briefly describe the generation of the knapsack instances. The follower’s data is generated according to Martello et al. (1999). The profits $p_i$ and the follower’s weights $w_i$ take uncorrelated integer values from the interval $[0, 100]$. For each instance size $n \in \{35, 40, 45, 50, 55\}$, 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 $N \in \{1, \ldots, 10\}$ 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 $[C - 10, C + 10]$. To study the effects of a $\Gamma$-robust follower, we consider four different uncertainty parameterizations. We assume that the deviations take either 10% or 25% of the nominal value. The parameter $\Gamma$ is set to either 10% or 50% of the instance size $n$. In the case of a fractional value for $\Gamma$, we then consider the closest integer. Hence, our test set contains 200 robustified knapsack instances.

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.

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

**MS-D/Ext-D:** The basic setting with the addition of dominance inequalities (19) regarding the leader’s decision.

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

**MS-LD/Ext-LD:** Like MS-L or Ext-L but with the addition of dominance inequalities (19).
Figure 1. Log-scaled ECDF plots of the runtimes (in s) for Ext (left) and MS (right) using lifted cuts and/or dominance inequalities.

Table 1. Mean and median runtimes (in s) and the number of solved instances for the variants with lifted cuts and/or dominance inequalities.

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>median</th>
<th>solved</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS</td>
<td>236.85</td>
<td>34.40</td>
<td>164</td>
</tr>
<tr>
<td>MS-D</td>
<td>27.02</td>
<td>9.28</td>
<td>193</td>
</tr>
<tr>
<td>MS-L</td>
<td>62.37</td>
<td>28.16</td>
<td>190</td>
</tr>
<tr>
<td>MS-LD</td>
<td>20.18</td>
<td>6.88</td>
<td>198</td>
</tr>
<tr>
<td>Ext</td>
<td>116.58</td>
<td>8.39</td>
<td>175</td>
</tr>
<tr>
<td>Ext-D</td>
<td>4.03</td>
<td>1.71</td>
<td>200</td>
</tr>
<tr>
<td>Ext-L</td>
<td>11.57</td>
<td>4.46</td>
<td>193</td>
</tr>
<tr>
<td>Ext-LD</td>
<td>2.80</td>
<td>1.40</td>
<td>200</td>
</tr>
</tbody>
</table>

As a default for MS, we consider the cut separation strategy in which all violated (lifted) scenario interdiction cuts are added to the problem formulation. Figure 1 shows the empirical cumulative distribution functions (ECDFs) w.r.t. the runtimes 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 (log-scaled x-axis). Note that, to have a fair comparison, we only consider the instances that every variant can solve within the time limit. While lifted scenario interdiction cuts and dominance inequalities on its own clearly improve the performance of both MS and Ext, the combination of these two ingredients yields only minor further improvement of MS-D and Ext-D. Nevertheless, MS-LD and Ext-LD slightly 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 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 (12) or the scenario sub-problems (13). 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 $(y, z, \theta)$ 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” MS-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 MS-LD-Max and Ext-LD-Max. Again, we focus on the instances that every variant can solve within the time limit for a fair comparison. Based on Figure 2, it can be seen that adding only the interdiction cuts corresponding to maximal packings of the follower significantly improves the performance of the overall solution method in the multi-scenario case. When following the approach using an extended formulation, maximal packings seem to have a slightly smaller impact on the performance of the method. This is due to the additional and rather strong requirement on the follower’s variable $\theta$ to obtain a feasible maximal packing for the extended formulation. The previous observations are also supported by the mean and median runtimes for the four considered variants summarized in Table 2. It can be seen that MS-LD-Max has considerably smaller mean and median runtimes compared to MS-LD. The same can be observed for the extended formulation, although the benefits of maximal packings are not as significant as in the multi-scenario case. To sum up, the observations drawn from Figure 2 and Table 2 clearly suggest that MS-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. We now compare the performance of MS-LD-Max and Ext-LD-Max with and without warmstarts. To this end, we consider the following two options.
**MS-Heur/Ext-Heur:** We consider a modified leader’s problem similar to the one proposed in Fischetti, Ljubić, et al. (2019) to introduce a heuristic for the \( \Gamma \)-robust knapsack interdiction problem. The idea is to add invalid upper-level constraints on the leader’s variables \( x \) to allow for an analytic expression of the optimal follower’s decision. For this purpose, we consider the follower’s objective function that is minimized over the set that is described by all upper- and lower-level constraints. However, each follower’s variable \( y_i \) is replaced by \( 1 - x_i \) in the problem formulation. Hence, the modified problem only contains the leader’s variables \( x \) in the multi-scenario case. In particular, we need to solve \( n - \Gamma + 2 \) many modified scenario sub-problems to obtain an optimal leader’s decision. When considering the extended formulation, the additional follower’s variables \((z, \theta)\), which are not subject to interdiction, remain in the formulation of the modified leader’s problem. The leader now minimizes over the extended variable space \((x, z, \theta)\). However, we need to flip the signs of the terms depending on the variables \((z, \theta)\) in the objective function to account for the min-max structure of the interdiction problem and to ensure that the modified leader’s problem is bounded. Let \( x \) and \((x, z, \theta)\) be an optimal solution of the modified leader’s problem for the multi-scenario approach and for the extended formulation, respectively. Then, the optimal reaction of the follower is given by \( y_i = 1 - x_i \) for all \( i \in [n] \) in either case. Note, however, that the modified problem might also be infeasible.

**MS-Nom/Ext-Nom:** Preliminary computational results revealed that the optimal leader’s decision of the nominal knapsack interdiction problem turns out to be also the \( \Gamma \)-robust solution for some instances. As a consequence, we also implemented warmstarting the problem with the nominal solution. Note that the latter options are labeled as an abbreviation of MS-LD-Max-Heur, Ext-LD-Max-Heur, MS-LD-Max-Nom, and Ext-LD-Max-Nom. To have a fair comparison, we only consider the instances that every variant can solve within the time limit. The ECDF plots for MS-LD-Max and Ext-LD-Max with and without warmstarts are displayed in Figure 3. For the multi-scenario approach, it can be seen that warmstarting the method using either the heuristic or the nominal leader’s decision slightly improves the performance of the solution approach. However, the mean and median runtimes shown in Table 3 clearly suggest that MS-Heur should be preferred over MS-Nom. Nevertheless, it can be seen that MS-Nom has a slightly smaller median runtime compared to MS-LD-Max, which shows that warmstarts using the nominal solution slightly improve the performance of the method for easier instances.

For the extended formulation, Figure 3 and Table 3 reveal that warmstarting the solution method seems to have no improving effect. On the one hand, this is due to higher computational costs for solving the modified leader’s problem. On the other hand, Figure 3 suggests that using the nominal solution to warmstart the method particularly increases the overall runtime for the easier knapsack instances. In these cases, the benefits of Gurobi’s integrated enhancement techniques outweigh the effect of warmstarting since solving the nominal problem is relatively costly. Nevertheless, the previous observation indicates the advantage of the extended formulation over the scenario-based approach on easier instances. To sum up, we choose MS-Heur and Ext-LD-Max as our “winner settings” among the considered variants.

**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-scenario 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 scenario 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-scenario 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 MS-Heur as a prefix for the considered variants when we focus on the following cut separation strategies.

**All-In:** The default setting in which all scenario 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 (18) that is maximally violated by the current leader’s decision.

**Scenario-Sorting:** In Álvarez-Miranda, Fernández, et al. (2015), the authors propose a learning mechanism to identify the scenarios 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 the scenarios $\ell \in \{1, \ldots, n+1\}$, add the first scenario interdiction cut that is violated by the current leader’s decision, and then break the loop.

**Random:** Among the violated scenario 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 scenario sub-problems in parallel. For each instance, the idealized runtimes are computed after solving the problem sequentially by taking the maximum over the runtime of each scenario sub-problem.

In Figure 4, we compare the aforementioned cut separation strategies w.r.t. sequential and idealized runtimes. Note that we only consider the 194 instances that every variant can solve.
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. This might be due to the following. By adding the first violated scenario interdiction cut, later scenarios that might produce stronger cuts are neglected. In particular, it is possible that the cut corresponding to the same scenario is added in each iteration of the algorithm. To overcome this situation, we consider Random as a modification of First-In. Based on Figure 4, it can be observed that this variant performs significantly better. In particular, Random clearly 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 Scenario-Sorting, we need to update the information about the cumulative violation of the scenario interdiction cuts and the frequency of generation in each iteration of the algorithm. Then, we order the scenarios accordingly, which causes overall higher computational costs compared to Random. When considering Most-Violated, we need to solve all scenario sub-problems to determine the most violated scenario interdiction cut, which seems to be rather expensive. The previous observations are underlined by the results in Table 4, since the sequential mean and median runtimes for Random are considerably smaller compared to all other cut separation strategies. Note that we label the sequential and idealized mean and median runtimes with the superscripts seq and ideal. Focusing on idealized runtimes, however, it is noteworthy that Random is no longer the “winning” cut separation strategy. Based on Figure 4 and Table 4, Most-Violated considerably dominates the randomized first insertion strategy. 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 scenario sub-problems. In particular, the previous observations suggest that the higher computational costs for solving all scenario sub-problems—even if this is done in parallel—are compensated by the strength of the added cuts. To sum up, Random is the “winning” cut separation strategy when considering sequential runtimes. However, provided that the necessary capacities are available to solve all scenario sub-problems in parallel, we prefer Most-Violated.

4.6. Comparison of the Solution Approaches. We now compare the “winning” parameterizations of the extended formulation and the multi-scenario formulation. For the scenario-based approach, we particularly distinguish between the sequential and the idealized setting. Hence, we consider MS-Heur-Random, MS-Heur-Most-Violated, and Ext-LD-Max, which we abbreviate with MS-seq, MS-ideal, and Ext in the following. Figure 5 shows the ECDF plots of the three considered variants. Note that we consider sequential and idealized...
Table 4. Sequential and idealized mean and median runtimes (in s) and the number of solved instances for the variants with different cut separation strategies.

<table>
<thead>
<tr>
<th></th>
<th>mean$^{\text{seq}}$</th>
<th>median$^{\text{seq}}$</th>
<th>mean$^{\text{ideal}}$</th>
<th>median$^{\text{ideal}}$</th>
<th>solved</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-In</td>
<td>36.45</td>
<td>4.03</td>
<td>10.10</td>
<td>0.70</td>
<td>198</td>
</tr>
<tr>
<td>Most-Violated</td>
<td>27.54</td>
<td>3.21</td>
<td>1.69</td>
<td>0.50</td>
<td>200</td>
</tr>
<tr>
<td>Scenario-Sorting</td>
<td>34.43</td>
<td>2.52</td>
<td>9.53</td>
<td>0.84</td>
<td>200</td>
</tr>
<tr>
<td>First-In</td>
<td>91.40</td>
<td>14.32</td>
<td>24.27</td>
<td>2.96</td>
<td>194</td>
</tr>
<tr>
<td>Random</td>
<td>3.63</td>
<td>1.15</td>
<td>2.18</td>
<td>0.60</td>
<td>200</td>
</tr>
</tbody>
</table>

Figure 5. Log-scaled ECDF plots of the runtimes (in s) for the “winner settings” of Ext and MS.

runtimes for MS-seq and MS-ideal, respectively. Moreover, we would like to mention that all three variants solve all of the 200 robustified knapsack instances to global optimality. We observe that MS-ideal clearly outperforms the remaining two approaches. This particularly affirms that the strength of the scenario-based approach lies in the possibility to parallelize the solution of the follower’s scenario sub-problems. The previous observation is also underlined by the mean and median runtimes in Table 5. It can be seen that MS-ideal has significantly smaller mean and median runtimes compared to MS-seq and Ext. If, however, the capacity is not available to have an idealized parallelization, the scenario-based approach MS-seq still performs slightly better than Ext. Based on Figure 5, Ext seems to have an advantage over MS-seq on the easier instances, whereas it seems to have a minor impact whether one considers MS-seq or Ext for the harder instances. Nevertheless, MS-seq has slightly smaller mean and median runtimes compared to Ext according to the results in Table 5. Overall, MS-seq is thus the better method in the sequential setting.

4.7. The Computational Price of Robustness. To conclude our computational study, we address the so-called price of robustness (Bertsimas and Sim 2004) in our specific bilevel setting. This expression captures the effect of robustification, e.g., on the overall runtimes of the solution methods. To evaluate the price of robustness, we compare the three “winning” approaches MS-seq, MS-ideal, and Ext for the considered uncertainty parameterizations, which are referred to as $(\Delta, \Gamma)$. 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
Table 5. Mean and median runtimes (in s) for the “winner settings” of Ext and MS.

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>median</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS-seq</td>
<td>12.88</td>
<td>1.26</td>
</tr>
<tr>
<td>MS-ideal</td>
<td>4.49</td>
<td>0.48</td>
</tr>
<tr>
<td>Ext</td>
<td>15.71</td>
<td>1.44</td>
</tr>
</tbody>
</table>

the parameter $\Gamma$ takes of the instance size. Based on Table 6, it can be seen that for the sequential solution methods, i.e., MS-seq and Ext, the mean and median runtimes to solve the $\Gamma$-robust knapsack interdiction problem increase with increasing values of $\Delta$ and $\Gamma$. For MS-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 7. For both the multi-scenario approach and the extended formulation, we compare the nominal runtimes to the mean runtimes obtained from all considered uncertainty parameterizations.

To further assess the price of robustness w.r.t. the runtimes, we measure the relative performance of the method in the $\Gamma$-robust and in the nominal case. This is done by determining the coefficient of runtimes $q_i = t_{i,\text{rob}}/t_{i,\text{nom}}$ for each knapsack instance $i \in [200]$. Here, $t_{i,\text{rob}}$ and $t_{i,\text{nom}}$ denote the runtimes of the considered solution method for instance $i$ in the $\Gamma$-robust and in the nominal case, respectively. In Figure 6, we show box-plots for the coefficients of runtimes corresponding to MS-seq, MS-ideal, and Ext. Each box in Figure 6 represents the distribution of the determined coefficients $q$ over all instances for the considered solution methods. It can be seen that the median coefficients of runtimes differ only slightly for MS-seq and Ext. For MS-ideal, however, the median is significantly smaller and particularly lies outside of the boxes for MS-seq and Ext. This observation emphasizes the difference between the methods, namely that MS-seq and Ext are sequential solution methods, while MS-ideal exploits parallelization. Moreover, we observe that the boxes for the sequential approaches are considerably larger than the one for MS-seq. This means that the coefficients in the idealized setting are less dispersed, which suggests that the performance of MS-ideal is more stable compared to the other two approaches. This is further illustrated by the spread of the outliers. Note that, in Figure 6, we use a logarithmic scaling of the $y$-axis for runtime coefficients greater than 50 such as to capture the spread of the outliers in a detailed and comprehensive way. For the sequential approaches, the outliers are widely scattered, which shows that the performance of these methods is rather volatile. In contrast to that, smaller ranges of the outliers can be observed for MS-ideal. Based on Figure 6, however, it can be seen that the overall range of the runtime coefficients for Ext is much larger compared to the scenario-based variants. In particular, this justifies a tendency to prefer the multi-scenario approach over the extended formulation—even in the sequential setting—since it seems to be the more stable method. Nevertheless, it is noteworthy that the coefficients of runtimes for MS-seq are strictly greater than 1 for all instances. Hence, the robustification always leads to increased computational costs in this setting. For the other two approaches, we observe that some of the instances can be solved faster compared to the nominal case. However, the mean and median coefficients of runtimes shown in Table 8 emphasize that the robustification is not “for free” but that the price of robustness for MS-ideal is comparatively small.

To sum up, $\Gamma$-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’s scenario sub-problems in parallel, the price of robustness w.r.t. runtimes is comparatively small. Thus, MS-ideal clearly outperforms MS-seq and Ext. In the sequential setting, the results in Tables 5 and 8 suggest that the performance of MS-seq is slightly better than the one of Ext w.r.t. the stability of the method as well as mean and median runtimes. Based on Figure 6, however, Ext seems to have an advantage on easier instances, which also justifies the use of the extended formulation.
Table 6. Mean and median runtimes (in s) for the uncertainty parameterizations given by \((\Delta, \Gamma)\).

<table>
<thead>
<tr>
<th>uncertainty parameterization</th>
<th>((\Delta, \Gamma) = )</th>
<th>(10,10)</th>
<th>(10,50)</th>
<th>(25,10)</th>
<th>(25,50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean MS-seq</td>
<td>5.65</td>
<td>14.68</td>
<td>10.53</td>
<td>20.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MS-ideal</td>
<td>4.90</td>
<td>3.97</td>
<td>5.75</td>
<td>3.35</td>
</tr>
<tr>
<td></td>
<td>Ext</td>
<td>8.88</td>
<td>13.70</td>
<td>10.30</td>
<td>29.95</td>
</tr>
<tr>
<td>median MS-seq</td>
<td>1.09</td>
<td>1.19</td>
<td>1.22</td>
<td>1.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MS-ideal</td>
<td>0.57</td>
<td>0.40</td>
<td>0.55</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>Ext</td>
<td>0.63</td>
<td>1.57</td>
<td>0.89</td>
<td>2.71</td>
</tr>
</tbody>
</table>

Figure 6. Box-plots of the coefficients of runtimes.

5. Conclusion

In this paper, we consider mixed-integer min-max problems with a follower facing uncertain lower-level data. We exploit a \(\Gamma\)-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-scenario 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 \(\Gamma\)-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 show that the extended formulation (Ext) performs better on easier knapsack instances. However, slightly smaller overall mean and median runtimes and a more stable performance of the method compared to Ext can be observed for the multi-scenario formulation (MS). In particular, we can exploit parallelization for the multi-scenario formulation, which is a major strength of this solution approach. Nevertheless, the study clearly justifies the extended formulation as well as the multi-scenario approach.

Despite the contribution of this paper, there are still several interesting research questions that require further investigation. We briefly sketch three of them.
Table 7. Runtime results (in s) for the nominal knapsack instances and mean runtimes for their robust counterparts.

<table>
<thead>
<tr>
<th>size</th>
<th>instance</th>
<th>MS-seq nominal</th>
<th>MS-ideal nominal</th>
<th>Ext nominal</th>
<th>MS-seq robust</th>
<th>MS-ideal robust</th>
<th>Ext robust</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>1</td>
<td>0.12</td>
<td>0.12</td>
<td>0.18</td>
<td>0.28</td>
<td>0.28</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.28</td>
<td>0.31</td>
<td>0.43</td>
<td>0.77</td>
<td>1.36</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.86</td>
<td>1.13</td>
<td>0.69</td>
<td>0.84</td>
<td>1.36</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.85</td>
<td>2.45</td>
<td>0.18</td>
<td>0.62</td>
<td>1.25</td>
<td>2.41</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.64</td>
<td>2.36</td>
<td>0.33</td>
<td>0.18</td>
<td>0.34</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.20</td>
<td>0.54</td>
<td>0.07</td>
<td>0.19</td>
<td>0.30</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.17</td>
<td>0.36</td>
<td>0.08</td>
<td>0.16</td>
<td>0.36</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.03</td>
<td>0.08</td>
<td>0.07</td>
<td>0.03</td>
<td>0.08</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.05</td>
<td>0.13</td>
<td>0.05</td>
<td>0.05</td>
<td>0.13</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.04</td>
<td>0.08</td>
<td>0.05</td>
<td>0.04</td>
<td>0.08</td>
<td>0.18</td>
</tr>
<tr>
<td>40</td>
<td>1</td>
<td>0.20</td>
<td>1.13</td>
<td>0.29</td>
<td>0.20</td>
<td>0.26</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.36</td>
<td>2.03</td>
<td>0.38</td>
<td>0.35</td>
<td>0.46</td>
<td>1.49</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.99</td>
<td>4.97</td>
<td>1.89</td>
<td>0.98</td>
<td>1.78</td>
<td>4.67</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.71</td>
<td>4.11</td>
<td>1.36</td>
<td>0.71</td>
<td>1.07</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.53</td>
<td>2.99</td>
<td>0.23</td>
<td>0.53</td>
<td>1.49</td>
<td>1.89</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.14</td>
<td>0.62</td>
<td>0.16</td>
<td>0.14</td>
<td>0.37</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.16</td>
<td>0.83</td>
<td>0.07</td>
<td>0.16</td>
<td>0.34</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.08</td>
<td>0.44</td>
<td>0.06</td>
<td>0.08</td>
<td>0.14</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.12</td>
<td>0.75</td>
<td>0.08</td>
<td>0.12</td>
<td>0.44</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.02</td>
<td>0.14</td>
<td>0.02</td>
<td>0.02</td>
<td>0.04</td>
<td>0.02</td>
</tr>
<tr>
<td>45</td>
<td>1</td>
<td>0.23</td>
<td>1.49</td>
<td>0.31</td>
<td>0.23</td>
<td>0.33</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.29</td>
<td>1.56</td>
<td>0.48</td>
<td>0.29</td>
<td>0.44</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.49</td>
<td>8.44</td>
<td>1.00</td>
<td>1.46</td>
<td>2.78</td>
<td>3.69</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3.40</td>
<td>12.27</td>
<td>0.49</td>
<td>3.37</td>
<td>4.90</td>
<td>7.49</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.72</td>
<td>5.29</td>
<td>0.61</td>
<td>1.71</td>
<td>3.34</td>
<td>7.54</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.29</td>
<td>1.28</td>
<td>0.12</td>
<td>0.29</td>
<td>0.60</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.10</td>
<td>0.57</td>
<td>0.08</td>
<td>0.10</td>
<td>0.28</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.16</td>
<td>0.59</td>
<td>0.08</td>
<td>0.15</td>
<td>0.30</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.06</td>
<td>0.67</td>
<td>0.09</td>
<td>0.06</td>
<td>0.26</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.06</td>
<td>0.56</td>
<td>0.09</td>
<td>0.06</td>
<td>0.20</td>
<td>0.88</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>0.27</td>
<td>2.22</td>
<td>0.48</td>
<td>0.25</td>
<td>0.40</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.60</td>
<td>13.88</td>
<td>5.67</td>
<td>2.50</td>
<td>4.72</td>
<td>15.80</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3.38</td>
<td>13.71</td>
<td>3.94</td>
<td>3.31</td>
<td>5.67</td>
<td>11.27</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.15</td>
<td>6.39</td>
<td>2.06</td>
<td>1.15</td>
<td>2.35</td>
<td>11.27</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5.45</td>
<td>14.30</td>
<td>16.69</td>
<td>5.55</td>
<td>5.34</td>
<td>16.69</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.33</td>
<td>1.06</td>
<td>1.40</td>
<td>0.32</td>
<td>0.49</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.27</td>
<td>1.15</td>
<td>1.89</td>
<td>0.27</td>
<td>0.64</td>
<td>1.89</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.12</td>
<td>0.57</td>
<td>0.56</td>
<td>0.12</td>
<td>0.26</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.19</td>
<td>1.04</td>
<td>5.22</td>
<td>0.19</td>
<td>0.62</td>
<td>5.22</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.05</td>
<td>0.51</td>
<td>0.26</td>
<td>0.06</td>
<td>0.13</td>
<td>0.26</td>
</tr>
<tr>
<td>55</td>
<td>1</td>
<td>0.51</td>
<td>4.03</td>
<td>2.17</td>
<td>0.52</td>
<td>0.87</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.64</td>
<td>10.16</td>
<td>5.54</td>
<td>2.72</td>
<td>3.09</td>
<td>3.57</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>113.95</td>
<td>413.26</td>
<td>99.16</td>
<td>112.48</td>
<td>128.28</td>
<td>536.64</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>14.07</td>
<td>84.75</td>
<td>79.55</td>
<td>13.98</td>
<td>38.12</td>
<td>19.57</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5.62</td>
<td>8.52</td>
<td>18.57</td>
<td>5.51</td>
<td>5.23</td>
<td>18.57</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.36</td>
<td>1.24</td>
<td>11.86</td>
<td>0.35</td>
<td>0.94</td>
<td>11.86</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.44</td>
<td>1.63</td>
<td>5.50</td>
<td>0.43</td>
<td>0.88</td>
<td>5.50</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.31</td>
<td>1.32</td>
<td>13.14</td>
<td>0.31</td>
<td>0.90</td>
<td>13.14</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.07</td>
<td>0.66</td>
<td>2.12</td>
<td>0.07</td>
<td>0.29</td>
<td>2.12</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.07</td>
<td>0.74</td>
<td>2.04</td>
<td>0.07</td>
<td>0.42</td>
<td>2.04</td>
</tr>
</tbody>
</table>
Table 8. Minimum, mean, median, and maximum values for the coefficients of runtimes.

<table>
<thead>
<tr>
<th></th>
<th>min</th>
<th>mean</th>
<th>median</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS-seq</td>
<td>2.02</td>
<td>23.14</td>
<td>8.64</td>
<td>786.49</td>
</tr>
<tr>
<td>MS-ideal</td>
<td>0.71</td>
<td>9.29</td>
<td>3.49</td>
<td>191.96</td>
</tr>
<tr>
<td>Ext</td>
<td>0.41</td>
<td>48.53</td>
<td>9.19</td>
<td>1458.82</td>
</tr>
</tbody>
</table>

(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-scenario 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


REFERENCES


APPENDIX A. OMITTED PROOFS

Proof of Proposition 1. Let $x$ be a feasible upper-level decision. Further, let $y \in Y(x)$ and let $y' \in Y$ be such that $y' \leq y$ holds. Due to $B \in \mathbb{R}_+^{m \times n}$, we obtain

$$By' \leq By \leq b,$$

$$y'_i \leq y_i \leq u_i(1 - x_i), \quad i \in [n],$$

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

Proof of Proposition 2. Let $x$ be a feasible upper-level decision. First, we show that the extended formulation (12) satisfies the monotonicity property. To this end, let $(y, z, \theta)$ be a feasible follower’s decision for Problem (12) for the given $x$. Further, let $y' \in Y$ be such that $y' \leq y$ holds. Due to $B \in \mathbb{R}_+^{m \times n}$ and $\Delta_i \geq 0$ for all $i \in [n]$, we obtain

$$z_i + \theta \leq \Delta_i y_i \leq \Delta_i y'_i, \quad i \in [n],$$

$$By' \leq By \leq b,$$

$$y'_i \leq y_i \leq u_i(1 - x_i), \quad i \in [n],$$

i.e., the follower’s decision $(y', z, \theta)$ is feasible for Problem (12) for the given $x$. Second, we show that each scenario sub-problem (13) satisfies the monotonicity property. Note that the feasible set of (13) is scenario-independent. Hence, there is no need to specify
the scenario \( \ell \in \{ \Gamma_d, \ldots, n + 1 \} \). For the given \( x \), let \( y \) be a feasible follower’s decision for sub-problem (13). 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-scenario 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 (13) for the given \( x \). Consequently, the \( \Gamma_d \)-robust follower’s problems (12) and (13) satisfy the monotonicity property. \( \square \)

**Proof of Proposition 3.** Let \( x \) be a feasible upper-level decision. For notational convenience, let \( \Psi(x) \) and \( \Psi \) denote the feasible sets of Problems (12) and (14), respectively. Further, let \((y^*, z^*, \theta^*)\) be an optimal solution of Problem (12) for the given leader’s decision \( x \). Then, \((y^*, z^*, \theta^*)\) is also feasible for Problem (14), 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 \((y^*, z^*, \theta^*)\). Thus, we obtain

\[
\max \left\{ \sum_{i=1}^{n} d_i y_i - \Gamma_d \theta - \sum_{i=1}^{n} z_i \in \Psi(x) \right\} \leq \max \left\{ \sum_{i=1}^{n} d_i y_i (1 - x_i) - \Gamma_d \theta - \sum_{i=1}^{n} z_i \in \Psi \right\}. \tag{28}
\]

Let \((\hat{y}, \hat{z}, \hat{\theta})\) be an optimal solution of Problem (14) 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 \((y', z', \theta)\) 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, \((y', z', \theta)\) is feasible for Problem (14) and satisfies all interdiction constraints. Moreover, we obtain

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

i.e., the alternative follower’s decision is optimal for Problem (14). In particular, it is also feasible for Problem (12), i.e., \((y', z', \theta) \in \Psi(x)\), and we have \( y'_i x_i = 0 \) for all \( i \in [n] \). Hence,
we obtain

\[
\max \left\{ \sum_{i=1}^{n} d_i y_i (1 - x_i) - \Gamma_d \theta - \sum_{i=1}^{n} z_i : (y, z, \theta) \in \Psi \right\} = \sum_{i=1}^{n} d_i y_i (1 - x_i) - \Gamma_d \theta - \sum_{i=1}^{n} z_i
\]

\[
\leq \max \left\{ \sum_{i=1}^{n} d_i y_i - \Gamma_d \theta - \sum_{i=1}^{n} z_i : (y, z, \theta) \in \Psi(x) \right\}.
\]

Due to (28) and (29), Problem (12) and (14) admit the same optimal value. \( \square \)

**Proof of Proposition 4.** Let \( x \) be a feasible upper-level decision and let the scenario \( \ell \in \{\Gamma_d, \ldots, n + 1\} \) be arbitrary but fixed. Further, let \( y^* \) be an optimal solution of the \( \ell \)th sub-problem (13) for the given leader’s decision \( x \). Then, \( y^* \) is also feasible for the \( \ell \)th sub-problem (15), 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

\[
\max \left\{ \sum_{i=1}^{n} \tilde{d}(\ell)_i y_i : y \in Y(x) \right\} \leq \max \left\{ \sum_{i=1}^{n} \tilde{d}(\ell)_i (1 - x_i) : y \in Y \right\}.
\] (30)

Let \( \tilde{y} \) be an optimal solution of the \( \ell \)th sub-problem (15) 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 \( \tilde{y}_k \leq 1 - x_k \) is not satisfied, i.e., \( \tilde{y}_k = 1 - x_k \). Then, we consider the alternative follower’s decision

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

By construction, \( y' \) is feasible for the \( \ell \)th sub-problem (15) and satisfies all interdiction constraints. Moreover, we obtain

\[
\max \left\{ \sum_{i=1}^{n} \tilde{d}(\ell)_i (1 - x_i) : y \in Y \right\} = \tilde{d}(\ell)_k \tilde{y}_k (1 - x_k) + \sum_{i \in [n] \setminus \{k\}} \tilde{d}(\ell)_i \tilde{y}_i (1 - x_i)
\]

\[
= \tilde{d}(\ell)_k \tilde{y}_k (1 - x_k) + \sum_{i \in [n] \setminus \{k\}} \tilde{d}(\ell)_i \tilde{y}'_i (1 - x_i)
\]

\[
= \sum_{i=1}^{n} \tilde{d}(\ell)_i y'_i (1 - x_i),
\]

i.e., the alternative follower’s decision is optimal for Problem (15). In particular, it is also feasible for Problem (13), i.e., \( y' \in Y(x) \), and we have \( y'_i x_i = 0 \) for all \( i \in [n] \). Hence, we obtain

\[
\max \left\{ \sum_{i=1}^{n} \tilde{d}(\ell)_i y_i (1 - x_i) : y \in Y \right\} = \sum_{i=1}^{n} \tilde{d}(\ell)_i y'_i (1 - x_i)
\]

\[
= \sum_{i=1}^{n} \tilde{d}(\ell)_i y'_i
\]

\[
\leq \max \left\{ \sum_{i=1}^{n} \tilde{d}(\ell)_i y_i : y \in Y(x) \right\}.
\] (31)

Due to (30) and (31), Problem (13) and (15) admit the same optimal value. \( \square \)
Proof of Theorem 1. 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 \((x, \eta)\) with respect to either the interdiction cuts (16) or the scenario interdiction cuts (17) implies \(\eta \geq \Phi(x)\). To this end, suppose that \((x, \eta)\) satisfies \(Ax \geq a\) and either the interdiction cuts
\[
\eta \geq \sum_{i=1}^{n} d_i \bar{y}_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i=1}^{n} \hat{z}_i \quad \text{for all} \ (\bar{y}, \hat{z}, \hat{\theta}) \in \bar{\Psi}
\]
or the scenario interdiction cuts
\[
\eta \geq -\Gamma_d \Delta d_{\ell} + \sum_{i=1}^{n} \bar{d}(\ell)_i \bar{y}_i (1 - x_i) \quad \text{for all} \ \bar{y} \in \bar{Y}, \ \ell \in \{\Gamma_d, \ldots, n + 1\}.
\]
By Propositions 3 and 4, this is equivalent to
\[
\eta \geq \max \left\{ \sum_{i=1}^{n} d_i \bar{y}_i (1 - x_i) - \Gamma_d \hat{\theta} - \sum_{i=1}^{n} \hat{z}_i : (\bar{y}, \hat{z}, \hat{\theta}) \in \bar{\Psi} \right\} = \Phi(x)
\]
and
\[
\eta \geq \max_{\ell \in \{\Gamma_d, \ldots, n + 1\}} \left\{ -\Gamma_d \Delta d_{\ell} + \max_{\bar{y} \in \bar{Y}} \left( \sum_{i=1}^{n} \bar{d}(\ell)_i \bar{y}_i (1 - x_i) \right) \right\} = \Phi(x),
\]
which concludes the proof.

Proof of Proposition 5. This follows immediately from \(d_i > 0\) and from the fact that \(x_i \in \{0, 1\}\) implies \(\bar{y}_i (1 - x_i) \leq y'_i (1 - x_i)\) for all \(i \in [n]\).

Proof of Proposition 6. Let \((x, \eta)\) be feasible for Problem (2) with the lower-level optimal-value function (3). Further, let the scenario \(\ell \in \{\Gamma_d, \ldots, n + 1\}\) be arbitrary but fixed. Due to \(x_i \in \{0, 1\}\) for all \(i \in [n]\), we have \(\bar{d}(\ell)_i (1 - x_i) \leq 0\) for all \(i \notin D^+_\ell\). Hence, all follower’s variables \(y_i\) with \(i \notin D^+_\ell\) could be omitted in this scenario sub-problem, i.e., we obtain
\[
\Phi^\ell(x) = \max \left\{ \sum_{i \in D^+_\ell} \bar{d}(\ell)_i \bar{y}_i (1 - x_i) : y \in \hat{Y} \right\}.
\]
The validity of the new scenario interdiction cuts (18) can be shown in a similar way as it is done in Section 3. In particular,
\[
-\Gamma_d \Delta d_{\ell} + \sum_{i=1}^{n} \bar{d}(\ell)_i \bar{y}_i (1 - x_i)
\]
\[
= -\Gamma_d \Delta d_{\ell} + \sum_{i \in D^+_\ell} \bar{d}(\ell)_i \bar{y}_i (1 - x_i) + \sum_{i \notin D^+_\ell} \bar{d}(\ell)_i \bar{y}_i (1 - x_i) \leq 0
\]
holds for all \(\bar{y} \in \hat{Y}\) and \(\ell \in \{\Gamma_d, \ldots, n + 1\}\). Thus, the cuts (18) dominate the basic scenario interdiction cuts (17).

Proof of Proposition 7. This can be shown in analogy to the proof of Proposition 5.