Connections and Reformulations between Robust and Bilevel Optimization

Marc Goerigk, Jannis Kurtz, Martin Schmidt, Johannes Thürauf

Abstract. Robust and bilevel optimization share the common feature that they involve a certain multilevel structure. Hence, although they model something rather different when used in practice, they seem to have a similar mathematical structure. In this paper, we analyze the connections between different types of robust problems (strictly robust problems with and without decision-dependence of their uncertainty sets, min-max-regret problems, and two-stage robust problems) as well as of bilevel problems (optimistic problems, pessimistic problems, and robust bilevel problems). It turns out that bilevel optimization seems to be more general in the sense that for most types of robust problems, one can find proper reformulations as bilevel problems but not necessarily the other way around. We hope that these results pave the way for a stronger connection between the two fields—in particular to use both theory and algorithms from one field in the other and vice versa.

1. Introduction

Both robust and bilevel optimization have been highly active fields of research in mathematical optimization and operations research over the last years and decades. When seen from a more general point of view, both types of problems involve a certain kind of multilevel structure. In bilevel optimization, the decision maker at the first level (also called the leader) decides first while anticipating the optimal reaction of the second-level player (also called the follower), whose optimization problem is parameterized by the leader's decision. Moreover, the leader's problem itself depends on the reaction of the follower so that the first and second level are interdependent and cannot be solved in a sequential way. For a general overview of the field, we refer to the books by Dempe (2002) and Dempe et al. (2015), the annotated bibliography by Dempe (2020), or the recent survey by Kleinert et al. (2021). In robust optimization, there are also two types of players acting. The first one is the decision maker, who usually wants to take a decision so that feasibility is guaranteed for all possible "choices" of the other agent and that, among all these so-called robust feasible points, the best one is chosen. We use quotation marks here since in robust optimization, the second agent is usually not an actual human decision maker but is used to model the realization of uncertain parameters in some prescribed uncertainty set. The key feature of robust optimization is that the latter realization is studied in some kind of worst-case paradigm. For more details we refer to the seminal paper by Soyster (1973) as well as to the books and surveys by Ben-Tal and Nemirovski (1998), Ben-Tal et al. (2009), Bertsimas et al. (2011), Bertsimas and den Hertog (2022), and Buchheim and Kurtz (2018).

This informal discussion already highlights the similarities of robust and bilevel optimization: two "decision makers" are acting and their problems depend on each

Date: June 22, 2023.

2020 *Mathematics Subject Classification.* 90C70, 91A65.

Key words and phrases. Bilevel optimization, Robust optimization, Reformulations.

other. The usual setting of robust optimization is that the objective of the "uncertainty player" is simply to harm the other decision maker. With respect to this, bilevel optimization is more general by allowing arbitrary objective functions of the second player that can, but do not need to, aim for harming the other player. We will later see that this feature of bilevel optimization renders this class more general than robust optimization.

For quite some time, the communities of robust optimization on the one hand and bilevel optimization on the other hand both had the vague intuition that there is some strong connection between these two fields. However, and maybe because the mentioned two communities have been rather disjoint, no systematic study of the connections of these two fields has been carried out. This changed at the 2022 Dagstuhl workshop "Optimization at the second level" that had the explicit aim to bring together researchers of the two fields to discuss the commonalities and the differences of their fields and to understand how one field can benefit from the theory and algorithms from the other field and vice versa; see also the workshop report by Brotcorne et al. (2023).

The aim of this paper is to follow the spirit of this workshop and to shed some first light on the connections of robust and bilevel optimization. To formally study these connections, throughout the paper we provide answers to questions of the following form:

> *If P is an instance of problem class $\mathcal{P}$ and if A is an algorithm for solving instances of problem class $\mathcal{Q}$, can then A also be used to solve P?*

Let us make this more clear using the example of one of the basic results provided in this paper: One can use an algorithm A for solving optimistic bilevel optimization problems $\mathcal{Q}$ for solving a strictly robust optimization problem P. The technique to study this is to provide proper reformulations—e.g., in the example above of a strictly robust problem as an optimistic bilevel problem.

So far, we explained the motivation and contribution using the example of strictly robust optimization problems as well as of optimistic bilevel problems. However, we also go a few steps further and additionally present connections between strictly robust optimization problems with decision-dependent uncertainty sets, two-stage robust optimization, and min-max regret problems as well as pessimistic bilevel problems and robust bilevel problems. For the latter and particularly new class of robust bilevel problems, we refer to the two recent papers by Beck et al. (2022, 2023). For the other, more established, problem classes we give some pointers to the classic literature when discussing the respective class.

To the best of our knowledge, there are only two papers in the literature that explicitly mention and discuss a connection between robust and bilevel optimization. First, Leyffer et al. (2020) survey nonlinear and robust optimization and mention what they call the "bilevel approach to robust optimization". By doing so, the authors highlighted that robust optimization problems with decision-dependent uncertainty can be written as bilevel problems. Second, Wiesemann et al. (2013) study pessimistic bilevel optimization and exploit the key idea that the standard pessimistic bilevel problem can be written as a robust optimization problem with decision-dependent uncertainty sets.

Finally, we also would like to mention that there exists another highly related class of problems, namely (generalized) semi-infinite optimization problems, which are equivalent to strictly robust optimization problems (with decision-dependent uncertainty sets in the generalized setting). We do not explicitly analyze the connections of semi-infinite optimization to robust and bilevel optimization but refer to the book by Stein (2013), in which the connection to bilevel optimization is studied in detail and where also robust optimization is mentioned as a special case; see also Ben-Tal and Nemirovski (1998) for the latter relation.

The remainder of the paper is structured as follows. In Section 2 we formally introduce all problem classes that we study afterward. Then, we discuss the connections between (optimistic and pessimistic) bilevel optimization and strictly robust optimization (with and without decision-dependence of the uncertainty sets) in Section 3. Afterward, in Section 4 we study the relations between bilevel optimization and min-max-regret problems. Lastly, the connections between two-stage robust and robust bilevel problems is considered in Section 5. The paper closes with some discussion of the results and future research directions in Section 6. There, we also summarize our findings in Figure 1.

2. Problem Statements and Reformulations

For the remainder of this paper, we denote a feasible and globally optimal point of an optimization problem as a solution and make the following assumption.

Assumption 1. *For all of the following optimization problems, there exists a solution.*

2.1. **Bilevel Optimization.** We consider bilevel optimization problems given by

$$\text{“}\min_{x}\text{”} \quad F(x,y) \tag{1a}$$
$$\text{s.t.} \quad G(x,y) \leq 0, \tag{1b}$$
$$y \in S(x), \tag{1c}$$

where $S(x)$ denotes the set of solutions of the x-parameterized problem

$$\min_{y} \quad f(x,y) \tag{2a}$$
$$\text{s.t.} \quad g(x,y) \leq 0. \tag{2b}$$

Problem (1) is referred to as the upper-level (or the leader's) problem and Problem (2) is the so-called lower-level (or the follower's) problem. In the literature, one often finds further upper- as well as lower-level constraints $x \in X$ and $y \in Y$. We do not state them here explicitly but consider them as being part of the feasible set described by $G(x,y) \leq 0$ and $g(x,y) \leq 0$, respectively. The objective functions are given by $F, f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}$ and the constraint functions by $G : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^m$ as well as $g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^\ell$. In the case that the lower-level problem does not have a unique solution, the bilevel problem (1) and (2) is ill-posed. This ambiguity is expressed by the quotation marks in (1a). To overcome this issue, it is common to pursue either an optimistic or a pessimistic approach to bilevel optimization; see, e.g., Dempe (2002). In the optimistic setting, the leader chooses the follower's response among the multiple solutions of the lower-level problem such that it favors the leader's objective function value. Hence, the leader also minimizes over the solutions of the follower, i.e., we consider the problem

$$\min_{x,y} \quad F(x,y) \tag{3a}$$
$$\text{s.t.} \quad G(x,y) \leq 0, \tag{3b}$$
$$y \in S(x). \tag{3c}$$

In the pessimistic setting, the leader anticipates that, among the multiple solutions of the follower, the worst possible response w.r.t. the upper-level objective function will be chosen by the follower. Thus, one studies the problem

$$\min_{x} \quad \max_{y \in S(x)} \; F(x, y) \tag{4a}$$
$$\text{s.t.} \quad G(x, y') \leq 0 \quad \forall y' \in S(x); \tag{4b}$$

see also Wiesemann et al. (2013).

A point (x, y) is called (bilevel) feasible for the bilevel problem (1) if x, y satisfy the upper level constraints of (1) and y is a solution of the follower's problem (2) for the given x. A point is bilevel optimal for Problem (1), i.e., a solution of (1), if it is bilevel feasible and obtains the smallest upper-level objective value among all bilevel feasible points. Consequently, from Assumption 1 it follows that the use of "min" instead of "inf" in Problem (1) is legitimate.

2.2. **Strictly Robust Optimization.** We also discuss strictly robust optimization problems of the form

$$\min_{x} \quad H(x) \tag{5a}$$
$$\text{s.t.} \quad h_i(x, u_i) \leq 0 \quad \forall i \in I, \ \forall u_i \in U_i, \tag{5b}$$
$$h_j(x) \leq 0 \quad \forall j \in J, \tag{5c}$$

with $I, J \subset \mathbb{N}$, $I \cap J = \emptyset$, $|I| < \infty$, $|J| < \infty$ as well as $H : \mathbb{R}^{n_x} \to \mathbb{R}$, $h_i : \mathbb{R}^{n_x} \times \mathbb{R}^{n_{u_i}} \to \mathbb{R}$ for all $i \in I$, and $h_j : \mathbb{R}^{n_x} \to \mathbb{R}$ for all $j \in J$. We suppose that the uncertainty sets $U_i \subseteq \mathbb{R}^{n_{u_i}}$ are described by finitely many constraints. For the ease of notation, we assume that the objective function H does not depend on the uncertainty u, which is without loss of generality. Otherwise, we may introduce a new variable to represent the objective value and add an uncertain epigraph constraint to the model; see, e.g., Bertsimas et al. (2011). Note that for modeling strictly robust optimization, we use scalar constraint functions h_k for all $k \in I \cup J$, which is mainly based on the following reasons. Using scalar constraint functions is a rather standard notion in strictly robust optimization problems since we can w.l.o.g. consider the uncertainty constraint-wise, i.e., we can consider a separate uncertainty set per constraint. For a more detailed discussion we refer to Bertsimas et al. (2011). In addition, the use of scalar constraint functions will simplify the presentation and proofs of some of the following results.

We further denote Problem (5) as a strictly robust optimization problem with decision-dependent uncertainty set if for at least one $i \in I$, the corresponding uncertainty set depends on the decision variables x, i.e., for $i \in I$, the uncertainty sets $U_i(x) \subseteq \mathbb{R}^{n_{u_i}}$ are given by finitely many constraints that can additionally depend on the decision variables x.

A point $x \in \mathbb{R}^{n_x}$ is strictly robust feasible for Problem (5) if it is feasible for Constraints (5c) and also satisfies (5b) for all realizations within the uncertainty sets. Note that a robust feasible point does not include any realization of the uncertainty. A robust feasible point that obtains the smallest objective value among all robust feasible points is then robust optimal, i.e., a solution of Problem (5). Again the use of "min" instead of "inf" in Problem (5) is legitimate due to Assumption 1.

2.3. **Regret Optimization.** While the classic robust counterpart considers the worst-case performance of solutions over all possible scenarios, alternative decision criteria have been studied as well. Indeed, an axiomatic consideration of decision criteria, see, e.g., French (1986), reveals that there is no perfect criterion that can fulfill all desired properties simultaneously.

The min-max regret criterion is most commonly defined with uncertainty only in the objective; see, e.g., Aissi et al. (2009), Kasperski and Zieliński (2016), and Kouvelis and Yu (2013). The idea is to find a solution that minimizes the difference to the best possible objective value over all scenarios. Formally, for an optimization problem with uncertain objective function $H(x, u)$ and uncertain constraints $h(x, u)$, we define the regret optimization problem as

$$\min_{x} \quad \max_{u \in U} \left\{ H(x,u) - \min_{\{y :\, h(y,u) \leq 0\}} H(y,u) \right\} \tag{6a}$$

$$\text{s.t.} \quad h(x,u) \leq 0 \quad \forall u \in U \tag{6b}$$

with $H : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}$ denoting the uncertain objective function and $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{\ell}$ denoting the uncertain constraints. Further, the set $U \subseteq \mathbb{R}^{n_u}$ denotes the uncertainty set that is described by finitely many constraints. As in the strictly robust case, a feasible point to the regret problem usually does not include any realization of the uncertainty. Note that here we do not assume constraint-wise uncertainty and thus use a vector-based notation for the constraints. Further note that the decision maker's solution needs to be feasible in every scenario, which is the same requirement as in strictly robust optimization. In the objective function, a normalization term has been added to represent the best possible objective value in each scenario u.

2.4. **Two-Stage Robust Optimization.** We also consider two-stage robust optimization in which in addition to here-and-now decisions x (also called first-stage decisions), there are wait-and-see decisions y (also called second-stage decisions) that can be decided after the uncertainty is revealed. We denote by

$$Y(x,u) = \{y \in \mathbb{R}^{n_y} :\, h(x,y,u) \leq 0\}$$

the set of all feasible wait-and-see decisions for a given here-and-now decision x and scenario u. The two-stage robust problem is then defined as

$$\min_{x \in X} \max_{u \in U} \min_{y \in Y(x,u)} \; H(x,y) \tag{7}$$

Here, $X \subseteq \mathbb{R}^{n_x}$ is the set of feasible here-and-now decisions and $U \subseteq \mathbb{R}^{n_u}$ denotes the uncertainty set. We suppose that the uncertainty set $U \subseteq \mathbb{R}^{n_u}$ is described by finitely many constraints. Further, we assume that $H : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}$ and $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^{\ell}$ holds. Problem (7) is a two-stage robust optimization problem with uncertainty set U. Again, we consider w.l.o.g. an objective function H independent from the uncertainty; see, e.g., Bertsimas et al. (2011).

Analogously to the case of strictly robust optimization, the uncertainty set can also depend on the decision variables x. For a given here-and-now decision x, we then describe the decision-dependent uncertainty set $U(x) \subseteq \mathbb{R}^{n_u}$ by finitely many constraints that can additionally depend on x. Note that for two-stage robust optimization, we generally cannot consider a separate uncertainty set for each constraint as it is possible in strictly robust optimization since both variants are not equivalent; see, e.g., Marandi and den Hertog (2018).

We denote a point x as (two-stage) robust feasible for Problem (7) if for each uncertainty $u \in U(x)$, there is a feasible point y that satisfies the second-stage constraints $h(x, y, u) \leq 0$. Thus, a two-stage robust feasible point does neither include the second-stage decisions y nor any realization of the uncertainty. A two-stage robust feasible point $x \in X$ that attains the minimum objective value among all two-stage robust feasible points is then optimal, i.e., a solution of Problem (7).

2.5. **Robust Bilevel Problems.** We now consider robust versions of optimistic bilevel problems. More precisely, we consider robust bilevel problems with "here-and-now follower" and with "wait-and-see follower".

In the variant with a here-and-now follower, first the leader and the follower have to make their decisions. Afterward, the uncertain parameters are revealed. This setting can be modeled as

$$\min_{x,y} \quad F(x,y) \tag{8a}$$

$$\text{s.t.} \quad G(x,y) \leq 0, \tag{8b}$$

$$y \in S(x), \tag{8c}$$

where $S(x)$ is the set of solutions of the x-parameterized problem

$$\min_y \left\{ f(x,y) \colon g(x,y,u) \leq 0 \ \forall u \in U(x) \right\}, \tag{9}$$

which models a here-and-now follower. The decision-dependent uncertainty set $U(x) \subseteq \mathbb{R}^{n_u}$ is given by finitely many constraints that can depend on the decisions x. For the constraint functions we suppose $G : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}^m$ and $g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^\ell$. As in robust optimization, we can w.l.o.g. consider a lower-level objective function $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \to \mathbb{R}$ that is independent from the uncertainty.

We call a point (x,y) (bilevel) feasible for the robust bilevel problem (8) with here-and-now follower if (x,y) satisfy $G(x,y) \leq 0$ and if y is a strictly robust solution of the follower's problem (9) for the given x. The point (x,y) is a solution of Problem (8) if it is bilevel feasible and if it obtains the smallest upper-level objective value among all bilevel feasible points. Consequently, from Assumption 1 it follows that the use of "min" in Problem (9) is legitimate.

In the variant with a wait-and-see follower, first the leader takes a here-and-now-decision, then the uncertainty is revealed, and afterward the follower makes a wait-and-see decision. The corresponding bilevel problem is given by

$$\min_{x \in X} \max_{u \in U(x)} \min_y \ \left\{ F(x,y) \colon y \in S(x,u) \right\} \tag{10}$$

with $S(x,u)$ being the set of solutions of the (x,u)-parameterized problem

$$\min_y \quad f(x,y) \quad \text{s.t.} \quad g(x,y,u) \leq 0. \tag{11}$$

The decision-dependent uncertainty set $U(x) \subseteq \mathbb{R}^{n_u}$ is again given by finitely many constraints, which may depend on the decisions x. In addition, the set $X \subseteq \mathbb{R}^{n_x}$ represents the set of feasible points of the leader and the lower-level constraints are given by $g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^\ell$.

We denote a point x as a bilevel feasible point of the optimistic robust bilevel problem (10) with wait-and-see follower if $x \in X$ and for every $u \in U(x)$ there exists $y \in S(x,u)$ that solves

$$\min_y \left\{ F(x,y) \colon y \in S(x,u) \right\}.$$

Again, a point x is optimal for Problem (10) if it is bilevel feasible for Problem (10) and it attains the smallest objective value among all bilevel feasible points.

Note that the representation of a solution of the robust bilevel problem with wait-and-see follower (10) only consists of the first-stage decision variables x. We explicitly do not include the wait-and-see decisions y as part of a solution since these decision depend on the realization of the uncertainty. This is in line with two-stage robust optimization, in which the solution also only consists of the first-stage decisions.

2.6. **Representation of Solutions.** We note that in the literature on bilevel and robust optimization, the representation of feasible points and solutions varies, which we briefly discuss in the following.

For deterministic bilevel problems of the form (1), we denote a solution by (x, y), i.e., we explicitly state the optimal response of the follower in the solution of the bilevel problem. This representation of solutions is common in the literature and allows us to better illustrate the links between bilevel and robust optimization. However, we also note that a solution of the bilevel problem (1) is also given only in terms of the upper-level decision x in the literature; see, e.g., Tahernejad et al. (2020).

For strictly robust optimization, a solution consists of the here-and-now decision x throughout the literature, i.e., a realization of the uncertainty is not part of the solution. For two-stage robust optimization, the representation of solutions in the literature is not as consistent as for strictly robust optimization. In general, there are two different representations of solutions in the literature. First, a two-stage robust solution is given by the first-stage decisions x and the second-stage variables y are not considered as part of the solution; see Ben-Tal et al. (2004). Second, a two-stage robust solution can also be given in terms of first- and second-stage variables by representing the second-stage variables by so-called decision rules, i.e., a solution is given by $(x, y(\cdot))$ in which $y : U(x) \to \mathbb{R}^{n_y}$ is a mapping from the uncertainty set to the image space of the second-stage variables. However, the representation of these decision rules is often only possible if they are restricted to certain classes such as affine decision rules; see, e.g., Yanıkoğlu et al. (2019). Since we focus on rather general functions in this article, we use the first representation of solutions in terms of the first-stage decisions only.

We now turn to the representation of solutions for robust bilevel problems. Since this field of research is very young, a consistent representation of solutions is not yet finally established in the literature. In robust bilevel problems with here-and-now follower as in (8), the follower's problem is a strictly robust optimization problem. Consequently, we represent a solution of the robust bilevel problem (8) by the leader's and follower's decisions (x, y), but do not consider any realization of the uncertainty as part of a solution. A robust bilevel problem with wait-and-see follower as in (10) is a bilevel problem containing so-called second-stage variables y whose values are determined after the realization of the uncertainty u. Thus, this class of problems combines aspects of bilevel and two-stage robust optimization. Consequently, we represent our solution in terms of the first-stage decisions x in line with our representation of two-stage robust solutions.

3. Connections Between Bilevel and Strictly Robust Optimization

For the results in this section, we need one more assumption regarding the uncertainty set and the constraint functions of (5).

Assumption 2. *For every robust feasible point x of Problem (5) and $i \in I$, the constraint functions $h_i(x, \cdot)$ are continuous. Additionally, for $i \in I$, the uncertainty set $U_i(x)$ is non-empty and compact.*

This assumption is rather natural in robust optimization and guarantees that optimizing the uncertain constraints over the uncertainty has a finite optimum that is attained. We start with a result that is also mentioned by Leyffer et al. (2020).

Theorem 1. *Let Assumption 2 be satisfied. Let further (x^*, u^*) be a solution of the optimistic bilevel problem*

$$\min_{x,u} \quad H(x) \tag{12a}$$
$$\text{s.t.} \quad h_i(x, u_i) \le 0 \quad \forall i \in I, \tag{12b}$$
$$h_j(x) \le 0 \quad \forall j \in J, \tag{12c}$$
$$u \in S(x), \tag{12d}$$

where $S(x)$ is the set of solutions of the x-parameterized lower-level problem

$$\max_{u=(u_i)_{i \in I}} \quad \sum_{i \in I} h_i(x, u_i) \quad \text{s.t.} \quad u_i \in U_i(x) \ \forall i \in I. \tag{13}$$

Then, x^* is a solution of the strictly robust optimization problem (5) with decision-dependent uncertainty sets $U_i(x)$, $i \in I$.

Proof. Let (x^*, u^*) be a solution of Problem (12). To prove feasibility of x^* for Problem (5) note that the lower-level problem is separable w.r.t. the index i. Hence, for the solution $u^* = (u_i^*)_{i \in I}$ we have that u_i^* is a solution of

$$\max_{u_i} \quad h_i(x^*, u_i) \quad \text{s.t.} \quad u_i \in U_i(x^*).$$

Since (x^*, u^*) is feasible for (12b), it follows $h_i(x^*, u_i) \le 0$ for all $u_i \in U_i(x)$, which proves that x^* is feasible for Constraints (5b). Furthermore, Constraints (5c) and (12c) are equivalent, which proves feasibility of x^* for Problem (5).

We now prove optimality of x^* for Problem (5). Due to Assumption 2, it follows that for each robust feasible point $\bar{x}$ of Problem (5) the $\bar{x}$-parameterized lower-level problem (13) admits a solution $\bar{u}$. Consequently, $(\bar{x}, \bar{u})$ is a bilevel feasible point of Problem (12). Since the objective functions of (12) and (5) are the same, it follows that x^* is also optimal for (5). □

Note that the compactness of the uncertainty set, as required in Assumption 2, is necessary for the validity of the theorem, which is shown in the following two examples. Furthermore, the examples show a significant difference between robust and bilevel optimization. Since in the classic bilevel setting a feasible solution involves the follower's solution, this solution has to be attained in the follower's problem. In contrast, in the robust setting the uncertain parameters only restrict the feasible region of the decision variables x but the worst-case parameter does not have to be attained in the robust constraint.

Example 1. *Consider the strictly robust problem*

$$\min_x \ x \quad \text{s.t.} \ -ux \le 1 \quad \forall u \in U,$$

where the open uncertainty set is given by $U = (-2, 2)$. Note that Assumption 2 does not hold in this case since U is not compact. The strictly robust solution of the latter problem is $x^* = -1/2$. The bilevel formulation in Theorem 1 is given as

$$\min_{x,u} \ x \quad \text{s.t.} \ -ux \le 1, \ u \in S(x),$$

where $S(x)$ is the set of solutions of the x-parameterized lower-level problem

$$\sup_u \ -ux \quad \text{s.t.} \quad u \in (-2, 2).$$

Note that we use sup *instead of* max *since the uncertainty set is not compact and, thus, it is not necessarily guaranteed that the maximum is attained. The latter problem* *has a solution if and only if $x = 0$. Hence, the solutions of the bilevel problem are given by $(x, u) = (0, u)$ with $u \in U$, which is a contradiction to the result in Theorem 1 if one would neglect that the uncertainty set is compact.*

Example 2. *Consider the strictly robust problem*

$$\min_x \ -x \quad s.t. \quad x \in [0,2], \ -\frac{1}{u}x \le 0 \quad \forall u \in U,$$

where the unbounded uncertainty set is given by $U = [1, \infty)$. Note that Assumption 2 does not hold in this case since U is not compact. The strictly robust solution of the latter problem is $x^* = 2$. The bilevel formulation in Theorem 1 is given as

$$\min_{x,u} \ -x \quad s.t. \quad x \in [0,2], \ -\frac{1}{u}x \le 0 \qquad u \in S(x),$$

where $S(x)$ is the set of solutions of the x-parameterized lower-level problem

$$\sup_u \ -\frac{1}{u}x \quad s.t. \quad u \in [1, \infty).$$

Note that we use sup *instead of* max *since the uncertainty set is not compact and, thus, it is not necessarily guaranteed that the maximum is attained. The latter problem has a solution that is attained if and only if $x = 0$. Hence, the solutions of the bilevel problem are given by $(x, u) = (0, u)$ with $u \in U$, which is a contradiction to the result in Theorem 1 if one would neglect the compactness assumption of the theorem.*

Remark 1.
- *(i) Theorem 1 shows that one can solve decision-dependent and strictly robust optimization problems by solving appropriately chosen optimistic bilevel problems in which the follower computes the required worst-case uncertainties by optimizing over the respective constraints of the uncertainty sets.*
- *(ii) Note that we actually prove a stronger statement than presented in Theorem 1. In the proof we show that for every bilevel feasible point (x, y) of Problem (12), the same x is a robust feasible point for (5). Under Assumption 2 every robust feasible point x can also be extended to a bilevel feasible point (x, u).*
- *(iii) In general, it is desired in bilevel optimization to have convex lower-level problems that satisfy a constraint qualification since these usually allow for single-level reformulations. For Problem (12) this is the case if the uncertain constraints depend on the uncertainty in a concave way and if the uncertainty sets $U_i(x)$ are convex (and have an interior point) as it is often an assumption in robust optimization. If the strictly robust optimization problem is even linear with polyhedral uncertainty sets, the corresponding bilevel problem (12) has both a linear upper- and lower-level problem.*

If we simply consider $U_i(x)$ being independent of x, we obtain the following result with the same proof.

Corollary 1. *Let Assumption 2 be satisfied. Let further (x^*, u^*) be a solution of the bilevel problem (12) with $U_i(x) = U_i$. Then, x^* is a solution of the strictly robust optimization problem (5).*

In the following, we show that each strictly robust problem can be reformulated as a pessimistic bilevel problem without requiring Assumption 2.

Remark 2. *A strictly robust solution of Problem (5) can also be computed by a pessimistic bilevel problem with a constant objective function in the lower-level problem.*

To this end, let (x^*, u^*) be a solution of the pessimistic bilevel problem

$$\begin{aligned} \min_{x} \max_{u \in S(x)} \quad & H(x) \\ \text{s.t.} \quad & h_i(x, u_i) \le 0 \quad \forall i \in I, \ \forall u = (u_i)_{i \in I} \in S(x), \\ & h_j(x) \le 0 \quad \forall j \in J, \end{aligned}$$

where $S(x)$ is the set of solutions of the lower-level problem

$$\min_{u=(u_i)_{i \in I}} \ 42 \quad \text{s.t.} \quad u_i \in U_i(x) \quad \forall i \in I. \tag{14}$$

Then, x^* is a solution of the strictly robust optimization problem (5) with decision-dependent uncertainty sets $U_i(x)$. Note that in Problem (14) the objective function is constant. Consequently, the set of solutions of the lower-level problem coincides with the feasible region of the follower, which is the uncertainty set. Thus, the pessimistic problem is directly equivalent to the strictly robust optimization problem and it is not necessary to assume that the uncertainty set is compact. Note that in the optimistic case of Theorem 1, the lower-level player computes the realization of the uncertainty that violates the feasibility of the upper-level player the most. To ensure that this most violating uncertainty exists, we have to require that the uncertainty set is compact in the optimistic formulation of Theorem 1.

From the latter results we can conclude that strictly robust optimization problems (both with and without decision-dependent uncertainty sets) can be written as bilevel problems—both in the optimistic and the pessimistic sense. For the classes of optimistic bilevel optimization problems and strictly robust optimization problems with decision-dependent uncertainty sets, we now also prove the reverse direction.

Theorem 2. *Let (x^*, y^*) be a solution of the strictly robust problem*

$$\min_{x,y} \quad F(x, y) \tag{15a}$$
$$\text{s.t.} \quad f(x, y) \le f(x, \tilde{y}) \quad \forall \tilde{y} \in U(x), \tag{15b}$$
$$G(x, y) \le 0, \tag{15c}$$
$$g(x, y) \le 0, \tag{15d}$$

where the decision-dependent uncertainty set is given by

$$U(x) := \{\tilde{y} \in \mathbb{R}^{n_y} : g(x, \tilde{y}) \le 0\}.$$

Then, (x^*, y^*) is a bilevel solution of the optimistic bilevel problem (3).

Proof. Let (x^*, y^*) be a solution of the strictly robust problem (15). The point (x^*, y^*) satisfies the upper-level constraints of (3) due to (15c). For fixed decisions x, the decision-dependent uncertainty set $U(x)$ equals the feasible region of the x-parameterized follower's problem (2). Consequently, Constraints (15b) and (15d) ensure that y^* is feasible and optimal for the x^*-parameterized lower-level problem (2). Thus, (x^*, y^*) is a bilevel feasible point for the optimistic bilevel problem (3).

To prove optimality of (x^*, y^*) for Problem (3), we note that every bilevel feasible point of (3) is also feasible for Problem (15). Since the objective function of (15) equals the leader's objective function in (3), it follows that (x^*, y^*) is also optimal for the optimistic bilevel problem (3). □

Remark 3. *(i) The latter theorem shows that one can solve optimistic bilevel problems by solving appropriately chosen strictly robust optimization problems with decision-dependent uncertainty sets, in which the decision-dependent* *uncertainty set represents the feasible region of the x-parameterized follower's problem.*

(ii) Note that in Theorem 2 we do not assume that the uncertainty set is compact. While common uncertainty sets studied in the robust optimization literature are compact, it can be of interest to relax this assumption; see, e.g., Buchheim and Kurtz (2017), where unbounded polyhedral uncertainty sets are considered.

(iii) We actually prove a stronger statement than presented in the theorem. In the proof, we show that every strictly robust solution of (15) is a bilevel feasible point of (3) and the other way around.

(iv) If the optimistic bilevel problem (3) consists of linear objective functions and constraints, then the corresponding strictly robust problem (15) has a linear objective function and linear constraints as well. The decision-dependent uncertainty set is then an x-parameterized polyhedral set depending on continuous "here-and-now" variables x.

(v) Note that for bilevel problems in which the follower's constraints g do not depend on the leader's decision x, Problem (15) becomes a strictly robust optimization problem with classic (non-decision dependent) uncertainty set U.

In line with Wiesemann et al. (2013), we can draw the following connection for pessimistic bilevel and robust optimization.

Remark 4. *A solution of the pessimistic bilevel problem (4) can be computed by solving the following strictly robust problem with decision-dependent uncertainty set*

$$\begin{aligned}
\min_{x,y} \quad & F(x,y) && \text{(16a)}\\
\text{s.t.} \quad & F(x,y) \geq F(x,y') \quad \forall y' \in U(x), && \text{(16b)}\\
& G(x,y') \leq 0 \quad \forall y' \in U(x), && \text{(16c)}\\
& f(x,y) \leq f(x,y') \quad \forall y' \in U(x), && \text{(16d)}\\
& G(x,y) \leq 0, && \text{(16e)}\\
& g(x,y) \leq 0, && \text{(16f)}
\end{aligned}$$

with uncertainty set

$$U(x) := \{\tilde{y} \colon f(x,\tilde{y}) \leq \chi(x),\ g(x,\tilde{y}) \leq 0\}.$$

Here, $\chi(x)$ is the optimal-value function of the x-parameterized lower-level problem (2). Constraints (16b) and (16c) model the pessimistic view of the follower by enforcing the worst-case objective value and the feasibility of the coupling constraints for all lower-level solutions. The remaining constraints model that the computed decisions y are a solution of the lower-level problem.

We stress that, in general, the optimal-value function of an optimization problem is not known or cannot be represented in a compact way. However, for specific classes of problems such as linear or convex problems (under an additional constraint qualification), there exist compact representations of the corresponding optimal-value function. These explicit formulations have been exploited to derive solution methods in bilevel and robust optimization; see, e.g., the classic textbooks by Dempe (2002) and Ben-Tal et al. (2009). Exemplarily, for linear problems and convex problems under additional constraint qualifications, the optimal-value function can be represented by finitely many constraints and variables using the corresponding Karush–Kuhn–Tucker (KKT) conditions or the strong-duality theorem. Alternatively, we now describe how one of the main reformulation techniques of robust optimization can be used to derive an explicit description of the constraint $f(x, \tilde{y}) \le \chi(x)$ of Remark 4 under specific assumptions.

Remark 5. *For a given point $x \in X$, let $\chi(x)$ be the optimal-value function of the x-parameterized lower-level problem (2). For the variable $y \in \mathbb{R}^{n_y}$, we now consider the constraint*

$$f(x, y) \le \chi(x). \tag{17}$$

We further assume that the dual of the follower's problem is given by

$$\max_{\alpha} \quad p(x, \alpha) \quad s.t. \quad v(x, \alpha) \le 0$$

with the functions $p : \mathbb{R}^{n_x} \times \mathbb{R}^{\ell} \to \mathbb{R}$ and $v : \mathbb{R}^{n_x} \times \mathbb{R}^{\ell} \to \mathbb{R}^k$. If strong duality holds, i.e., the optimal value of the dual of the follower's problem equals the optimal value of the follower's problem (2), we can equivalently reformulate Inequality (17) as

$$f(x, y) \le p(x, \alpha), \quad v(x, \alpha) \le 0.$$

This reformulation is one of the main reformulation techniques in robust optimization and can be generalized to nonlinear uncertain inequalities that are concave in the uncertainty under specific assumptions; see Ben-Tal et al. (2015).

4. Connections Between Bilevel and Regret Optimization

We now consider regret problems through the lens of bilevel optimization.

Theorem 3. *Let (x^*, y^*) be a solution of the pessimistic bilevel problem*

$$\min_{x} \left\{ \max_{(y_1, y_2) \in S(x)} H(x, y_1) - H(y_2, y_1) \colon h(x, y_1') \le 0 \ \forall y' = (y_1', y_2') \in S(x) \right\} \tag{18}$$

with $y = (y_1, y_2)$ and $S(x) = \arg\min_y \{42 : y_1 \in U, h(y_2, y_1) \le 0\}$. Then, x^* is a solution to the regret problem (6).

Proof. Let (x^*, y^*) be a solution of Problem (18). It then holds that $h(x^*, y_1) \le 0$ for all $y = (y_1, y_2) \in S(x)$. By Assumption 1, Problem (6) is feasible. Hence, $S(x)$ projected onto y_1 equals U and $h(x^*, u) \le 0$ for all $u \in U$.

Furthermore, note that for all x we have

$$\begin{aligned}
&\max_{(y_1, y_2) \in S(x)} (H(x, y_1) - H(y_2, y_1)) \\
&= \max_{y_1 \in U} \max_{y_2 : h(y_2, y_1) \le 0} (H(x, y_1) - H(y_2, y_1)) \\
&= \max_{u \in U} \left(H(x, u) - \min_{y_2 : h(y_2, u) \le 0} H(y_2, u) \right),
\end{aligned}$$

so x^* is an optimal solution to the corresponding regret problem. □

Remark 6. *(i) Theorem 3 shows that we can solve robust regret problems by solving pessimistic bilevel problems with a specific structure.*

(ii) We actually prove a stronger result that the set of feasible first-stage decisions x of the bilevel problem is the same as for the regret problem, and each solution has the same objective value under Assumption 1.

(iii) Similar to the discussion presented in Remark 2, the lower-level problem uses a constant objective function (set to the arbitrary value 42 in this case).

(iv) Note that we used Assumption 1 to ensure that Problem (6) is feasible. Indeed, without this assumption it is possible that the bilevel problem (18) is feasible, whereas the regret problem is infeasible. Imagine a "very bad" scenario $u \in U$, for which no feasible solution exists. Then, there is no $(y_1, y_2) \in S(x)$ with $y_1 = u$ and the bilevel problem can ignore this scenario, whereas the regret problem becomes infeasible.

If the uncertainty only affects the objective function, we can compute a solution of the regret problem (6) by an optimistic bilevel problem.

Theorem 4. *Let (x^*, y^*) be a solution to the optimistic bilevel problem*

$$\min_{x, y \in S(x)} \{H(x, y_1) - H(y_2, y_1) \colon h(x) \leq 0\}$$

with $y = (y_1, y_2)$ and $S(x) = \arg\min_y \{H(y_2, y_1) - H(x, y_1) : y_1 \in U, h(y_2) \leq 0\}$. Then, x^* is a solution to the regret problem (6) without uncertainty in the constraints.

Proof. As the set of feasible solutions $X = \{x : h(x) \leq 0\}$ is the same for both problems, we only need to consider the objective function. It holds for all feasible x that

$$\begin{aligned}
&\min_{y \in S(x)} (H(x, y_1) - H(y_2, y_1)) \\
&= \min_y \{H(x, y_1) - H(y_2, y_1) : y \in \arg\min_{y'} \{H(y_2', y_1') - H(x, y_1') : y_1' \in U, h(y_2') \leq 0\}\} \\
&= \max_y \{H(x, y_1) - H(y_2, y_1) : y_1 \in U, h(y_2) \leq 0\} \\
&= \max_{y_1 \in U} \{H(x, y_1) - \min_{y_2 : h(y_2) \leq 0} H(y_2, y_1)\},
\end{aligned}$$

which completes the proof. □

Most commonly, regret problems are considered for combinatorial problems with interval uncertainty; see, e.g., Kasperski and Zieliński (2016). The reason is the following result of the literature, which is crucial to treat these regret problems.

Lemma 1 (Aissi et al. (2009))**.** *Let $X \subseteq \{0, 1\}^n$ and consider the regret problem with uncertain linear objective:*

$$\min_{x \in X} \operatorname{reg}(x) \quad \textit{with} \quad \operatorname{reg}(x) = \max_{c \in U} \left\{ \sum_{i \in [n]} c_i x_i - \min_{y \in X} \sum_{i \in [n]} c_i y_i \right\}.$$

For interval uncertainty $U = \times_{i \in [n]} [\underline{c}_i, \bar{c}_i]$ with $d_i = \bar{c}_i - \underline{c}_i \geq 0$, it holds that

$$\operatorname{reg}(x) = \sum_{i \in [n]} \bar{c}_i x_i - \min_{y \in X} \{\underline{c}_i + d_i x_i\} y_i.$$

We can now derive the following result for the special case of combinatorial problems with interval uncertainty affecting only the objective function.

Theorem 5. *Let (x^*, y^*) be a solution to the optimistic bilevel problem*

$$\min_{x \in X, y \in S(x)} \sum_{i \in [n]} \bar{c}_i x_i - \sum_{i \in [n]} (\underline{c}_i + d_i x_i) y_i$$

for $X \subseteq \{0, 1\}^n$ with $d_i = \bar{c}_i - \underline{c}_i \geq 0$ and

$$S(x) = \arg\min_y \left\{ \sum_{i \in [n]} (\underline{c}_i + d_i x_i) y_i \colon y \in X \right\}.$$

Then, x^* is a solution to the regret problem

$$\min_{x\in X}\max_{c\in U}\left\{\sum_{i\in[n]} c_i x_i - \min_{y\in X}\sum_{i\in[n]} c_i y_i\right\}$$

with interval uncertainty $U = \times_{i\in[n]}[\underline{c}_i, \bar{c}_i]$.

Proof. By construction, $x^* \in X$ and, thus, it is feasible for the regret problem. Using Lemma 1, for all $x \in X$ we obtain

$$\begin{aligned}
&\min_{y\in S(x)}\sum_{i\in[n]}\bar{c}_i x_i - \sum_{i\in[n]}(\underline{c}_i + d_i x_i)y_i\\
&= \sum_{i\in[n]}\bar{c}_i x_i - \max_{y\in S(x)}\sum_{i\in[n]}(\underline{c}_i + d_i x_i)y_i\\
&= \sum_{i\in[n]}\bar{c}_i x_i - \max_{y\in X}\left\{\sum_{i\in[n]}(\underline{c}_i + d_i x_i)y_i \colon y \in \arg\min_{y'\in X}\sum_{i\in[n]}(\underline{c}_i + d_i x_i)y'_i\right\}\\
&= \sum_{i\in[n]}\bar{c}_i x_i - \min_{y\in X}\sum_{i\in[n]}(\underline{c}_i + d_i x_i)y_i
\end{aligned}$$

and x^* is therefore an optimal regret solution. □

5. Connections Between Robust Bilevel Optimization and Two-Stage Robust Optimization

In this section, we consider the connection between optimistic robust bilevel problems and two-stage robust optimization.

5.1. Wait-and-See Follower.

We first show that we can cast any two-stage robust optimization problem as a robust bilevel problem with wait-and-see follower.

Theorem 6. *Let x^* be a solution of the optimistic robust bilevel problem with wait-and-see follower*

$$\min_{x\in X}\max_{u\in U(x)}\min_{y}\{H(x,y)\colon y\in S(x,u)\} \tag{19}$$

where $X \subseteq \mathbb{R}^{n_x}$, $U(x) \subseteq \mathbb{R}^{n_u}$, and $S(x,u)$ is the set of solutions of the (x,u)-parameterized lower-level problem

$$\min_{y} \quad H(x,y) \quad \text{s.t.} \quad h(x,y,u) \le 0.$$

Then, x^* is a solution of the two-stage robust problem (7) with decision-dependent uncertainty set $U(x)$.

Proof. Let x^* be a solution of Problem (19). Robust feasibility of x^* for the two-stage robust problem (7) directly follows since the first-stage feasible region X coincides with the feasible region of the leader, the uncertainty set $U(x)$ coincides with the uncertainty set in (19), and the second-stage feasible region $Y(x,u) = \{y\colon h(x,y,u) \le 0\}$ of Problem (7) coincides with the feasible region of the follower.

From Assumption 1, it follows that there is a solution x of the two-stage robust problem (7) that is also a robust bilevel feasible point of Problem (19) since the leader and follower minimize the same objective function in Problem (19). From the same observation, the optimality of x^* follows. Hence, x^* has the same objective value in (7) as in (19), which proves optimality. □

Remark 7. *(i) The latter theorem shows that one can solve two-stage robust problems by solving appropriately chosen optimistic robust bilevel problems with wait-and-see follower, where the follower minimizes the same objective function as the leader. In this bilevel problem, the leader's model takes the role of the first-stage and the follower's model takes the role of the second-stage of the corresponding two-stage robust problem.*

(ii) Note that we actually prove a stronger statement than presented in the theorem. In the proof we show that for every robust and bilevel feasible point x of Problem (19) the same x is feasible for the two-stage robust problem (7). Moreover, every robust feasible point x of Problem (7) is also a robust bilevel feasible point x of Problem (10).

(iii) If the first and second-stage of the two-stage robust problem have only linear constraints, the corresponding bilevel problem has both a linear upper- and lower-level problem.

(iv) The result of the latter theorem also holds if we replace $U(x)$ by a decision-independent uncertainty set U.

Note that by using the optimal-value function of the follower's problem (11), we can reformulate an optimistic robust bilevel problem with wait-and-see follower as the following two-stage robust problem

$$\min_{x\in X} \max_{u\in U(x)} \min_{y} \left\{F(x,y)\colon f(x,y)\leq \chi(x,u),\, g(x,y,u)\leq 0\right\},$$

where $\chi(x,u)$ is the optimal-value function of the (x,u)-parameterized follower's problem, i.e.,

$$\chi(x,u) := \min_{y} \left\{f(x,y)\colon g(x,y,u)\leq 0\right\}.$$

However, we again stress that, in general, the optimal-value function χ is not known or cannot be represented in a compact way. We refer to Remark 5 for an exemplary case in which we can reformulate inequalities of the type $f(x,y)\leq\chi(x,u)$.

5.2. **The Worst-Case Perspective of Two-stage Robust Optimization and Robust Bilevel Optimization with Wait-and-See Follower.** The following example shows an important difference between two-stage robust optimization and robust bilevel optimization with a wait-and-see follower on the one hand and general trilevel problems on the other hand. It especially highlights the "worst-case" perspective of two-stage robust optimization and robust bilevel optimization with wait-and-see follower, which compute a solution that is feasible for all realizations of the uncertainty.

Consider the two-stage robust problem

$$\min_{z\in\{0,1\}} \max_{u\in\{0,1\}, u\leq 1-z} \min_{y} \left\{z\colon y\in\{0,1\},\; 2u\leq y\right\} \tag{20}$$

and the robust bilevel-problem with a wait-and-see follower

$$\min_{z\in\{0,1\}} \max_{u\in\{0,1\},\; u\leq 1-z} \min_{y} \left\{z\colon y\in S(z,u)\right\}, \tag{21}$$

where $S(z,u)$ is the set of solutions of the (z,u)-parameterized problem

$$\min_{y\in\{0,1\}} z \quad \text{s.t.} \quad 2u\leq y.$$

From Theorem 6 and Remark 7, it follows that Problems (20) and (21) have the same feasible "here-and-now" decisions z.

We now compare the two previous optimization problems with the trilevel problem

$$\min_{z\in\{0,1\}} z \quad \text{s.t.} \quad (u,y)\in S(z), \tag{22}$$

where $S(z)$ is the set of solutions of the z-parameterized bilevel problem

$$\max_{u\in\{0,1\}} z \quad \text{s.t.} \quad u \le 1-z,\ y\in \hat{S}(u),$$

where $\hat{S}(u)$ is the of solutions of the u-parameterized lower-level problem

$$\min_{y\in\{0,1\}} \quad z \quad \text{s.t.} \quad 2u \le y.$$

We now make a case distinction for the two possible assignments of $z \in \{0,1\}$.

If $z=1$ holds, then $u=0$ is the only possible option for the maximization player due to $u \le 1-z=0$ and $u\in\{0,1\}$. Furthermore, both possible assignments $y\in\{0,1\}$ satisfy $2u=0\le y$. Note that $z=1$ is a two-stage robust feasible point of Problem (20) and of the robust bilevel problem (21) since for each uncertainty $u\in\{0,1\}$ there exists a feasible second-stage decision y. Moreover, there exist feasible points of the trilevel problem with $z=1$, $u=0$, and $y\in\{0,1\}$. Note that the objective value is not depending on y. Thus, we obtain 1 as the overall objective value if we choose $z=1$ for all of the three problems.

If $z=0$ holds, then $u=0$ and $u=1$ both satisfy $u\le 1-z=1$. For $u=1$, there is no feasible decision $y\in\{0,1\}$ that satisfies $2u=2\le y$. Consequently, $z=0$ is neither a two-stage robust feasible point of Problem (20) nor a robust bilevel feasible point of Problem (21). However, since for $u=0$ both assignments for $y\in\{0,1\}$ satisfy $2u=0\le y$, we obtain the trilevel feasible point with $z=0$, $u=0$, and $y\in\{0,1\}$, having an objective value of 0.

Overall, the two-stage robust solution and the robust bilevel feasible solution satisfy $z=1$ and, thus, have objective value 1. However, for $z=0$, the trilevel problem (22) has solutions with objective value of 0.

Consequently, the optimal value of the two-stage robust problem (20), respectively of the robust bilevel problem (21), differs from the corresponding optimal value of the trilevel problem (22). Note that this difference regarding the solutions and objective values stems from a different interpretation of the "maximization" player. In two-stage robust optimization and robust bilevel optimization with a wait-and-see follower, the maximization player seeks to find an uncertainty $u\in U(x)$ so that there is no feasible wait-and-see decision y. Thus, the maximization player interprets the infeasibility of the last minimization problem as having an objective function value of $+\infty$, i.e., as the best possible case to achieve. Contrarily, in the corresponding trilevel optimization problem, the maximization player interprets this infeasibility of the third-level player due to the choice of u as not being feasible since maximizing over the empty set leads to $-\infty$. Thus, if possible, the maximization player avoids to choose a point $u\in U(x)$ that leads to an infeasible third-level problem.

To describe it in other words, in two-stage robust and robust bilevel optimization the maximization player actually acts in a worst-case sense for the minimization player since this player even would choose a worst-case uncertainty that leads to the overall infeasibility of the problem. For example, nature can be such a counterplayer that does not care about the needs of the leader at all. In the trilevel setup, the counterplayer is still a competitor of the minimization player since this player again optimizes in the contrary direction compared to the leader. However, the counterplayer always tries to avoid the overall infeasibility of the problem. For example, in a market environment, the counterplayer can be a competitor that works against the leader, but still tries to avoid the collapse of the market, e.g., modeled by the last minimization problem, since the competitor still participates in the same market as well.

5.3. **Here-and-Now Follower.** In this section, we briefly discuss optimistic but robust bilevel problems with a here-and-now follower. We first note that we can reformulate any strictly robust problem with decision-dependent uncertainty set as an optimistic robust bilevel problem with a here-and-now-follower. This directly follows from Theorem 1 since optimistic bilevel problems are a special case of robust bilevel problem with a here-and-now follower.

Second, we note that using the optimal-value function of the follower's problem (9), i.e.,

$$\chi(x) := \min_y \{f(x,y) \colon g(x,y,u) \leq 0 \ \forall u \in U(x)\},$$

we can reformulate an optimistic robust bilevel problem with a here-and-now follower as the following strictly robust problem

$$\min_{x,y} \{F(x,y) \colon x \in X,\ f(x,y) \leq \chi(x),\ g(x,y,u) \leq 0 \ \forall u \in U(x)\}.$$

Up to now, we have seen that robust bilevel problems with a wait-and-see follower are closely connected to two-stage robust problems. It is possible to reformulate each two-stage robust problem as a robust bilevel problem with a wait-and-see follower in which the follower imitates the second-stage. Using optimal-value functions, we also provide a reformulation for the reverse direction. Further, robust bilevel problems with a here-and-now follower are closely connected to strictly robust problems with decision-dependent uncertainty sets. It is possible to reformulate each strictly robust problem with decision-dependent uncertainty set as a robust bilevel problem with a here-and-now follower in which the follower imitates the uncertainty set. However again, we were able to show the reverse direction only if we are allowed to use optimal-value functions.

6. Conclusion

In this paper we shed some first light on the connections between robust and bilevel optimization. We summarize our findings in Figure 1. In a nutshell, our results state that bilevel optimization is more general since for most of the robust problems we found proper reformulations as a bilevel problem but not necessarily the other way around. However, we did not formally prove that the other way around is not possible. We indeed also give reformulations for some given bilevel problem as a robust problem but use optimal-value functions in the constraints of these reformulations. Our intuition is that it is not possible to state reformulations in these cases that do not use such optimal-value functions. These functions are usually not known or cannot be stated in closed and compact form, i.e., by only using a polynomial (in the number of variables and constraints of the respective problem) number of additional variables and constraints. Despite the fact that in some situations compact closed-form representations are available, using optimal-value functions for general optimizations as "usual" constraints will most likely lead to merging two levels of the polynomial hierarchy—and by this to an overall collapse of the latter. We consequently think that a proof of the impossibility of such reformulations need to use complexity-theoretic arguments, which are out of the scope of the present paper but a reasonable topic of future research. Let us also mention the brief discussion by Cerulli (2021) in this context, where a small collection of polynomial-time solvable special cases of

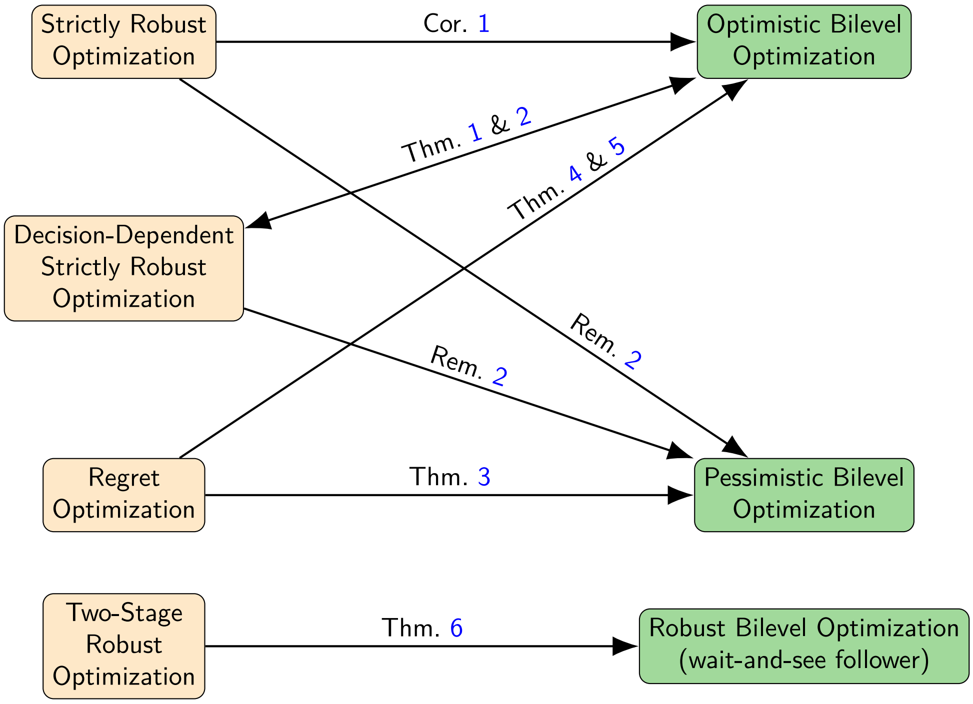

Figure 1. Summary of the relations between bilevel and robust optimization. An arrow from A to B stands for the statement that we provided a reformulation without using optimal-value functions so that an instance of A can be solved by solving an instance of B. Note that we do not include arrows between two robust or two bilevel problems to simplify the overview. However, introducing the latter would lead to further connections between robust and bilevel problems.

bilevel optimization is given, which maybe contain the cases that might have a proper reformulation as a respective robust problem without using optimal-value functions.

We also want to point to the algorithmic consequences of our results. For all the cases in which we identify a proper reformulation, this paves the way for using the theory and algorithms from one field in the other, which may open the door to many new and hybrid techniques for solving the respective problems.

Robust and bilevel optimization problems exhibit many similarities. In this paper, we have made some steps towards a better integration of these two disciplines and hope to inspire the respective research communities to work together more closely. This collaboration can be expected to bring tangible benefits. As an example, we note the seminal paper on decision-dependent uncertainty by Nohadani and Sharma (2018), where a big-M formulation is introduced. Similar formulations have also been studied in the bilevel setting, where it is shown that choosing an appropriate big-M value is as hard as solving the bilevel problem in general; see Kleinert et al. (2020). More joint workshops such as the event held in Dagstuhl will be a great boon to both communities.

Acknowledgements

The first three authors thank all other participants of the 2022 Dagstuhl workshop "Optimization at the second level" for the deep and inspiring discussion during that workshop.

References

Aissi, H., C. Bazgan, and D. Vanderpooten (2009). "Min–max and min–max regret versions of combinatorial optimization problems: A survey." In: *European Journal of Operational Research* 197.2, pp. 427–438. DOI: 10.1016/j.ejor.2008.09.012.

Beck, Y., I. Ljubić, and M. Schmidt (2022). "A Brief Introduction to Robust Bilevel Optimization." In: *SIAG on Optimization Views and News. A Forum for the SIAM Activity Group on Optimization* 30.2, pp. 1–10. URL: https://siagoptimization.github.io/assets/views/ViewsAndNews-30-2.pdf.

– (2023). "A Survey on Bilevel Optimization Under Uncertainty." In: *European Journal on Operational Research*. DOI: 10.1016/j.ejor.2023.01.008.

Ben-Tal, A. and A. Nemirovski (1998). "Robust Convex Optimization." In: *Mathematics of Operations Research* 23.4, pp. 769–805. DOI: 10.1287/moor.23.4.769.

Ben-Tal, A., D. den Hertog, and J.-P. Vial (2015). "Deriving robust counterparts of nonlinear uncertain inequalities." In: *Mathematical Programming* 149.1-2, pp. 265–299. DOI: 10.1007/s10107-014-0750-8.

Ben-Tal, A., L. El Ghaoui, and A. Nemirovski (2009). *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, pp. xxii+542. DOI: 10.1515/9781400831050.

Ben-Tal, A., A. Goryashko, E. Guslitzer, and A. Nemirovski (2004). "Adjustable robust solutions of uncertain linear programs." In: *Mathematical Programming* 99.2, pp. 351–376. DOI: 10.1007/s10107-003-0454-y.

Bertsimas, D., D. B. Brown, and C. Caramanis (2011). "Theory and Applications of Robust Optimization." In: *SIAM Review* 53.3, pp. 464–501. DOI: 10.1137/080734510.

Bertsimas, D. and D. den Hertog (2022). *Robust and Adaptive Optimization*. Dynamic Ideas LLC. URL: https://www.dynamic-ideas.com/books/robust-and-adaptive-optimization.

Brotcorne, L., C. Buchheim, D. den Hertog, and D. Henke (2023). "Optimization at the Second Level (Dagstuhl Seminar 22441)." In: *Dagstuhl Reports*. Vol. 12. 10. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. DOI: 10.4230/DagRep.12.10.207.

Buchheim, C. and J. Kurtz (2017). "Min–max–min robust combinatorial optimization." In: *Mathematical Programming* 163, pp. 1–23. DOI: 10.1007/s10107-016-1053-z.

– (2018). "Robust combinatorial optimization under convex and discrete cost uncertainty." In: *EURO Journal on Computational Optimization* 6.3, pp. 211–238. DOI: 10.1007/s13675-018-0103-0.

Cerulli, M. (2021). "Bilevel optimization and applications." PhD thesis. Institut Polytechnique de Paris. URL: https://theses.hal.science/tel-03587548/document.

Dempe, S. (2002). *Foundations of Bilevel Programming*. Springer US. DOI: 10.1007/b101970.

Dempe, S. (2020). "Bilevel Optimization: Theory, Algorithms, Applications and a Bibliography." In: *Bilevel Optimization: Advances and Next Challenges*. Ed. by S. Dempe and A. Zemkoho. Springer International Publishing, pp. 581–672. DOI: 10.1007/978-3-030-52119-6_20.

Dempe, S., V. Kalashnikov, G. A. Pérez-Valdés, and N. Kalashnykova (2015). *Bilevel Programming Problems*. Springer. DOI: 10.1007/978-3-662-45827-3.

French, S. (1986). *Decision theory: an introduction to the mathematics of rationality*. Halsted Press.

Kasperski, A. and P. Zieliński (2016). "Robust Discrete Optimization Under Discrete and Interval Uncertainty: A Survey." In: *Robustness analysis in decision aiding, optimization, and analytics*, pp. 113–143. DOI: 10.1007/978-3-319-33121-8_6.

Kleinert, T., M. Labbé, I. Ljubić, and M. Schmidt (2021). "A Survey on Mixed-Integer Programming Techniques in Bilevel Optimization." In: *EURO Journal on Computational Optimization* 9, p. 100007. DOI: 10.1016/j.ejco.2021.100007.

Kleinert, T., M. Labbé, F. Plein, and M. Schmidt (2020). "There's No Free Lunch: On the Hardness of Choosing a Correct Big-M in Bilevel Optimization." In: *Operations Research* 68.6, pp. 1716–1721. DOI: 10.1287/opre.2019.1944.

Kouvelis, P. and G. Yu (2013). *Robust Discrete Optimization and Its Applications*. Vol. 14. Springer Science & Business Media. DOI: 10.1007/978-1-4757-2620-6.

Leyffer, S., M. Menickelly, T. Munson, C. Vanaret, and S. M. Wild (2020). "A survey of nonlinear robust optimization." In: *INFOR: Information Systems and Operational Research* 58.2, pp. 342–373. DOI: 10.1080/03155986.2020.1730676.

Marandi, A. and D. den Hertog (2018). "When are static and adjustable robust optimization problems with constraint-wise uncertainty equivalent?" In: *Mathematical Programming* 170, pp. 555–568. DOI: 10.1007/s10107-017-1166-z.

Nohadani, O. and K. Sharma (2018). "Optimization under Decision-Dependent Uncertainty." In: *SIAM Journal on Optimization* 28.2, pp. 1773–1795. DOI: 10.1137/17M1110560.

Soyster, A. L. (1973). "Technical Note—Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming." In: *Operations Research* 21.5, pp. 1154–1157. DOI: 10.1287/opre.21.5.1154.

Stein, O. (2013). *Bi-Level Strategies in Semi-Infinite Programming*. Springer New York, NY. DOI: 10.1007/978-1-4419-9164-5.

Tahernejad, S., T. K. Ralphs, and S. T. DeNegre (2020). "A branch-and-cut algorithm for mixed integer bilevel linear optimization problems and its implementation." In: *Mathematical Programming Computation* 12. DOI: 10.1007/s12532-020-00183-6.

Wiesemann, W., A. Tsoukalas, P.-M. Kleniati, and B. Rustem (2013). "Pessimistic Bilevel Optimization." In: *SIAM Journal on Optimization* 23.1, pp. 353–380. DOI: 10.1137/120864015.

Yanıkoğlu, I., B. L. Gorissen, and D. den Hertog (2019). "A survey of adjustable robust optimization." In: *European Journal of Operational Research* 277.3, pp. 799–813. DOI: 10.1016/j.ejor.2018.08.031.

(M. Goerigk) University of Passau, Business Decisions and Data Science, Dr.-Hans-Kapfinger Str. 30, 94032 Passau, Germany
Email address: marc.goerigk@uni-passau.de

(J. Kurtz) Amsterdam Business School, University of Amsterdam, 1018 TV Amsterdam, Netherlands
Email address: j.kurtz@uva.nl

(M. Schmidt, J. Thürauf) Trier University, Department of Mathematics, Universitätsring 15, 54296 Trier, Germany
Email address: martin.schmidt@uni-trier.de, johannes.thuerauf@uni-trier.de