



HAL
open science

Near-optimal robust bilevel optimization

Mathieu Besançon, Miguel F Anjos, Luce Brotcorne

► **To cite this version:**

Mathieu Besançon, Miguel F Anjos, Luce Brotcorne. Near-optimal robust bilevel optimization. 2019. hal-02414848v1

HAL Id: hal-02414848

<https://hal.science/hal-02414848v1>

Preprint submitted on 16 Dec 2019 (v1), last revised 3 Jan 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

NEAR-OPTIMAL ROBUST BILEVEL OPTIMIZATION

MATHIEU BESANÇON*, MIGUEL F. ANJOS†, AND LUCE BROTCORNE ‡

Abstract. Bilevel optimization studies problems where the optimal response to a second mathematical optimization problem is integrated in the constraints. Such structure arises in a variety of decision-making problems in areas such as market equilibria, policy design or product pricing. We introduce near-optimal robustness for bilevel problems, protecting the upper-level decision-maker from bounded rationality at the lower level and show it is a restriction of the corresponding pessimistic bilevel problem. Essential properties are derived in generic and specific settings. This model finds a corresponding and intuitive interpretation in various situations cast as bilevel optimization problems. We develop a duality-based solution method for cases where the lower level is convex, leveraging the methodology from robust and bilevel literature. The models obtained are tested numerically using different solvers and formulations, showing the successful implementation of the near-optimal robust bilevel problem.

Key words. bilevel optimization, robust optimization, game theory, bounded rationality, duality, bilinear constraints, extended formulation

AMS subject classifications. 90C33 90C46 91A65 90C26 90C34

1. Introduction. In this paper, we consider the near-optimal robust versions of bilevel problems *NORBiP*, which we define as:

$$(1.1a) \quad \min_{x,v} F(x, v)$$

s.t.

$$(1.1b) \quad G_k(x, v) \leq 0 \quad \forall k \in \llbracket m_u \rrbracket$$

$$(1.1c) \quad x \in \mathcal{X}$$

$$(1.1d) \quad v \in \arg \min_y \{f(x, v) \text{ s.t. } g(x, v) \leq 0, y \in \mathcal{Y}\}$$

$$(1.1e) \quad G_k(x, z) \leq 0 \quad \forall k \in \llbracket m_u \rrbracket, \forall z \in \mathcal{Z}(x; \delta)$$

$$(1.1f) \quad \text{with } \mathcal{Z}(x; \delta) = \{y \mid g(x, y) \leq 0, y \in \mathcal{Y}, f(x, y) \leq f(x, v) + \delta\}$$

where n_l, n_u are the number of lower- and upper-level variables respectively, m_l, m_u are the number of lower- and upper-level constraints respectively, $\mathcal{X} \subseteq \mathbb{R}^{n_u}$, $\mathcal{Y} \subseteq \mathbb{R}^{n_l}$. We use the notation $\llbracket a \rrbracket \equiv \{1, 2..a\}$ for any natural number $a > 0$. The upper- and lower-level objective functions are noted $F, f : \mathcal{X} \times \mathcal{Y} \mapsto \mathbb{R}$ respectively. Constraint (2.1b) and $g(x, y) \leq 0$ are the upper- and lower-level constraints respectively.

Bilevel optimization problems embed the optimality conditions of a sub-problem into the constraints of a decision-making process. They can efficiently model various decision-making problems such as Stackelberg games, market equilibria, price-setting problems. A recent review of methods and applications of bilevel problems is presented in [11]. In this paper, we consider an extension of bilevel problems where the

*Department of Mathematics & Industrial Engineering, Polytechnique Montréal, Montréal, QC, Canada H3C 3A7 (mathieu.besancon@polymtl.ca).

†School of Mathematics, University of Edinburgh, Edinburgh, Scotland, UK EH9 3FD, and Department of Mathematics and Industrial Engineering, Polytechnique Montréal, Montréal, Quebec, Canada H3C 3A7 (anjos@stanfordalumni.org).

‡INOCS, INRIA Lille-Nord Europe, France (luce.brotcorne@inria.fr).

lower-level is solved to near-optimality.

Solving bilevel problems under limited deviations of the lower-level variables was introduced in [39] under the term “ ε -approximation” of the pessimistic bilevel problem. The authors define special properties and a solution method for this variant in the so-called independent case, *i.e.* where the lower-level feasible set is independent of the upper-level decision. We generalize the approach of [39] to problems involving upper- and lower-level variables in the constraints at both levels. Since this variant consists in protecting the upper-level feasibility against uncertainty of near-optimal solutions of the lower-level, we next use the terms *near-optimal robustness* and *near-optimal robust bilevel problem (NORBiP)* to qualify this extension.

In the context of bilevel problems modelling Stackelberg games, when optimizing their objective function, the leader (upper-level problem) anticipates an optimal reaction of the follower (lower-level problem) to their decisions. However, in many practical cases, the follower makes near-optimal decisions [39], meaning decisions resulting in a limited deviation of the lower-level objective. An important issue in this setting is the definition of the robustness of the leader’s decisions with respect to near-optimal followers ones.

An interpretation for a near-optimal decision of the follower has been developed in game theory as *bounded rationality*. The concept was initially proposed in [36], and is sometimes referred to as ε -rationality [2]. Bounded rationality defines an economic and behavioural interpretation of a decision-making process where an agent aims to take *any* solution associated with a “satisfactory” objective value instead of the optimal value.

In some engineering and control applications using bilevel formulations [10, 37, 28, 13], the decision-maker optimizes an outcome over a system which is modelled as the lower level in a bilevel setting. Lower-level near-optimal solutions find an intuitive interpretation as interruptions near the equilibrium of the system, for instance with friction. The designer of the system would require that the upper-level constraints be feasible for all near-optimal lower-level solutions.

The introduction of robustness in games or more broadly decision-making under uncertainty in game theory has been thoroughly explored in the literature. In [1], the authors prove the existence of robust counterparts of Nash equilibria under standard assumptions and without knowledge of probability distributions associated with uncertainty for simultaneous game. In [29], the robust version of a network congestion problem is developed, considering users as making decisions under bounded rationality and leading to a robust Wardrop equilibrium found via a column generation scheme building path candidates. Robust versions of bilevel problems modelling specific Stackelberg games have been studied in [19, 32], using concepts of robustness to protect the leader against non-rationality or partial rationality of the follower.

The main contributions of the paper are:

1. The generalization of the “ ε -approximation” from [39] to generic bilevel problems as *NORBiP*, resulting in a generalized semi-infinite problem and its interpretation as a special case of robust optimization applied to bilevel problems.

2. The study of duality-based reformulations of *NORBiP* where the lower-level problem is convex and linear, resulting in a finite-dimensional single-level reformulation.
3. An extended formulation for the linear-linear *NORBiP*, linearizing the single-level model with disjunctive constraints.
4. A solution algorithm for the linear-linear *NORBiP* using the extended formulation.

The paper is organized as follows. In [section 2](#), we define the concepts of near-optimal set and near-optimal robust bilevel problem. In [section 3](#) we study special properties of near-optimal robust bilevel problems in the general setting while we consider the cases of convex and linear lower-level problems in [section 4](#) and [section 5](#) respectively. These cases where the follower problems are either convex or linear respectively allow the reformulation of the bilevel problem as a single level problem. A solution algorithm and computational experiments are defined and commented for the linear case in [section 6](#). Finally, in [section 7](#) we draw some conclusions and highlight research perspectives on near-optimal robust bilevel problems.

2. Near-optimal set and near-optimal robust bilevel problem. In this section we first define the near-optimal set of the lower level and their extensions to near-optimal robust bilevel problems. Next we illustrate the concepts on an example. The generic bilevel problem is classically defined as:

$$\begin{aligned}
 (2.1a) \quad & \min_x F(x, v) \\
 & \text{s.t.} \\
 (2.1b) \quad & G_k(x, v) \leq 0 \quad \forall k \in \llbracket m_u \rrbracket \\
 (2.1c) \quad & x \in \mathcal{X} \\
 (2.1d) \quad & v \in \arg \min_{y \in \mathcal{Y}} \{f(x, y) \text{ s.t. } g_i(x, y) \leq 0 \forall i \in \llbracket m_l \rrbracket\}.
 \end{aligned}$$

The upper- and lower-level objective functions are noted $F, f : \mathcal{X} \times \mathcal{Y} \mapsto \mathbb{R}$ respectively. Constraint [\(2.1b\)](#) and $g_i(x, y) \leq 0 \forall i \in \llbracket m_l \rrbracket$ are the upper- and lower-level constraints respectively. In the rest of this section, we assume $\mathcal{Y} = \mathbb{R}^{n_l}$ such that the lower-level feasible set is determined only by the g_i functions. The optimal value function $\phi(x)$ defined as follows is used for compact notation:

$$\begin{aligned}
 (2.2) \quad & \phi : \mathbb{R}^{n_u} \rightarrow \{-\infty\} \cup \mathbb{R} \cup \{+\infty\} \\
 & \phi(x) = \min_y \{f(x, y) \text{ s.t. } g(x, y) \leq 0\}.
 \end{aligned}$$

To keep the notation succinct, the indices of the lower-level constraints g_i are omitted when not needed as in [\(2.2\)](#). Throughout the paper, it is assumed that the lower-level problem is feasible and bounded for any given upper-level decision.

When, for a feasible upper-level decision, the solution to the lower-level problem is not unique, the bilevel problem is not well-defined and further assumptions are required [\[11\]](#). In the optimistic assumption, the lower-level solution selected is the most favourable to the upper-level. This means the upper-level can pick the lower-level solution, as long as it is optimal for the lower level (i.e. that it is bilevel-feasible). The classical pessimistic bilevel model assumes that the lower level picks the optimal solution that is worst for the upper-level. We refer the reader to [\[14, Chapter 1\]](#) for further information on the two assumptions and their implications.

The near-optimal set of the lower level $\mathcal{Z}(x; \delta)$ is defined for a given upper-level decision x and tolerance δ as:

$$\mathcal{Z}(x; \delta) = \{y \text{ s.t. } g(x, y) \leq 0, f(x, y) \leq \phi(x) + \delta\}.$$

The semicolon in $\mathcal{Z}(x; \delta)$ separates decision variables and parameters of the problem. The special case $\mathcal{Z}(x; 0)$ corresponds to the set of optimal solutions to the original lower-level problem and therefore to the pessimistic bilevel problem as formulated in [39]:

$$f(x, y) \leq \phi(x) \quad \forall y \in \mathcal{Z}(x; 0).$$

A Near-Optimal Robust Bilevel Problem, *NORBiP*, of parameter δ is defined as a bilevel problem where the upper-level constraints are respected for any lower-level solution z in the near-optimal set $\mathcal{Z}(x; \delta)$.

$$\begin{aligned} (2.3a) \quad & \min_{x, v} F(x, v) \\ & \text{s.t.} \\ (2.3b) \quad & G_k(x, v) \leq 0 \quad \forall k \in \llbracket m_u \rrbracket \\ (2.3c) \quad & f(x, v) \leq \phi(x) \\ (2.3d) \quad & g(x, v) \leq 0 \\ (2.3e) \quad & G_k(x, z) \leq 0 \quad \forall z \in \mathcal{Z}(x; \delta) \quad \forall k \in \llbracket m_u \rrbracket \\ (2.3f) \quad & x \in \mathcal{X}. \end{aligned}$$

Note that each k constraint in (2.3b) is satisfied if the corresponding constraint set in (2.3e) holds and is therefore redundant, since $v \in \mathcal{Z}(x; \delta)$. However, we mention (2.3b) in the formulation to highlight the structure of the optimistic bilevel problem in the near-optimal robust formulation. Constraint (2.3e) is a generalized semi-infinite constraint, following the terminology from [38]. The dependence of the set of constraints $\mathcal{Z}(x; \delta)$ on the decision variable leads to the characterization of the problem as a robust problem with decision-dependent uncertainty [18]. Each constraint in the set (2.3e) can be replaced by the corresponding worst-case second-level decision z_k obtained by solving the *adversarial problem*, parametrized by (x, v, δ) :

$$\begin{aligned} (2.4a) \quad & z_k \in \arg \max_y G_k(x, y) \\ & \text{s.t.} \\ (2.4b) \quad & f(x, y) \leq \phi(x) + \delta \\ (2.4c) \quad & g(x, y) \leq 0. \end{aligned}$$

Finally, the near-optimal robust bilevel optimization problem can then be expressed

as:

$$\begin{aligned}
(2.5a) \quad & \min_{x,v} F(x,v) \\
& \text{s.t.} \\
(2.5b) \quad & f(x,v) \leq \phi(x) \\
(2.5c) \quad & g(x,v) \leq 0 \\
(2.5d) \quad & 0 \geq \max_y \{G_k(x,y) \text{ s.t. } g(x,y) \leq 0, f(x,y) \leq f(x,v) + \delta\} \quad \forall k \in \llbracket m_u \rrbracket. \\
(2.5e) \quad & x \in \mathcal{X}.
\end{aligned}$$

The literature on robust optimization distinguishes uncertainty on constraints and on the objective function [16]. The first class of problems corresponds to remaining feasible for any value of the uncertain parameter, the second corresponds to optimizing for the worst case, with respect to the objective, of the uncertain parameter. In the case of bilevel optimization, the first case corresponds to *NORBiP*, studying the impact of near-optimal lower-level solutions on the upper-level constraints. The second case corresponds to the impact of near-optimal lower-level decision on the upper-level objective value. We next show the model including uncertainty on the objective, named Objective-Robust Near-Optimal Robust Bilevel Problem (*ORNOBiP*), is a special case of *NORBiP*.

ORNOBiP is defined as follows:

$$\begin{aligned}
(2.6a) \quad & \min_{x \in \mathcal{X}} \sup_{z \in \mathcal{Z}(x;\delta)} F(x,z) \\
& \text{s.t.} \\
(2.6b) \quad & \mathcal{Z}(x;\delta) = \{y \text{ s.t. } g(x,y) \leq 0, f(x,y) \leq \phi(x) + \delta\}.
\end{aligned}$$

In contrast with most objective-robust problem formulations, the uncertainty set \mathcal{Z} depends on the upper-level solution x , qualifying (2.6) as a problem with decision-dependent uncertainty.

PROPOSITION 2.1. *ORNOBiP is a special case of NORBiP.*

Proof. The reduction of the objective-uncertain robust problem to a constraint-uncertain robust formulation is detailed in [4]. In particular, Problem (2.6) is equivalent to:

$$\begin{aligned}
& \min_{x,\tau} \tau \\
& \text{s.t.} \\
& x \in \mathcal{X} \\
& \tau \geq F(x,z) \quad \forall z \in \mathcal{Z}(x,\delta),
\end{aligned}$$

this formulation is a special case of *NORBiP*. □

PROPOSITION 2.2. *The pessimistic bilevel optimization problem as formulated in [25] is both a special case and a relaxation of ORNOBiP.*

Proof. With the special case $\delta = 0$, the inner problem of *ORNOBiP* is equivalent to finding the worst lower-level decision with respect to the upper-level objective

amongst the lower-level-optimal solutions. For any $\delta > 0$, the inner problem can select worse or equal solutions with respect to the upper-level. The pessimistic bilevel problem is therefore a relaxation of *ORNOBiP*. \square

We next illustrate the concept of near-optimal set and near-optimal robust solution on an example. Let us consider the following linear bilevel problem:

$$\begin{aligned}
 (2.7a) \quad & \min_{x,v} x \\
 & \text{s.t.} \\
 (2.7b) \quad & x \geq 0 \\
 (2.7c) \quad & v \geq 1 - \frac{x}{10} \\
 (2.7d) \quad & v \in \arg \max_y \{y \text{ s.t. } y \leq 1 + \frac{x}{10}\}.
 \end{aligned}$$

A graphical representation of the problem is provided in [Figure 1](#). The high-point relaxation of a bilevel problem is obtained by relaxing the optimality constraint of the lower-level, while maintaining feasibility. The high-point relaxation of Problem (2.7) is:

$$\begin{aligned}
 & \min_{x,v} x \\
 & \text{s.t.} \\
 & x \geq 0 \\
 & v \geq 1 - \frac{x}{10} \\
 & v \leq 1 + \frac{x}{10}.
 \end{aligned}$$

The shaded area represents the interior of the polytope which is feasible for the high-point relaxation. The induced set, obtained by taking into account the optimal lower-level reaction, is given by:

$$\{(x, y) \in (\mathbb{R}_+, \mathbb{R}) \text{ s.t. } y = 1 + \frac{x}{10}\},$$

the unique optimal point is $(\hat{x}, \hat{y}) = (0, 1)$.

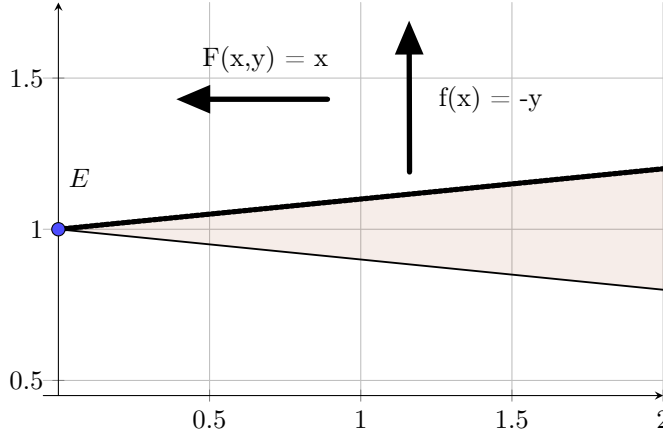


FIG. 1. *Linear bilevel problem*

Consider now a near-optimal tolerance of the follower with $\delta = 0.1$. If the upper-level decision is \hat{x} , the lower level can take any value between $1 - \delta = 0.9$ and 1 leading to infeasible upper-level except for 1. The problem can be reformulated as:

$$\begin{aligned}
 & \min_{x,v} x \\
 & \text{s.t.} \\
 & x \geq 0 \\
 & v \geq 1 - \frac{x}{10} \\
 & v \in \arg \max_y \{y \text{ s.t. } y \leq 1 + \frac{x}{10}\} \\
 & z \geq 1 - \frac{x}{10} \quad \forall z \text{ s.t. } \{z \leq 1 + \frac{x}{10}, z \geq v - \delta\}.
 \end{aligned}$$

Figure 2 illustrates the near-optimal equivalent of the problem.

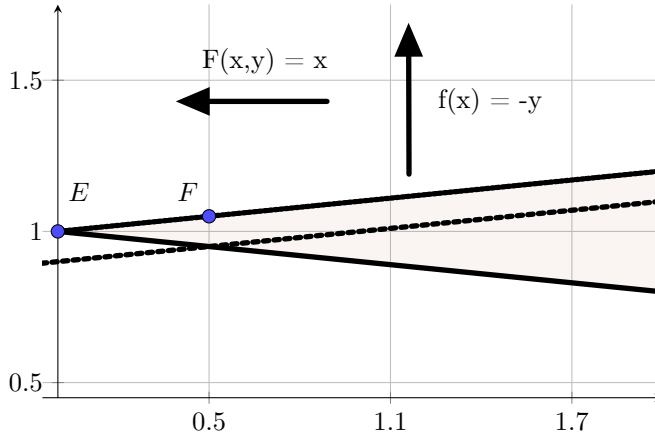


FIG. 2. *Linear bilevel problem with a near-optimality robustness constraint*

The dashed line represents the constraint of robustness to near-optimality. The optimal upper-level decision is $x = 0.5$, for which the optimal lower-level reaction is $y = 1 + 0.1 \cdot 0.5 = 1.05$. The boundary of the near-optimal set is $y = 1 - 0.1 \cdot 0.5 = 0.95$.

3. Special properties of near-optimal robust bilevel optimization problems. In this section, we define some properties of the near-optimal robust problem.

PROPOSITION 3.1. *If the second-level optimization problem is convex, then the near-optimal set $\mathcal{Z}(x; \theta)$ is convex.*

Proof. $\mathcal{Z}(x; \delta)$ is the intersection of two sets:

- $\mathcal{F} = \{y, g(x, y) \leq 0\}$
- $\mathcal{N} = \{y, f(x, y) \leq \phi(x) + \delta\}$

\mathcal{F} is the intersection of sublevel sets of convex functions g_i , \mathcal{N} is the sublevel set of a convex function f ; hence both are convex sets [7]. Thus $\mathcal{Z}(x; \delta)$ being the intersection of two convex sets is also convex (possibly empty). \square

In robust optimization, the characteristics of the uncertainty set sharply influences the difficulty of solving the problem. The near-optimal set of the lower-level is not systematically bounded; this can lead to infeasible or ill-defined near-optimal robust counterparts of bilevel problems. In the next proposition we define conditions under which the uncertainty set $\mathcal{Z}(x; \delta)$ is bounded.

PROPOSITION 3.2. *For a given pair (x, δ) , any of the following properties is sufficient for $\mathcal{Z}(x; \delta)$ to be a bounded set:*

1. *The lower-level feasible domain is bounded.*
2. *$f(x, \cdot)$ is radially unbounded with respect to y .*
3. *$f(x, \cdot)$ is radially bounded, such that:*

$$\lim_{r \in \mathbb{R}, r \rightarrow +\infty} f(x, rs) > f(x, v) + \delta \quad \forall s \in \mathcal{S},$$

With \mathcal{S} the unit sphere in the space of lower-level decisions.

Proof. The first case is trivially satisfied since $\mathcal{Z}(x; \delta)$ is the intersection of sets including the lower-level feasible set. If $f(x, \cdot)$ is radially unbounded, for any finite $\delta > 0$, there is a maximum radius around v beyond which any value of the objective function is greater than $f(x, v) + \delta$. The third case follows the same line of reasoning as the second, with a lower bound in any direction $\|y\| \rightarrow \infty$, such that this lower bound is above $f(x, v) + \delta$. \square

The radius of robust feasibility is defined as the maximum “size” of the uncertain set [24, 26]. In the case of near-optimal robustness, the radius can be interpreted as the maximum deviation of the lower-level objective from its optimal value.

DEFINITION 3.3. *For a given optimization problem BiP , let $\mathcal{NO}(BiP; \delta)$ be the optimum value of the near-optimal robust problem constructed from BiP with a tolerance δ . The radius of near-optimal feasibility $\hat{\delta}$ is defined by:*

$$(3.1a) \quad \hat{\delta} = \arg \max_{\delta} \delta$$

s.t.

$$(3.1b) \quad \mathcal{NO}(BiP; \delta) < \infty.$$

PROPOSITION 3.4. *The standard optimistic bilevel problem BiP is a relaxation of the equivalent near-optimal robust bilevel problem for any $\delta > 0$.*

Proof. Additional variables $z_i, i \in \llbracket m_i \rrbracket$ are introduced in the optimistic bilevel problem, resulting in the following model:

$$\begin{aligned} & \min_{x,v,z} F(x,v) \\ & \text{s.t.} \\ & G_k(x,v) \leq 0 \quad \forall k \in \llbracket m_u \rrbracket \\ & f(x,v) \leq \phi(x) \\ & g(x,v) \leq 0 \\ & x \in \mathcal{X}. \end{aligned}$$

This model is strictly equivalent to the optimistic bilevel problem and is a relaxation of Problem (2.5), which is additionally restricted by constraint (2.5d). Furthermore, at each point where the bilevel problem is feasible, the objective value of the two problems are the same or NORBiP is infeasible. \square

PROPOSITION 3.5. *The pessimistic bilevel problem formulated in [39] as:*

$$\begin{aligned} & \min_{x \in \mathcal{X}} F(x) \\ & G(x,z) \leq 0 \quad \forall z \in \mathcal{Z}(x) = \underset{y}{\operatorname{argmin}} \{f(x,y), y \in Y(x)\} \end{aligned}$$

is both a relaxation and a special case of NORBiP. $Y(x)$ is the feasible set of the lower-level problem, depending on the upper-level decision x .

Proof. For $\delta = 0$, NORBiP can be re-written as:

$$\begin{aligned} & \min_{x \in \mathcal{X}} F(x,v) \\ & f(x,v) \leq \min_y \{f(x,y), y \in Y(x)\} \\ & v \in Y(x) \\ & G(x,v) \leq 0 \\ & G(x,z) \leq 0 \quad \forall z \in \mathcal{Z}(x; \delta = 0) \\ & \mathcal{Z}(x; \delta = 0) = \{y, f(x,y) \leq f(x,v) + 0, y \in Y(x)\} \\ & g(x,y) \leq 0 \Leftrightarrow y \in Y(x). \end{aligned}$$

For any $\delta > 0$, the feasible domain of the adversarial problem allows for worse-case values with respect to the upper-level constraints to be taken at the lower-level decision, for which the upper-level optimal value can only be worse than the optimistic formulation. \square

PROPOSITION 3.6. *If the bilevel problem is feasible, then the adversarial problem (2.4) is feasible.*

Proof. If the bilevel problem is feasible, then the solution $z = v$ is feasible for the primal adversarial problem. \square

PROPOSITION 3.7. *If G_k is K_k -Lipschitz continuous for a given $k \in \llbracket m_u \rrbracket$, and (\hat{x}, \hat{y}) is a bilevel-feasible point, such that:*

$$G_k(\hat{x}, \hat{y}) < 0,$$

then the constraint $G_k(\hat{x}, y) \leq 0$ is satisfied for all $y \in \mathcal{F}_L^{(k)}$ such that:

$$\mathcal{F}_L^{(k)}(\hat{x}, \hat{y}) = \{y \in \mathbb{R}^{n_l} \text{ s.t. } \|y - \hat{y}\| \leq \frac{|G_k(\hat{x}, \hat{y})|}{K_k}\}.$$

Proof. As $G_k(\hat{x}, \hat{y}) < 0$, and $G_k(\hat{x}, \cdot)$ is continuous, there exists a ball $\mathcal{B}_r(\hat{y})$ in \mathbb{R}^{n_l} centered on (\hat{y}) of radius $r > 0$, such that

$$G(\hat{x}, y) \leq 0 \quad \forall y \in \mathcal{B}_r(\hat{y}).$$

Let us define:

$$\begin{aligned} r_0 &= \arg \max_r r \\ &\text{s.t.} \\ &G(\hat{x}, y) \leq 0 \quad \forall y \in \mathcal{B}_r(\hat{y}). \end{aligned}$$

By continuity, this problem always admits a feasible solution. If it is bounded, there exists a point y_0 on the boundary of the ball, such that $G_k(\hat{x}, y_0) = 0$. It follows from the Lipschitz continuity property that:

$$\begin{aligned} |G_k(\hat{x}, \hat{y}) - G_k(\hat{x}, y_0)| &\leq K_k \|y_0 - \hat{y}\| \\ \frac{|G_k(\hat{x}, \hat{y})|}{K_k} &\leq \|y_0 - \hat{y}\|. \end{aligned}$$

$G_k(\hat{x}, y) \leq G_k(\hat{x}, y_0) \quad \forall y \in \mathcal{B}_{r_0}(\hat{y})$, therefore all lower-level solutions in the set:

$$\mathcal{F}_L^{(k)}(\hat{x}, \hat{y}) = \{y \in \mathbb{R}^{n_l} \text{ s.t. } \|y - \hat{y}\| \leq \frac{|G_k(\hat{x}, \hat{y})|}{K_k}\}$$

satisfy constraint k . □

COROLLARY 3.8. *Let (\hat{x}, \hat{y}) be a bilevel-feasible solution of a near-optimal robust bilevel problem of tolerance δ , and*

$$\mathcal{F}_L(\hat{x}, \hat{y}) = \bigcap_{k=1}^{m_u} \mathcal{F}_L^{(k)}(\hat{x}, \hat{y}),$$

then $\mathcal{Z}(x; \delta) \subseteq \mathcal{F}_L(\hat{x}, \hat{y})$ is a sufficient condition for near-optimal robustness of (\hat{x}, \hat{y}) .

Proof. Any lower-level solution $y \in \mathcal{F}_L(\hat{x}, \hat{y})$ satisfies all m_u upper-level constraints, $\mathcal{Z}(x; \delta) \subseteq \mathcal{F}_L(\hat{x}, \hat{y})$ is therefore a sufficient condition for the near-optimality robustness of (\hat{x}, \hat{y}) . □

4. Near-optimal robust bilevel problems with a convex lower level.

In this section, we study near-optimal robust bilevel problems where the lower-level problem (2.1d) is a parametric convex optimization problem with both a differentiable objective function and differentiable constraints. If Slater's constraint qualifications hold, the Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient for the optimality of the lower-level problem and strong duality holds for the adversarial sub-problems. These two properties are leveraged to reformulate *NORBIP* as a one-level closed-form problem.

Given a bilevel solution (x, v) , the adversarial problem associated with constraint k can be formulated as:

$$\begin{aligned}
(4.1a) \quad & \max_y G_k(x, y) \\
& \text{s.t.} \\
(4.1b) \quad & g(x, y) \leq 0 \\
(4.1c) \quad & f(x, y) \leq f(x, v) + \delta.
\end{aligned}$$

Even if the upper-level constraints are convex with respect to y , Problem (4.1) is in general not convex since the function to maximize is convex over a convex set. First-order optimality conditions may thus yield several critical points and the definition of a solution method needs to rely on global optimization techniques [30, 3].

By assuming that the constraints of the upper-level problem $G_k(x, y)$ are decomposable and that the projection onto the lower variable space is affine, the adversarial problem becomes convex:

$$(4.2) \quad G_k(x, y) \leq 0 \Leftrightarrow G_k(x) + H_k^T y \leq q_k.$$

The k -th adversarial problem is then expressed as:

$$\begin{aligned}
(4.3a) \quad & \max_y H_k^T y \\
& \text{s.t.} \\
(4.3b) \quad & g_i(x, y) \leq 0 \quad \forall i \in \llbracket m_l \rrbracket \quad (\alpha_i) \\
(4.3c) \quad & f(x, y) \leq f(x, v) + \delta \quad (\beta)
\end{aligned}$$

and is convex for a fixed pair (x, v) . Satisfying the upper-level constraint in the worst-case requires that the objective value of Problem (4.3) be lower than $q_k - G_k(x)$. We denote by \mathcal{A}_k and \mathcal{D}_k the objective value of the adversarial problem (4.3) and its dual or *dual adversarial problem*. \mathcal{D}_k takes values in the extended-real set to account for infeasible and unbounded cases. Proposition 3.6 holds for Problem (4.3). The feasibility of the upper-level constraint with the dual adversarial objective value as formulated in (4.4) is by weak duality of convex problems a sufficient condition for the feasibility of the near-optimal solution. If Slater's constraint qualifications hold, it is also a necessary condition due to strong duality [7].

$$(4.4) \quad \mathcal{A}_k \leq \mathcal{D}_k \leq q_k - G_k(x)$$

The generic form for the single-level reformulation of the near-optimal robust problem can then be expressed as:

$$\begin{aligned}
(4.5a) \quad & \min_{x, v, \alpha, \beta} F(x, v) \\
(4.5b) \quad & \text{s.t.} \\
(4.5c) \quad & G(x) + Hv \leq q \\
(4.5d) \quad & f(x, v) \leq \phi(x) \\
(4.5e) \quad & g(x, v) \leq 0 \\
(4.5f) \quad & \mathcal{D}_k \leq q_k - G_k(x) \quad \forall k \in \llbracket m_u \rrbracket \\
(4.5g) \quad & x \in \mathcal{X}.
\end{aligned}$$

(α, β) are certificates for the near-optimality robustness of the solution. In order to write Problem (4.5) in a closed form, the lower-level problem (4.5d)–(4.5e) is reduced to its KKT conditions:

$$\begin{aligned}
(4.6a) \quad & \nabla_v f(x, v) - \sum_{i=1}^{m_l} \lambda_i \cdot \nabla_v g_i(x, v) = 0 \\
(4.6b) \quad & g_i(x, v) \leq 0 \quad \forall i \in \llbracket m_l \rrbracket \\
(4.6c) \quad & \lambda_i \geq 0 \quad \forall i \in \llbracket m_l \rrbracket \\
(4.6d) \quad & \lambda_i \cdot g_i(x, v) = 0 \quad \forall i \in \llbracket m_l \rrbracket.
\end{aligned}$$

This direct formulation of the KKT conditions cannot be tackled by non-linear solvers [12] and requires transformations or specific techniques, such as relaxations of the equality constraints in (4.6d) into inequalities as developed in [35, 34] or branching on combinations of variables allowed to take non-zero values.

We consider in the rest of this section problems where the lower-level is a conic convex optimization problem:

$$\begin{aligned}
(4.7a) \quad & \min_y \langle d, y \rangle \\
(4.7b) \quad & \text{s.t.} \\
(4.7c) \quad & Ax + By = b \\
(4.7d) \quad & y \in \mathcal{K}
\end{aligned}$$

where $\langle \cdot, \cdot \rangle$ is the inner product associated with the space of the lower-level variables. This class encompasses a broad class of convex optimization problems of practical interest [27, Chapter 4], while the dual problem can be written in a closed-form if the dual cone is known. \mathcal{K} is considered to be a proper cone in the sense of [7, Chapter 2]. The k -th adversarial problem is given by:

$$\begin{aligned}
(4.8a) \quad & \max_y \langle H_k, y \rangle \\
(4.8b) \quad & \text{s.t.} \\
(4.8c) \quad & By = b - Ax \\
(4.8d) \quad & \langle d, y \rangle + r = \langle d, v \rangle + \delta \\
(4.8e) \quad & y \in \mathcal{K} \\
(4.8f) \quad & r \geq 0
\end{aligned}$$

where r is a slack variable used to formulate the near-optimality constraint in standard form. With the following change of variables:

$$\hat{y} = \begin{bmatrix} y \\ r \end{bmatrix} \quad \hat{B} = \begin{bmatrix} B & 0 \end{bmatrix} \quad \hat{d} = \begin{bmatrix} d & 1 \end{bmatrix} \quad \hat{H}_k = \begin{bmatrix} H_k \\ 0 \end{bmatrix}$$

$$\hat{\mathcal{K}} = \{(y, r), y \in \mathcal{K}, r \geq 0\}.$$

$\hat{\mathcal{K}}$ is a cone as the Cartesian product of \mathcal{K} and the nonnegative orthant. Problem

(4.8) is reformulated as:

$$\begin{aligned}
& \max_{\hat{y}} \hat{H}_k^T \hat{y} \\
& \text{s.t.} \\
& (\hat{B}\hat{y})_i = b_i - (Ax)_i \quad \forall i \in \llbracket m_l \rrbracket \quad (\alpha_i) \\
& \langle \hat{d}, \hat{y} \rangle = \langle d, v \rangle + \delta \quad (\beta) \\
& \hat{y} \in \hat{\mathcal{K}}
\end{aligned}$$

which is a conic optimization problem, for which the dual problem is:

$$\begin{aligned}
(4.9a) \quad & \min_{\alpha, \beta, s_k} \sum_{i=1}^{m_L} (b - Ax)_i \alpha_i + (\langle d, v \rangle + \delta) \beta \\
(4.9b) \quad & \text{s.t.} \\
(4.9c) \quad & \hat{B}^T \alpha + \beta \hat{d} + s = \hat{H}_k \\
(4.9d) \quad & s \in \hat{\mathcal{K}}^*,
\end{aligned}$$

with $\hat{\mathcal{K}}^*$ the dual cone of $\hat{\mathcal{K}}$. In the worst case (maximum number of non-zero coefficients), there are $(m_l \cdot n_u + n_l)$ of these terms in m_u non-linear non-convex constraints. This number of bilinear terms can be reduced by introducing the following variables (p, o) , along with the corresponding constraints:

$$\begin{aligned}
(4.10a) \quad & \min_{\alpha, \beta, s, p, o} \sum_{i=1}^{m_L} p_i \alpha_i + (o + \delta) \beta \\
& \text{s.t.} \\
(4.10b) \quad & p = b - Ax \\
(4.10c) \quad & o = \langle d, v \rangle \\
(4.10d) \quad & \hat{B}^T \alpha + \beta \hat{d} + s = \hat{H}_k \\
(4.10e) \quad & s \in \hat{\mathcal{K}}^*.
\end{aligned}$$

The number of bilinear terms in the set of constraints is thus reduced from $n_u \cdot m_l + n_l$ to $m_l + 1$ terms in (4.10a). Problem (4.9) or equivalently Problem (4.10) have a convex feasible set but a bilinear non-convex objective function. The KKT conditions of the follower problem (4.7a)–(4.7d) are given for the primal-dual pair (x, λ) :

$$\begin{aligned}
(4.11a) \quad & By = b - Ax \\
(4.11b) \quad & y \in \mathcal{K} \\
(4.11c) \quad & d - B^T \lambda \in \mathcal{K}^* \\
(4.11d) \quad & \langle d - B^T \lambda, y \rangle = 0.
\end{aligned}$$

The upper-level problem is thus expressed as:

$$\begin{aligned}
(4.12a) \quad & \min_{x,v,\lambda,\alpha,\beta,s} F(x,v) \\
& \text{s.t.} \\
(4.12b) \quad & G(x) + Hv \leq q \\
(4.12c) \quad & Ax + Bv = b \\
(4.12d) \quad & d - B^T \lambda \in \mathcal{K}^* \\
(4.12e) \quad & \langle d - B^T \lambda, v \rangle = 0 \\
(4.12f) \quad & \sum_{i=1}^{m_L} (Ax - b)_i \cdot \alpha_{ik} + \beta_k \cdot (\langle v, d \rangle + \delta) \leq q_k - (Gx)_k \quad \forall k \in \llbracket m_u \rrbracket \\
(4.12g) \quad & \hat{B}^T \alpha_k + \hat{d} \cdot \beta_k + s_k = \hat{H}_k \quad \forall k \in \llbracket m_u \rrbracket \\
(4.12h) \quad & x \in \mathcal{X}, v \in \mathcal{K} \\
(4.12i) \quad & s_k \in \hat{\mathcal{K}}^* \quad \forall k \in \llbracket m_u \rrbracket.
\end{aligned}$$

The Mangasarian-Fromovitz constraint qualification is violated at every feasible point of Constraint (4.12e) [33]. In non-linear approaches to complementarity constraints [35, 12], relaxed versions parametrized by a tolerance on the complementarity constraint are computed:

$$\begin{aligned}
(4.13a) \quad & \langle d - B^T \lambda, v \rangle \leq \varepsilon \\
(4.13b) \quad & -\langle d - B^T \lambda, v \rangle \leq \varepsilon.
\end{aligned}$$

For $\varepsilon > 0$, the Mangasarian-Fromovitz constraint qualification is respected. A sequence of solutions of the relaxed problem with $\varepsilon \rightarrow 0$ converges to a stationary point of the initial problem. Constraints (4.12f) and (4.13) are both bilinear non-convex inequalities, the other constraints added by the near-optimal robust model are conic and linear constraints. Near-optimal robustness has thus not changed the complexity class of the solution method proposed in [12].

5. Linear near-optimal robust bilevel problem. In this section, we consider a special case of the convex case presented in section 4. More precisely, we study the near-optimal robust version of the linear-linear bilevel problem. The structure of the lower-level problem is exploited to derive an extended formulation leading to an efficient solution algorithm. In the linear case, the inner product of two vectors in \mathbb{R}^n , $\langle a, b \rangle$ is equivalently written $a^T b$.

The near-optimal robust linear bilevel problem is formulated as:

$$\begin{aligned}
(5.1a) \quad & \min_{x,v} c_x^T x + c_y^T v \\
& \text{s.t.} \\
(5.1b) \quad & Gx + Hv \leq q \\
(5.1c) \quad & d^T v \leq \phi(x) \\
(5.1d) \quad & Ax + Bv \leq b \\
(5.1e) \quad & Gx + Hz \leq q \quad \forall z \in \mathcal{Z}(x; \delta) \\
(5.1f) \quad & v \in \mathbb{R}_+^{m_l} \\
(5.1g) \quad & x \in \mathcal{X}.
\end{aligned}$$

For a given pair (x, v) , each semi-infinite robust constraint (5.1e) can be reformulated as the objective value of the following adversarial problem:

$$\begin{aligned}
(5.2a) \quad & \max_y H_k^T y \\
& \text{s.t.} \\
(5.2b) \quad & (By)_i \leq b_i - (Ax)_i \quad \forall i \in \llbracket m_i \rrbracket \quad (\alpha_i) \\
(5.2c) \quad & d^T y \leq d^T v + \delta \quad (\beta) \\
(5.2d) \quad & y \in \mathbb{R}_+^{n_i}.
\end{aligned}$$

Let (α, β) be the dual variables associated with each group of constraints (5.2b)–(5.2c). The near-optimal robustness version of Problem (5.1) is feasible only if the objective value of each k -th adversarial subproblem (5.2) is lower than $q_k - (Gx)_k$. The dual of problem (5.2) is defined as:

$$\begin{aligned}
(5.3a) \quad & \min_{\alpha, \beta} \alpha^T (b - Ax) + \beta \cdot (d^T v + \delta) \\
& \text{s.t.} \\
(5.3b) \quad & B^T \alpha + \beta d \geq H_k \\
(5.3c) \quad & \alpha \in \mathbb{R}_+^{m_i} \beta \in \mathbb{R}_+.
\end{aligned}$$

Based on Proposition 3.6 and weak duality results, the dual problem is either infeasible or feasible and bounded. By strong duality, the objective value of the dual and primal problems are equal. This value must be smaller than $q_k - (Gx)_k$ to satisfy constraint (5.1e). This is equivalent to the existence of a feasible dual solution (α, β) certifying the feasibility of (x, v) within the near-optimal set $\mathcal{Z}(x; \delta)$. We obtain one pair of certificates for each upper-level constraint in $\llbracket m_u \rrbracket$, resulting in the following problem:

$$\begin{aligned}
(5.4a) \quad & \min_{x, v, \alpha, \beta} c_x^T x + c_y^T v \\
& \text{s.t.} \\
(5.4b) \quad & Gx + Hv \leq q \\
(5.4c) \quad & d^T v \leq \phi(x) \\
(5.4d) \quad & Ax + Bv \leq b \\
(5.4e) \quad & \alpha_k^T (b - Ax) + \beta_k \cdot (d^T v + \delta) \leq q_k - (Gx)_k \quad \forall k \in \llbracket m_u \rrbracket \\
(5.4f) \quad & B^T \alpha_k + \beta_k d \geq H_k \quad \forall k \in \llbracket m_u \rrbracket \\
(5.4g) \quad & \alpha_k \in \mathbb{R}_+^{m_i} \beta_k \in \mathbb{R}_+ \quad \forall k \in \llbracket m_u \rrbracket \\
(5.4h) \quad & v \in \mathbb{R}_+^{n_i} \\
(5.4i) \quad & x \in \mathcal{X}.
\end{aligned}$$

The number of bilinear terms can be reduced by introducing variables (p, o) as detailed in section 4. Lower-level optimality is guaranteed by the corresponding Karush-Kuhn-Tucker conditions:

$$\begin{aligned}
(5.5a) \quad & d_j + \sum_i B_{ij} \lambda_i - \sigma_j = 0 \quad \forall j \in \llbracket n_i \rrbracket \\
(5.5b) \quad & 0 \leq b_i - (Ax)_i - (Bv)_i \perp \lambda_i \geq 0 \quad \forall i \in \llbracket m_i \rrbracket \\
(5.5c) \quad & 0 \leq v_j \perp \sigma_j \geq 0 \quad \forall j \in \llbracket n_i \rrbracket \\
(5.5d) \quad & \sigma \geq 0, \lambda \geq 0
\end{aligned}$$

where \perp defines a complementarity constraint. Constraints (5.5b)–(5.5c) are non-linear, non-convex constraints for which every feasible point violates the Mangasarian–Fromovitz constraint qualification [33]. A common technique to linearize them is the “big-M” reformulation, introducing auxiliary binary variables with primal and dual upper bounds. The resulting formulation has a weak continuous relaxation. Furthermore, the correct choice of bounds is itself a NP-hard problem [20], and the introduction of these bounds can lead to cutting valid and potentially optimal solutions [31]. Other modelling and solution approaches, such as special ordered sets of type 1 (SOS1) or indicator constraints avoid the need of specifying bounds.

The aggregated formulation of the near-optimal robust bilevel problem is:

(5.6a)

$$\min_{x,v,\lambda,\sigma,\alpha,\beta} c_x^T x + c_y^T v$$

s.t.

$$(5.6b) \quad Gx + Hv \leq q$$

$$(5.6c) \quad Ax + Bv \leq b$$

$$(5.6d) \quad d_j + \sum_i \lambda_i B_{ij} - \sigma_j = 0 \quad \forall j \in \llbracket n_l \rrbracket$$

$$(5.6e) \quad 0 \leq \lambda_i \perp A_i x + B_i v - b_i \leq 0 \quad \forall i \in \llbracket m_l \rrbracket$$

$$(5.6f) \quad 0 \leq \sigma_j \perp v_j \geq 0 \quad \forall j \in \llbracket n_l \rrbracket \llbracket n_l \rrbracket$$

$$(5.6g) \quad x \in \mathcal{X}$$

$$(5.6h) \quad \sum_{i=1}^{m_l} \alpha_{ki} (b - Ax)_i + \beta_k \cdot (d^T v + \delta) \leq q_k - (Gx)_k \quad \forall k \in \llbracket m_u \rrbracket$$

$$(5.6i) \quad \sum_{i=1}^{m_l} B_{ij} \alpha_{ki} + \beta_k d_j \geq H_{kj} \quad \forall k \in \llbracket m_u \rrbracket, \forall j \in \llbracket n_l \rrbracket$$

$$(5.6j) \quad \alpha_k \in \mathbb{R}_+^{m_l}, \beta_k \in \mathbb{R}_+ \quad \forall k \in \llbracket m_u \rrbracket.$$

Constraints (5.6h) contain bilinear terms, for which several solution approaches have been developed. These include linear inequalities iteratively tightened using variable bounds or mixed-integer formulations [9, 8, 21].

Regarding the complementarity constraints, one solution approach is to apply the same relaxation from $\langle a, b \rangle = 0$ to $\langle a, b \rangle \leq t$ as in section 4, yielding a sequence of non-linear problems converging to a local minimum. Another method is developed in the following subsection, exploiting properties of the dual feasible space of the lower level to construct a completely linear formulation.

Extended formulation. The bilinear constraints involve products of variables from the optimistic formulation (x, v) with dual variables of each of the k dual-adversarial problems. For fixed values of (x, v) , m_u dual adversarial sub-problems (5.3) are defined. For each k of these sub-problems, the optimal value must be lower than $q_k - (Gx)_k$. The feasible region of each sub-problem is defined by (5.3b)–(5.3c) and is independent of (x, v) . The objective functions are linear in (α, β) . Following Proposition 3.6, Problem (5.3) is bounded; if it is feasible, a vertex of the polytope (5.3b)–(5.3c) is an optimal solution. Constraints (5.6h)–(5.6j) can be replaced by dis-

junctive constraints, such that for each constraint (k), at least one extreme vertex is feasible. Let \mathcal{V}_k be the number of vertices of the k -th sub-problem and α_k^l, β_k^l be the l -th vertex of the k -th sub-problem. The constraints (5.6h)–(5.6j) can be written as:

$$(5.7) \quad \bigvee_{l=1}^{\mathcal{V}_k} \sum_{i=1}^{m_l} \alpha_{ki}^l (b - Ax)_i + \beta_k^l \cdot (d^T v + \delta) \leq q_k - (Gx)_k \quad \forall k \in \llbracket m_u \rrbracket$$

where $\bigvee_{i=1}^N \mathcal{C}_i$ is the disjunction (logical “OR”) operator, expressing the constraint that at least one of the constraints \mathcal{C}_i must be respected. These disjunctions are equivalent to indicator constraints [6].

This reformulation of bilinear constraints based on the polyhedral description of the (α, β) feasible space is similar to the Benders decomposition paradigm. In the near-optimal robust extended formulation, at least one of the vertices must satisfy a constraint while Benders decomposition consists in satisfying a set of constraints for all extreme vertices and rays of the dual polyhedron. Disjunctive constraints (5.7) are equivalent to the following formulation, using a set cover constraint and special ordered sets of type 1:

$$(5.8a) \quad \theta_k^l \in \mathbb{B} \quad \forall k, \forall l$$

$$(5.8b) \quad \omega_k^l \geq 0 \quad \forall k, \forall l$$

$$(5.8c) \quad (b - Ax)^T \alpha_k^l + \beta_k^l (d^T v + \delta) - \omega_k^l \leq q_k - (Gx)_k \quad \forall k, \forall l$$

$$(5.8d) \quad \sum_{l=1}^{\mathcal{V}_k} \theta_k^l \geq 1 \quad \forall k$$

$$(5.8e) \quad \text{SOS1}(\theta_k^l, \omega_k^l) \quad \forall k \forall l.$$

In conclusion, using disjunctive constraints over the extreme vertices of each dual polyhedron, along with SOS1 constraints for the complementarity constraints of the lower-level KKT conditions leads to an exact reformulation of Problem (5.6) that can be solved by MILP solvers with standard branch-and-cut algorithms. Although the SOS1-based formulation (5.8) is equivalent to the indicator constraints, the latter enable MILP solvers to separate more inequalities in the branch-and-cut process¹.

Bounded example. An example is provided to illustrate the solution concept derived in this section. Consider the bilevel linear problem defined by the following data:

$$x \in \mathbb{R}_+, y \in \mathbb{R}_+$$

$$G = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad H = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \quad q = \begin{bmatrix} 11 \\ 13 \end{bmatrix} \quad c_x = [1] \quad c_y = [-10]$$

$$A = \begin{bmatrix} -2 \\ 5 \end{bmatrix} \quad B = \begin{bmatrix} -1 \\ -4 \end{bmatrix} \quad b = \begin{bmatrix} -5 \\ 30 \end{bmatrix} \quad d = [1]$$

¹SCIP documentation on indicator constraints.

The optimal solution of the high-point relaxation is $(x, v) = (5, 4)$ which is not bilevel-feasible. The optimal value of the optimistic bilevel problem is at $(x, v) = (1, 3)$. These two points are respectively represented by the blue diamond and red cross on [Figure 3](#). The dotted segments represent the upper-level constraints and the solid lines represent the lower-level constraints.

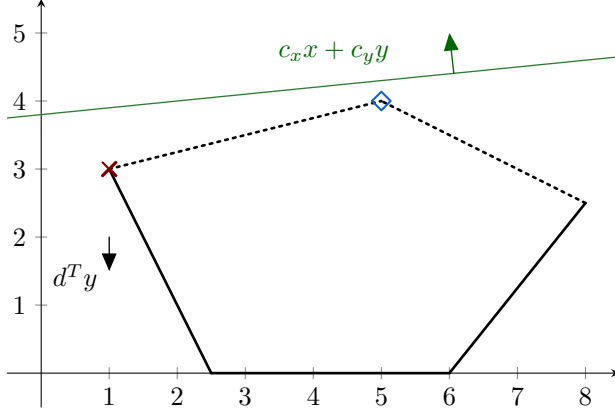


FIG. 3. Representation of the bilevel problem.

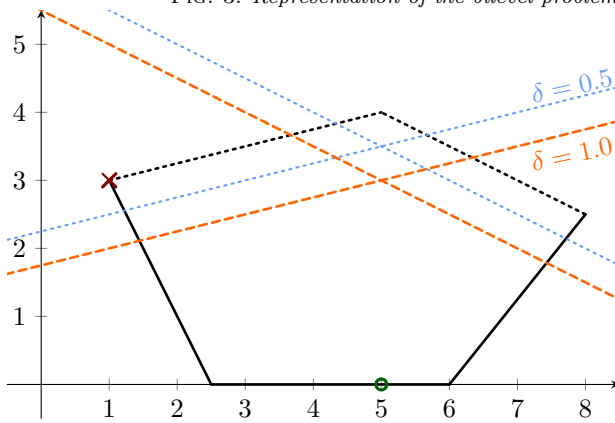


FIG. 4. Near-optimal robustness constraints.

The (α, β) feasible space is defined as:

$$\begin{aligned} -1\alpha_{1,1} - 4\alpha_{1,2} + \beta_1 &\geq 4 \\ -1\alpha_{2,1} - 4\alpha_{2,2} + \beta_2 &\geq 2 \\ \alpha_{ki} &\geq 0, \beta_k \geq 0. \end{aligned}$$

This feasible space can be described as a set of extreme points and rays. It consists in this case of one extreme point $(\alpha_{ki} = 0, \beta_1 = 4, \beta_2 = 2)$ and 4 extreme rays. The (x, v) solution needs to be valid for the corresponding near-optimality conditions:

$$\begin{aligned} \beta_1 (v + \delta) &\leq 11 + x \\ \beta_2 (v + \delta) &\leq 13 - x. \end{aligned}$$

The cuts generated are represented in [Figure 4](#) for $\delta = 0.5$ and $\delta = 1.0$ in dotted blue and dashed orange respectively. The radius of near-optimal feasibility can be

computed using the formulation provided in [Definition 3.3](#), a radius of $\hat{\delta} = 5$ can be computed, for which the feasible domain at the upper-level is reduced to the point $x = 5$, for which $v = 0$, represented as a green circle at $(5, 0)$.

Solution algorithm. We present the solution approach for the linear-linear case from [section 5](#). [Algorithm 5.1](#) uses the extended formulation to solve the linear-linear case of *NORBiP*. One central principle in its design is to prove optimality or infeasibility early in the resolution process, and only then build and solve the extended formulation model.

Algorithm 5.1 Near-Optimal Robust Vertex Enumeration Procedure (NORVEP)

```

1: function near_optimal_bilevel(BiP,  $\delta$ )
2:  $s_0 \leftarrow \text{solve}(\mathcal{P}_0(\text{BiP}))$ 
3: if  $s_0 = \text{Infeasible}$  then
4:   return HighPointInfeasible
5: end if
6:  $s_1 \leftarrow \text{solve}(\mathcal{P}_1(\text{BiP}))$ 
7: if  $s_1 = \text{Infeasible}$  then
8:   return OptimisticInfeasible
9: end if
10: for  $k \in \llbracket m_u \rrbracket$  do
11:    $\text{feas}_k \leftarrow \text{solve}(\text{FEAS}_k((\text{BiP})); \delta)$ 
12:   if  $\text{feas}_k = \text{Infeasible}$  then
13:     return DualAdversarialInfeasible( $k$ )
14:   end if
15: end for
16:  $C_k \leftarrow (\alpha_k^l, \beta_k^l)_{l \in \mathcal{V}_k} \forall k \in \llbracket m_u \rrbracket$ 
17:  $s_{no}(x, v, \alpha, \beta) \leftarrow \text{solve}(\mathcal{P}_{no}(\text{BiP}, (C_k)_{k \in \llbracket m_u \rrbracket}); \delta)$ 
18: return  $s_{no}$ 
19: end function

```

$\mathcal{P}_0(\text{BiP})$, $\mathcal{P}_1(\text{BiP})$, $\text{FEAS}_k((\text{BiP}))$, $\mathcal{P}_{no}(\text{BiP}; \delta)$ are respectively the high-point relaxation, optimistic bilevel problem, dual feasibility and near-optimal robust problem. C_k is the list of extreme vertices of the k -th dual adversarial polyhedron. The function returns an algebraic data type, providing the information on the final status and available data conditioned on this status. The algorithm can be split in a first and second phase, corresponding respectively to a pre-solve (solving relaxations and enumerating vertices) and solve (solve the complete near-optimal robust problem) step.

6. Computational experiments. In this section, we present the application of the solution method to solve instances of the linear-linear near-optimal robust bilevel problem. A total number of 1000 small, 100 medium and 100 large random instances are considered and characterized as follows:

$$\begin{aligned}
(m_u, m_l, n_l, n_u) &= (5, 5, 5, 5) && \text{(small)} \\
(m_u, m_l, n_l, n_u) &= (10, 10, 10, 10) && \text{(medium)} \\
(m_u, m_l, n_l, n_u) &= (20, 10, 20, 20) && \text{(large)}
\end{aligned}$$

All matrices are randomly generated with uniform coefficients in $[0, 1]$ with a sparsity of 60% (all matrix entries have a 0.4 probability of being non-zero). High-point feasibility and the vertex enumeration procedures are run after generating each tuple of random parameters to discard infeasible problems. Collecting 1000 small instances requires generating 10532 trials, the 100 medium-sized instances are obtained with 8830 trials and the 100 large instances are built after 90855 trials.

We first tested the non-extended formulation including bilinear constraints as presented Problem (5.1), which often fails to converge even for small problems. This result was observed with both the SCIP and CPLEX solvers, modelling complementarity constraints using special ordered sets of type 1. The two solvers handle the bilinearities using different reformulations; SCIP uses McCormick envelopes while CPLEX casts the problem as a QCQP.

Algorithm 5.1 is implemented in Julia [5] using the JuMP modelling framework [15]; the MILP used is SCIP 6.0 [17] with SoPlex 4.0, both with default solving parameters. SCIP handles indicator constraints in the form of linear inequality constraints activated only if a binary variable is equal to one. *Polyhedra.jl* [22] is used to model the dual subproblem polyhedra with CDDLib [23] as a solver running the double-description algorithm, producing the list of extreme vertices and rays from the constraint-based representation. The exact rational representation of numbers is used in CDDLib instead of floating-point types to avoid rounding errors. All experiments are performed on a consumer-end laptop with 15.5GB of RAM and an Intel i7 1.9GHz CPU running Ubuntu 16.04LTS.

Table 1 summarizes the results for different values of δ .

δ	0.01	0.1	0.2	1.0	3.0	5.0	7.0	10.0	12.0
Small (/1000)	0	8	22	86	185	236	248	257	260
Medium (/100)	0	0	0	6	16	17	17	17	17

TABLE 1
Number of infeasible problems for various tolerance levels δ

Statistics on the computation times of the two phases of Algorithm 5.1 are provided in Table 2 and Table 3.

Size	mean	10% quant.	50% quant.	90% quant.
Small	0.023	0.014	0.019	0.046
Medium	1.098	0.424	0.956	2.148
Large	44.011	18.740	39.144	76.903

TABLE 2
Runtime statistics for the vertex enumeration (s).

Out of the 100 large instances, one was stopped before completion for time limit purposes and was not included in the runtime statistics of Table 3. Figure 5 shows the distribution of the upper-level objective values across small and medium-sized instances. The number of problems solved to optimality monotonically decreases with δ (Table 1), corresponding to greater δ values making *NORBiP* harder. The optimal values also shift upwards as δ increases, higher δ values imply more conservativeness

Size	# optimized	mean	10% quant.	50% quant.	90% quant.
Small	639	0.042	0.007	0.013	0.076
Medium	63	3.686	0.269	1.600	7.139
Large	78	634.486	0.037	0.128	1789.982

TABLE 3

Runtime statistics for the optimization phase (s).

and potentially a greater distance to the optimal optimistic value. Figure 6 shows the distribution of the lower-level objective for different values of δ .

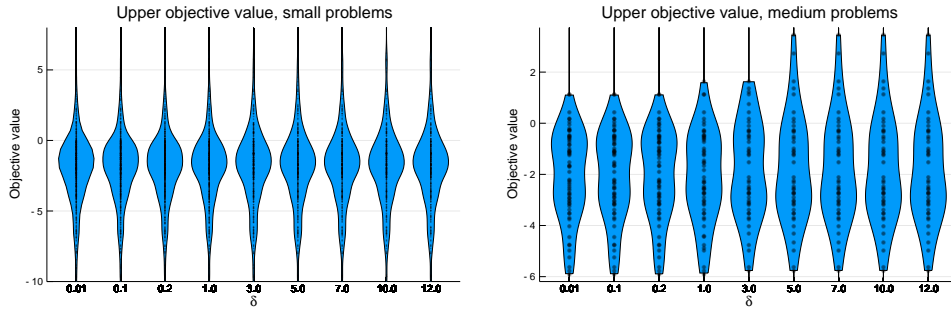


FIG. 5. Upper objective distributions.

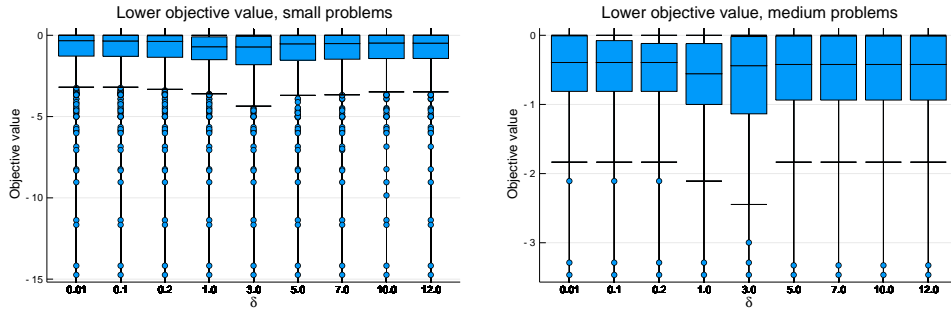


FIG. 6. Lower objective distributions.

Scaling up the dimension of the tackled problems is mostly limited by memory, since the formulation of the problem requires allocating variables and constraints for all vertices of the dual polyhedron of each of the $k \in \llbracket m_u \rrbracket$ subproblems.

7. Conclusion. In this paper, near-optimality robustness was developed as a robust model extending bilevel optimization, protecting the decision-maker from limited deviations of the lower-level. From a robust optimization perspective, the interest of near-optimality is the specific interpretation of the uncertainty set as a near-optimal set of the lower level and of the uncertainty budget as a tolerance δ on the lower-level objective.

Closed-form, single-level expressions of the linear-linear and convex problems were

developed, based on KKT conditions characterizing lower-level optimality and dual certificates of feasibility for the robustness. In the linear case, an extended formulation is applied, leveraging the boundedness of the dual to re-write the problem as a MIP with disjunctive constraints, a structure leveraging the performance of MILP solvers.

Future work will consider heuristics to combine with the extended formulation for memory-limited approaches and efficient solution techniques for the convex case.

REFERENCES

- [1] M. AGHASSI AND D. BERTSIMAS, *Robust game theory*, Mathematical Programming, 107 (2006), pp. 231–273, <https://doi.org/10.1007/s10107-005-0686-0>.
- [2] R. J. AUMANN, *Rationality and bounded rationality*, in Cooperation: Game-Theoretic Approaches, Springer, 1997, pp. 219–231.
- [3] H. BENSON AND R. HORST, *A branch and bound-outer approximation algorithm for concave minimization over a convex set*, Computers & Mathematics with Applications, 21 (1991), pp. 67 – 76, [https://doi.org/10.1016/0898-1221\(91\)90161-V](https://doi.org/10.1016/0898-1221(91)90161-V).
- [4] D. BERTSIMAS, D. B. BROWN, AND C. CARAMANIS, *Theory and Applications of Robust Optimization*, SIAM Review, 53 (2011), pp. 464–501, <https://doi.org/10.1137/080734510>.
- [5] J. BEZANSON, A. EDELMAN, S. KARPINSKI, AND V. SHAH, *Julia: A Fresh Approach to Numerical Computing*, SIAM Review, 59 (2017), pp. 65–98, <https://doi.org/10.1137/141000671>.
- [6] P. BONAMI, A. LODI, A. TRAMONTANI, AND S. WIESE, *On mathematical programming with indicator constraints*, Mathematical Programming, 151 (2015), pp. 191–223, <https://doi.org/10.1007/s10107-015-0891-4>.
- [7] S. BOYD AND L. VANDENBERGHE, *Convex optimization*, Cambridge university press, 2004.
- [8] P. M. CASTRO, *Tightening piecewise McCormick relaxations for bilinear problems*, Computers & Chemical Engineering, 72 (2015), pp. 300–311, <https://doi.org/10.1016/J.COMPCHEMENG.2014.03.025>.
- [9] P. M. CASTRO AND I. E. GROSSMANN, *Optimality-based bound contraction with multiparametric disaggregation for the global optimization of mixed-integer bilinear problems*, Journal of Global Optimization, 59 (2014), pp. 277–306, <https://doi.org/10.1007/s10898-014-0162-6>.
- [10] S. DEMPE, *Foundations of bilevel programming*, Springer Science & Business Media, 2002, <https://doi.org/10.1007/b101970>.
- [11] S. DEMPE, *Bilevel optimization: theory, algorithms and applications*, (2018).
- [12] S. DEMPE, *Computing locally optimal solutions of the bilevel optimization problem using the kkt approach*, in International Conference on Mathematical Optimization Theory and Operations Research, Springer, 2019, pp. 147–157.
- [13] S. DEMPE, F. HARDER, P. MEHLITZ, AND G. WACHSMUTH, *Solving inverse optimal control problems via value functions to global optimality*, Journal of Global Optimization, 74 (2019), pp. 297–325, <https://doi.org/10.1007/s10898-019-00758-1>.
- [14] S. DEMPE, V. KALASHNIKOV, G. A. PREZ-VALDS, AND N. KALASHNYKOVA, *Bilevel Programming Problems: Theory, Algorithms and Applications to Energy Networks*, Springer Publishing Company, Incorporated, 2015.
- [15] I. DUNNING, J. HUCHETTE, AND M. LUBIN, *JuMP: A Modeling Language for Mathematical Optimization*, SIAM Review, 59 (2017), pp. 295–320, <https://doi.org/10.1137/15M1020575>.
- [16] V. GABREL, C. MURAT, AND A. THIELE, *Recent advances in robust optimization: An overview*, European Journal of Operational Research, 235 (2014), pp. 471–483, <https://doi.org/10.1016/J.EJOR.2013.09.036>.
- [17] A. GLEIXNER, M. BASTUBBE, L. EIFLER, T. GALLY, G. GAMRATH, R. L. GOTTWALD, G. HENDEL, C. HOJNY, T. KOCH, M. E. LÜBBECKE, S. J. MAHER, M. MILTENBERGER, B. MÜLLER, M. E. PFETSCH, C. PUCHERT, D. REHFELDT, F. SCHLÖSSER, C. SCHUBERT, F. SERRANO, Y. SHINANO, J. M. VIERNICKEL, M. WALTER, F. WEGSCHEIDER, J. T. WITT, AND J. WITZIG, *The SCIP Optimization Suite 6.0*, technical report, Optimization Online, July 2018, http://www.optimization-online.org/DB_HTML/2018/07/6692.html.
- [18] V. GOEL AND I. E. GROSSMANN, *A class of stochastic programs with decision dependent uncertainty*, Mathematical Programming, 108 (2006), pp. 355–394, <https://doi.org/10.1007/s10107-006-0715-7>.
- [19] M. JAIN, F. ORDÓÑEZ, J. PITA, C. PORTWAY, M. TAMBE, C. WESTERN, P. PARUCHURI, AND

- S. KRAUS, *Robust solutions in stackelberg games: Addressing boundedly rational human preference models*, 2008.
- [20] T. KLEINERT, M. LABBÉ, F. PLEIN, AND M. SCHMIDT, *There's no free lunch: On the hardness of choosing a correct big-m in bilevel optimization*, (2019).
- [21] S. KOŁODZIEJ, P. M. CASTRO, AND I. E. GROSSMANN, *Global optimization of bilinear programs with a multiparametric disaggregation technique*, *Journal of Global Optimization*, 57 (2013), pp. 1039–1063, <https://doi.org/10.1007/s10898-012-0022-1>.
- [22] B. LEGAT, R. DEITS, O. EVANS, G. GORETKIN, T. KOOLEN, J. HUCHETTE, D. OYAMA, M. FORETS, GUBERGER, R. SCHWARZ, E. SABA, AND C. COLEMAN, *JuliaPolyhedra/Polyhedra.jl: v0.5.1*, may 2019, <https://doi.org/10.5281/zenodo.3066026>.
- [23] B. LEGAT, R. DEITS, M. FORETS, D. OYAMA, F. PACAUD, AND E. SABA, *JuliaPolyhedra/CDDLib.jl: v0.5.2*, may 2019, <https://doi.org/10.5281/zenodo.2671760>.
- [24] F. LIERS, L. SCHEWE, AND J. THÜRAUF, *Radius of robust feasibility for mixed-integer problems*, (2019).
- [25] P. LORIDAN AND J. MORGAN, *Weak via strong stackelberg problem: New results*, *Journal of Global Optimization*, 8 (1996), pp. 263–287, <https://doi.org/10.1007/BF00121269>.
- [26] M.A. GOBERNA AND V. JEYAKUMAR AND G. LI AND N. LINH, *Radius of robust feasibility formulas for classes of convex programs with uncertain polynomial constraints*, *Operations Research Letters*, 44 (2016), pp. 67 – 73, <https://doi.org/10.1016/j.orl.2015.11.011>.
- [27] Y. NESTEROV AND A. NEMIROVSKII, *Interior-point polynomial algorithms in convex programming*, vol. 13, Siam, 1994.
- [28] M. G. NICHOLLS, *Optimizing the operations of an aluminium smelter using non-linear bi-level programming*, in *Multilevel Optimization: Algorithms and Applications*, Springer, 1998, pp. 117–148.
- [29] F. ORDÓÑEZ AND N. E. STIER-MOSES, *Robust wardrop equilibrium*, in *Network Control and Optimization*, T. Chahed and B. Tuffin, eds., Berlin, Heidelberg, 2007, Springer Berlin Heidelberg, pp. 247–256, <https://doi.org/10.1007/978-3-540-72709-5>.
- [30] P. PARDALOS AND J. ROSEN, *Methods for global concave minimization: A bibliographic survey*, *SIAM Review*, 28 (1986), pp. 367–379, <https://doi.org/10.1137/1028106>.
- [31] S. PINEDA AND J. M. MORALES, *Solving linear bilevel problems using big-ms: Not all that glitters is gold*, *IEEE Transactions on Power Systems*, 34 (2019), pp. 2469–2471, <https://doi.org/10.1109/TPWRS.2019.2892607>.
- [32] J. PITA, M. JAIN, M. TAMBE, F. ORDÓÑEZ, AND S. KRAUS, *Robust solutions to stackelberg games: Addressing bounded rationality and limited observations in human cognition*, *Artificial Intelligence*, 174 (2010), pp. 1142–1171, <https://doi.org/10.1016/j.artint.2010.07.002>.
- [33] H. SCHEEL AND S. SCHOLTES, *Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity*, *Mathematics of Operations Research*, 25 (2000), pp. 1–22.
- [34] L. SCHEWE AND M. SCHMIDT, *Computing feasible points for binary MINLPs with MPECs*, *Mathematical Programming Computation*, 11 (2019), pp. 95–118, <https://doi.org/10.1007/s12532-018-0141-x>.
- [35] S. SCHOLTES, *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*, *SIAM Journal on Optimization*, 11 (2001), pp. 918–936.
- [36] H. A. SIMON, *Theories of bounded rationality*, *Decision and organization*, 1 (1972), pp. 161–176.
- [37] G. E. STAVROULAKIS AND H. GÜNZEL, *Optimal structural design in nonsmooth mechanics*, in *Multilevel Optimization: Algorithms and Applications*, Springer, 1998, pp. 91–115.
- [38] G. STILL, *Generalized semi-infinite programming: Theory and methods*, *European Journal of Operational Research*, 119 (1999), pp. 301–313.
- [39] W. WIESEMANN, A. TSOUKALAS, P.-M. KLENIATI, AND B. RUSTEM, *Pessimistic Bilevel Optimization*, *SIAM Journal on Optimization*, 23 (2013), pp. 353–380, <https://doi.org/10.1137/120864015>.