Institute of Computer Science
Academy of Sciences of the Czech Republic

On the computation of relaxed pessimistic solutions to MPECs (revised version)

M. Červinka, C. Matonoha, J.V. Outrata

Technical report No. 1088

December 2009, revised December 2010

Pod Vodárenskou věží 2, 182 07 Prague 8 phone: +420 2 688 42 44, fax: +420 2 858 57 89,
e-mail:e-mail:ics@cs.cas.cz

Institute of Computer Science
Academy of Sciences of the Czech Republic

On the computation of relaxed pessimistic solutions to MPECs (revised version) [1]

M. Červinka, C. Matonoha, J.V. Outrata [2]

Technical report No. 1088

December 2009, revised December 2010

Abstract:

In this paper, we propose a new numerical method to compute approximate and so-called relaxed pessimistic solutions to mathematical programs with equilibrium constraints (MPECs), where the solution map arising in the equilibrium constraints is not single-valued. This method combines two types of existing codes, a code for derivative-free optimization under box constraints, BFO or BOBYQA, and a method for solving special parametric MPECs from the interactive system UFO. We report on numerical performance in several small-dimensional test problems.

Keywords:
MPEC, equilibrium constraints, pessimistic solution, value function, relaxed and approximate solutions.

---

[1]This work was supported by the Grant Agency of the Czech Republic, project No. 201/09/1957, and the institutional research plan No. AV0Z10300504.

[2]M. Červinka (cervinka@utia.cas.cz), J.V. Outrata (outrata@utia.cas.cz): Institute of Information Theory and Automation AS CR, Pod Vodárenskou věží 4, 182 08 Prague 8, Czech Republic; C. Matonoha (matonoha@cs.cas.cz): Institute of Computer Science AS CR, Pod Vodárenskou věží 2, 182 07 Prague 8, Czech Republic.

1 Introduction

In the last twenty years researchers have paid a lot of attention to optimization problems where, among the constraints, there is a special one in the form of a variational inequality or a complementarity problem. One speaks about an *equilibrium constraint*, and the overall optimization problem coined the name *mathematical program with equilibrium constraints* (MPEC). As an early version of an MPEC one can consider the Stackelberg game of two players ([26]), and we use the respective terminology very often also in the MPEC setting.

Let us consider an abstract MPEC in the form

$$\begin{aligned} &\underset{x}{\text{minimize}}\; f(x,y)\\ &\text{subject to}\\ &\qquad y\in S(x)\\ &\qquad x\in\omega. \end{aligned} \tag{1}$$

In (1), $x \in \mathbb{R}^n$ is the strategy of the dominant player called *Leader*, who acts first and aims to minimize his objective f by using strategies from a closed set $\omega \subset \mathbb{R}^n$. The so-called *solution map* $S[\mathbb{R}^n \rightrightarrows \mathbb{R}^m]$, arising in the equilibrium constraint $y \in S(x)$, assigns x the set of possible responses of his opponent(s) called *Follower(s)*. So, $y \in \mathbb{R}^m$ stands for the cumulative strategy of all Followers and S describes their decision rule. Unfortunately, problem (1) is not well-posed, whenever S is not single-valued on ω. Then, namely, the Leader can hardly optimize his choice of x, not knowing the response of his opponent(s).

To avoid this hurdle, in some situations one imposes an additional hypothesis specifying the response of the Follower(s) at those $x \in \omega$, where $S(x)$ is not a singleton. We usually assume that he (they) behave(s) with respect to the Leader's objective either in a *cooperative* or in a *non-cooperative* way. In the former case one speaks about the *optimistic* solution concept in which the MPEC (1) is replaced by a hierarchical optimization problem where, on the upper level, one minimizes the value function

$$\mu(x) := \inf_{y\in S(x)} f(x,y)$$

over $x \in \omega$. This allows us to convert (1) to the (well-defined) optimization problem

$$\begin{aligned} &\underset{x,y}{\text{minimize}}\; f(x,y)\\ &\text{subject to}\\ &\qquad y\in S(x)\\ &\qquad x\in\omega, \end{aligned} \tag{2}$$

provided we accept the fact that (2) may possess more local solutions than the minimization of μ over ω.

In (2) one minimizes f with respect to both variables x and y. A vast majority of the MPEC literature, including the monographs [17], [21] and [9], is devoted mainly to problem (2) and its numerous variants. To introduce its counterpart, the *pessimistic* solution concept, one usually employs the value function $\vartheta[\mathbb{R}^n \to \bar{\mathbb{R}}]$ defined by

$$\vartheta(x) := \sup_{y \in S(x)} f(x, y).$$

A pair $(\hat{x}, \hat{y}) \in \omega \times \mathbb{R}^m$ is declared a (local) *pessimistic* solution to (1), provided

$$\begin{aligned} &\vartheta(\hat{x}) = f(\hat{x}, \hat{y}) \\ &\vartheta(\hat{x}) \leq \vartheta(x) \text{ for all } x \in \mathcal{O} \cap \omega, \end{aligned} \tag{3}$$

where $\mathcal{O}$ is a neighborhood of $\hat{x}$.

Such a pair exists, however, only under special, rather restrictive assumptions on f and S, such as inner semicontinuity of S, cf. [3, Corollary 3.2.2.1]. In numerous papers by Loridan and Morgan (see e.g. [11], [12], [13]), a lot of attention has been paid to various relaxations of condition (3), leading to more workable solution concepts for the non-cooperative case. Such an effort is very important because a non-cooperative behavior of the Follower(s) can frequently be observed in applications.

To illustrate the intrinsical difference between the above two solution concepts, we present an academic example of an MPEC with a multi-valued solution map.

Example 1. Consider the problem

$$\begin{aligned} &\text{minimize } |y| \\ &\text{subject to} \\ &\qquad y \in S(x), \end{aligned} \tag{4}$$

where S (see Figure 1) is the solution map of the *nonlinear complementarity problem* (NCP):

$$\text{For a given } x \text{ find } y \text{ such that } \min\{F(x,y), G(x,y)\} = 0,$$

with

$$F(x,y) = \begin{cases} x & \text{if } y \geq -\frac{3}{2}, \\ x + (y + \frac{3}{2})^2 & \text{otherwise,} \end{cases}$$

$$G(x,y) = (x-1)^2 - y + \frac{1}{2}.$$

The (optimistic) value function μ is discontinuous at $\bar{x} = 0$ which is also the first component of the (unique) optimistic solution of the MPEC (4). On the other hand, the (pessimistic) value function ϑ is continuous everywhere and its minimum is attained at $\hat{x} = 1$. $\triangle$

This example clearly demonstrates that when interested in pessimistic solutions of MPECs one has to analyze primarily the value function ϑ.

Our aim in this paper is to develop a numerical procedure to the computation of an approximate pessimistic solution to (1). As mentioned by Dempe in [8], to find a pessimistic solution to (1) one either has to minimize a discontinuous, implicitly given value function

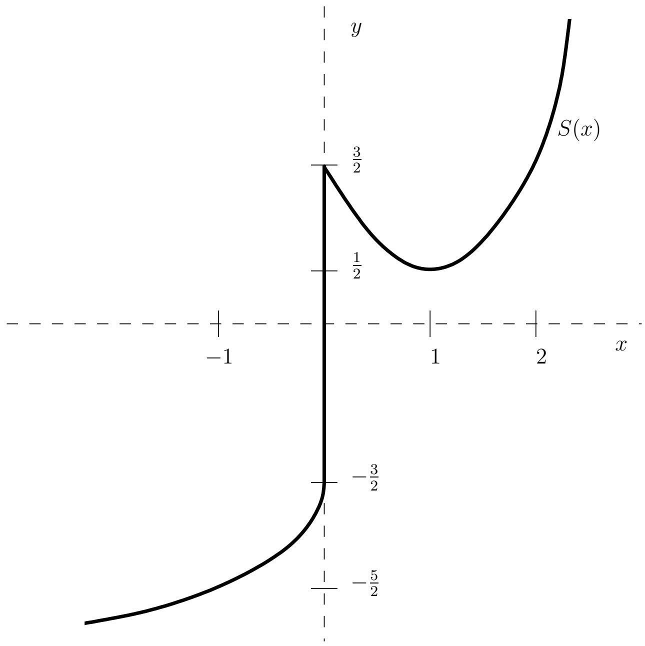

Figure 1: Multifunction S from Example 1.

which is generally not lower semicontinuous, or one has to minimize its special relaxation constructed via a modification of the equilibrium constraint. In this paper we address the first option. To this end we propose a procedure, where we compute the values of ϑ needed for its approximate minimization without using a first-order information.

The plan of the paper is as follows. In the next section we provide a preliminary analysis of the problem and describe two "*relaxed*" solution concepts which are suitable and reasonable to consider when a local pessimistic solution to MPEC does not exist. In Section 3, we give a brief description of our proposed numerical method, particularly its already existing components, BFO [22] and UFO [16]. In the final section we summarize our numerical experience on test MPECs and comment also on combination of BOBYQA [23] and UFO.

The following notation is employed: $\operatorname{dist}_\Omega(\cdot)$ is the distance function to a set Ω. By $x \xrightarrow{\Omega} \bar{x}$ we mean that $x \to \bar{x}$ with $x \in \Omega$ and by $x \xrightarrow{g} \bar{x}$ we mean that $x \to \bar{x}$ with $g(x) \to g(\bar{x})$. For a real-valued function f we use the notation epi f, hypo f, Gph f and $[f \leq a]$ to denote its epigraph, hypograph, graph and level sets, respectively.

For the readers' convenience we now state the definitions of several basic notions from modern variational analysis.

For a set Ω and a point $\bar{x} \in \mathrm{cl}\Omega$, the *Fréchet normal cone* to Ω at $\bar{x}$ is defined by

$$\widehat{N}_\Omega(\bar{x}) := \left\{ x^* \in \mathbb{R}^n \;\middle|\; \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x - \bar{x} \rangle}{\| x - \bar{x} \|} \leq 0 \right\}.$$

The *limiting normal cone* to Ω at $\bar{x}$ is given by

$$N_\Omega(\bar{x}) = \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}_\Omega(x),$$

where the "Lim sup" stands for the Painlevé-Kuratowski upper (or outer) limit. This limit is defined for a set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ at a point $\bar{x}$ by

$$\operatorname*{Lim\,sup}_{x \to \bar{x}} F(x) := \{y \in \mathbb{R}^m \mid \exists x_k \to x, \exists y_k \to y \text{ with } y_k \in F(x_k)\}.$$

For a convex set Ω, both normal cones N_Ω and $\widehat{N}_\Omega$ reduce to the normal cone of convex analysis, for which we use simply the notation N_Ω.

For a function $f[\mathbb{R}^n \to \mathbb{R}]$, and a point $\bar{x} \in \mathbb{R}^n$, the sets

$$\widehat{\partial} f(\bar{x}) = \{y \in \mathbb{R}^n \mid (y, -1) \in \widehat{N}_{\text{epi}f}(\bar{x}, f(\bar{x}))\}$$

and

$$\partial f(\bar{x}) = \{y \in \mathbb{R}^n \mid (y, -1) \in N_{\text{epi}f}(\bar{x}, f(\bar{x}))\}$$

are the (lower) *Fréchet* and the (lower) *limiting subdifferentials* of f at $\bar{x}$, respectively. The *upper Fréchet subdifferential* of f at $\bar{x}$ is given by

$$\widehat{\partial}^+ f(\bar{x}) = \{y \in \mathbb{R}^n \mid (-y, 1) \in \widehat{N}_{\text{hypo}f}(\bar{x}, f(\bar{x}))\}.$$

Given a set-valued mapping $F[\mathbb{R}^n \rightrightarrows \mathbb{R}^m]$ and a point $(\bar{x}, \bar{y})$ from its graph

$$\text{Gph}F := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m | y \in F(x)\},$$

the *Fréchet coderivative* $\widehat{D}^* F(\bar{x}, \bar{y})[\mathbb{R}^m \rightrightarrows \mathbb{R}^n]$ of F at $(\bar{x}, \bar{y})$ is defined by

$$\widehat{D}^* F(\bar{x}, \bar{y})(y^*) := \{x^* \in \mathbb{R}^n | (x^*, -y^*) \in \widehat{N}_{\text{Gph}F}(\bar{x}, \bar{y})\},$$

and the *(normal) coderivative* $D^* F(\bar{x}, \bar{y})[\mathbb{R}^m \rightrightarrows \mathbb{R}^n]$ of F at $(\bar{x}, \bar{y})$ is defined by

$$D^* F(\bar{x}, \bar{y})(y^*) := \{x^* \in \mathbb{R}^n | (x^*, -y^*) \in N_{\text{Gph}F}(\bar{x}, \bar{y})\}.$$

When F is single-valued at $\bar{x}$, we omit $\bar{y}$ in the notation $\widehat{D}^* F(\bar{x}, \bar{y})$ or $D^* F(\bar{x}, \bar{y})$.

Finally, in this paper we use the notion of calmness. A single-valued mapping $f[\mathbb{R}^n \to \mathbb{R}^m]$ is said to be *calm* at $\bar{x}$ with modulus $L \geq 0$ if there is a neighborhood $\mathcal{U}$ of $\bar{x}$ such that

$$|f(x) - f(\bar{x})| \leq L\|x - \bar{x}\| \text{ for all } x \in \mathcal{U}.$$

2 Problem analysis

Before we proceed with relaxations of the pessimistic solution concept, let us start with the application of the upper Fréchet subdifferential condition for local minima under geometric constraints [18, Proposition 5.2] to derive a new type of necessary optimality conditions for pessimistic solutions of (1) only in terms of regular normal cones to Gph$S$ and ω.

Theorem 1. *Let $(\hat{x}, \hat{y})$ be a pessimistic solution to (1) such that $\vartheta(\hat{x})$ is finite, f be Fréchet differentiable at $(\hat{x}, \hat{y})$ and suppose that the map*

$$M(x) = \{y \in S(x) | \vartheta(x) = f(x, y)\}$$

admits a selection that is calm at $(\hat{x}, \hat{y})$. Then one has the following inclusion

$$-\nabla_x f(\hat{x}, \hat{y}) + \widehat{D}^* S(\hat{x}, \hat{y})(-\nabla_y f(\hat{x}, \hat{y})) \subset \widehat{N}_\omega(\hat{x}). \tag{5}$$

Proof. Applying [18, Proposition 5.2] to problem

$$\begin{aligned} &\text{minimize } \vartheta(x) \\ &\text{subject to} \\ &\qquad x \in \omega, \end{aligned} \tag{6}$$

we arrive at inclusions

$$\begin{aligned} \widehat{\partial}(-\vartheta)(\hat{x}) &\subset \widehat{N}_\omega(\hat{x}), \\ \widehat{\partial}(-\vartheta)(\hat{x}) &\subset N_\omega(\hat{x}). \end{aligned}$$

It remains to apply [19, Theorem 2] which yields

$$\widehat{\partial}(-\vartheta)(\hat{x}) = -\nabla_x f(\hat{x}, \hat{y}) + \widehat{D}^* S(\hat{x}, \hat{y})(-\nabla_y f(\hat{x}, \hat{y})).$$

This concludes the proof. □

The optimistic solution of (1) is in general easier to compute. Among possible applications, Theorem 1 can be used to check whether the optimistic solution of (1) satisfying the imposed assumptions, is at the same time also the pessimistic solution of that problem. For illustration, see the following example.

Example 2. Consider the MPEC

$$\begin{aligned} &\underset{x}{\text{minimize}}\ x + y_1 \\ &\text{subject to } y \in S(x), \\ &\qquad\qquad x \in [0, 2], \end{aligned}$$

where

$$\begin{aligned} S(x) &= \left\{ y \in \mathbb{R}^2 \middle| 0 \in \begin{bmatrix} y_2 \\ 2(x-1)^2 - 2y_1 + 3y_2 \end{bmatrix} + N_{\mathbb{R}^2_+}(y) \right\} \\ &= \{y \in \mathbb{R}^2 | 0 \le y_1 \le (x-1)^2, y_2 = 0\}, \end{aligned}$$

see Figure 2.

The optimistic solution of this MPEC is attained at $(\bar{x}, \bar{y}) = (0, 0, 0)$. On the other hand, 0 is not the first component of a pessimistic solution. Indeed, $(0, 1, 0)$ is not a pessimistic solution since $\nabla_x f(0, 1, 0) = 1$, $\widehat{D}^* S(0, 1, 0)(-1, 0) = D^* S(0, 1, 0)(-1, 0) = \{2\}$ and thus (5) yields

$$\{1\} \subset \mathbb{R}_-.$$

The pessimistic solution is attained at $(\hat{x}, \hat{y}) = (1/2, 1/4, 0)$. △

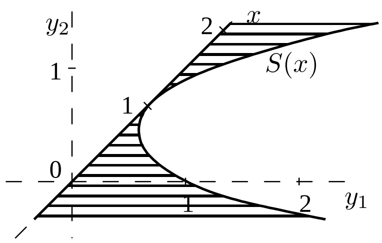

Figure 2: Multifunction S from Example 2.

As we mentioned in the introduction, a (local) pessimistic solution to (1) exists only under restrictive assumptions on problem data. Therefore we consider the following relaxation of the pessimistic solution concept.

Definition 1. *(relaxed pessimistic solution to MPEC)*
The pair $(\hat{x}, \hat{y}) \in \omega \times \mathbb{R}^m$ is called a (local) relaxed pessimistic *solution to (1), provided $\exists x_i \xrightarrow{\omega} \hat{x}, y_i \to \hat{y}, y_i \in S(x_i)$, and a neighborhood $\mathcal{O}$ of $\hat{x}$ such that $\vartheta(x_i) = f(x_i, y_i)$ and $\vartheta(x_i) \to \inf_{x \in \omega \cap \mathcal{O}} \vartheta(x)$.*

Clearly, possible accumulation points $\tilde{y}$ of $\{y_i\}$ do not generally fulfill the relation $\vartheta(\hat{x}) = f(\hat{x}, \hat{y})$ due to the possible lack of continuity of ϑ.

Assume for simplicity throughout the whole sequel that

Assumption 1. $\omega := \{x \in \mathbb{R}^n | a_i \leq x_i \leq b_i\}$, where $a_i, b_i \in \bar{\mathbb{R}}, i = 1, \ldots, n$.

Assumption 2. S is nonempty and convex-valued over ω.

Assumption 3. S is compact-valued and outer semicontinuous over ω, cf [25, Def. 5.4].

In numerous equilibrium problems, namely, the solution map S may have a difficult structure with disconnected images, which would make the computation of the values of ϑ impracticable. Assumption 2 is intended to prevent such situations. Assumption 3 ensures by [2, Theorem 1.4.16] that μ is lower-semicontinuous (lsc) over ω, ϑ is upper-semicontinuous (usc) over ω and that for all $x \in \omega$ one has

$$\mu(x) = \min_{y \in S(x)} f(x, y),$$
$$\vartheta(x) = \max_{y \in S(x)} f(x, y).$$

In the text below we describe a class of equilibria satisfying Assumptions 2 and 3.

Let us denote by $\widehat{\vartheta}$ the lsc regularization of ϑ, i.e., the largest lsc minorant of ϑ. Then it is clear that under the imposed assumption $\hat{x}$ is a relaxed pessimistic solution to (1) if and only if it is a local minimum of $\widehat{\vartheta}$ over ω. In this way $\widehat{\vartheta}(\hat{x})$ provides us with a lower estimate for all values of ϑ at points x near $\hat{x}$ feasible to (1). Further, under the imposed assumptions, this type of "solution" to (1) exists whenever ω is compact.

However, the relaxed pessimistic solutions are typically not pessimistic solutions to MPEC (1). Therefore, the Leader is usually forced to deviate slightly from his relaxed optimal strategy and has to be content with an approximate solution.

Definition 2. *((δ,ε)-pessimistic solution to MPEC)*
*Let $\hat{x}$ be the first component of a relaxed pessimistic solution to (1) and $\delta, \varepsilon > 0$ be given. We say that $(\tilde{x}, \tilde{y}) \in \omega \times \mathbb{R}^m$ is a (δ,ε)-*pessimistic *solution to (1), provided*

$$\begin{aligned} \widehat{\vartheta}(\tilde{x}) &= \vartheta(\tilde{x}) = f(\tilde{x}, \tilde{y}) \\ \widehat{\vartheta}(\tilde{x}) &\leq \vartheta(\hat{x}) + \varepsilon \\ \|\tilde{x} - \hat{x}\| &< \delta. \end{aligned}$$

This notion corresponds to the so-called η-solutions by Loridan and Morgan ([11]) when $\delta = +\infty$. We include a parameter δ to this concept because in some cases the choice of δ directly corresponds to the trust-region radius or the accuracy level for the variables in numerical method described below.

When approximating the first component $\hat{x}$ of a (relaxed) pessimistic solution we are interested in a relationship between parameters δ and ε. Especially important is the case when there is a real $L \geq 0$ such that

$$|\vartheta(x) - \widehat{\vartheta}(\hat{x})| \leq L\|x - \hat{x}\| \tag{7}$$

for all $x \in \omega$ close to $\hat{x}$ with $\vartheta(x)$ close to $\widehat{\vartheta}(\hat{x})$. In the following we show how inequality (7) can be verified in a special class of MPECs.

Suppose that $\omega = \mathbb{R}^n$ and

$$S(x) = \{y | 0 \in F(x,y) + N_C(y)\}, \tag{8}$$

where $C \subset \mathbb{R}^m$ is a convex compact set, $F[\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m]$ is continuously differentiable and $F(x, \cdot)$ is monotone for all x. Then S is a nonempty convex- and compact-valued outer semicontinuous multifunction and thus equilibria, governed by the generalized equation in (8), satisfy both Assumptions 2 and 3.

To enforce the possible satisfaction of inequality (7), however, we will further simplify the structure of S by assuming that F is affine and C is polyhedral. Then we know from [24, Lemma 4], cf also [21, Theorem 2.7], that there exists a finite number, say k, of convex polyhedra Ξ_i such that

$$\operatorname{Gph} S = \bigcup_{i=1}^{k} \Xi_i.$$

For $i = 1, \dots, k$, let us introduce the (polyhedral) sets

$$\Omega_i := \{x | \exists y \in \mathbb{R}^m \text{ such that } (x, y) \in \Xi_i\},$$

and with each $x \in \mathbb{R}^n$ let us associate the index set

$$I(x) := \{i \in \{1, \dots, k\} | x \in \Omega_i\}.$$

It follows that

$$\vartheta(x) = \max_{i \in I(x)} \vartheta_i(x),$$

where

$$\vartheta_i(x) := \max_y \{f(x, y) | (x, y) \in \Xi_i\}.$$

By defining the maps $S_i[\mathbb{R}^n \rightrightarrows \mathbb{R}^m]$ via Gph $S_i = \Xi_i$, we easily infer that $\mathrm{dom} S_i = \Omega_i$ and

$$\vartheta_i(x) = \max_{y \in S_i(x)} f(x, y).$$

By virtue of [25, Example 9.35] S_i happens to be continuous relative to its domain and thus application of [2, Theorem 1.4.16] yields continuity of function ϑ_i relative to Ω_i.

Consider now the first component $\hat{x}$ of a relaxed pessimistic solution of a respective MPEC and put $a := \widehat{\vartheta}(\hat{x})$. It follows that there is a subset of $I(x)$, say $I_0(\hat{x})$, such that

$$(\hat{x}, a) \in \text{Gph } \vartheta_i \text{ for all } i \in I_0(\hat{x}).$$

Since $\vartheta(x) \geq a$ for all x close to $\hat{x}$, to prove (7), in fact, it only suffices to verify that for each $i \in I_0(\hat{x})$, ϑ_i is calm at $(\hat{x}, a)$ *from above*, i.e., the one-sided inequality

$$\vartheta_i(x) \leq a + L_i \|x - \hat{x}\| \;\; \forall x \in \Omega_i \text{ and close to } \hat{x} \tag{9}$$

holds true.

For the calmness of ϑ_i from above we have the following theorem at our disposal, where

$$g(u) := \begin{cases} \mathrm{dist}_{\{\vartheta_i^{-1}(u)\}}(\hat{x}) = \inf\{\|x - \hat{x}\| | \vartheta_i(x) = u\} & \text{if } u > a \\ 0 & \text{otherwise,} \end{cases}$$

and for $u > a$

$$\mathcal{N}(u) := \text{ argmin}\{\|x - \hat{x}\| | \vartheta_i(x) = u\}.$$

Theorem 2. *Let $\hat{x}$ be the first component of a relaxed pessimistic solution of (1) and $i \in I_0(\hat{x})$. Then inequality (9) is fulfilled whenever*

$$0 \notin \mathit{Limsup}_{\substack{u \xrightarrow{g} a \\ g(u) > 0}} \bigcup_{\tilde{x} \in \mathcal{N}(u)} \left\{ \frac{1}{\alpha} \middle| \alpha \frac{\tilde{x} - \hat{x}}{\|\tilde{x} - \hat{x}\|} \in \partial \vartheta_i(\tilde{x}) \right\}. \tag{10}$$

If condition (10) holds for all $i \in I_0(\hat{x})$, then inequality (7) holds true.

Proof. Using the convention that infimum of a function over the empty set amounts to $+\infty$, it is easy to see that the calmness of ϑ_i at $(\hat{x}, a)$ from above (inequality (9)) is equivalent with the condition

$$\operatorname{dist}_{[g\leq 0]}(u) \leq L_i g(u) \;\; \forall u \text{ close to } a. \tag{11}$$

To this end, we can now invoke [10, Theorem 2.1] and conclude that inequality (11) is implied by the condition

$$0 \notin \operatorname*{Limsup}_{\substack{u \xrightarrow{g} a \\ g(u)>0}} \; \partial g(u). \tag{12}$$

As $\|x - \hat{x}\|$ is differentiable for $x \in \vartheta^{-1}(u)$, $u > a$, we obtain from [19, Theorem 7] that

$$\partial g(u) \subset \bigcup_{\tilde{x} \in \mathcal{N}(u)} \left\{ y^* \,\middle|\, \frac{\tilde{x} - \hat{x}}{\|\tilde{x} - \hat{x}\|} \in D^* \vartheta_i(\tilde{x})(y^*) \right\}.$$

Thus, condition (12) is implied by

$$0 \notin \operatorname*{Limsup}_{\substack{u \xrightarrow{g} a \\ g(u)>0}} \bigcup_{\tilde{x} \in \mathcal{N}(u)} \left\{ y^* \,\middle|\, \frac{\tilde{x} - \hat{x}}{\|\tilde{x} - \hat{x}\|} \in D^* \vartheta_i(\tilde{x})(y^*) \right\}.$$

It remains to express the coderivative $D^* \vartheta_i(\tilde{x})$ in terms of the limiting subdifferential $\partial \vartheta_i(\tilde{x})$. Clearly, by the definition of the coderivative,

$$\left\{ y^* \,\middle|\, \frac{\tilde{x} - \hat{x}}{\|\tilde{x} - \hat{x}\|} \in D^* \vartheta_i(\tilde{x})(y^*) \right\} = \left\{ \frac{1}{\alpha} \,\middle|\, \begin{bmatrix} \alpha \frac{\tilde{x} - \hat{x}}{\|\tilde{x} - \hat{x}\|} \\ -1 \end{bmatrix} \in N_{\operatorname{Gph} \vartheta_i}(\tilde{x}, \vartheta_i(\tilde{x})) \right\}.$$

We claim that if a vector of the form $(d, -1)^\top$ belongs to $N_{\operatorname{Gph}\vartheta_i}(\tilde{x}, \vartheta_i(\tilde{x}))$, then d belongs to $\partial \vartheta_i(\tilde{x})$. This follows from the fact that, by the continuity of ϑ_i relative to Ω_i, there exist sequences $x_k \to \tilde{x}$, $d_k \to d$, $\mu_k \to -1$ such that

$$(d_k, \mu_k)^\top \in \widehat{N}_{\operatorname{Gph} \vartheta_i}(x_i, \vartheta_i(x_i)) \;\; \forall k.$$

Using [18, Theorem 1.80] we now infer that for all k sufficiently large one has

$$(d_k, \mu_k)^\top \in \widehat{N}_{epi \vartheta_i}(x_i, \vartheta_i(x_i))$$

and so, consequently,

$$(d, -1)^\top \in N_{epi \vartheta_i}(\tilde{x}, \vartheta_i(\tilde{x})),$$

i.e., $d \in \partial \vartheta_i(\tilde{x})$. The statement has been established. □

Remark. Note that (11) amounts to the (local) *error bound property* of g at a and the set on the right-hand side of (12) is the (limiting) outer subdifferential of g at a introduced in [10].

We illustrate the application of Theorem 2 by means of the following example.

Example 3. Consider the MPEC

$$\begin{aligned} &\underset{x}{\text{minimize}}\ (x^2 - x)y_1 + y_2 \\ &\text{subject to} \\ &\qquad y \in S(x), \end{aligned} \tag{13}$$

where

$$S(x) = \left\{ y \in \mathbb{R}^2 \,\middle|\, 0 \in \begin{bmatrix} x \\ 0 \end{bmatrix} + N_\Omega(y) \right\},$$

with

$$\Omega = \{(y_1, y_2) \in \mathbb{R}^2_+ | y_1 \leq 1, y_1 + 2y_2 \leq 2\}.$$

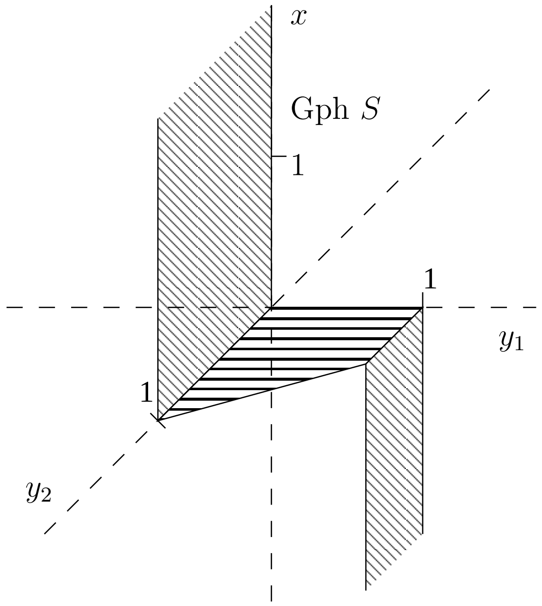

Figure 3: The graph of multifunction S in Example 3.

Clearly, the graph of S consists of three polyhedral pieces (see Figure 3)

$$\begin{aligned} \Xi_1 =& \{(x, y_1, y_2) | x \leq 0, y_1 = 1, y_2 \in [0, 0.5]\}; \\ \Xi_2 =& \{(x, y_1, y_2) | x = 0, (y_1, y_2) \in \Omega\}; \\ \Xi_3 =& \{(x, y_1, y_2) | x \geq 0, y_1 = 0, y_2 \in [0, 1]\} \end{aligned}$$

and, consequently,

$$\vartheta_1(x) = \begin{cases} x^2 - x + \frac{1}{2} & \text{for } x \leq 0; \\ -\infty & \text{otherwise}; \end{cases}$$

$$\vartheta_2(x) = \begin{cases} 1 & \text{for } x = 0; \\ -\infty & \text{otherwise}; \end{cases}$$

$$\vartheta_3(x) = \begin{cases} 1 & \text{for } x \geq 0; \\ -\infty & \text{otherwise}. \end{cases}$$

In this case, one has $\hat{x} = 0$ with $a = \frac{1}{2}$, $I(\hat{x}) = \{1, 2, 3\}$ and $I_0(\hat{x}) = \{1\}$. Since S_1 corresponding to Ξ_1 is continuous relative to $\mathbb{R}_-$, and for $u > \frac{1}{2}$ one has $\mathcal{N}(u) = \sqrt{u - \frac{1}{4}} + \frac{1}{2}$, the set on the right-hand side of (10) amounts to

$$\operatorname*{Limsup}_{\substack{u \searrow \frac{1}{2} \\ u \neq \frac{1}{2}}} \left\{ \frac{1}{\alpha} \,\middle|\, \alpha = 2x - 1, u = x^2 - x + \frac{1}{2} \right\}$$

$$= \operatorname*{Limsup}_{\substack{x \nearrow 0 \\ x \neq 0}} \left\{ \frac{1}{2x - 1} \right\} \cup \operatorname*{Limsup}_{\substack{x \searrow 1 \\ x \neq 1}} \left\{ \frac{1}{2x - 1} \right\} = \{-1, 1\}.$$

Theorem 2 thus yields the calmness of ϑ_1 at $(0, \frac{1}{2})$ from above and we infer that the respective function $\vartheta(\cdot) = \max_{i=1,2,3} \vartheta_i(\cdot)$ satisfies inequality (7). $\triangle$

We have included the above example to our collection of test problems for the numerical method proposed in Section 3. For numerical results see Table 2.

Inequality (7) signalizes a numerically important fact that by decreasing δ we may theoretically compute a (δ, ε)-pessimistic solution, whose objective value is arbitrarily close to the unattainable value $\widehat{\vartheta}(\hat{x})$. Note that the (restrictive) assumptions imposed on F and C are needed only to achieve the favorable disjunctive structure of Gph S with the respective functions ϑ_i continuous relative to Ω_i. Such a structure can be obtained, however, also in other situations.

3 Numerical method

Our aim is to suggest a numerical procedure for the computation of (δ, ε)-pessimistic solutions to (1), an approximation of relaxed pessimistic solution. To this end we split the pessimistically formulated MPEC into the outer and the inner optimization problems.

For solving the inner optimization problem

$$\begin{aligned} &\underset{y}{\text{maximize}}\ f(x, y) \\ &\text{subject to} \\ &\qquad y \in S(x) \end{aligned} \tag{14}$$

with a fixed x we use a suitable optimization method from the interactive system UFO. As explained in the previous sections, the optimal value function of this problem, $\vartheta(x)$, is generally an usc function. Thus, for the outer optimization problem

$$\begin{aligned} &\underset{x}{\text{minimize}}\ \vartheta(x) \\ &\text{subject to} \\ &\qquad\qquad x \in \omega, \end{aligned} \tag{15}$$

we use the code BFO by P. Toint for derivative free minimization of (possibly discontinuous) functions. To find (ε, δ)-pessimistic solutions to MPECs, we have combined these two algorithms into one code. Alternatively, we also replaced BFO by the algorithm BOBYQA developed by M.J.D. Powell. A brief description of these algorithms follows.

UFO [16] is an interactive system for universal functional optimization written in Fortran that serves for solving both dense medium-size and sparse large-scale optimization problems. It can be used for formulation and solution of particular optimization problems, for preparation of specialized optimization routines and for designing and testing new optimization methods. One can generate a large number of modifications of a given method and find the most suitable implementation. The optimization methods can be implemented with various strategies for a step-size selection. It contains line-search methods, general trust-region methods, special trust-region methods for nonlinear least squares, Marquardt-type methods for nonlinear least squares and filter-type methods for nonlinear programming problems including Fletcher-Leyffer filters, barrier filters and Markov filters. Moreover, various direct solvers for different matrix representations and various iterative solvers with different preconditioners can be used for the computation of a descent direction.

The inner problem, typically an optimization problem with variational inequality constraints, is difficult to solve by standard methods since the Mangasarian-Fromowitz constraint qualification is not satisfied at any feasible point. For the numerical experiments described in the next section we have used two approaches. First, we have considered the inner problem as a nonlinear program and used the standard interior-point method [14]. Trust-region realization and line-search approach with suitable restarts were used for the direction determination, cf. [1], [4], [6]. After setting up several parameters, mainly the maximal stepsize and the trust-region radius, we have managed to obtain quite a good solution of the inner problems for fixed x. In the second approach we considered the inner constraints as complementarity constraints and used a recently developed method which is based on the interior-point approach and uses an exact penalty function to remove complementarity constraints, cf. [15].

BFO [22] is a "Brute-Force Optimizer", written in Matlab, for unconstrained or bound-constrained optimization in continuous and/or discrete variables, where the number of variables is small (not larger than 10). The derivatives of the objective are assumed to be unavailable or inexistent. Objective function values and a starting point x^0 must be provided by the user.

The algorithm proceeds by evaluating the objective function at points differing from the current iterate by a positive (forward) and a negative (backward) step in each variable. The corresponding stepsizes are computed on a grid given by varying fractions of the user-specified increments. For continuous variables, these fractions are decreased (yielding a finer grid) as soon as no progress can be made from the current point and until the desired accuracy is reached. For discrete variables, the user-supplied increment may not be reduced.

The algorithm is stopped as soon as no progress can be made from the current iterate by taking forward and backward steps of length of the specified accuracy levels for continuous variables and of length of the specified increments for discrete variables. However, this may be insufficient to guarantee that the computed point is a local minimizer when the objective function is not differentiable.

The Fortran code BOBYQA [23] is a bound-constrained optimization algorithm for computing a local minimum of a function F of several variables. The function values of F can also be specified by a "black box" and the information about its derivatives need not be available.

BOBYQA is based on finding interpolation points $u_1, \dots, u_m$ and computing quadratic approximations Q_k of F that satisfy $Q_k(u_i) = F(u_i),\ i = 1, \dots, m$. At each iteration, a new point $x_{k+1} = x_k + d_k$ is computed and one of the interpolation points, say u_j, is replaced by x_{k+1}. Thus only one interpolation point is altered on each iteration. A direction vector d_k is chosen by minimizing $Q_k(x_k + d)$ subject to the prescribed bounds on variables under the condition $d \leq \Delta_k$, where Δ_k is the current trust-region radius. At each iteration, as a new point of a minimizing sequence x_k^* we take the point which minimizes F among all current interpolation points.

BOBYQA consists of a very accurate and efficient system of updating the approximation models and it maintains a "good" set of interpolation points. This makes BOBYQA numerically very stable and not sensitive to a reasonable level of computational errors in values of the objective. However, BOBYQA does not make use of the problem structure and the established local convergence rate is closer to linear than to quadratic. For this reason, the algorithm sometimes prefers the *early termination*, i.e., it stops when we are still far from an optimal solution but the cost for maintaining the "good" set of interpolation points is too high or the approximation is poor, see [7, Section 1.3].

From the above discussion it is clear that there are no guarantees for convergence either for the combination of BFO and UFO or for the combination of BOBYQA and UFO. If the computational precision in UFO is maintained high enough (by several orders higher than in BFO or BOBYQA), the convergence rate of the combination of codes depends mainly on performance of the derivative-free optimization tool. Then, e.g. in cases when no early termination occurs in BOBYQA, we are able to compute an approximation of the relaxed pessimistic solution by choosing the final trust-region radius as δ.

Our final note is about the special situations when the map S happens to be continuous over ω (in the set-valued sense). Then ϑ is continuous over ω as well and the notions of relaxed pessimistic and (δ, ε)-pessimistic solutions become superfluous. Our proposed procedure will then generate pessimistic solutions in the sense of (3).

4 Numerical experiments

We have performed tests on several examples of small dimension by using the codes BFO and BOBYQA for the outer problem and the UFO system for the inner problem. Examples 4 and 5 refer to the example in [9, Section 5.1], the former being the pessimistic and the latter being the optimistic formulation of the same problem. By including Example 5 in our collection of test problems we intend to show that our proposed method could be used also for computation of optimistic solutions to (1). Example 6 is a simple MPEC from [20] and [21] where the solution map is single-valued and continuous at each point relative to ω and by this example we test whether our method can compute the solutions (in the original sense) to (1). Example 7 is of similar nature as Examples 3 and 4 but with $x \in \mathbb{R}^3$. In Example 8 we propose a more general framework for constructing test problems in the form of pessimistically formulated MPECs of arbitrary dimension.

In all our computations we have considered δ, the final precision level given consecutively by the values $10^{-2}, 10^{-3}, 10^{-4}$ and 10^{-5} and the accuracy level 10^{-7} for computations of the value functions by UFO.

For each reported results on test problems, we also include the number of objective function evaluations `neval` (number of UFO calls), see Tables 2-6. Note that the higher dimension, the more UFO calls is required by BFO and hence this procedure might be untractable for large complicated problems for which each UFO computation takes more then just a fraction of a second.

Example 4. ([9]) (pessimistic formulation)

$$\min_{x\in[-2,2]} \max_{y\in S(x)} x^2 + y^2, \tag{16}$$

where

$$S(x) = \{y | 0 \in -x + N_{[0,1]}(y)\}.$$

For this problem, the pessimistic value function has the form

$$\vartheta(x) = \begin{cases} x^2 & \text{with } y = 0 \text{ for } x < 0; \\ x^2 + 1 & \text{with } y = 1 \text{ for } x \geq 0. \end{cases}$$

We can see that there is no solution of problem (16) in the sense of (3). However, $\hat{x} = 0$ is the first component of the relaxed pessimistic solution. The results are displayed in Table 3. △

Example 5. ([9]) (optimistic formulation)

$$\min_{x\in[-2,2]} \min_{y\in S(x)} x^2 + y^2, \tag{17}$$

where S is the same multifunction as in Example 4. The optimistic value function has in this problem the form

$$\mu(x) = \begin{cases} x^2 & \text{with } y = 0 \text{ for } x \leq 0; \\ x^2 + 1 & \text{with } y = 1 \text{ for } x > 0. \end{cases}$$

Thus its global optimistic solution is attained at $(\bar{x}, \bar{y}) = (0, 0)$. The numerical results are the same as in Example 4 and can also be found in Table 3. This follows from the fact that the respective optimistic and pessimistic value functions differ only at $x = 0$ and thus from the same starting point BFO proceeds in both problems identically, provided it avoids the origin where $\varphi(0) \neq \vartheta(0)$. For comparison, in Table 3 we display the results of both these problems for two different starting points. As expected, the choice of initial point x^0 influences significantly the computation of a solution. △

Example 6. ([20], [21]) (An MPEC with a single-valued solution map at each feasible point)

Consider an *oligopolistic market* model with 5 firms producing a homogeneous product and attempting to maximize their profits; see, e.g., [20] and [21]. Let $x \in \mathbb{R}$ denote the *production* of the Leader and let $y_i \in \mathbb{R},\ i = 1, \ldots, 4$, be the production of the ith Follower.

Let

$$T = x + \sum_{i=1}^{4} y_i$$

denote the *overall production on the market*, and let $p : \operatorname{int} \mathbb{R}_+ \to \operatorname{int} \mathbb{R}_+$ be the so-called *inverse demand curve* that assigns T the price at which consumers are willing to purchase. The MPEC formulation of the problem of the Leader can be written in the form

$$\begin{aligned} &\text{minimize } c_0(x) - xp(T) \\ &\text{subject to } 0 \in F(x, y) + N_{\mathbb{R}^4_+}(y) \\ &\qquad\qquad x \geq 0, \end{aligned}$$

where

$$F(x, y) = \begin{pmatrix} \nabla c_1(y_1) - p(T) - y_1 \nabla p(T) \\ \vdots \\ \nabla c_4(y_4) - p(T) - y_4 \nabla p(T) \end{pmatrix}.$$

Let the production cost functions $c_i, i = 0, \ldots, 4$, be in the form

$$c_i(z) = b_i z + \frac{\beta_i}{1 + \beta_i} K_i^{-\frac{1}{\beta_i}} (z)^{\frac{1+\beta_i}{\beta_i}},$$

where b_i, K_i and $\beta_i,\ i = 0, \ldots, 4$, are positive parameters given by Table 1.

Further, let

$$p(T) = 5000^{\frac{1}{\gamma}} T^{-\frac{1}{\gamma}},$$

with a parameter $\gamma \geq 1$ termed demand elasticity. The numerical results, cf Table 4, can be compared to [21, Table 12.4] where the chosen accuracy is 5×10^{-4}. △

	Leader	Follower 1	Follower 2	Follower 3	Follower 4
b_i	2	8	6	4	2
K_i	5	5	5	5	5
β_i	1.2	1.1	1.0	0.9	0.8

Table 1: Parameter specification for the production costs

Example 7. (A pessimistically formulated MPEC with a relaxed pessimistic solution)

$$\min_{x\in\mathbb{R}^3} \max_{y\in S(x)} \frac{1}{2}\|x\|^2 + \frac{1}{2}\|y\|^2, \tag{18}$$

where

$$S(x) = \{y \in \mathbb{R}^3 | 0 \in By + x + N_{\Delta^3}(y)\}. \tag{19}$$

In (19) Δ^3 is the standard 3-simplex in $\mathbb{R}^3$ and B is the symmetric positive semidefinite matrix

$$\begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{pmatrix}.$$

The solution map S is multi-valued only at $x = (0,0,0)$ which is thus the only point of discontinuity of ϑ. There is no solution of problem (18) in the sense of (3). However, $\hat{x} = (0,0,0)$ is the first component of the relaxed pessimistic solution of (18). For numerical results for Example 7, see Table 5. △

Example 8. (A pessimistically formulated MPEC of an arbitrary dimension with a relaxed pessimistic solution)

$$\min_{x\in\mathbb{R}^n} \max_{y\in S(x)} \frac{1}{2}\|x-a\|^2 + y_1, \tag{20}$$

where

$$\begin{aligned} S(x) =& \{y \in \mathbb{R}^n | 0 \in x + N_C(y), C = \{y | a^\top y = c, 0 \le y \le \xi\}\} \\ =& \arg\min_{y\in\mathbb{R}^n} \{x^\top y | a^\top y = c, 0 \le y \le \xi\}, \end{aligned} \tag{21}$$

$a \in \mathbb{R}^n_+, \xi \in \mathbb{R}^n_+$ and $c \in \mathbb{R}_+$ are given constants. In our numerical simulations, for a chosen dimension $n \ge 3$ of the problem, we randomly generated $a_i > 0, i = 1,\dots,n$, and $c > 0$ and set $\xi_i = \frac{7c}{10a_i}, i = 1,\dots,n$. Clearly, in each such a problem of arbitrary dimension $n \ge 3$ the mapping S is multivalued at $x = a$ which is also the first component of the relaxed pessimistic solution of (20).

Tests on randomly generated pessimistic MPECs (20) were performed with the final precision level $\delta = 10^{-5}$ for dimensions $n = 5, 7$ and 10, respecting the upper bound for dimension of x in BFO. In Table 6 we report results for one randomly generated problem for each chosen dimension.

Observe that this example also fulfills all assumptions imposed in Theorem 2. △

	$x^0 = -0.1$		
accuracy level	$\tilde{x}$	$\vartheta(\tilde{x})$	neval
10^{-2}	-1.3188E-03	5.013205E-01	35
10^{-3}	-4.7665E-04	5.004769E-01	39
10^{-4}	-1.9561E-05	5.000196E-01	44
10^{-5}	-4.8948E-06	5.000050E-01	49

Table 2: Results for Example 3 using BFO and UFO

	$x^0 = 1$			$x^0 = -1$		
accuracy level	$\tilde{x}$	$\vartheta(\tilde{x})$	neval	$\tilde{x}$	$\vartheta(\tilde{x})$	neval
10^{-2}	-2.6602E-03	7.0768E-06	41	-4.8018E-03	2.3057E-05	45
10^{-3}	-3.0889E-04	9.5415E-08	53	-3.4369E-04	1.1812E-07	57
10^{-4}	-2.0862E-05	4.3521E-10	65	-6.7267E-06	4.5248E-11	72
10^{-5}	-2.1996E-06	4.8381E-12	70	-5.1346E-07	2.6364E-13	73

Table 3: Results for Examples 4 and 5 using BFO and UFO

	$x^0 = 150$		
accuracy level	$\tilde{x}$	$\vartheta(\tilde{x})$	neval
10^{-2}	99.532339	958.634749	54
10^{-3}	99.534400	958.634749	70
10^{-4}	99.534471	958.634749	70
10^{-5}	99.534471	958.634749	72

Table 4: Results for Example 6 using BFO and UFO

	$x^0 = (1, 1, 1)$				
accuracy level	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$	$\vartheta(\tilde{x})$	neval
10^{-2}	4.0030E-03	2.8674E-03	7.1296E-04	1.2377E-05	192
10^{-3}	-6.8044E-05	3.1937E-04	-7.1007E-06	5.9399E-08	260
10^{-4}	1.1433E-06	-1.9192E-06	3.1330E-05	4.9512E-10	294
10^{-5}	2.6704E-06	1.2923E-06	2.3969E-06	7.2732E-12	376

Table 5: Results for Example 7 using BFO and UFO

For the combination of BOBYQA and UFO, the choices of values of `RHOBEG` and `RHOEND`, the initial and final values of a trust-region radius for BOBYQA, and of initial point x^0, are crucial. In all our computations we have set the number of interpolation points of BOBYQA to $m = 2n+1$, where n is the dimension of x. Since BOBYQA can be used only for problems with $n \geq 2$, we have introduced, where necessary, an artificial variable which, however, does not enter the objective. Even though BOBYQA was designed primarily for minimization of continuous objectives, for all our test problems we have obtained merely the same (satisfactory) results as with the combination of BFO and UFO. Hence we do not report these results in separate tables here.

	$n = 5$	$n = 7$	$n = 10$
$\vartheta(\tilde{x})$	4.7291E-11	3.5892E-11	1.4816E-10
$\tilde{x}_1 - a_1$	6.6631E-06	2.1495E-06	9.4760E-06
$\tilde{x}_2 - a_2$	7.4962E-07	3.7293E-06	2.0207E-06
$\tilde{x}_3 - a_3$	2.0791E-06	-4.5862E-06	-5.3025E-06
$\tilde{x}_4 - a_4$	-5.5591E-06	7.9849E-07	-1.4967E-06
$\tilde{x}_5 - a_5$	-3.7941E-06	-2.3226E-06	-4.5371E-06
$\tilde{x}_6 - a_6$	-	-3.3658E-06	-1.6041E-07
$\tilde{x}_7 - a_7$	-	-3.8577E-06	3.8176E-06
$\tilde{x}_8 - a_8$	-	-	-2.5545E-06
$\tilde{x}_9 - a_9$	-	-	9.4421E-06
$\tilde{x}_{10} - a_{10}$	-	-	6.3944E-06
$\tilde{y}_1$	2.5960E-16	2.8052E-16	1.6464E-13
neval	729	1450	2198

Table 6: Results for Example 8 using BFO and UFO

The iteration process of BOBYQA is different from that of BFO. We also observed that for a given MPEC, the minimization process of BOBYQA for an usc value function ϑ and a lsc value function φ differs and it is much faster for the latter case. Thus, unlike reported in Table 2, the algorithm composed of BOBYQA and UFO, achieved the prescribed accuracy level in Example 5 in significantly less number of iterations than in Example 4, given the same starting point. A possible explanation may lie in the construction of the quadratic interpolation of the objective.

5 Conclusion

A numerical procedure has been proposed for the computation of approximate pessimistic solutions to a class of MPECs. The two main blocks of this procedure consist of standard codes for derivative-free optimization and for the solution of special MPECs. They may be replaced by different codes serving the same purpose. We have tested our procedure on several small-dimensional academic examples of MPECs.

By using tools of modern variational analysis, a method has been suggested for local analysis of the pessimistic value function ϑ around the relaxed pessimistic solution. This method enables a post-optimal analysis of the behavior of ϑ in simple examples of a special structure. Since local analysis of usc functions is a rather new topic, this result has only a preliminary character. It offers, however, an interesting new research area in variational analysis, not restricted only to the pessimistic solution concept for MPECs.

Acknowledgment

We are grateful to M.J.D. Powell and Ph.L. Toint for their kind help in providing the codes of BOBYQA and BFO. We also thank M. Kočvara for numerous fruitful suggestions and discussions and to Alexandra Schwartz for a careful reading of preliminary versions of our manuscript.

References

[1] A. Antoniou and W.S. Lu, *Practical Optimization*, Springer, 2007.

[2] J.-P. Aubin and H. Frankowska, *Set-valued Analysis*, Birkhäuser, Boston, 1990.

[3] B. Bank, J. Guddat, D. Klatte, B. Kummer and K. Tammer, *Non-linear Parametric Optimization*, Akademie-Verlag, Berlin, 1982.

[4] M.S. Bazaraa, H.D. Sherali and C.M. Shetty, *Nonlinear Programming, Theory and Algorithms*, Wiley, 2006.

[5] C.G. Broyden, *The convergence of a class of double rank minimization algorithms. Part 1 - general considerations. Part 2 - the new algorithm.* J. Inst. Math. Appl. 6 (1970), pp. 76–90, pp. 222–231.

[6] A.R. Conn, N.I.M. Gould and Ph.L. Toint, *Trust-region methods*, SIAM, 2000.

[7] A.R. Conn, K. Scheinberg and L.N. Vicente, *Introduction to Derivative-Free Optimization*, SIAM, 2009.

[8] S. Dempe, *A bundle algorithm applied to bilevel programming problems with non-unique lower level solutions*, Computational Optimization and Applications 15 (2000), pp. 145–166.

[9] S. Dempe, *Foundations of bi-level programming*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

[10] A.D. Ioffe and J. Outrata, *On metric and calmness qualification conditions in subdifferential calculus*, Set-Valued Analysis 16 (2008), pp. 199–227.

[11] P. Loridan and J. Morgan, *New results on Approximate Solutions in Two-Level Optimization*, Optimization 20(6) (1989), pp. 819–836.

[12] P. Loridan and J. Morgan, *ε-regularized two-level optimization problems: approximation and existence results*, Optimization - Fifth French-German Conference Castel Novel 1988, Lecture Notes in Mathematics 1405, Springer Verlag, 1989, pp. 99–113.

[13] P. Loridan and J. Morgan, *Weak Via Strong Stackelberg Problem: New Results*, J. of Global Optimization 8 (1996), pp. 263–287.

[14] L. Lukšan, C. Matonoha, J. Vlček, *Interior-point method for nonlinear nonconvex optimization*, Numerical Linear Algebra with Applications 11 (2004), pp. 431–453.

[15] L. Lukšan, C. Matonoha and J. Vlček, *Interior-Point Method for Nonlinear Programming with Complementarity Constraints*, Technical Report V-1039, Institute of Computer Science, Academy of Sciences of the Czech Republic, 2008.

[16] L. Lukšan, M. Tůma, J. Vlček, N. Ramešová, M. Šiška, J. Hartman and C. Matonoha, *UFO 2008 - Interactive system for universal functional optimization*, Technical Report V-1040, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, 2008.

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

[18] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation*, Volume 1: Theory, Volume 2: Applications, Springer, 2006.

[19] B.S. Mordukhovich, N.M. Nam and N.D. Yen, *Subgradients of marginal functions in parametric mathematical programming*, Mathematical Programming 116 (2009), pp. 369–396.

[20] F.H. Murphy, H.D. Sherali and A.L. Soyster, *A mathematical programming approach for determining oligopolistic market equilibrium*, Mathematical Programming 24 (1982), pp. 92–106.

[21] J.V. Outrata, M. Kočvara and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, Kluwer Academic Publisher, Dordrecht, The Netherlands, 1998.

[22] M. Porcelli and Ph.L. Toint, *BFO, a Brute Force Optimizer for mixed integer nonlinear bound-constrained optimization and its self tuning*, Technical Report, Namur Research Center on Complex Systems (NAXYS), in preparation, 2010.

[23] M.J.D. Powell, *The BOBYQA algorithm for bound constrained optimization without derivatives*, Technical report DAMTP 2009/NA06, Cambridge, 2009.

[24] S.M. Robinson, *An implicit function theorem for generalized variational inequalities*, Technical Summary Report 1672, Mathematics Research Center, University of Wisconsin-Madison, 1976.

[25] R.T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

[26] H. von Stackelberg, *Marktform und Gleichgewicht*, Springer, Berlin, 1934.