On generalized semi-infinite optimization and bilevel optimization

Oliver Stein* and Georg Still†

Abstract

The paper studies the connections and differences between bilevel problems (BL) and generalized semi-infinite problems (GSIP). Under natural assumptions (GSIP) can be seen as a special case of a (BL). We consider the so-called reduction approach for (BL) and (GSIP) leading to optimality conditions and Newton-type methods for solving the problems. We show by a structural analysis that for (GSIP)-problems the regularity assumptions for the reduction approach can be expected to hold generically at a solution but for general (BL)-problems not. The genericity behavior of (BL) and (GSIP) is in particular studied for linear problems.

Keywords: Semi-infinite programming, bilevel programming, optimality conditions, genericity behavior, numerical methods.

Mathematical Subject Classification 1991: 90C34, 90C30, 90C31

1 Introduction

We consider generalized semi-infinite problems,

$$\text{(GSIP)}: \quad \min_x F_{\text{GSIP}}(x) \qquad \text{s.t. } x \in M_{\text{GSIP}} = \{x \in \mathbb{R}^n \mid G(x,y) \geq 0, \text{ for all } y \in Y(x)\}$$
$$\text{where} \quad Y(x) = \{y \in \mathbb{R}^m \mid g(x,y) \geq 0\}.$$

For the special case, that the set $Y(x) = Y$ does not depend on the variable x, this problem is a common semi-infinite problem (SIP). Bilevel problems are of the form

$$\text{(BL)}: \quad \min_{x,y} \; F(x,y) \qquad \text{subject to } G(x,y) \geq 0 \text{ and } y \text{ is a solution of} \tag{1}$$
$$\text{Q(x)}: \quad \min_y \; f(x,y) \quad \text{s.t.} \quad y \in Y(x),$$

with $Y(x)$ defined as in (GSIP).

Throughout the paper we assume $F_{\text{GSIP}} \in C(\mathbb{R}^n, \mathbb{R})$, $F \in C(\mathbb{R}^n \times \mathbb{R}^m, \mathbb{R})$, $G \in C(\mathbb{R}^n \times \mathbb{R}^m, \mathbb{R}^p)$, $g \in C(\mathbb{R}^n \times \mathbb{R}^m, \mathbb{R}^q)$. We use the abbreviation $J = \{1, \ldots, p\}$ and $I = \{1, \ldots, q\}$ for the index sets of the constraints G and g.

*Department of Mathematics, University of Aachen, Germany

†Faculty of Mathematical Sciences, University of Twente, The Netherlands

There is an extensive literature on bilevel optimization (see [16] and the references in this book). Semi-infinite programming (SIP) is an important field of research as well (cf. for example the survey article [7] with more than 300 references). Generalized semi-infinite problems are studied only recently (cf. e.g. [8, 10, 12, 14, 15, 17, 18, 19, 20, 21, 23]).

Bilevel problems often arise as Operations Research problems in an economic context. They can be interpreted as a game between two players. Player 1 (upper level player) tries to minimize his object F depending on (x, y) and player 2 (lower level player) who for given x chooses y as a solution of the lower level problem $Q(x)$. Applications of (SIP) and (GSIP) mostly appear in technical sciences. For applications of (SIP) we refer to [7]. Applications of (GSIP) are e.g. the *maneuverability* problem in robotics, the *reverse Chebyshev approximation* (see e.g. [20]) and time minimal control problems (see [12]).

In this paper we will show that there is a strong connection between bilevel and generalized semi-infinite problems. Under certain assumptions (GSIP) can be seen as a special instance of a (BL). We will discuss the connections but also the differences between (GSIP) and (BL). The paper is organized as follows.

In Section 2 we study the structure of the feasible sets of (BL) and (GSIP) and consider a natural condition under which (GSIP) becomes a special case of (BL). In Section 3 we apply the so-called *local reduction approach.* This technique leads to optimality conditions for (BL)- and (GSIP)-problems and gives the basis for (Quasi-) Newton methods for solving the problems. We then ask whether the regularity assumptions used in the reduction approach are natural, i.e. assumptions which are generic. We analyze the difference between the structure of typical classes of (BL) problems and the class of (GSIP). It appears that for classes of bilevel problems the regularity assumptions are not generic. This shows that the regularity assumptions used in bilevel programming are often not valid at the solution of (BL). The structural discussion however leads to the conjecture, that for (GSIP) the regularity assumptions for the 'reduction' can be expected to hold generically.

In Section 4 we give a detailed analysis of linear problems and prove the genericity conjecture for linear generalized semi-infinite problems.

Section 5 briefly describes the Kuhn-Tucker approach from bilevel programming for solving linear (GSIP) problems.

2 Relations between GSIP- and BL-problems

In this section we compare the structure of (GSIP) and (BL). We introduce some notation with $x \in I\!R^n$, $y \in I\!R^m$:

$S(x)$	$= \{y \mid \text{y is a (global) solution of } Q(x)\}$	set of solutions of$Q(x)$
S	$= \{(x,y) \mid y \in S(x)\}$	solution graph of Q
M_G	$= \{(x,y) \mid G(x,y) \geq 0\}$	constrained set of upper level
X	$= \{x \mid Y(x) \neq \emptyset\} \quad (\equiv dom\ Y)$	domain of the mapping Y
M_{BL}	$= \{(x,y) \mid (x,y) \in M_G,\ y \in S(x)\}$	feasible set of (BL)

We assume that the set-valued mapping $Y : I\!R^n \to 2^{I\!R^m}$ is uniformly compact on $I\!R^n$, i.e. for any $\overline{x} \in I\!R^n$ there exists a ball $B_\rho(\overline{x}) = \{x \in I\!R^n \mid ||x - \overline{x}|| \leq \rho\},\ \rho >$

0, such that clos $(\cup_{x\in B_\rho(\overline{x})}Y(x))$ is compact. Then, under our assumptions, the mapping Y is closed and upper semi-continuous in the sense of Berge and the set X $(=\text{dom}\, Y)$ is a closed set. Moreover under this condition, the lower level problems $Q(x),\ x\in X$, always have solutions.

Now we try to transform GSIP into a problem of bilevel type. Let us define

$$f(x,y) := \min_{1\le j\le p} G_j(x,y)\ , \tag{2}$$

where the minimum is taken over the components G_j of the vector-valued function G. Consider the parametric problem

$$Q(x):\quad \min_y\ f(x,y) \qquad \text{s.t.}\quad y\in Y(x).$$

Then, for $x\in I\!R^n$ such that $Y(x)\neq\emptyset$ we have $f(x,y)\ge 0,\ \forall y\in Y(x)$, if and only if a solution y of $Q(x)$ (and thus all solutions) satisfies $f(x,y)\ge 0$. Observe that $f(x,y)\ge 0$ is equivalent with the inequality $G(x,y)\ge 0$. So, by the relation between G and f, here for GSIP in contrast to the general BL-case, the feasibility condition $G(x,y)\ge 0$ for the upper level holds for all solutions y of $Q(x)$ if it holds for one solution. Thus, for x with $Y(x)\neq\emptyset$ the condition $x\in M_{\text{GSIP}}$ is equivalent with $x\in \text{pr}_x(\ M_{\text{BL}_{GSIP}})$ where

$$M_{\text{BL}_{GSIP}} = \{(x,y)\mid G(x,y)\ge 0 \text{ and } y \text{ is a solution of } Q(x)\ \}\ .$$

Here, pr_x denotes the orthogonal projection onto the space $I\!R^n$ (x-variable). Summarizing, the bilevel formulation of the generalized semi-infinite problem is given by

$$\begin{array}{lll} \text{BL}_{\text{GSIP}} & \min\ F_{\text{GSIP}}(x)\quad \text{s.t. } G(x,y)\ge 0 \quad \text{and } y \text{ is a solution of} \\ & \qquad Q(x):\ \min\limits_y\ f(x,y)\ \text{ s.t. }\quad y\in Y(x)\ , \end{array} \tag{3}$$

with the function f in (2). We have shown that if $Y(x)\neq\emptyset$ holds for all $x\in I\!R^n$, then (GSIP) is equivalent with BL_{GSIP}, i.e. the problem (GSIP) can be seen as a special instance of a (BL).

Note however that for $S(x)=\emptyset$, x belongs to M_{GSIP}, (no constraints for x) but not to $pr_x(M_{BL_{GSIP}})$. With the set

$$\hat{M}_{\text{GSIP}} = \{x\in \text{dom}\, Y\mid G(x,y)\ge 0\ \text{ for (all)}\quad y\in S(x)\} \tag{4}$$

we actually have $M_{GSIP} = \hat{M}_{\text{GSIP}}\cup(\text{dom}\ Y)^c$ where A^c we denotes the complement of the set A in the corresponding space. Thus, we have shown part (b) of the following lemma which provides different representations for the feasible sets of (GSIP) and (BL). Part (c) has been shown in [15], whereas part (a) follows directly from the definition.

Lemma 1 *The following holds.*

(a) $M_{BL} = M_G\cap S$.

(b) $M_{GSIP} = \hat{M}_{GSIP}\cup(\text{dom}\ Y)^c = \text{pr}_x(\ M_{BL_{GSIP}})\cup(\text{dom}\ Y)^c$.

(c) $M_{GSIP} = \left(pr_x\ (M_G^c \cap S)\right)^c$

In view of Lemma 1a, since M_G is closed, the set M_{BL} is closed if S is closed. The set S is closed if the mapping Y is (lower semi-) continuous on X ($=$ dom Y). In view of Lemma 1b, since $(\text{dom}\, Y)^c$ is open, the set M_{GSIP} need not be closed, even when $\text{pr}_x\ M_{\text{BL}_{GSIP}}$ is closed. For further details on the feasible set of (GSIP) we refer to [15].

Let be given (x,y), $y \in Y(x)$. We say that at y the *Linear Independency Constraint Qualification* (LICQ) is satisfied for the lower level problem $Q(x)$ if

$$D_y g_i(x,y),\ i \in I(x,y) := \{i \in I \mid g_i(x,y) = 0\} \quad \text{are linearly independent.}$$

At $y \in Y(x)$, the weaker *Mangasarian Fromovitz Constraint Qualification* (MFCQ) is said to hold for $Q(x)$ if there exists a vector ξ such that

$$D_y g_i(x,y)\xi > 0, \quad \text{for all } \ i \in I(x,y).$$

The next lemma lists some standard sufficient conditions for the continuity of Y.

Lemma 2 *Under our assumptions on the set valued mapping Y we have.*

- *(a) If the function $g(x,y) = Ax + By - b$ is affine linear then Y is continuous on X.*
- *(b) Let $U \subset X$ be open. Let for any $x \in U$ the function $-g(x,y)$ be convex in y and let for any $x \in U$ the Slater condition hold: There exist $y = y(x) \in I\!R^m$ such that $g(x,y) > 0$. Then Y is continuous on U.*
- *(c) Let $U \subset X$ be open. Let for any $x \in U$ the condition (MFCQ) be fulfilled at all $y \in Y(x)$. Then Y is continuous on U.*

Lemma 2a shows that if g is affine linear then the feasible set M_{BL} is closed (see also Theorem 1 a). We give an example which shows that this need no more be true if g is not affine linear.

Example 1. Consider the bilevel problem

$$\begin{aligned} \min \quad & F(x,y) := y - x \qquad \text{subject to } x \le 1 \text{ and } \quad y \text{ is a solution of} \\ & \text{Q(x):}\ \ \min\ f(x,y) = -x - y \ \text{ s.t. } \ -xy \le 0,\ 0 \le y \le 2\ . \end{aligned}$$

Then,

$$Y(x) = \begin{cases} \{0\} & \text{if } x < 0 \\ [0,2] & \text{if } x \ge 0 \end{cases} \qquad \text{and} \qquad S(x) = \begin{cases} \{0\} & \text{if } x < 0 \\ \{2\} & \text{if } x \ge 0 \end{cases}$$

These mappings are not continuous at the point $x = 0$. We find $M_{\text{BL}} = \{(x,0) \mid x < 0\} \cup \{(x,2) \mid 0 \le x \le 1\}$. Obviously, M_{BL} is not closed and a global solution of (BL) does not exist. A local minimizer is $(\overline{x},\overline{y}) = (1,2)$.

A similar counterexample is given for (GSIP).

Example 2. Consider the generalized semi-infinite problem

$$\min\ F(x,y) := -x \qquad \text{s.t.}\ \ x \le 1 \text{ and } G(x,y) := -x - y \ge 0 \quad \text{for all } y \in Y(x)$$
$$\text{where } Y(x) = \{y \in I\!R \mid -xy \le 0,\ 0 \le y \le 2\}$$

Then, for the bilevel problem $\mathrm{BL}_{\mathrm{GSIP}}$ we find with the sets $Y(x)$, $S(x)$ in Example 1, $M_{\mathrm{BL}_{GSIP}} = \{(x,0) \mid x < 0\}$ and $M_{\mathrm{GSIP}} = \mathrm{pr}_x\ M_{\mathrm{BL}_{GSIP}} = (-\infty, 0)$. Again, M_{GSIP} is not closed and a solution of (GSIP) does not exist.

In view of these negative examples it seems natural to assume that Y is continuous on M_{BL}.

For (GSIP) we have to sharpen this condition slightly. Let be given a point $\overline{x} \notin \mathrm{dom}\, Y$, i.e. $Y(\overline{x}) = \emptyset$. Then, since $(\mathrm{dom}\, Y)^c$ is open, around $\overline{x}$ the problem (GSIP) can be regarded as an unconstrained problem, i.e. around $\overline{x}$ the problem does not have the structure of a real 'infinitely constrained problem'. Thus, to exclude this degenerate situation, in the sequel, we will assume that

$$M_{\mathrm{GSIP}} \subset \ \mathrm{int}\, X \quad \mathrm{where} X = \mathrm{dom}\, Y \tag{5}$$

This condition can always be satisfied by adding to the original constraints $G \ge 0$ appropriate extra conditions (such as $|x_i| \le \rho$). Remember that by the discussion above, assumption (5) implies that the problem (GSIP) is a (BL) problem with the special structure that lower level object function coincide with the the upper level constraint. In addition to (5) we assume that $Y(x)$ is continuous on M_{GSIP}. A natural sufficient condition for the continuity of Y is the condition (see Lemma 2(b)): For some open set U we have $M_{\mathrm{GSIP}} \subset U$ and

$$\text{for all } x \in U \ \text{(MFCQ) is satisfied for all } y \in Y(x), \tag{6}$$

This condition also implies (5).

In the following we are going to compare the structure of a general (BL) with the structure of a ($\mathrm{BL}_{\mathrm{GSIP}}$) satisfying (6).

3 Local reduction approach

A possible theoretical and practical approach for solving (BL) and (GSIP) is the so-called local reduction. For (SIP) and (GSIP) this is a standard approach introduced in [5] (see also [7] and [20]). It is also used in bilevel programming (see e.g. [2]). The idea is to transform the problem locally into a common finite optimization problem. Such a transformation is possible if certain regularity assumptions hold. Under these assumptions we obtain a system of optimality conditions for a minimizer of (BL) and (GSIP) and the solution can be computed by applying a (Quasi-) Newton method for solving this system of equations.

In the present section we derive the optimality conditions and discuss the question whether the regularity assumptions are natural conditions which can be expected to hold at the solution in the generic case.

<u>Local reduction for general (BL):</u> Let be given $\overline{x} \in I\!R^n$, $\overline{y} \in S(\overline{x})$, $(\overline{x}, \overline{y}) \in M_G$, i.e. $(\overline{x}, \overline{y})$ is feasible for (BL). Let the following assumption hold.

A1$_{\mathrm{BL}}$ There exist a neighborhood $U(\overline{x})$ of $\overline{x}$ and a C^1-function $y : U(\overline{x}) \to I\!R^m$ such that $y(\overline{x}) = \overline{y}$ and for any $x \in U(\overline{x})$ the vector $y(x)$ is the (unique, global) solution of $Q(x)$.

The following is a standard assumption in parametric optimization and sufficient for A1$_{\mathrm{BL}}$. (This assumption is often used in nonlinear bilevel programming see e.g. [2, Assumpt. A2]).

A2$_{\mathrm{BL}}$ All problem functions of (BL) are C^2-functions and at the unique solution $\overline{y}$ of $Q(\overline{x})$ we have

(1) (LICQ) is satisfied and the Kuhn-Tucker condition with multipliers $\overline{\gamma}_i > 0$ (strict complementary slackness):

$$D_y \mathcal{L}^{\overline{y}}(\overline{x}, \overline{y}, \overline{\gamma}) := D_y f(\overline{x}, \overline{y}) - \sum_{i \in I(\overline{x}, \overline{y})} \overline{\gamma}_i D_y g_i(\overline{x}, \overline{y}) = 0 \ .$$

(2) The standard second order sufficiency optimality condition: the Hessian $D_y^2 \mathcal{L}^{\overline{y}}(\overline{x}, \overline{y}, \overline{\gamma})$ is positive definite on the tangent space $\{\xi \in I\!R^q \mid D_y g_i(\overline{x}, \overline{y})\xi = 0, \ i \in I(\overline{x}, \overline{y})\}$.

Obviously, under A1$_{\mathrm{BL}}$, on $U(\overline{x})$, the problem (BL) is equivalent with the so-called *locally reduced problem*

$$\mathrm{BL}_{\overline{x}}: \quad \min_x \ \hat{F}(x) := F(x, y(x)) \qquad \text{subject to } \hat{G}(x) := G(x, y(x)) \geq 0.$$

BL$_{\overline{x}}$ is a finite optimization problem and standard optimality conditions applied to this problem lead to optimality conditions for (BL) as follows.

Suppose, the active gradients $D\hat{G}_j(\overline{x}) = D_x G_j(\overline{x}, \overline{y}) + D_y G_j(\overline{x}, \overline{y})\overline{\eta}$, $j \in J(\overline{x}) := \{j \in \{1, \ldots, p\} \mid G_j(\overline{x}, \overline{y}) = 0\}$ are linearly independent, where we set $\overline{\eta} = Dy(\overline{x})$. Then, a necessary optimality condition for $(\overline{x}, \overline{y})$ to solve (BL) is: There exist multipliers $\overline{\mu}_j \geq 0$ such that

$$\begin{aligned} D\hat{F}(\overline{x}) - \sum_{j \in J(\overline{x})} \overline{\mu}_j D\hat{G}_j(\overline{x}) \quad &= \quad D_x F(\overline{x}, \overline{y}) - \sum_{j \in J(\overline{x})} \overline{\mu}_j D_x G_j(\overline{x}, \overline{y}) \\ &+ \quad \Big(D_y F(\overline{x}, \overline{y}) - \sum_{j \in J(\overline{x})} \overline{\mu}_j D_y G_j(\overline{x}, \overline{y}) \Big) \overline{\eta} = 0. \end{aligned} \tag{7}$$

Consider now the Kuhn-Tucker conditions for a solution $y = y(x)$ of $Q(x)$,

$$H(x, y, \gamma) := \begin{array}{r} D_y \mathcal{L}^{\overline{y}}(x, y, \gamma) = 0 \\ g_i(x, y) = 0 \quad i \in I(\overline{x}, \overline{y}). \end{array} \tag{8}$$

Under assumption A2$_{\mathrm{BL}}$, by applying the Implicit Function Theorem to $H = 0$, it follows that there exist a neighborhood $U(\overline{x})$ of $\overline{x}$ and C^1-functions $y(x)$, $\gamma(x)$ such that $H(x, y(x), \gamma(x)) = 0$, $x \in U(\overline{x})$ and $y(x)$ is the solution function in A1$_{\mathrm{BL}}$ with corresponding multiplier $\gamma(x)$. By differentiating the relation $H(x, y(x), \gamma(x)) = 0$ w.r.t. x we find

$$D_x H(\overline{x}, \overline{y}, \overline{\gamma}) + D_y H(\overline{x}, \overline{y}, \overline{\gamma})\, \overline{\eta} + D_\gamma H(\overline{x}, \overline{y}, \overline{\gamma})\, \overline{\theta} = 0 \ ,$$

where $\overline{\theta} = D\gamma(\overline{x})$. Altogether, in view of (7) (8), we obtain the following system of equations for a solution $(\overline{x}, \overline{y})$ of (BL) and the corresponding multipliers and their derivatives $(\overline{\mu}, \overline{\gamma}, \overline{\eta}, \overline{\theta})$:

$$\begin{aligned} D_x F(x,y) - \sum_{j \in J(\overline{x})} \mu_j D_x G_j(x,y) \qquad & \\ + \big(D_y F(x,y) - \sum_{j \in J(\overline{x})} \mu_j D_y G_j(x,y)\big)\eta &= 0 \\ G_j(x,y) &= 0 \quad j \in J(\overline{x}) \\ D_y \mathcal{L}^{\overline{y}}(x,y,\gamma) &= 0 \\ g_i(x,y) &= 0 \quad i \in I(\overline{x},\overline{y}). \\ D_x H(x,y,\gamma) + D_y H(x,y,\gamma)\,\eta + D_\gamma H(x,y,\gamma)\,\theta &= 0 \end{aligned} \tag{9}$$

This is a system of $n + m + |J(\overline{x})| + |I(\overline{x}, \overline{y})| + n\ m + n\ |I(\overline{x}, \overline{y})|$ equations for the same number of unknowns $x, y, \mu, \gamma, \eta, \theta$. To compute a solution $(\overline{x}, \overline{y})$ of (BL), we could apply a (Quasi-) Newton procedure for solving the system (9).

We are now interested to know whether assumption $\mathrm{A2_{BL}}$ - essential for the local reduction - is a natural condition. We first give an illustrative example.

Example 3. Consider the bilevel problem

$$\begin{aligned} P_1: \qquad & \max\ F(x,y) := x + y \qquad \text{s.t. } G(x,y) := x + 2y - 8 \le 0 \text{ and } y \text{ solves} \\ & \mathrm{Q(x)}: \ \max\ f(x,y) := x + y \quad \text{ s.t. } \ 0 \le y \le 4\ . \end{aligned}$$

Here, $S(x) = \{y(x) = 4\}$ and the feasible set is $M_{\mathrm{BL}} = \{(x,4) \mid x \le 0\}$. The optimal solution $y(x) = 4$ of $Q(x)$ is feasible w.r.t. $G \ge 0$ only for $x \le 0$. Hence, the solution of (BL) is $(\overline{x}, \overline{y}) = (0,4)$ with value $F(\overline{x}, \overline{y}) = f(\overline{x}, \overline{y}) = 4$. It is easily checked, that at the solution $y(x) = 4$ of $Q(x)$ the conditions of $\mathrm{A2_{BL}}(1)$ is fulfilled (the second order condition (2) is superfluous since $Q(x)$ is a linear program).

Now let us consider the problem P_2 obtained from P_1 by only moving the condition $G \ge 0$ to the lower level,

$$\begin{aligned} P_2: \qquad & \max\ F(x,y) = x + y \qquad \text{s.t. } y \text{ solves} \\ & \mathrm{Q(x)}: \ \max\ f(x,y) = x + y \quad \text{ s.t. } \ 0 \le y \le 4, \quad x + 2y - 8 \le 0 \end{aligned}$$

Then for P_2 also the points $\{(x,y) \mid y = 4 - \frac{x}{2},\ x \in [0,8]\}$ become feasible and the point $(\hat{x}, \hat{y}) = (8,0)$ is optimal (with a better value $F(\hat{x}, \hat{y}) = f(\hat{x}, \hat{y}) = 8$). However now, for P_2 the solution $\hat{y} = 0$ of $Q(\hat{x})$ does not fulfill the assumptions $\mathrm{A2_{BL}}(1)$. The point $\hat{y} = 0$ is a degenerate vertex-solution of the linear program $Q(\hat{x})$. At $\hat{y} = 0$, even the (MFCQ) is not valid for $Q(\hat{x})$.

Remark. By definition of (LICQ) the assumption $\mathrm{A2_{BL}}$ (1) can only hold if there are not more than m lower level constraints active at a solution $(\hat{x}, \hat{y})$ of (BL). We will see in Section 4 for the linear case (cf. Theorem 2) that the number of active lower level constraints is directly related to the number of active upper level constraints. In fact, as in the preceeding examples, it will appear that (LICQ) is not valid in the lower level if there are not enought ($\geq n$) upper level constraints active at $(\hat{x}, \hat{y})$. In this case the 'reduction approach' is not possible.

Note, that a typical bilevel model in operations research need not have any upper level constraints. In fact viewing a (BL) in (1) as a game between an upper level player 1 and a lower level player 2, the player 2 could accept the upper level constraints $G \geq 0$ in his lower level problem $Q(x)$. In the example above, the strategy to pass the upper level constraints to the lower level problem even leads to a better object value for both players. The next Lemma shows that for the upper level player such a policy is always an advantage (for the lower level player it may be advantageous but also unfavourable depending whether his object is 'similar' or 'adverse' to the upper level object). Let $(\widetilde{BL})$ denote the bilevel problem obtained from (BL) by passing the constraints $G \geq 0$ to the lower level constraints $g \geq 0$ and let $\widetilde{M}_{\mathrm{BL}}$ denote the corresponding feasible set.

Lemma 3 *Let be given a bilevel problem (1). Then we have $M_{BL} \subset \widetilde{M}_{BL}$ and for the solutions $(\overline{x}, \overline{y})$ of (BL), $(\tilde{x}, \tilde{y})$ of $(\widetilde{BL})$, respectively, it follows $F(\tilde{x}, \tilde{y}) \leq F(\overline{x}, \overline{y})$.*

PROOF. Let (x, y) be feasible for (BL), i.e. y is a solution of $Q(x)$ and $G(x, y) \geq 0$. Then y is also feasible for the lower level problem of $(\widetilde{BL})$,

$$\widetilde{Q}(x): \quad \min_y f(x, y) \quad \text{s.t.} \quad G(x, y) \geq 0, \quad g(x, y) \geq 0.$$

Since the feasible set $\widetilde{Y}(x)$ of $\widetilde{Q}(x)$ is contained in the feasible set $Y(x)$ of $Q(x)$, y must also be a solution of $\widetilde{Q}(x)$, i.e. (x, y) is feasible for $(\widetilde{BL})$. □

The situation for (GSIP) and its bilevel formulation $\mathrm{BL}_{\mathrm{GSIP}}$ (see (3)) is quite different. Let $\overline{x}$ be feasible for (GSIP). We define the set of active points

$$Y_0(\overline{x}) = \{\overline{y} \in Y(\overline{x}) \mid \min_{1 \leq j \leq p} G_j(\overline{x}, \overline{y}) = 0\}.$$

Note that $\overline{x}$ is feasible if and only if $G(\overline{x}, y) \geq 0$ for all $y \in Y(\overline{x})$. Thus for $\overline{x} \in M_{\mathrm{GSIP}}$ every point $\overline{y} \in Y_0(\overline{x})$ is a global solution of $Q(\overline{x})$. Suppose now that $\overline{x}$ is a solution of (GSIP) and that $Y_0(\overline{x}) = \emptyset$. Then by continuity assumptions in Section 1, near $\overline{x}$ the problem (GSIP) is equivalent with the unconstrained problem $\min_x \; F_{\mathrm{GSIP}}(x)$. We exclude such a situation by assuming that in addition to (5) the following holds:

$$Y_0(\overline{x}) \neq \emptyset \qquad \text{for any local solution } \overline{x} \text{ of (GSIP)}.$$

Consequently, the bilevel problems $\mathrm{BL}_{\mathrm{GSIP}}$ related to generalized semi-infinite problems intrinsicly have at least one upper level constraint active in the solution. Moreover, typically in (GSIP) the degree of freedom in the variable x 'forces' the solution $\overline{x}$ of $\mathrm{BL}_{\mathrm{GSIP}}$ to a location such that as many active points $\overline{y}_l \in Y_0(\overline{x})$ occur, (i.e. solutions $\overline{y}_l$ of $Q(\overline{x})$) as the degree of freedom in the minimization model allows. This behavior is illustrated with the following geometrical interpretation of (GSIP) (cf. also [20]):

Given the region $M_G = \{(x,y) \mid G(x,y) \geq 0\}$ in $\mathbb{R}^n \times \mathbb{R}^m$. Then we have to find x such that the set $\{x\} \times Y(x)$ is contained in M_G and such that some functional $f(x)$ is maximized. Often the function $f(x)$ can be viewed as the volume of the set $Y(x)$.

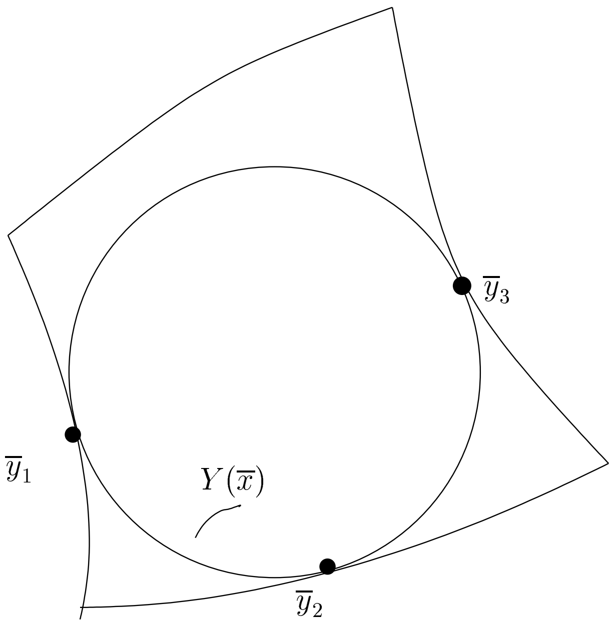

Figure 1: Illustration of a (GSIP) at the solution $\overline{x}$.

Figure 1 illustrates the situation of such a problem at the solution $\overline{x}$ of (GSIP). The different points $\overline{y}_1, \overline{y}_2, \overline{y}_3$ where the set $Y(\overline{x})$ touches the set $\{y \mid G(\overline{x}, y) = 0\}$ are the solutions of the lower level problem $Q(\overline{x})$, i.e. $Y_0(\overline{x}) = \{\overline{y}_1, \overline{y}_2, \overline{y}_3\}$. So typically in semi-infinite optimization we have to admit different solutions of $Q(x)$.

Local reduction for (GSIP): In semi-infinite optimization the local reduction is a standard technique. As motivated above, because of the special structure of (GSIP), in contrast to the general (BL) case, at a solution $\overline{x}$ of (GSIP) typically different solutions of $Q(\overline{x})$ must be considered, i.e. $Y_0(\overline{x})$ need not to be a singleton.

Let be given $\overline{x} \in M_{\text{GSIP}}$ and let $Y_0(\overline{x})$ consist of finitely many points, $Y_0(\overline{x}) = \{\overline{y}^1, \ldots, \overline{y}^r\}$, $r \geq 1$. We make the following assumption.

A1$_{\text{GSIP}}$ There exist a neighborhood $U(\overline{x})$ of $\overline{x}$ and r C^1-functions $y^l : U(\overline{x}) \to \mathbb{R}^m$, such that $y^l(\overline{x}) = \overline{y}^l$, and for any $x \in U(\overline{x})$ the values $y^l(x)$, $l = 1, \ldots, r$, include all (global) solution of $Q(x)$.

As in the (BL) case we give a natural sufficiency condition for A1$_{\text{GSIP}}$.

A2$_{\text{GSIP}}$ All problem functions of (GSIP) are C^2-functions. Let for $\overline{y} \in Y_0(\overline{x})$ be defined $J_{\overline{y}} := \{j \in J \mid G_j(\overline{x}, \overline{y}) = 0\}$ and for $j \in J_{\overline{y}}$,

$$Q_j(\overline{x}): \quad \min\ G_j(\overline{x}, y) \quad \text{s.t.} \quad y \in Y(\overline{x}) \ .$$

For all $\overline{y} \in Y_0(\overline{x})$ and all $j \in J_{\overline{y}}$ we have with the Lagrange function $\mathcal{L}_j^{\overline{y}}(\overline{x}, y, \gamma) = G_j(\overline{x}, y) - \sum_{i \in I(\overline{x},\overline{y})} \gamma_i g_i(\overline{x}, y)$:

(1) (LICQ) is satisfied at $\overline{y}$ for $Q_j(\overline{x})$ and the Kuhn-Tucker condition $D_y \mathcal{L}_j^{\overline{y}}(\overline{x}, \overline{y}, \overline{\gamma}) = 0$ with multipliers $\overline{\gamma}_i^j > 0$ (strict complementary slackness).

(2) The standard second order sufficiency optimality condition at $\overline{y}$ for $Q_j(\overline{x})$: the Hessian $D_y^2 \mathcal{L}_j^{\overline{y}}(\overline{x}, \overline{y}, \overline{\gamma})$ is positive definite on $\{\xi \in I\!R^q \mid D_y g_i(\overline{x}, \overline{y})\xi = 0,\ i \in I(\overline{x}, \overline{y})\}$.

Under A2$_{\text{GSIP}}$ the set $Y_0(\overline{x})$ must be finite, $Y_0(\overline{x}) = \{\overline{y}^l,\ l = 1, \ldots, r\}$, and locally in a neighborhood $U(\overline{x})$ of $\overline{x}$, the problem (GSIP) is equivalent with the *locally reduced problem*

$$\text{GSIP}_{\overline{x}}: \quad \min\ F_{\text{GSIP}}(x) \quad \text{s.t.} \quad \hat{G}_j^l(x) := G_j(x, y^l(x)) \geq 0,\ j \in J_{\overline{y}^l},\ l = 1, \ldots, r.$$

Again GSIP$_{\overline{x}}$ is a finite optimization problem and optimality conditions of finite optimization applied to this problem lead to optimality conditions for (GSIP). We only give the conditions for the case $p = 1$ (see also [8], [20]; the modification to the case $p > 1$ is straightforward).

Similar to (7) with $Y_0(\overline{x}) = \{\overline{y}^l,\ l = 1, \ldots, r\}$ we obtain for $p = 1$ the optimality condition (putting $y^l = y^l(x)$),

$$DF_{\text{GSIP}}(x) - \sum_{l=1}^{r} \mu_l \Big(D_x G(x, y^l) + D_y G(x, y^l) D y^l \Big) = 0 \tag{10}$$

with multipliers $\mu_l \geq 0$. For (GSIP) however the equations simplify. Consider the Kuhn-Tucker condition for $Q(x)$ at the solutions y^l:

$$D_y G(x, y^l) = \sum_{i \in I(\overline{x}, \overline{y}^l)} \gamma_i^l(x) D_y g_i(x, y^l)$$

and $g_i(x, y^l) = 0,\ i \in I(\overline{x}, \overline{y}^l)$. By differentiating the relation $g_i(x, y^l(x)) = 0$ we find $D_x g_i(x, y^l) = -D_y g_i(x, y^l) D y^l$ and

$$D_y G(x, y^l) D y^l = \sum_{i \in I(\overline{x}, \overline{y}^l)} \gamma_i^l D_y g_i(x, y^l) D y^l = - \sum_{i \in I(\overline{x}, \overline{y}^l)} \gamma_i^l D_x g_i(x, y^l) \ .$$

Substituting in (10) we obtain the following system of optimality conditions:

$$\begin{aligned}
DF_{\text{GSIP}}(x) - \sum_{l=1}^{r} \mu_l \Big(D_x G(x, y^l) - \sum_{i \in I(\overline{x}, \overline{y}^l)} \gamma_i^l D_x g_i(x, y^l) \Big) &= 0 \\
G(x, y^l) &= 0 \qquad l = 1, \ldots, r \\
\text{and for } l = 1, \ldots, r \qquad & \qquad\qquad\qquad (11) \\
D_y G(x, y^l) - \sum_{i \in I(\overline{x}, \overline{y}^l)} \gamma_i^j D_y g_i(x, y^l) &= 0 \\
g_i(x, y^l) &= 0 \qquad i \in I(\overline{x}, \overline{y}^l)
\end{aligned}$$

This system consists of $K := n + r + \sum_{l=1}^{r} \big(m + |I(\overline{x}, \overline{y}^l)|\big)$ equations for the K unknowns $x \in \mathbb{R}^n$, $\mu_l \in \mathbb{R}$, $y^l \in \mathbb{R}^m$, $\gamma^l \in \mathbb{R}^{|I(\overline{x},\overline{y}^l)|}, l = 1, \dots, r$. In [22] it has been shown that under the assumption $\mathrm{A2}_{\mathrm{GSIP}}$ at a solution $\overline{x}$ of (GSIP) and the additional assumptions that the gradients $\big(D_x G(x, y^l) - \sum_{i \in I(\overline{x}, \overline{y}^l)} \gamma_i^j D_x g_i(x, y^l)\big)$, $l = 1, \dots, r$, are linearly independent, the Jacobian of (11) is regular at the solution $(\overline{x}, \overline{y}^1, \overline{\gamma}^1, \dots, \overline{y}^r, \overline{\gamma}^r)$. Hence, to solve (GSIP) numerically, we can apply a (Quasi) Newton method to (11). See e.g. [7] for solving common (SIP) problems by Newton methods.

Note that in contrast to equation (9) for the (BL)-problems the optimality condition (11) for $\mathrm{BL}_{\mathrm{GSIP}}$ does not contain the derivatives $Dy^l, D\gamma^l$ as unknowns. The reason is that for the problem $\mathrm{BL}_{\mathrm{GSIP}}$ the upper level constraints G_j coincide with the lower level objects. So only the information of the value function of $Q(x)$ is really needed in the upper level and not the full information about the solution $y(x)$.

We are now going to discuss the question whether the assumption $\mathrm{A2}_{\mathrm{BL}}$ or $\mathrm{A2}_{\mathrm{GSIP}}$ for the local reduction can be expected to hold generically at a solution. By a generic subset $\mathcal{S}$ of a problem set $\mathcal{P}$ we roughly mean a subset which is open and dense in $\mathcal{P}$ (in some appropriate topology).

For the problem P_2 of Example 3 the assumption (LICQ) in $\mathrm{A2}_{\mathrm{BL}}$ (even (MFCQ)) is not valid. This negative behavior is stable w.r.t. smooth nonlinear (small) perturbations. Hence we can state.

> For the general class of (BL) problems the assumption $\mathrm{A2}_{\mathrm{BL}}$ is not generic at a local solution $(\overline{x}, \overline{y})$. For typical classes of bilevel problems, in particular problems without upper level constraints, (LICQ) or even (MFCQ) will not be satisfied at the solution $\overline{y}$ of $Q(\overline{x})$. For such problems we cannot expect a 'nice' system of optimality conditions for $(\overline{x}, \overline{y})$ which can be solved with smooth methods. Consequently in this situation the 'reduction approach' can only be used with caution.

As indicated above, the special class of bilevel problems $\mathrm{BL}_{\mathrm{GSIP}}$ related to (GSIP) may have a better genericity behavior. For the sub-class of common semi-infinite problems it has been shown in [11] that $\mathrm{A2}_{\mathrm{GSIP}}$ is generically fulfilled at each local solution. A similar genericity analysis for (GSIP) has not yet been done. In [17] some particular results are obtained. It has been proven for example that generically for (GSIP) the number $|Y_0(\overline{x})|$ of lower level local minima at a solution $\overline{x}$ is bounded by n, $|Y_0(\overline{x})| \leq n$. We will show for the linear case in the next section that generically $|Y_0(\overline{x})| = n$ holds. We formulate the

> **Conjecture:** In the class $\mathrm{BL}_{\mathrm{GSIP}}$ (appropriately defined) the assumptions $\mathrm{A2}_{\mathrm{GSIP}}$ holds generically at a solution $\overline{x}$ of a (GSIP) problem. In particular generically, all local minima $\overline{y}_l$, $l = 1, \dots, r$, of $Q(\overline{x})$ are nondegenerate minima.

In the next section this conjecture is proven for the special case of linear problems (see Theorem 3). We also will present a detailed analysis of the negative results for the general class of linear bilevel problems.

Summarizing, roughly speaking, generically for classes of general (BL) problems from operational research, at a solution $(\overline{x}, \overline{y})$, the minimizer $\overline{y}$ of $Q(\overline{x})$ will be a unique minimizer but (LICQ) (or even (MFCQ)) will not be satisfied at $\overline{y}$. In contrast, for (GSIP) generically we expect at a solution different minimizer $\overline{y}_l$ of $Q(\overline{x})$ but each solution will be non-degenerate, such that a smooth approach for solving (GSIP) is possible.

To analyze the difference between (BL) and (GSIP) in the next section, we have to modify the bilevel formulation of (GSIP) in (3). In the formulation (3), when the upper level contains different (smooth) constraints, i.e. if $p > 1$, then the object function $f(x, y) = \min_{1 \le j \le p} G_j(x, y)$ is not a C^1-function (only Lipschitz-continuous). To transform (GSIP) into a smooth (BL) we consider the following generalization of the bilevel problem (3),

$$\begin{aligned} \text{(BL):} \quad & \min_{x,y} \; F(x,y) \quad \text{s.t. } G(x, y_1, \ldots, y_r) \ge 0 \text{ and for } \; l = 1, \ldots, r \qquad (12) \\ & y_l \text{ solves } \quad Q_l(x) \colon \; \min_{y_l} \; f_l(x, y_l) \quad \text{s.t.} \quad g^l(x, y_l) \ge 0, \end{aligned}$$

where $x \in I\!R^n$, $y = (y_1, \ldots, y_r)$, $y_l \in R^{m_l}, l = 1, \ldots, r$, and the functions are defined accordingly. This problem can be seen as a game between an upper level player and r different lower level players with r different lower level problems Q_l. The problem BL_{GSIP} (and (GSIP)) can be written as

$$\begin{aligned} (\text{BL}_{\text{GSIP}}) \colon \quad & \min \; F_{\text{GSIP}}(x) \quad \text{s.t. } G_l(x, y_l) \ge 0 \quad l = 1, \ldots, p \;\text{ where } \qquad (13) \\ & y_l \text{ solves } Q_l(x) \colon \; \min_{y_l} \; G_l(x, y_l) \quad \text{s.t.} \quad g(x, y_l) \ge 0. \end{aligned}$$

This represents a bilevel problem (12) with the special conditions $f_l = G_l$, $m_l = m$, $r = p$ and $g^l = g$ not depending on l.

4 Linear problems

In this section we are concerned with linear (GSIP) and (BL), i.e. all problem functions are affine linear. We describe the structure of the feasible sets and analyze which kind of regularity can be expected at a solution $(\overline{x}, \overline{y})$ of (BL) or at a solution $\overline{x}$ of (GSIP). We will show that a general (BL) and the special case of a bilevel problem BL_{GSIP} arising from (GSIP) may have different generic behavior.

We consider the following linear bilevel problem (cf. (12)) with $y = (y_1, \ldots, y_r) \in I\!R^{m_0}$, $y_l \in I\!R^{m_l}$, $l = 1, \ldots, r$, $m_0 = m_1 + \cdots + m_r$,

$$\begin{aligned} \text{(LBL):} \quad & \min \; c_0^T x + d_0^T y \quad \text{subject to} \qquad (14) \\ & A_0 x + B_0 y - b_0 \ge 0 \quad \text{ and for } l = 1, \ldots, r, \; y_l \text{ is a solution of} \\ & Q_l(x) \colon \; \min_{y_l} \; c_l^T x + d_l^T y_l \quad \text{s.t.} \quad A_l x + B_l y_l - b_l \ge 0. \end{aligned}$$

Here, $c_0 \in I\!R^n$, $d_0 \in I\!R^{m_0}$, A_0 is a $p \times n$-matrix, B_0 is a $p \times m_0$-matrix, A_l are $q_l \times n$-matrices, B_l are $q_l \times m_l$-matrices etc. A linear (GSIP) is of the form

$$\begin{aligned} \text{(LGSIP):} \quad \min \; c_0^T x \quad & \text{s.t. } A_0 x + B_0 y - b_0 \ge 0 \quad \text{for all } y \text{ in Y(x)} \\ & \text{where } \; Y(x) = \{ y \in I\!R^m \mid Ax + By - b \ge 0 \}. \end{aligned}$$

Let in the sequel $a_l^0,\ b_l^0$ denote the rows of $A_0,\ B_0$, respectively and let β_l^0 be the components of b_0. Then (LGSIP) can be written in the form of a (LBL),

$$
\begin{aligned}
(\mathrm{BL}_{\mathrm{LGSIP}})\colon \quad & \min\ c_0^T x \quad \text{s.t. } G_l(x,y) := (a_l^0)^T x + (b_l^0)^T y_l - \beta_l^0 \geq 0,\ l = 1,\dots,p, \\
& \text{and for } l = 1,\dots,p,\ y_l \text{ is a solution of} \\
& Q_l(x)\colon\ \min_{y_l}\ (a_l^0)^T x + (b_l^0)^T y_l \quad \text{s.t.} \quad Ax + By_l - b \geq 0.
\end{aligned}
$$

with $A_l = A,\ B_l = B,\ b_l = b$ not depending on l and $p = r$. We have to complete our notation. With $y = (y_1,\dots,y_r)$ we define for $l = 1,\dots,r$:

$Y_l(x)$	$= \{y_l \in \mathbb{R}^{m_l},\ (y \in \mathbb{R}^{m_0}) \mid A_l x + B_l y_l - b_l \geq 0\}$	feasible set of $Q_l(x)$
Y_l	$= \{(x,y_l),((x,y)) \mid y_l \in Y_l(x)\}$	the graph of $Y_l(x)$
$S_l(x)$	$= \{y_l \in \mathbb{R}^{m_l},(y \in \mathbb{R}^{m_0}) \mid y_l \text{ solves } Q_l(x)\}$	set of solutions of$Q_l(x)$
M_G	$= \{(x,y) \mid G(x,y) := A_0 x + B_0 y - b_0 \geq 0\}$	upper level constraints
X	$= \cap_l \operatorname{dom} Y_l\ (= \cap_l \{x \mid Y_l(x) \neq \emptyset\})$	
M_{sem}	$= M_G \cap (\cap_l Y_l)$	the semi-feasible set
S	$= \{(x,y) \mid y_l \in S_l(x),\ l = 1,\dots,r\}$	the solution graph
M_{BL}	$= \{(x,y) \mid (x,y) \in M_{\mathrm{sem}} \cap S\}$	feasible set of (LBL)

We will regard the sets $Y_l(x),\ S_l(x)$ as sets in $\mathbb{R}^{m_l}$ or as sets in $\mathbb{R}^{m_0}$ depending on the context. We introduce the following assumptions.

AL$_1$ The sets $S_l(x)$ are compact subsets of $\mathbb{R}^{m_l}$, $l = 1,\dots,r$, for all $x \in R^n$. This assumption in particular implies that $Q_l(x)$ always has a vertex solution (if $S_l(\overline{x}) \neq \emptyset$).

AL$_2$ The polyhedron $M_{\mathrm{sem}} \subset \mathbb{R}^n \times \mathbb{R}^{m_0}$ is bounded (thus compact).

Recall that for $Y(x)$ the Slater condition is said to hold if there exists $\tilde{y} = \tilde{y}(x) \in \mathbb{R}^m$ such that $Ax + B\tilde{y} - b > 0$ (see Lemma 1(b)). In view of our regularity assumption (6) for (LGSIP) we consider the following assumption.

AL$_3$ For (LGSIP) let the Slater condition be satisfied for all $x \in M_{\mathrm{GSIP}}$. (Then in particular, (LGSIP) is equivalent with $\mathrm{BL}_{\mathrm{LGSIP}}$.)

The following theorem contains the main results on the structure of the feasible set and the solution of a (LBL).

Theorem 1 *For (LBL) the following holds.*

(a) The feasible set $M_{BL} = M_{sem} \cap S$ consist of a union of finitely many faces f^k of the polyhedron M_{sem},

$$M_{BL} = \cup_{k=1}^K f^k.$$

In particular, M_{BL} is a closed set in $R^n \times \mathbb{R}^{m_0}$.

(b) If no upper level constraints are present, then the set M_{BL} is path-wise connected.

If the assumptions AL_1 and AL_2 hold then we have.

(c) The solution of (LBL) occurs at a vertex of some face f^{k_0}, $k_0 \in \{1, \ldots, K\}$, and thus at a vertex of M_{sem}

(d) The value functions $v_l(x) := \min_{y_l \in Y_l(x)} c_l^T x + d_l^T y_l$ of $Q_l(x)$, $l = 1, \ldots, r$ are convex and Lipschitz continuous on dom Y_l.

PROOF. (For a proof of (a) and (c) for the case $r = 1$ we refer e.g. to [16]). For completeness we give a proof for the general case $r \geq 1$.
(a) The set M_{sem} is a polyhedron in $I\!R^n \times I\!R^{m_0}$. Let f^0 be a d-dimensional face of M_{sem}, with $d \geq 1$ (i.e. f^0 is not a vertex). Let (x^0, y^0) be a point in the relative interior of f^0 and let (x^0, y^0) belong to M_{BL}. We now show that the whole face f^0 belongs to M_{BL}. Then, since every point of M_{sem} which is not a vertex is contained in the relative interior of some face of M_{sem}, the proof is completed.

Let $(x^1, y^1) \in f^0$ be arbitrary. Since (x^0, y^0) is a point in the relative interior of f^0 there exists a point (x^2, y^2) in f^0 such that with some λ, $0 < \lambda < 1$,

$$(x^0, y^0) = \lambda(x^1, y^1) + (1 - \lambda)(x^2, y^2). \tag{15}$$

Given any point $(x^1, \overline{y})$ in $\cap_l Y_l$, i.e. $\overline{y}_l \in Y_l(x^1)$, we consider

$$(x^0, y^*) := \lambda(x^1, \overline{y}) + (1 - \lambda)(x^2, y^2).$$

It follows $y_l^* \in Y_l(x^0)$. Since $(x^0, y^0) \in M_{\text{BL}}$, i.e. y_l^0 are solutions of $Q_l(x^0)$, in view of (15) we have,

$$d_l^T y_l^0 = \lambda d_l^T y_l^1 + (1 - \lambda) d_l^T y_l^2 \leq d_l^T y_l^* = \lambda d_l^T \overline{y}_l + (1 - \lambda) d_l^T y_l^2$$

and $d_l^T y_l^1 \leq d_l^T \overline{y}_l$. This implies that y_l^1 are solutions of $Q_l(x^1)$ and $(x^1, y^1) \in M_{\text{BL}}$.
(b) See [22] for $r = 1$. The generalization to $r > 1$ is not difficult.
(c) In view of (a) the feasible set M_{BL} consists of the union of faces $f^1, \ldots, f^K$ of the compact set M_{sem} (cf. AL_2). A solution of (LBL) exists and must be contained in some of these faces say f^{k_0}. Thus the problem (LBL) can be replaced by the problem

$$\min_{x,y} \; c_0^T x + d_0^T y \quad \text{s.t.} \quad (x, y) \in f^{k_0}.$$

Since f^{k_0} is a bounded polyhedron the minimum is attained at a vertex $(\overline{x}, \overline{y})$ of f^{k_0}. Since f^{k_0} is a face of M_{sem} the point $(\overline{x}, \overline{y})$ is also a vertex of M_{sem}.
(d) See [16]. □

Note that the set M_{LBL}, as the union of faces of the polyhedron M_{sem}, is typically a non-convex set. The same holds for $\text{pr}_x \, M_{\text{BL}_{LGSIP}}$ which may have re-entrant corners (see e.g. [17]).

From Theorem 1 in view of Lemma 1(b) and using AL_3 we directly obtain the following corollary for (BL_{LGSIP}). Recall that F_{GSIP} does not depend on y and that M_{LGSIP} is a subset of $I\!R^n$.

Corollary 1 *Let be given (LGSIP) satisfying AL_3. Then the feasible set $M_{LGSIP} = pr_x\ M_{BL_{LGSIP}}$ is the subset of the polyhedron $pr_x\ M_{sem}$. given by a union of polyhedra $pr_x\ f^k$, $k = 1, \ldots, K$. In particular, M_{LGSIP} is closed.*

PROOF. We have only to note that the projection of the polyhedra M_{sem}, f^k are again polyhedra. $\square$

We are now going to describe the structural difference between a general (LBL) and a problem (BL_{LGSIP}).

Theorem 2 *Let $(\overline{x}, \overline{y})$ be a vertex solution of (LBL) in (14), i.e. $(\overline{x}, \overline{y})$ is a vertex of M_{sem}. Suppose, for the number p of upper level constraints we have $p < n$. Then at least one of the solutions $\overline{y}_l$ of $Q_l(\overline{x})$, say $\overline{y}_{l_0}$, does not fulfill the condition (LICQ) for $Q_{l_0}(\overline{x})$ (or even not (MFCQ)), i.e. $\overline{y}_{l_0}$ is a degenerate vertex solution of the linear problem $Q_{l_0}(\overline{x})$.*

PROOF. If (LICQ) is satisfied at $\overline{y}_l \in I\!R^{m_l}$ for $Q_l(\overline{x})$, then at most m_l of the inequalities $A_l\overline{x} + B_l\overline{y}_l - b_l \geq 0$ can be active ($l = 1, \ldots, r$). Together with maximally p active constraints in the upper level, the number of active constraints for $(\overline{x}, \overline{y})$ is less than or equal to $p + \sum_{l=1}^{r} m_l = p + m_0 < n + m_0$. Consequently $(\overline{x}, \overline{y})$ cannot be a vertex of the polyhedron M_{sem} in $I\!R^n \times I\!R^{m_0}$. $\square$

In view of Theorem 2, when the number p of constraints $G \geq 0$ in the upper level is too small, the regularity assumption A2_{BL} (1) in Section 2 cannot hold. (Note that this situation is stable under small smooth nonlinear perturbations.) In the extreme case, a general (BL) may have no constraints in the upper level (i.e. $p = 0$). By definition, as we have discussed in Section 2, a BL_{GSIP} always has at least one upper level constraint. This difference makes the generalized semi-infinite problems behave better. We give an illustrative example.

Consider the (BL) without constraints in the upper level,

$$\max\ x + y \qquad \text{subject to}\ \ y \text{ is a solution of}$$
$$\text{Q(x):}\ \ \max_{y}\ 2y \quad \text{s.t.} \quad 0 \leq y \leq \tfrac{1}{2}x\ , \quad 2y + x - 1 \leq 0$$

Here, the feasible set M_{LBL} is given by the union $f^1 \cup f^2$ of the faces $f^1 = \{(x, y)\ |y = \frac{1}{2}x,\ 0 \leq x \leq \frac{1}{2}\}$ and $f^2 = \{(x, y)\ |2y + x - 1 = 0,\ \frac{1}{2} \leq x \leq 1\}$ of M_{sem}. The solution is attained at the vertex $(\overline{x}, \overline{y}) = (1, 0)$. At the solution $\overline{y}$ of the one-dimensional problem $Q(\overline{x})$ two lower level constraints $y = 0,\ 2y + x - 1 = 0$ are active. Thus, $\overline{y}$ is a degenerate vertex of $Y(\overline{x})$. Since $Y(\overline{x}) = \{\overline{y}\}$ consists only of one point, the Slater condition (MFCQ) is not valid.

Consider a similar (LGSIP) problem.

$$\max_{x \geq 0}\ x \qquad \text{s.t.} \quad 2y + x - 1 \leq 1 \quad \text{for all}\ y \in Y(x) := \{y \mid 0 \leq y \leq \frac{1}{2}x\}$$

with the bilevel formulation

$$\max_{x \geq 0}\ x \qquad \text{s.t.} \quad 2y + x - 1 \leq 0\ \ \text{and}\ y \text{ is a solution of}$$
$$\text{Q(x):}\ \ \max_{y}\ 2y \quad \text{s.t.} \quad 0 \leq y \leq \frac{1}{2}x\ .$$

(The condition $x \geq 0$ is added to yield the assumption AL$_3$.) Here, the feasible set $M_{\mathrm{BL}_{\mathrm{LGSIP}}}$ consists of the face $f^1 = \{(x,y) \mid y = \frac{1}{2}x,\ 0 \leq x \leq \frac{1}{2}\}$ of M_{sem}. The solution of BL$_{\mathrm{LGSIP}}$ is attained at the vertex $(\overline{x}, \overline{y}) = (\frac{1}{2}, \frac{1}{2})$. In contrast to the solution of the (LBL) above, here, at the solution $\overline{y}$ of $Q(\overline{x})$ only one lower level constraint $y = \frac{1}{2}x$ is active and $\overline{y}$ is a non-degenerate vertex of $Q(\overline{x})$. The solution $\overline{x} = \frac{1}{2}$ of (GSIP) is a vertex of the feasible set $\mathrm{M}_{\mathrm{LGSIP}} = \mathrm{pr}_x\ f^1 = [0, \frac{1}{2}]$.

We no show that the regularity properties of this example hold generically in (LGSIP). We have to introduce some definitions and facts from genericity theory.

Firstly we define the problem set for (LBL) and BL$_{\mathrm{LGSIP}}$. Let us fix the vector
$s = (n, r, p, m_1, q_1, \ldots, m_r, q_r)$. A problem (LBL) in (14) can be seen as an element from

$$\mathcal{P}_s = \{P = (A_l, B_l, b_l, c_l, d_l,\ l = 0, \ldots, r)\}\ ,$$

where the dimensions of A_l, B_l etc. are defined by s. The set $\mathcal{P}_s$ can be identified with $I\!R^K$, where

$$K := (n+1)p + \sum_{l=1}^{r} n + (p+2)m_l + (n + m_l + 1)q_l\ .$$

For BL$_{\mathrm{LGSIP}}$ in view of $A_l = A, B_l = B, m_l = m, q_l = q$ and $r = p$ we define $s_{\mathrm{GSIP}} = (n, p, m, q)$ and the corresponding set of BL$_{\mathrm{LGSIP}}$ problems

$$P_{s_{\mathrm{GSIP}}} = \{P_{\mathrm{GSIP}} = (A_0, B_0, A, B, b, b_0, c_0)\} \equiv I\!R^{K_{\mathrm{GSIP}}}$$

with $K_{\mathrm{GSIP}} := (n + m + 1)(p + q) + n$.

In the sequel, by a generic subset $\mathcal{V}$ of $I\!R^K$ we mean a set which is open and has a complement $\mathcal{V}^c = I\!R^K \setminus \mathcal{V}$ of measure zero (notation $\mu(\mathcal{V}^c) = 0$). Note that $\mu(\mathcal{V}^c) = 0$ implies that the set $\mathcal{V}$ is dense in $I\!R^K$. For definitions and details in genericity and stratification theory we refer to [3].

The whole genericity analysis can be based on the following general result (see [3] for a proof).

Lemma 4 *Let $h : I\!R^K \to I\!R$ be a polynomial function, $h \not\equiv 0$. Then the solution set $h^{-1}(0) = \{w \in I\!R^K \mid h(w) = 0\}$ is a closed set of measure zero. Equivalently the complement $\mathcal{V} = I\!R^K \setminus h^{-1}(0)$ is a generic set in $I\!R^K$.*

This lemma will be used in a way indicated in the following lemma.

Lemma 5 *Let V_l denote the set of real $(l{\times}l)$-matrices, $V_l = \{A = (a_{ij})_{i,j=1,\ldots,l} \mid a_{ij} \in I\!R\} \equiv I\!R^{l \cdot l}$. Then, the set $V_l^0 = \{A \in V_l \mid \det A = 0\}$ is a closed set of measure zero in $I\!R^{l \cdot l}$. Equivalently the set $V_l^r = V_l \setminus V_l^0$ of regular matrices is generic in $I\!R^{l \cdot l}$.*

Proof. In view of the Laplace expansion $\det A = \sum_{\pi \in \Pi_l} \mathrm{sign}\pi\ a_{1\,\pi(1)} \cdots a_{l\,\pi(l)}$ the mapping $h : I\!R^{l \cdot l} \to I\!R$, $h(A) = \det A$, is a polynomial. Since $h(I) = 1$ we have $h \not\equiv 0$ (I denotes the unit matrix). The result now follows from Lemma 4.

$\square$

In the proofs later on we tacitly make use of the following simple facts:

> If V is a generic subset in $\mathbb{R}^q$, then $\mathbb{R}^s \times V$ is generic in $\mathbb{R}^s \times \mathbb{R}^q$. Let $V_1, \dots, V_r$ be generic subsets of $\mathbb{R}^q$. Then the intersection $V = \cap_{i=1}^r V_i$ is generic in $\mathbb{R}^q$.

We give the first genericity result.

Lemma 6 *The problem set $\mathcal{P}_s$ (or $P_{s_{GSIP}}$) contains a generic subset $\mathcal{V}$ such that for any problem P in $\mathcal{V}$ the following holds.*

(a) All vertices of the semi-feasible set M_{sem} of P are non-degenerate. All local solutions of P are locally unique and occur at vertices of M_{sem}. All local solutions have different object values. In particular, the problem P has a unique global (vertex-) solution.

(b) For any $x \in X$ and l, if $Q_l(x)$ has a solution, then this solution $y_l(x)$ is unique and occurs at a vertex of $Y_l(x)$.

Proof. (a) For $r = 1$, the result is proven in [22, Th.3(a),(c)]. The generalization to the case $r > 1$ is not difficult (we have only to take care of the fact that now the problem matrices have block-structure).

(b) Choose $\overline{x} \in X$ arbitrarily and $l \in \{1, \dots, r\}$. Consider the lower level problem

$$Q_l(\overline{x}): \quad \min_{y_l} \; d_l y_l \qquad \text{st. } B_l\, y_l \leq b_l - A_l \overline{x}, \qquad .$$

Suppose y_l is a solution of $Q_l(\overline{x})$. Then there exist I_l, $I_l \subset \{1, \dots, q_l\}$, $|I_l| \leq m_l$ (by Caratheodory's Theorem), $0 < u_l \in \mathbb{R}^{|I_l|}$ such that

$$u_l^T (B_l)_{I_l} = -d_l^T \;, \qquad (B_l)_j \; y_l = (b_l)_j - (A_l)_j \; \overline{x}, \quad j \in I_l \;. \tag{16}$$

Here $(B_l)_{I_l}$ denotes the sub-matrix of B_l only containing the rows with indices in I_l. Generically, $|I_l| \geq m_l$, i.e. we can assume $|I_l| = m_l$. In fact, if $|I_l| < m_l$ then in view of $u_l^T (B_l)_{I_l} = -d_l^T$ the $(|I_l| + 1) \times (|I_l| + 1)$-matrix (assume for brevity $I_l = \{1, \dots, |I_l|\}$ and we denote the elements of B_l by $(B_l)_{ij}$)

$$\hat{B} := [(B_l)_{ij} \;{}_{|\, \substack{i=1,\dots,|I_l|+1 \\ j=1,\dots,|I_l|}} \;\; \hat{b}] \quad \text{ with } \quad \hat{b} := ((d_l)_1, \dots, (d_l)_{|I_l|+1})^T$$

would satisfy $\det(\hat{B}) = 0$ which can generically be avoided.

Since generically (with $|I_l| = m_l$) the matrix $(B_l)_{I_l}$, is regular, a solution y_l of $Q_l(\overline{x})$ is generically a vertex of the polyhedron $Y_l(\overline{x})$. Since $u_l > 0$ this solution is unique. $\square$

For the analysis of (LGSIP) we define the set

$$P^r_{s_{\mathrm{GSIP}}} \quad = \quad \{P \in P_{s_{\mathrm{GSIP}}} \mid \text{the assumption AL}_3 \text{ hold} \} \;.$$

It is not difficult to show that the problem set $P^r_{s_{\mathrm{GSIP}}}$ is open in $\mathbb{R}^{K_{\mathrm{GSIP}}}$.

The next theorem describes the difference between general (BL) and (GSIP) problems (see also Theorem 2). It shows that for $\mathrm{BL}_{\mathrm{LGSIP}}$, in the generic case, n upper level constraints must be active at a solution $\overline{x}$ of (LGSIP) and that the regularity assumption $\mathrm{A2}_{\mathrm{LGSIP}}$ in Section 3 holds.

Theorem 3 *The problem set $P^r_{s_{GSIP}}$ contains a generic subset $\mathcal{V}$ such that for any BL_{LGSIP} problem P in $\mathcal{V}$ the following holds.*

If P has a local solution $(\overline{x}, \overline{y})$ (vertex solution see Lemma 6) then, precisely n upper-level constraints are active, i.e. there exist n indices in the index set $J = \{1, \ldots, p\}$, say $l = 1, \ldots, n$ such that with the solutions $\overline{y}_l$ of $Q_l(\overline{x})$ we have

$$(a_l^0)^T \overline{x} + (b_l^0)^T \overline{y}_l - \beta_l^0 = 0, \qquad l = 1, \ldots, n.$$

These solutions $\overline{y}_l$ are nondegenerate-vertices (i.e. (LICQ) holds). Moreover the local solution $\overline{x}$ of (LGSIP) is attained at a vertex of $pr_x\ M_{sem}$.

In particular, if $p < n$ holds, then generically BL_{LGSIP} and the corresponding (LGSIP) does not have a solution, i.e. the problem is unbounded.

PROOF. First we show that generically at least n constraints must be active at the local solution $(\overline{x}, \overline{y})$ of $\mathrm{BL}_{\mathrm{LGSIP}}$ or the local solution $\overline{x}$ of (LGSIP).

Suppose that $k < n$ points are active at $\overline{x}$. This means there are $k < n$ vertex solutions $\overline{y}_l$ of $Q_l(\overline{x})$, say $l = 1, \ldots, k$, active, i.e. we have

$$\begin{array}{ll} (a_l^0)^T \overline{x} + (b_l^0)^T \overline{y}_l - \beta_l^0 = 0, & l = 1, \ldots, k \\ (a_l^0)^T \overline{x} + (b_l^0)^T \overline{y}_l - \beta_l^0 > 0, & l = k+1, \ldots, p. \end{array} \tag{17}$$

We show that then generically AL_3 is violated (see the definition of $\mathcal{P}^r_{s_{\mathrm{GSIP}}}$).

With the value function $v_l(x)$ of $Q_l(x)$, the local solution $\overline{x}$ of (LGSIP) must be a local minimizer of the problem,

$$\min \quad c_0^T x \quad \text{s.t.} \quad v_l(x) \geq 0, \quad l = 1, \ldots, k. \tag{18}$$

Consider the optimality conditions (Kuhn-Tucker- and Complementary conditions) for the solutions $\overline{y}_l$ of $Q_l(\overline{x})$,

$$\begin{array}{rcl} B^T \lambda_l - b_l^0 & = & 0 \\ \lambda_l^T (A\overline{x} + By_l - b) & = & 0. \end{array} \tag{19}$$

Generically the solutions $\overline{y}_l$ of $Q_l(\overline{x})$ are unique (cf. Lemma 6(b)). Let D_l denote the set of Lagrange multipliers λ_l satisfying (19). Suppose AL_3 is satisfied (Slater condition). Then by a well-known theorem (see e.g. [13]) the value functions v_l are directionally differentiable and with the Lagrange function $\mathcal{L}^l(x, y, \lambda) := (a_l^0 x + b_l^0 y - \beta_l^0)^T - \lambda_l^T (Ax + By - b)$ the directional derivative $Dv_l(\overline{x}, d) := \lim_{t \downarrow 0} \frac{v_l(\overline{x}+td) - v_l(\overline{x})}{t}$ is given by

$$Dv_l(\overline{x}, d) = \max_{\lambda_l \in D_l} D_x \mathcal{L}^l(\overline{x}, y_l, \lambda_l) d = \max_{\lambda_l \in D_l} (a_l^0 - A^T \lambda_l)^T d.$$

Choosing one $\overline{\lambda}_l \in D_l$ arbitrarily (e.g. the multiplier $\overline{\lambda}_l = -(B_{I_l}^T)^{-1} b_l^0$; see (16) in the proof of Lemma 6(b)). then obviously

$$(a_l^0 - A^T \overline{\lambda}_l)^T d \leq Dv_l(\overline{x}, d).$$

Generically we can assume that the vectors c_0, $(a_l^0 - A^T\overline{\lambda}_l)$, $l = 1, \ldots k$, $(k < n)$ are linearly independent. Thus, there is a solution d of

$$c_0^T d = -1 \ , \quad (a_l^0 - A^T\overline{\lambda}_l)^T d = 1, \quad l = 1, \ldots k.$$

This implies that for $x_t = \overline{x} + td$, $t > 0$ small, we have $c_0^T x_t < c_0^T \overline{x}$ and in view of $v_l(\overline{x}) = 0$

$$v_l(x_t) = v_l(\overline{x}) + tDv_l(\overline{x}, d) + o(t) \geq t(a_l^0 - A^T\overline{\lambda}_l)^T d + o(t) > 0,$$

$l = 1, \ldots k$, contradicting the fact that $\overline{x}$ is a local solution of (18).

We now show that $k = n$ must be valid. In view of Lemma 6(a) generically the local solution $\overline{z} := (\overline{x}, \overline{y}_1, \ldots, \overline{y}_k)$ is a non-degenerate vertex of the semi-feasible set $\mathrm{M}_{\mathrm{sem}} \subset I\!R^{n+km}$ of the problem

$$\begin{array}{lll} (\mathrm{LBL}_k): & \min\ c_0^T x & \text{s.t. } G_l(x,y) := a_l^0 x + b_l^0 y - \beta_l^0 \geq 0,\ l = 1, \ldots, k, \\ & \text{and for } l = 1, \ldots, k, & y_l \text{ is a solution of} \\ & Q_l(x):\ \min\limits_{y_l}\ (a_l^0)^T x + (b_l^0)^T y_l & \text{s.t.} \quad Ax + By_l - b \geq 0. \end{array}$$

Thus $n + km$ constraints must be active in $\overline{z}$. So for the number N_a of active constraints we must have

$$N_a = n + km = k + \sum_{l=1}^{k} |I_l| \ .$$

Using $k \geq n$ and $|I_l| \geq m$ (y_l are vertices of $Q_l(\overline{x})$) we find $k \leq n$, $|I_l| \leq m$, i.e.

$$k = n \quad \text{and} \quad |I_l| = m,\ l = 1, \ldots, k \ .$$

In view of Lemma 5 generically the $(m \times m)$-matrices B_{I_l} are regular. Thus (LICQ) is fulfilled generically.

We now show that generically the solution $\overline{x}$ is attained at a vertex of $\mathrm{pr}_x\, M_{\mathrm{sem}}$. The Kuhn-Tucker condition for the solutions $\overline{y}_l$ of $Q_l(\overline{x})$ read,

$$B_{I_l}^T \lambda_l = b_l^0 \ , \qquad \lambda_l \geq 0.$$

Generically we must have $\lambda_l > 0$ (see the proof of Lemma 6(b)). By standard sensitivity analysis it follows that locally near $\overline{x}$ the solutions $y_l(x)$ of $Q_l(x)$, $l = 1, \ldots, n$, $(k = n)$ with $y_l(\overline{x}) = \overline{y}_l$ are given by

$$A_{I_l} x + B_{I_l} y_l(x) - b_{I_l} = 0 \quad \text{or} \quad y_l(x) = B_{I_l}^{-1}(b_{I_l} - A_{I_l} x).$$

By substituting this solution into (17) we see that a point x near $\overline{x}$ is feasible if and only if

$$\big((a_l^0)^T - (b_l^0)^T B_{I_l}^{-1} A_{I_l}\big)x - \big(\beta_l^0 - (b_l^0)^T B_{I_l}^{-1} b_{I_l}\big) \geq 0, \quad l = 1, \ldots, n. \tag{20}$$

Generically the vectors $(a_l^0)^T - (b_l^0)^T B_{I_l}^{-1} b_{I_l}$, $l = 1, \ldots n$, must be linearly independent. Thus the n inequalities in (20) define the vertex $\overline{x}$ of the polyhedron $\mathrm{pr}_x\, M_{\mathrm{sem}}$.

$\square$

5 Algorithm for linear GSIP

In the preceding sections we have seen that (GSIP) can be regarded as a special instance of a bilevel problem. Because of the special structure of (GSIP) not all approaches for semi-infinite programming are appropriate for general (BL) problems (for example the reduction approach). However any method for bilevel problems can be used to solve the bilevel formulation of (GSIP) problems. We refer to [16] for a survey of methods for solving (BL) (for the case $r = 1$).

Here we only consider the linear case and briefly outline the generalization to $r > 1$ (r lower level players) of an algorithm due to Bard and Moore (cf. [1]) which is based on a so-called Kuhn-Tucker approach. With this method, also (LGSIP) can be solved.

Consider the necessary and sufficient Kuhn-Tucker optimality conditions for a solution y_l of the linear program $Q_l(x)$: Introducing slack variables $v_l \in R^{q_l}$ with the Lagrange multiplier vectors $\lambda_l \in I\!R^{q_l}$ these conditions are,

$$\begin{aligned} A_l x + B_l y_l - b_l - v_l &= 0 \\ \lambda_l^T B_l - d_l &= 0 \\ \lambda_l \geq 0, \quad v_l &\geq 0 \\ \lambda_l^T v_l &= 0 \qquad (\textit{complementarity conditions}) \end{aligned}$$

It follows that $(\overline{x}, \overline{y})$ is a solution of (LBL) (cf. (14)) if and only if with slack vectors v_l and multipliers λ_l the point $(\overline{x}, \overline{y})$ solves the optimization problem

$$\begin{aligned} \min_{x,y} \ c_0^T x + d_0^T y \quad \text{s.t.} \qquad\qquad & A_0 x + B_0 y - b_0 \geq 0 \\ \text{and for } l = 1, \ldots, r \quad & A_l x + B_l y_l - b_l - v_l = 0 \\ & \lambda_l^T B_l - d_l = 0 \\ & \lambda_l \geq 0, \quad v_l \geq 0 \\ & \lambda_l^T v_l = 0 \end{aligned} \tag{21}$$

Apart from the complementarity conditions $\lambda_l^T v_l = 0$ this problem is linear.

A branch and bound method to solve (21) is as follows. We define $q := q_1 + \ldots + q_r$, the vectors $\Lambda := (\lambda_1, \ldots, \lambda_r)$, $V := (v_1, \ldots, v_r)$ in $I\!R^q$ and the index set $K := \{1, \ldots, q\}$. In view of $\Lambda,\ V \geq 0$, the complementarity condition $\Lambda^T V$ in (21) is equivalent with $\Lambda_i V_i = 0$ for all $i \in K$. For given index sets $K^+,\ K^- \subset K,\ K^+ \cap K^- = \emptyset$ we define the sets

$$\Lambda(K^+) = \{\Lambda \geq 0 \mid \Lambda_i = 0,\ i \in K^+\}, \quad V(K^-) = \{V \geq 0 \mid V_i = 0,\ i \in K^-\}.$$

For any pair $\Lambda,\ V$ with $\Lambda \in \Lambda(K^+)$, $V \in V(K^-)$ let LBL(K^+, K^-) denote the problem obtained by replacing in the (LBL) problem (21) the complementarity condition $\Lambda V = 0$ by the conditions $\Lambda \in \Lambda(K^+)$, $V \in V(K^-)$. The problems LBL(K^+, K^-) are linear programs and for the right choice of K^+, K^- the solution of LBL(K^+, K^-) coincides with the solution of (LBL). The idea of the Bard/Moore algorithm is to examine in a branch and bound search all possible choices of K^+, K^- (see [1] for further details):

The algorithm starts with $K^+ = K^- = \emptyset$. Obviously, the value of LBL$(\emptyset, \emptyset)$ (a relaxation of (LBL)) gives a lower bound for the value of (LBL).

BRANCH AND BOUND ALGORITHM:

start: Put $k=0$, $K_0^+=\emptyset$, $K_0^-=\emptyset$, val= ∞.

step $k\to k+1$: Given K_k^+, K_k^-, try to calculate a solution x^k, y^k, Λ^k, V^k of LBL(K_k^+, K_k^-) with value val$_k$.

1. If LBL(K_k^+, K_k^-) is infeasible or if val$_k \geq$ val goto 3.
   If $\Lambda_i V_i = 0$ for all $i \in K$ put val = val$_k$, goto 3.

2. *(Branching w.r.t. Λ)* Select an index $i_k \in K \setminus K_k^+$ such that $\Lambda_{i_k} V_{i_k} > 0$, put $K_{k+1}^+ = K_k^+ \cup \{i_k\}$, $K_{k+1}^- = K_k^-$, goto 4.

3. Perform backtracking (see [1]) for details), goto 4.

4. $k+1 \to k$.

With this methods problems of size up to $n = m = 100$ (for $r = 1$) can be solved (cf. [1] , [4] for numerical experiments.)

References

[1] Bard J.F., Moore J.T., *A branch and bound algorithm for the bilevel programming problem*, SIAM J. Sci. Stat. Comput. Vol.11, No. 2, 281-292, (1990).

[2] Falk E.J., Liu J., *On bilevel programming, Part I: general nonlinear cases*, Mathematical Programming 70, 47-72, (1995).

[3] Gibson C.G., Wirthmüller K., Du Plessis A.A., Looijenga E.J.N., *Topological stability of smooth mappings*, Lecture Notes in Math., vol. 552, Springer-Verlag, Berlin (1976).

[4] Hamming R., *On the bilevel programming problem*, Thesis in the Faculty of Mathematics, University of Twente (1998).

[5] R. Hettich, H.Th. Jongen, *Semi-infinite programming: conditions of optimality and applications* in J. Stoer, ed.: Optimization Techniques 2, 1-11, Springer, Berlin-Heidelberg-New York (1978).

[6] R. Hettich, G. Still, *Semi-infinite programming models in Robotics*, in 'Parametric Optimization and Related Topics II', Guddat et al. (eds.), Akademie Verlag, Berlin (1991).

[7] Hettich R., Kortanek K., Semi-infinite programming: Theory, methods and applications, SIAM Review, vol 35, No.3, 380-429, (1993).

[8] R. Hettich, G. Still, *Second order optimality conditions for generalized semi-infinite programming problems*, Optimization Vol. 34, pp. 195-211, (1995).

[9] Hoffmann A. and Reinhardt R., *On reverse Chebyshev approximation problems*, Technical University of Illmenau, Preprint No. M08/94, (1994).

[10] H.Th. Jongen, J.-J. Rückmann, O. Stein, *Generalized semi-infinite optimization: a first order optimality condition and examples,* Mathematical Programming, Vol. 83 (1998), 145-158.

[11] H. Th Jongen, G.. Zwier, *On the local structure of the feasible set in semi-infinite optimization*, in: Brosowski, Deutsch (eds.): Int. Ser. Num. Math. 72, Birkhäuser Verlag, Basel, 185-202, (1984).

[12] Kaplan A. and Tichatschke R., *On a class of terminal variational problems*, in 'Parametric Optimization and Related Topics IV', Guddat et al. (eds.), Peter Lang Verlag, Frankfurt a.M., (1997).

[13] Rockafellar R.T., *Directional differentiability of the optimal value function in nonlinear programming problem*, Math. Programming Study 21, 213-226, (1984).

[14] J.-J. Rückmann, A. Shapiro, *First-order optimality conditions in generalized semi-infinite programming,* Journal of Optimization Theory and Applications, Vol. 101 (1999), 677-691.

[15] Rückmann J.-J., Stein O., *On linear and linearized generalized semi-infinite problems*, Annals of Operations Research, to appear.

[16] Shimizu K., Ishizuka Y., Bard J., *Nondifferentiable and two-level mathematical programming*, Kluwer Academic Publishers, Boston, (1997).

[17] Stein O., *On level sets of marginal functions*, Optimization, Vol. 48 (2000), 43-67.

[18] O. Stein, *First order optimality conditions for degenerate index sets in generalized semi-infinite programming,* Preprint No. 91, Lehrstuhl C für Mathematik, RWTH Aachen, 2000.

[19] O. Stein, G. Still, *On optimality conditions for generalized semi-infinite programming problems,* Journal of Optimization Theory and Applications, Vol. 104, (2000), 443-458.

[20] Still G., *Generalized semi-infinite programming: Theory and methods*, European Journal of Operational Research 119, 301-313, (1999).

[21] Still G., *Generalized semi-infinite programming: Numerical aspects*, Optimization, Vol. 49, 223-242, (2001).

[22] Still G., *Linear bilevel problems: Genericity results and an efficient method for computing local minima*, to appear in Mathematical Methods of Operations Research.

[23] G.-W. Weber, *Generalized Semi-Infinite Optimization and Related Topics,* Habilitation Thesis, Darmstadt University of Technology, (1999).