Schedae Informaticae Vol. 21 (2012): 65–79
doi: 10.4467/20838476SI.12.005.0815

Combined Reformulation of Bilevel Programming Problems

Maria Pilecka
Department of Mathematics and Computer Science,
TU Bergakademie Freiberg, Akademiestraße 6, 09596 Freiberg, Germany
e-mail: *maria.pilecka@math.tu-freiberg.de*

Abstract. In [19] J.J. Ye and D.L. Zhu proposed a new reformulation of a bilevel programming problem which compounds the value function and KKT approaches. In [19] partial calmness condition was also adapted to this new reformulation and optimality conditions using partial calmness were introduced. In this paper we investigate above all local equivalence of the combined reformulation and the initial problem and how constraint qualifications and optimality conditions could be defined for this reformulation without using partial calmness.
Since the optimal value function is in general nondifferentiable and KKT constraints have MPEC-structure, the combined reformulation is a nonsmooth MPEC. This special structure allows us to adapt some constraint qualifications and necessary optimality conditions from MPEC theory using disjunctive form of the combined reformulation. An example shows, that some of the proposed constraint qualifications can be fulfilled.

Keywords: bilevel programming, value function reformulation, KKT reformulation, constraint qualifications, optimality conditions.

1. Introduction

Bilevel programming problem is a hierarchical optimization problem, which was introduced in 1973 by Bracken and McGill and nowadays is being intensively researched (see [4, 2, 19] and references therein). This problem consists of the upper level problem, whose feasible set contains optimal solutions of the parametric programming problem in the lower level. There are two approaches to deal with this problem, which is in general a set valued optimization problem: pessimistic and optimistic approach. The problem considered in this paper is related to the optimistic bilevel programming problem, because they have the same global solutions and the relationship between local solutions could also be stated under additional assumptions (see [5]). The following problem will be considered:

$$\begin{aligned} F(x,y) &\longrightarrow \min_{x,y} \\ G(x,y) &\leq 0, \\ y &\in \Psi(x), \end{aligned} \tag{1}$$

where $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$ and

$$\Psi(x) = \arg\min_y \left\{ f(x,y) : \ g(x,y) \leq 0 \right\},$$

with $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$. We assume in this paper that the functions F, G, f, and g are in general nonlinear, twice differentiable.

In order to deal with this problem it is common to reformulate it into a one level optimization problem. There are two most popular ways to do this. We can apply the optimal value function of the lower level problem

$$V(x) = \inf_y \left\{ f(x,y) : g(x,y) \leq 0 \right\} \tag{2}$$

and, using the following restriction in the upper level

$$f(x,y) - V(x) \leq 0, \tag{3}$$

together with the feasibility condition of the lower level problem ($g(x,y) \leq 0$), the lower level problem can be replaced. Another possibility is to consider the KKT optimality conditions of the lower level problem

$$\begin{aligned} \nabla_y f(x,y) + \sum_{i=1}^{p} \lambda_i \nabla_y g_i(x,y) &= 0, \\ g(x,y) &\leq 0, \\ \lambda &\geq 0, \\ \lambda^\top g(x,y) &= 0, \end{aligned} \tag{4}$$

instead of the lower level problem.

The second approach is mostly used for the bilevel programming problems with convex lower level problem, since then the KKT conditions are not only necessary but also sufficient for the lower level. However, the local solutions of KKT reformulation do not have to be local minimal with respect to the initial bilevel programming problem even if the lower level problem is convex (see [6]). For a nonconvex lower level problem global solutions of KKT reformulation do not need to coincide with solutions of the bilevel problem. In [19] Ye and Zhu proposed a new reformulation of a bilevel programming problem, which combines both described approaches. Using this combined reformulation we can transform the problem (1) as follows:

$$\begin{aligned} F(x,y) &\longrightarrow \min_{x,y,\lambda} \\ G(x,y) &\leq 0, \\ f(x,y) - V(x) &\leq 0, \\ \nabla_y f(x,y) + \sum_{i=1}^{p} \lambda_i \nabla_y g_i(x,y) &= 0, \\ g(x,y) &\leq 0, \\ \lambda &\geq 0, \\ \lambda^\top g(x,y) &= 0, \end{aligned} \tag{5}$$

where $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R},\ G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k,\ f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R},\ g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$.

It is obvious, that if the lower level problem is convex and consequently the KKT optimality conditions are sufficient, then one part of the restrictions of this problem (restrictions (4) or (3)) is redundant. All constraints are only useful if the lower level problem is not convex, because in this case they all can restrict the feasible set of the upper level problem.

The optimal value function $V(x)$ is assumed to be local Lipschitz continuous, what can be fulfilled under not too strong assumptions on the lower level problem (see [4, 12]).

Because of the last assumption we use several tools, which approximate local Lipschitz continuous function in a neighbourhood of a considered point. These include Clarke directional derivative:

$$f^\circ(x,d) = \limsup_{x' \to x,\, \lambda \to 0^+} \frac{f(x' + \lambda d) - f(x')}{\lambda}, \tag{6}$$

where $x',\ x,\ d \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$, Michel–Penot directional derivative:

$$f^\diamond(x,d) = \sup_{z \in \mathbb{R}^n} \limsup_{\lambda \to 0^+} \frac{f(x + \lambda d + \lambda z) - f(x + \lambda z)}{\lambda}, \tag{7}$$

where $z,\ x,\ d \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ and Dini directional derivative:

$$f^\uparrow(x,d) = \limsup_{\lambda \to 0^+} \frac{f(x + \lambda d) - f(x)}{\lambda}, \tag{8}$$

where $x,\ d \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.

Using the directional derivatives (6) and (7) it is possible to define Clarke (9) and Michel–Penot subdifferential (10):

$$\partial_C f(x) = \{\xi \in \mathbb{R}^n : f^\circ(x,d) \geq \langle \xi, d \rangle \ \text{ for all } d \in \mathbb{R}^n\}, \tag{9}$$

$$\partial_\diamond f(x) = \{\xi \in \mathbb{R}^n : f^\diamond(x,d) \geq \langle \xi, d \rangle \ \text{ for all } d \in \mathbb{R}^n\}. \tag{10}$$

For Lipschitz continuous functions the following relationships between the introduced tools can be stated:

$$f^\diamond(x,d) \leq f^\circ(x,d), \tag{11}$$

$$\partial_\diamond f \subseteq \partial_C f.$$

For that reason optimality conditions using Michel–Penot subdifferential are more restrictive than the same conditions with Clarke subdifferential.

For the function $G : \mathbb{R}^n \to \mathbb{R}$ the gradient $\nabla G(x)$ is a column vector. If the function is defined as follows: $G : \mathbb{R}^n \to \mathbb{R}^p$, then $\nabla G(x) \in \mathbb{R}^{p\times n}$ denotes its Jacobian.

In the second section we consider some important properties of the combined reformulation, which include above all equivalence between the problem (1) and the combined reformulation (5) as equivalence of the local and global solutions.

Section 3 involves consideration of the combined reformulation as a nonsmooth mathematical program with equilibrium constraints (MPEC) including definition of some regularity conditions and stationarity conditions. An example demonstrates satisfiability of these conditions.

In order to show the desired effects as clear as possible, we use simple examples involving mostly linear constraints. More general examples can be found, but the calculations are usually much more difficult.

2. Properties of the combined reformulation

One of the most important questions is whether the optimal solutions of the combined reformulation and the solutions of the initial bilevel programming problem are equivalent. This question will now be considered concerning separately global and local solutions of both optimization problems (cf. [19]).

THEOREM 1. *Assume that $(\bar{x}, \bar{y})$ is a global solution of bilevel programming problem (1) and the KKT conditions hold for the lower level problem at $(\bar{x}, \bar{y})$. Then the point $(\bar{x}, \bar{y}, \bar{\lambda})$ is a global optimal solution of the combined reformulation (5) for every $\bar{\lambda} \in \mathbb{R}^p$, $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y})$, where*

$$\Lambda(\bar{x}, \bar{y}) = \left\{ \lambda \geq 0 : \nabla_y f(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \lambda_i \nabla_y g_i(\bar{x}, \bar{y}) = 0, \lambda^\top g(\bar{x}, \bar{y}) = 0, g(\bar{x}, \bar{y}) \leq 0 \right\}.$$

Conversely, let $(\bar{x}, \bar{y}, \bar{\lambda})$ be a global solution of the combined reformulation (5) and the KKT conditions hold for the lower level problem at all points (x, y) with $y \in \Psi(x)$, then the point $(\bar{x}, \bar{y})$ is a global solution of the bilevel programming problem (1).

The proof is analogous to proof in [19].

In order to formulate a theorem regarding local solutions of problems (1) and (5) we need some additional assumptions. Definition of the term “local solution” has also a great meaning. Let us consider the following example:

EXAMPLE 1. The bilevel programming problem is defined as follows:

$$F(x,y) = y_2 - x \to \min_{x,y}$$
$$y = (y_1, y_2)^\top \in \Psi(x),$$

with $\Psi(x) = \arg\min_y \{-y_1^3 : x(y_1-1) - y_2 \le 0,\ x(y_1-1) + y_2 \le 0,\ y_1 - x - 1 \le 0\}$. We can observe, that for $x < 0$ the feasible set of the lower level problem is empty, therefore we restrict our consideration to the case when $x \ge 0$. Regarding $x \ge 0$ we can observe, that the optimal solution of the lower level problem is a KKT point $\bar{y}^\top = (1,0)$ and for this reason if $x \to \infty$, $(\bar{x}, \bar{y})$ is always feasible for the bilevel programming problem, but $F(\bar{x}, \bar{y}) \to -\infty$, i.e. there exist neither global nor local solution of this problem.

Let us now consider the combined reformulation of this problem. Because of the fact, that for every solution of the lower level problem we have $\bar{y}_1 = 1$, it follows, that $V(x) = -1$ (for $x \ge 0$). The new reformulation has the following form:

$$\begin{aligned} F(x,y) = y_2 - x &\to \min_{x,y,\lambda} \\ -y_1^3 + 1 &\le 0, \\ \begin{pmatrix} -3y_1^2 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} x \\ -1 \end{pmatrix} + \lambda_2 \begin{pmatrix} x \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= 0, \\ x(y_1-1) - y_2 &\le 0, \\ x(y_1-1) + y_2 &\le 0, \\ y_1 - x - 1 &\le 0, \\ \lambda_1(x(y_1-1) - y_2) &= 0, \\ \lambda_2(x(y_1-1) + y_2) &= 0, \\ \lambda_3(y_1 - x - 1) &= 0, \\ \lambda_1, \lambda_2, \lambda_3 &\ge 0. \end{aligned}$$

This problem has the following feasible set:

$$M(x) = \left\{ \left(x, 1, 0, \frac{3}{2x}, \frac{3}{2x}, 0 \right) \ \text{ for } \ x > 0,\ (0,1,0,\lambda_1,\lambda_1,3) \ \text{ with } \lambda_1 \ge 0 \right\}.$$

We can notice, that the constraint $-y_1^3 + 1 \le 0$ is not redundant and restricts the feasible set of the reformulation.

For this problem we have also no global solution, but if we consider a neighbourhood of the point $(\bar{x}, \bar{y}, \bar{\lambda}) = (0,1,0,\lambda_1,\lambda_1,3)$ i.e. $\mathbb{B}_{\frac{1}{2}}(\bar{x}, \bar{y}, \bar{\lambda})$ ($\mathbb{B}_{\frac{1}{2}}(z)$ denotes the ball of radius $\frac{1}{2}$ and centre at the point $\bar{z}$), than due to complementarity conditions we have $y_1 = x + 1$, which is satisfied only if $x = 0$, so the point $(\bar{x}, \bar{y}, \bar{\lambda})$ is a local solution of the combined reformulation. On the other hand if we define a neighbourhood of $(\bar{x}, \bar{y}, \bar{\lambda})$ as $\mathbb{B}_\delta(0,1,0) \times \mathbb{R}^3$, then we cannot find any δ such that $(\bar{x}, \bar{y}, \bar{\lambda})$ is local optimal for the combined reformulation.

In the next theorem we consider locality of solutions only with respect to the variables x and y:

Theorem 2. *[19] Let $(\bar{x}, \bar{y})$ be a local solution of the problem (1). Suppose, that the KKT conditions hold for the lower level problem at $(\bar{x}, \bar{y})$. Then there exists $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y})$ such that $(\bar{x}, \bar{y}, \bar{\lambda})$ is a local solution of the combined reformulation (5). Conversely, let $(\bar{x}, \bar{y}, \bar{\lambda})$ be a solution to the combined reformulation (5) in the neighbourhood $U(\bar{x}, \bar{y}) \times \mathbb{R}^p$ and the KKT conditions hold for the lower level problem at all points $(x, y) \in U(\bar{x}, \bar{y})$ with $y \in \Psi(x)$, then the point $(\bar{x}, \bar{y})$ is a local solution of the bilevel programming problem (1).*

We can notice, that in the converse statement of the Theorem 2 a global solution of the combined reformulation in the neighbourhood of the point $(\bar{x}, \bar{y}, \bar{\lambda})$ defined as $U(\bar{x}, \bar{y}) \times \mathbb{R}^p$ is considered. Compared with common notion of the local solution of a programming problem, the neighbourhood of the point $(\bar{x}, \bar{y}, \bar{\lambda})$ is not restricted with respect to all variables.

It is also worth to note that the set of KKT multipliers $\Lambda(x, y)$ in Example 1 is not upper semicontinuous in sense of Berge [1]. For $x \to 0^+$ it follows $\lambda_1, \lambda_2 \to \infty, \lambda_3 = 0$ but at $x = 0$ we have $\lambda_1, \lambda_2 \in \mathbb{R}_+$, $\lambda_1 = \lambda_2$, $\lambda_3 = 3$. That is the reason why the local optimal solutions of the combined reformulation with respect to x, y and λ do not coincide with local solutions of the initial bilevel programming problem.

If we assume that a Mangasarian Fromowitz Constraint Qualification (MFCQ) is satisfied at the local solution of the lower level problem we know, that the set of KKT multipliers is compact for all points in a neighbourhood of $(\bar{x}, \bar{y})$ and that $\Lambda(x, y)$ is upper semicontinuous [15]. (Notice, that in Example 1 MFCQ does not hold only at the point (0,1,0).) But even if MFCQ is satisfied for the lower level problem at the considered point, we still need some additional assumptions for the local solutions of the bilevel programming problem and the local solutions of the combined reformulation (with respect to all variables including λ) to be equivalent. In this case in order to know wether this local solutions coincide, we have to examine if $(\bar{x}, \bar{y}, \bar{\lambda})$ is local optimal for the combined reformulation for all $\lambda \in \Lambda(\bar{x}, \bar{y})$. The following result has been adapted from [6].

Proposition 1. *Assume that MFCQ is satisfied for the lower level problem at the point $(\bar{x}, \bar{y})$ and let $(\bar{x}, \bar{y}, \bar{\lambda})$ be a local solution of the problem (5) for all $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y})$. Then the point $(\bar{x}, \bar{y})$ is local optimal for the bilevel programming problem (1).*

The proof is analogous to the proof in [6], we only need to use MFCQ instead of the Slater Constraint Qualification due to lack of convexity of the lower level problem.

Obviously it follows from the the Proposition 1, that not every local solution of the combined reformulation is also local optimal for the bilevel programming problem.

The following example (see [6] slightly modified) illustrates the Proposition 1.

Example 2. Consider following (nonconvex) lower level problem:

$$-y^3 \longrightarrow \min_y$$
$$x + y \leq 1,$$
$$-x + y \leq 1.$$

It can be easily noticed, that the optimal solution of this problem, depending on x, can be stated as follows:

$$y(x) = \begin{cases} x+1 & \text{for } x \leq 0, \\ -x+1 & \text{for } x \geq 0. \end{cases}$$

We can also determine the set of KKT multipliers for each KKT point:

$$\Lambda(x,y) = \begin{cases} \left\{(3y^2, 0)\right\}; \text{ if } x > 0, y = -x+1 \\ \left\{(0, 3y^2)\right\}; \text{ if } x < 0, y = x+1 \\ \{(0,0)\}; \text{ if } x \in (-1,1), y = 0 \\ \operatorname{conv}\{(3,0),(0,3)\}; \text{ falls } x = 0, y = 1. \end{cases}$$

The bilevel programming problem has the following form:

$$(x-1)^2 + (y-1)^2 \longrightarrow \min_{x,y}$$
$$y \in \Psi(x).$$

This optimization problem has a global optimal solution at the point $(\bar{x}, \bar{y}) = (0.5, 0.5)$ and no local solutions.

Let us now consider the combined reformulation (5) of this problem with optimal value function $V(x) = -y(x)^3$:

$$(x-1)^2 + (y-1)^2 \rightarrow \min_{x,y,\lambda}$$
$$-y^3 + (y(x))^3 \leq 0,$$
$$-3y^2 + \lambda_1 + \lambda_2 = 0,$$
$$x + y \leq 1,$$
$$-x + y \leq 1,$$
$$\lambda_i \geq 0, \ i = \{1,2\},$$
$$\lambda_1(x + y - 1) + \lambda_2(-x + y - 1) = 0.$$

In this case there exist not only one global solution $(\bar{x}, \bar{y}, \bar{\lambda}) = (0.5, 0.5, (0.75, 0))$ but also a local solution at the point $(\hat{x}, \hat{y}, \hat{\lambda}) = (0, 1, (0, 3))$. This can be noticed if we choose an appropriate neighbourhood of the point $(\hat{x}, \hat{y}, \hat{\lambda})$ with $\lambda_2 > 0$, for example $U = \mathbb{B}_1(0, 1, (0, 3))$. This implies, that for all feasible points in this neighbourhood due to complementarity condition the second restriction of the lower level problem need to be active. For that reason we cannot find in this neighbourhood any feasible point (x, y, λ) with $F(x, y) < F(\hat{x}, \hat{y})$.

Of course it is difficult to check if one point is local optimal for a combined reformulation for every $\lambda \in \Lambda(\bar{x}, \bar{y})$. It is even possible to find an example (cf. [6] with convex lower level problem) such that for only one KKT multiplier $\hat{\lambda} \in \mathrm{int}\Lambda(\bar{x}, \bar{y})$ the point $(\bar{x}, \bar{y}, \hat{\lambda})$ is not local optimal for the combined reformulation. Hence it is also not a local solution of the initial bilevel problem. For that reason it is important to consider, wether it is possible to restrict the set of regarded KKT multipliers under appropriate assumptions. Some regularity conditions such as Constant Rank Constraint Qualification (CRCQ) (cf. [9]) are helpful in order to achieve a better result. We can state the following conclusion similar to the result from [6].

Corollary 1. *Assume that at the point $(\bar{x}, \bar{y})$, $\bar{y} \in \Psi(\bar{x})$ both MFCQ and CRCQ are satisfied for the lower level problem. Moreover, let $(\bar{x}, \bar{y}, \bar{\lambda})$ be a local solution of the combined reformulation (5) for all vertices $\bar{\lambda}$ of the set of KKT multipliers $\Lambda(\bar{x}, \bar{y})$. Then the point $(\bar{x}, \bar{y})$ is a local solution of the bilevel programming problem.*

If we consider Example 2, then we can state, that both MFCQ an CRCQ are satisfied at the point $(0, 1)$ for the lower level problem and that is why we need only to check if two points $(0, 1, (0, 3))$ and $(0, 1, (3, 0))$ are local solutions of the combined reformulation. We can easily state, that the point $(0, 1, (3, 0))$ is not local optimal for the new reformulation and consequently it is also not local solution of the initial bilevel programming problem.

3. Combined reformulation as a nonsmooth MPEC

After equivalence considerations we now need to find out which constraint qualifications can be satisfied for this problem and how to define the necessary optimality conditions.

Since combined reformulation has a structure of a nonsmooth MPEC (see Section 1) it is possible to apply results from Movahedian and Nobakhtian [13] to obtain some regularity and stationarity conditions.

At first we define the combined reformulation in disjunctive form (cf. [13,8]):

$$\begin{aligned} F(x, y) &\to \min_{x,y,\lambda} \\ H(x, y, \lambda) &\in \Gamma, \end{aligned} \tag{12}$$

where all restrictions are aggregated to one function:

$$\begin{aligned} H(x, y, \lambda) := \Big(&G_{I_G}(x, y), f(x, y) - V(x), \nabla_y f(x, y) + \sum_{i=1}^{p} \lambda_i \nabla_y g_i(x, y), \\ &-g_\alpha(x, y), \lambda_\alpha, -g_\beta(x, y), \lambda_\beta, -g_\gamma(x, y), \lambda_\gamma \Big), \end{aligned} \tag{13}$$

where

$$I_G := I_G(\bar{x},\bar{y}) := \{i = 1,...,k : G_i(\bar{x},\bar{y}) = 0\}$$

and the right hand side of the constraints is defined as follows:

$$\Gamma := \bigcup_{(\beta_1,\beta_2)\in\mathcal{P}(\beta)} \Gamma_{\beta_1,\beta_2}. \tag{14}$$

In order to define the complementarity constraint $\lambda^\top g(x,y) = 0$ we need to introduce the following index sets:

$$\begin{aligned} \alpha := \alpha(\bar{x},\bar{y},\bar{\lambda}) &:= \left\{i : g_i(\bar{x},\bar{y}) = 0,\ \bar{\lambda}_i > 0\right\}, \\ \beta := \beta(\bar{x},\bar{y},\bar{\lambda}) &:= \left\{i : g_i(\bar{x},\bar{y}) = 0,\ \bar{\lambda}_i = 0\right\}, \\ \gamma := \gamma(\bar{x},\bar{y},\bar{\lambda}) &:= \left\{i : -g_i(\bar{x},\bar{y}) > 0,\ \bar{\lambda}_i = 0\right\}. \end{aligned} \tag{15}$$

The index set β, for which both $g_i(x,y)$ as well as λ_i are active, can be divided into two partitions in many various ways. The set of all partitions of β is defined as follows:

$$\mathcal{P}(\beta) := \{(\beta_1,\beta_2) : \beta_1 \cup \beta_2 = \beta,\ \beta_1 \cap \beta_2 = \emptyset\}. \tag{16}$$

With the aid of this set we can define the convex polyhedra:

$$\Gamma_{\beta_1,\beta_2} := \mathbb{R}_-^{|I_G|} \times \mathbb{R}_- \times 0_m \times 0_{|\alpha|} \times \mathbb{R}_+^{|\alpha|} \times \Delta_{\beta_1,\beta_2} \times \Delta_{\beta_2,\beta_1} \times \mathbb{R}_+^{|\gamma|} \times 0_{|\gamma|}, \tag{17}$$

where $|A|$ denotes the cardinality of the set A and

$$(\Delta_{\mu,\nu})_j := \begin{cases} 0 & : j \in \mu, \\ \mathbb{R}_+ & : j \in \nu. \end{cases} \tag{18}$$

Notice that this disjunctive problem depends on vector $(\bar{x},\bar{y},\bar{\lambda})$ and hence it is equivalent to the combined reformulation (i.e. the feasible sets of these two problems are equal) in a neighbourhood of the considered point $(\bar{x},\bar{y},\bar{\lambda})$.

There are several constraint qualifications like LICQ or MFCQ, that were adapted for MPEC (cf. [7]): MPEC-LICQ, MPEC-SMFCQ, piecewise MPEC-MFCQ. The latter one is also the weakest of this group of regularity conditions. For this reason we define this condition for a nonsmooth MPEC (see [10]):

DEFINITION 1. *Consider the following optimization problem:*

$$\begin{aligned} f(z) &\longrightarrow \min_z \\ G_i(z) = 0,\ i \in \alpha,\ H_i(z) &= 0\ i \in \gamma, \\ G_i(z) \geq 0,\ \ H_i(z) &= 0\ i \in \beta_1, \\ G_i(z) = 0,\ \ H_i(z) &\geq 0\ i \in \beta_2, \\ g(z) \leq 0,\ h(z) &= 0. \end{aligned} \tag{19}$$

where f, g, G, H are local Lipschitz continuous. The piecewise MPEC-NMFCQ holds at the feasible point $\bar{z}$ if NMFCQ is satisfied at this point for every partition $(\beta_1,\beta_2) \in \mathcal{P}(\beta)$.

NMFCQ as a nonsmooth version of MFCQ can be found in [4].

It turns out, that not only constraint qualifications like nonsmooth LICQ or MFCQ cannot be satisfied for the combined reformulation [19], but also the nonsmooth MPEC counterparts of this conditions fail to be fulfilled at any feasible point of (5). This effect was noticed simultaneously by Ye in [17].

To establish this result we need the well known property of MFCQ, which holds also for NMFCQ:

LEMMA 1. *[18] Let x be a feasible point of the following problem:*

$$\begin{aligned} f \longrightarrow \min_x \\ g(x) \leq 0, \\ h(x) = 0, \end{aligned} \tag{20}$$

with $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^p$, $h : \mathbb{R}^n \to \mathbb{R}^q$ local Lipschitz continuous. Then NMFCQ is satisfied at this point if and only if the set of abnormal multipliers $\Lambda^0(x) = \{0\}$, with

$$\Lambda^0(x) = \left\{(\lambda, \mu) \in \mathbb{R}^p_+ \times \mathbb{R}^q : 0 \in \partial_C g(x)^\top \lambda + \partial_C h(x)^\top \mu, \lambda^\top g(x) = 0\right\}. \tag{21}$$

THEOREM 3. *[14] Let $(\bar{x}, \bar{y}, \bar{\lambda})$ be a feasible point of the optimization problem (5), then for every partition of the index set β there exists a nonzero abnormal multiplier.*

This result can be shown using the idea, that having a feasible point of the combined reformulation $(\bar{x}, \bar{y}, \bar{\lambda})$, the optimal solution of the following optimization problem:

$$\begin{aligned} f(x, y) - V(x) \longrightarrow \min_{x,y} \\ g(x, y) \leq 0 \end{aligned}$$

is the point $(\bar{x}, \bar{y})$. Taking now the Fritz–John multipliers for this problem, we can easily find abnormal multipliers for the combined reformulation (see also [14,19]). Consequently it can be stated, that the piecewise MPEC-NMFCQ cannot be satisfied because of the restriction including the optimal value function.

As this regularity condition fail to be satisfied at any feasible point of the combined reformulation, it is important to check if some weaker regularity conditions can be satisfied for this problem. The next definition involves some regularity conditions adapted from [13], that are one of the weakest for the general nonlinear optimization problems: Abadie and Guignard Constraint Qualifications. These conditions are defined with the aid of the following cones, which are adapted for the disjunctive problem (cf. [16,11,13]):

- Tangent cone (Bouligand cone):

$$\begin{aligned} T_B((\bar{x}, \bar{y}, \bar{\lambda}), H^{-1}(\Gamma)) = \big\{d \in \mathbb{R}^{n+m+p} : \exists\, t_k \downarrow 0,\ \exists d^k \to d \ \text{ with} \\ \forall k : (\bar{x}, \bar{y}, \bar{\lambda}) + t_k d^k \in H^{-1}(\Gamma)\big\}, \end{aligned} \tag{22}$$

- Linearized tangent cone:

$$T^{lin}_{MPEC}(\bar{x},\bar{y},\bar{\lambda}) := \big\{(d,v) \in \mathbb{R}^{n+m} \times \mathbb{R}^p : \nabla G_{I_G}(\bar{x},\bar{y})d \le 0, \quad (23)$$
$$\nabla f(\bar{x},\bar{y})^\top d - V^\uparrow(\bar{x}; d_x) \le 0,$$
$$\nabla(\nabla_y f + \sum_{i=1}^p \lambda_i \nabla_y g_i)(\bar{x},\bar{y})d + \nabla_y g(\bar{x},\bar{y})^\top v = 0,$$
$$\nabla g_\alpha(\bar{x},\bar{y})d = 0,$$
$$v_\gamma = 0,$$
$$(-\nabla g_\beta(\bar{x},\bar{y})d, v_\beta) \in \mathcal{C}\big\},$$

with $\mathcal{C}$ defined as follows:

$$\mathcal{C} := \left\{(a,b) \in \mathbb{R}^{2|\beta|} : a \ge 0, b \ge 0, a^\top b = 0\right\} \quad (24)$$

and $\nabla h(x,y) := \begin{pmatrix} \nabla_x h(x,y) \\ \nabla_y h(x,y) \end{pmatrix}$,

- Fréchet normal cone:

$$N_F((\bar{x},\bar{y},\bar{\lambda}), H^{-1}(\Gamma)) = \big\{d \in \mathbb{R}^{n+m+p} : \langle d, (x,y,\lambda) - (\bar{x},\bar{y},\bar{\lambda})\rangle \le 0 \quad (25)$$
$$\forall (x,y,\lambda) \in T_B((\bar{x},\bar{y},\bar{\lambda}),\ H^{-1}(\Gamma))\big\}.$$

Definition 2. *Let $(\bar{x},\bar{y},\bar{\lambda})$ be a feasible solution of the problem (5).*

(i) The nonsmooth Abadie Constraint Qualification for MPEC (MPEC-NACQ) is satisfied at the point $(\bar{x},\bar{y},\bar{\lambda})$ if the following condition holds:

$$T_B((\bar{x},\bar{y},\bar{\lambda}), H^{-1}(\Gamma)) = T^{lin}_{MPEC}(\bar{x},\bar{y},\bar{\lambda}),$$

with H and Γ defined in (13) and (14) and if there exists $K > 0$ such that for every $d \in \mathbb{R}^{n+m+p}$ it follows:

$$\rho_{T^{lin}_{MPEC}(\bar{x},\bar{y},\bar{\lambda})}(d) \le K\Big(\|max\{0, (\nabla G_{I_G}(\bar{x},\bar{y}))d\}\|^2$$
$$+ \left\|max\left\{0, (\nabla f(\bar{x},\bar{y}))^\top d - V^\uparrow(\bar{x}; d_x)\right\}\right\|^2$$
$$+ \left\|\nabla(\nabla_y f + \sum_{i=1}^p \lambda_i \nabla_y g_i)(\bar{x},\bar{y},\bar{\lambda})d\right\|^2 + \|\nabla g_\alpha(\bar{x},\bar{y})d\|^2$$
$$+ \left\|d_{\lambda_\gamma}\right\|^2 + (\rho_C(-\nabla g_\beta(\bar{x},\bar{y})d, d_{\lambda_\beta}))^2\Big)^{\frac{1}{2}}, \quad (26)$$

where $\rho_A(x) = \inf\{\|x - a\| : a \in A\}$ denotes the distance function.

(ii) The nonsmooth Guignard Constraint Qualification for MPEC (MPEC-NGCQ) is satisfied at the point $(\bar{x},\bar{y},\bar{\lambda})$ if it holds:

$$N_F((\bar{x},\bar{y},\bar{\lambda}), H^{-1}(\Gamma)) = (T^{lin}_{MPEC}(\bar{x},\bar{y},\bar{\lambda}))^\circ,$$

(where C° denotes a polar cone to the cone C) and if there exists $K > 0$ such that for every $d \in \mathbb{R}^{n+m+p}$ the inequality (26) is satisfied. Additionally the following condition should be also satisfied:

$$\nabla F(\bar{x},\bar{y})^{\top} d \geq 0 \quad \forall d \in conv\, T_B((\bar{x},\bar{y},\bar{\lambda}), H^{-1}(\Gamma)). \tag{27}$$

(iii) The weak Abadie Constraint Qualification for MPEC (MPEC-WACQ) is satisfied at the point $(\bar{x},\bar{y},\bar{\lambda})$, if there exists $K_1 > 0$ such that for every $d \in \mathbb{R}^{n+m+p}$ it follows:

$$\begin{aligned} 0 \leq \nabla F(\bar{x},\bar{y})^{\top} d + K_1 \Big(&\left\| max\left\{0, (\nabla G_{I_G}(\bar{x},\bar{y}))d\right\}\right\|^2 \\ &+ \left\| max\left\{0, (\nabla f(\bar{x},\bar{y}))^{\top} d - V^{\uparrow}(\bar{x}; d_x)\right\}\right\|^2 \\ + \left\| \nabla(\nabla_y f + \sum_{i=1}^{p} \lambda_i \nabla_y g_i)(\bar{x},\bar{y},\bar{\lambda}) d \right\|^2 &+ \left\| \nabla g_{\alpha}(\bar{x},\bar{y}) d\right\|^2 \\ &+ \left\| d_{\lambda_\gamma}\right\|^2 + (\rho_C(-\nabla g_{\beta}(\bar{x},\bar{y})d, d_{\lambda_\beta}))^2 \Big)^{\frac{1}{2}}. \end{aligned}$$

Under these regularity conditions we can obtain the following stationarity conditions, which were adapted from [13].

DEFINITION 3. *Assume that (x,y,λ) is a local solution of the combined reformulation (5) and let the sets α, β, and γ be defined as in (15). Then the point (x,y,λ) is:*

1. *MPEC M-stationary, if it holds:*

$$\begin{aligned} 0 \in \nabla_x F(x,y) + \sum_{i=1}^{k} \lambda_i^G \nabla_x G_i(x,y) + \lambda^V (\nabla_x f(x,y) - \partial_{\diamond} V(x)) \\ + \sum_{i=1}^{m} \lambda_i^{KKT} \nabla_x (\nabla_{y_i} f + \sum_{j=1}^{p} \lambda_j \nabla_{y_i} g_j)(x,y) + \sum_{i=1}^{p} \lambda_i^g \nabla_x g_i(x,y) \end{aligned} \tag{28}$$

$$\begin{aligned} 0 = \nabla_y F(x,y) + \sum_{i=1}^{k} \lambda_i^G \nabla_y G_i(x,y) + \lambda^V \nabla_y f(x,y) \\ + \sum_{i=1}^{m} \lambda_i^{KKT} \nabla_y (\nabla_{y_i} f + \sum_{j=1}^{p} \lambda_j \nabla_{y_i} g_j)(x,y) + \sum_{i=1}^{p} \lambda_i^g \nabla_y g_i(x,y) \end{aligned}$$

$$0 = \sum_{i=1}^{m} \lambda_i^{KKT} \nabla_{y_i} g(x,y) - \lambda^{\lambda}, \tag{29}$$

$$\lambda^G \geq 0,\ G(x,y)^{\top} \lambda^G = 0,\ \lambda^V \geq 0,\ \lambda_{\gamma}^g = 0,\ \lambda_{\alpha}^{\lambda} = 0, \tag{30}$$

$$(\lambda_i^g > 0 \wedge \lambda_i^{\lambda} > 0) \vee \lambda_i^g \lambda_i^{\lambda} = 0, \quad \forall i \in \beta, \tag{31}$$

for $(\lambda^G, \lambda^V, \lambda^{KKT}, \lambda^g, \lambda^{\lambda}) \in \mathbb{R}^k \times \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^p$.

2. *MPEC strong-stationary, if the conditions (28-30) are satisfied and additionally*
$$\lambda_i^g \geq 0 \wedge \lambda_i^\lambda \geq 0, \quad \forall i \in \beta.$$

The weakest of the constraint qualifications from the Definition 2 is MPEC-WACQ. The following theorem states the M-stationarity of a local optimal point under fulfilment of MPEC-WACQ (cf. [13]).

THEOREM 4. *[14] Let $(\bar{x}, \bar{y}, \bar{\lambda})$ be a local solution of the combined reformulation (5) and assume that the function H and the set Γ are defined as in (13) and (14). If MPEC-WACQ holds at the point $(\bar{x}, \bar{y}, \bar{\lambda})$, then $(\bar{x}, \bar{y}, \bar{\lambda})$ is M-stationary.*

The proof is analogous to proof in [13] (cf. [14]).

Following example shows, that MPEC-NACQ, MPEC-NGCQ and consequently also MPEC-WACQ can be satisfied for the combined reformulation.

EXAMPLE 3. Consider the bilevel optimization problem:

$$\begin{aligned} (x-1)^2 + (y-1)^2 &\to \min_{x,y} \\ -xy &\leq 0, \\ y &\in \Psi(x), \end{aligned}$$

with $\Psi(x) = \arg\min_y \{-y^3 : x + y \leq 1, -x + y \leq -1\}$.

This problem has only one global optimal point $(\bar{x}, \bar{y}) = (1, 0)$ and a local solution at the point $(\hat{x}, \hat{y}) = (0, -1)$.

We are now going to examine if MPEC-NACQ and MPEC-NGCQ are satisfied at the global optimal point of the combined reformulation. Let us consider the combined reformulation of this bilevel programming problem:

$$\begin{aligned} (x-1)^2 + (y-1)^2 &\to \min_{x,y,\lambda} \\ -xy &\leq 0, \\ -y^3 - V(x) &\leq 0, \\ -3y^2 + \lambda_1 + \lambda_2 &= 0, \\ -x + y &\leq -1, \\ x + y &\leq 1, \\ \lambda_1, \lambda_2 &\geq 0, \\ \lambda_1(-x + y + 1) &= 0, \\ \lambda_2(x + y - 1) &= 0, \end{aligned}$$

with $V(x) = \begin{cases} (x-1)^3, & \text{for} \quad x \geq 1, \\ -(x-1)^3, & \text{for} \quad x \leq 1. \end{cases}$

Let us consider the point $(\bar{x}, \bar{y}, \bar{\lambda}) = (1, 0, 0, 0)$. This point is a global solution of this programming problem, the local solution is at the point $(\hat{x}, \hat{y}, \hat{\lambda}) = (0, -1, 0, 3)$.

Because of the fact, that the optimal solution is a separated point with respect to all variables, the tangent cone consists only of a zero vector. Now we need to determine the linearized cone by solving a system of linear inequalities and equalities (23) with $\alpha = \gamma = \emptyset$ and $\beta = \{1, 2\}$:

$$\begin{aligned} T^{lin}_{MPEC}(\bar{x}, \bar{y}, \bar{\lambda}) = \big\{(d_1, d_2, v_1, v_2) \in \mathbb{R}^4 : -d_2 \leq 0, \\ v_1 + v_2 = 0, \\ d_1 - d_2 \geq 0, \quad v_1 \geq 0, \quad v_1(d_1 - d_2) = 0 \\ -d_1 - d_2 \geq 0, \quad v_2 \geq 0, \quad v_2(-d_1 - d_2) = 0\big\} = \\ \big\{(d_1, d_2, v_1, v_2) \in \mathbb{R}^4 : d_1 = d_2 = v_1 = v_2 = 0\big\}. \end{aligned}$$

That means: $T_B((\bar{x}, \bar{y}, \bar{\lambda}),\, H^{-1}(\Gamma)) = T^{lin}_{MPEC}(\bar{x}, \bar{y}, \bar{\lambda})$.

Because of the fact, that in this example the optimal value function is differentiable, we can state, that the condition (26) is satisfied (see [13] for more details). Consequently MPEC-NACQ holds at the point $(\bar{x}, \bar{y}, \bar{\lambda})$.

If we consider MPEC-NGCQ we find out, that because of the definition of the Fréchet normal cone (25) and due to the fact, that $T_B((\bar{x}, \bar{y}, \bar{\lambda}),\, H^{-1}(\Gamma)) = \{0\}$, the condition (27) is trivially satisfied and the condition MPEC-NGCQ holds also at the considered point. Consequently MPEC-WACQ is also satisfied at $(\bar{x}, \bar{y}, \bar{\lambda})$.

Therefore we can state, that the point $(\bar{x}, \bar{y}, \bar{\lambda})$ is MPEC M-stationary with e.g. the following KKT multipliers $(\lambda^G, \lambda^V, \lambda^{KKT}, \lambda^g, \lambda^\lambda) = (0, 0, 0, 1, 1, 0, 0)$.

4. Conclusion

The combined reformulation is a convenient proposal how to deal with bilevel optimization problems if the lower level problem is not assumed to be convex. It is possible to find some regularity conditions and to derive optimality conditions, that can be satisfied for this nonsmooth MPEC. The obtained stationarity conditions without using partial calmness are actually the same as with using this regularity condition (cf. [19]). In the future it would be interesting to examine whether there exists a relationship between partial calmness and MPEC-NACQ, MPEC-WACQ or MPEC-NGCQ.

5. References

[1] Bank B., Guddat J., Klatte D., Kummer B., Tammer K.; *Non-Linear Parametric Optimization*, Akademie-Verlag, Berlin 1982.

[2] Bard J.F.; *Practical Bilevel Optimization: Algorithms and Applications*, Kluwer Academic Publishers, Dordrecht 1998.

[3] Clarke F.H.; *Optimization and Nonsmooth Analysis*, SIAM, Philadelphia 1990.

[4] Dempe S.; *Foundations of Bilevel Programming*, Kluwer Academic Publishers, Dordrecht 2002.

[5] Dempe S., Dutta J.; *Bilevel programming with convex lower level problems,* in: Dempe S., Kalashnikov V. (ed.), *Optimization with Multivalued Mappings: Theory, Applications and Algorithms*, Springer Science+Business Media, LLC, 2006, pp. 51–71.

[6] Dempe S., Dutta J.; *Is Bilevel Programming a Special Case of a Mathematical Program with Complementarity Constraints?,* Mathematical Programming, Ser A. 131(1–2), 2012, pp. 37–48.

[7] Flegel M.L.; *Constraint Qualifications and Stationarity Concepts for Mathematical Programs with Equilibrium Constraints*, Doktorarbeit (PhD thesis), University of Würzburg, 2005. Available via http://opus.bibliothek.uni-wuerzburg.de/volltexte/2005/1245/pdf/dissertation.pdf

[8] Flegel M.L., Kanzow C., Outrata J.V.; *Optimality Conditions for Disjunctive Programs with Application to Mathematical Programs with Equilibrium Constraints,* Set-Valued Analysis 15(2), 2007, pp. 139–162.

[9] Lu S.; *Implications of the constant rank constraint qualification,* Mathematical Programming 126(2), 2011, pp. 365–392.

[10] Luo Z.Q., Pang J.S., Ralph D.; *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge 1996.

[11] Mordukhovich B.S.; *Variational Analysis and Generalized Differentiation I: Basic Theory*, Springer Verlag, Berlin, Heidelberg 2006.

[12] Mordukhovich B.S., Nam N.M.; *Variational Stability and Marginal Functions via Generalized Differentiation,* Mathematics of Operations Research 30(4), 2005, pp. 800–816.

[13] Movahedian N., Nobakhtian S.; *Constraint Qualifications for Nonsmooth Mathematical Programs with Equilibrium Constraints,* Set-Valued and Variational Analysis 17(1), 2009, pp. 63–95.

[14] Pilecka M.; *Eine spezielle Transformation der Zwei-Ebenen-Optimierungsaufgabe*, Masterarbeit (Master's thesis), TU Bergakademie Freiberg 2011.

[15] Robinson S.M.; *Generalized equations and their solutions, part II: applications to nonlinear programming,* Math. Program. Stud. 19, 1982, pp. 200–221.

[16] Rockafellar R.T., Wets R.J.B.; *Variational Analysis. A Series of Comprehensive Studies in Mathematics*, Springer, Berlin, Heidelberg 2009.

[17] Ye J.J.; *Necessary Optimality Conditions for Multiobjective Bilevel Programs,* Mathematics of Operations Research 36(1), 2011, pp. 165–184.

[18] Ye J.J., Zhu D.L.; *Optimality conditions for bilevel programming problems,* Optimization 33(1), 1995, pp. 9–27.

[19] Ye J.J., Zhu D.L.; *New Necessary Optimality Conditions for Bilevel Programs by Combining the MPEC and Value Function Approaches,* SIAM J. on Optimization 20(4), 2010, pp. 1885–1905.