

---

Faculty of Mathematical Sciences

University of Twente

University for Technical and Social Sciences

---

---

P.O. Box 217  
7500 AE Enschede  
The Netherlands

Phone: +31-53-4893400

Fax: +31-53-4893114

Email: [memo@math.utwente.nl](mailto:memo@math.utwente.nl)

---

MEMORANDUM No. 1538

Linear bilevel problems: Genericity results  
and an efficient method for  
computing local minima

T.J. FREDERIKS AND G.J. STILL

AUGUST 2000

ISSN 0169-2690

# **Linear Bilevel Problems: Genericity results and an efficient method for computing local minima**

Theo Frederiks and Georg Still

## **Abstract**

The paper is concerned with linear bilevel problems. These nonconvex problems are known to be NP-complete. So, no efficient method for solving the global bilevel problem can be expected. In this paper we give a genericity analysis of linear bilevel problems and present a new algorithm for computing efficiently local minimizers. The method is based on the given structural analysis and combines ideas of the Simplex method with projected gradient steps.

**Keywords:** linear bilevel programming, genericity results, numerical methods

**Mathematical Subject Classification 1991:** 90C26

# 1 Introduction

This paper deals with linear bilevel problems of the form

$$\begin{aligned} \text{(LBL): } \min_{x,y} a^1 x + b^1 y \quad \text{s.t.} \quad & A^1 x + B^1 y \leq c^1 \\ & \text{and } y \text{ is a solution of} \\ Q(x): \min_y a^2 x + b^2 y \quad \text{s.t.} \quad & A^2 x + B^2 y \leq c^2 . \end{aligned}$$

with given matrices  $A^1 \in \mathbb{R}^{k_1 \cdot n}$ ,  $B^1 \in \mathbb{R}^{k_1 \cdot m}$ ,  $A^2 \in \mathbb{R}^{k_2 \cdot n}$ ,  $B^2 \in \mathbb{R}^{k_2 \cdot m}$ , vectors  $a^1, a^2 \in \mathbb{R}^n$ ,  $b^1, b^2 \in \mathbb{R}^m$ ,  $c^1 \in \mathbb{R}^{k_1}$ ,  $c^2 \in \mathbb{R}^{k_2}$  and variables  $x \in \mathbb{R}^n$ ,  $y \in \mathbb{R}^m$ .

Throughout the paper we omit the transposed sign in some expressions. For example  $a x$  denotes the inner product  $a^T x$  in  $\mathbb{R}^n$  and  $u B^2$  the left multiplication of the matrix  $B^2$  by the vector  $u \in \mathbb{R}^{k_2}$ .

The linear bilevel problem can be considered as a game between an upper level player and a lower level player which for given  $x \in \mathbb{R}^n$  has to solve the lower level problem  $Q(x)$ . The constraints  $A^1 x + B^1 y \leq c^1$ , resp.  $A^2 x + B^2 y \leq c^2$  are called upper- resp. lower level constraints. For a theoretical and practical introduction into bilevel programming the reader is referred to [8]. LBL-problems are non-convex problems which are NP-complete (cf. [2]).

The aim of the present paper is twofold. Firstly we develop genericity results for linear bilevel problems. By genericity results we roughly speaking mean statements which assert that for *almost all* LBL-problems certain nice properties are fulfilled. Secondly, since the LBL-problem is NP-complete, it could be of interest to develop an algorithm which is able to compute at least a local minimizer efficiently. We present such an algorithm for the LBL-problems without upper level constraints. The algorithm is based on the genericity analysis and combines ideas of the Simplex method in linear programming with projected gradient steps.

The paper is organized as follows. In the second section we give an overview on the structure of the LBL-problems. Section 3 is concerned with genericity results. In Section 4 we introduce our new algorithm for computing local minimizer of LBL and discuss complexity questions. In the last section we report on numerical experiments by comparing the performance of our local minimization algorithm with a Kuhn-Tucker method of Bard/Moore for solving the global LBL-problem.

## 2 Preliminary results

We firstly introduce some notation. With  $x \in \mathbb{R}^n$ ,  $y \in \mathbb{R}^m$  we define,

$M_2(x)$	$= \{y \mid A^2x + B^2y \leq c^2\}$	feasible set of $Q(x)$
$M_2$	$= \{(x, y) \mid A^2x + B^2y \leq c^2\}$	the graph of $M_2(x)$
$X_2$	$= \{x \mid M_2(x) \neq \emptyset\}$	projection of $M_2$ onto $\mathbb{R}^n$
$S(x)$	$= \{y \mid y \text{ solves } Q(x)\}$	set of solutions of $Q(x)$
$S$	$= \{(x, y) \mid y \in S(x)\}$	the graph of $S(x)$
$v(x)$	$= a^1x + b^1y$ with $y \in S(x)$	value function of $Q(x)$
$M_1$	$= \{(x, y) \mid A^1x + B^1y \leq c^1\}$	constraints of the upper level
$M_{\text{sem}}$	$= M_1 \cap M_2$	the semi-feasible set
$X$	$= \{x \mid (x, y) \in M_{\text{sem}}, \text{ for some } y\}$	projection of $M_{\text{sem}}$ onto $\mathbb{R}^n$
$M$	$= \{(x, y) \mid (x, y) \in M_{\text{sem}}, y \in S(x)\}$	feasible set of LBL

**Remark 1** *The polyhedra  $M_2$ ,  $M_{\text{sem}}$  and their projections  $X_2$ ,  $X$  are closed sets.*

Throughout the paper the following abbreviations are used: We put  $N = n + m$ ,  $z = (x, y) \in \mathbb{R}^N$  and

$$A = \begin{pmatrix} A^1 \\ A^2 \end{pmatrix}, B = \begin{pmatrix} B^1 \\ B^2 \end{pmatrix}, c = \begin{pmatrix} c^1 \\ c^2 \end{pmatrix}, d^1 = \begin{pmatrix} a^1 \\ b^1 \end{pmatrix}, d^2 = \begin{pmatrix} a^2 \\ b^2 \end{pmatrix}, C = [A \ B].$$

We define  $J^1 = \{1, \dots, k_1\}$ ,  $J^2 = \{k_1 + 1, \dots, k_1 + k_2\}$ ,  $J = J^1 \cup J^2$  and denote by  $C_j, A_j, B_j$  the  $j^{\text{th}}$  rows of  $C, A, B$ ,  $j \in J$ . Then, the semi-feasible polyhedron can be written as

$$M_{\text{sem}} = \{z \in \mathbb{R}^N \mid C_j z \leq c_j, j \in J\}.$$

For a given index set  $J_0 \subset J$  let  $C_{J_0}$  be the sub-matrix of  $C$  only containing the rows  $C_j$  with indices  $j \in J_0$ .  $A_{J_0}, B_{J_0}, c_{J_0}$  are defined accordingly.

A subset  $f_0 \subset M_{\text{sem}}$  is called a face of the polyhedron  $M_{\text{sem}}$  if there exists an index set  $J_0 \subset J$  such that

$$f_0 = f(J_0) := \{z \in M_{\text{sem}} \mid C_j z = c_j, j \in J_0\}. \quad (1)$$

Let be given  $J_0 \subset J$  and the related face  $f_0 = f(J_0)$  of  $M_{\text{sem}}$ . We say that  $f_0$  has dimension  $d$ ,  $0 \leq d \leq N$ , if there exists an element  $z_0 \in f_0$  such that

$$C_j z_0 < c_j, \quad j \in J \setminus J_0 \quad \text{and} \quad \dim \text{span}\{C_j, j \in J_0\} = N - d. \quad (2)$$

The  $d$ -dimensional face  $f_0 = f(J_0)$  is said to be non-degenerate if  $|J_0| = N - d$ . A vertex of  $M_{\text{sem}}$  is a face of dimension 0. If  $\text{int } M_{\text{sem}} \neq \emptyset$  then  $M_{\text{sem}}$  is a face of dimension  $N$ .

The following assumptions will play an important role.

**A1:** For all  $x \in X_2$  the solution set  $S(x)$  is bounded (and thus compact).

**A2:** The polyhedron  $M_{\text{sem}}$  is bounded (thus compact).

The following theorem contains the main results on the structure of linear bilevel problems.

**Theorem 1** *Let be given an LBL-problem. Then the following holds.*

(a) *The set  $X_2 \subset \mathbb{R}^n$  is a polyhedron (thus closed and convex).*

(b) *The feasible set  $M$  consist of a union  $f_1 \cup f_2 \cup \dots \cup f_l$  of faces of the polyhedron  $M_{\text{sem}}$ . In particular,  $M$  is a closed set in  $\mathbb{R}^{n+m}$ .*

*If moreover the assumptions A1 and A2 hold we have:*

(c) *The value function  $v(x)$  of  $Q(x)$  is convex and Lipschitz continuous on  $X_2$ .*

(d) *A global solution of LBL occurs at a vertex of  $M_{\text{sem}}$ .*

(e) *For any local minimizer  $(x_k, y_k)$  of LBL on a face  $f_k$  there exist a local minimizer  $(\bar{x}_k, \bar{y}_k) \in f_k$  which is a vertex of  $M_{\text{sem}}$  with the same value  $a^1 x_k + b^1 y_k = a^1 \bar{x}_k + b^1 \bar{y}_k$ .*

**Proof.** For a proof of (a)-(d) we refer to [8] (or to [9] for a slightly more general bilevel problem).

(e) Let  $(x_k, y_k) \in f_k$  be a local minimizer of LBL ( $f_k \subset M$ ,  $f_k$  a face of  $M_{\text{sem}}$ ). Then  $(x_k, y_k)$  is a global solution of the linear program

$$\min_{x,y} a^1 x + b^1 y \quad \text{st.} \quad (x, y) \in f_k .$$

Since  $f_k$  is a bounded polyhedron, the solution of this problem occurs at a vertex  $(\bar{x}_k, \bar{y}_k)$  of  $f_k$ . Since a vertex of  $f_k$  is a vertex of  $M_{\text{sem}}$  the proof is completed.  $\square$

When the BL-problem has upper level constraints, then the feasible set need not be connected in general. As an example consider the LBL with  $x, y \in \mathbb{R}$ :

$$\begin{aligned} \min x + y \quad \text{st.} \quad & y \geq 1 \text{ and } y \text{ solves} \\ & Q(x) : \min y \quad \text{st.} \quad y - x \geq 0, \quad y + x \geq 0 . \end{aligned}$$

The feasible set consists of the set  $\{(x, y) \mid y = |x|, y \geq 1\}$  which obviously is not connected. If the upper level constraint  $y \geq 1$  is omitted then the feasible set  $\{(x, y) \mid y = |x|\}$  becomes connected.

More generally the following holds.

**Theorem 2** *Let for all  $x \in X_2$  the set  $S(x)$  be non-empty (fulfilled if A2 holds). Then if no upper level constraints are present, the feasible set  $M$  of the LBL-problem is path-wise connected.*

**Proof.** We have to show that for any two points  $(x^1, y^1), (x^2, y^2)$  in  $M$  there is a path in  $M$  from  $(x^1, y^1)$  to  $(x^2, y^2)$ . Suppose that this is not the case. Consider the maximal path-connected component  $M^1$  in  $M$  containing  $(x^1, y^1)$  and suppose that  $M^1$  does not coincide with  $M$ . Since  $M$  consists of the union of say  $K$  faces and any face is convex (thus path-wise connected),  $M^1$  consists of a number of these faces say  $f_1, \dots, f_{K_1}$ ,  $K_1 < K$ , i.e.  $M^1 = \cup_{k=1}^{K_1} f_k$ . It now follows that

$$\left(\cup_{k=1}^{K_1} f_k\right) \cap f_j = \emptyset, \quad j = K_1 + 1, \dots, K.$$

In fact,  $\bar{x} \in f_j \cap M^1$  implies that  $f_j \in M^1$ . Consequently, with the closed set  $M^2 := \cup_{j=K_1+1}^K f_j$ , we must have

$$M^1 \cup M^2 = M, \quad M^1 \cap M^2 = \emptyset.$$

Let  $X^i, i \in \{1, 2\}$ , denote the projections of  $M^i$  onto  $\mathbb{R}^n$ . We have  $X = X^1 \cup X^2$  and  $X^1, X^2$  are closed (projections of polyhedra are polyhedra). Since there are no upper level constraints it follows  $M_{\text{sem}} = M_2$  and  $X = X_2$ . Moreover for any  $x \in X$  we have  $S(x) \neq \emptyset$ . Thus the projection onto  $\mathbb{R}^n$  of  $M$  coincides with  $X$  and  $X$  is a polyhedron (in particular convex). Consider the line segment  $L$  between the points  $x^1 \in X^1$  and  $x^2 \in X^2$ . Since  $L \subset X = X^1 \cup X^2$  and  $X^1, X^2$  are closed, there must exist a point  $\bar{x} \in L$  belonging to both sets  $X^1$  and  $X^2$ . Consequently, there are points  $(\bar{x}, \bar{y}^1) \in M^1$  and  $(\bar{x}, \bar{y}^2) \in M^2$ . Since  $S(\bar{x})$  is convex, the whole line segment between  $(\bar{x}, \bar{y}^1)$  and  $(\bar{x}, \bar{y}^2)$  lies in  $M$ . This contradicts the assumption that  $M^1$  is not path-wise connected with  $M^2$ .  $\square$

In Section 4 we develop a new algorithm for computing local solutions of linear bilevel problems without upper level constraints. This raises the question whether it is possible to avoid upper level constraints in the LBL-model. If there are upper level constraints then the the players could change the model by passing the upper level constraints to the lower level. We briefly discuss such a strategy and begin with an illustrative example.

**Example 1.** Consider

$$\begin{aligned} \max F(x, y) := x + y \quad & \text{s.t. } x + 2y - 8 \leq 0 \quad \text{and } y \text{ is a solution of} \\ Q(x): \max f(x, y) := x + y \quad & \text{s.t. } 0 \leq y \leq 4 \end{aligned}$$

Here,  $S(x) = \{y(x) = 4\}$  for all  $x \in \mathbb{R}$ . The feasible set is given by  $M = \{(x, 4) \mid x \leq 0\}$ . The optimal solution is  $(\bar{x}, \bar{y}) = (0, 4)$  with values  $F(\bar{x}, \bar{y}) = f(\bar{x}, \bar{y}) = 4$ . Consider now the bilevel problem obtained by passing the upper level constraints  $x + 2y - 8 \leq 0$  to the lower level  $Q(x)$ . Then the solution set becomes  $S(x) = \{4\}$  for  $x \leq 0$ ,  $S(x) = \{-\frac{x}{2} + 4\}$  for  $x \in [0, 8]$ ,  $S(x) = \emptyset$  otherwise, and for the feasible set we find  $M_0 = M \cup \{(x, -\frac{x}{2} + 4) \mid 0 \leq x \leq 8\}$  with optimal solution  $(\hat{x}, \hat{y}) = (8, 0)$  and values  $F(\hat{x}, \hat{y}) = f(\hat{x}, \hat{y}) = 8$ .  $\square$

The following lemma shows that this strategy, to pass the upper level constraints to the lower level, may change the model but it can only be an advantage for the upper level player. For the lower level player it can lead to a better but also to an inferior object value depending whether his object is 'similar' or 'opposite' to the upper level object.

Let  $\text{LBL}_0$  be the bilevel problem obtained from LBL by passing the constraints  $A^1x + B^1y \leq c^1$  to the lower level and let  $M_0$  be the feasible set of  $\text{LBL}_0$ .

**Lemma 1** *For the the feasible sets  $M_0$  and  $M$  of  $\text{LBL}_0$  and LBL we have*

$$M \subset M_0.$$

*In particular, for the corresponding minimal values the inequality  $v_0 \leq v$  holds.*

**Proof.** Let  $\bar{x} \in X$  be given. If  $S(\bar{x}) \cap \{(\bar{x}, y) \mid A^1\bar{x} + B^1y \leq c^1\} \neq \emptyset$  then  $S_0(\bar{x}) = S(\bar{x}) \cap \{(\bar{x}, y) \mid A^1\bar{x} + B^1y \leq c^1\}$ . If  $S(\bar{x}) \cap \{(\bar{x}, y) \mid A^1\bar{x} + B^1y \leq c^1\} = \emptyset$  then there is no feasible point  $(\bar{x}, y)$  for LBL. This implies  $M \subset M_0$ .  $\square$

### 3 Genericity results for LBL

In this section we study the structure of the feasible set and the set of local minimizers of LBL from a generic viewpoint.

Throughout the paper, by a generic subset  $G$  of  $\mathbb{R}^K$  we mean a set which is open and has a complement  $G^c = \mathbb{R}^K \setminus G$  of measure zero (notation  $\mu(G^c) = 0$ ). Note that  $\mu(G^c) = 0$  implies that the set  $G$  is dense in  $\mathbb{R}^K$ . For details on genericity and stratification theory we refer to [4].

Our genericity analysis will be based on the following 'non-trivial' result (see [4]).

**Lemma 2** *Let  $p : \mathbb{R}^K \rightarrow \mathbb{R}$  be a polynomial function,  $p \not\equiv 0$ . Then, the solution set  $p^{-1}(0) = \{w \in \mathbb{R}^K \mid p(w) = 0\}$  is a closed set of measure zero. Equivalently the complement  $G = \mathbb{R}^K \setminus p^{-1}(0)$  is a generic set in  $\mathbb{R}^K$ .*

This result will be used repeatedly in a way illustrated by the following lemma.

**Lemma 3** *Let  $V_l$  denote the set of real  $(l \times l)$ -matrices,  $V_l = \{A = (a_{ij})_{i,j=1,\dots,l} \mid a_{ij} \in \mathbb{R}\} \cong \mathbb{R}^{l \cdot l}$ . Then, the set  $V_l^0 = \{A \in V_l \mid \det A = 0\}$  is a closed set of measure zero in  $\mathbb{R}^{l \cdot l}$ . Equivalently the set  $V_l^r = V_l \setminus V_l^0$  of regular matrices is generic in  $\mathbb{R}^{l \cdot l}$ .*

**Proof.** In view of the Laplace expansion,  $\det A = \sum_{\pi \in \Pi_l} \text{sign} \pi a_{1 \pi(1)} \cdots a_{l \pi(l)}$ , the mapping  $p : \mathbb{R}^{l \cdot l} \rightarrow \mathbb{R}$ ,  $p(A) = \det A$ , is polynomial. Since  $p(I) = 1$  we have  $p \not\equiv 0$  ( $I$  denotes the unit matrix). The result now follows from Lemma 2.  $\square$

**Remark 2** *In the proof of Theorem 3 later on we implicitly make use of the following elementary facts:*

*Let  $V$  be 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  $V = \bigcap_{i=1}^r V_i$  is generic in  $\mathbb{R}^q$ .*

It is well-known that for common linear programs generically all vertices of the feasible set are non-degenerate. This result gives a theoretical explanation why the simplex method works well in general without a special rule for avoiding cycling. In the following we generalize such a genericity result to linear bilevel problems.

Firstly we introduce a set which formally describes a linear bilevel problem as a point in  $\mathbb{R}^K$ . Let  $p = (n, m, k_1, k_2)$  be fixed ( $p$  gives the 'size' of the LBL). A problem LBL can be seen as an element from the set

$$P_p = \{P = (A^1, A^2, B^1, B^2, c^1, c^2, a^1, a^2, b^1, b^2)\}$$

with arrays  $A^1, \dots, b^2$  as defined in Section 1. The set  $P_p$  can be identified with  $\mathbb{R}^K$ ,

$$P_p \equiv \mathbb{R}^K, \quad K := (k_1 + k_2 + 2)(n + m) + k_1 + k_2.$$

The following theorem contains the main genericity results.

**Theorem 3** *The problem set  $P_p \equiv \mathbb{R}^K$  contains a generic subset  $V$  such that for any LBL-problem  $P$  in  $V$  the following holds.*

- (a) *All faces of the semi-feasible polyhedron  $M_{sem}$  of problem  $P$  are non-degenerate faces. In particular all vertices of  $M_{sem}$  are non-degenerate.*
- (b) *For any  $x \in X$ , if  $Q(x)$  has a solution, then there is a unique solution  $y(x)$  of  $Q(x)$  attained at a vertex of  $M_2(x)$ .*
- (c) *All local minimizers  $z_v = (x_v, y_v)$ ,  $v = 1, \dots, q$ , of  $P$  are locally unique minimizers and (non-degenerate) vertices of  $M_{sem}$ . All values  $v_v = d^1 z_v$ ,  $v = 1, \dots, q$ , are different. In particular,  $P$  has a unique global minimizer.*
- (d) *The feasible set  $M$  of  $P$  consist of a union  $f_1 \cup f_2 \cup \dots \cup f_l$  of non-degenerate faces  $f_i$  of  $M_{sem}$  of dimension  $n$ .*

**Proof.** (a) Suppose  $f_0 = f(J_0)$  is a face of  $M_{sem}$  of dimension  $d$ ,  $0 \leq d \leq N$ , i.e. with an index  $J_0 \subset J$  and a point  $z_0 \in f_0$  we have

$$C_j z_0 < c_j, \quad j \in J \setminus J_0 \quad \text{and} \quad \dim \text{span}\{C_j, j \in J_0\} = N - d. \quad (3)$$

We now show that generically the face is nondegenerate, i.e.  $|J_0| = N - d$ .

Case  $d = 0$ : Then  $f_0 = \{z_0\}$  is a vertex of  $M_{sem}$ . Suppose that  $|J_0| > N$  holds. Then there is some subset  $J_1 \subset J_0$  such that  $|J_1| = N$ ,

$$C_{J_1} z_0 = c_{J_1} \quad \text{and} \quad C_{J_1} \text{ is a regular } (N \times N)\text{-matrix}.$$

Choose  $j_0 \in J_0 \setminus J_1$  arbitrarily. Then for the vertex  $z_0$  we have

$$C_{j_0} z_0 \neq c_{j_0} \iff C_{J_1, j_0} := \begin{pmatrix} C_{J_1} & c_{J_1} \\ C_{j_0} & c_{j_0} \end{pmatrix} \text{ is a regular matrix.} \quad (4)$$

By Lemma 2 the set  $V_{J_1, j_0} = \{P \in P_p \mid \det C_{J_1, j_0} \neq 0\}$  is generic in  $P_p$ . By Remark 2 also the intersection  $V_1 := \bigcap_{j_0 \in J_0 \setminus J_1} V_{J_1, j_0}$  is generic. By construction, in this set  $V_1$  the vertex  $z_0$  is non-degenerate.



Case  $d > 0$ : Suppose that  $|J_0| > N - d$ . Assume for brevity  $J_0 = \{1, \dots, k\}$ , ( $k = |J_0| > n - d$ ). Let  $C$  be the  $(N \times k)$ -matrix  $C = [C_1, \dots, C_k]$ . Relation (3) implies  $\text{rank } C = N - d$  and then

$$\det(C_{ij})_{i,j=1,\dots,N-d+1} = 0.$$

By Lemma 2 this can generically be excluded.

For the proof of (c) we moreover now show that for a generic subset of problems in  $P_p$  all vertices  $\bar{z} = (\bar{x}, \bar{y})$  of  $M_{\text{sem}}$  have different function values  $\bar{v} = d^1 \bar{z}$ . To do so let us assume that  $z_0$  and  $z_1$ ,  $z_0 \neq z_1$ , are (non-degenerate) vertices of  $M_{\text{sem}}$ . Then with corresponding index sets  $J_0, J_1 \subset J$ ,  $J_0 \neq J_1$ ,  $|J_0| = |J_1| = n + m$  we have

$$z_0 = C_{J_0}^{-1} c_{J_0}, \quad z_1 = C_{J_1}^{-1} c_{J_1}.$$

With the adjoint  $C_{J_0}^{\text{ad}}$  of  $C_{J_0}$  we can write  $C_{J_0}^{-1} = \frac{1}{\det C_{J_0}} C_{J_0}^{\text{ad}}$  and accordingly  $C_{J_1}^{-1} = \frac{1}{\det C_{J_1}} C_{J_1}^{\text{ad}}$ . Now, the values  $v_0$  and  $v_1$  are the same, i.e.  $d^1 z_0 - d^1 z_1 = 0$ , if and only if

$$p(C_{J_0}, c_{J_0}, C_{J_1}, c_{J_1}, d^1) := \det C_{J_1} \cdot d^1 C_{J_0}^{\text{ad}} c_{J_0} - \det C_{J_0} \cdot d^1 C_{J_1}^{\text{ad}} c_{J_1} = 0.$$

This relation represents a polynomial equation  $p = 0$  with a non-vanishing polynomial  $p$ . In View of Lemma 2 the set  $S_{J_0, J_1} := p^{-1}(0)$  is closed and of measure zero in  $P_p$ . Thus, the complement  $V_{J_0, J_1} = P_p \setminus S_{J_0, J_1}$  is generic. By construction, for  $P$  in  $V_{J_0, J_1}$  the vertices  $z_0, z_1$  have different values. Since there are only finitely many such sets  $J_0, J_1 \subset J$  the intersection of all corresponding sets  $V_{J_0, J_1}$  is generic in  $P_p$ .

(b) Choose  $x_0 \in X$  arbitrarily and consider the lower level problem

$$Q(x_0) : \quad \min_y d^2 y \quad \text{st. } B_j y \leq b_j - A_j x_0, \quad j \in J^2.$$

Suppose  $y_0$  is a solution of  $Q(x_0)$ . Then there exist  $J_0, J_0 \subset J^2$ ,  $|J_0| \leq m$  (by Caratheodory's Theorem),  $0 < u_0 \in \mathbb{R}^{|J_0|}$  such that

$$u B_{J_0} = -b^2, \quad B_j y_0 = b_j - A_j x_0, \quad j \in J_0.$$

Generically,  $|J_0| \geq m$ , i.e. we can assume  $|J_0| = m$ . In fact, if  $|J_0| < m$  then in view of  $u B_{J_0} = -b^2$  the  $(|J_0| + 1) \times (|J_0| + 1)$ -matrix (assume for brevity  $J_0 = \{1, \dots, |J_0|\}$ )

$$\hat{B} := [(B_{ij})_{\substack{i=1,\dots,|J_0|+1 \\ j=1,\dots,|J_0|}} \hat{b}] \quad \text{with} \quad \hat{b} := (b_1^2, \dots, b_{|J_0|+1}^2)^T$$

would satisfy  $\det(\hat{B}) = 0$  which can generically be avoided. Since generically (with  $|J_0| = m$ ) the matrix  $B_{J_0}$  is regular, a solution  $y_0$  of  $Q(x_0)$  is generically a vertex of the polyhedron  $M_2(x_0)$ . Moreover since the multiplier vector  $u_0$  is positive it is not difficult to show that  $y_0$  is the unique solution.

(c) Let  $z_0 = (x_0, y_0)$  be a local minimizer of the bilevel problem  $P$ . The feasible point  $z_0$  belongs to a face  $f_0 = f(J_0)$  of  $M_{\text{sem}}$  and by Theorem 1(b) we can assume  $f_0 \in M$ . Since  $f_0$  is a polyhedron,  $z_0$  is a global minimizer of the linear program

$$\min d^1 z \quad \text{st. } z \in f_0 = f(J_0) := \{z \in \mathbb{R}^{n+m} \mid Cz \leq c, C_{J_0} z = c_{J_0}\}. \quad (5)$$

With the same arguments as in part (b) we can show that generically the solution of this program occurs at a vertex  $z_1$  of the polyhedron  $f_0$  and that the solution is unique. The vertex  $z_0$  of the face  $f_0$  is also a (non-degenerate) vertex of  $M_{\text{sem}}$ . By the arguments in the proof of (a) all vertices have different values.

(d) Choose  $z_1 = (x_1, y_1) \in M$ . Since  $y_1$  solves  $Q(x_1)$  there exist  $J_1 \subset J$ ,  $J_1^2 \subset J_1 \cap J^2$ ,  $|J_1^2| = m$ ,  $u_1 \geq 0$ ,  $u_1 \in \mathbb{R}^m$  such that

$$C_j z_1 < c_j, \quad j \in J \setminus J_1, \quad C_j z_1 = c_j, \quad j \in J_1 \quad u_1 B_{J_1^2} = -b^2. \quad (6)$$

We now show that generically  $z_1$  is contained in a face  $f_0$  of dimension  $n$  given by (1) with  $J_0 = J_1^2$ . (Since  $|J_1^2| = m$  by definition (see Section 2) this face is non-degenerate.)

We firstly notice that for a generic subset of  $P_p$  we have  $|J_1| \leq n + m$ . Otherwise  $|J_1| > n + m$  and with some  $J_1^0 \subset J_1$ ,  $|J_1^0| = n + m$ ,  $j_0 \in J_1 \setminus J_1^0$  the quadratic matrix  $C_{J_1^0 j_0}$  in (4) would be singular which generically can be avoided.

Moreover, for a generic subset in  $P_p$  we have

$$\text{rank } C_{J_1} = |J_1| \quad \text{and} \quad \text{rank } C_{J_1^2} = |J_1^2| = m. \quad (7)$$

This holds since the condition  $\text{rank } C_{J_1} < |J_1|$  or  $\text{rank } C_{J_1^2} < |J_1^2|$  would imply that

$$\det(C_{ij})_{i,j \in J_1} = 0 \quad \text{or} \quad \det(C_{ij})_{i,j \in J_1^2} = 0$$

which by Lemma 2 can generically be excluded.

We now show that  $z_1$  is contained in an  $n$ -dimensional (non-degenerate) face. Using  $\text{rank } C_{J_1} = |J_1| \leq n + m$  (see (7)) there exist a vector  $\xi \in \mathbb{R}^{n+m}$  satisfying

$$C_{J_1^2} \xi = 0, \quad C_j \xi = -1, \quad j \in J_1 \setminus J_1^2.$$

Then, for  $z_0 := z_1 + t\xi$ ,  $t > 0$  small enough, we have (cf. (6))

$$C_{J_1^2} z_0 = c_{J_1^2}, \quad C_j z_0 < c_j, \quad j \in J_1 \setminus J_1^2, \quad u_1 B_{J_1^2} = -b^2. \quad (8)$$

Thus  $z_1$  and  $z_0$  are contained in the feasible face

$$f_0 = \{z \in M_{\text{sem}} \mid C_{J_1^2} z = c_{J_1^2}\}$$

of dimension  $d = n + m - |J_1^2| = n$ .

Suppose now that  $z_1$  is contained in a feasible face  $f_2$  of dimension  $> n$ . Then by definition there exist a feasible point  $z_2 = (x_2, y_2) \in f_2$  and index sets  $J_2 \subset J$ ,  $J_2^2 := J_2 \cap J^2$  and  $u_2 \geq 0$ ,  $u_2 \in \mathbb{R}^{J_2^2}$  such that

$$C_{J_2^2} z_2 = c_{J_2^2}, \quad C_j z_2 < c_j, \quad j \in J \setminus J_2, \quad u_2 B_{J_2^2} = -b^2$$

and  $\dim \text{span} \{C_j, j \in J_2\} = n + m - d < m$  (i.e.  $d > n$ ). Generically we can assume that  $C_{J_2}$  has full rank  $|J_2|$  (see (7)). This implies  $|J_2^2| \leq |J_2| < m$  and  $y_2$  is not a vertex solution of  $Q(x_2)$ . However this can generically be excluded as shown in the proof of part (b).  $\square$

We say that the semi-feasible set  $M_{\text{sem}}$  satisfies the Slater condition if there is a point  $z_0$  such that

$$Cz_0 < c .$$

Such a point  $z_0$  is an inner point of  $M_{\text{sem}}$ . For the numerical computations we want to restrict the problem set to the following set of linear bilevel problems.

$$P_p^r = \{P \in P_p \mid M_{\text{sem}} \text{ fulfills the Slater condition , } M_{\text{sem}} \text{ compact, } \\ S(x) \text{ is compact } \forall x \in X\} .$$

In this set, for any  $x \in X$  a solution of  $Q(x)$  exists. The following stability statement holds.

**Lemma 4** *The problem set  $P_p^r$  is open in  $\mathbb{R}^K$ .*

**Proof.** Let for  $\bar{P} \in P_p$  the Slater condition be satisfied with  $\bar{z}$ , i.e.  $\bar{C}\bar{z} < \bar{c}$  ( $\bar{C}$ ,  $\bar{c}$  defining the constraints of  $\bar{P}$ ). Then, obviously for a whole neighbourhood of problems  $P \in P_p$  the condition  $C\bar{z} < c$  holds.

To show that  $M_{\text{sem}}$  is compact it suffice to prove boundedness. We show: Given  $\bar{P} \in P_p^r$  with bounded  $M_{\text{sem}}(\bar{P})$  there exists some  $\varepsilon$  such that

$$\bigcup_{\|P-\bar{P}\|<\varepsilon} M_{\text{sem}}(P) \quad \text{is bounded} . \quad (9)$$

Suppose (9) does not hold. Then there exists a sequence of problems  $P_k \in P_p$  and vectors  $z_k \in M_{\text{sem}}(P_k)$  such that (with  $C_k, c_k$  corresponding to  $P_k$ )

$$P_k \rightarrow \bar{P}, \quad C_k z_k \leq c_k, \quad \text{and} \quad \|z_k\| \rightarrow \infty \quad \text{for } k \rightarrow \infty . \quad (10)$$

By dividing the constraints by  $\|z_k\|$  and assuming (take a subsequence)  $\frac{z_k}{\|z_k\|} \rightarrow \hat{z}$  we find for  $k \rightarrow \infty$ ,

$$\bar{C}\hat{z} \leq 0 .$$

Choosing a point  $\bar{z} \in M_{\text{sem}}(\bar{P})$  also  $z(t) := \bar{z} + t\hat{z} \in M_{\text{sem}}(\bar{P})$  for all  $t > 0$  contradicting the boundedness of  $M_{\text{sem}}(\bar{P})$ .

We finally prove the statement for  $S(x)$ . Let us assume that we have given a problem  $\bar{P} \in P_p^r$  such that the corresponding sets  $\bar{S}(x)$  are compact for all  $x \in \bar{X}$ . We have to show that there exists some  $\varepsilon$  such that for all  $P$ ,  $\|P - \bar{P}\| < \varepsilon$  with corresponding sets  $S$  and  $X$  the property

$$S(x) \quad \text{is compact for all } x \in X \quad (11)$$

holds. We only have to prove boundedness since the solution sets  $S(x)$  are always (closed) faces of  $M_2(x)$ . Suppose now that (11) is not true in a neighborhood of  $\bar{P}$ . Then there exists a sequence of problems  $P_k \in P_p$ ,  $P_k \rightarrow \bar{P}$  and points  $x_k \in X_k$ ,  $y_k \in S_k(x_k)$  such that

$$\|y_k\| \rightarrow \infty \quad \text{for } k \rightarrow \infty .$$

In view of  $x_k \in X_k$  we can choose elements  $(x_k, \tilde{y}_k) \in M_{\text{sem}}(P_k)$ . Using (9) we can assume (taking a subsequence)

$$(x_k, \tilde{y}_k) \rightarrow (\bar{x}, \tilde{y}) \in M_{\text{sem}}(\bar{P}). \quad (12)$$

By assumption,  $S(\bar{x})$  is bounded. Since the sequence  $\|(x_k, \tilde{y}_k)\|$  is bounded the following inequalities hold with some  $\rho > 0$ ,

$$a_k^2 x_k + b_k^2 y_k \leq a_k^2 x_k + b_k^2 \tilde{y}_k \leq \rho.$$

Consequently

$$A_k^2 x_k + B_k^2 y_k \leq c_k^2, \quad a_k^2 x_k + b_k^2 y_k \leq \rho.$$

Dividing these relations by  $\|y_k\|$  and assuming  $\frac{y_k}{\|y_k\|} \rightarrow \hat{y}$ ,  $\|\hat{y}\| = 1$  we find using  $\|y_k\| \rightarrow \infty$  that

$$\bar{B}^2 \hat{y} \leq 0 \quad \text{and} \quad \bar{b}^2 \hat{y} \leq 0. \quad (13)$$

We choose some  $\bar{y} \in \bar{S}(\bar{x})$  and define  $y(t) := \bar{y} + t\hat{y}$ . In view of (13), for all  $t > 0$ ,

$$\bar{A}^2 \bar{x} + \bar{B}^2 y(t) \leq \bar{c}^2 \quad \text{and} \quad \bar{a}^2 \bar{x} + \bar{b}^2 y(t) \leq \bar{a}^2 \bar{x} + \bar{b}^2 \bar{y},$$

i.e.  $y(t) \in \bar{S}(\bar{x})$ . This contradicts the fact that  $\bar{S}(\bar{x})$  is bounded.  $\square$

These genericity results in particular mean that given a LBL-problem  $P$  which does not have the nice properties in Theorem 3 (i.e.  $P \notin V$ ) by almost all arbitrarily small perturbations we obtain a problem in  $V$ . However, in contrast to the situation for linear programs, where a 'small' perturbation of the problem data leads to a 'small' perturbation of the minimal value, here for LBL-problems an arbitrarily small perturbation of the problem may lead to a large perturbation in the minimal value.

To illustrate this phenomenon we give an example:

$$\begin{aligned} \min x + y_1 \quad & \text{s.t. } y = (y_1, y_2) \text{ is a solution of} \\ \text{Q(x): } \min \varepsilon y_1 + y_2 \quad & \text{s.t. } 0 \leq x, y_1, y_2 \leq 1 \end{aligned}$$

For  $\varepsilon = 0$  the situation is degenerate. We find  $S(x) = \{(x, y_1, 0) \mid 0 \leq y_1 \leq 1\}$  for all  $x$  and the feasible set  $M = \{(x, y_1, 0) \mid 0 \leq x, y_1 \leq 1\}$  has dimension 2. The minimizer is  $(\bar{x}, \bar{y}_1, \bar{y}_2) = (0, 0, 0)$  with value  $v = 0$ . For any  $\varepsilon < 0$  the situation is non-degenerate. We find  $S(x) = \{(x, 1, 0)\}$  for all  $x$  and the feasible set  $M = \{(x, 1, 0) \mid 0 \leq x \leq 1\}$  has dimension 1. The minimizer of the LBL has completely changed to  $(\bar{x}, \bar{y}_1, \bar{y}_2) = (0, 1, 0)$  with value  $v = 1$ .

## 4 A new algorithm for computing local minima of LBL

Different methods for solving linear bilevel problems have been designed. For example the algorithm of Bard/Moore in [1] which combines a Kuhn-Tucker approach with a branch and bound method, the penalty method (see e.g. White/Anandalingam [10]) and the subgradient method (see e.g. Falk/Liu [3]). An overview of numerical methods is to be found in [8].

It is well-known that the LBL-problem (also the problem without upper level constraints) is NP-complete (see [2]). So, (unless P=NP) no efficient (polynomial) algorithm can be expected to solve the global minimization problem for LBL. Therefore it could be of interest to have a method which is able to compute at least a local minimizer of LBL efficiently.

In this section we describe such an algorithm for the bilevel problem without upper level constraints: With  $z = (x, y) \in \mathbb{R}^{n+m}$ ,  $C = [A \ B]$

$$\begin{aligned} \text{(LBL}_0\text{): } \min_{x,y} a^1 x + b^1 y \quad (\text{or } d^1 z) \quad \text{s.t. } y \text{ is a solution of} \\ Q(x): \min_y b^2 y \quad \text{s.t.} \quad Ax + By \leq c \quad (\text{or } Cz \leq c). \end{aligned}$$

Again,  $C_j$  denotes the  $j$ -th row of  $C$ ,  $j \in J := \{1, \dots, k\}$ . As usual, for  $z$  satisfying  $Cz \leq c$  the active index set is defined by  $J(z) = \{j \in J \mid C_j z = c_j\}$ . For  $J_k \subset J$  we introduce the linear subspace

$$S(J_k) = \{z \in \mathbb{R}^{n+m} \mid C_{J_k} z = 0\}.$$

In every step of the algorithm below we have to compute the projection of the object vector  $-d^1$  onto a space  $S(J_k)$  corresponding to a face  $f(J_k)$  of  $M_{\text{sem}}$ .

Our method is based on the analysis of the structure of the feasible set in Section 2 and combines projected gradient steps with ideas of the Simplex method. The conceptual method is as follows:

- Phase I: Compute a starting feasible point  $z_0 = (x_0, y_0)$  of  $\text{LBL}_0$ .
- Phase II: Compute a local minimizer  $\bar{z} = (\bar{x}, \bar{y})$  by proceeding with projected gradient steps along feasible faces of dimension  $(n - \kappa)$ ,  $\kappa = \kappa_0, \dots, n$ .

We now describe our algorithm in detail.

**Phase I:** (*Computation of a feasible point  $z_0$  and a descent direction  $d_0$  in  $z_0$* )

1. Compute a solution  $\hat{z} = (\hat{x}, \hat{y})$  of the LP-relaxation of  $\text{LBL}_0$  :

$$\min_z d^1 z \quad \text{s.t.} \quad Cz \leq c.$$

(If  $\hat{z}$  is feasible for  $\text{LBL}_0$ , i.e. if  $\hat{y}$  solves  $Q(\hat{x})$  then stop:  $\hat{z}$  is a solution of  $\text{LBL}_0$ .)

2. Compute a solution  $y_0$  of  $Q(\hat{x})$ ;  $z_0 := (\hat{x}, y_0)$  is a feasible point.
3. Put  $J_0 = J(z_0)$  and compute the projection  $s_0$  of  $-d^1$  onto  $S(J_0)$ .

Phase II : (Computation of a local minimizer)

We start with the feasible point  $z_0$  and the direction  $s_0$  computed in Phase I and end up with a local minimizer  $\bar{z}$  of  $LBL_0$ .

Step  $k \rightarrow k + 1$  : We have given

a feasible point	$z_k = (x_k, y_k) \in \mathbb{R}^{n+m}$
a feasible descent direction	$s_k \in \mathbb{R}^{n+m}$
a multiplier	$u_k \geq 0$
an index set	$J_k \subset J(z_k), m \leq  J_k  \leq n + m - 1$

such that

1.  $z_k$  and  $z_k(t) := z_k + ts_k$  ( $t \geq 0$  small) are contained in  $f(J_k)$  and

$$C_j z_k(t) < c_j, \quad j \in J \setminus J_k, \quad \text{for all } t > 0 \text{ small.}$$

2.  $d^1 s_k < 0$

3.  $u_k B_{J_k} = -b^2$

(i) : Move along the (feasible) ray  $z_k(t) := z_k + ts_k, t \geq 0$  to the 'boundary' of  $M_{\text{sem}}$ . The maximum step-length is

$$t_k := \min_{j \notin J_k, C_j s_k > 0} \frac{c_j - C_j z_k}{C_j s_k} \quad \text{with} \quad j_k^+ \in \operatorname{argmin}\{t_k\}.$$

Put  $z_{k+1} := z_k + t_k s_k$ .

(Since  $M_{\text{sem}}$  is bounded we must have  $t_k < \infty$ .)

(ii) : Change to a new feasible face depending on the number  $|J_k|, m \leq |J_k| \leq n + m - 1$ . We distinguish between three cases

- |     |                            |                                    |
|-----|----------------------------|------------------------------------|
| (A) | $ J_k  = m$                | (face of 'maximum' dimension $n$ ) |
| (B) | $m <  J_k  \leq n + m - 2$ |                                    |
| (C) | $ J_k  = n + m - 1$        | (feasible edge)                    |

(A)  $|J_k| = m$  (try to move to a new feasible face of dimension  $n$ )

Compute the solution  $\bar{u}$  of

$$\bar{u} B_{J_k} = B_{j_k^+}.$$

case  $\bar{u} \not\leq 0$  :

(a) (feasibility test w.r.t. the multipliers of the lower level) Compute

$$\rho_k = \min_{j \in J_k, \bar{u}_j > 0} \left\{ \frac{(u_k)_j}{\bar{u}_j} \right\}, \quad j_k^- \in \operatorname{argmin}\{\rho_k\}.$$

and put  $J_* = J_k \cup \{j_k^+\} \setminus \{j_k^-\}$ .

(Note that by construction  $B_{J_*}^T (u_k - \rho_k \bar{u}) + \rho_k B_{j_k^+} = -b^2$ .)

(b) (*feasibility test w.r.t. constraints*) Compute the projection  $s_{k+1}$  of  $-d^1$  onto  $S(J_*)$ .

If  $s_{k+1}C_{j_k^-} < 0$  put  $J_{k+1} = J_k \cup \{j_k^+\} \setminus \{j_k^-\}$ .

( $s_{k+1}$  is a feasible descent direction in  $z_{k+1}$  on the face  $f(J_{k+1})$  of dimension  $n$ )

If  $s_{k+1}C_{j_k^-} \geq 0$  put  $J_{k+1} = J_k \cup \{j_k^+\}$  and compute the projection  $s_{k+1}$  of  $-d^1$  onto  $f(J_{k+1})$ .

( $s_{k+1}$  is feasible direction of descent in  $z_{k+1}$  on the face  $f(J_{k+1})$  of dimension  $n - 1$ )

case  $\bar{u} \leq 0$ : (face  $f(J_*)$  is not feasible)

Put  $J_{k+1} = J_k \cup \{j_k^+\}$  and compute the projection  $s_{k+1}$  of  $-d^1$  onto  $S(J_{k+1})$ .

(B)  $m < |J_k| \leq n + m - 2$  (move to a face of dimension  $n + m - |J_k| - 1$ .)

Put  $J_{k+1} = J_k \cup \{j_k^+\}$ . and compute the projection  $s_{k+1}$  of  $-d^1$  onto  $S(J_{k+1})$ .

(C)  $|J_k| = n + m - 1$  (*find a feasible descent edge emanating from the vertex  $z_{k+1}$* )

Put  $J_* = J_k \cup \{j_k^+\}$  (and assume for brevity  $J_* = \{1, \dots, n + m\}$ ).

For  $i \in J_k$ :

- (1) Compute the solution  $s^i$  of  $C_{J_*} s = -e_i$   
( $s^i$  is the direction of the edge emanating from  $z_{k+1}$ )  
if  $d^1 s^i \geq 0$  goto next  $i$ , if  $d^1 s^i < 0$  goto (2).
- (2) Put  $J^i = J_* \setminus \{i\}$  and solve the feasibility condition

$$u^i B_{J^i} = -b^2, \quad u^i \geq 0.$$

If a solution exists then put  $s_{k+1} = s^i$ ,  $J_{k+1} = J_k \cup \{j_k^+\} \setminus \{i\}$ ,  $k \rightarrow k + 1$ .  
otherwise goto next  $i$ .

If no feasible edge of descent is found in  $z_{k+1}$  then the vertex  $z_{k+1}$  is a local minimizer of  $LBL_0$  (see Lemma 5).

**Lemma 5** *Given a non-degenerate vertex  $z_{k+1} \in M_{sem}$ , and let  $J_* := J(z_{k+1})$ . Suppose for all directions  $s^i$  (of the edges emanating from  $z_{k+1}$ ),  $i \in J_*$ , at least one of the following holds:*

1. The vector  $s^i$  is not a descent direction ( $s^i d^1 \geq 0$ )
2. The points  $z_{k+1}(t) := z_{k+1} + t s^i$ , with  $t > 0$  small, are not feasible.

*Then the vertex  $z_{k+1}$  is a local minimum of  $LBL_0$ .*

**Proof.** Assume  $z_{k+1}$  is not a local minimizer. Then a descent direction  $d$  must exist, such that

$$\text{For } t > 0 \text{ small, } z_{k+1}(t) = z_{k+1} + t d \text{ is feasible and } d^1 d < 0. \quad (14)$$

Direction  $d$  can be written as positive combination of the directions  $s^i$ . So,

$$d = \sum_{i \in \bar{J}} \alpha_i s^i$$

with some  $\bar{J} \subset J_*$  and  $\alpha_i > 0, i \in \bar{J}$ . (Again, assume for brevity  $J_* = \{1, \dots, n+m\}$ .) Consequently, in view of  $C_i s^i = -1, i \in \bar{J}$ , the indices in  $\bar{J}$  are no longer active for  $z_{k+1}(t)$  for  $t > 0$  small, i.e.  $J(z_{k+1}(t)) = J_* \setminus \bar{J} =: J_0$ . Moreover, since  $z_{k+1}(t)$  is feasible, with some  $\bar{u}$  we have

$$B_{J_0}^T \bar{u} = -b^2, \quad \bar{u} \geq 0. \quad (15)$$

Thus, for all  $i \in \bar{J} = J_* \setminus J_0$  the multiplier  $\bar{u}$  in (15) gives a solution of

$$B_{J_* \setminus \{i\}}^T u = -b^2, \quad u \geq 0.$$

In other words,  $s^i$  is a feasible direction in  $z_{k+1}$ . By assumption  $s^i d^1 \geq 0$ . Therefore,

$$d d^1 = \sum_{i \in \bar{J}} \alpha_i s^i d^1 \geq 0.$$

in contradiction to the second condition in (14).  $\square$

**Remark 3** (*Finiteness of the algorithm*)

If all feasible faces attained during the algorithm are non-degenerate and  $s_k \neq 0$  for all projections (which is generically satisfied) the algorithm above computes a local minimizer after finitely many steps. This can be seen as follows. Suppose that in step (ii)A of the algorithm we arrive at point  $z_{k+1}$  with active indices  $J(z_{k+1}) = J_k \cup \{j_k^+\}$ . Then if  $s_{k+1} C_{j_k^-} < 0$  holds we pass to a new face  $f(J_{k+1})$  with  $J_{k+1} = J_k \cup \{j_k^+\} \setminus \cup \{j_k^-\}$  of dimension  $n$  and we never can come back to a point in the relative interior of the face  $f(J(z_{k+1}))$ . In the other case  $s_{k+1} C_{j_k^-} \geq 0$  we pass successively to faces of dimension  $n-1, n-2, \dots, 0$ . Finally we end up with steps proceeding from vertex to vertex of the polyhedron  $M_{\text{sem}}$  with strictly decreasing object value. So, during the algorithm we never can reach two points  $z_k, z_l$  with the same active index sets  $J(z_k) = J(z_l)$ . The result follows since there are only finitely many possible active sets.

**Remark 4** For brevity we did not indicate how the different steps of our algorithm can be implemented efficiently (using appropriate update-formulas).

**Remark 5** (*Complexity of the algorithm*)

Every step  $k$  of our algorithm has a complexity which is polynomial in the problem size  $\sigma = (n, m, k)$ . In Phase I we have to solve two linear programs (which could be done by some polynomial algorithm). In Phase II either a projection has to be calculated (linear system) in iiA, iiB, or in iiC we have to solve a linear system and a linear feasibility problem (linear program).

However, the algorithm runs along feasible faces of  $M_{\text{sem}}$  similar to the strategy of the Simplex method. It is well-known that the worst-case behavior of the Simplex



method is not polynomial (cf. Klee/Minty [6]). Thus we expect that also for our algorithm in the worst case the number of iterations to find a local minimizer will grow exponentially with the problem size. On the other hand it is also well-known that the average behavior of the simplex method is much better (in average, the number of Simplex-steps grows proportionally with the number  $k$  of constraints; see for example Shamir [7]). So we also expect a polynomial average behavior of our method. This means, that in practice our algorithm could behave polynomially. This hope is supported by the numerical results in the next section.

**Remark 6** We restricted our algorithm to problems  $LBL_0$  without upper level constraints. The reason is that for these problems a feasible starting point can be found efficiently (by solving two LP's in Phase I).

Unfortunately, if upper level constraints are present, then the point  $z_0$  computed in Phase I need not satisfy the upper level constraints. In this case  $z_0$  is not feasible for LBL. We did not succeed in finding an efficient Phase I procedure for problems with upper level constraints. We fear, that for general LBL, the problem to find an initial feasible point has the same complexity as solving the LBL-problem to optimality (NP-complete). The modification of Phase II to general linear bilevel problems with upper level constraints does not make any problems.

We emphasize that our algorithm could be used to 'accelerate' branch and bound methods (for example, the Bard/Moore algorithm).

The 'pivot-strategy' for selecting a new feasible face can be modified in various directions. By Theorem 2, since the feasible set is path-wise connected, in principle we can reach the global minimizer of  $LBL_0$  from our starting point  $z_0$ .

## 5 Computational experiments

In this section we report on some numerical experiments with our algorithm. We compare the computing time for our local search with the time needed for the global minimization by an implementation of the Bard/Moore method in Hamming [5] on the same machine and on the same randomly generated problems. Some results are presented in table (6.1), in which we use the abbreviations:

$n$	: Number of leader variables.
$m$	: Number of follower variables.
$k$	: Number of constraints (in addition we added the constraints $x, y \geq 0$ ).
$ J_0 $	: Number of active indices at the feasible starting point $z_0$ (see Phase I).
$N_{it}$	: Number of iterations $k$ in Phase II.
$v_0 = d^1 z_0$	: Object value of the leader in the feasible starting point.
$v_{loc}$	: Object value of the leader in the local minimum.
$v_{glob}$	: Objective function value of the leader in the global minimum.
$t_{loc}$	: time (in sec.) needed for computing the local minimizer (our algorithm).
$t_{glob}$	: time (in sec.) needed for computing the global minimizer (implementation of the Bard/Moore method).

$n$	$m$	$k$	$ J_0 $	$N_{it}$	$v_0$	$v_{loc}$	$v_{glob}$	$t_{loc}$	$t_{glob}$
2	6	8	8	1	7.68		7.68	0.3	8
2	8	10	10	1	-3.90		-3.90	0.4	19
2	10	12	11	4	-44.88	-61.20	-61.20	1.3	25
4	2	6	6	1	-21.67		-21.67	0.3	2
4	4	8	7	2	-1.69	-8.53	-9.08	0.5	7
4	6	10	9	2	-0.13	-5.59	-7.16	0.7	15
4	10	14	13	3	80.72	65.36	-15.22	1.9	220
6	2	8	8	1	-186.08		-186.08	0.3	3
6	4	10	8	4	-100.99	-121.58	-148.51	0.8	21
6	6	12	11	2	-155.01	-159.39	-159.39	0.9	62
6	8	14	14	1	-328.92		-328.92	0.9	154
6	10	16	15	3	-30.11	-32.30	-34.77	1.9	133
8	2	10	10	1	-129.85		-129.85	0.5	3
8	4	12	11	2	-8.89	-30.10	-38.22	0.8	28
8	6	14	14	1	-98.12		-98.12	0.8	15
8	8	16	14	3	-128.92	-131.08	-142.52	1.7	150
8	10	18	16	3	-69.48	-96.81	-102.55	2.5	691
10	2	12	12	1	-23.95		-23.95	0.6	9
10	4	14	14	1	-131.04		-131.04	0.8	24
10	6	16	11	6	-76.81	-131.69	-149.34	1.8	213
10	8	18	15	4	-101.81	-112.57	-122.42	2.3	429
10	10	20	19	2	-88.42	-88.76	-88.76	3.6	763
12	12	24	23	2	-1.26	-30.82	-113.65	5.6	3318
16	16	32	29	4	-6.73	-18.38	-47.28	13	11413
20	20	40	36	6	-11.09	-16.94	-17.57	21	13938

**Table 6.1** Results of the computation of local minimum versus global minimum.

In 12 of the 25 test problems our local method ended up with the global solution.

The next two tables contain results with problems for constant  $n + m$  and different  $n, m$ . In the first table, for 8 of the 15, and in the second, for 6 of the 9 problems, the local method computed the global solution. In table 6.3,  $N_{ver}$  gives the number of vertex to vertex steps in Phase II *iiC*.

$(n, m, k)$	$ J_0 $	$v_0$	$v_{loc}$	$v_{glob}$	$t_{loc}$	$t_{glob}$
(4,8,16)	11	-58.07	-61.52	-61.52	1	78
	10	-107.53	-127.18	-141.50	1	133
	11	-33.01	-33.74	-33.74	1	123
	11	-18.44	-40.29	-40.29	1	169
	11	-39.40	-73.24	-78.41	2	243
(8,4,16)	12	-107.82		-107.82	1	66
	11	-44.19	-56.84	-95.90	1	62
	11	12.79	1.63	-70.82	2	43
	12	-1.08		-1.08	1	12
	12	-1.02		-1.02	1	10
(6,6,16)	12	-8.89		-8.89	1	51
	11	-14.98	-19.62	-43.29	1	92
	11	-115.81	-116.36	-126.66	1	78
	11	-116.12	-123.14	-123.14	1	187
	10	-102.42	-129.60	-129.60	1	209

**Table 6.2** Results for computing local minima versus global minima for constant  $n + m = 12$ .

$(n, m, k)$	$ J_0 $	$N_{ver}$	$v_0$	$v_{loc}$	$v_{glob}$	$t_{loc}$	$t_{glob}$
(8,8,20)	16	0	-135.96		-135.96	2	594
	13	6	-139.56	-173.84	-173.84	4	1005
	14	0	-60.40	-60.80	-60.80	2	923
(4,12,20)	15	0	151.87	150.50	-0.66	4	3366
	15	3	-85.15	-130.98	-140.70	4	1531
	14	0	-36.12	-42.10	-59.03	3	1781
(12,4,20)	16	0	-84.65		-84.65	2	139
	14	5	-60.98	-89.25	-89.25	3	209
	12	0	-121.34	-162.97	-162.97	2	75

**Table 6.3** Results of computation of local minimum versus global minimum for for constant  $n + m = 16$ .

The next table gives the computation time for the local search for increasing problem size (average of 3 randomly generated problems). The results suggest a polynomial behavior; doubling the problem size leads to a factor of about 10 in the computing time.

$(n, m, k)$	mean $t_{loc}$	$(n, m, k)$	mean $t_{loc}$
(2,2,6)	0.3	(14,14,42)	22
(4,4,12)	0.7	(16,16,48)	38
(6,6,18)	1.6	(18,18,54)	34
(8,8,24)	3.0	(20,20,60)	97
(10,10,30)	7.8	(24,24,72)	140
(12,12,36)	11	(28,28,84)	302

**Table 6.4** Mean computing times for local minimum.

Surprisingly, in all our computations we never had to start after Phase I on a feasible face of maximal dimension  $n$ . In many cases the starting feasible point  $z_0$  in Phase I, coincides with the global minimizer. In most of the other cases the point  $z_0$  was situated 'near' the local (or even global) solution, such that only few steps in Phase II had to be performed. This explains why in our experiments the computing time of our local search seems to behaves polynomial in contrast to the drastic increase in the computing time for the global search (compare for example the results in Tables 6.2 and 6.3 for  $(n, m, k) = (4, 8, 16)$  and  $(n, m, k) = (4, 12, 20)$ ; and also the experiments in [5]).

## 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] Ben-Ayed O., Blair C.E., *Computational difficulties of bilevel linear programming*, Oper. Res., 38 (3), 556-560, (1990).

- [3] Falk J.F., Liu J., *On bilevel programming, Part I*, Mathematical Programming, 70 (1), 47-72, (1995).
- [4] 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).
- [5] Hamming R., *Linear Bi-Level Optimization Problems*, Master's Thesis, Department Applied Mathematics at University of Twente, (1998).
- [6] Klee V., Minty G.J., *How good is the Simplex algorithm?*, in O. Shisha (ed.), Inequalities III, 159-175, Academic Press, New York, (1972).
- [7] Shamir R., *The efficiency of the simplex method: a survey*, Management Science 33, 301-334, (1987).
- [8] Shimizu K., Ishizuka Y., Bard J., *Nondifferentiable and two-level mathematical programming*, Kluwer Academic Publishers, Boston, (1997).
- [9] Stein O. and Still G., *Generalized semi-infinite optimization and bilevel optimization*, in preparation
- [10] White D.J., Anandalingam, *A penalty function approach for solving bilevel linear problems*, J. Global Optim. , 3 (4), 397-419, (1993)