J Glob Optim (2010) 47:557–582
DOI 10.1007/s10898-009-9479-y

Global solution of nonlinear mixed-integer bilevel programs

Alexander Mitsos

Received: 3 April 2008 / Accepted: 13 October 2009 / Published online: 30 October 2009

Abstract An algorithm for the global optimization of nonlinear bilevel mixed-integer programs is presented, based on a recent proposal for continuous bilevel programs by Mitsos et al. (J Glob Optim 42(4):475–513, 2008). The algorithm relies on a convergent lower bound and an optional upper bound. No branching is required or performed. The lower bound is obtained by solving a mixed-integer nonlinear program, containing the constraints of the lower-level and upper-level programs; its convergence is achieved by also including a parametric upper bound to the optimal solution function of the lower-level program. This lower-level parametric upper bound is based on Slater-points of the lower-level program and subsets of the upper-level host sets for which this point remains lower-level feasible. Under suitable assumptions the KKT necessary conditions of the lower-level program can be used to tighten the lower bounding problem. The optional upper bound to the optimal solution of the bilevel program is obtained by solving an augmented upper-level problem for fixed upper-level variables. A convergence proof is given along with illustrative examples. An implementation is described and applied to a test set comprising original and literature problems. The main complication relative to the continuous case is the construction of the parametric upper bound to the lower-level optimal objective value, in particular due to the presence of upper-level integer variables. This challenge is resolved by performing interval analysis over the convex hull of the upper-level integer variables.

Keywords Bilevel program · Nonconvex · Global optimization · Mixed-integer · MINLP · Parametric optimization

1 Introduction

Bilevel programs are programs where an upper-level program (or outer program, superscript u) is constrained by an embedded lower-level program (or inner program, superscript l

A. Mitsos (✉)
Aachen Institute for Advanced Study in Computational Engineering Science, RWTH Aachen, Pauwelsstr. 12, 52074 Aachen, Germany
e-mail: amitsos@alum.mit.edu

or m). Here, mixed-integer nonlinear bilevel programs are considered without any convexity assumptions:

$$\begin{aligned} f^{u,*} = \min_{\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l} & f^u\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \\ \text{s.t. } & \mathbf{g}^u\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \leq \mathbf{0} \\ & \left(\mathbf{x}^l,\mathbf{y}^l\right) \in \arg\min_{\mathbf{x}^m,\mathbf{y}^m} f^l\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^m,\mathbf{y}^m\right) \\ & \text{s.t. } \mathbf{g}^{l,1}\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^m,\mathbf{y}^m\right) \leq \mathbf{0} \\ & \mathbf{g}^{l,2}\left(\mathbf{x}^m,\mathbf{y}^m\right) \leq \mathbf{0} \\ & \mathbf{x}^u \in X^u \subset \mathbb{R}^{n_x^u}, \quad \mathbf{y}^u \in Y^u \subset \mathbb{Z}^{n_y^u}, \\ & \mathbf{x}^l,\mathbf{x}^m \in X^l \subset \mathbb{R}^{n_x^l}, \quad \mathbf{y}^l,\mathbf{y}^m \in Y^l \subset \mathbb{Z}^{n_y^l}. \end{aligned} \tag{1}$$

The co-operative (or optimistic, weak) formulation [9] is assumed where if for a given choice of upper-level variables $(\mathbf{x}^u, \mathbf{y}^u)$ the lower-level program has multiple optimal solutions $\left(\mathbf{x}^l, \mathbf{y}^l\right)$, the upper-level optimizer can choose among them. Dummy variables (superscript m instead of l) are used in the lower-level program for clarity.

There are many algorithms for bilevel programs, see, e.g., [4,7,9,10,25,29,34] for reviews. The majority of publications assume continuous variables and linear or convex nonlinear objective functions and constraints. There are several contributions to the mixed-integer case, however, most of these are restricted to mixed-integer linear bilevel programs. Only the most relevant contributions are discussed here. Moore and Bard [4,25] discuss difficulties of the mixed-integer case, in particular that removing the integrality constraints does not necessarily give a relaxation of the bilevel program. Jan and Chern [19] consider bilevel programs without upper-level constraints and with functions that are sums of univariate nondecreasing terms ("separable monotone bilevel programs"). Thirwani and Arora [33] consider bilevel programs with integer variables, linear constraints and fractional linear objective functions. Sahin and Ciric [27] propose an algorithm based on simulated annealing and test its performance on examples including mixed-integer nonlinear bilevel problems. Gümus and Floudas [18] propose deterministic algorithms for the special case that the functions in the lower-level program are convex with respect to the continuous lower-level variables and multilinear with respect to the integer lower-level variables.

Nonconvexity in the lower-level program is a major complication in all embedded optimization problems. Deterministic algorithms that allow nonconvexity exist for some embedded problems, such as min–max programs [12,36], semi-infinite programs (SIP) [5,6,13,23], and generalized semi-infinite programs [21]. Recently Mitsos et al. [22] proposed the first deterministic algorithm for the global solution of continuous bilevel programs with a nonconvex lower-level program. This algorithm proceeds similar to the algorithm by Blankenship and Falk [6] for SIPs, by adding cuts to the lower bounding problem until the upper bound is guaranteed to generate an $\varepsilon-$optimal point. The cuts used are generated by parametric upper bounds to the lower-level problem. Bilevel programs with 0–1 variables in the upper-level program could in principle be solved by the aforementioned algorithm by replacing the integrality constraint with an MPEC constraint, e.g., $y^u \in \{0, 1\} \Leftrightarrow y^u\ (y^u - 1) \leq 0,\ y^u \in [0, 1]$. This however would have two drawbacks: (i) since these constraints violate the constraint qualification, this reformulation would limit the applicability of solvers for the subproblems; (ii) the generation of parametric upper bounds to the lower-level objective function would not be as efficient as in the proposal herein. Integer variables in the lower-level program do not cause a significant burden to the basic variant of the algorithm by Mitsos et al. [22], but could not be handled by the KKT-based lower bounding problem proposed therein.

In this article, the aforementioned algorithm is extended to the mixed-integer case. The same principle is used, namely a successively tighter lower bounding procedure. The simpler variant without branching is used, mainly for notational simplicity. All subtasks of the algorithm are carefully examined and generalized to allow for integer variables. The main challenge is the influence of upper-level variables on the generation of parametric upper bounds to the lower-level program. Similar to the continuous case these bounds are based on pairs comprising a subset of the upper-level host set $X^{u,k} \times Y^{u,k}$ and a lower-level point $\left(\mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right)$, that remains feasible in the lower-level program for all $(\mathbf{x}^u, \mathbf{y}^u) \in X^{u,k} \times Y^{u,k}$. To avoid enumeration of the upper-level integer feasible solutions, it is required that $\left(\mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right)$ remains feasible in the lower-level program for any $(\mathbf{x}^u, \mathbf{y}^u) \in X^{u,k} \times \operatorname{conv}\left(Y^{u,k}\right)$.

In the remainder of this article first the definitions and assumptions made are given in Sect. 2. Then, in Sect. 3 a convergent lower bounding procedure is described, followed by a convergent upper bounding procedure in Sect. 4. In Sect. 5 an algorithm combining these procedures is proposed, followed by a proof of finite termination and illustrative examples. In Sect. 6 an implementation is described and numerical results from a test set comprising original and literature problems are given. Section 7 gives conclusions and potential for future work.

2 Definitions and assumptions

In many places of this article the continuous and integer variables are considered jointly to simplify notation:

Definition 1 (*Joint variables*) The upper-level variables are defined as $\mathbf{z}^u \equiv (\mathbf{x}^u, \mathbf{y}^u)$ and their host set is $Z^u \equiv X^u \times Y^u$. Similarly in the lower-level program $\mathbf{z}^l \equiv \left(\mathbf{x}^l, \mathbf{y}^l\right)$ and $Z^l \equiv X^l \times Y^l$.

Throughout the article the optimal solution value to the lower-level problem (or value function of the lower-level problem [35]) is used:

Definition 2 (*Parametric optimal solution function*) The parametric optimal solution value of the lower-level program as a function of the upper-level variables is denoted $f^{l,*}(\mathbf{x}^u, \mathbf{y}^u)$. For points $(\mathbf{x}^u, \mathbf{y}^u)$ for which the lower-level program is infeasible, the convention $f^{l,*}(\mathbf{x}^u, \mathbf{y}^u) = +\infty$ is used.

Note that the set of optimal solutions to the lower-level program can be empty, a singleton, a finite set or an infinite set. Moreover, the cardinality of this set can depend on the upper-level variables.

It is convenient to define the set of candidate upper level points:

Definition 3 (*Candidate upper-level points*) The subset of Z^u which is admissible in the upper-level and lower-level program is denoted:

$$Z^{u,\infty} = \Big\{\mathbf{z}^u \in Z^u : \exists\, \mathbf{z}^l \in Z^l : \mathbf{g}^u\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) \le \mathbf{0},$$
$$\mathbf{g}^{l,1}\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) \le \mathbf{0},\ \mathbf{g}^{l,2}\left(\mathbf{x}^l, \mathbf{y}^l\right) \le \mathbf{0}\Big\}.$$

It is convenient to also define the following (potentially nonconvex) level set:

Definition 4 (*Level set*) For a given $\bar{f}^u \in \mathbb{R}$ define the level set

$$Z^{u,s}\left(\bar{f}^u\right) = \left\{\mathbf{z}^u \in Z^u : \exists\, \mathbf{z}^l \in Z^l :\ \mathbf{g}^u\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) \leq \mathbf{0},\ \mathbf{g}^{l,1}\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) \leq \mathbf{0}, \right.$$
$$\left. \mathbf{g}^{l,2}\left(\mathbf{x}^l, \mathbf{y}^l\right) \leq \mathbf{0},\ f^u\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) \leq \bar{f}^u\right\}.$$

By definition $Z^{u,\infty} = Z^{u,s}(\infty)$. As is explained in the sequel, the algorithm by construction visits only points in $Z^{u,s}(f^{u,*})$.

2.1 Assumptions

In the following the assumptions required for convergence of the algorithm proposed are discussed. As is typical in global optimization box-constrained host sets are considered:

Assumption 1 (*Host sets*) Explicit bounds are known for all variables i.e., we have $X^u = \left[\mathbf{x}^{u,\mathrm{LO}}, \mathbf{x}^{u,\mathrm{UP}}\right]$, $Y^u = \mathbb{Z}^{n_y^u} \cap \left[\mathbf{y}^{u,\mathrm{LO}}, \mathbf{y}^{u,\mathrm{UP}}\right]$, $X^l = \left[\mathbf{x}^{l,\mathrm{LO}}, \mathbf{x}^{l,\mathrm{UP}}\right]$ and $Y^l = \mathbb{Z}^{n_y^l} \cap \left[\mathbf{y}^{l,\mathrm{LO}}, \mathbf{y}^{l,\mathrm{UP}}\right]$ with $\mathbf{x}^{u,\mathrm{LO}}, \mathbf{x}^{u,\mathrm{UP}} \in \mathbb{R}^{n_x^u}$, $\mathbf{x}^{l,\mathrm{LO}}, \mathbf{x}^{l,\mathrm{UP}} \in \mathbb{R}^{n_x^l}$, $\mathbf{y}^{u,\mathrm{LO}}, \mathbf{y}^{u,\mathrm{UP}} \in \mathbb{Z}^{n_y^u}$ and $\mathbf{y}^{l,\mathrm{LO}}, \mathbf{y}^{l,\mathrm{UP}} \in \mathbb{Z}^{n_y^l}$.

It would also be possible to consider more general compact host sets. Another typical assumption used is continuity:

Assumption 2 (*Basic properties of functions*) All functions in (1) are assumed to be continuous on $X^u \times X^l$ for all pairs $\left(\mathbf{y}^u, \mathbf{y}^l\right) \in Y^u \times Y^l$.

If some upper-level variables are continuous the following additional assumption is required for convergence:

Assumption 3 (*Lower-level problem*) If $n_x^u > 0$, there exists some $\bar{\varepsilon}_f^u > 0$ such that for each point $\bar{\mathbf{z}}^u \in Z^{u,\infty}$ at least one of the following two conditions holds:

1. For any $\varepsilon_{f1}^l > 0$ there exists a lower-level vector $\tilde{\mathbf{z}}^l \in Z^l$ such that

$$\mathbf{g}^{l,1}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right) < \mathbf{0}, \quad \mathbf{g}^{l,2}\left(\tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right) \leq \mathbf{0},$$
$$f^l\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right) \leq f^{l,*}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u\right) + \varepsilon_{f1}^l. \tag{2}$$

2. The upper-level objective value is $\bar{\varepsilon}_f^u$ worse than the optimal objective value f^*, i.e., $\bar{\mathbf{z}}^u \notin Z^{u,s}\left(f^* + \bar{\varepsilon}_f^u\right)$.

This assumption is needed, since the algorithm proposed uses estimates on the parametric optimal solution function of the lower-level program. For a thorough discussion of this assumption the reader is referred to [22]. Finally, it should be mentioned that the algorithm proposed requires the global solution of mixed-integer nonlinear programs (MINLPs) as subproblems. Therefore, an implicit assumption is that the functions involved in (1) satisfy the requirements by the MINLP solvers, see, e.g., [20,32].

3 Lower bounding procedure

The bilevel program (1) is equivalent to [3,22,35]:

$$\begin{aligned} f^{u,*} = \min_{\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l} & \; f^u\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \\ \text{s.t. } & \mathbf{g}^u\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \le \mathbf{0} \\ & \mathbf{g}^{l,1}\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \le \mathbf{0} \\ & \mathbf{g}^{l,2}\left(\mathbf{x}^l,\mathbf{y}^l\right) \le \mathbf{0} \\ & f^l\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \le f^{l,*}\left(\mathbf{x}^u,\mathbf{y}^u\right) \\ & \mathbf{x}^u \in X^u, \quad \mathbf{y}^u \in Y^u, \\ & \mathbf{x}^l \in X^l, \quad \mathbf{y}^l \in Y^l. \end{aligned} \tag{3}$$

Similar to the continuous case [22] the advantage of using this form is that while the lower-level program may have infinitely many optimal solution points, it always has a unique optimal objective value.

Let now K be an index set for a finite collection of pairs $\left(Z^{u,k},\mathbf{z}^{l,k}\right)$, or equivalently $\left(\left(X^{u,k},Y^{u,k}\right),\left(\mathbf{x}^{l,k},\mathbf{y}^{l,k}\right)\right)$, composed of sets $Z^{u,k} \subset Z^u$ and points $\mathbf{z}^{l,k} \in Z^l$, such that for each $\mathbf{z}^{l,k}$ the lower-level constraints are satisfied for all $\bar{\mathbf{z}}^u \in Z^{u,k}$, i.e.,

$$\begin{aligned} \mathbf{g}^{l,1}\left(\bar{\mathbf{x}}^u,\bar{\mathbf{y}}^u,\mathbf{x}^{l,k},\mathbf{y}^{l,k}\right) & \le \mathbf{0}, \qquad \forall\left(\bar{\mathbf{x}}^u,\bar{\mathbf{y}}^u\right) \in X^{u,k} \times Y^{u,k} \\ \mathbf{g}^{l,2}\left(\mathbf{x}^{l,k},\mathbf{y}^{l,k}\right) & \le \mathbf{0}. \end{aligned} \tag{4}$$

Then, it is easy to verify that a valid relaxation of (3) is given by:

$$\begin{aligned} \text{LBD} = \min_{\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l} & \; f^u\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \\ \text{s.t. } & \mathbf{g}^u\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \le \mathbf{0} \\ & \mathbf{g}^{l,1}\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \le \mathbf{0} \\ & \mathbf{g}^{l,2}\left(\mathbf{x}^l,\mathbf{y}^l\right) \le \mathbf{0} \\ \left(\mathbf{x}^u,\mathbf{y}^u\right) \in X^{u,k} \times Y^{u,k} \Rightarrow & \; f^l\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \le f^l\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^{l,k},\mathbf{y}^{l,k}\right), \quad \forall k \in K \\ & \mathbf{x}^u \in X^u, \quad \mathbf{y}^u \in Y^u, \\ & \mathbf{x}^l \in X^l, \quad \mathbf{y}^l \in Y^l. \end{aligned} \tag{5}$$

The relaxation implies that $\text{LBD} \le f^{u,*}$. The logical constraints can be implemented using a mixed-integer reformulation as described in detail in Sect. 6. Note that to ensure closed feasible sets of (3) the logical constraints are enforced in the interior of $X^{u,k}$.

In the following it is described how to efficiently generate pairs $\left(Z^{u,k},\mathbf{z}^{l,k}\right)$, which will ensure convergence of the lower bound. These steps are similar to the continuous case [22], but some additional considerations are needed. The first step is to fix the upper-level variables $(\mathbf{x}^u,\mathbf{y}^u)$ to the values of the optimal solution $(\bar{\mathbf{x}}^u,\bar{\mathbf{y}}^u)$ obtained by the lower bounding problem (5) and to solve the lower-level problem

$$\begin{aligned}\bar{f}^{l,*} = \min_{\mathbf{x}^m,\mathbf{y}^m} & f^l\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^m, \mathbf{y}^m\right) \\ \text{s.t. } & \mathbf{g}^{l,1}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^m, \mathbf{y}^m\right) \le \mathbf{0} \\ & \mathbf{g}^{l,2}\left(\mathbf{x}^m, \mathbf{y}^m\right) \le \mathbf{0} \\ & \mathbf{x}^m \in X^l, \quad \mathbf{y}^m \in Y^l,\end{aligned} \tag{6}$$

to global optimality. Feasibility of this MINLP is guaranteed since $(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u)$ solve the lower bounding problem (5). The optimal objective value of (6) is also used for the upper bounding procedure, see Sect. 4.

The second step is to pick $\varepsilon^l_{f2} > 0$ and to find a pair $\mathbf{z}^{l,k} \in Z^l$, such that $\mathbf{g}^{l,1}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) < \mathbf{0}$, $\mathbf{g}^{l,2}\left(\mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) \le \mathbf{0}$, and $f^l\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) \le \bar{f}^{l,*} + \varepsilon^l_{f2}$, e.g., by solution of the MINLP

$$\begin{aligned}\psi^* = \min_{\mathbf{x}^m,\mathbf{y}^m} & \psi \\ \text{s.t. } & f^l\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^m, \mathbf{y}^m\right) \le \bar{f}^{l,*} + \varepsilon^l_{f2} \\ & g_i^{l,1}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^m, \mathbf{y}^m\right) \le \psi, \quad i = 1, \ldots, n_g^{l,1} \\ & \mathbf{g}^{l,2}\left(\mathbf{x}^m, \mathbf{y}^m\right) \le \mathbf{0} \\ & \mathbf{x}^m \in X^l, \quad \mathbf{y}^m \in Y^l,\end{aligned} \tag{7}$$

This problem is feasible by the solution of (6). Provided that condition (2) of Assumption 3 is satisfied, ψ^* is negative and the point $\mathbf{z}^{l,k}$ obtained satisfies the required properties. With a slightly different assumption it would be possible to fix the integer variables in the lower-level problem $\mathbf{y}^m$ to the point found by the solution of (6).

The final step is to construct a set $Z^{u,k} \subset Z^u$ such that (4) is satisfied. The set $Z^{u,k} = X^{u,k} \times Y^{u,k}$ should be constructed such that $\bar{\mathbf{x}}^u$ is in the interior of $X^{u,k}$. In the case that $\bar{\mathbf{x}}^u$ is close to the boundary of X^u, the set $X^{u,k}$ should be constructed such that some of its faces coincide with the corresponding faces of the boundary of X^u. Similar to the continuous case [22], the basic methodology by Oluwole et al. [26] can be used to successively guess smaller boxes until a conservative estimate of (4) is satisfied. For a given point $\mathbf{z}^{l,k}$ and a candidate set $Z^{u,k}$ interval analysis [1,24] is employed to overestimate

$$\psi = \max_{\mathbf{x}^u \in X^{u,k},\ \mathbf{y}^u \in Y^{u,k}} \max_i g_i^{l,1}\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right).$$

If $\psi \le 0$, (4) is satisfied. A consequence of this overestimation is that the largest possible $Z^{u,k}$ is not necessarily obtained.

Compared to the continuous case, the presence of integer variables adds an additional degree of complexity. Since the lower-level variables are fixed in the construction of $Z^{u,k}$, no special treatment is needed for the integer variables in the lower-level problem. On the other hand, for the integer variables of the upper-level problem $\mathbf{y}^u$ there are several alternatives. A simple alternative is to fix these variables $\mathbf{y}^u = \bar{\mathbf{y}}^u$ and construct the box $Z^{u,k}$ by varying only the $\mathbf{x}^u$ components. This alternative has the disadvantage that the logical constraints are only valid for this particular value of the integer variables and this could lead to unnecessary iterations, especially when the objective functions and constraints are relatively insensitive to some integer variables. Another alternative is to construct multiple $Z^{u,k}$, each for a different choice of integer variables. This alternative however, leads to an explosion of logical constraints, which is undesired, since it increases the computational requirement for the lower bounding problems. Therefore, here a third alternative is taken, namely to construct $Y^{u,k} \subset \mathbb{Z}^{n_y^u}$, but evaluate the interval extension over the convex hull of $Y^{u,k}$, denoted conv $\left(Y^{u,k}\right)$. This relaxation leads to weaker estimates and therefore smaller $Z^{u,k}$ than possible. Note that $Y^{u,k} \subset \mathbb{Z}^{n_y^u}$ is required for the validity of the logical constraint introduced.

Subroutine 1 (Calculating $Z^{u,k}$)
Given a point $\bar{\mathbf{z}}^u = (\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u)$, a point $\bar{\mathbf{z}}^{l,k} = \left(\mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right)$, and the set Z^u, bounds for the box $Z^{u,k} = \left[\mathbf{z}^{u,k,\mathrm{LO}}, \mathbf{z}^{u,k,\mathrm{UP}}\right]$ are calculated. For simplicity, successively smaller boxes are guessed by scaling the box by $d \in (0, 1]$.

1. Set $d = 1$.
2. **LOOP**

   (a) **FOR** $j = 1, \ldots, n_x + n_y$ **DO**
   - **IF** $\bar{z}_j^u - \frac{d}{2}\left(z_j^{u,\mathrm{UP}} - z_j^{u,\mathrm{LO}}\right) < z_j^{u,\mathrm{LO}}$ **THEN**
     - Set $z_j^{u,k,\mathrm{LO}} = z_j^{u,\mathrm{LO}}$.
     - Set $z_j^{u,k,\mathrm{UP}} = z_j^{u,\mathrm{LO}} + d\left(z_j^{u,\mathrm{UP}} - z_j^{u,\mathrm{LO}}\right)$.
   - **ELSE IF** $\bar{z}_j^u + \frac{d}{2}\left(z_j^{u,\mathrm{UP}} - z_j^{u,\mathrm{LO}}\right) > z_j^{u,\mathrm{UP}}$ **THEN**
     - Set $z_j^{u,k,\mathrm{LO}} = z_j^{u,\mathrm{UP}} - d\left(z_j^{u,\mathrm{UP}} - z_j^{u,\mathrm{LO}}\right)$.
     - Set $z_j^{u,k,\mathrm{UP}} = z_j^{u,\mathrm{UP}}$.
   - **ELSE**
     - Set $z_j^{u,k,\mathrm{LO}} = \bar{z}_j^u - \frac{d}{2}\left(z_j^{u,\mathrm{UP}} - z_j^{u,\mathrm{LO}}\right)$.
     - Set $z_j^{u,k,\mathrm{UP}} = \bar{z}_j^u + \frac{d}{2}\left(z_j^{u,\mathrm{UP}} - z_j^{u,\mathrm{LO}}\right)$.

     END

   (b) **FOR** $j = 1, \ldots, n_y$ **DO**
   - Set $y_j^{u,k,\mathrm{LO}} = \lceil y_j^{u,k,\mathrm{LO}} \rceil$.
   - Set $y_j^{u,k,\mathrm{UP}} = \lfloor y_j^{u,k,\mathrm{UP}} \rfloor$.

   (c) Check (4) by evaluating the interval extension of $\mathbf{g}^{l,1}\left(\cdot, \cdot, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right)$ on $X^{u,k} \times$ conv $\left(Y^{u,k}\right)$.
   IF (4) is satisfied **THEN** terminate **ELSE** Set $d = d/2$ **END**.

   END

At the end of the subroutine $Z^{u,k} = X^{u,k} \times Y^{u,k}$ is obtained, where $X^{u,k} = \left[\mathbf{x}^{u,k,\mathrm{LO}}, \mathbf{x}^{u,k,\mathrm{UP}}\right] \subset X^u$ and $Y^{u,k} = \left[\mathbf{y}^{u,k,\mathrm{LO}}, \mathbf{y}^{u,k,\mathrm{UP}}\right] \subset Y^u$. For the pair $\left(Z^{u,k}, \mathbf{z}^{l,k}\right)$ (4) is satisfied. The reason for using the floor ($\lfloor \cdot \rfloor$) and ceil ($\lceil \cdot \rceil$) functions is to ensure that $\mathbf{y}^{u,k,\mathrm{LO}}$, $\mathbf{y}^{u,k,\mathrm{UP}}$ are integer-valued and also to accelerate convergence of the interval extension.

Here, it is assumed that d is reduced by half at each iteration. In general it is sufficient to consider a sequence of $d > 0$ converging to zero. Then, Subroutine 1 is finite by the same arguments as in the continuous case [22]. The computational requirement for this subroutine is typically insignificant compared to the lower bounding problems. For an efficient implementation the details of this procedure are important and should be tuned for the instance considered.

3.1 KKT-based tightening of lower bounding problem

In the following a tightening of the lower bounding problem is proposed based on the KKT necessary conditions in the lower-level program for the case that $n_x^l > 0$. This requires additional assumptions on the lower-level program:

Assumption 4 (*Assumptions for the KKT-based lower bound*) Let

$$Q^\infty = \Big\{\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{y}^l\right) \in X^u \times Y^u \times Y^l : \exists \mathbf{x}^l \in X^l : \mathbf{g}^u\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) \leq \mathbf{0}, \\ \mathbf{g}^{l,1}\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) \leq \mathbf{0}, \mathbf{g}^{l,2}\left(\mathbf{x}^l, \mathbf{y}^l\right) \leq \mathbf{0}\Big\}.$$

For all $\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \bar{\mathbf{y}}^l\right) \in Q^\infty$ the following three conditions hold: (i) differentiability of $f^l(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \cdot, \bar{\mathbf{y}}^l)$, $\mathbf{g}^{l,1}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \cdot, \bar{\mathbf{y}}^l\right)$ and $\mathbf{g}^{l,2}\left(\cdot, \bar{\mathbf{y}}^l\right)$ on some open set embedding X^l, (ii) a constraint qualification for the lower-level program (restricted to fixed $\bar{\mathbf{y}}^l$), and (iii) a priori known upper bounds for the KKT multipliers.

The first two parts of Assumption 4 are standard for smooth NLP solvers. The third part is not easy to verify for a general lower-level program, but there are many interesting applications for which it is satisfied. For instance, in feasibility and flexibility problems [17] the KKT multipliers are bounded above by one. Also, for semi-infinite programs bounds for the KKT multipliers of the lower-level program can be readily estimated [23]. In the case of nonunique KKT multipliers, the a priori specified upper bounds $\boldsymbol{\mu}^{\max}$ need not permit the entire set of possible multipliers, but rather at least one member of the set. For problems for which any of the three parts is violated, the KKT-based lower bounding problem is not applicable, see also [22], where the importance of the a priori bounds is discussed.

Consider for simplicity the combination of constraints $\left(\mathbf{g}^{l,1}, \mathbf{g}^{l,2} \text{ and } \mathbf{x}^l \in \left[\mathbf{x}^{l,\mathrm{LO}}, \mathbf{x}^{l,\mathrm{UP}}\right]\right)$ of the lower-level program, i.e.:

$$\tilde{g}_j\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) = \begin{cases} g_j^{l,1}\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right), & j = 1, \ldots, n_g^{l,1} \\ g_{j-n_g^{l,1}}^{l,2}\left(\mathbf{x}^l, \mathbf{y}^l\right), & j = n_g^{l,1}+1, \ldots, n_g^{l,1}+n_g^{l,2} \\ x_{j-n_g^{l,1}-n_g^{l,2}}^l - x_{j-n_g^{l,1}-n_g^{l,2}}^{l,\mathrm{UP}}, & j = n_g^{l,1}+n_g^{l,2}+1, \ldots, n_g^{l,1}+n_g^{l,2}+n_x^l \\ -x_{j-n_g^{l,1}-n_g^{l,2}-n_x^l}^l + x_{j-n_g^{l,1}-n_g^{l,2}-n_x^l}^{l,\mathrm{LO}}, & j = n_g^{l,1}+n_g^{l,2}+n_x^l+1, \ldots, n_g^{l,1}+n_g^{l,2}+2\,n_x^l. \end{cases}$$

If Assumption 4 is satisfied, the lower bounding problem (5) can be tightened by further requiring that $\mathbf{x}^l$ satisfies the KKT necessary conditions for the lower-level program (for fixed $\bar{\mathbf{y}}^l$). The KKT multipliers $\boldsymbol{\mu}$ are added to the set of variables

$$\begin{aligned} \mathrm{LBD} = \min_{\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l, \boldsymbol{\mu}} \; & f^u\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) \\ \text{s.t. } & \mathbf{g}^u\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) \leq \mathbf{0} \\ & \mathbf{g}^{l,1}\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) \leq \mathbf{0} \\ & \mathbf{g}^{l,2}\left(\mathbf{x}^l, \mathbf{y}^l\right) \leq \mathbf{0} \qquad (8) \\ & \left(\mathbf{x}^u, \mathbf{y}^u\right) \in X^{u,k} \times Y^{u,k} \Rightarrow f^l\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) \leq f^l\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right), \quad \forall k \in K \end{aligned}$$

$$
\begin{aligned}
\nabla_{\mathbf{x}^l} f^l \left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) + \boldsymbol{\mu}^{\mathrm{T}} \nabla_{\mathbf{x}^l} \tilde{\mathbf{g}} \left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) = \mathbf{0} \\
\mu_j \tilde{g}_j \left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) = 0, \quad j = 1, \ldots, n_p + n_q + 2n_x^l \\
\mu_j \in \left[0, \mu_j^{\max}\right], \quad j = 1, \ldots, n_p + n_q + 2n_x^l \\
\mathbf{x}^u \in X^u, \quad \mathbf{y}^u \in Y^u, \\
\mathbf{x}^l \in X^l, \quad \mathbf{y}^l \in Y^l.
\end{aligned}
$$

Unlike the continuous case, even if the functions in the lower-level program are convex, using the KKT-based tightening does not ensure convergence in the first iteration. Note also that no multipliers are needed for the constraints which do not depend on the continuous lower-level variables.

4 Upper bounding problem

The upper bounding procedure of the continuous case [22] can essentially be also applied to the mixed-integer case. The only significant difference is that the resulting optimization problem is a MINLP as opposed to a NLP. Recall that in the lower bounding procedure the lower-level program (6) is solved for a candidate $\bar{\mathbf{z}}^u$. By the selection of $\bar{\mathbf{z}}^u$, this MINLP is feasible. Therefore its optimal objective value $\bar{f}^{l,*}$ is finite. To obtain an upper bound the following MINLP can be solved for fixed $\bar{\mathbf{z}}^u$

$$
\begin{aligned}
\mathrm{UBD} = \min_{\mathbf{x}^l, \mathbf{y}^l} \; & f^u \left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^l, \mathbf{y}^l\right) \\
\text{s.t. } & \mathbf{g}^u \left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^l, \mathbf{y}^l\right) \leq \mathbf{0} \\
& \mathbf{g}^{l,1} \left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^l, \mathbf{y}^l\right) \leq \mathbf{0} \\
& \mathbf{g}^{l,2} \left(\mathbf{x}^l, \mathbf{y}^l\right) \leq \mathbf{0} \\
& f^l \left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^l, \mathbf{y}^l\right) \leq \bar{f}^{l,*} + \varepsilon_f^l \\
& \mathrm{LBD} \leq f^u \left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^l, \mathbf{y}^l\right) \\
& \mathbf{x}^l \in X^l, \quad \mathbf{y}^l \in Y^l,
\end{aligned}
\tag{9}
$$

allowing an ε_f^l-violation of the lower-level objective. This step is performed, because, due to potential non-uniqueness of the solutions of the lower-level program, a valid upper bound may be obtained even if the solution to (6) does not satisfy the upper-level constraints. If (9) is infeasible then no solution exists for $(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u)$; otherwise an upper bound is obtained. The inequality LBD $\leq f^u \left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^l, \mathbf{y}^l\right)$ is added to accelerate convergence of (9) and to alleviate partially the consequences of allowing ε_f^l-optimality in the lower-level program. Note that the solution of (9) is only an upper bound in the sense of an $\varepsilon-$feasible point, i.e., a point that satisfies the constraints of the lower- and upper-level programs and is a ε_f^l-optimal point in the lower-level program.

5 Algorithm

The overall algorithm is very similar to the continuous case [22]. Recall though that the subproblems solved by the algorithm are changed as described in the previous sections. The basic principle of the algorithm is that the addition of parametric upper bounds on the optimal solution value of the lower-level program via the pairs $\left(Z^{u,k}, \mathbf{z}^{l,k}\right)$ makes the lower bounding problems successively tighter. Recall that the generation of parametric upper bounds is possible due to condition (2) of Assumption 3. Finite termination is guaranteed, since, at the worse case, a finite number of sets $Z^{u,k}$ is constructed that covers $Z^{u,s}\ (f^{u,*})$.

Input to the algorithm are the optimality tolerances $\varepsilon_{\text{MINLP}}$, ε_f^u and ε_f^l, satisfying the assumptions of Theorem 1.

Algorithm 1

1. (**Initialization**)
   Set LBD $= -\infty$, UBD $= +\infty$, $k = 1$, $K = \emptyset$.
2. (**Lower bounding**)
   Solve (5) or (8) to $\varepsilon_{\text{MINLP}}$ optimality.
   IF Feasible **THEN**
   - Set LBD to the optimal objective value (final lower bound).
   - Set $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ equal to the solution point ($\varepsilon_{\text{MINLP}}-$optimal point).

   ELSE (**Infeasible problem**)
   - Terminate.

   END
3. (**Termination**)
   IF LBD $\geq$ UBD $- \varepsilon_f^u$ **THEN** Terminate.
4. (**Lower-level problem**)
   Solve MINLP (6) to $\varepsilon_{\text{MINLP}}$ optimality. (Recall that feasibility of this program is guaranteed.)
   Set $f^{l,*}$ equal to the optimal objective value (final lower bound).
5. (**Populate parametric upper bounds to lower-level problem**)
   Solve (7) to $\varepsilon_{\text{MINLP}}$ optimality. (Recall that feasibility of this program is guaranteed.)
   - Set $\left(\mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right)$ equal to the solution point.
   - Apply Subroutine 1 to obtain a set $Z^{u,k}$.
   - Insert k to K.
   - Set $k = k + 1$.
6. (**Upper bounding**)
   Solve MINLP (9) and (if feasible) obtain an $\varepsilon_{\text{MINLP}}-$optimal point $\left(\bar{\mathbf{x}}^l, \bar{\mathbf{y}}^l\right)$.
   IF Feasible and $f^u\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \bar{\mathbf{x}}^l, \bar{\mathbf{y}}^l\right) <$ UBD **THEN** set UBD $= f\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \bar{\mathbf{x}}^l, \bar{\mathbf{y}}^l\right)$ and $\left(\mathbf{x}^{u,*}, \mathbf{y}^{u,*}, \mathbf{x}^{l,*}, \mathbf{y}^{l,*}\right) = \left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \bar{\mathbf{x}}^l, \bar{\mathbf{y}}^l\right)$.
7. (**Loop**)
   IF LBD $\geq$ UBD $- \varepsilon_f^u$ **THEN** Terminate **ELSE** Goto step 2.

On termination of the algorithm, if UBD $= +\infty$, the instance is infeasible. Otherwise, UBD is an ε_f^u-estimate of the optimal solution value $\left(\text{UBD} \leq f^{u,*} + \varepsilon_f^u\right)$ and $\left(\mathbf{x}^{u,*}, \mathbf{y}^{u,*}, \mathbf{x}^{l,*}, \mathbf{y}^{l,*}\right)$ is an $\varepsilon-$optimal point at which UBD is attained.

5.1 Convergence proof

In this section a convergence proof for Algorithm 1 is given. Note again that no convexity or uniqueness assumptions are made for either the upper- or lower-level programs. The proof outline follows the continuous case [22]. However, the addition of the integrality constraints mandates alteration of the proofs.

Lemma 1 (Continuity of optimal solution function of lower-level problem) *For any* $(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u) \in Z^{u,\infty}$ *satisfying* (2) *the optimal objective function* $f^{l,*}(\cdot, \bar{\mathbf{y}}^u)$ *of the lower-level problem is continuous at* $\bar{\mathbf{x}}^u$.

Proof Consider any fixed $(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u) \in Z^{u,\infty}$. By (2) for any $\varepsilon^l_{f1} > 0$, there exists $\left(\tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right) \in X^l \times Y^l$ such that

$$\mathbf{g}^{l,1}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right) < \mathbf{0}, \qquad \mathbf{g}^{l,2}\left(\tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right) \leq \mathbf{0} \tag{10}$$

$$f^l\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right) \leq f^{l,*}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u\right) + \varepsilon^l_{f1}. \tag{11}$$

By continuity of the lower-level objective $f^l\left(\cdot, \bar{\mathbf{y}}^u, \tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right)$ on X^u, for any $\varepsilon^l_{f3} > 0$ there exists $\delta_1 > 0$ such that

$$f^l\left(\mathbf{x}^u, \bar{\mathbf{y}}^u, \tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right) < f^l\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right) + \varepsilon^l_{f3}, \quad \forall \mathbf{x}^u \in X^u : ||\bar{\mathbf{x}}^u - \mathbf{x}^u|| < \delta_1. \tag{12}$$

Combining inequalities (11) and (12) it follows:

$$f^l\left(\mathbf{x}^u, \bar{\mathbf{y}}^u, \tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right) < f^{l,*}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u\right) + \varepsilon^l_{f1} + \varepsilon^l_{f3}, \quad \forall \mathbf{x}^u \in X^u : ||\bar{\mathbf{x}}^u - \mathbf{x}^u|| < \delta_1. \tag{13}$$

By (10) and continuity of $\mathbf{g}^{l,1}\left(\cdot, \bar{\mathbf{y}}^u, \tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right)$, there exists $\delta_2 > 0$ such that

$$\mathbf{g}^{l,1}\left(\mathbf{x}^u, \bar{\mathbf{y}}^u, \tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right) \leq \mathbf{0}, \quad \forall \mathbf{x}^u \in X^u : ||\bar{\mathbf{x}}^u - \mathbf{x}^u|| < \delta_2.$$

Together with $\mathbf{g}^{l,2}\left(\tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right) \leq \mathbf{0}$, $\left(\tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right)$ is feasible in the lower-level program for all $\mathbf{x}^u \in X^u : ||\bar{\mathbf{x}}^u - \mathbf{x}^u|| < \delta_2$ and fixed $\bar{\mathbf{y}}^u$. By the definition of $f^{l,*}$ we therefore have

$$f^{l,*}\left(\mathbf{x}^u, \bar{\mathbf{y}}^u\right) \leq f^l\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \tilde{\mathbf{x}}^l, \tilde{\mathbf{y}}^l\right), \quad \forall \mathbf{x}^u \in X^u : ||\bar{\mathbf{x}}^u - \mathbf{x}^u|| < \delta_2.$$

With (13) we obtain

$$f^{l,*}\left(\mathbf{x}^u, \bar{\mathbf{y}}^u\right) < f^{l,*}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u\right) + \varepsilon^l_{f1} + \varepsilon^l_{f3}, \quad \forall \mathbf{x}^u \in X^u : ||\bar{\mathbf{x}}^u - \mathbf{x}^u|| < \min\{\delta_1, \delta_2\}$$

which proves that $f^{l,*}(\cdot, \bar{\mathbf{y}}^u)$ is upper semi-continuous at $\bar{\mathbf{x}}^u$.

By Theorem 4.2.1 in Bank et al. [2] for a fixed $\bar{\mathbf{y}}^u$ for all $\mathbf{x}^u \in X^u : (\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u) \in Z^{u,\infty}$ the optimal objective function $f^{l,*}(\cdot, \bar{\mathbf{y}}^u)$ of the lower-level problem is lower semi-continuous.

Lemma 2 (Minimum of bilevel program exists) *Under Assumptions 1, 2 and 3, either* (1) *is infeasible or the minimum of* (1) *exists.*

Proof Let for now $f^{u,*}$ denote the infimum of (1) without asserting that the minimum is attained. By Definition 4 of the level sets and the definition of the infimum, the bilevel program (1) is equivalent to

$$
\begin{aligned}
&\min_{\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l} f^u\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right)\\
&\text{s.t. } f^u\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \le f^{u,*}\\
&\quad \mathbf{g}^u\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \le \mathbf{0}\\
&\quad f^l\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \le f^{l,*}\left(\mathbf{x}^u,\mathbf{y}^u\right)\\
&\quad \mathbf{g}^{l,1}\left(\mathbf{x}^u,\mathbf{y}^u,\mathbf{x}^l,\mathbf{y}^l\right) \le \mathbf{0}\\
&\quad \mathbf{g}^{l,2}\left(\mathbf{x}^l,\mathbf{y}^l\right) \le \mathbf{0}\\
&\quad \mathbf{x}^u \in X^u \subset \mathbb{R}^{n_x^u},\quad \mathbf{y}^u \in Y^u \subset \mathbb{Z}^{n_y^u},\\
&\quad \mathbf{x}^l \in X^l \subset \mathbb{R}^{n_x^l},\quad \mathbf{y}^l \in Y^l \subset \mathbb{Z}^{n_y^l}.
\end{aligned}
\tag{14}
$$

Only points $(\mathbf{x}^u,\mathbf{y}^u) \in Z^{u,s}\left(f^{u,*}\right)$ are candidates of the feasible set in this augmented problem. For these points condition (2) is satisfied under Assumption 3. Therefore, by Lemma 1 $f^{l,*}$ is upper-semi continuous on the feasible set of (14). Therefore, by Weierstrass' theorem either (14) is infeasible or its minimum is attained. As a consequence either (1) is infeasible or the minimum of (1) exists. □

The following Lemma shows that the algorithm generates points $\left(\mathbf{x}^{l,k},\mathbf{y}^{l,k}\right)$ for which it is possible to construct sets $X^{u,k} \times Y^{u,k}$ such that $X^{u,k}$ have a nonempty interior. Therefore, Subroutine 1 leads to finite coverage of the level-set $Z^{u,s}\left(f^{u,*}\right)$.

Lemma 3 (Existence of sets $X^{u,k}$ with nonempty interior) *Consider a fixed* $\bar{\mathbf{y}}^u$. *For any (arbitrary but fixed)* $\varepsilon^l_{f2} > 0$ *there exists* $\delta_1 > 0$ *such that for any* $\bar{f}^u \le f^{u,*} + \bar{\varepsilon}^u_f$ *and for each point* $\bar{\mathbf{x}}^u \in X^u : (\bar{\mathbf{x}}^u,\bar{\mathbf{y}}^u) \in Z^{u,s}\left(\bar{f}^u\right)$ *the points* $\left(\mathbf{x}^{l,k},\mathbf{y}^{l,k}\right)$ *generated in Step 5 of Algorithm 1 satisfy*

$$
\begin{aligned}
f^l\left(\bar{\mathbf{x}}^u,\bar{\mathbf{y}}^u,\mathbf{x}^{l,k},\mathbf{y}^{l,k}\right) &\le f^{l,*}\left(\bar{\mathbf{x}}^u,\bar{\mathbf{y}}^u\right) + \varepsilon^l_{f2}\\
\mathbf{g}^{l,1}\left(\mathbf{x}^u,\bar{\mathbf{y}}^u,\mathbf{x}^{l,k},\mathbf{y}^{l,k}\right) &\le \mathbf{0},\quad \forall \mathbf{x}^u \in X^u : ||\mathbf{x}^u - \bar{\mathbf{x}}^u|| < \delta_1\\
\mathbf{g}^{l,2}\left(\mathbf{x}^{l,k},\mathbf{y}^{l,k}\right) &\le \mathbf{0}.
\end{aligned}
$$

Note that existence of a δ_1 valid for all $\bar{\mathbf{x}}^u$ is shown.

Proof Let $\bar{X}^{u,s} = \{\bar{\mathbf{x}} \in X^u : (\bar{\mathbf{x}}^u,\bar{\mathbf{y}}^u) \in Z^{u,s}\left(\bar{f}^u\right)\}$. Since $Z^{u,s}\left(\bar{f}^u\right)$ is compact, so is $\bar{X}^{u,s}$. Since $\bar{f}^u \le f^{u,*} + \bar{\varepsilon}^u_f$, all points $(\bar{\mathbf{x}}^u,\bar{\mathbf{y}}^u) \in Z^{u,s}\left(\bar{f}^u\right)$ satisfy (2). Therefore, by Lemma 1 the optimal objective function of the lower-level problem $f^{l,*}(\cdot,\bar{\mathbf{y}}^u)$ is continuous at all $\bar{\mathbf{x}} \in \bar{X}^{u,s}$.

Let ψ^* denote the parametric optimal solution value of (7) for fixed $\bar{\mathbf{y}}^u$. By the continuity of the functions and the compactness of $\bar{X}^{u,s}$, ψ^* is continuous and its maximum over $\bar{\mathbf{x}}^u \in \bar{X}^{u,s}$ is attained. Since $\varepsilon^l_{f2} > 0$, by (2) ψ^* is strictly negative on $\bar{X}^{u,s}$. Therefore, there exists $\bar{\psi}^* < 0$ such that for all $\bar{\mathbf{x}} \in \bar{X}^{u,s}$:

$$
\begin{aligned}
g_i^{l,1}\left(\bar{\mathbf{x}}^u,\bar{\mathbf{y}}^u,\mathbf{x}^{l,k},\mathbf{y}^{l,k}\right) \le \bar{\psi}^* < 0,\ i = 1,\ldots,n_g^{l,1}\quad \mathbf{g}^{l,2}\left(\bar{\mathbf{x}}^{l,k},\mathbf{y}^{l,k}\right) \le \mathbf{0},\\
f^l\left(\bar{\mathbf{x}}^u,\bar{\mathbf{y}}^u,\mathbf{x}^{l,k},\mathbf{y}^{l,k}\right) \le f^{l,*}\left(\bar{\mathbf{x}}^u,\bar{\mathbf{y}}^u\right) + \varepsilon^l_{f2}.
\end{aligned}
$$

Since $\mathbf{g}^{l,1}\left(\cdot, \bar{\mathbf{y}}^u, \mathbf{x}^l, \mathbf{y}^l\right)$ is continuous, and $\bar{X}^{u,s}$ is compact, $\mathbf{g}^{l,1}\left(\cdot, \bar{\mathbf{y}}^u, \mathbf{x}^l, \mathbf{y}^l\right)$ is uniformly continuous on $\bar{X}^{u,s}$. Therefore, there exists $\delta_1 > 0$ (independent of $\bar{\mathbf{x}}^u$) such that for any $\bar{\mathbf{x}}^u \in \bar{X}^{u,s}$:

$$f^l\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) \leq f^{l,*}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u\right) + \varepsilon^l_{f2}$$
$$\mathbf{g}^{l,1}\left(\mathbf{x}^u, \bar{\mathbf{y}}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) \leq \mathbf{0}, \quad \forall \mathbf{x}^u \in X^u : ||\mathbf{x}^u - \bar{\mathbf{x}}^u|| < \delta_1$$
$$\mathbf{g}^{l,2}\left(\mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) \leq \mathbf{0}. \qquad \square$$

Note that Subroutine 1 underestimates δ_1, but it still furnishes a positive value.
The following Lemma is proved in [22]:

Lemma 4 *Let $X_t \subset X$ be compact, and $\delta > 0$. Consider any infinite sequence of points $\mathbf{x}^i \in X_t$. There exists a finite index $I > 0$, such that*

$$||\mathbf{x}^I - \mathbf{x}^i|| \leq \delta, \qquad \textit{for some } i < I.$$

Theorem 1 (Finite termination) *If the tolerances of the optimization subproblems $\varepsilon_{\text{MINLP}}$ and ε^l_{f2} in* (7) *satisfy*

$$0 < \varepsilon_{\text{MINLP}} \leq \min\{\varepsilon^u_f/2, \bar{\varepsilon}^u_f, \varepsilon^l_{f1}\}$$
$$0 < \varepsilon^l_{f2} < \varepsilon^l_f - \varepsilon_{\text{MINLP}},$$

then Algorithm 1 terminates finitely.

Note that the validity of the upper- and lower-bounding problem guarantee that at finite termination a global solution (within tolerances) of (1) is obtained.

Since the proof of Theorem 1 is lengthy, first an outline of the proof is presented. Since the lower bounding problem is solved globally, it only visits points $\bar{\mathbf{z}} \in Z^{u,s}\left(f^{u,*} + \varepsilon_{\text{MINLP}}\right)$. By Assumption 3 at these points, parametric upper bounds to the optimal solution value of the inner program can be constructed via the pairs $\left(Z^{u,k}, \mathbf{z}^{l,k}\right)$. These pairs lead to a successive tightening of the lower bounding problem. After a finite number of iterations it either becomes infeasible or it furnishes a point inside an existing $Z^{u,k}$ which is also a $\varepsilon-$optimal point. The proof follows the outline of the continuous case (Theorem 1 in [22]). Therein, continuity with respect to the upper-level variables was used, but not with respect to the lower-level variables. The set Y^u is finite and the algorithm only visits points $\bar{\mathbf{y}} \in Y^u$. The proof considers fixed $\bar{\mathbf{y}}$, since at the worst case explicit enumeration of the integer variables will be performed. In typical instances convergence will occur without explicit enumeration of the integer variables.

Proof Let $\bar{f}^u = f^{u,*} + \bar{\varepsilon}^u_f$.

The lower bounding problem is a valid relaxation of the bilevel program and is solved globally. Therefore, only points $\bar{\mathbf{z}} \in Z^{u,s}\left(f^{u,*} + \varepsilon_{\text{MINLP}}\right)$ are furnished by the lower bounding problem. By assumption $\varepsilon_{\text{MINLP}} < \bar{\varepsilon}^u_f$ and therefore, $\bar{\mathbf{z}} \in Z^{u,s}\left(\bar{f}^u\right)$. By Assumption 3 condition (2) is satisfied and this allows the generation of logical constraints via the pairs $\left(Z^{u,k}, \mathbf{z}^{l,k}\right)$.

Let $\bar{\mathbf{z}}^u \in Z^{u,s}\left(\bar{f}^u\right)$ be furnished by the lower bounding problem. By definition $\bar{\mathbf{z}}^u = \left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u\right)$. We will show that if in a subsequent iteration the lower bounding problem furnishes a point $\left(\hat{\mathbf{x}}^u, \hat{\mathbf{y}}^u, \hat{\mathbf{x}}^l, \hat{\mathbf{y}}^l\right)$, with $\hat{\mathbf{y}}^u = \bar{\mathbf{y}}^l$ and $\hat{\mathbf{x}}^u$ sufficiently close to $\bar{\mathbf{x}}^u$, this point $\left(\hat{\mathbf{x}}^u, \hat{\mathbf{y}}^u, \hat{\mathbf{x}}^l, \hat{\mathbf{y}}^l\right)$ is $\varepsilon-$optimal.

By Lemma 3 there exists $\delta_1 > 0$ such that for $\bar{\mathbf{x}}^u$ the points $\mathbf{z}^{l,k} = \left(\bar{\mathbf{x}}^{l,k}, \bar{\mathbf{y}}^{l,k}\right)$ generated in Step 5 of Algorithm 1 satisfy

$$\mathbf{g}^{l,1}\left(\mathbf{x}^u, \bar{\mathbf{y}}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) \leq \mathbf{0}, \quad \forall \mathbf{x}^u \in X^u : ||\mathbf{x}^u - \bar{\mathbf{x}}^u|| < \delta_1, \quad \mathbf{g}^{l,2}\left(\mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) \leq \mathbf{0}$$
$$f^l\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) \leq f^{l,*}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u\right) + \varepsilon^l_{f2}. \tag{15}$$

By assumption $\varepsilon^l_f - \varepsilon_{\text{MINLP}} - \varepsilon^l_{f2} > 0$. By continuity of $f^{l,*}\left(\cdot, \bar{\mathbf{y}}^u\right)$ at $\bar{\mathbf{x}}^u$ there exists $\delta_2 > 0$ such that

$$f^{l,*}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u\right) \leq f^{l,*}\left(\mathbf{x}^u, \bar{\mathbf{y}}^u\right) + \left(\varepsilon^l_f - \varepsilon^l_{f2} - \varepsilon_{\text{MINLP}}\right)/2, \quad \forall \mathbf{x}^u \in X^u : ||\mathbf{x}^u - \bar{\mathbf{x}}^u|| < \delta_2.$$

By continuity of $f^l\left(\cdot, \bar{\mathbf{y}}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right)$ on X^u, there exists $\delta_3 > 0$ such that

$$f^l\left(\mathbf{x}^u, \bar{\mathbf{y}}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) \leq f^l\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) + \left(\varepsilon^l_f - \varepsilon^l_{f2} - \varepsilon_{\text{MINLP}}\right)/2,$$
$$\forall \mathbf{x}^u \in X^u : ||\mathbf{x}^u - \bar{\mathbf{x}}^u|| < \delta_3.$$

Combining these last three inequalities gives

$$f^l\left(\mathbf{x}^u, \bar{\mathbf{y}}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) \leq f^{l,*}\left(\mathbf{x}^u, \bar{\mathbf{y}}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) + \varepsilon^l_f - \varepsilon_{\text{MINLP}},$$
$$\forall \mathbf{x}^u \in X^u : ||\mathbf{x}^u - \bar{\mathbf{x}}^u|| < \min\{\delta_2, \delta_3\}.$$

Therefore, together with (15), $\mathbf{x}^{l,k}, \mathbf{y}^{l,k}$ is ε_h-optimal in the lower-level problem for $\mathbf{y}^u = \bar{\mathbf{y}}^u$ and all $\mathbf{x}^u \in X^u : ||\mathbf{x}^u - \bar{\mathbf{x}}^u|| < \delta$, where $\delta = \min\{\delta_1, \delta_2, \delta_3\} > 0$. Note that these $\left(\mathbf{x}^u, \bar{\mathbf{y}}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right)$ are not necessarily feasible with respect to the upper-level constraints, and therefore termination does not occur immediately.

Recall that $Z^{u,s}\left(\bar{f}^u\right) = \left(X^{u,s}\left(\bar{f}^u\right), Y^{u,s}\left(\bar{f}^u\right)\right)$ and note that $Y^{u,s}\left(\bar{f}^u\right)$ is finite and $X^{u,s}\left(\bar{f}^u\right)$ compact. Since $\delta > 0$, by Lemma 4 after a finite number of iterations either the lower bounding problem becomes infeasible, in which case the algorithm terminates, or the lower bounding problem furnishes a point $\left(\hat{\mathbf{x}}^u, \hat{\mathbf{y}}^u, \hat{\mathbf{x}}^l, \hat{\mathbf{y}}^l\right)$, with $\hat{\mathbf{y}}^u = \bar{\mathbf{y}}^u$ and $\hat{\mathbf{x}}^u$ sufficiently close to $\bar{\mathbf{x}}^u$, i.e., $||\hat{\mathbf{x}}^u - \bar{\mathbf{x}}^u|| < \delta$. By construction of the lower bounding problem, this $\left(\hat{\mathbf{x}}^u, \hat{\mathbf{y}}^u, \hat{\mathbf{x}}^l, \hat{\mathbf{y}}^l\right)$ satisfies the lower- and upper-level constraints and also $f^l\left(\hat{\mathbf{x}}^u, \hat{\mathbf{y}}^u, \hat{\mathbf{x}}^l, \hat{\mathbf{y}}^l\right) \leq f^l\left(\hat{\mathbf{x}}^u, \hat{\mathbf{y}}^u, \hat{\mathbf{x}}^{l,k}, \hat{\mathbf{y}}^{l,k}\right)$ (by the logical constraint) and as a consequence

$$f^l\left(\hat{\mathbf{x}}^u, \hat{\mathbf{y}}^u, \hat{\mathbf{x}}^l, \hat{\mathbf{y}}^l\right) \leq f^{l,*}\left(\hat{\mathbf{x}}^u, \hat{\mathbf{y}}^u\right) + \varepsilon^l_f - \varepsilon_{\text{MINLP}}$$

or $\left(\hat{\mathbf{x}}^l, \hat{\mathbf{y}}^l\right)$ is ε^l_f-optimal in the lower-level problem for $\left(\hat{\mathbf{x}}^u, \hat{\mathbf{y}}^u\right)$. Note that the $\varepsilon_{\text{MINLP}}$ tolerance is included here, because the global solution of the lower-level problem only gives a $\varepsilon_{\text{MINLP}}-$ estimate of $f^{l,*}\left(\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u\right)$. The lower bound LBD obtained satisfies LBD $\geq f^u\left(\hat{\mathbf{x}}^u, \hat{\mathbf{y}}^u, \hat{\mathbf{x}}^l, \hat{\mathbf{y}}^l\right) - \varepsilon_{\text{MINLP}}$. The point $\left(\hat{\mathbf{x}}^u, \hat{\mathbf{y}}^u, \hat{\mathbf{x}}^l, \hat{\mathbf{y}}^l\right)$ is feasible in the bilevel program and therefore the upper bounding problem (Step 6) furnishes an upper bound UBD, satisfying UBD $\leq f\left(\hat{\mathbf{x}}^u, \hat{\mathbf{y}}^u, \hat{\mathbf{x}}^l, \hat{\mathbf{y}}^l\right) + \varepsilon_{\text{MINLP}}$. Noting now that the optimization problems are solved with tolerance $\varepsilon_{\text{MINLP}} < \varepsilon^u_f/2$ it follows UBD $-$ LBD $\leq 2\,\varepsilon_{\text{MINLP}} \leq \varepsilon^u_f$ and the algorithm terminates. □

5.2 Illustrative examples

Example 5.1 As a first illustrative example consider the following box-constrained bilevel program:

$$\min_{x^u, y^l} x^u - 2\,x^u\,y^l - 0.1y^l$$
$$\text{s.t. } y^l \in \arg\min_{y^m} y^m$$
$$x^u \in [-1, 1], \quad y^l, y^m \in \{0, 1\}$$

with the unique optimal solution $x^u = -1$, $y^l = 0$ and an objective value of -1. This example demonstrates how the algorithm handles integer variables in the lower-level program (in the worse case). Note that the bilinear term $2\,x^u\,y^l$ could be reformulated exactly [16], but there is no need to do so.

Consider the application of Algorithm 1. The lower bounding problem without KKT-based tightening (5) is used since the lower-level program does not have any continuous variables. At the first iteration the lower bounding problem

$$\min_{x^u \in [-1,1], y^l \in \{0,1\}} x^u - 2\,x^u\,y^l - 0.1y^l$$

is solved and $\bar{x}^u = 1$, $\bar{y}^l = 1$, LBD $= -1.1$ is obtained. Then the lower-level problem is solved for $\bar{x}^u = 1$

$$\min_{y^m \in \{0,1\}} y^m$$

and $\bar{y}^l = 0$, $\bar{f}^{l,*} = 0$ is obtained. Since the constraints of the lower-level problem do not depend on x^u, the pair $([-1, 1], 0)$ is used for the parametric upper bounds of the lower-level problem. The first iteration is concluded by solving the augmented upper bounding problem for $\bar{x}^u = 1$:

$$\min_{y^l \in \{0,1\}} 1 - 2\,y^l - 0.1\,y^l,$$
$$\text{s.t. } y^l \leq 0$$

obtaining the upper bound $\bar{x}^u = 1$, $\bar{y}^l = 0$ with UBD $= 1$.

A second iteration is performed. Now the lower bounding problem contains a parametric upper bound to the lower-level problem

$$\min_{x^u \in [-1,1], y^l \in \{0,1\}} x^u - 2\,x^u\,y^l - 0.1y^l,$$
$$\text{s.t. } y^l \leq 0.$$

The solution of this problem gives the unique optimal point of the bilevel problem $x^u = -1$, $y^l = 0$ with LBD $= -1$. Then the lower-level problem is solved for $\bar{x}^u = -1$

$$\min_{y^m \in \{0,1\}} y^m$$

and $\bar{y}^l = 0$, $\bar{f}^{l,*} = 0$ is obtained. Since the constraints of the lower-level problem do not depend on x^u, the pair $([-1, 1], 0)$ is used for the parametric upper bounds of the lower-level problem. The second iteration is concluded by solving the augmented upper bounding problem for $\bar{x}^u = -1$:

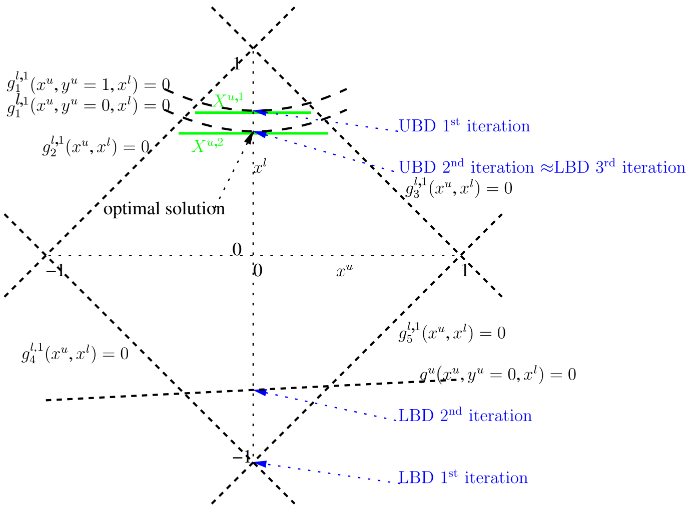

Fig. 1 Graphical illustration of Example 5.2

$$
\begin{aligned}
&\min_{y^l \in \{0,1\}} \; -1 + 2\, y^l - 0.1\, y^l, \\
&\text{s.t. } y^l \leq 0
\end{aligned}
$$

obtaining the upper bound $\bar{x}^u = 1$, $\bar{y}^l = 0$ with UBD $= -1$. Since LBD $=$ UBD the algorithm terminates.

Example 5.2 Consider now the bilevel program

$$
\begin{aligned}
&\min_{x^u, y^u, x^l} \; 10 \left(x^u\right)^2 + x^l \\
&\text{s.t.} \; -0.7 + 0.1 x^u - y^u \leq 0 \\
&\qquad x^l \in \arg\min_{x^m} -\left(x^m\right)^3 \\
&\qquad \text{s.t.} \; -0.5 - \left(x^u\right)^2 - 0.1\, y^u + x^m \leq 0 \\
&\qquad -1 \pm x^u \pm x^m \leq 0 \\
&\qquad x^u \in [-1, 1], \quad y^l \in \{0, 1\}, \quad x^l, x^m \in [-1, 1]
\end{aligned}
$$

with the unique optimal solution $x^u = 0$, $y^u = 0$, $x^l = 0.5$ and the optimal objective value 0.5. The notation $\pm$ is used for compactness of the presentation of four linear constraints on x^u and x^m. This example demonstrates how the algorithm handles integer variables in the upper-level program (in the worse case). The example also demonstrates the generation of parametric bounds when the constraints depend on the upper-level variables. Figure 1 shows the feasible set and the points returned by the iterations of the algorithm.

For simplicity, the lower bounding problem without KKT-based tightening (5) is used here. At the first iteration the lower bounding problem

$$
\begin{aligned}
&\min_{x^u, y^u, x^l} 10\ \left(x^u\right)^2 + x^l \\
&\text{s.t.}\ -0.7 + 0.1x^u - y^u \le 0 \\
&-0.5 - \left(x^u\right)^2 - 0.1\, y^u + x^l \le 0 \\
&\quad -1 \pm x^u \pm x^l \le 0 \\
&\quad x^u \in [-1, 1], \quad y^u \in \{0, 1\}, \quad x^l \in [-1, 1]
\end{aligned}
$$

is solved obtaining $x^u = 0$, $y^u = 1$, $x^l = -1$ and LBD $= -1$. The lower-level program is then solved

$$
\begin{aligned}
&\min_{x^l} - \left(x^l\right)^3 \\
&\text{s.t.}\ -0.5 - (0)^2 - 0.1 \times 1 + x^l \le 0 \\
&\qquad -1 \pm 0 \pm x^l \le 0 \\
&\qquad x^l \in [-1, 1]
\end{aligned}
$$

obtaining the unique solution $x^l = 0.6$, $\bar{f}^{l,*} = -0.216$. Since the lower-level constraints depend on the upper-level variables, pairs $Z^{u,k}$, $\mathbf{z}^{l,k}$ need to be identified. The auxiliary problem (7) is solved for $\varepsilon^l_{f2} = 0.001$ obtaining the point $x^{l,1} \approx 0.5992$. Then Subroutine 1 is applied obtaining $X^{u,1} = [-0.387, 0.387]$, $Y^{u,1} = \{1\}$ (note the degenerate interval for the binary variable). The relatively large value for $\varepsilon^l_{f2} = 0.001$ is chosen so that the interval $X^{u,1}$ is sufficiently large for quick convergence of the lower bound; in the computational results (Sect. 6) a smaller value is used for consistency with the other examples. To conclude the first iteration, the upper bounding problem is solved

$$
\begin{aligned}
&\min_{x^l} x^l \\
&\text{s.t.}\ -0.5 - (0)^2 - 0.1 \times 1 + x^l \le 0 \\
&\qquad -1 \pm 0 \pm x^l \le 0 \\
&\qquad - \left(x^l\right)^3 \le -0.216 \\
&\qquad x^l \in [-1, 1]
\end{aligned}
$$

obtaining an upper bound of UBD $= 0.6$ attained at $x^u = 0$, $y^u = 1$, $x^l = 0.6$.

At the second iteration the lower bounding problem

$$
\begin{aligned}
&\min_{x^u, y^u, x^l} 10\ \left(x^u\right)^2 + x^l \\
&\text{s.t.}\ -0.7 + 0.1x^u - y^u \le 0 \\
&-0.5 - \left(x^u\right)^2 - 0.1\, y^u + x^l \le 0 \\
&-1 \pm x^u \pm x^l \le 0 \\
&-0.387 \le x^u \le 0.387 \text{ and } y^u = 1 \Rightarrow - \left(x^l\right)^3 \le - (0.5992)^3 \\
&x^u \in [-1, 1], \quad y^u \in \{0, 1\}, \quad x^l \in [-1, 1]
\end{aligned}
$$

is solved obtaining $x^u \approx -0.005$, $y^u = 0$, $x^l \approx -0.7005$ and LBD ≈ -0.70025. Similar to the first iteration $x^{l,2} = 0.4989$, $X^{u,2} = [-0.483, 0.4733]$, $Y^{u,1} = \{0\}$ are obtained for the parametric upper bounds to the lower-level objective. The upper bound is updated to

UBD ≈ 0.5003 attained at $x^u \approx -0.005$, $y^u = 0$, $x^l \approx 0.5$ and this is very close to the global minimum.

At the third iteration the lower bounding problem

$$\begin{aligned}
&\min_{x^u, y^u, x^l} 10 \left(x^u\right)^2 + x^l \\
&\text{s.t.} -0.7 + 0.1x^u - y^u \leq 0 \\
&-0.5 - \left(x^u\right)^2 - 0.1\, y^u + x^l \leq 0 \\
&-1 \pm x^u \pm x^l \leq 0 \\
&-0.387 \leq x^u \leq 0.387 \text{ and } y^u = 1 \Rightarrow -\left(x^l\right)^3 \leq -(0.5992)^3 \\
&-0.483 \leq x^u \leq 0.4733 \text{ and } y^u = 0 \Rightarrow -\left(x^l\right)^3 \leq -(0.4989)^3 \\
&x^u \in [-1, 1], \quad y^u \in \{0, 1\}, \quad x^l \in [-1, 1]
\end{aligned}$$

is solved obtaining $x^u \approx 0$, $y^u = 0$, $x^l \approx 0.4989$ and LBD ≈ 0.5. The lower and upper bound have essentially converged.

6 Implementation

The algorithm is implemented in C++ and tested on a 32-bit 1.8 GHz Pentium M processor running Linux 2.6.18. As is typical with optimization codes, both an absolute and relative termination criterion are used and termination occurs if either of the criteria is satisfied. The MINLPs are all solved globally with BARON version 7.8.1 [28] using GAMS version 22.5 [8] through system calls.

The complementarity conditions of the KKT-based lower bounding problem are implemented using the big-M formulation [14,15]. The logical constraints in the lower bounding problem are also implemented using additional integer variables as described in the following.

For boxes $Z^{u,k}$ that satisfy $Z^{u,k} = Z^u$ no logical constraint is needed. Otherwise up to two binary variables and constraints are introduced for each component of $\mathbf{z}^u$, as described in Subroutine 2. Therefore up to $2\left(n_x^u + n_y^u\right) + 1$ binary variables are required to formulate a logical constraint. Note that the bounds $y_j^{u,k,\text{LO}}$, $y_j^{u,k,\text{UP}}$ are integer-valued for all $j = 1, \ldots, n_y^u$ since in the construction of the boxes the floor and ceil functions are used.

Subroutine 2 (Implementation of logical constraints)

- Set $c = 0$
- **FOR** $j = 1, \ldots, n_x^u$ **DO**
  - **IF** $x_j^{u,k,\text{LO}} > x_j^{u,\text{LO}}$ **THEN**
    - Set $c = c + 1$ and introduce a binary variable $w_c \in \{0, 1\}$ corresponding to $x_j > x_j^{u,k,\text{LO}}$.
    - Introduce a constraint

$$w_c \geq \frac{x_j^u - x_j^{u,k,\text{LO}}}{x_j^{u,\text{UP}} - x_j^{u,\text{LO}}}.$$

  - **IF** $x_j^{u,k,\text{UP}} < x_j^{u,\text{UP}}$ **THEN**

Set $c = c + 1$ and introduce a binary variable $w_c \in \{0, 1\}$ corresponding to $x_j < x_j^{u,k,\mathrm{UP}}$.
Introduce a constraint

$$w_c \geq \frac{x_j^{u,k,\mathrm{UP}} - x_j^u}{x_j^{u,\mathrm{UP}} - x_j^{u,\mathrm{LO}}}.$$

- **FOR** $j = 1, \ldots, n_y^u$ **DO**
  - **IF** $y_j^{u,k,\mathrm{LO}} > y_j^{u,\mathrm{LO}}$ **THEN**
    * Set $c = c + 1$ and introduce a binary variable $w_c \in \{0, 1\}$ corresponding to $y_j \geq y_j^{u,k,\mathrm{LO}}$.
    * Introduce a constraint

$$w_c \geq \frac{y_j^u - y_j^{u,k,\mathrm{LO}} + 0.5}{y_j^{u,\mathrm{UP}} - y_j^{u,\mathrm{LO}}}.$$

  - **IF** $y_j^{u,k,\mathrm{UP}} < y_j^{u,\mathrm{UP}}$ **THEN**
    * Set $c = c + 1$ and introduce a binary variable $w_c \in \{0, 1\}$ corresponding to $y_j \leq y_j^{u,k,\mathrm{UP}}$.
    * Introduce a constraint

$$w_c \geq \frac{y_j^{u,k,\mathrm{UP}} - y_j^u + 0.5}{y_j^{u,\mathrm{UP}} - y_j^{u,\mathrm{LO}}}.$$

  END FOR
- Introduce the logical constraint as

$$f^l\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l\right) \leq f^l\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right) + \sum_{i=1}^{c}(1 - w_i)\left(f^{l,\max} - f^l\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right)\right),$$

  where $f^{l,\max} \geq f^l\left(\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^{l,k}, \mathbf{y}^{l,k}\right)$ for all $(\mathbf{x}^u, \mathbf{y}^u) \in Z^u$. This value can be easily estimated using interval analysis, see, e.g., [24]. Note that unless $w_i = 1$, for all $i = 1, \ldots, c$ this constraint is redundant.

For the continuous variables the constraint is only introduced in the interior of $Z^{u,k}$ while for the integer variables also for the boundary (by the addition of 0.5 in the numerator). Note also that for $j > n_x^u$ (i.e., for the integer variables) $z_j^{u,k,\mathrm{UP}} = z_j^{u,k,\mathrm{LO}}$ is possible, but still both additional binary variables are needed.

6.1 Test set

To test the algorithm literature examples are used along with a few new test problems. The problem formulations are given in Appendix A. Table 1 contains a summary of the problem properties. The first column is the label of the example. The second through fifth columns $\left(n_x^u, n_y^u, n_x^l, n_y^l\right)$ contain the number of variables. The sixth through ninth columns $\left(f^u, g^u, f^l, g^l\right)$ contain the functional form of the upper-level objective, the upper-level constraints, the lower-level objective and the lower-level constraints for fixed values of the integer

Table 1 Summary of problem properties

Label	n_x^u	n_y^u	n_x^l	n_y^l	f^u	g^u	f^l	g^l
am_1_0_0_1_01	1	0	0	1	C	N/A	N/A	N/A
am_1_1_1_0_01	1	1	1	0	C	L	N	L
am_1_1_1_1_01	1	1	1	1	C	L	N	L
am_1_1_1_1_02	1	1	1	1	N	L	N	N
am_3_3_3_3_01	3	3	3	3	N	N	N	L
Edmunds_Bard	1	0	0	1	C	N/A	N/A	N/A
Jan_Chern	0	2	0	3	N/A	N/A	N/A	N/A
Moore_Bard	0	1	0	1	N/A	N/A	N/A	N/A
Sahin_Ciric	2	0	0	2	N	N/A	N/A	N/A
Thirwani_Arora	0	1	0	1	N/A	N/A	N/A	N/A

Table 2 Numerical results without KKT based tightening

Label	$\bar{f}^u$	$\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \bar{\mathbf{x}}^l, \bar{\mathbf{y}}^l$	UBD	#UBD	#LBD	Time
am_1_0_0_1_01	−1	−1, 0	2	2	2	0.03 0.67
am_1_1_1_0_01	0.500029	−0.00195395, 0, 0.49999	7	7	8	0.30 3.1
am_1_1_1_1_01	−1.00365	0.00364962, 1, −0.0003648, 0	8	8	8	0.20 2.7
am_1_1_1_1_02	0.206899	−0.568658, 0, 0.45378, 0	6	6	6	0.83 3.5
am_3_3_3_3_01	−2.50001	3.1e−06, −1, 0, 1, 0, 1, 1, 1, −0.299997, 1, 0, 0	2	2	2	0.27 1.2
Edmunds_Bard	0.444444	1.33333, 2	1	1	1	0.02 0.44
Jan_Chern	−51	1, 3, 2, 2, 2	2	6	7	0.68 3.35
Moore_Bard	5	3, 1	2	2	2	0.01 0.75
Sahin_Ciric	−400	0, 10, 1, 1	3	3	3	0.12 1.01
Thirwani_Arora	−0.666667	0, 3	2	2	2	0.03 0.60

variables: N/A stands for not applicable (no continuous variables, or no constraints), A stands for affine linear, C stands for convex nonlinear, N stands for nonconvex nonlinear; for the upper-level functions the characterization is joint in $\mathbf{x}^u$ and $\mathbf{x}^l$ while for the lower-level functions the characterization is only for the $\mathbf{x}^l$—dependence, e.g., convex means partially convex on X^l. Note that the integrality constraints make all lower-level and upper-level problems nonconvex.

The optimality and feasibility tolerances for BARON are set to $\varepsilon_{\text{MINLP}} = 10^{-6}$ for all problems. The optimality tolerance for the lower-level problem is set to $\varepsilon_f^l = 10^{-5}$ and the absolute and relative termination criteria to $\varepsilon_f^u = 10^{-4}$. Note that for all problems the tolerances used satisfy the assumptions in Theorems 1. For all problems $\varepsilon_{f2}^l = 0.8\varepsilon_f^l$ is used for all iterations. To guess the boxes $Z^{u,k}$, a decreasing sequence is used; each time the interval diameter is set to one and decreased by a factor of 0.9 until the interval extensions show feasibility. Natural interval extensions are used for the overestimation.

Tables 2 and 3 contain the numerical results with and without KKT based tightening (whenever applicable). The first column (Label) has the label of the problem. The second through ninth columns contain the results obtained: $\bar{f}^u$ shows the optimal objective value

Table 3 Numerical results with KKT based tightening

Label	$\bar{f}^u$	$\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \bar{\mathbf{x}}^l, \bar{\mathbf{y}}^l$	UBD	#UBD	#LBD	Time		$\mu^{\max}$
am_1_1_1_0_01	0.499989	−0.000493355,0,0.499987	2	2	3	0.04	0.93	2
am_1_1_1_1_01	−1.00365	0.00364733,1,−0.000182366,0	8	8	8	0.46	2.9	10
am_1_1_1_1_02	0.209503	−0.554475,0,0.455445,0	2	2	2	0.30	1.3	2
am_3_3_3_3_01	−2.5	0,−1,0,1,0,1,1,1,−0.300002,1,0,0	1	1	1	0.40	0.80	100

obtained; $\bar{\mathbf{x}}^u, \bar{\mathbf{y}}^u, \bar{\mathbf{x}}^l, \bar{\mathbf{y}}^l$ shows the optimal solution obtained; UBD shows the iteration at which the optimal solution is first obtained; #UBD shows the number of upper bounding calls; #LBD shows the number of lower bounding calls; the first time column shows the sum of CPU time reported by GAMS and spent in the main program, while the second time column shows the time obtained by the timing function. Note that there is a significant difference between these two times requirements, presumably due to the system calls and processing time for GAMS. Because CPU times are small, an average of 10 runs is presented. Additionally, the last column in Table 3 contains the bound for the KKT multipliers used.

All the test problems are solved without significant numerical difficulties. Similarly to single-level optimization appropriate choice of tolerances is necessary for computational efficiency and accuracy of solutions. The KKT-based heuristic for the lower bound, when applicable, can reduce the number of iterations required. However, the cost per iteration is higher because of the complementarity conditions. For one of the literature test problems (Jan_Chern) the point reported in the literature is found infeasible (lower-level suboptimal) and for another suboptimal (Sahin_Ciric).

7 Conclusions

A deterministic algorithm for the solution of mixed-integer bilevel programs was presented along with a convergence proof and illustrative examples. An implementation was performed and the algorithm tested with numerical examples. The algorithm is guaranteed to finitely find a point which satisfies $\varepsilon-$optimality in both the upper-level and lower-level programs. The algorithm uses a lower bound which converges based on parametric upper bounds to the lower-level problem and an upper bound based on probing the solutions furnished by the lower bounding problem.

An interesting extension of the algorithm is to a branch-and-bound framework. This would result to a somewhat cumbersome notation, but there is no conceptual problem expected in performing the extension. Relative to the continuous case, there are additional branching alternatives, such as branching only on the integer variables. Also, it is worthwhile exploring alternatives to the MINLP formulation of the KKT-based lower bounds as in the MPEC formulation [31,30]. In NLPs and MINLPs domain reduction methods have been very successful [32] and it would be interesting to also consider domain reduction for bilevel programs.

Acknowledgments I would like to thank the anonymous reviewers for the detailed comments that considerably improved this article. Financial support from the Deutsche Forschungsgemeinschaft (German Research Association) through grant GSC 111 is gratefully acknowledged.

A Test problems

The bilevel program

$$\begin{aligned}
&\min_{x^u, y^l} x^u - 2\,x^u y^l - 0.1\,y^l \\
&\text{s.t. } y^l \in \arg\min_{y^m} y^m \\
&x^u \in [-1, 1], \quad y^l, y^m \in \{0, 1\},
\end{aligned} \tag{am_1_0_0_1_01}$$

has the unique solution $x^u = -1$, $y^l = 0$ with an objective value of -1.

The bilevel program

$$\begin{aligned}
&\min_{x^u, y^u, y^l} 10\,\left(x^u\right)^2 + x^l \\
&\text{s.t. } -0.7 + 0.1\,x^u - y^u - x^l \le 0 \\
&x^l \in \arg\min_{x^m} -\left(x^m\right)^3 \\
&\text{s.t. } -0.5 - \left(x^u\right)^2 - 0.1\,y^u + x^m \le 0 \\
&\quad -1 - x^u + x^m \le 0 \\
&\quad -1 + x^u + x^m \le 0 \\
&\quad -1 - x^u - x^m \le 0 \\
&\quad -1 + x^u - x^m \le 0 \\
&\quad x^u \in [-1, 1], \quad y^u \in \{0, 1\}, \qquad x^l, x^m \in [-1, 1]
\end{aligned} \tag{am_1_1_1_0_01}$$

has the unique optimal solution $x^u = 0$, $y^u = 0$, $x^l = 0.5$ and the optimal objective value 0.5.

The bilevel program

$$\begin{aligned}
&\min_{x^u, y^u, x^l, y^l} -y^u - x^u - y^l + x^u\,x^l + 10\,\left(x^l\right)^2 \\
&\text{s.t. } -0.5 y^u + x^u \le 0 \\
&\quad \left(x^l, y^l\right) \in \arg\min_{x^m, y^m} y^m - x^u \left(x^m\right)^2 + 0.5 \left(x^m\right)^4 \\
&\quad \text{s.t. } -0.2\,y^m + x^m \le 0 \\
&\quad x^u \in [-1, 1], \quad y^u \in \{0, 1\}, \qquad x^l, x^m \in [-1, 1], y^l, y^m \in \{0, 1\}
\end{aligned} \tag{am_1_1_1_1_01}$$

has the unique optimal solution point is $x^u = 0$, $y^u = 1$, $x^l = 0$, $y^l = 0$ and an objective function of -1.

The bilevel program

$$\min_{x^u, y^u, x^l, y^l} y^u + \left(x^u + 0.6\right)^2 - y^l + \left(x^l\right)^2$$
$$\text{s.t.} \; - 0.3y^u + x^u + y^l \leq 0$$
$$\left(x^l, y^l\right) \in \arg\min_{x^m, y^m} y^m + \left(x^m\right)^4 + 0.4/3\,\left(-x^u + 1\right)\left(x^m\right)^3$$
$$+ \left(-0.02\left(x^u\right)^2 + 0.16x^u - 0.4\right)\left(x^m\right)^2 + \left(0.004\,\left(x^u\right)^3 - 0.036\left(x^u\right)^2 + 0.08\,\left(x^u\right)\right)x^m$$
$$\text{s.t. } 0.01\,\left(1 + \left(x^u\right)^2\right) - \left(x^m\right)^2 \leq 0 \qquad \text{(am\_1\_1\_1\_1\_02)}$$
$$x^u \in [-1, 1], \quad y^u \in \{0, 1\}, \qquad x^l, x^m \in [-1, 1], \quad y^l, y^m \in \{0, 1\}$$

has the best known solution $x^u = -0.5545$, $y^u = 0$, $x^l = 0.4554$, $y^l = 0$ with an objective value of 0.209.

The bilevel program

$$\min_{\mathbf{x}^u, \mathbf{y}^u, \mathbf{x}^l, \mathbf{y}^l} y_1^u + y_2^u - y_3^u + 2\left(x_1^u - 0.5\right)^2\, y_1^l - 2y_1^l + x_2^u\, x_1^l\, x_2^l + \left(x_3^u\right)^2 + \left(x_3^l + 0.3\right)^2$$
$$\text{s.t.} \; - 0.7 + 0.1\, x_1^u \leq 0$$
$$0.1x_2^u - y_1^u + y_1^l \leq 0$$
$$0.1x_3^u + y_2^u - y_3^l \leq 0$$
$$x_2^u x_3^u + x_1^l - x_2^l \leq 0$$
$$\left(\mathbf{x}^l, \mathbf{y}^l\right) \in \arg\min_{\mathbf{x}^m, \mathbf{y}^m} -y_1^m + 2y_2^m + 3\left(y^l\right)^3 - \left(x_1^m\right)^3\, x_2^m y_1^u - 2x_3^m\, x_1^u \quad \text{(am\_3\_3\_3\_3\_01)}$$
$$\text{s.t. } x_1^u - y_1^m \leq 0$$
$$x_2^u - x_1^m \leq 0$$
$$x_3^u - x_2^m \leq 0$$
$$\mathbf{x}^u \in [-1, 1]^3, \quad y^u \in \{0, 1\}^3, \qquad \mathbf{x}^l, \mathbf{x}^m \in [-1, 1]^3, \quad \mathbf{y}^l, \mathbf{y}^m \in \{0, 1\}^3$$

has the best known solution $\mathbf{x}^u = (0, -1, 0)$, $\mathbf{y}^u = (1, 0, 1)$, $\mathbf{x}^l = (1, 1, -0.3)$, $\mathbf{y}^l = (1, 0, 0)$ with an objective value of -2.5.

The bilevel program by Edmunds and Bard [11] (Eq. 3 therein)

$$\min_{x^u, y^l} \left(x^u - 2\right)^2 + \left(y^l - 2\right)$$
$$\text{s.t. } y^l \in \arg\min_{y^m} \left(y^m\right)^2$$
$$\text{s.t.} \; - 2x^u - 2y^m + 5 \leq 0 \qquad \text{(Edmunds\_Bard)}$$
$$x^u - y^m - 1 \leq 0$$
$$3\,x^u + 2\,y^m - 8 \leq 0$$
$$x^u \in [0, 3], \quad y^l, y^m \in \{0, 1, 2\},$$

has the unique solution $x^u = 4/3$, $y^l = 2$ with an objective value of 4/9.

The bilevel program by Jan and Chern [19]

$$\min_{\mathbf{y}^u, \mathbf{y}^l} -3\,y_1^u - 6\,y_2^u - 2\,y_1^l - 4\,y_2^l - 9\,y_3^l$$

$$\text{s.t. } \mathbf{y}^l \in \arg\min_{\mathbf{y}^m} -\left(1-(1-0.9)^{y_1^u}\right)\left(1-(1-0.875)^{y_2^u}\right)\left(1-(1-0.79)^{y_1^m}\right)$$
$$\left(1-(1-0.8)^{y_2^m}\right)\left(1-(1-0.93)^{y_3^m}\right)$$
$$\text{s.t. } 5\,y_1^u + 4\,y_2^u + 9\,y_1^m + 7\,y_2^m + 7\,y_3^m - 64 \le 0 \tag{Jan\_Chern}$$
$$6\,y_1^u + 7\,y_2^u + 8\,y_1^m + 9\,y_2^m + 6\,y_3^m - 76 \le 0$$
$$\mathbf{y}^u \in \{0,1,2,3\}^2, \qquad \mathbf{y}^l, \mathbf{y}^m \in \{0,1,2,3,4,5\}^3$$

has the unique solution $y_1^u = 1$, $y_2^u = 3$, $y_1^l = 2$, $y_2^l = 2$, $y_l^3 = 2$ with an objective value of -51. The optimality of this point can be verified by explicit enumeration. Jan and Chern [19] report $\mathbf{y}^u = (2,3)$, $\mathbf{y}^l = (1,2,2)$ as the optimal solution with an objective value of -52. However, this point is lower-level suboptimal and therefore not feasible: for $\mathbf{y}^u = (2,3)$ the point $\mathbf{y}^l = (2,2,1)$ gives $f^l = -0.84$ whereas $\mathbf{y}^l = (1,2,2)$ gives -0.75.

The bilevel program by Moore and Bard [25] (Example 2 therein)

$$\min_{y^u, y^l} y^u + 2\,y^l$$
$$\text{s.t. } y^l \in \arg\min_{y^m} -y^m$$
$$\text{s.t. } -y^u + 2.5\,y^m - 3.75 \le 0 \tag{Moore\_Bard}$$
$$-y^u - 2.5\,y^m + 3.75 \le 0$$
$$2.5y^u + y^m - 8.75 \le 0$$
$$y^u \in \{0,1,2,3\}, \quad y^l, y^m \in \{1,2\}$$

has the unique optimal solution $y^u = 3$, $y^l = 1$ with an objective value of -5.

The bilevel program by Sahin and Ciric [27] (Example 4 therein)

$$\min_{\mathbf{x^u}, \mathbf{y}^l} -\left(\left(-0.4\,\left(x_1^u\right)^2\,x_2^u + 4\,\left(x_2^u\right)^2\right)\,y_1^l\,y_2^l + \left(-\left(x_2^u\right)^3 + 3\,\left(x_1^u\right)^2\,x_2^u\right)\left(1-y_1^l\right)\,y_2^l\right.$$
$$\left. + \left(2\,\left(x_2^u\right)^2 - x_1^u\right)\left(1-y_2^l\right)\right)$$
$$\text{s.t. } \mathbf{y}^l \in \arg\min_{\mathbf{y}^m} -\left(\left(\left(x_1^u\right)^2\,\left(x_2^u\right)^2 + 8\,\left(x_2^u\right)^3 - 14\,\left(x_1^u\right)^2 - 5\,x_1^u\right)\,y_1^m\,y_2^m\right.$$
$$\left. + \left(-x_1^u\,\left(x_2^u\right)^2 + 5\,x_1^u\,x_2^u + 4\,x_2^u\right)\left(1-y_1^m\right)\,y_2^m + 8\,x_1^u\,y_1^m\,\left(1-y_2^m\right)\right) \tag{Sahin\_Ciric}$$
$$\text{s.t. } 1 - y_1^m - y_2^m \le 0$$
$$\mathbf{x}^u \in [0,10]^2, \quad \mathbf{y}^l, \mathbf{y}^m \in \{0,1\}^2$$

has the unique optimal solution $x_1^u = 0$, $x_2^u = 10$, $y_1^l = 1$, $y_2^l = 1$ with an objective value of -400. There are two simple ways to verify this solution. The first possibility is to reformulate (Sahin_Ciric) as a single-level problem. This is possible because the lower-level program only has three feasible points, which are independent of the upper-level variables. Alternatively, the lower-level program can be solved parametrically as a function of the upper-level variables and the optimal solution plugged into the upper-level variable. The optimality regions of each $\mathbf{y}^l$ are: (i) the point $\mathbf{y}^l = (1,0)$ is optimal for large x_1^u and small x_2^u; (ii) the point $\mathbf{y}^l = (0,1)$ is optimal in a narrow region of small x_1^u and intermediate x_2^u, and (iii) the point $\mathbf{y}^l = (1,1)$ is optimal for large x_1^u and large x_2^u. In Fig. 2 the optimality regions are shown, along with the upper-level objective function, which results by fixing the lower-level variables to the lower-level optimal solution point. Note the pronounced local minimum at $x_1^u \approx 6.03836$,

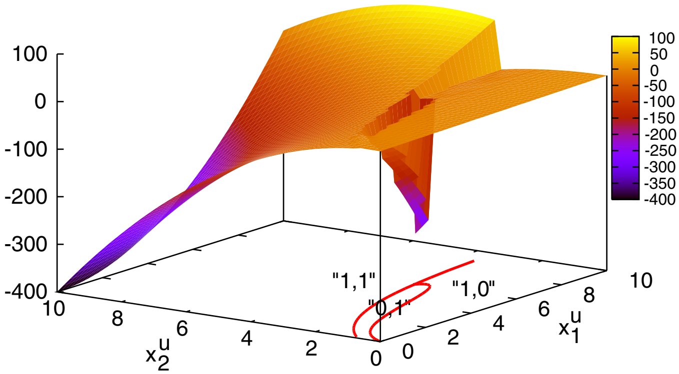

Fig. 2 Graphical illustration of Example Sahin_Ciric. In the $x_1^u - x_2^u$ plane the optimality regions for the lower-level program are illustrated. "0,1" indicates the region where $\mathbf{y}^l = (0, 1)$ is optimal, "1,0" indicates the region where $\mathbf{y}^l = (1, 0)$ is optimal, and "1,1" indicates the region where $\mathbf{y}^l = (1, 1)$ is optimal. The surface plot shows the upper-level objective function with the lower-level variables fixed to the solution to the lower-level problem

$x_2^u \approx 2.95684$, $y_1^l = 0$, $y_2^l = 1$ with the upper-level objective value -297.58451. Previous articles [18,27] wrongfully report this as the global minimum.

The bilevel program by Thirwani and Arora [33]

$$\begin{aligned}
&\min_{y^u, y^l} -\frac{2\, y_1^l + 3\, y_1^u}{y_1^l + 4\, y_1^u + 6}\\
&\text{s.t. } y_1^u + y_1^l - 5 \leq 0\\
&y_1^u + 3\, y_1^l - 10 \leq 0\\
&y^l \in \arg\min_{y^m} -\frac{3\, y_1^m + 4\, y_1^u}{6\, y_1^m + 4\, y_1^u + 3}\\
&y^u \in \{0, 1, 2, 3, 4, 5\}, \quad y^l, y^m \in \{0, 1, 2, 3\}
\end{aligned} \tag{Thirwani_Arora}$$

has the unique optimal solution $y^u = 0$, $y^l = 3$ with an objective value of $-2/3$.

References

1. Alefeld, G., Mayer, G.: Interval analysis: theory and applications. J. Comput. Appl. Math. **121**(1-2), 421–464 (2000)
2. Bank, B., Guddat, J., Klatte, D., Kummer, B., Tammer, K.: Non-Linear Parametric Optimization. Birkhäuser Verlag, Stuttgart (1983)
3. Bard, J.F.: An algorithm for solving the general bilevel programming problem. Math. Oper. Res. **8**(2), 260–272 (1983)
4. Bard, J.F.: Practical Bilevel Optimization: Algorithms and Applications. Nonconvex Optimization and Its Applications. Kluwer, Dordrecht (1998)
5. Bhattacharjee, B., Lemonidis, P., Green, W.H. Jr., Barton, P.I.: Global solution of semi-infinite programs. Math. Program. Ser. B **103**(2), 283–307 (2005)
6. Blankenship, J.W., Falk, J.E.: Infinitely constrained optimization problems. J. Optim. Theory Appl. **19**(2), 261–281 (1976)
7. Bracken, J., McGill, J.T.: Mathematical programs with optimization problems in constraints. Oper. Res. **21**(1), 37–44 (1973)
8. Brooke, A., Kendrick, D., Meeraus, A.: GAMS: A User's Guide. Scientific, Redwood City (1988)

9. Dempe, S.: Foundations of Bilevel Programming. Nonconvex Optimization and its Applications. Kluwer, Dordrecht (2002)
10. Dempe, S.: Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. Optimization **52**(3), 333–359 (2003)
11. Edmunds, T.A., Bard, J.F.: An algorithm for the mixed-integer nonlinear bilevel programming problem. Ann. Oper. Res. **34**(1), 149–162 (1992)
12. Falk, J.E., Hoffman, K.: A nonconvex max- min problem. Naval Res. Logist. **24**(3), 441–450 (1977)
13. Floudas, C.A., Stein, O.: The adaptive convexification algorithm: a feasible point method for semi-infinite programming. SIAM J. Optim. **18**(4), 1187–1208 (2007)
14. Fortuny-Amat, J., McCarl, B.: A representation and economic interpretation of a two-level programming problem. J. Oper. Res. Soc. **32**(9), 783–792 (1981)
15. Glover, F.: Improved linear integer programming formulations of nonlinear integer programs. Manag. Sci. **22**(4), 455–460 (1975)
16. Glover, F., Woolsey, E.: Further reduction of zero-one polynomial programming problems to zero-one linear programming problems. Oper. Res. **21**(1), 156–161 (1973)
17. Grossmann, I.E., Halemane, K.P.: Decomposition strategy for designing flexible chemical-plants. AIChE J. **28**(4), 686–694 (1982)
18. Gümüs, Z.H., Floudas, C.A.: Global optimization of mixed-integer bilevel programming problems. Comput. Manag. Sci. **2**(3), 181–212 (2005)
19. Jan, R.H., Chern, M.S.: Nonlinear integer bilevel programming. Eur. J. Oper. Res. **72**(3), 574–587 (1994)
20. Kesavan, P., Allgor, R.J., Gatzke, E.P., Barton, P.I.: Outer approximation algorithms for separable nonconvex mixed-integer nonlinear programs. Math. Program. **100**(3), 517–535 (2004)
21. Lemonidis, P., Barton, P.I.: Interval methods for generalized semi-infinite programs. International conference on parametric optimization and related topics (PARAOPT VIII), Cairo, Egypt, 27 Nov 2005–1 Dec 2005
22. Mitsos, A., Lemonidis, P., Barton, P.I.: Global solution of bilevel programs with a nonconvex inner program. J. Glob. Optim. **42**(4), 475–513 (2008)
23. Mitsos, A., Lemonidis, P., Lee, C.K., Barton, P.I.: Relaxation-based bounds for semi-infinite programs. SIAM J. Optim. **19**(1), 77–113 (2008)
24. Moore, R.: Methods and Applications of Interval Analysis. SIAM, Philadelphia (1979)
25. Moore, J.T., Bard, J.F.: The mixed integer linear bilevel programming problem. Oper. Res. **38**(5), 911–921 (1990)
26. Oluwole, O.O., Bhattacharjee, B., Tolsma, J.E., Barton, P.I., Green, W.H. Jr.: Rigorous valid ranges for optimally reduced kinetic models. Combustion Flame **146**(1–2), 348–365 (2006)
27. Sahin, K.H., Ciric, A.R.: A dual temperature simulated annealing approach for solving bilevel programming problems. Comput. Chem. Eng. **23**(1), 11–25 (1998)
28. Sahinidis, N., Tawarmalani M.: BARON. http://www.gams.com/solvers/baron.pdf (2005)
29. Shimizu, K., Ishizuka, Y., Bard, F.J.: Nondifferentiable and Two-Level Mathematical Programming. Kluwer, Boston (1997)
30. Stein, O., Still, G.: Solving semi-infinite optimization problems with interior point techniques. SIAM J. Control Optim. **42**(3), 769–788 (2003)
31. Stein, O., Oldenburg, J., Marquardt, W.: Continuous reformulations of discrete-continuous optimization problems. Comput. Chem. Eng. **28**(10), 1951–1966 (2004)
32. Tawarmalani, M., Sahinidis, N.V.: Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming. Nonconvex Optimization and its Applications. Kluwer, Boston (2002)
33. Thirwani, D., Arora, S.R.: An algorithm for the integer linear fractional bilevel programming problem. Optimization **39**, 53–67 (1997)
34. Vicente, L.N., Calamai, P.H.: Bilevel and multilevel programming—a bibliography review. J. Glob. Optim. **5**(3), 291–306 (1994)
35. Ye, J.J.: Constraint qualifications and KKT conditions for bilevel programming problems. Math. Oper. Res. **31**(4), 811–824 (2006)
36. Zuhe, S., Neumaier, A., Eiermann, M.C.: Solving minimax problems by interval-methods. BIT, **30**(4), 742–751 (1990)