There's No Free Lunch: On the Hardness of Choosing a Correct Big-M in Bilevel Optimization

Thomas Kleinert, Martine Labbé, Fränk Plein, Martin Schmidt

There's No Free Lunch:
On the Hardness of Choosing a
Correct Big-M in Bilevel Optimization

Thomas Kleinert[1,2], Martine Labbé[3,4],
Fränk Plein[3,4], and Martin Schmidt[5]

Abstract. One of the most frequently used approaches to solve linear bilevel optimization problems consists in replacing the lower-level problem with its Karush–Kuhn–Tucker (KKT) conditions and by reformulating the KKT complementarity conditions using techniques from mixed-integer linear optimization. The latter step requires to determine some big-M constant in order to bound the lower level's dual feasible set such that no bilevel-optimal solution is cut off. In practice, heuristics are often used to find a big-M although it is known that these approaches may fail. In this paper, we consider the hardness of two proxies for the above mentioned concept of a bilevel-correct big-M. First, we prove that verifying that a given big-M does not cut off any feasible vertex of the lower level's dual polyhedron cannot be done in polynomial time unless $\mathsf{P} = \mathsf{NP}$. Second, we show that verifying that a given big-M does not cut off any optimal point of the lower level's dual problem (for any point in the projection of the high-point relaxation onto the leader's decision space) is as hard as solving the original bilevel problem.

1. Introduction

A bilevel optimization problem consists in a constrained optimization problem in which some constraints specify that a subset of variables constitutes an optimal solution of a second (auxiliary) optimization problem. Since the publication of the first and seminal paper [7], research on the subject has become increasingly important. Indeed, the bilevel structure allows the modeling of a large number of real-life problems involving two types of decision makers, a leader and a follower (or several followers) interacting hierarchically. Such optimization problems appear in many fields of application like energy markets [1, 10, 18–21, 23, 25], critical infrastructure defense [8, 9, 14, 29], or pricing [26, 27, 31].

Due to their ability of modeling hierarchical decision processes, bilevel optimization problems are inherently hard to solve. In [13, 22] it is shown that even the easiest instantiation, i.e., bilevel problems with linear upper and lower level, is strongly NP-hard. Moreover, even checking local optimality for a given point is NP-hard as well [32]. For other hardness results we refer to, e.g., [4]. For general surveys of bilevel optimization see [3, 11–13] and [33] for a survey focusing on linear-linear (LP-LP) bilevel problems.

1

In this paper, we consider the LP-LP bilevel problem

$$\min_{x\in\mathbb{R}^n, y\in\mathbb{R}^m} \quad c^\top x + d^\top y \tag{1a}$$

$$\text{s.t.} \quad Ax + By \geq a, \tag{1b}$$

$$y \in \arg\min_{\bar{y}}\{f^\top \bar{y} : Cx + D\bar{y} \geq b\} \tag{1c}$$

with $c \in \mathbb{R}^n$, $d, f \in \mathbb{R}^m$, $A \in \mathbb{R}^{k\times n}$, $B \in \mathbb{R}^{k\times m}$, $a \in \mathbb{R}^k$, $C \in \mathbb{R}^{\ell\times n}$, $D \in \mathbb{R}^{\ell\times m}$, and $b \in \mathbb{R}^\ell$. Problem (1a)–(1b) is the so-called upper level and (1c) is the lower level. Here, we consider the optimistic version of bilevel optimization [12]. This means that whenever the lower-level problem has multiple solutions y, the leader chooses the most favorable one in terms of the upper level's problem.

LP-LP bilevel problems are often solved by a reformulation to an equivalent single-level problem. Usually, this is done by one of the following two approaches. One can replace the lower level with its primal and dual feasibility conditions as well as the strong-duality equation or one replaces the lower level with its Karush–Kuhn–Tucker (KKT) conditions. Both approaches have their drawbacks: The strong-duality based technique yields nonconvex bilinear terms whereas the KKT approach leads to the following mathematical program with complementarity constraints (MPCC):

$$\min_{x,y,\lambda} \quad c^\top x + d^\top y \tag{2a}$$

$$\text{s.t.} \quad Ax + By \geq a, \ D^\top \lambda = f, \tag{2b}$$

$$0 \leq Cx + Dy - b \perp \lambda \geq 0. \tag{2c}$$

Here, for two non-negative vectors $0 \leq a, b \in \mathbb{R}^n$, $a \perp b$ abbreviates $a_i b_i = 0$ for all $i = 1, \ldots, n$. The single-level reformulation (2) of the bilevel problem (1) explicitly models the optimistic version of bilevel optimization by construction. Consequently, there is no need anymore to distinguish between y and $\bar{y}$. In what follows, we thus call x the upper-level and y the lower-level (primal) variables. The lower-level dual variables are given by $\lambda \in \mathbb{R}^\ell$.

The hardness of Problem (2) stems from the complementarity conditions (2c). Since these constraints can be reformulated in a mixed-integer linear way, the KKT approach is often preferred in practice. Typically, one applies the reformulation introduced in [15], which requires an additional binary variable $z_i \in \{0, 1\}$ for every $i \in \{1, \ldots, \ell\}$ and the additional constraints

$$\lambda_i \leq M_\mathrm{d} z_i, \quad (Cx + Dy - b)_i \leq (1 - z_i)M_\mathrm{p}, \quad i \in \{1, \ldots, \ell\}, \tag{3}$$

where M_d and M_p are sufficiently large constants, called big-Ms. In this note, we focus on the big-M for bounding the lower-level dual variables, i.e., on $M = M_\mathrm{d}$. Since we show that finding M_d is hard, this obviously implies that finding M_d and M_p is hard as well. Applying (3) requires to bound the lower level's dual polyhedron such that no point λ^* that is part of an optimal solution (x^*, y^*, λ^*) of (2) is cut off. Stated differently, one needs to choose an M that preserves all bilevel-optimal points (x^*, y^*). We call an M with this property a *bilevel-correct* big-M.

When the dual of the lower level has a finite optimal value, there exists an optimal solution λ^* that is a vertex of the associated feasible polyhedron. Hence, it is sufficient to obtain bounds on the dual variables that—independently of the upper-level decision—do not cut off (i) any feasible vertex of the lower level's dual polyhedron or (ii) any optimal vertex of the lower level's dual polyhedron. Note that these requirements do not take into account bilevel optimality but still preserve all optimal solutions (x^*, y^*, λ^*) of (2), i.e., (i) and (ii) still yield bilevel-correct big-M's. This means, that (i) and (ii) are sufficient (but not necessary) conditions for a big-M to be bilevel-correct.

The choice of the big-M is often done heuristically, which may result in a severe issue: If the big-M is not chosen large enough, a "solution" of (2) with (2c) replaced by (3) does not need to be a bilevel-optimal point. In fact, this point does not even need to be bilevel-feasible. See, e.g., [30], where a common heuristic for computing a big-M is shown to deliver wrong results.

The contribution of this note is twofold. First, in Section 2, we consider the hardness of verifying that a given big-M does not cut off any feasible vertex of the lower level's dual polyhedron. We show that there is no polynomial-time algorithm for this verification unless $\mathsf{P} = \mathsf{NP}$. Second, in Section 3, we show that validating that a given big-M does not cut off any optimal point of the lower level's dual problem (for any given feasible upper-level variable x) is as hard as solving the original bilevel problem. Both results together imply that there is no hope for an efficient, i.e., polynomial-time, general-purpose method for verifying or computing a correct big-M in bilevel optimization unless $\mathsf{P} = \mathsf{NP}$. Thus, our results strongly indicate that problem-specific bounds on the lower level's dual variables need to be investigated if the given bilevel problem is going to be solved using the KKT approach combined with the classical big-M linearization of KKT complementarity conditions.

2. Hardness of Bounding the Vertices of an Unbounded Polyhedron

Whenever the bilevel problem (1) is feasible, the lower-level primal and dual problem have a finite optimal solution. In particular, there is a vertex of the feasible region of the lower-level dual problem at which the optimal dual solution is attained. Thus, one way of preserving every bilevel-optimal solution in the KKT reformulation (2) is to choose a big-M such that no lower-level dual vertex is cut off. This bounding approach yields a bilevel-correct big-M. In this section, we show—even more generally—that bounding the vertices of an unbounded polyhedron is hard. This result is then applied to the lower level's dual polyhedron. Since the hardness result is mainly based on the unboundedness of this polyhedron, the question arises whether this situation frequently appears in practical LP-LP bilevel problems. It turns out that this is the case for almost all instances of bilevel test sets from the literature—which is also supported by the theoretical results in [34].

To obtain a hardness result in the Turing model of computation, we assume that all problem data are rational and thus are Turing representable. Let $P(A, b) := \{x \in \mathbb{Q}^n : Ax \le b\}$ be an unbounded polyhedron defined by $A \in \mathbb{Q}^{k \times n}$ and $b \in \mathbb{Q}^k$. For $M \in \mathbb{Q}$ and $j \in \{1, \ldots, n\}$, let $Q_j(A, b, M) := \{x \in \mathbb{Q}^n : Ax \le b, \, x_j \le M\}$ be the polyhedron obtained from adding the bound $x_j \le M$ to $P(A, b)$. To validate a given big-M, we need to verify that for every $j \in \{1, \ldots, n\}$ the bound $x_j \le M$ is satisfied by all vertices of $P(A, b)$. This results in the following decision problem.

Component-wise valid bound for the vertices of a polyhedron (CVBVP).
Input: $A \in \mathbb{Q}^{k \times n}$, $b \in \mathbb{Q}^k$, $j \in \{1, \ldots, n\}$, $M \in \mathbb{Q}$.
Question: Does $v \in Q_j(A, b, M)$ hold for every vertex v of $P(A, b)$?

We will see in the following that validating a big-M is related to the problem of finding an optimal vertex v in an unbounded polyhedron with respect to a linear objective function $h^\top v$. If the polyhedron is bounded at least in the direction of optimization, then this problem is equivalent to linear optimization. However, in the general case of polyhedra that are unbounded in the direction of optimization, this is a difficult task. As shown in [16], the decision problem that studies the existence of a vertex of a given polyhedron such that the corresponding objective function value is larger or equal to a certain threshold K is strongly NP-complete. The proof is based on a reduction from the Hamiltonian path problem [17, Problem GT39]

and can easily be extended to the decision problem that decides whether a vertex with an objective function value strictly larger than a certain threshold exists:

Optimal vertex of a polyhedron (OVP).

Input: $A \in \mathbb{Q}^{k \times n}$, $b \in \mathbb{Q}^k$, $h \in \mathbb{Q}^n$, $K \in \mathbb{Q}$.

Question: Is there a vertex v of $P(A, b)$ with $h^\top v > K$?

As pointed out above, w.r.t. the linearization (3), we are interested in the special case $h = e_j$. The related decision problem is the following:

Component-wise optimal vertex of a polyhedron (COVP).

Input: $A \in \mathbb{Q}^{k \times n}$, $b \in \mathbb{Q}^k$, $j \in \{1, \ldots, n\}$, $K \in \mathbb{Q}$.

Question: Is there a vertex v of $P(A, b)$ with $v_j > K$?

We now show that even for this subclass of instances, the decision problem is strongly NP-complete. In what follows, $\mathrm{vert}(P(A, b))$ denotes the set of vertices of the polyhedron $P(A, b)$.

Theorem 1. *COVP is strongly NP-complete.*

Proof. We prove the result for $j = 1$. For any other $j' \in \{2, \ldots, n\}$, e_1 can be replaced with $e_{j'}$ in the proof.

It is clear that the problem is in NP. We prove its hardness by reduction from OVP. Let $A \in \mathbb{Q}^{k \times n}$, $b \in \mathbb{Q}^k$, $h \in \mathbb{Q}^n$, $K \in \mathbb{Q}$ be a given OVP instance and assume $h \neq 0$. Otherwise, the corresponding instance is trivial. Now take $j \in \{1, \ldots, n\}$ with $h_j \neq 0$. We construct a basis of $\mathbb{Q}^n$ by replacing e_j with h. If we put h as the first basis vector, the corresponding linear transformation is given by the inverse of matrix $B = [h, e_1, \ldots, e_{j-1}, e_{j+1}, \ldots, e_n]$ and can be computed in polynomial time. Using this basis change, we now linearly transform the hyperplanes defining the polyhedron $P(A, b)$ and the objective function vector h of the given OVP instance.

We construct an instance of COVP by $\tilde{A} = AB^{-\top}$, $\tilde{b} = b$, $\tilde{K} = K$. Note that $B^{-1}h = e_1$ holds. It remains to show that there exists a vertex v of $P(A, b)$ with $h^\top v > K$ if and only if there exists a vertex $\tilde{v}$ of $P(\tilde{A}, \tilde{b})$ with $\tilde{v}_1 > \tilde{K}$. Let $v \in \mathrm{vert}(P(A, b))$ such that $h^\top v > K$ and define $\tilde{v} := B^\top v$. Then,

$$\tilde{A}\tilde{v} = AB^{-\top}B^\top v = Av \leq b = \tilde{b}, \quad \tilde{v}_1 = h^\top B^{-\top}B^\top v = h^\top v > K = \tilde{K}.$$

Thus, $\tilde{v} \in P(\tilde{A}, \tilde{b})$ and it is clear that $\tilde{v}$ is also a vertex of $P(A, b)$.

Conversely, let $\tilde{v} \in \mathrm{vert}(P(\tilde{A}, \tilde{b}))$ with $\tilde{v}_1 > \tilde{K}$ and define $v := B^{-\top}\tilde{v}$. Then,

$$Av = AB^{-\top}\tilde{v} = \tilde{A}\tilde{v} \leq \tilde{b} = b, \quad h^\top v = h^\top B^{-\top}\tilde{v} = \tilde{v}_1 > \tilde{K} = K.$$

Thus, v is a vertex of $P(A, b)$. $\qquad\qquad\square$

Using problem COVP, we can deduce the complexity of CVBVP.

Theorem 2. *CVBVP is strongly coNP-complete.*

Proof. The decision problem $\overline{\mathrm{CVBVP}}$, i.e., the complement of CVBVP, is to find a vertex $v \in \mathrm{vert}(P(A, b))$ such that $v_j > M$ holds. This is equivalent to COVP with $K = M$. $\qquad\qquad\square$

Finally, we can state the main result of this section.

Corollary 1. *Let $A \in \mathbb{Q}^{k \times n}$, $b \in \mathbb{Q}^k$, and $M \in \mathbb{Q}$. Then, there exists no polynomial-time algorithm for checking whether*

$$\mathrm{vert}(P(A, b)) \subseteq \bigcap_{j=1}^n Q_j(A, b, M),$$

unless P = NP.

Proof. Assume a polynomial-time algorithm exists. Then, for every $j \in \{1, \ldots, n\}$ we can efficiently decide whether $\mathrm{vert}(P(A, b)) \subseteq Q_j(A, b, M)$ holds. This implies that we can decide CVBVP in polynomial time, and thus $\mathsf{P} = \mathsf{coNP}$ must also hold. Since P is closed under taking the complement, it follows that $\mathsf{P} = \mathsf{NP}$. $\qquad\square$

As a final remark, note that to compute the tightest possible big-M such that no vertex of $P(A, b)$ is cut off, we can set

$$M := \max_{j \in \{1, \ldots, n\}} \left\{ \max_{x \in \mathrm{vert}(P(A,b))} x_j \right\}. \tag{4}$$

It is equivalent to solving COVP for $j \in \{1, \ldots, n\}$ and taking the maximum value. Thus, (4) cannot be computed in polynomial time, unless $\mathsf{P} = \mathsf{NP}$.

3. Valid Bounds for Bilevel-Feasible Solutions

Recall that a big-M is bilevel-correct, if it does not cut off any bilevel-optimal solution. For this, it is sufficient to find a big-M that maintains at least one optimal lower-level dual vertex for every feasible upper-level decision. This means that, in contrast to the big-M of Section 2, we now allow to cut off lower-level dual vertices that do not correspond to an optimal solution.

Here and in what follows we denote the high-point relaxation of the bilevel problem (1) as

$$H := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \colon Ax + By \geq a, \, Cx + Dy \geq b\}$$

and the corresponding projection onto the space of x-variables is defined as

$$H_x := \{x \in \mathbb{R}^n \colon \exists y \text{ with } (x, y) \in H\}.$$

For the sake of simplicity, we make the following assumption.

Assumption 1. *For every upper-level decision $x \in H_x$, the lower-level problem (1c) admits a unique solution y and satisfies the linear independence constraint qualifications (LICQ) at y.*

The assumption of a unique lower-level solution is justified in this scope given that the bilevel problem becomes even more difficult to analyze otherwise; see, e.g., Chapter 7 in [12]. Moreover, the LICQ guarantees the uniqueness of the lower-level dual optimal solution for every upper-level decision $x \in H_x$; see, e.g., Chapter 12 of [28]. Let us note that the assumption that LICQ holds at the unique primal solution of the lower level can be very strong. However, we will show the hardness of choosing a big-M that does not cut off any optimal vertex of the lower level's dual polyhedron under the simplifying Assumption 1. Thus, the problem of choosing such a big-M is hard also for the situation in which Assumption 1 is dropped.

We start by introducing a validity criterion for the big-M proxy discussed in this section. To this end, define the lower-level optimal value function $\varphi(x)$ for any upper-level decision $x \in H_x$ by means of its dual as

$$\varphi(x) := \max_\lambda \, \{(b - Cx)^\top \lambda \colon D^\top \lambda = f, \, \lambda \geq 0\}. \tag{5}$$

Further, for any upper-level decision $x \in H_x$, $M \in \mathbb{R}$, and $i \in \{1, \ldots, \ell\}$, let $\varphi_i(x, M)$ be the optimal value function of the lower level's dual problem with the additional bound $\lambda_i \leq M$, i.e.,

$$\varphi_i(x, M) := \max_\lambda \, \{(b - Cx)^\top \lambda \colon D^\top \lambda = f, \, \lambda \geq 0, \, \lambda_i \leq M\}, \tag{6}$$

where we formally set $\varphi_i(x, M) = -\infty$ if Problem (6) is infeasible. Under Assumption 1, all bilevel-feasible solutions remain the same after adding the big-M bounds to the lower level's dual problem if and only if for every upper-level decision $x \in H_x$

and for every $i \in \{1, \ldots, \ell\}$, the lower-level optimal value stays unchanged, i.e., if $\varphi(x) = \varphi_i(x, M)$ holds.

We now collect some simple observations on these two optimal value functions that are used afterward.

Observation 1. *Given an upper-level decision $x \in H_x$ and $i \in \{1, \ldots, \ell\}$, the following properties hold:*

(a) *$\varphi(x) \geq \varphi_i(x, M)$ for every $M \in \mathbb{R}$.*
(b) *$\varphi_i(x, \cdot)$ is monotonically increasing.*
(c) *Suppose there exists an $M \in \mathbb{R}$ with $\varphi(x) = \varphi_i(x, M)$. Then, $\varphi(x) = \varphi_i(x, \tilde{M})$ holds for every $\tilde{M} \geq M$.*

Lemma 1. *Suppose that Assumption 1 holds and let an upper-level decision $x \in H_x$ and $M = M(x) \in \mathbb{R}$ be given. Then, for every $i \in \{1, \ldots, \ell\}$, $\varphi(x) = \varphi_i(x, M(x))$ holds if and only if $M(x) \geq \max\{\lambda_i^*(x) \colon i \in \{1, \ldots, \ell\}\}$, where $\lambda^*(x)$ is the unique optimal solution of the lower level's dual problem (5) corresponding to x.*

Proof. If $M(x) < \max\{\lambda_i^*(x) \colon i \in \{1, \ldots, \ell\}\}$, then there is an $i \in \{1, \ldots, \ell\}$ such that the optimal solution of the lower level's dual problem is cut off by the bound $\lambda_i \leq M(x)$, which again is equivalent to $\varphi(x) > \varphi_i(x, M(x))$. $\qquad\square$

In particular, this implies that for every fixed upper-level decision, we can validate a given big-M by computing the corresponding unique optimal solution of the lower level's dual problem and by verifying that it satisfies the bounds $\lambda_i \leq M$ for all $i \in \{1, \ldots, \ell\}$.

For the case that all upper-level decisions are taken into account, the next result gives a necessary and sufficient condition for the property that a big-M does not cut off any bilevel-feasible point.

Theorem 3. *Let $M \in \mathbb{R}$ be given and suppose that Assumption 1 holds. Then, for every upper-level decision $x \in H_x$ and for every $i \in \{1, \ldots, \ell\}$, $\varphi(x) = \varphi_i(x, M)$ holds if and only if*

$$M \geq \max_{i \in \{1, \ldots, \ell\}} \left\{ \max_{x, y, \lambda} \ \{\lambda_i \colon (2b), (2c)\} \right\}. \tag{7}$$

Proof. Observe that the first constraint of (2b) defines the domain of the upper-level decisions x, whereas the second constraint together with (2c) determine the lower-level primal-dual optimal solution (y, λ) corresponding to x. The final result then follows by Lemma 1 and Property (c) in Observation 1. $\qquad\square$

Theorem 3 implies that validating a big-M requires optimizing different objective functions over a set of constraints that are equivalent to feasibility of the original bilevel problem. In [24], linear 0-1-feasibility has been shown to be NP-complete. It is thus possible to adapt the techniques from [2] to show the NP-completeness of LP-LP bilevel feasibility by reduction from linear 0-1-feasibility. Similarly to Corollary 1, we can thus state that there is no polynomial-time validation of a given big-M w.r.t. (7) unless P = NP. On the other hand, computing the tightest big-M w.r.t. the proxy considered in this section requires solving a maximization problem over all bilevel-feasible solutions for every $i \in \{1, \ldots, \ell\}$ and taking the maximum objective value. Computing this big-M is therefore as hard as solving the initial problem and there is little hope of doing it efficiently, unless the original bilevel problem (1) can be solved in polynomial time.

4. CONCLUSION

Many applications of LP-LP bilevel optimization make use of the KKT reformulation of the lower-level problem together with a big-M linearization of the KKT

complementarity constraints. This results in a single-level mixed-integer linear problem that can, in principle, be solved with state-of-the-art solvers. However, to guarantee bilevel feasibility of a solution obtained by this approach, one needs to validate the bilevel-correctness of the big-M that is used to bound the lower level's dual variables—a necessary task that is not always carried out in practice. In general, such a big-M is bilevel-correct if it does not cut off any bilevel-optimal point. In this note we considered two proxies for this type of correctness and proved that even validating that a given big-M does not cut off any feasible or optimal vertex of the lower level's dual polyhedron cannot be done in polynomial time unless $\mathsf{P} = \mathsf{NP}$. Both proxies abstract from upper-level optimality and, thus, are only sufficient but not necessary conditions for a big-M to be bilevel-correct. Hence, validating that a given big-M preserves all bilevel-optimal points can be expected to be at least as hard since one needs to take into account another—i.e., the upper level's—optimization problem on top of what needs to be considered for the two proxies.

Our results strongly suggest that the popular big-M approach needs to be applied very carefully. If the bilevel-correctness of the chosen big-M is not guaranteed by problem-specific insights, it cannot be formally guaranteed that the obtained "solutions" are indeed bilevel-optimal. In such cases, we suggest to better resort to exact approaches that do not rely on big-M's like, e.g., the kth best algorithm ([6] or [3, Chapter 5.3.1]) or branch-and-bound methods ([5, 22], or [3, Chapter 5.3.2]). Moreover, identifying reasonably generic sub-classes of bilevel optimization problems for which it is easy to determine a bilevel-correct big-M is subject to future research.

Acknowledgments

References

[1] M. Ambrosius, V. Grimm, T. Kleinert, F. Liers, M. Schmidt, and G. Zöttl. *Endogenous Price Zones and Investment Incentives in Electricity Markets: An Application of Multilevel Optimization with Graph Partitioning.* Tech. rep. Oct. 2018. URL: http://www.optimization-online.org/DB_HTML/2018/10/6868.html.

[2] C. Audet, P. Hansen, B. Jaumard, and G. Savard. "Links Between Linear Bilevel and Mixed 0–1 Programming Problems." In: *Journal of Optimization Theory and Applications* 93.2 (1997), pp. 273–300. DOI: 10.1023/A:1022645805569.

[3] J. F. Bard. *Practical bilevel optimization: algorithms and applications.* Vol. 30. Springer Science & Business Media, 2013.

[4] J. F. Bard. "Some properties of the bilevel programming problem." In: *Journal of Optimization Theory and Applications* 68.2 (1991), pp. 371–378. DOI: 10.1007/BF00941574.

[5] J. F. Bard and J. T. Moore. "A branch and bound algorithm for the bilevel programming problem." In: *SIAM Journal on Scientific and Statistical Computing* 11.2 (1990), pp. 281–292. DOI: 10.1137/0911017.

[6]   W. F. Bialas and M. H. Karwan. "Two-level linear programming." In: *Management Science* 30.8 (1984), pp. 1004–1020. DOI: `10.1287/mnsc.30.8.1004`.

[7]   J. Bracken and J. T. McGill. "Mathematical programs with optimization problems in the constraints." In: *Operations Research* 21.1 (1973), pp. 37–44. DOI: `10.1287/opre.21.1.37`.

[8]   A. Caprara, M. Carvalho, A. Lodi, and G. J. Woeginger. "Bilevel Knapsack with Interdiction Constraints." In: *INFORMS Journal on Computing* 28.2 (2016), pp. 319–333. DOI: `10.1287/ijoc.2015.0676`.

[9]   C. Casorrán, B. Fortz, M. Labbé, and F. Ordóñez. *A study of general and security Stackelberg game formulations.* Tech. rep. Nov. 2018. URL: `https://hal.inria.fr/hal-01917798`.

[10]   O. Daxhelet and Y. Smeers. "The EU regulation on cross-border trade of electricity: A two-stage equilibrium model." In: *European Journal of Operational Research* 181.3 (2007), pp. 1396–1412. DOI: `10.1016/j.ejor.2005.12.040`.

[11]   S. Dempe. *Bilevel Optimization: Theory, Algorithms, and Applications.* Tech. rep. 2018. URL: `http://www.optimization-online.org/DB_HTML/2018/08/6773.html`.

[12]   S. Dempe. *Foundations of bilevel programming.* Springer, 2002. DOI: `10.1007/b101970`.

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

[14]   S. DeNegre. "Interdiction and discrete bilevel linear programming." PhD thesis. Lehigh University, 2011.

[15]   J. Fortuny-Amat and B. McCarl. "A Representation and Economic Interpretation of a Two-Level Programming Problem." In: *The Journal of the Operational Research Society* 32.9 (1981), pp. 783–792. JSTOR: `2581394`.

[16]   K. Fukuda, T. M. Liebling, and F. Margot. "Analysis of Backtrack Algorithms for Listing All Vertices and All Faces of a Convex Polyhedron." In: *Computational Geometry* 8.1 (1997), pp. 1–12. DOI: `10.1016/0925-7721(95)00049-6`.

[17]   M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* New York, NY, USA: W. H. Freeman & Co., 1979.

[18]   V. Grimm, J. Grübel, L. Schewe, M. Schmidt, and G. Zöttl. "Nonconvex equilibrium models for gas market analysis: Failure of standard techniques and alternative modeling approaches." In: *European Journal of Operational Research* 273.3 (2019), pp. 1097–1108. DOI: `10.1016/j.ejor.2018.09.016`.

[19]   V. Grimm, T. Kleinert, F. Liers, M. Schmidt, and G. Zöttl. "Optimal price zones of electricity markets: a mixed-integer multilevel model and global solution approaches." In: *Optimization Methods and Software* 34.2 (2019), pp. 406–436. DOI: `10.1080/10556788.2017.1401069`.

[20]   V. Grimm, A. Martin, M. Schmidt, M. Weibelzahl, and G. Zöttl. "Transmission and Generation Investment in Electricity Markets: The Effects of Market Splitting and Network Fee Regimes." In: *European Journal of Operational Research* 254.2 (2016), pp. 493–509. DOI: `10.1016/j.ejor.2016.03.044`.

[21]   V. Grimm, L. Schewe, M. Schmidt, and G. Zöttl. "A Multilevel Model of the European Entry-Exit Gas Market." In: *Mathematical Methods of Operations Research* (2018). DOI: `10.1007/s00186-018-0647-z`.

[22]   P. Hansen, B. Jaumard, and G. Savard. "New branch-and-bound rules for linear bilevel programming." In: *SIAM Journal on Scientific and Statistical Computing* 13.5 (1992), pp. 1194–1217. DOI: `10.1137/0913069`.

[23] X. Hu and D. Ralph. "Using EPECs to Model Bilevel Games in Restructured Electricity Markets with Locational Prices." In: *Operations Research* 55.5 (2007), pp. 809–827. DOI: 10.1287/opre.1070.0431.

[24] R. M. Karp. "On the Computational Complexity of Combinatorial Problems." In: *Networks* 5.1 (1975), pp. 45–68. DOI: 10.1002/net.1975.5.1.45.

[25] T. Kleinert and M. Schmidt. "Global Optimization of Multilevel Electricity Market Models Including Network Design and Graph Partitioning." In: *Discrete Optimization* 33 (2019), pp. 43–69. DOI: 10.1016/j.disopt.2019.02.002.

[26] M. Labbé, P. Marcotte, and G. Savard. "A bilevel model of taxation and its application to optimal highway pricing." In: *Management Science* 44.12-part-1 (1998), pp. 1608–1622. DOI: 10.1287/mnsc.44.12.1608.

[27] M. Labbé and A. Violin. "Bilevel programming and price setting problems." In: *4OR* 11.1 (2013), pp. 1–30. DOI: 10.1007/s10288-012-0213-0.

[28] J. Nocedal and S. J. Wright. *Numerical Optimization*. 2nd ed. Springer Series in Operations Research and Financial Engineering. New York: Springer-Verlag, 2006. DOI: 10.1007/978-0-387-40065-5.

[29] P. Paruchuri, J. P. Pearce, J. Marecki, M. Tambe, F. Ordonez, and S. Kraus. "Playing games for security: An efficient exact algorithm for solving Bayesian Stackelberg games." In: *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Volume 2*. International Foundation for Autonomous Agents and Multiagent Systems. 2008, pp. 895–902.

[30] S. Pineda and J. M. Morales. "Solving Linear Bilevel Problems Using Big-Ms: Not All That Glitters Is Gold." In: *IEEE Transactions on Power Systems* (2019). DOI: 10.1109/TPWRS.2019.2892607.

[31] P. Rusmevichientong, B. Van Roy, and P. W. Glynn. "A nonparametric approach to multiproduct pricing." In: *Operations Research* 54.1 (2006), pp. 82–98. DOI: 10.1287/opre.1050.0252.

[32] L. Vicente, G. Savard, and J. Júdice. "Descent approaches for quadratic bilevel programming." In: *Journal of Optimization Theory and Applications* 81.2 (1994), pp. 379–399. DOI: 10.1007/BF02191670.

[33] U.-P. Wen and S.-T. Hsu. "Linear Bi-Level Programming Problems – A Review." In: *The Journal of the Operational Research Society* 42.2 (1991), pp. 125–133. URL: http://www.jstor.org/stable/2583177.

[34] A. Williams. "Boundedness relations for linear constraint sets." In: *Linear Algebra and its Applications* 3.2 (1970), pp. 129–141. DOI: 10.1016/0024-3795(70)90009-1.

(T. Kleinert) [1] Friedrich-Alexander-Universität Erlangen-Nürnberg, Discrete Optimization, Cauerstr. 11, 91058 Erlangen, Germany; [2] Energie Campus Nürnberg, Fürther Str. 250, 90429 Nürnberg, Germany

E-mail address: thomas.kleinert@fau.de

(M. Labbé, F. Plein) [3] Université Libre de Bruxelles, Department of Computer Science, Boulevard du Triomphe, CP212, 1050 Brussels, Belgium; [4] Inria Lille - Nord Europe, Parc scientifique de la Haute Borne, 40, av. Halley - Bât A - Park Plaza, 59650 Villeneuve d'Ascq, France

E-mail address: {martine.labbe,frank.plein}@ulb.ac.be

(M. Schmidt) [5] Trier University, Department of Mathematics, Universitätsring 15, 54296 Trier, Germany

E-mail address: martin.schmidt@uni-trier.de