Math. Program., Ser. A (2013) 138:309–332
DOI 10.1007/s10107-012-0535-x

FULL LENGTH PAPER

Solving bilevel programs with the KKT-approach

Gemayqzel Bouza Allende · Georg Still

Received: 8 September 2009 / Accepted: 6 March 2012 / Published online: 3 April 2012

Abstract Bilevel programs (BL) form a special class of optimization problems. They appear in many models in economics, game theory and mathematical physics. BL programs show a more complicated structure than standard finite problems. We study the so-called KKT-approach for solving bilevel problems, where the lower level minimality condition is replaced by the KKT- or the FJ-condition. This leads to a special structured mathematical program with complementarity constraints. We analyze the KKT-approach from a generic viewpoint and reveal the advantages and possible drawbacks of this approach for solving BL problems numerically.

Keywords Bilevel problems · KKT-condition · FJ-condition · Mathematical programs with complementarity constraints · Genericity · Critical points

Mathematics Subject Classification 90C30 · 90C31

1 Introduction

In the present article we consider bilevel problems (BL) of the form:

$$P_{BL}: \quad \min_{x,y} f(x, y) \quad \text{s.t. } (x, y) \in \mathcal{M}_{BL} \tag{1.1}$$

where $x \in \mathbb{R}^n,\ y \in \mathbb{R}^m$ and the feasible set is given by

G. B. Allende
University of Havana, Havana, Cuba
e-mail: gema@matcom.uh.cu

G. Still (✉)
University of Twente, Enschede, The Netherlands
e-mail: g.still@math.utwente.nl

$$\mathcal{M}_{BL} = \left\{ (x, y) \in \mathbb{R}^{n+m} \;\middle|\; \begin{array}{l} g_j(x, y) \geq 0, \; j \in J = \{1, \ldots, q\} \text{ and} \\ y \text{ is a global minimizer of } Q(x) \end{array} \right\}.$$

The so-called *lower level* problem

$$Q(x): \quad \min_y \phi(x, y) \quad \text{s.t. } y \in Y(x) = \left\{ y \in \mathbb{R}^m \mid v_i(x, y) \geq 0, \; i \in I = \{1, \ldots, l\} \right\}$$

represents a parametric program. Throughout the article we assume $(f, g_1, \ldots, g_q) \in [C^2]^{1+q}_{n+m} := C^2(\mathbb{R}^{n+m}, \mathbb{R}^{1+q})$ and $(\phi, v_1, \ldots, v_l) \in [C^3]^{1+l}_{n+m}$. Note that we consider local minimizers in the upper and global minimizers in the lower level.

Bilevel problems form an important class of mathematical programs. They appear for example in equilibrium models, in Stackelberg Games (*cf.*, [2]), and in semi-infinite programming (see [19,20]). The bilevel structure makes BL difficult to solve. Even for the feasibility check, obviously, a finite program (in y) has to be solved. During the last 20 years, books and many papers are dedicated to this topic, see e.g., [2,6,13] and the references therein.

Also from a topological viewpoint, BL is more complicated than standard finite programming. The feasible set of a BL may for example not be closed (see e.g., [20]). This phenomenon arises when the feasible set $Y(x)$ of the lower level problem does not depend continuously on x, and non-closedness can even be stable with respect to (wrt.) small, smooth perturbations of the problem functions.

An appealing way to deal with general BL's is the so called *Karush-Kuhn-Tucker* (KKT) approach where the lower level constraint, that y is a global minimizer of the program $Q(x)$, is firstly relaxed to the condition that y is a local minimizer of $Q(x)$. The latter condition is then replaced by the KKT-conditions

$$\begin{aligned} \nabla_y \phi(x, y) - \sum_{i=1}^{l} \lambda_i \nabla_y v_i(x, y) = 0, & \\ \lambda_i \geq 0, \quad v_i(x, y) \geq 0, \quad i = 1, \ldots, l, & \\ \lambda_i v_i(x, y) = 0, \quad i = 1, \ldots, l. & \end{aligned} \tag{1.2}$$

In this exposition, we more generally replace the lower level constraint by the *Fritz-John* (FJ) (necessary) conditions

$$\begin{aligned} \lambda_0 \nabla_y \phi(x, y) - \sum_{i=1}^{l} \lambda_i \nabla_y v_i(x, y) = 0, & \\ \lambda_0 \geq 0 \text{ and } \lambda_i \geq 0, \quad v_i(x, y) \geq 0, \quad i = 1, \ldots, l, & \\ \lambda_i v_i(x, y) = 0, \quad i = 1, \ldots, l & \\ \sum_{i=0}^{l} \lambda_i = 1. & \end{aligned} \tag{1.3}$$

Then instead of P_{BL}, we consider the program

$$P_{\text{FJBL}}: \quad \min_{x,y,\lambda} f(x, y) \quad \text{s.t. } (x, y, \lambda) \in \mathcal{M}_{\text{FJBL}}, \tag{1.4}$$

where $\mathcal{M}_{\text{FJBL}} = \left\{(x, y, \lambda) \in \mathbb{R}^{n+m+l+1} \mid (1.3) \text{ holds and } g_j(x, y) \geq 0,\ j \in J\right\}$. Similarly we define the program P_{KKTBL} with corresponding feasible set $\mathcal{M}_{\text{KKTBL}} = \left\{(x, y, \lambda) \in \mathbb{R}^{n+m+l} \mid (1.2) \text{ holds and } g_j(x, y) \geq 0,\ j \in J\right\}$. Note that each minimizer $y(x)$ of $Q(x)$ necessarily has to solve the FJ-conditions (1.3), so that the inclusion

$$\mathcal{M}_{BL} \subset \mathcal{M}_{\text{FJBL}}|_{\mathbb{R}^n \times \mathbb{R}^m} \tag{1.5}$$

must hold. Here, $\mathcal{M}_{\text{FJBL}}|_{\mathbb{R}^n \times \mathbb{R}^m}$ denotes the projection of $\mathcal{M}_{\text{FJBL}}$ into the space $\mathbb{R}^n \times \mathbb{R}^m$.

So, the main purpose of the KKT-approach is to find (local) minimizers of the original BL program by computing (local) minimizers of the relaxation P_{FJBL}.

Remark 1.1 Note that in case, the lower level problem $Q(x)$ is convex and satisfies a constraint qualification, y is a global minimizer of $Q(x)$ iff the KKT-condition (or equivalently the FJ-condition) is satisfied. So, under this condition, P_{BL} and P_{KKTBL} (as well as P_{FJBL}) are equivalent.

However, in general, (since (1.4) does not guarantee that y is a solution of $Q(x)$) P_{BL} and P_{FJBL} are not equivalent. So, the genericity results on P_{FJBL} in this paper do not (directly) allow conclusions on the generic structure of bilevel programming.

By (1.5) however, P_{FJBL} does yield a valid relaxation of the original problem P_{BL}. We emphasize that in general a solution of $Q(x)$ does not necessarily satisfy the KKT-conditions (1.2). So, the inclusion $\mathcal{M}_{BL} \subset \mathcal{M}_{\text{KKTBL}}|_{\mathbb{R}^n \times \mathbb{R}^m}$ is not true in general. It is therefore preferable to consider P_{FJBL} instead of P_{KKTBL}. Note that in [12] it has been shown (for $n = 1$) that the inclusion $\mathcal{M}_{BL} \subset \mathcal{M}_{\text{KKTBL}}|_{\mathbb{R}^n \times \mathbb{R}^m}$ holds generically.

Both problems P_{KKTBL}, P_{FJBL} represent specially structured *mathematical programs with complementarity constraints* (MPCC). These MPCC problems have a less complicated structure than the original BL. In particular the feasible sets $\mathcal{M}_{\text{KKTBL}}$, $\mathcal{M}_{\text{FJBL}}$ are always closed. For literature on MPCC we refer the reader e.g., to , [4,5,7,14] and [16]. To solve P_{KKTBL}, P_{FJBL} numerically, we can e.g., apply a smoothing procedure, where the complementarity constraints $\lambda_i v_i = 0$ are replaced by the perturbed relations $\lambda_i v_i = \tau$ with small $\tau > 0$. In this way , we obtain a perturbed problem $P_{\text{KKTBL}}(\tau)$ or $P_{\text{FJBL}}(\tau)$ which can be solved with methods from standard nonlinear programming. This approach has been successfully applied to the numerical solution of semi-infinite programs (see [21]).

The aim of the present article is to analyze this KKT-approach for solving BL. For that purpose the generic structure of the MPCC problem P_{FJBL} is studied. It appears that some of the difficulties of BL disappear in the KKT formulation P_{FJBL}, but a part of the singular behavior also persists in P_{FJBL}. A main result of our paper (Theorem 3.2), however, reveals that at a local solution of BL where the KKT approach leads to a singular system, generically, the minimizer can be computed by a (non-singular) reduced system. This suggests a (conceptual) computational approach which is able to overcome the singular behavior of the original BL. We emphasize that our genericity analysis exhibits the intrinsic structure of the KKT approach and precisely describes the situations any generic solution method for P_{FJBL} should be able to deal with.

The article is organized as follows. In Sect. 2 we sketch the results on MPCC problems needed later on. Section 3 considers the MPCC program P_{FJBL} and analyzes the generical properties of its feasible set and the critical points. The consequences of these results in terms of the original BL problem are discussed in Sect. 4. The whole investigations lead to an algorithmic approach for solving BL which is described in Sect. 5 along with some numerical experiments.

2 Preliminaries

In this section, we sketch concepts and results on MPCC needed later on, such as stationarity, optimality conditions, the smoothing method and certain genericity results (see [3] for details). Let us firstly recall some definitions for standard finite programs:

$$(P): \quad \min_x f(x) \quad \text{s.t.} \quad x \in \mathcal{M} := \left\{ x \in \mathbb{R}^n \mid \begin{array}{ll} h_k(x) = 0, & k \in K \\ g_j(x) \geq 0, & j \in J \end{array} \right\} \tag{2.1}$$

where $K = \{1, \ldots, q_0\}$, $J = \{1, \ldots q\}$, and $f, h_1, \ldots, h_{q_0}, g_1, \ldots, g_q \in C^2$. For example, the active index set $J_0(x)$, the Lagrangian function $L(x, \lambda, \mu)$ with Lagrangian multipliers (λ, μ), the LICQ- and MFCQ-condition, the tangent space $T_x\mathcal{M}$, the Karush-Kuhn-Tucker condition as well as SC (strict complementarity) and the second order condition (SOC), $(\nabla_x^2 L(x, \lambda, \mu)|_{T_x\mathcal{M}}$ is regular). We refer the reader to [10], [3] for details. We like to notice that in this section we use the symbols f, g_j to denote functions in a different context than in the other sections.

We now consider MPCC problems of the form

$$(P_{CC}): \quad \min_x f(x) \text{ s.t. } x \in \mathcal{M}_{CC} = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{r} h_k(x) = 0,\ k \in K \\ g_j(x) \geq 0,\ j \in J \\ r_i(x)s_i(x) = 0,\ i \in I \\ r_i(x), s_i(x) \geq 0,\ i \in I \end{array} \right\} \tag{2.2}$$

with $I = \{1, \ldots, l\}$. For this class of programs the standard concept has to be adapted as follows (see e.g., [13,15]). For a feasible point $\overline{x} \in \mathcal{M}_{CC}$ we introduce the active index sets

$$\begin{array}{ll} J_0(\overline{x}) = \{ j \mid g_j(\overline{x}) = 0 \}, & I_{rs}(\overline{x}) = \{ i \mid r_i(\overline{x}) = s_i(\overline{x}) = 0 \}, \\ I_r(\overline{x}) = \{ i \mid r_i(\overline{x}) = 0,\ s_i(\overline{x}) > 0 \}, & I_s(\overline{x}) = \{ i \mid s_i(\overline{x}) = 0,\ r_i(\overline{x}) > 0 \}. \end{array}$$

The Lagrangian function of FJ-type (near $\overline{x}$) is given by:

$$\begin{aligned} L(x, \mu_0, \lambda, \mu, \rho, \sigma) = \mu_0 f(x) &- \sum_{j \in J_0(\overline{x})} \mu_j g_j(x) - \sum_{k=1}^{q_0} \lambda_k h_k(x) \\ &- \sum_{i \in I_r(\overline{x}) \cup I_{rs}(\overline{x})} \rho_i r_i(x) - \sum_{i \in I_s(\overline{x}) \cup I_{rs}(\overline{x})} \sigma_i s_i(x). \end{aligned}$$

We say that MPCC-LICQ holds at $\overline{x} \in \mathcal{M}_{CC}$ if the vectors $\nabla h_k(\overline{x}), k \in K, \nabla g_j(\overline{x})$, $j \in J_0(\overline{x}), \nabla r_i(\overline{x}), i \in I_r(\overline{x}) \cup I_{rs}(\overline{x}), \nabla s_i(\overline{x}), i \in I_s(\overline{x}) \cup I_{rs}(\overline{x})$ are linearly independent. In the sequel, h, g, r and s stands for $(h_1, \ldots, h_{q_0})$, $(g_1, \ldots, g_q)$, $(r_1, \ldots, r_l)$

and $(s_1, \ldots, s_l)$ respectively. For $y \in \mathbb{R}^m$ and some index set $I_0 \subset \{1, \ldots, m\}$ we use the abbreviation y_{I_0} to denote the subvector $(y_i, i \in I_0)$ and write ∇s_{I_0} instead of $[\nabla s_i, i \in I_0]$.

Definition 2.1 Let $\overline{x} \in \mathcal{M}_{CC}$.

- *We call* $\overline{x}$ *a* critical point *if MPCC-LICQ* is satisfied at $\overline{x}$ and with (unique) multipliers $(\lambda, \mu, \rho, \sigma)$ the relation $\nabla_x L(\overline{x}, 1, \lambda, \mu, \rho, \sigma) = 0$ holds.
- We call $\overline{x}$ a strongly stationary point *if there are multipliers* $(\lambda, \mu, \rho, \sigma)$ *satisfying* $\nabla_x L(\overline{x}, 1, \lambda, \mu, \rho, \sigma) = 0$ *and* $\mu \geq 0$ *as well as* $\sigma_i,\ \rho_i \geq 0, \forall i \in I_{rs}(\overline{x})$.

Note that in the MPCC literature mostly the concept of weakly stationary points is used ("critical points that need not to satisfy MPCC-LICQ").

Proposition 2.1 (cf. [8,15]) *If* $\overline{x}$ *is a local minimizer where MPCC-LICQ is satisfied, then* $\overline{x}$ *is a strongly stationary point.*

Other first and second order optimality conditions can be found in [4] and [15].

Definition 2.2 *Let* $\overline{x}$ *be a critical point of* P_{CC} *with associated multiplier* $(1, \lambda, \mu, \rho, \sigma)$. *We say that* the MPCC-strict complementarity condition (MPCC-SC) holds *if*

$$\mu_j \neq 0,\ \forall j \in J_0(\overline{x}), \quad \rho_i \neq 0, \sigma_i \neq 0,\ \forall i \in I_{rs}(\overline{x}). \tag{2.3}$$

The MPCC-second order condition (MPCC-SOC) *is satisfied if*

$$\nabla_x^2 L(\overline{x}, 1, \lambda, \mu, \rho, \sigma)|_{T_{\overline{x}}\mathcal{M}_{CC}} := V^T \nabla_x^2 L(\overline{x}, 1, \lambda, \mu, \rho, \sigma) V \quad \text{is nonsingular.} \tag{2.4}$$

V is a matrix with as columns a basis of the *tangent space* $T_{\overline{x}}\mathcal{M}_{CC} := \{d \mid \nabla h_k(\overline{x})d = 0, k \in K, \nabla g_j(\overline{x})d = 0, j \in J_0(\overline{x}), \nabla s_i(\overline{x})d = 0, i \in I_s(\overline{x}) \cup I_{rs}(\overline{x}), \nabla r_i(\overline{x})d = 0, i \in I_r(\overline{x}) \cup I_{rs}(\overline{x})\}$.
A critical point $\overline{x} \in \mathcal{M}_{CC}$ *such that MPCC-LICQ, MPCC-SC and MPCC-SOC hold is called* a non-degenerate critical point in the MPCC-sense. *We say that the problem* P_{CC} *in* (2.2) is regular in the MPCC-sense *if for any feasible point* $\overline{x} \in \mathcal{M}_{CC}$ *the MPCC-LICQ condition is satisfied and each critical point is non-degenerate in the MPCC-sense.*

If $\overline{x}$ is a non-degenerate critical point in the MPCC-sense, such that $\mu_j > 0,\ j \in J_0(\overline{x}),\ \rho_i,\ \sigma_i > 0,\ i \in I_{rs}(\overline{x})$ are fulfilled and the matrix $\nabla_x^2 L(\overline{x}, 1, \lambda, \mu, \rho, \sigma)|_{T_{\overline{x}}\mathcal{M}_{CC}}$ is positive definite, then it can be seen that $\overline{x}$ is a local minimizer of P_{CC}.

Remark 2.1 To show that our MPCC program P_{FJBL} in Sect. 1 generically satisfies MPCC-SOC we will need the following fact: It is well-known (see e.g., [9]) that

$$M = \begin{pmatrix} Q & B \\ B^T & 0 \end{pmatrix} \text{ is nonsingular} \quad \Leftrightarrow \begin{cases} V^T Q V = Q|_{ker(B^T)} \text{ is nonsingular} \\ \qquad\quad \text{and } B \text{ has full rank} \end{cases}$$

If MPCC-LICQ is fulfilled, then it is easy to see that MPCC-SOC holds (at $(\overline{x}, 1, \lambda, \mu, \rho, \sigma)$) for P_{CC} if and only if the following matrix is non-singular ($J_0 := J_0(\overline{x})$ etc.):

$$\begin{pmatrix} \nabla_x^2 L(\overline{x}, 1, \lambda, \mu, \rho, \sigma) & \nabla h_K(\overline{x}) & \nabla g_{J_0}(\overline{x}) & \nabla s_{I_s \cup I_{rs}}(\overline{x}) & \nabla r_{I_r \cup I_{rs}}(\overline{x}) \\ [\nabla h_K(\overline{x})]^T & 0 & 0 & 0 & 0 \\ [\nabla g_{J_0}(\overline{x})]^T & 0 & 0 & 0 & 0 \\ [\nabla s_{I_s \cup I_{rs}}(\overline{x})]^T & 0 & 0 & 0 & 0 \\ [\nabla r_{I_r \cup I_{rs}}(\overline{x})]^T & 0 & 0 & 0 & 0 \end{pmatrix}$$

We also recall some basics in genericity theory. To denote the space $C^\kappa(\mathbb{R}^N, \mathbb{R}^M)$ we use the shorthand notation $[C^\kappa]_N^M$. This space can be endowed with the so-called strong C_S^τ-topology ($\tau \le \kappa$) (see [10] for details). We say that a property is generically fulfilled in $[C^\kappa]_N^M$ wrt. the C_S^τ-topology if there is a set $\mathcal{P}_0 \subset [C^\kappa]_N^M$ such that the property holds for all functions in $\mathcal{P}_0$ and where $\mathcal{P}_0 = \cap_{i=1}^\infty \mathcal{P}_i$ with subsets $\mathcal{P}_i \subset [C^\kappa]_N^M$ which are open and dense sets wrt. the C_S^τ-topology. By identifying the MPCC program with its problem functions (f, h, r, s, g) the set of MPCC's can be identified, e.g., with the set $[C^2]_n^{1+q_0+2l+q}$. We now give genericity results for MPCC in two different forms:

Theorem 2.1 ([4,17]) *Fix* $(f, h, r, s, g) \in [C^2]_n^{1+q_0+2l+q}$. *Then for almost all* $(b, C_h, d_h, C_r, d_r, , C_s, d_s, C_g, d_g) \in \mathbb{R}^{n+q_0n+q_0+ln+l+ln+l+qn+q}$, *the problem defined by* $(f + b^T x, h + C_h x + d_h, r + C_r x + d_r, s + C_s x + d_s, g + C_g x + d_g)$ *is regular in the MPCC-sense (almost all is to be understood in the sense of the Lebesgue measure).*

Moreover, generically in $[C^2]_n^{1+q_0+2l+q}$ *wrt. the* C_S^2*-topology, the problems* P_{CC} *are regular in the MPCC-sense.*

The second statement is proven in [17] (and [4]) based on the genericity results of Jongen-Jonker-Twilt for standard programs (see e.g., [9]). The first statement can be shown by using

Lemma 2.1 (Parameterized Sard Lemma, *cf.* [9]) *Let* $F(z, u)$ *be in* $[C^\kappa]_{n+p}^l$, *with* $\kappa > \max\{0, n - l\}$ *and* $z \in \mathbb{R}^n$, $u \in \mathbb{R}^p$. *Let us assume that 0 is a regular value of F (i.e.,* $\forall (z, u) : F(z, u) = 0 \Rightarrow \nabla_{(z,u)} F(z, u)$ *has rank l). Then for almost every* $u \in \mathbb{R}^p$, *0 is a regular value of the function* $\hat{F}_u : \mathbb{R}^n \to \mathbb{R}^l$, $\hat{F}_u(z) = F(z, u)$.

Finally, we consider the smoothing approach for solving the MPCC problem, where instead of (2.2), we solve the perturbed program P_τ,

$$\mathrm{P}_\tau : \quad \min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad x \in \mathcal{M}_\tau := \left\{ x \;\middle|\; \begin{array}{r} h_k(x) = 0,\ k \in K \\ g_j(x) \ge 0,\ j \in J \\ r_i(x) s_i(x) = \tau,\ i \in I \\ r_i(x), s_i(x) \ge 0,\ i \in I \end{array} \right\} \tag{2.5}$$

where $\tau > 0$ is a (small) perturbation parameter. We refer to [4] for convergence results for P_τ, $\tau \to 0$.

3 Genericity analysis of the KKT approach

We now consider the KKT formulation P_{FJBL} (*cf.,* (1.4)) of the bilevel problem P_{BL}. Since it is a complementarity constrained program with a special structure, the genericity results of Theorem 2.1, valid for generic MPCC programs, have to be adjusted to the special structure of P_{FJBL}.

Remark 3.1 Note that in [18] a genericity analysis wrt. MPCC-LICQ has been done for a program (MPCC's with stationary constraints) similar to the KKT formulation. However, our FJ approach leads to a program with a (slightly) different structure so that (unfortunately) a separate genericity analysis is needed. Besides, our approach does not lead to an (artificial) condition $(\phi, v, g) \in [C^\ell]$, $\ell > \max\{1, n\}$ as in [18].

It appears that also for P_{FJBL} generically MPCC-LICQ is satisfied at all feasible points. But for the local minimizers $(\overline{x}, \overline{y}, \overline{\lambda})$, the situation is more complicated than for general MPCC problems. We will show that in the generic situation the conditions MPCC-SC and MPCC-SOC may fail at local minimizers $(\overline{x}, \overline{y}, \overline{\lambda})$ of P_{FJBL}.

With respect to a feasible point $(\overline{x}, \overline{y}, \overline{\lambda})$ of problem P_{FJBL}, we introduce the active index sets:

$$
\begin{aligned}
&J_{0v}(\overline{x},\overline{y}) = \{i \in I \mid v_i(\overline{x},\overline{y}) = 0\} && J_0(\overline{x},\overline{y},\overline{\lambda}) = \left\{i \in I \mid v_i(\overline{x},\overline{y}) = 0,\ \overline{\lambda}_i > 0\right\}\\
&J_{0g}(\overline{x},\overline{y}) = \{j \in J \mid g_j(\overline{x},\overline{y}) = 0\} && J\Lambda_0(\overline{x},\overline{y},\overline{\lambda}) = \left\{i \in I \mid v_i(\overline{x},\overline{y}) = \overline{\lambda}_i = 0\right\}\\
&J_{\overline{\lambda}_0} = \begin{cases} \{0\}, & \text{if } \overline{\lambda}_0 = 0\\ \varnothing, & \text{otherwise}\end{cases} && \Lambda_0(\overline{x},\overline{y},\overline{\lambda}) = \left\{i \in I \mid v_i(\overline{x},\overline{y}) > 0,\ \overline{\lambda}_i = 0\right\}
\end{aligned}
\tag{3.1}
$$

Note that $J_{0v}(\overline{x},\overline{y}) = J_0(\overline{x},\overline{y},\overline{\lambda}) \cup J\Lambda_0(\overline{x},\overline{y},\overline{\lambda})$ does not depend on $\overline{\lambda}$. We begin by showing that, MPCC-LICQ is generically fulfilled for P_{FJBL}. The density part is proven by applying the Sard Lemma 2.1 to an appropriately chosen perturbation of (fixed) problem functions $\hat{\phi}, \hat{v}_i, \hat{g}_j$ of a given BL program. We define perturbations of these functions:

$$
(\star)\quad
\begin{aligned}
\phi(x,y) &= \hat{\phi}(x,y) + x^T[C_\phi^x]y + \frac{y^T[C_\phi^y]y}{2} + d_\phi^T y\\
v_i(x,y) &= \hat{v}_i(x,y) + x^T[C_i^x]y + \frac{y^T[C_i^y]y}{2} + b_i^{x^T} x + b_i^{y^T} y + d_i, \quad i \in I\\
g_j(x,y) &= \hat{g}_j(x,y) + C_{g_j}^T(x,y) + d_{g_j}, \quad j \in J.
\end{aligned}
$$

The matrices $C_\phi^x, C_i^x \in \mathbb{R}^{nm}$, the $m \times m$ symmetric matrices $C_\phi^y, C_i^y \in \mathbb{R}^{\frac{m(m+1)}{2}}$ and the vectors $d_\phi \in R^m, b_i^x \in \mathbb{R}^n, b_i^y \in R^m, d_i \in \mathbb{R}; C_{g_j} \in \mathbb{R}^{(n+m)}, d_{g_j} \in \mathbb{R}$ in $(\star)$ define the perturbations. So, $(\star)$ defines perturbed problem functions depending on the parameters

$$
\begin{aligned}
&Q := (C_\phi^x, C_\phi^y, d_\phi, C_i^x, C_i^y, b_i^x, b_i^y, d_i; i \in I, C_{g_j}, d_{g_j}, j \in J) \in \mathbb{R}^{N_{\text{FJBL}}} \text{ where}\\
&N_{\text{FJBL}} := nm + \tfrac{m(m+1)}{2} + m + l\big(nm + \tfrac{m(m+1)}{2} + n+m+1\big) + q(n+m+1).
\end{aligned}
$$

Theorem 3.1 *Let* $(\hat{\phi}, \hat{v}_1, \ldots, \hat{v}_l) \in [C^3]_{n+m}^{1+l}$ *and* $(\hat{g}_1, \ldots, \hat{g}_q) \in [C^2]_{n+m}^q$ *be fixed. Then for almost all parameters* $Q \in \mathbb{R}^{N_{FJBL}}$ *the condition MPCC-LICQ holds at all* *points of the feasible set $\mathcal{M}_{\text{FJBL}}$ defined by the perturbed problem functions $(\phi, v, g) = (\phi, v_1, \dots, v_l, g_1, \dots, g_q)$.*

Moreover, generically in the set $\{(\phi, v, g)\} \equiv [C^3]^{1+l}_{n+m} \times [C^2]^q_{n+m}$, wrt. the $C^3_S \times C^2_S$-topology, MPCC-LICQ holds at all points of the feasible set $\mathcal{M}_{\text{FJBL}}$.

Proof In [3] this result has been proven for $\mathcal{M}_{\text{KKTBL}}$. However, since in the proof for $\mathcal{M}_{\text{FJBL}}$ additional (non-trivial) technical difficulties appear we give the first part of the proof in detail. This first part is shown by using the Parameterized Sard Lemma. To do so, we consider a feasible point $(\overline{x}, \overline{y}, \overline{\lambda})$ of the problem P_{FJBL} defined by the problem functions (ϕ, v, g). There is a partition $J_0 = J_0(\overline{x}, \overline{y}, \overline{\lambda})$, $J\Lambda_0 = J\Lambda_0(\overline{x}, \overline{y}, \overline{\lambda})$, $\Lambda_0 = \Lambda_0(\overline{x}, \overline{y}, \overline{\lambda})$ of $I = \{1, \dots, l\}$ and a set $J_{0g} = J_{0g}(\overline{x}, \overline{y}) \subset \{1, \dots, q\}$, such that $(\overline{x}, \overline{y}, \overline{\lambda})$ solves the feasibility conditions:

$$
\begin{array}{lrll}
(1) & \lambda_0 \nabla_y \phi(x, y) - \sum_{i=1}^{l} \lambda_i \nabla_y v_i(x, y) = 0, & & \\
(2) & v_i(x, y) = 0, & i \in J_0 \cup J\Lambda_0, & \\
(3) & \lambda_i = 0, & i \in J\Lambda_0 \cup \Lambda_0, & \\
(3) & \lambda_0 = 0, & \text{in this case } J_{\lambda_0} := \{0\}, & (3.2) \\
 & & \text{or } J_{\lambda_0} := \emptyset \text{ if } \lambda_0 \neq 0 & \\
(4) & \sum_{i=0}^{l} \lambda_i = 1, & & \\
(5) & g_j(x, y) = 0, & j \in J_{0g}. &
\end{array}
$$

Let us fix a (possible) active index set $(J_0, J\Lambda_0, \Lambda_0, J_{0g}, J_{\lambda_0})$. For any solution (x, y, λ) of (3.2) the MPCC-LICQ fails, if and only if there is a vector $0 \neq (\alpha, \beta, \mu, \gamma, \rho) \in \mathbb{R}^\kappa$, $\kappa = m + |J_0 \cup J\Lambda_0| + |J_{0g}| + |J\Lambda_0 \cup \Lambda_0| + |J_{\lambda_0}| + 1$ such that $(x, y, \lambda, \alpha, \beta, \mu, \gamma, \rho)$ solves the equations:

$$
\begin{array}{ll}
(6) & \left[\lambda_0 \nabla_{(x,y)} [\nabla_y \phi(x, y)]^T - \sum_{i=1}^{l} \lambda_i \nabla_{(x,y)} [\nabla_y v_i(x, y)]^T\right] \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} \\
 & \quad + \sum_{i \in J_0 \cup J\Lambda_0} \beta_i \nabla_{(x,y)} v_i(x, y) + \sum_{j \in J_{0g}} \mu_j \nabla_{(x,y)} g_j(x, y) = 0, \\
 & \nabla_y [\phi(x, y)]^T \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} + \gamma_0 + \rho = 0, \ (\gamma_0 \neq 0 \Rightarrow \lambda_0 = 0) \\
(7) & -\nabla_y [v_i(x, y)]^T \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} + \rho = 0, \ i \in J_0, \\
 & -\nabla_y [v_i(x, y)]^T \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} + \gamma_i + \rho = 0, \ i \in J\Lambda_0 \cup \Lambda_0, \\
 & \text{and (3.2)}
\end{array}
\qquad (3.3)
$$

To apply the Sard Lemma we have to show that for all solutions of (3.3) its Jacobian (wrt. the variables and appropriate parameters) has full row rank. In order to simplify the analysis we consider different cases.

Case $\alpha = 0$: If $\alpha = 0$, then $\rho = \gamma = 0$, and thus the gradients $\nabla_{(x,y)} g_{J_{0g}}$, $\nabla_{(x,y)} v_{J_0 \cup J\Lambda_0}$ must be linearly dependent at (x, y). In other words, LICQ fails in the feasible set $\mathcal{M}^0$,

$$\mathcal{M}^0 = \left\{ (x, y) \in \mathbb{R}^{n+m} \;\middle|\; \begin{array}{ll} v_i(x, y) = 0, & i \in J_0 \cup J\Lambda_0, \\ g_j(x, y) \geq 0, & j \in J_{0g}. \end{array} \right\}$$

But $\mathcal{M}^0$ is the feasible set of a standard nonlinear program, and it is well-known (see [10]) that for almost every linear perturbation of $(\hat{v}_{J_0 \cup J\Lambda_0}, \hat{g}_{J_{0g}})$, the LICQ condition holds for all $(x, y) \in \mathcal{M}^0$. So, for almost every $(b_i^x, b_i^y, d_i; i \in I, C_{g_j}, d_{g_j}, j \in J)$ there is no point (x, y, λ) in the feasible set $\mathcal{M}_{\text{FJBL}}$ (defined by the perturbed functions (ϕ, v, g)) where LICQ fails at a (nontrivial) solution $(\alpha, \beta, \mu, \gamma)$ of (3.3) with $\alpha = 0$.

Case $\alpha \neq 0$: Without loss of generality (wlog.) we assume $\alpha_1 = 1$ in (3.3). Note that now the system (3.3) only depends on the variables $(x, y, \lambda, \alpha_2, \ldots, \alpha_m, \beta, \mu, \gamma, \rho) \in \mathbb{R}^{n+m+l+\kappa-1}$. We define the following vectors and matrices: $\hat{y} = [y_1(1|\alpha) + (0, \; y_2, \ldots, y_m)]$

$$\Lambda_1 = \begin{pmatrix} -(1, \alpha_2, \ldots, \alpha_m) & 0 & \ldots & 0 \\ 0 & (1, \alpha_2, \ldots, \alpha_m) & 0 \ldots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ldots & (1, \alpha_2, \ldots, \alpha_m) \end{pmatrix}$$
$$\in \mathbb{R}^{(l+1) \times m(l+1)},$$

$$\Lambda_2^i = \begin{pmatrix} -\lambda_i + \beta_i y_1 & \ldots & -\lambda_i \alpha_m + \beta_i y_m \\ 0 & \ddots & 0 \\ 0 & 0 & -\lambda_i + \beta_i y_1 \end{pmatrix} \in \mathbb{R}^{m \times m},$$

$$Y = \begin{pmatrix} y_1 & y_2 & \ldots & y_m \\ 0 & y_1 & 0 \ldots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ldots & y_1 \end{pmatrix} \in \mathbb{R}^{m \times m},$$

$$Y_l = \begin{pmatrix} -\hat{y} & 0 & \ldots & 0 \\ 0 & \hat{y} & 0 \ldots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ldots & \hat{y} \end{pmatrix} \in \mathbb{R}^{(l+1) \times m(l+1)},$$

and for a matrix A we put $\Lambda(A) := (\lambda_0 A, -\lambda_1 A, \ldots, -\lambda_l A)$, $\Theta(A) := (\beta_1 A, \ldots, \beta_l A)$.

Sub-case $\mu \neq 0$: The Jacobian of the system (3.3) ($\alpha_1 = 1$) with respect to the variables $z = (x, y, \lambda, \alpha_2, \ldots, \alpha_m, \beta, \mu, \gamma, \rho)$ and the parameters $u = (d_\phi, b_i^y, d_i; i \in I,$

$C_{g_j}, d_{g_j}, j \in J$) (*cf.*, Lemma 2.1) is (in the sequel $\otimes$ denotes matrices of appropriate dimensions and (6), (7) etc. indicates the derivatives of the corresponding equations in (3.2), (3.3)):

	$\partial_{(x,y,\lambda)}$	$\partial_{\alpha,\beta,\mu}$	∂_ρ	∂_γ	$\partial_{d_\phi} \vert \partial_{b_i^y}$	∂_{d_i}	$\partial_{C_{g_j}}$	$\partial_{d_{g_j}}$
(6)	$\otimes$	$\otimes$	0	0	$0 \vert \otimes$	0	$\mu_1 I_{n+m} .. \mu_q I_{n+m}$	0
(7)	$\otimes$	$\otimes$	$\begin{matrix}1\\ \vdots\\ 1\end{matrix}$	$\otimes$	$-\Lambda_1$	0	0	0
(1)	$\otimes$	0	0	0	$\Lambda(I_m)$	0	0	0
(2)	$\otimes$	0	0	0	$\otimes$	$I_{\vert J_0 \cup J\Lambda_0 \vert} \vert 0$	0	0
(5)	$\otimes$	0	0	0	0	0	$\otimes$	$I_{J_{0g}} \vert 0$
(3)	$0 \vert I_{\vert J\Lambda_0 \cup \Lambda_0 \cup J_{\lambda_0} \vert}$	0	0	0	0	0	0	0
(4)	$0 \vert 1, \ldots, 1$	0	0	0	0	0	0	0

By checking the rows of this matrix and using the fact that by $\sum_{i=0}^{l} \lambda_i = 1$, (at least) one of the numbers $\lambda_i, i = 0, \ldots, l$, is non-zero, we can see that the Jacobian has full row rank.

Sub-case $\mu = 0, \beta \neq 0$*:* Denoting by C_ϕ^{y0} and C_i^{y0} the first column of C_ϕ^y and C_i^y, respectively, the Jacobian matrix wrt. the given variables and parameters reads: $\big((6_x), (6_y)$ represent the equations (6) in (3.3) corresponding to the partial derivatives wrt. the variables x, y, respectively$\big)$

	$\partial_{(x,y,\lambda,\alpha,\beta)}$	$\partial_{\rho,\gamma}$	$\partial_{C_\phi^{y0}} \vert \partial_{C_i^{y0}}$	$\partial_{d_\phi} \vert \partial_{b_i^y}$	$\partial_{b_i^x}$	∂_{d_i}	$\partial_{d_{g_j}}$
(6_x)	$\otimes$	0	0	0	$\Theta(I_n)$	0	0
(6_y)	$\otimes$	0	$\lambda_0 \begin{pmatrix} 1 & \alpha \\ 0 & I \end{pmatrix} \Lambda_2^1, \ldots, \Lambda_2^l$	$0 \vert \Theta(I_m)$	0	0	0
(7)	$\otimes$	$\left(\begin{array}{c\|c} 1 & \\ \vdots & I \\ & 0 \\ 1 & \end{array}\right)$	$-Y_l$	$-\Lambda_1$	0	0	0
(1)	$\otimes$	0	$\Lambda(Y)$	$\Lambda(I_m)$	0	0	0
(5)	$\otimes$	0	0	0	0	0	$I_{J_{0g}} \vert 0$
(2)	$\otimes$	0	$0 \vert \otimes$	$0 \vert \otimes$	$\otimes$	$I_{\vert J_0 \cup J\Lambda_0 \vert} \vert 0$	0
(3)	$0 \vert I_{\vert J\Lambda_0 \cup \Lambda_0 \cup J_{\lambda_0} \vert} \vert 0$	0	0	0	0	0	0
(4)	$0 \vert 1, \ldots, 1 \vert 0$	0	0	0	0	0	0

We now show that the rows of this matrix are linearly independent (l.i.). Obviously, the rows corresponding to row-block 5,2 are l.i. with respect to the other rows. As $\beta \neq 0$, also matrix $\Theta(I_n)$ in the first block has full row rank. To show the linear independence of the row blocks 6_y, 7, 1 we show that the sub-matrix formed by row blocks 6_y, 7, 1 and columns corresponding to the derivatives with respect to ρ, C_ϕ^{y0}, C_i^{y0}, d_ϕ, b_i^y has full row rank.

Let us suppose that this does not hold. Then there is a vector $(a, b, c) \neq 0$ such that for the corresponding combination of the rows we find:

$$\lambda_0 \begin{pmatrix} 1 & 0 \\ \alpha & I \end{pmatrix} a + b_0 \hat{y} + \lambda_0 Y^T c = 0 \quad [\Lambda_2^i]^T a - b_i \hat{y} - \lambda_i Y^T c = 0$$
$$\begin{pmatrix} 1 \\ \alpha \end{pmatrix} b_0 + \lambda_0 c = 0 \qquad \beta_i a - \begin{pmatrix} 1 \\ \alpha \end{pmatrix} b_i - \lambda_i c = 0 \tag{3.4}$$

and (see column ∂_ρ) $\sum_i b_i = 0$. Taking the first equation of each system in (3.4) yields

$$a_1\lambda_0 + y_1 b_0 + \lambda_0 c_1 y_1 = 0 \quad (\lambda_i + \beta_i y_1)a_1 - y_1 b_i - y_1 \lambda_i c_1 = 0$$
$$b_0 + \lambda_0 c_1 = 0 \qquad \beta_i a_1 - b_i - \lambda_i c_1 = 0 \tag{3.5}$$

Multiplying the second row by y_1 and subtracting from the first, we obtain $\lambda_i a_1 = 0, \forall i$ and $\sum_{i=0}^{l} \lambda_i = 1$ implies $a_1 = 0$. So from (3.5) we get

$$b_0 + \lambda_0 c_1 = 0 \qquad b_i + \lambda_i c_1 = 0.$$

Summing up, using $\sum_{i=0}^{l} \lambda_i = 1$ and $\sum_i b_i = 0$, we find $c_1 = 0$ and then $b_i = 0$ for all $i = 1, \ldots, l$. So, the system (3.4) reduces to

$$\lambda_0 a_j + \lambda_0 y_1 c_j = 0 \quad -\lambda_i a_j + \beta_i y_1 a_j - \lambda_i y_1 c_j = 0$$
$$\lambda_0 c_j = 0 \qquad \beta_i a_j - \lambda_i c_j = 0 \tag{3.6}$$

for $j = 1, \ldots, m$. By repeating the same trick, we multiply the second row by y_1 and subtract it from the first and obtain: $\lambda_0 a_j = 0$, $-\lambda_i a_j = 0$. Finally using $\sum_{i=0}^{l} \lambda_i = 1$ again we conclude $a_j = 0$ for all j and analogously $c = 0$, contradicting $(a, b, c) \neq 0$.

So, we have shown that row blocks 6_y, 7, 1 are l.i. with respect to the other blocks. Now the independence of blocks 3, 4 is a consequence of part ∂_λ.

Sub-case $\mu = 0$, $\beta = 0$: Here we consider the Jacobian

	$\partial_{(x,y,\lambda,\alpha)}$	$\partial_{\rho,\gamma}$	$\partial_{C_\phi^x} \vert \partial_{C_i^x}$	$\partial_{C_\phi^{y0}} \vert \partial_{C_i^{y0}}$	$\partial_{d_\phi} \vert \partial_{b_i^y}$	∂_{d_i}	$\partial_{d_{g_j}}$
(6_x)	$\otimes$	0	$\Lambda(I)$	0	0	0	0
(6_y)	$\otimes$	0	0	$\Lambda \begin{pmatrix} 1 & \alpha \\ 0 & I \end{pmatrix}$	0	0	0
(7)	$\otimes$	$\left(\begin{array}{c\vert c} 1 & \\ \vdots & \otimes \\ 1 & \end{array}\right)$	$\otimes$	$\otimes$	$-\Lambda_1$	0	0
(1)	$\otimes$	0	$\otimes$	$\otimes$	$\Lambda(I_m)$	0	0
(2)	$\otimes$	0	$0\vert\otimes$	$0\vert\otimes$	$0\vert\otimes$	$I_{\vert J_0 \cup J\Lambda_0\vert}\vert 0$	0
(5)	$\otimes$	0	0	0	0	0	$I_{J_{0g}}\vert 0$
(3)	$0\vert I_{\vert J\Lambda_0 \cup \Lambda_0 \cup J_{\lambda_0}\vert}\vert 0$	0	0	0	0	0	0
(4)	$0\vert 1, \ldots, 1\vert 0$	0	0	0	0	0	0

Again the blocks 5, 2 are linearly independent. Next we take a combination (b, c) of blocks 7, 1 wrt. columns $\partial_{\rho, d_\phi, C_i^{y0}}$. This yields

$$b_i \begin{pmatrix} 1 \\ \alpha \end{pmatrix} + \lambda_i c = 0 \quad \forall i, \quad \sum_i b_i = 0 \tag{3.7}$$

Summing up and using $\sum_i \lambda_i = 1$, $\sum_i b_i = 0$ gives $\sum_i b_i \binom{1}{\alpha} + c = c = 0$. Reconsidering the first equations $b_i + \lambda_i c_1 = 0$ of system (3.7), implies $b_i = 0$. So, row block 7,1 are linearly independent with respect to the other rows. The independency of block 6_y is clearly a consequence of the fact that the matrix $\Lambda \begin{pmatrix} 1 & \alpha \\ 0 & I_{m-1} \end{pmatrix}$ has full row rank since $\begin{pmatrix} 1 & \alpha \\ 0 & I_{m-1} \end{pmatrix}$ is non-singular and $\sum_{i=0}^{l} \lambda_i = 1$. The same holds for 6_x. Block 3,4 are now evidently independent and we can conclude that the whole matrix has full row rank.

So in all cases the hypothesis of the Sard Lemma 2.1 are fulfilled and thus for almost every perturbation $u = (C_\phi^x, C_\phi^y, d_\phi, C_i^x, C_i^y, b_i^x, b_i^y, , d_i; i \in I, C_{g_j}, d_{g_j}, j \in J)$ the sub-matrix of the Jacobian with columns corresponding to the variables $z = (x, y, \lambda, \alpha_2, \ldots, \alpha_m, \beta, \mu, \gamma, \rho)$ also has full row rank. Consequently, the number $E = n+m+l+1+m+|J_0 \cup J\Lambda_0| + |J\Lambda_0 \cup \Lambda_0| + 1 + |J_{\lambda_0}| + |J_{0g}|$ of rows (equations) cannot exceed the number $V = n+m+l+1+m-1+|J_0 \cup J\Lambda_0| + 1 + |J_{\lambda_0}| + |J_{0g}| + |J\Lambda_0 \cup \Lambda_0|$ of columns (variables) leading to the inequality $1 \leq 0$, which is impossible. So, for almost every parameter, there is no solution of the system (3.3) with $\alpha_1 = 1$ (i.e., with $\alpha \neq 0$).

The perturbation arguments hold for any choice of active index sets J_0, $J\Lambda_0$, Λ_0, J_{λ_0}, J_{0g}. By considering the (finite) intersection of all corresponding parameters $(C_\phi^x, C_\phi^y, d_\phi, C_i^x, C_i^y, b_i^x, b_i^y, d_i; i \in I, C_{g_j}, d_{g_j}, j \in J)$ such that *MPCC-LICQ* holds for (3.2) we obtain our first statement.

We only give a sketch of the second general genericity statement. (See [3] for the complete proof for M_{KKTBL}.) It is firstly shown that for any fixed $N \in \mathbb{N}$, the set of functions (ϕ, v, g) where MPCC-LICQ holds at all feasible points (x, y, λ) of $\mathcal{M}_{\text{FJBL}}$ with $\|\lambda\| \leq N$, is open and dense in $[C^3]_{n+m}^{1+l} \times [C^2]_{n+m}^{q}$ wrt. the $C_S^3 \times C_S^2$-topology. The density part of the statement follows, as usual, directly from the perturbation result above. (For details we refer to [17] where such a genericity result has been proven for another class of programs). The openness property is shown by using stability arguments. If we finally consider the intersection of the open and dense sets for $N = 1, 2, \ldots$, we obtain the generic set of functions where MPCC-LICQ holds at all feasible points. $\square$

We now study the structure near the critical points (*cf.*, Definition 2.1) of problem P_{FJBL} (see (1.4)) in the generic case. Note that by Theorem 3.1 generically MPCC-LICQ holds for P_{FJBL}. The critical points (x, y, λ) are feasible points such that with multipliers $(\alpha, \beta, \mu, \gamma, \rho)$ (corresponding to the constraints $\lambda_0[\nabla_y \phi] - \sum_{i=1}^{l} \lambda_i [\nabla_y v_i] = 0$, $v_i \geq 0$, $g_j \geq 0$ $\lambda_i \geq 0$ and $\sum_{i=0}^{l} \lambda_i = 1$) the vector $(x, y, \lambda, \alpha, \beta, \mu, \gamma, \rho)$ solves the system:

$$\nabla_{(x,y)} f(x, y) - \sum_{i \in J_0 \cup J\Lambda_0} \beta_i \nabla_{(x,y)} v_i(x, y) - \sum_{j \in J_{0g}} \mu_j \nabla_{(x,y)} g_j(x, y)$$
$$-\nabla_{(x,y)} \left[\lambda_0 \nabla_y \phi(x, y)^T - \sum_{i=1}^{l} \lambda_i \nabla_y v_i(x, y)^T \right] \alpha = 0,$$
$$-\nabla_y \phi(x, y)^T \alpha - \rho - \gamma_0 = 0, \quad \text{where } J_{\lambda_0} = \varnothing \Rightarrow \gamma_0 = 0$$

$$\nabla_y v_i(x,y)^T \alpha - \rho = 0, \quad i \in J_0,$$
$$\nabla_y v_i(x,y)^T \alpha - \gamma_i - \rho = 0, \quad i \in J\Lambda_0 \cup \Lambda_0 \tag{3.8}$$

with J_0, $J\Lambda_0$, Λ_0, J_{λ_0}, J_{0g}, the active index sets at (x, y, λ), as defined in (3.1). Note that, as generically MPCC-LICQ holds, generically, any solution of (1.4) must satisfy the KKT conditions (3.8). For numerical purposes, it would be desirable that the conditions MPCC-SC and MPCC-SOC hold at critical points. MPCC-SC for (3.8) means $\mu_j \neq 0$, $j \in J_{0g}$, $\beta_i, \gamma_i \neq 0$ for all $i \in J\Lambda_0$. Unfortunately, as the following example shows, MPCC-SC may fail for P_{FJBL} even in the generic situation.

Example 3.1

$$\min -x - y \quad \text{s.t. } y \text{ solves } Q(x): \quad \begin{array}{rl} & \min y \\ \text{s.t.} & v_1(x,y) := -x + y \geq 0, \\ & v_2(x,y) := -y \geq 0. \end{array} \tag{3.9}$$

It can easily be seen that the point $(\overline{x}, \overline{y}) = (0, 0)$ is the minimizer of this bilevel problem and that at $\overline{y}$ with $\nabla_y v_1(\overline{x}, \overline{y}) = -\nabla_y v_2(\overline{x}, \overline{y}) = 1$ the condition MFCQ fails for $Q(\overline{x})$. The solution $(\overline{x}, \overline{y})$ can be found as the solution of the relaxation:

$$\min -x - y \quad \text{s.t.} \quad v_1(x,y) = -x + y \geq 0, \quad v_2(x,y) = -y \geq 0$$

by computing the unique solution of the equations:

$$v_1(x,y) = -x + y = 0, \quad v_2(x,y) = -y = 0.$$

The KKT approach leads to the program P_{FJBL} of the form

$$\min -x - y \quad \text{s.t.} \quad \begin{array}{rr} -x + y \geq 0, & -y \geq 0, \\ \lambda_0 - \lambda_1 + \lambda_2 = 0, & \lambda_0, \lambda_1, \lambda_2 \geq 0, \\ (-x + y)\lambda_1 = 0, & -y\lambda_2 = 0. \\ \lambda_0 + \lambda_1 + \lambda_2 = 1, & . \end{array} \tag{3.10}$$

The points $(x, y, \lambda_0, \lambda_1, \lambda_2) = (0, 0, \lambda_0, 1/2, 1/2 - \lambda_0)$, $0 \leq \lambda_0 \leq 1/2$ are the global minimizers of (3.10). If we choose $z^1 := (0, 0, 1/2, 1/2, 0)$, the associated multipliers (see (3.8)) are $\alpha = \gamma = \rho = 0$ and $\beta_1 = 1$, $\beta_2 = 2$. As $J\Lambda_0(z^1) = \{2\}$, *MPCC-SC* is violated at z^1. For a minimizer $z = (0, 0, \lambda)$ with $\lambda = (\lambda_0, \lambda_1, \lambda_2) > 0$, MPCC-SC holds ($J\Lambda_0(z) = \varnothing$) but the condition MPCC-SOC fails. To see this, note that $T_z\mathcal{M}_{\text{KKTBL}}$ is generated by the vector $(0, 0, -1, 0, 1)$ while $\nabla^2_{(x,y,\lambda)} L(0, 0, \lambda, 1, 2, 0, 0) = 0$, so that $\nabla^2_{(x,y,\lambda)} L(0, 0, \lambda, 1, 2, 0, 0)|_{T_z\mathcal{M}_{\text{KKTBL}}}$ is a singular matrix.

To show that the failure of MPCC-SC is stable against any small smooth perturbation we consider (small smooth) perturbations $f(x, y) = -x - y + \varepsilon_1(x, y)$, $\phi(x, y) = y + \varepsilon_2(x, y)$, $v_1(x, y) = -x + y + \varepsilon_3(x, y)$, $v_2(x, y) = -y + \varepsilon_4(x, y)$ of the problem functions in (3.9). The unique intersection point (x^*, y^*) of the constraints,

$$v_1(x,y) = -x + y + \epsilon_3(x,y) = 0 \quad \text{and} \quad v_2(x,y) = -y + \epsilon_4(x,y) = 0$$

is still the unique minimizer of the perturbed BL-problem and the corresponding FJ-condition in (3.10) is $\sum_i \lambda_i = 1$ and

$$\lambda_0\left(1+\frac{\partial\epsilon_2}{\partial y}(x^*,y^*)\right)-\lambda_1\left(1+\frac{\partial\epsilon_3}{\partial y}(x^*,y^*)\right)-\lambda_2\left(-1+\frac{\partial\epsilon_4}{\partial y}(x^*,y^*)\right)=0$$

and allows different solutions $(x^*, y^*, \lambda_0, \lambda_1, \lambda_2)$ of (3.10). If we choose again a solution $z^* = (x^*, y^*, \lambda_0^*, \lambda_1^*, \lambda_2^*)$ with $J_0(z^*) = \{1\}$, $J\Lambda_0(z^*) = \{2\}$, $\Lambda_0(z^*) = \emptyset$ we see that z^* is a solution of (3.8) with corresponding multipliers $\alpha^* = \gamma^* = \rho^* = 0$ and $\beta^* = (\beta_1, \beta_2)$ is given as the unique solution of the equation

$$\begin{pmatrix} -1+\frac{\partial\epsilon_3}{\partial x}(x^*,y^*) & \frac{\partial\epsilon_4}{\partial x}(x^*,y^*) \\ 1+\frac{\partial\epsilon_3}{\partial y}(x^*,y^*) & -1+\frac{\partial\epsilon_4}{\partial y}(x^*,y^*)\end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}=\begin{pmatrix}-1+\frac{\partial\epsilon_1}{\partial x}(x^*,y^*)\\ -1+\frac{\partial\epsilon_1}{\partial y}(x^*,y^*)\end{pmatrix}$$

Note that for small $\frac{\partial\epsilon_i}{\partial x}, \frac{\partial\epsilon_i}{\partial y}$ the system matrix is non-singular. So again, in view of $J\Lambda_0(z^*) = \{2\}$ $(\gamma_2 = \gamma^* = 0)$ at the solution $(z^*, \alpha^*, \gamma^*, \beta^*)$ of (3.8) MPCC-SC is violated.

The next result describes the generic properties of the critical points of P_{FJBL}.

Theorem 3.2 *Given $(\hat{f}, \hat{\phi}, \hat{v}_1, \ldots, \hat{v}_l) \in [C^3]^{l+2}_{n+m}$, $(\hat{g}_1, \ldots, \hat{g}_q) \in [C^2]^q_{n+m}$, let us consider the perturbed problem functions $f = \hat{f} + b^T(x, y)$, $\phi = \hat{\phi}(x, y) + c_\phi^T y$, $v_i = \hat{v}_i + c_{v_i}^T y + d_{v_i}$, $g_j = \hat{g}_j + d_{g_j}$. Then for almost every $(b, c_\phi, c_v, d_v, d_g) \in \mathbb{R}^{(n+m)+m+ml+l+q}$, at all solutions $(\overline{x}, \overline{y}, \overline{\lambda}, \alpha, \beta, \mu, \gamma, \rho)$ of the corresponding system (3.8) (i.e., $(\overline{x}, \overline{y}, \overline{\lambda})$ is a critical point of the MPCC program P_{FJBL} in (1.4)) the following holds: MPCC-LICQ is satisfied and depending on the cases:*

- BL-1: *If $\alpha \neq 0$, MPCC-SC and MPCC-SOC are fulfilled so that $(\overline{x}, \overline{y}, \overline{\lambda})$ is an isolated non-degenerate critical point of P_{FJBL} (in the MPCC-sense).*
- BL-2: *If $\alpha = 0$, then the multipliers μ_j, β_i associated with $g_j(\overline{x}, \overline{y})$, $v_i(\overline{x}, \overline{y})$, $j \in J_{0g}(\overline{x}, \overline{y})$, $i \in J_{0v}(\overline{x}, \overline{y})$, are not equal to zero and the inequality $|J_0(\overline{x}, \overline{y}, \overline{\lambda})| \geq m + |J_{\overline{\lambda}_0}(\overline{x}, \overline{y}, \overline{\lambda})|$ holds. If $\overline{\lambda}$ is such that $rank\begin{pmatrix}\nabla_y v_{J_0(\overline{x},\overline{y},\overline{\lambda})\cup\{0\}\backslash J_{\overline{\lambda}_0}(\overline{x},\overline{y},\overline{\lambda})}\\ 1,\ldots,1\end{pmatrix} = |J_0(\overline{x}, \overline{y}, \overline{\lambda}) \cup \{0\}\backslash J_{\overline{\lambda}_0}(\overline{x}, \overline{y}, \overline{\lambda})| = m + 1$ (where v_0 stands for ϕ), then MPCC-SOC holds.*
  Moreover, (given the critical point $(\overline{x}, \overline{y}, \overline{\lambda})$) there always exists a vertex solution $\overline{\lambda}^*$ (of (3.14) below) such that $(\overline{x}, \overline{y}, \overline{\lambda}^*)$ is a critical point of P_{FJBL} satisfying

$$rank\begin{pmatrix}\nabla_y v_{J_0(\overline{x},\overline{y},\overline{\lambda}^*)\cup\{0\}\backslash J_{\overline{\lambda}_0^*}(\overline{x},\overline{y},\overline{\lambda}^*)}\\ 1,\ldots,1\end{pmatrix} = |J_0(\overline{x}, \overline{y}, \overline{\lambda}^*) \cup \{0\}\backslash J_{\overline{\lambda}_0^*}(\overline{x}, \overline{y}, \overline{\lambda}^*)| = m+1. \tag{3.11}$$

Proof We sketch the proof and refer to [3] for the detailed analysis for P_{KKTBL}. (Recall that by Theorem 3.1, MPCC-LICQ generically holds on the whole feasible set.) Let us now consider a critical point $(\overline{x}, \overline{y}, \overline{\lambda})$ of (1.4) with multipliers

$(\alpha, \beta, \mu, \gamma, \rho)$ (*cf.* (3.8)). The following system describes the critical point condition together with the possibility that some of the multipliers are equal to zero, say $\beta_i, \gamma_j, \mu_k = 0$, $i \in J\Lambda^*_{0\beta} \subset J\Lambda_0$, $j \in J\Lambda^*_{0\gamma} \subset J\Lambda_0 \cup J_{\lambda_0}$, $k \in J^*_{0g} \subset J_{0g}$:

$$\text{(3.8), (3.2) holds and} \quad \begin{aligned} \beta_i &= 0, \quad i \in J\Lambda^*_{0\beta} \subset J\Lambda_0, \\ \gamma_j &= 0, \quad j \in J\Lambda^*_{0\gamma} \subset J\Lambda_0 \cup J_{\lambda_0}, \\ \mu_k &= 0, \quad k \in J^*_{0g} \subset J_{0g}. \end{aligned} \tag{3.12}$$

Here, we again skip the arguments (x, y, λ). With this setting, MPCC-SC means that $J^*_{0g} = J\Lambda^*_{0\beta} = J\Lambda^*_{0\gamma} = \varnothing$.

We consider solutions $(x, y, \lambda, \alpha, \beta, \gamma, \mu, \rho)$ of (3.12) for the perturbed functions f, ϕ, v_i, g_j and distinguish between the two cases, $\alpha = 0$ and $\alpha \neq 0$.

Case $\boldsymbol{\alpha \neq 0}$: For fixed $N \in \mathbb{N}$ we consider solutions of (3.12) with $\|\alpha\| > \frac{1}{N}$ and apply the Parameterized Sard Lemma as in the proof of Theorem 3.1 as follows. We compute the Jacobian of the system (3.12) wrt. the variables $(x, y, \lambda, \alpha, \beta, \mu, \gamma, \rho)$ and the parameters $(b, c_\phi, c_v, d_v, d_g)$. This gives a matrix similar to the Jacobian matrices in the proof of Theorem 3.1. It can be checked that this Jacobian has full row rank. The Parameterized Sard Lemma then implies that for almost every $(b, c_\phi, c_v, d_v, d_g)$, the Jacobian matrix of the system (3.12), with respect to the variables $(x, y, \lambda, \alpha, \beta, \gamma, \mu, \rho)$, has full row rank $E := n + m + l + 1 + m + |J_{0v}| + |J\Lambda_0| + |\Lambda_0| + 1 + |J_{\overline{\lambda}_0}| + |J_{0g}| + |J\Lambda^*_{0\beta}| + |J\Lambda^*_{0\gamma}| + |J^*_{0g}|$. But this rank cannot exceed the number $V := n + m + l + 1 + m + |J_0| + |J\Lambda_0| + |J\Lambda_0| + |\Lambda_0| + |J_{\overline{\lambda}_0}| + |J_{0g}| + 1$ of involved variables. So, in view of $J_0 \cup J\Lambda_0 = J_{0v}$ we must have

$$|J\Lambda^*_{0\beta}| + |J\Lambda^*_{0\gamma}| + |J^*_{0g}| = 0,$$

i.e. *MPCC-SC* holds. With similar arguments, using the full rank condition for the Jacobian, one shows that for almost all parameters $(b, c_\phi, c_v, d_v, d_g)$ MPCC-SOC and MPCC-LICQ holds (the last also follows more generally from Theorem 3.1). In particular this implies that the critical point $(\overline{x}, \overline{y}, \overline{\lambda})$ is an isolated non-degenerate critical point of P_{FJBL}.

By taking all finitely many possible combinations of active index sets into account, we conclude that for almost every linear perturbation of $(\hat{f}, \hat{\phi}, \hat{v}, \hat{g})$, the solutions of the system (3.12) with $\|\alpha\| > \frac{1}{N}$ are non-degenerate critical points of P_{FJBL}. Taking the intersection $\cap_{N\in\mathbb{N}}$ of all these function sets, we conclude that for almost every linear perturbation the non-degeneracy condition holds at all critical points with $\alpha \neq 0$.

Case $\boldsymbol{\alpha = 0}$: For a solution of (3.12) this assumption implies $\rho = 0$ and $\gamma_i = 0$, $i \in J\Lambda_0 \cup \Lambda_0$ (see (3.8)). As the set $J_{0v}(\overline{x}, \overline{y}) = J_0(\overline{x}, \overline{y}, \overline{\lambda}) \cup J\Lambda_0(\overline{x}, \overline{y}, \overline{\lambda})$ does not depend on the particular choice of λ, the critical point condition for $(\overline{x}, \overline{y}, \lambda)$ decomposes into a system in (x, y, β, μ),

$$\begin{aligned}\nabla f(x,y) - \sum_{i\in J_{0v}} \beta_i \nabla v_i(x,y) - \sum_{j\in J_{0g}} \mu_j \nabla g_j(x,y) &= 0,\\ v_i(x,y) = 0, \quad i \in J_0 \cup J\Lambda_0 &\\ g_j(x,y) = 0, \quad j \in J_{0g} &\end{aligned} \tag{3.13}$$

and for fixed $(\overline{x}, \overline{y})$ a system in λ,

$$\lambda_0 \nabla_y \phi(\overline{x}, \overline{y}) - \sum_{i\in J_{0v}(\overline{x},\overline{y})} \lambda_i \nabla_y v_i(\overline{x}, \overline{y}) = 0, \qquad \sum_{i\in J_{0v}(\overline{x},\overline{y})\cup\{0\}} \lambda_i = 1, \quad \lambda_i \geq 0. \tag{3.14}$$

Note that any solution $(\overline{x}, \overline{y}, \overline{\lambda}, \beta, \mu)$ of the system (3.13), (3.14) yields a critical point $(\overline{x}, \overline{y})$ of the standard program,

$$\begin{aligned}\min f(x,y) \quad \text{s.t. } v_i(x,y) &\geq 0, \quad i = 1,\dots,l,\\ g_j(x,y) &\geq 0, \quad j = 1,\dots,q,\end{aligned} \tag{3.15}$$

with corresponding multipliers β and μ. So, for almost all d_{v_i}, d_{g_j} the point $(\overline{x}, \overline{y})$ is a non-degenerate critical point of (3.15), i.e., $\beta_i, \mu_j \neq 0$ for all i, j and the Hessian of (3.13) (∇ denotes $\nabla_{(x,y)}$ and we again skip the arguments (x, y)),

$$A = \begin{pmatrix} \nabla^2 f - \sum\limits_{i\in J_0\cup J\Lambda_0} \beta_i \nabla^2 v_i - \sum\limits_{j\in J_{0g}} \mu_j \nabla^2 g_j & \nabla v_{J_0\cup J\Lambda_0} & \nabla g_{J_{0g}} \\ \nabla^T v_{J_0\cup J\Lambda_0} & 0 & 0 \\ \nabla^T g_{J_{0g}} & 0 & 0 \end{pmatrix} \tag{3.16}$$

is nonsingular. This follows by the genericity results for standard programs (see [10]). We now define

$$L_s = f - \sum_{i\in J_0\cup J\Lambda_0} \beta_i v_i - \sum_{j\in J_{0g}} \mu_j g_j \quad \text{and} \quad L_Q = \lambda_0 \nabla_y \phi - \sum_{i\in J_{0v}} \lambda_i \nabla_y v_i .$$

The application of the Sard Lemma to the system (3.13), (3.14) implies that, for almost every $(b, d_{v_1}, \dots, d_{v_l}, d_g, c_\phi, c_{v_1}, \dots, c_{v_l})$, the matrix

$$\begin{pmatrix} \nabla^2 L_s & 0 & \nabla v_{J_{0v}} & \nabla g_{J_{0g}} \\ \nabla^T v_{J_{0v}} & 0 & 0 & 0 \\ \nabla^T g_{J_{0g}} & 0 & 0 & 0 \\ \nabla^T L_Q & \nabla_y \phi, \nabla_y v_{J_{0v}} & 0 & 0 \\ 0 & 1, \dots, 1 & 0 & 0 \\ 0 & 0 \mid I_{J\Lambda_0 \cup J_{\overline{\lambda}_0}} & 0 & 0 \end{pmatrix} \tag{3.17}$$

has full row rank which in particular yields (by comparing the number of rows and columns)

$$m + |J_{\overline{\lambda}_0}| \leq |J_0|. \tag{3.18}$$

Note that by Charatheodory's Lemma, for any given critical point $(\bar{x}, \bar{y}, \bar{\lambda})$ we can choose a solution $\bar{\lambda}^*$ of (3.14) such that $(\bar{x}, \bar{y}, \bar{\lambda}^*)$ is a critical point satisfying (3.11) (*cf.*, (3.18)).

Now we wish to prove that MPCC-SOC is fulfilled if

$$(\star) \quad rank \begin{pmatrix} \nabla_y v_{(J_0 \cup \{0\}) \setminus J_{\bar{\lambda}_0}} \\ 1, \dots, 1 \end{pmatrix} = |J_0 \cup \{0\} \setminus J_{\bar{\lambda}_0}| = m + 1 \ \text{(full rank).}$$

Recall that $v_0 = \phi$. Using Remark 2.1, as MPCC-LICQ holds, we only need to prove the regularity of

$$M = \begin{pmatrix} \nabla^2 L_s & 0 & \nabla L_Q & \nabla v_{J_{0v}} & \nabla g_{J_{0g}} & 0 & 0 \\ 0 & 0 & \begin{matrix} \nabla_y^T v_0 \\ \nabla_y^T v_1 \\ \vdots \\ \nabla_y^T v_l \end{matrix} & 0 & 0 & \begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} & \begin{matrix} 0 \\ \vdots \\ I_{J\Lambda_0 \cup J_{\bar{\lambda}_0} \cup \Lambda_0} \end{matrix} \\ \nabla^T v_{J_0 \cup J\Lambda_0} & 0 & 0 & 0 & 0 & 0 & 0 \\ \nabla^T g_{J_{0g}} & 0 & 0 & 0 & 0 & 0 & 0 \\ \nabla^T L_Q & \nabla_y v_0, \nabla_y v_1, \dots, \nabla_y v_l & 0 & 0 & 0 & 0 & 0 \\ 0 & 1, \dots, 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 \,|\, I_{J\Lambda_0 \cup J_{\bar{\lambda}_0} \cup \Lambda_0} & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.19}$$

Here we assume that the vectors $\nabla_y v_0, \nabla_y v_1, \dots, \nabla_y v_l$ are ordered according to the index sets $J_0 \cup \{0\} \setminus J_{\bar{\lambda}_0}$, $J\Lambda_0 \cup J_{\bar{\lambda}_0}$, Λ_0. Now if $(\star)$ holds, then the columns of

$$\begin{matrix} \nabla_y^T v_0 & 1 & 0 \\ \nabla_y^T v_1 & 1 & \vdots \\ \vdots & \vdots & 0 \\ \nabla_y^T v_l & 1 & I_{J\Lambda_0 \cup J_{\bar{\lambda}_0} \cup \Lambda_0} \end{matrix}$$

are l.i. Hence, the corresponding columns in M are l.i. of the other columns in M. Using $(\star)$ we see that the number of these columns is:

$$m+1+|J\Lambda_0| + |J_{\bar{\lambda}_0}| + |\Lambda_0| = |J_0| + |\{0\}| - |J_{\bar{\lambda}_0}| + |J\Lambda_0| + |J_{\bar{\lambda}_0}| + |\Lambda_0| = l + 1.$$

Deleting these $l + 1$ columns ($[c_1, \dots, c_{l+1}]$) in M we obtain a matrix $\bar{M}$ (with N rows and $N - l - 1$ columns) that contains the matrix (3.17) as submatrix and has additional $l + 1$ zero-rows. Since (3.17) (as shown above) has full row rank the matrix $\bar{M}$ has row rank $N - l - 1$ and the same column rank. By adding again the $l + 1$ l.i. columns $[c_1, \dots, c_{l+1}]$ it follows that the matrix M in (3.19), has full rank N. □

Remark 3.2 Note that in the case where $(\star)$ (see proof above) is satisfied the vector $\bar{\lambda}$ is a vertex of the polyhedron (3.14).

Remark 3.3 For the special case that $Q(x)$ does'nt contain any constraints, i.e., $Q(x)$ is an unconstrained problem, then $P_{\text{FJBL}} = P_{\text{KKTBL}}$ reduces to a standard finite program (with constraint $\nabla_y \phi(x, y) = 0$) and the genericity results for standard finite programs in [10] can directly be applied to find that P_{FJBL} generically satisfies *LICQ*, *SOC* and *SC*.

Remark 3.4 An important subclass of critical points are the so called C-stationary points, see [11] (i.e., critical points such that $\beta_i \gamma_i \geq 0 \forall i \in J\Lambda_0$). Cases BL-1 and BL-2 will appear (generically) even for this class and corresponding genericity results can be similarly obtained.

We combine the generic properties of Theorems 3.1, 3.2 in a definition.

Definition 3.1 A bilevel problem P_{BL} is called KKT-regular if its corresponding FJ-relaxation P_{FJBL} has the regularity properties of the generic class in Theorems 3.1, 3.2.

Obviously, this definition directly yields

Corollary 3.1 *For almost all perturbations of* $(f, \phi, v_1, \ldots, v_l, g_1, \ldots, g_q)$*, linear in* $(f, g_1, \ldots, g_q) \in [C^2]$ *and quadratic in* $(\phi, v_1, \ldots, v_l) \in [C^3]$*,the corresponding problems* P_{BL} *are KKT-regular.*

From Theorem 3.2 we conclude that generically P_{FJBL} (see (1.4)) may have singular critical points only for solutions with $\alpha = 0$. We now describe the possible singular behavior at critical points $(\overline{x}, \overline{y}, \overline{\lambda})$ in this case $\alpha = 0$ in Theorem 3.2, case BL-2, more precisely. In this case (generically) the lower level problem partially vanishes. By (the proof of) Theorem 3.2, $(\overline{x}, \overline{y})$ is a critical point of the nonlinear program (3.15) (with the upper and lower level constraints). Moreover $\overline{\lambda}$ must be a solution of system (3.14) (for $(\overline{x}, \overline{y})$). Recall (see proof of Theorem 3.2), that for $(\overline{x}, \overline{y}, \overline{\lambda})$ we can construct a critical point $(\overline{x}, \overline{y}, \overline{\lambda}^*)$ of P_{FJBL} which also satisfies (3.11). For this particular critical point, (generically) one of the following sub-cases will hold in BL-2:

Case 1: $\overline{\lambda}_0^* \neq 0$ then $J_0(\overline{x}, \overline{y}, \overline{\lambda}^*) \subset J_0(\overline{x}, \overline{y}, \overline{\lambda})$ and $rank \begin{pmatrix} \nabla_y \phi, \nabla_y v_{J_0(\overline{x}, \overline{y}, \overline{\lambda}^*)} \\ 1, \ldots, 1 \end{pmatrix} = |J_0| + 1 = m + 1$. The following subcases can occur.

(a) in case $|J_{0v}(\overline{x}, \overline{y})| = m$: LICQ holds at $\overline{y}$ for $Q(\overline{x})$ and $\overline{\lambda} = \lambda^*$ is the unique solution of system (3.14). The point $(\overline{x}, \overline{y}, \overline{\lambda})$ is an isolated non-degenerate critical point of P_{FJBL} in the MPCC-sense, i.e., *MPCC-SOC* and *MPCC-SC* are fulfilled.

(b) in case $|J_{0v}(\overline{x}, \overline{y})| > m$: $J\Lambda_0(\overline{x}, \overline{y}, \lambda^*) \neq \varnothing$ holds for the vertex solutions λ^* of the system (3.14). So, *MPCC-SC* fails for $(\overline{x}, \overline{y}, \lambda^*)$ since $\gamma_i = 0$, $\forall i \in J\Lambda_0(\overline{x}, \overline{y}, \lambda^*)$.

Case 2: if $\overline{\lambda}_0^* = 0$ then $rank \begin{pmatrix} \nabla_y v_{J_0(\overline{x}, \overline{y}, \overline{\lambda}^*)} \\ 1, \ldots, 1 \end{pmatrix} = |J_0| = m + 1$. The following sub-cases are possible:

(a) in case $|J_{0v}(\overline{x}, \overline{y})| = m + 1$: $\overline{\lambda} = \lambda^*$ is the unique solution of system (3.14). The point $(\overline{x}, \overline{y}, \overline{\lambda})$ is an isolated non-degenerate critical point of P_{FJBL} in the MPCC-sense, i.e., *MPCC-SOC* and *MPCC-SC* are fulfilled.
(b) in case $|J_{0v}(\overline{x}, \overline{y})| > m + 1$: $J\Lambda_0(\overline{x}, \overline{y}, \lambda^*) \neq \varnothing$ holds for the vertex solutions λ^* of the system (3.14). So, *MPCC-SC* fails for $(\overline{x}, \overline{y}, \lambda^*)$ since $\gamma_i = 0$, $\forall i \in J\Lambda_0(\overline{x}, \overline{y}, \lambda^*)$.

Remark 3.5 Only in the cases BL-1, BL-2 case 1(a) or case 2 (a) we can guarantee that the solution $(\overline{x}, \overline{y}, \overline{\lambda})$ above is an isolated, critical point of P_{FJBL}.

4 Interpretation of the results in terms of P_{BL}

In this section we analyze the relation between the original program P_{BL} and the corresponding relaxation P_{FJBL} (or P_{KKTBL}) in the generic case, assuming that P_{BL} is KKT-regular (see Definition 3.1).

We begin with the case that $(\overline{x}, \overline{y}, \overline{\lambda})$ is a local minimizer of P_{FJBL} which satisfies the conditions BL-1 or BL-2 in Case 1(a), Case 2(a) above (see also Theorem 3.2). Then, according to Theorem 3.2 the point $(\overline{x}, \overline{y}, \overline{\lambda})$ is an isolated non-degenerate local minimizer satisfying MPCC-LICQ, MPCC-SC, and MPCC-SOC. By the results in [4, Theorems 3.3,3.4] generically $(\overline{x}, \overline{y}, \overline{\lambda})$ is an (isolated) local minimizer of P_{FJBL} either of order $p = 1$ or of order $p = 2$. This means that with constants $\varepsilon > 0, \kappa > 0$ the inequality holds:

$$f(x, y, \lambda) - f(\overline{x}, \overline{y}, \overline{\lambda}) \geq \kappa \|(x, y, \lambda) - (\overline{x}, \overline{y}, \overline{\lambda})\|^p \quad \text{for all } (x, y, \lambda) \in \mathcal{M}_{\text{FJBL}} \tag{4.1}$$

satisfying $\|(x, y, \lambda) - (\overline{x}, \overline{y}, \overline{\lambda})\| < \varepsilon$. Note that the point $(\overline{x}, \overline{y})$ need not be feasible for the original problem P_{BL}, i.e., $\overline{y}$ need not be a local minimizer of $Q(\overline{x})$. However, if $(\overline{x}, \overline{y})$ is feasible for P_{BL}, then it is also an isolated local minimizer of P_{BL}. This is stated in

Corollary 4.1 *Let P_{BL} be a KKT-regular problem and let $(\overline{x}, \overline{y}, \overline{\lambda})$ be an (isolated, non-degenerate) local minimizer of order $p = 1$ or $p = 2$ of the corresponding program P_{FJBL} in (1.4). Then, the solution $\overline{\lambda}$ of (3.14) is uniquely determined. Moreover under these conditions, if $(\overline{x}, \overline{y}) \in M_{BL}$ (feasible) then it is also a local minimizer of P_{BL} of (the same) order $p = 1$ or $p = 2$.*

Proof Assume now, that (3.14) has two solutions $\overline{\lambda} \neq \hat{\lambda}$. Then for $\delta \in [0, 1]$ also $(\overline{x}, \overline{y}, (1 - \delta)\overline{\lambda} + \delta\hat{\lambda})$ are feasible points of problem (1.4) with the same minimal objective value $f(\overline{x}, \overline{y})$. So, for small $\delta > 0$, $(\overline{x}, \overline{y}, (1 - \delta)\overline{\lambda} + \delta\hat{\lambda})$ is a local minimizer of P_{FJBL}, contradicting the fact that $(\overline{x}, \overline{y}, \overline{\lambda})$ is an isolated critical point of P_{FJBL}.

Now, let $(\overline{x}, \overline{y})$ be feasible for P_{BL}, i.e., $\overline{y}$ solves $Q(\overline{x})$, and consider any $(x, y, \lambda) \in \mathcal{M}_{\text{FJBL}}$ with $(x, y) \approx (\overline{x}, \overline{y})$. In view of the fact that $\overline{\lambda}$ is the unique solution of (3.14) a continuity argument shows that also $\lambda \approx \overline{\lambda}$ must hold. Hence, in view of the inclusion (1.5), from (4.1) we can conclude that the point $(\overline{x}, \overline{y})$ is a (locally unique) minimizer of P_{BL} of the same order $p = 1$ or $p = 2$. □

Next, we consider the situation where $(\overline{x}, \overline{y}, \overline{\lambda})$ is a local minimizer of P_{FJBL} such that the condition BL-2, of Theorem 3.2 holds and MPCC-SC is not fulfilled (BL-2, Case 1(b), Case 2(b)). In this case $\overline{\lambda}$ is not the unique solution of (3.14). Let furthermore $(\overline{x}, \overline{y})$ be feasible for P_{BL}, i.e., $\overline{y}$ is a solution of $Q(\overline{x})$. In this situation we cannot expect that around $\overline{x}$ the solution $y(x)$ of $Q(x)$ can be described by a smooth function $y(x)$ so that in this case, the original program P_{BL} cannot be solved by a reduction approach. We give an illustrative (generic) example for the case BL-2, case 1(b).

Example 4.1

$$\min\ 3x_1 + x_2 + y \quad s.t. \quad x_1 \geq 0, \quad y \textit{ solves } Q(x): \begin{array}{l} \min y \\ s.t.\ x_1 + y \geq 0, \\ \quad\ x_2 + y \geq 0. \end{array}$$

The solution $y(x)$ of the lower level problem $Q(x)$ is given by $y(x) = -x_1$ if $x_1 \leq x_2$ and $y(x) = -x_2$ otherwise. So the function $y(x)$ is not C^1 around $\overline{x} = (0, 0)$. The point $(\overline{x}, \overline{y}) = (0, 0, 0)$ is the global minimizer of the problem. Since $Q(x)$ is linear, wlog. we can consider the KKT formulation P_{KKTBL} instead of P_{FJBL}, and the point $(\overline{x}, \overline{y}, \lambda_1, \lambda_2) = (0, 0, 0, 1, 0)$ can be shown to be the minimizer of P_{KKTBL} with associated multipliers $\alpha = 0$, $\mu = 1$, $\beta = (1, 1)$, $\gamma = (0, 0)$ and $|J_{0v}(\overline{x}, \overline{y})| = 2 > m$.

Let us discuss the structure for the case BL-2, cases 1, 2 (b) further. We again analyze only the first subcase, where at a critical point $(\overline{x}, \overline{y}, \overline{\lambda})$ of P_{FJBL} we consider the solution set of (3.14) with $|J_{0v}(\overline{x}, \overline{y})| > m$. This solution set is a polyhedron of dimension d, $d \leq |J_{0v}(\overline{x}, \overline{y})| - m$. We denote this polyhedron by $R(\overline{x}, \overline{y})$. For each $\lambda^* \in R(\overline{x}, \overline{y})$, the point $(\overline{x}, \overline{y}, \lambda^*)$ is a critical point of P_{KKTBL}. The vertices of $R(\overline{x}, \overline{y})$ are given by those solutions λ^* such that $rank(\nabla_y v_{J_0(\overline{x},\overline{y},\lambda^*)}(\overline{x}, \overline{y})) = |J_0(\overline{x}, \overline{y}, \lambda^*)| = m$. In the present situation the following bad behavior may occur: The points $(\overline{x}, \overline{y}, \lambda)$ with $\lambda \in R(\overline{x}, \overline{y})$ and $J\Lambda_0(\overline{x}, \overline{y}, \lambda) = \varnothing$ (i.e., λ is in the relative interior of $R(\overline{x}, \overline{y})$) may be local minimizers of P_{FJBL}, but for a vertex $\overline{\lambda}$ of $R(\overline{x}, \overline{y})$, the point $(\overline{x}, \overline{y}, \overline{\lambda})$ is no longer a local minimizer. This means, in particular, that the set of local minimizers may not be closed. We give an example:

Example 4.2

$$\min\ -x + y \quad s.t.\ y \textit{ solves } Q(x): \min y \quad s.t. \quad x \geq 0, y \geq 0.$$

The corresponding KKT relaxation P_{KKTBL} is

$$\begin{aligned} \min\ -x + y \quad \text{s.t. } 1 - \lambda_1 &= 0, & y &\geq 0, \\ x &\geq 0, & \lambda_1, \lambda_2 &\geq 0, \\ y\lambda_1 &= 0, & x\lambda_2 &= 0. \end{aligned}$$

Obviously the points $(x, y, \lambda_1, \lambda_2) = (0, 0, 1, \lambda_2)$, with $\lambda_2 > 0$, are feasible with $|J_{0v}(x, y)| = 2 > 1 = m$ and have the same objective value $f(x, y) = 0$. It is not difficult to see that these points are local minimizers of P_{KKTBL}. However, for the vertex solution $(\lambda_1, \lambda_2) = (1, 0)$ of (3.14) (with $\lambda_0 = 1$) the corresponding point $(0, 0, 1, 0)$ is no longer a local minimizer. Indeed, the feasible points $(x, 0, 1, 0)$, $x > 0$, have a smaller objective value $f(x, 0) = -x$.

The preceding example also shows that in contrast to the other cases (see Corollary 4.1) in these cases BL-2, subcases 1, 2(b), the fact that $(\overline{x}, \overline{y}, \overline{\lambda})$ is a local minimizer of P_{FJBL}, and that $(\overline{x}, \overline{y})$ is feasible for BL, does not imply that $(\overline{x}, \overline{y})$ is a local minimizer of the original BL program. In these cases only a weaker statement than in Corollary 4.1 can be proven (see [3] for details).

Corollary 4.2 *Let $(\overline{x}, \overline{y}) \in M_{BL}$, such that $(\overline{x}, \overline{y}, \overline{\lambda})$ is a critical point of P_{FJBL} satisfying BL-2, subcases 1, 2(b). Assume that for all possible lower level multiplier vertices $\overline{\lambda}^*$ of (3.14), the condition $\beta_{J\Lambda_0(\overline{x},\overline{y},\overline{\lambda}^*)} > 0$ and $A|_{T_{\overline{x}}M} \succ 0$ holds (with A as in (3.16), M the feasible set of (3.15)). Then $(\overline{x}, \overline{y})$ is a local minimizer of the bilevel problem.*

Proof The proof is similar to that of Corollary 5.2.5 in [3] for the case of P_{KKTBL}. □

Remark 4.1 For semi-infinite programs it is known that generically for a solution $(\overline{x}, \overline{y})$ of the BL formulation the condition LICQ is satisfied at $\overline{y}$ wrt. $Q(\overline{x})$. So for SIP we can restrict the KKT approach to P_{KKTBL} and in Theorem 3.2 only the cases BL-1 and BL-2 subcase 1(a) can occur.

5 A numerical approach for solving BL

This section deals with the numerical aspects of the KKT approach for solving the original BL problem P_{BL}. In particular, we discuss the consequences of the preceding genericity results for this approach. The results suggest that we have to distinguish between the cases BL-1, BL-2 Case 1, 2(a) and the cases BL-2 Case 1, 2(b). At a minimizer $(\overline{x}, \overline{y}, \overline{\lambda})$ of P_{FJBL} satisfying BL-1 (or BL-2,(a)) the regularity conditions MPCC-LICQ,-SC,-SOC, are fulfilled so that this minimizer can be computed numerically with methods from MPCC, e.g., with the smoothing approach, where the problem P_{FJBL} is replaced by the perturbed version (see (2.5)):

$$P(\tau): \text{ program (1.4) with } \lambda_i v_i(x, y) = 0 \text{ replaced by } \lambda_i v_i(x, y) = \tau,\ i \in I, \quad (5.1)$$

where $\tau > 0$ is a (small) perturbation parameter. The program $P(\tau)$ represents an ordinary finite program and can numerically be solved by using software for standard programs. From [4, Theorem 5] we obtain

Proposition 5.1 *Let $(\overline{x}, \overline{y}, \overline{\lambda})$ be a minimizer of P_{FJBL} such that MPCC-LICQ,-SC,-SOC hold. Then for any $\tau > 0$ small enough there is a (locally unique) minimizer $(x_\tau, y_\tau, \lambda_\tau)$ of problem (5.1). Moreover the convergence $\|(x_\tau, y_\tau, \lambda_\tau) - (\overline{x}, \overline{y}, \overline{\lambda})\| = O(\sqrt{\tau})$, for $\tau \to 0$, takes place.*

At minimizers $(\overline{x}, \overline{y}, \overline{\lambda})$ of P_{FJBL} in cases BL-2,1(b), 2(b), a degenerate structure occurs, so that for the computation of these minimizers the KKT approach may not work. We give a generic example for the case BL-2,1(b) where the minimizer of P_{KKTBL} cannot be approximated by minimizers of $P(\tau)$.

Example 5.1

$$\min x + y \quad s.t. \quad y \, solves \quad Q(x) : \min \; y \quad \text{s.t.} \quad \begin{array}{r} y \geq 0, \\ 1 - y \geq 0, \\ x \geq 0. \end{array}$$

with corresponding problem P_{KKTBL}:

$$\min x + y \quad \text{s.t.} \quad \begin{array}{rr} 1 - \lambda_1 + \lambda_2 = 0, & y \geq 0, \\ 1 - y \geq 0, & x \geq 0, \\ \lambda_1, \lambda_2, \lambda_3 \geq 0, & \lambda_1 y = 0, \\ \lambda_2(1 - y) = 0, & \lambda_3 x = 0. \end{array}$$

This problem has the solution $(\overline{x}, \overline{y}, \overline{\lambda}) = (0, 0, 1, 0, 0)$ with multipliers $\beta_1 = \beta_2 = 1$, $\alpha = \gamma_1 = \gamma_2 = \gamma_3 = 0$, and $|J\Lambda_0(\overline{x}, \overline{y}, \overline{\lambda})| = 1$. It can be shown (see [3]) that near $(\overline{x}, \overline{y}, \overline{\lambda})$ there do not exist minimizers (even not critical points) of $P(\tau)$ for (small) $\tau > 0$.

However, fortunately our analysis in Sect. 3 has revealed that in case BL-2, generically, the corresponding minimizer $(\overline{x}, \overline{y})$ of P_{BL} can directly be found by computing a minimizer $(\overline{x}, \overline{y})$ of the reduced (standard) problem (3.15) and then by checking whether $(\overline{x}, \overline{y})$ is feasible for P_{BL} (i.e., $\overline{y}$ solves $Q(\overline{x})$). Recall that for minimizers of P_{FJBL} in the case BL-2, 1, 2(a) both approaches are possible. The results obtained so far suggest the

Conceptual method for solving P_{BL} **(***in the generic case***):** According to our analysis a solution of a (KKT-regular) program P_{BL} can be obtained as a solution of P_{FJBL}. The latter is either a (nondegenerate) solution of (3.15) (case BL-2) or a nondegenerate solution of P_{FJBL} (case BL-1). So we try both alternatives:

1. Try to compute the minimizer $(\overline{x}, \overline{y}, \overline{\lambda})$ of P_{FJBL} which satisfy BL-2 as a solution $(\overline{x}, \overline{y})$ of the relaxation (3.15) (such that also (3.14) holds) and check whether $\overline{y}$ solves $Q(\overline{x})$. If so, $(\overline{x}, \overline{y})$ is a minimizer of P_{BL}.

2. Try to compute a (nondegenerate) solution $(\overline{x}, \overline{y}, \overline{\lambda})$ of P_{FJBL} by applying the smoothing approach $P(\tau)$ in (5.1) (or some other method) for solving the MPCC program (1.5). In case the procedure converges to a nondegenerate solution $(\overline{x}, \overline{y}, \overline{\lambda})$ (case $\alpha \neq 0$), check whether $\overline{y}$ solves $Q(\overline{x})$. If so, $(\overline{x}, \overline{y})$ is a minimizer of P_{BL}.
If the case $\alpha = 0$ is detected, i.e., the method generates $(x_k, y_k) \to (\overline{x}, \overline{y})$, $\alpha_k \to 0$, we switch to step 1 (with last iterate (x_k, y_k) as starting point).

Remark 5.1 Note that not any solution $(\overline{x}, \overline{y}, \overline{\lambda})$ of P_{FJBL} computed by the above method need to lead to a solution $(\overline{x}, \overline{y})$ of P_{BL}. We have to additionally verify sufficient conditions (second order conditions wrt. $Q(\overline{x})$) (see Corollary 4.1 for step 2, Corollary 4.2 for step 1).

To illustrate our solution method we consider 3 (simple) bilevel problems (all are taken from http://www-unix.mcs.anl.gov/~leyffer/MacMPEC). These problems are solved numerically with the help of the corresponding program P_{KKTBL} using the smoothing approach $P(\tau)$. The (standard) finite programs $P(\tau)$ have been computed with the MATLAB procedure *fmincon*. In the following the numerical results are given in 2 decimal places, i.e., 12.00009 is written as 12.00.

The first example is due to Bard [1]:

$$
\begin{aligned}
\min\, & (x-5)^2 + (2y+1)^2 \\
\text{s.t. } & x \geq 0, \quad y \geq 0, \\
\text{and } & \; y \text{ solves } Q(x) : \min (y-1)^2 - 1.5xy \\
& \qquad\qquad\qquad \text{s.t. } \; 3x - y - 3 \geq 0, \\
& \qquad\qquad\qquad\quad -x + 0.5y + 4 \geq 0, \\
& \qquad\qquad\qquad\quad -x - y + 7 \geq 0.
\end{aligned}
$$

The reported local minimizer is $(\overline{x}, \overline{y}, \overline{\lambda}) = (1, 0, 3.5, 0, 0)$. To solve the program $P(\tau)$ (corresponding to the MPCC problem P_{KKTBL}) we started with $(x_0, y_0) = (0, 0)$ and lower level multiplier $\lambda^0 = (1, 1, 1)$. We computed (in 0.2 s CPU) the approximate solution $(x, y, \lambda) = (1.00, 0.00, 3.50, 0.00, 0.00)$ with error $0.39\,e$-06 for $\tau = 1.00\,e$-06.

To study the dependence of the procedure on the starting point (x_0, y_0) we tried to solve the same problem using 50 random starting points first in $[-1, 5]^2$ and then in $[-20, 20]^2$. Our smoothing approach succeeded for 33 starting points in the first and for 16 in the second case.

The second example (see also [1]) is:

$$
\begin{aligned}
\min\, & -x_1^2 - 3x_2 - 4y_1 + y_2^2 \\
\text{s.t. } & x \geq 0, \; y \geq 0, \; -x_1^2 - 2x_2 + 4 \geq 0, \\
\text{and } & \; y \text{ solves } Q(x) : \min y_1^2 - 5y_2 \\
& \qquad\qquad \text{s.t. } \; x_1^2 - 2x_1 + x_2^2 - 2y_1 + y_2 + 3 \geq 0, \\
& \qquad\qquad x_2 + 3y_1 - 4y_2 - 4 \geq 0.
\end{aligned}
$$

Here the reported minimizer is $(\overline{x}, \overline{y}, \overline{\lambda}) = (0, 2, 1.875, 0.9062, 0, 1.25)$. Starting our KKT approach with $(x_0, y_0) = (0, 0, 0, 0)$ and multiplier $\lambda^0 = (1, 1)$ we obtained (in 33 s CPU) the approximate solution $(x, y) = (0.00, 2.00, 1.88, 0.91)$, $\lambda = (0.00, 1.25)$ of $P(\tau)$. For $\tau = 1\,e$-06 the error $\|(x, y) - (\overline{x}, \overline{y})\|$ was $4.97\,e$-05.

By choosing 50 starting points (x_0, y_0) randomly from $[-1, 5]^4$ and $[-20, 20]^4$ the smoothing procedure succeeded 44 times in the first cases and 45 times in the second.

We end with a degenerate (non-generic) example, where *MPCC-LICQ* fails at the solution point (*cf.*, [7]). The problem is

$$
\begin{aligned}
\min\, & -x_1^2 - 2x_1 + x_2^2 - 2x_2 + y_1^2 + y_2^2 \\
\text{s.t. } & x \geq 0, \; y \geq 0, \; -x_1 + 2 \geq 0, \\
\text{and } & \; y \text{ solves } Q(x) : \min y_1^2 - 2x_1y_1 + y_2^2 - 2x_2y_2 \\
& \qquad\qquad \text{s.t. } \; .25 - (y_1 - 1)^2 \geq 0, \\
& \qquad\qquad\qquad .25 - (y_2 - 1)^2 \geq 0.
\end{aligned}
$$

In this case the minimizer is $(.5, .5, .5, .5)$ and the lower level multipliers, $\lambda = (0, 0)$. At that point, *MPCC-LICQ* is not satisfied. Our approach behaved surprisingly stable with error of order $\sqrt{\tau}$ for $\tau \geq 8.6736\,e$-19.

Acknowledgments We would like to thank both referee's for their many valuable comments and for their effort to make the paper more readable.

References

1. Bard, J.F.: Convex two-level optimization. Math. Program. **40**(1), 15–27 (1988)
2. Bard, J.F.: Practical Bilevel Optimization. Algorithms and Applications. Nonconvex Optimization and Its Applications, vol. 30. Kluwer, Dordrecht (1998)
3. Bouza, G.: Mathematical programs with equilibrium constraints: solution techniques from parametric optimization. PhD thesis, University of Twente, 2006 (PDF-file see: http://wwwhome.math.utwente.nl/~stillgj/opt.html)
4. Bouza, G., Still, G.: Mathematical programs with complementarity constraints: convergence properties of a smoothing method. Math. Oper. Res. **32**(2), 467–483 (2007)
5. Chen, X., Fukushima, M.: A smoothing method for a mathematical program with P-matrix linear complementarity constraints. Comput. Optim. Appl. **27**(3), 223–246 (2004)
6. Dempe, S.: Annotated bibliography on bilevel programming and mathematical programs with equilibrium constrains. Optimization **52**(3), 333–359 (2003)
7. Facchinei, F., Houyuan, J., Qi, L.: A smoothing method for mathematical programs with equilibrium constraints. Math. Program. **85**(1), 107–134 (1999)
8. Flegel, M., Kanzow, C.: A Fritz John approach to first order optimality conditions for mathematical programming with equilibrium constraints. Optimization **52**(3), 277–286 (2003)
9. Gómez, W., Guddat, J., Jongen, H.Th., Rückmann, J.J., Solano, C.: Curvas criticas y saltos en la optimizacion no lineal. http://www.emis.de/monographs/curvas/index.html (2000)
10. Guddat, J., Guerra, F., Jongen, H.Th.: Parametric Optimization: Singularities, Pathfollowing and Jumps. Teubner and Wiley, Chichester (1990)
11. Jongen, H.Th., Rueckmann, J.J., Shikhman, V.: MPCC: critical point theory. SIAM J. Optim. **20**(1), 473–484 (2009)
12. Jongen, H.Th., Shikhman, V.: Bilevel Optimization: on the structure of the feasible set. Math. Program. (2012, submitted)
13. Luo, Z.Q., Pang, J.S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge (1996)
14. Ralph, D., Wright, S.J.: Some properties of regularization and penalization schemes for MPEC's. Optim. Methods Softw. **19**(5), 527–556 (2004)
15. Scheel, H., Scholtes, S.: Mathematical programs with complementarity constraints: stationarity, optimality and sensitivity. Math. Oper. Res. **25**(1), 1–22 (2000)
16. Scholtes, S.: Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. SIAM J. Optim. **11**(4), 918–936 (2001)
17. Scholtes, S., Stoehr, M.: Exact penalization of mathematical programs with equilibrium constraints. SIAM J. Control Optim. **37**(2), 617–652 (1999)
18. Scholtes, S., Stöhr, M.: How stringent is the linear independence assumption for mathematical programs with stationary constraints? Math. Oper. Res. **26**(4), 851–863 (2001)
19. Stein, O.: Bi-Level Strategies in Semi-Infinite Programming. Nonconvex Optimization and its Applications, vol. 71. Kluwer, Boston, MA (2003)
20. Stein, O., Still, G.: On generalized semi-infinite optimization and bilevel optimization Eur. J. Oper. Res. **142**(3), 444–462 (2002)
21. Stein, O., Still, G.: Solving semi-infinite optimization problems with interior point techniques SIAM J. Control Optim. **42**(3), 769–788 (2003)