Optimization Methods and Software
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/goms20

On the computation of relaxed pessimistic solutions to MPECs

M. Červinka [a], C. Matonoha [b] & J. V. Outrata [a]

[a] Institute of Information Theory and Automation AS CR, Pod Vodárenskou věží 4, 182 08, Prague, 8, Czech Republic

[b] Institute of Computer Science AS CR, Pod Vodárenskou věží 2, 182 07, Prague, 8, Czech Republic

Available online: 28 Oct 2011

To cite this article: M. Červinka, C. Matonoha & J. V. Outrata (2011): On the computation of relaxed pessimistic solutions to MPECs, Optimization Methods and Software, DOI:10.1080/10556788.2011.627585

To link to this article: http://dx.doi.org/10.1080/10556788.2011.627585

Optimization Methods & Software
iFirst, 2011, 1–21

Taylor & Francis
Taylor & Francis Group

On the computation of relaxed pessimistic solutions to MPECs

M. Červinka[a]*, C. Matonoha[b] and J.V. Outrata[a]

[a]*Institute of Information Theory and Automation AS CR, Pod Vodárenskou věží 4, 182 08 Prague 8, Czech Republic;* [b]*Institute of Computer Science AS CR, Pod Vodárenskou věží 2, 182 07 Prague 8, Czech Republic*

(*Received 3 December 2010; final version received 23 September 2011*)

In this paper, we propose a new numerical method to compute approximate and the 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. Introduction

In the last 20 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 [28], and we use the respective terminology very often also in the MPEC setting.

Let us consider an abstract MPEC in the form

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

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

*Corresponding author. Email: cervinka@utia.cas.cz

ISSN 1055-6788 print/ISSN 1029-4937 online

http://dx.doi.org/10.1080/10556788.2011.627585
http://www.tandfonline.com

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

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

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 [8,17,23], 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

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

where $\mathcal{O}$ is a neighbourhood 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 [11–13], 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 behaviour 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

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

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

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

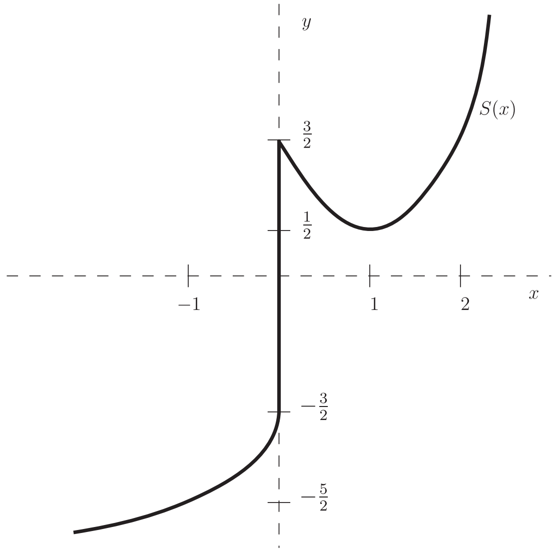

Figure 1. Multifunction *S* from Example 1.

with

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

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

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 $\hat{x} = 1$.

This example clearly demonstrates that when interested in pessimistic solutions of MPECs, one has to analyse 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 [7], to find a pessimistic solution to (1), one either has to minimize a discontinuous, implicitly given value function which is generally not lower-semicontinuous (lsc), 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 Sections 3 and 4, we give a brief description of our proposed numerical method and summarize our numerical experience on test MPECs. Information on already existing components of our numerical method, in particular the system UFO [16] and the numerical codes BOBYQA [25] and BFO [24], can be found in the appendix.

The following notation is employed: $\text{dist}_\Omega(\cdot)$ 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, dom f, Gph f and $[f \leq a]$ to denote its epigraph, hypograph, domain, 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 $\bar{x} \in \text{cl}\Omega$, the *Fréchet normal cone* to Ω at $\bar{x}$ is defined by

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

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

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

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

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

For a convex set Ω, both normal cones N_Ω and $\hat{N}_\Omega$ 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

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

and

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

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

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

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

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

and the *limiting* (*Mordukhovich*) *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 Lipschitz continuity *at* a point. A single-valued mapping $f[\mathbb{R}^n \to \mathbb{R}^m]$ is said to be *Lipschitz* at $\bar{x}$ with modulus $L \geq 0$ if there is a neighbourhood $\mathcal{U}$ of $\bar{x}$ such that

$$|f(x) - f(\bar{x})| \leq L\|x - \bar{x}\| \quad \text{for all } x \in \mathcal{U}. \tag{5}$$

Evidently, a function satisfying (5) need not be Lipschitz on a neighbourhood. In [27], the authors use the name 'calmness' for this type of behaviour.

2. Problem analysis

Before we proceed with relaxations of the pessimistic solution concept, let us start with the application of the upper Fréchet subdifferential optimality condition for local minima under geometric constraints [19, 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 2 *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 Lipschitz at $(\hat{x}, \hat{y})$. Then one has the inclusion

$$-\nabla_x f(\hat{x}, \hat{y}) + \hat{D}^* S(\hat{x}, \hat{y})(-\nabla_y f(\hat{x}, \hat{y})) \subset \hat{N}_\omega(\hat{x}). \tag{6}$$

Remark 3 A single-valued mapping σ is called a *selection* of a multifunction F if $\sigma(x) \in S(x)$ for all $x \in \operatorname{dom} F$.

Proof Applying [19, Proposition 5.2] to problem

$$\begin{aligned} &\text{minimize} && \vartheta(x) \\ &\text{subject to} && x \in \omega, \end{aligned} \tag{7}$$

we arrive at inclusions

$$\hat{\partial}(-\vartheta)(\hat{x}) \subset \hat{N}_\omega(\hat{x}).$$

It remains to apply [20, Theorem 2] which yields

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

This concludes the proof. ∎

For the computation of $\hat{D}^* S(\hat{x}, \hat{y})$ in a frequently arising class of equilibria, the reader is referred to [9]. In most cases, the optimistic solution of (1) is easier to compute. Among possible applications, Theorem 2 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 4 Consider the MPEC

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

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.

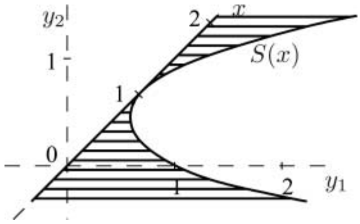

Figure 2. Multifunction S from Example 4.

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 $\nabla_x f(0, 1, 0) = 1$, $\hat{D}^* S(0, 1, 0)(-1, 0) = D^* S(0, 1, 0)(-1, 0) = \{2\}$ and thus (6) yields

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

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

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 5 (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 neighbourhood* $\mathcal{O}$ *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 fulfil 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 6 $\omega := \{x \in \mathbb{R}^n | a_i \le x_i \le b_i\}$, *where* $a_i, b_i \in \bar{\mathbb{R}}, i = 1, \ldots, n$.

Assumption 7 *S is non-empty and convex-valued over* ω.

Assumption 8 *S is compact-valued and outer semicontinuous over* ω, *cf.* [27, Definition 5.4].

In numerous equilibrium problems, the solution map S may have a difficult structure with disconnected images, which would make the computation of the values of ϑ impracticable. Assumption 7 is intended to prevent such situations. Assumption 8 ensures by [2, Theorem 1.4.16] that μ is 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 7 and 8.

Let us denote by $\hat{\vartheta}$ 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 $\hat{\vartheta}$ over ω. In this way, $\hat{\vartheta}(\hat{x})$ 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 9 ((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*

$$\hat{\vartheta}(\tilde{x}) = \vartheta(\tilde{x}) = f(\tilde{x}, \tilde{y})$$
$$\hat{\vartheta}(\tilde{x}) \le \hat{\vartheta}(\hat{x}) + \varepsilon$$
$$\|\tilde{x} - \hat{x}\| < \delta.$$

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 the 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 $L \ge 0$ such that

$$|\vartheta(x) - \hat{\vartheta}(\hat{x})| \le L\|x - \hat{x}\| \tag{8}$$

for all $x \in \omega$ close to $\hat{x}$ with $\vartheta(x)$ close to $\hat{\vartheta}(\hat{x})$. In the following, we show how inequality (8) can be verified in a special class of MPECs.

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

$$S(x) = \{y | 0 \in F(x, y) + N_C(y)\}, \tag{9}$$

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 non-empty convex- and compact-valued outer semicontinuous multifunction and thus equilibria, governed by the generalized equation in (9), satisfy both Assumptions 2 and 3.

To enforce the possible satisfaction of inequality (8), however, we will further simplify the structure of S by assuming that

ASSUMPTION 10 *F is affine and C is polyhedral.*

Then we know from [26, Lemma 4], cf. also [23, Theorem 2.7], that there exists a finite number, say k, of convex polyhedra Ξ_i such that

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

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

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

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 dom$S_i = \Omega_i$ and

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

By virtue of [27, 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 $a := \hat{\vartheta}(\hat{x})$. It follows that there is a subset of $I(x)$, say $I_0(\hat{x})$, such that

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

Since $\vartheta(x) \geq a$ for all x close to $\hat{x}$, to prove (8), in fact, it only suffices to verify that for each $i \in I_0(\hat{x})$, ϑ_i is Lipschitz at $(\hat{x}, a)$ *from above*, i.e., ϑ_i satisfies at $(\hat{x}, a)$ the following one-sided inequality

$$\vartheta_i(x) \leq a + L_i\|x - \hat{x}\| \;\; \forall x \in \Omega_i \text{ and close to } \hat{x}. \tag{10}$$

This property is sometimes called calmness from above, see [27, Section 8.F].

For the calmness of ϑ_i from above, we dispose with the following theorem, where

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

and for $u > a$

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

THEOREM 11 *Let $\hat{x}$ be the first component of a relaxed pessimistic solution of* (1) *and $i \in I_0(\hat{x})$. Then inequality* (10) *is fulfilled whenever*

$$0 \notin \operatorname*{Limsup}_{\substack{u \xrightarrow{g} a \\ g(u) > 0}} \bigcup_{\tilde{x} \in \mathcal{M}(u)} \left\{ \frac{1}{\alpha} \middle| \alpha \frac{\tilde{x} - \hat{x}}{\|\tilde{x} - \hat{x}\|} \in \partial \vartheta_i(\tilde{x}) \right\}. \tag{11}$$

If condition (11) *holds for all $i \in I_0(\hat{x})$, then inequality* (8) *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 Lipschitz continuity of ϑ_i at $(\hat{x}, a)$ from above (inequality (10)) is equivalent with the condition

$$\operatorname{dist}_{[g\leq 0]}(u) \leq L_i g(u) \quad \forall u \text{ close to } a. \tag{12}$$

To continue, we can now invoke [10, Theorem 2.1] and conclude that inequality (12) is implied by the condition

$$0 \notin \operatorname*{Limsup}_{\substack{u \xrightarrow{g} a \\ g(u)>0}} \partial g(u). \tag{13}$$

As the norm $\|x - \hat{x}\|$ is differentiable for $x \in \vartheta^{-1}(u)$, $u > a$, we obtain from [20, Theorem 7] that

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

Thus, condition (13) is implied by

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

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

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

Using [19, Theorem 1.80], we now infer that for all k sufficiently large, one has

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

and so, consequently,

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

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

Remark 12 Note that (12) amounts to the (local) *error bound property* of g at a and the set on the right-hand side of (13) is the (limiting) outer subdifferential of g at a introduced in [10].

We illustrate the application of Theorem 11 by means of the following example.

Example 13 Consider the MPEC

$$\begin{aligned} &\underset{x}{\text{minimize}} \quad (x^2 - x)y_1 + y_2 \\ &\text{subject to } y \in S(x), \end{aligned} \tag{14}$$

where

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

with

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

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

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

and, consequently,

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

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

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

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 $\mathcal{M}(u) = \sqrt{u - \frac{1}{4}} + \frac{1}{2}$, the set on the

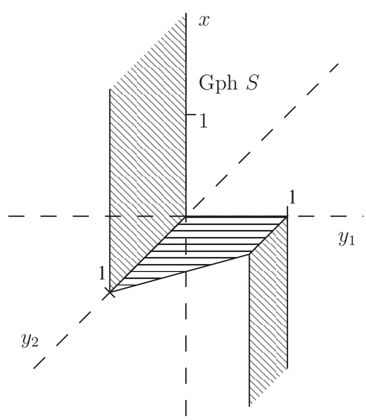

Figure 3. The graph of multifunction S in Example 13.

Table 1. Results for Example 13 using BFO and UFO.

	$x^0 = -0.1$			
Accuracy level	$\tilde{x}$	$\vartheta(\tilde{x})$	neval	niter
10^{-2}	−1.3188E−03	5.013209E−01	35	2689
10^{-3}	−4.7665E−04	5.004769E−01	39	3627
10^{-4}	−1.9561E−05	5.000196E−01	44	4093
10^{-5}	−4.8948E−06	5.000050E−01	49	5476

Table 2. Results for Example 13 using BOBYQA and UFO.

	$x^0 = -0.1$			
Accuracy level	$\tilde{x}$	$\vartheta(\tilde{x})$	neval	niter
10^{-2}	−3.5723E−03	5.035851E−01	27	4243
10^{-3}	−6.7684E−04	5.006773E−01	37	5088
10^{-4}	−6.7684E−04	5.006773E−01	45	5770
10^{-5}	−1.4815E−06	5.000017E−01	95	9923

right-hand side of (11) 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 11 thus yields the calmness of ϑ_1 at $(0, 1/2)$ from above and we infer that the respective function $\vartheta(\cdot) = \max_{i=1,2,3} \vartheta_i(\cdot)$ satisfies inequality (8).

We have included the above example to our collection of test problems for the numerical method proposed in Section 3. For numerical results, see Tables 1 and 2.

Inequality (8) 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 favourable 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. Recall that we consider Assumptions 6–10 and 8 to be satisfied.

For solving the inner optimization problem

$$\begin{aligned} \underset{y}{\text{maximize}} \quad & f(x, y) \\ \text{subject to} \quad & y \in S(x) \end{aligned} \tag{15}$$

with a fixed x, we use a suitable optimization method from the interactive system UFO [16]. 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

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \vartheta(x) \\ \text{subject to} \quad & x \in \omega, \end{aligned} \tag{16}$$

we use the code BFO [24] by Ph.L. Toint for derivative-free minimization of (possibly discontinuous) functions. To find (δ, ε)-pessimistic solutions to MPECs, we have combined these two algorithms into one code. Alternatively, we also replaced BFO by the algorithm BOBYQA [25] developed by M.J.D. Powell.

We give a brief description of our algorithm in the form of a pseudo code.

Data: An initial point x^0, the accuracy level ε_{out} for the outer problem, the accuracy level ε_{in} for the inner problem.
Step 1: Set $k := 0$.
Step 2: If the stopping conditions of the derivative-free method for the point x^k are satisfied with the tolerance ε_{out}, then STOP.
Step 3: Use an optimization method from UFO to find a solution y^k to the inner problem for fixed x^k with the precision ε_{in}.
Step 4: Set $\vartheta(x^k) = f(x^k, y^k)$ and compute a new iteration x^{k+1} of the outer problem using the derivative-free method.
Step 5: Set $k \leftarrow k + 1$; go to Step 2.

Our method is based on using a derivative-free solver, namely either BFO or BOBYQA, to minimize function $\vartheta(x)$. As soon as the derivative-free method requires a function evaluation, a suitable routine from the UFO system is called to obtain an approximate value for $\vartheta(x)$. Various tolerances ε_{out} are used in our numerical experiments. Concerning ε_{in} for the inner problem, adaptive tolerance levels can be considered. The reason is that if we are far from the solution to the main problem, than inaccurate solution to the inner problem may be sufficient. A more accurate solution is meaningful when we approach the solution to the outer problem. However, when we have the solution to the inner problem with smaller tolerance, then the function value $\vartheta(x)$ can be too inaccurate, which can be confusing for the derivative-free solver. For this reason, we make efforts to obtain a solution to the inner problem with high precision.

For reader's convenience, we summarize information on UFO, BFO and BOBYQA in the appendix. For detailed information, please see the above-mentioned references.

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 14 and 15 refer to the example in [8, Section 5.1], the former being the pessimistic and the latter being the optimistic formulation of the same problem. By including Example 15 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 16 is a simple MPEC from [21,23] 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 17 is of similar nature as Examples 13 and 14 but with $x \in \mathbb{R}^3$. In Example 18, we propose a more general

Table 3. Results for Example 14 using BFO and UFO.

	$x^0 = 1$				$x^0 = -1$			
Accuracy level	$\tilde{x}$	$\vartheta(\tilde{x})$	neval	niter	$\tilde{x}$	$\vartheta(\tilde{x})$	neval	niter
10^{-2}	−2.6602E−03	7.0768E−06	41	789	−4.8018E−03	2.3057E−05	45	820
10^{-3}	−3.0889E−04	9.5415E−08	53	1118	−3.4369E−04	1.1812E−07	57	1175
10^{-4}	−2.0862E−05	4.3521E−10	65	1523	−6.7267E−06	4.5248E−11	72	1698
10^{-5}	−2.1996E−06	4.8381E−12	70	1743	−5.2315E−06	2.7369E−11	85	2437

Table 4. Results for Example 14 using BOBYQA and UFO.

	$x^0 = 1$				$x^0 = -1$			
Accuracy level	$\tilde{x}$	$\vartheta(\tilde{x})$	neval	niter	$\tilde{x}$	$\vartheta(\tilde{x})$	neval	niter
10^{-2}	−8.2740E−02	6.8459E−03	14	244	−3.3103E−02	1.0958E−03	18	301
10^{-3}	−1.5838E−03	2.5084E−06	32	652	−3.1050E−02	9.6409E−04	23	409
10^{-4}	−6.1240E−05	3.7503E−09	51	1204	−2.0194E−04	4.0780E−08	45	957
10^{-5}	−7.0333E−06	4.9468E−11	67	1840	−3.2076E−06	1.0289E−11	60	1451

framework for constructing test problems in the form of pessimistically formulated MPECs of arbitrary dimension.

We have considered values 10^{-2}, 10^{-3}, 10^{-4} and 10^{-5} of the precision ε_{out}. In BFO, this parameter is denoted as δ and in BOBYQA as `RHOEND`. In all our computations, we set the accuracy level for UFO computations to 10^{-6}.

Each test result also includes the number of objective function evaluations `neval` (number of UFO calls) and the total number of inner iterations `niter` (Tables 1–6 and 8–13). Note that the higher dimension, the more UFO calls is required by the derivative-free algorithm and hence this procedure might be untractable for large complicated problems for which each UFO computation takes more than just a fraction of a second.

Example 14 [8] (pessimistic formulation)

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

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 (17) in the sense of (3). However, $\hat{x} = 0$ is the first component of the relaxed pessimistic solution. The results are displayed in Tables 3 and 4.

Example 15 [8] (optimistic formulation)

$$\min_{x\in[-2,2]} \min_{y\in S(x)} x^2 + y^2, \tag{18}$$

where S is the same multifunction as in Example 14. The optimistic value function has in this

Table 5. Results of Example 15 using BFO and UFO.

Accuracy level	$x^0 = 1$ $\tilde{x}$	$\vartheta(\tilde{x})$	neval	niter	$x^0 = -1$ $\tilde{x}$	$\vartheta(\tilde{x})$	neval	niter
10^{-2}	−2.6602E−03	7.0768E−06	41	732	−4.8018E−03	2.3057E−05	45	787
10^{-3}	−3.0889E−04	9.5415E−08	53	979	−3.4369E−04	1.1812E−07	57	1010
10^{-4}	−2.0862E−05	4.3521E−10	65	1274	−6.7267E−06	4.5248E−11	72	1366
10^{-5}	−2.1996E−06	4.8381E−12	70	1419	−5.1345E−07	2.6364E−13	73	1419

Table 6. Results for Example 15 using BOBYQA and UFO.

Accuracy level	$x^0 = 1$ $\tilde{x}$	$\vartheta(\tilde{x})$	neval	niter	$x^0 = -1$ $\tilde{x}$	$\vartheta(\tilde{x})$	neval	niter
10^{-2}	0.0000E+00	5.1272E−10	14	217	0.0000E+00	5.1272E−10	14	241
10^{-3}	0.0000E+00	5.1272E−10	18	288	0.0000E+00	5.1272E−10	18	313
10^{-4}	0.0000E+00	5.1272E−10	24	412	−2.4372E−06	5.9399E−12	23	418
10^{-5}	−6.7285E−06	4.5272E−11	28	520	−2.4372E−06	5.9399E−12	28	520

problem the form

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

Thus its global optimistic solution is attained at $(\bar{x}, \bar{y}) = (0, 0)$. The numerical results can be found in Tables 5 and 6.

Example 16 [21,23] (An MPEC with a single-valued solution map at each feasible point) Consider an *oligopolistic market* model with five firms producing a homogeneous product and attempting to maximize their profits; see, e.g. [21,23]. Let $x \in \mathbb{R}$ denote the *production* of the Leader and let $y_i \in \mathbb{R},\ i = 1, \dots, 4$, 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

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

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 $c_i, i = 0, \dots, 4$, be in the form

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

Table 7. Parameter specification for the production costs.

	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 8. Results of Example 16 using BFO and UFO.

	$x^0 = 150$			
Accuracy level	$\tilde{x}$	$\vartheta(\tilde{x})$	neval	niter
10^{-2}	99.53233	958.634749	54	394
10^{-3}	99.53440	958.634750	70	506
10^{-4}	99.53452	958.634750	71	513
10^{-5}	99.53452	958.634750	80	576

Table 9. Results for Example 16 using BOBYQA and UFO.

	$x^0 = 150$			
Accuracy level	$\tilde{x}$	$\vartheta(\tilde{x})$	neval	niter
10^{-2}	99.53535	958.634750	204	1532
10^{-3}	99.53473	958.634750	220	1644
10^{-4}	99.53452	958.634750	268	1980
10^{-5}	99.53452	958.634750	292	2148

where b_i, K_i and β_i, $i = 0, \ldots, 4$, are positive parameters given by Table 7.

Further, let

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

with a parameter $\gamma \geq 1$ termed demand elasticity. The numerical results, see Tables 8 and 9, can be compared with [23, Table 12.4] where the chosen accuracy is 5×10^{-4}.

Example 17 (A pessimistically formulated MPEC with a relaxed pessimistic solution)

$$\min_{x \in \mathbb{R}^3} \max_{y \in S(x)} \frac{1}{2}\|x\|^2 + \frac{1}{2}\|y\|^2, \tag{19}$$

where

$$S(x) = \{y \in \mathbb{R}^3 | 0 \in By + x + N_{\Delta^3}(y)\}. \tag{20}$$

In (20) Δ^3 is the standard three-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 (19) in the sense of (3). However, $\hat{x} = (0, 0, 0)$ is the first

Table 10. Results for Example 17 using BFO and UFO.

	$x^0 = (1,1,1)$					
Accuracy level	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$	$\vartheta(\tilde{x})$	neval	niter
10^{-2}	4.0030E−03	2.8674E−03	7.1296E−04	1.2377E−05	192	5854
10^{-3}	−6.8044E−05	3.1937E−04	−7.1007E−06	5.9399E−08	260	11,894
10^{-4}	−1.9698E−06	−5.2347E−07	1.3221E−05	9.5927E−11	302	21,104
10^{-5}	2.8807E−06	3.9539E−06	2.0290E−06	1.4024E−11	404	47,541

Table 11. Results for Example 17 using BOBYQA and UFO.

	$x^0 = (1,1,1)$					
Accuracy level	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$	$\vartheta(\tilde{x})$	neval	niter
10^{-2}	2.8246E−04	0.0000E+00	0.0000E+00	3.9892E−08	21	585
10^{-3}	0.0000E+00	0.0000E+00	3.3178E−05	5.5043E−10	26	979
10^{-4}	0.0000E+00	0.0000E+00	3.3179E−05	5.5043E−10	30	1283
10^{-5}	0.0000E+00	3.9427E−06	2.5245E−05	3.2680E−10	33	1312

component of the relaxed pessimistic solution of (19). For numerical results of Example 17, see Tables 10 and 11.

Example 18 (A pessimistically formulated MPEC of an arbitrary dimension with a relaxed pessimistic solution)

$$\min_{x \in \mathbb{R}^n} \max_{y \in S(x)} \frac{1}{2}\|x - a\|^2 + y_1, \tag{21}$$

where

$$\begin{aligned} S(x) &= \{y \in \mathbb{R}^n | 0 \in x + N_C(y), C = \{y | a^\top y = c, 0 \le y \le \xi\}\} \\ &= \operatorname{argmin}_{y \in \mathbb{R}^n}\{x^\top y | a^\top y = c, 0 \le y \le \xi\}, \end{aligned} \tag{22}$$

$a \in \mathbb{R}^n_+, \xi \in \mathbb{R}^n_+$ 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 $\xi_i = 7c/10a_i,\ i = 1, \dots, n$. Clearly, in each such a problem of arbitrary dimension $n \ge 3$ the mapping S is multi-valued at $x = a$ which is also the first component of the relaxed pessimistic solution of (21).

Tests on randomly generated pessimistic MPECs (21) were performed with the accuracy level $\varepsilon_{\text{out}} = 10^{-3}$ for dimensions $n = 5, 7$ and 10, respecting the upper bound for dimension of x in BFO. In Tables 12 and 13, we display results for one randomly generated problem for each chosen dimension.

Observe that this example also fulfills all assumptions imposed in Theorem 11.

Remark 19 MPEC solvers from the UFO system solve the inner problem only to local optimality. Note, that the objective functions of the inner problems of Examples 13 and 18 are concave and hence the achieved local maximum coincides with global maximum. In Examples 14, 15 and 17, the objective of the inner problems is non-concave, nevertheless monotone. Recall that for outer iterations only the objective function values are required, see Step 4 of the pseudo code. The feasible set of the inner problem of Example 16 is a singleton. Hence, in all our examples, we

Table 12. Results of Example 18 using BFO and UFO.

	$n = 5$	$n = 7$	$n = 10$
$\vartheta(\tilde{x})$	2.2332E−08	3.6194E−08	6.3720E−08
$\lvert\tilde{x}_1 - a_1\rvert$	3.1436E−06	9.7398E−05	1.4094E−04
$\lvert\tilde{x}_2 - a_2\rvert$	1.7766E−04	1.1088E−04	1.2510E−04
$\lvert\tilde{x}_3 - a_3\rvert$	4.8796E−05	3.2391E−05	1.3370E−04
$\lvert\tilde{x}_4 - a_4\rvert$	2.9030E−05	2.9904E−05	2.9079E−05
$\lvert\tilde{x}_5 - a_5\rvert$	9.9338E−05	8.0485E−06	1.3239E−05
$\lvert\tilde{x}_6 - a_6\rvert$	–	2.1372E−04	6.0411E−05
$\lvert\tilde{x}_7 - a_7\rvert$	–	5.4058E−05	7.6219E−05
$\lvert\tilde{x}_8 - a_8\rvert$	–	–	2.0728E−04
$\lvert\tilde{x}_9 - a_9\rvert$	–	–	1.0901E−04
$\lvert\tilde{x}_{10} - a_{10}\rvert$	–	–	9.3382E−05
Neval	560	728	1568
Niter	291,252	497,440	1,222,010

Table 13. Results for Example 18 using BOBYQA and UFO.

	$n = 5$	$n = 7$	$n = 10$
$\vartheta(\tilde{x})$	1.0457E−05	7.5686E−06	7.9517E−08
$\lvert\tilde{x}_1 - a_1\rvert$	2.9207E−03	1.0668E−03	3.5089E−05
$\lvert\tilde{x}_2 - a_2\rvert$	1.3966E−03	5.0643E−04	9.7482E−05
$\lvert\tilde{x}_3 - a_3\rvert$	2.7919E−04	4.2053E−04	6.5710E−05
$\lvert\tilde{x}_4 - a_4\rvert$	2.3091E−03	8.3652E−04	1.2207E−04
$\lvert\tilde{x}_5 - a_5\rvert$	2.2413E−03	6.9819E−06	2.2455E−06
$\lvert\tilde{x}_6 - a_6\rvert$	–	7.1518E−04	1.0833E−05
$\lvert\tilde{x}_7 - a_7\rvert$	–	3.5149E−03	5.3241E−05
$\lvert\tilde{x}_8 - a_8\rvert$	–	–	1.2408E−05
$\lvert\tilde{x}_9 - a_9\rvert$	–	–	7.5868E−05
$\lvert\tilde{x}_{10} - a_{10}\rvert$	–	–	3.4672E−04
Neval	106	137	153
Niter	120,702	159,750	187,567

have not encountered any difficulties with computation of just local optima of the inner problem. For general examples that do not fall into any of the three cases above, one should use a global MPEC solver. In general, to find a global solution to an MPEC is an NP-hard problem. There has been some recent development in this topic, see e.g. [18].

In all our computations reported in Tables 1–6 and 8–13, we do not consider any kind of warm starts. We have tested the option to include warm starts, e.g. to choose the initial point y^0 for the kth start of the inner iteration process as the solution to the inner problem from preceding step $k - 1$. This modification of the code did not yield any significant improvement in the test results.

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, and `RHOBEG` $= 1$. 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.

Notice that the reported results in Tables 3 and 5 coincide (with the exception of `niter`). Given the same starting parameters, the iteration process of BFO is the same regardless of optimistic or pessimistic formulation. On the other hand, the iteration process of BOBYQA is different from that of BFO. We observed that for a given MPEC, the minimization process of BOBYQA for an usc value function ϑ and an lsc value function μ differs and it is much faster for the latter case. As can be seen from Tables 4 and 6, the algorithm composed of BOBYQA and UFO achieved the prescribed accuracy level in Example 15 in less number of iterations than in Example 14, given the same starting parameters. A possible explanation may lie in the construction of the quadratic interpolation of the objective, see [25].

Note also that in most examples the number of function evaluations is significantly less for BOBYQA than for BFO.

For comments on performance of used optimization methods from UFO, please see appendix.

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 behaviour 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.

Both used derivative-free optimization methods, BFO and BOBYQA, do not benefit much from the structure of problems. We expect improvement in performance of the proposed algorithm when using other available derivative-free optimization codes that (at least partially) benefit from the structure of the specific class of problems. This will be subject to our future investigation.

Our final note concerns 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).

Acknowledgements

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. The authors would also like to thank the anonymous referees for their careful review of the manuscript and for many constructive comments and suggestions. This work was supported by the Grant Agency of the Czech Republic, project No. 201/09/1957, by the Ministry of Education of the Czech Republic, grant 1M0572, and by the institutional research plan No. AV0Z10300504.

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] A.R. Conn, N.I.M. Gould, and Ph.L. Toint, *Trust-region Methods*, SIAM, 2000.
[6] A.R. Conn, K. Scheinberg, and L.N. Vicente, *Introduction to Derivative-free Optimization*, SIAM, 2009.

[7] S. Dempe, *A bundle algorithm applied to bilevel programming problems with non-unique lower level solutions*, Comput. Optim. Appl. 15 (2000), pp. 145–166.
[8] S. Dempe, *Foundations of Bi-level Programming*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.
[9] R. Henrion, J.V. Outrata, and T. Surowiec, *On regular coderivatives in parametric equilibria with non-unique multipliers*, (2010), accepted to Mathematical Programming B.
[10] A.D. Ioffe and J. Outrata, *On metric and calmness qualification conditions in subdifferential calculus*, Set-Valued Anal. 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. Global Optim. 8 (1996), pp. 263–287.
[14] L. Lukšan, C. Matonoha, and J. Vlček, *Interior-point method for nonlinear nonconvex optimization*, Numer. Linear Algebra Appl. 11 (2004), pp. 431–453.
[15] L. Lukšan, C. Matonoha, and J. Vlček, *Interior-point method for nonlinear programming with complementarity constraints*, Tech. Rep. 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*, Tech. Rep. V-1083, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, 2010.
[17] Z.-Q. Luo, J.-S. Pang, and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.
[18] A. Mitsos, P. Lemonidis, and P.I. Barton, *Global solution of bilevel programs with a nonconvex inner program*, J. Global Optim. 42 (2008), pp. 475–513.
[19] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation*, Vol. 1: Theory, Vol. 2: Applications, Springer, 2006.
[20] B.S. Mordukhovich, N.M. Nam, and N.D. Yen, *Subgradients of marginal functions in parametric mathematical programming*, Math. Program. 116 (2009), pp. 369–396.
[21] F.H. Murphy, H.D. Sherali, and A.L. Soyster, *A mathematical programming approach for determining oligopolistic market equilibrium*, Math. Program. 24 (1982), pp. 92–106.
[22] J. Nocedal and S.J. Wright, *Numerical Optimization*, 2nd ed., Springer, 2006.
[23] J.V. Outrata, M. Kočvara, and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, Kluwer Academic Publisher, Dordrecht, The Netherlands, 1998.
[24] M. Porcelli, and Ph.L. Toint, *BFO, a Brute force optimizer for mixed integer nonlinear bound-constrained optimization and its self tuning*, Tech. Rep., Namur Research Center on Complex Systems (NAXYS), in preparation, 2010.
[25] M.J.D. Powell, *The BOBYQA algorithm for bound constrained optimization without derivatives*, Tech. Rep. DAMTP 2009/NA06, Cambridge, 2009.
[26] S.M. Robinson, *An implicit function theorem for generalized variational inequalities*, Tech. Summary Rep. 1672, Mathematics Research Center, University of Wisconsin-Madison, 1976.
[27] R.T. Rockafellar and R.J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
[28] H. von Stackelberg, *Marktform und Gleichgewicht*, Springer, Berlin, 1934.

Appendix

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 a lot of iterative solvers with several preconditioners can be used for the computation of a descent direction.

Inner problem (15) is typically an optimization problem with variational inequality constraints. After introducing the Karush-Kuhn-Tucker (KKT) conditions we obtain a minimization problem containing nonlinear equality and inequality constraints and complementarity constraints which can be written in the general form

$$\min \varphi(y), \quad \text{subject to} \quad c_E(y) = 0, \quad c_I(y) \le 0, \quad c_K^{\mathrm{T}}(y) c_L(y) = 0, \tag{A1}$$

where $E = \{1, \dots, m_E\}$, $I = \{1, \dots, m_I\}$, and $K, L \subseteq I$ are index sets. This problem is difficult to solve by standard methods since the Mangasarian–Fromowitz constraint qualification is not satisfied at any feasible point. Therefore, special methods for solving problem (A1) have been developed. For the numerical experiments reported in Section 4 we have considered the following two approaches that are implemented in the UFO system.

In the first approach, inner problem (A1) is treated as a general nonlinear programming problem in which the complementarity constraints are considered as general equality constraints. Thus complementarity constraints $c_K^{\mathrm{T}}(y) c_L(y) = 0$ are included into $c_E(y)$ and the function $\varphi(y)$ is minimized on the set given by constraints $c_E(y) = 0$ and $c_I(y) \le 0$. We use the interior-point method, see e.g. [1,4], whose main idea lies in introduction of a slack vector $s \in \mathbb{R}^{m_I}$ and transformation of the original problem to a sequence of problems with the logarithmic barrier function. For a given barrier coefficient $\nu > 0$, we search for a minimum of function $\varphi(y) - \nu e^{\mathrm{T}} \ln(S) e$ on the set of equality constraints $c_E(y) = 0,\ c_I(y) + s = 0$. Here e is the vector with unit elements, $S = \operatorname{diag}(s_i,\ i = 1, \dots, m_I)$, and $\nu \to 0$. We introduce the Lagrange multipliers $u = (u_E, u_I)^{\mathrm{T}}$ and use the Newton method to solve the necessary KKT conditions. In each iteration of the Newton method we obtain a system of linear equations to determine direction vectors d_y, d_s, d_u, where the Hessian matrix is approximated by gradient differences. This system is symmetrized and the solution is obtained from the indefinitely preconditioned conjugate gradient method. A step-length $0 < \alpha < \bar{\alpha}$ is determined using a line-search technique with maximum step-length $\bar{\alpha}$ described in [14]. It is necessary to fulfil conditions $s_i, u_i > 0,\ i = 1, \dots, m_I$. A further requirement for the selection of the step-length is satisfaction of a suitable goal criterion. This criterion is usually a merit function which is a combination of the barrier function and a measure of constraint violation. Finally, the barrier coefficient $\nu \to 0$ is updated, for this purpose we use heuristic formulas, see [14] for details. After setting up several line-search parameters, mainly the value $\bar{\alpha}$, we have been able to obtain quite a good solution to inner problems for fixed x.

The second approach for solving problem (A1) consists in using a recently developed method for solving nonlinear programming problems with complementarity constraints. It is also based on the interior-point approach and uses an exact penalty function to remove complementarity constraints, cf. [15]. In this case, which is similar to the previous one, inner problem (A1) is replaced with a sequence of problems

$$\min\{\varphi(y) + \varrho s_K^{\mathrm{T}} s_L - \nu e^{\mathrm{T}} \ln(S) e\}, \quad \text{subject to} \quad c_E(y) = 0, \quad c_I(y) + s = 0, \tag{A2}$$

where $\varrho > 0$ is a penalty parameter, which have the same solution as (A1) if ϱ is sufficiently large. The algorithm for solving problem (A2) is now the same as in the first approach. Constraints of transformed problem (A2) usually satisfy the Mangasarian–Fromowitz constraint qualification. On the other hand, the corresponding matrix of the system of linear equations contains extra non-zero entries and can be indefinite. A merit function contains an additional penalty term with parameter ϱ and one needs also a heuristic procedure for updating ϱ. As in the previous case, after setting up several parameters, mainly a parameter for increase of ϱ and the maximal step-size $\bar{\alpha}$, we have been able to compute (an approximation of) the same solution as by the first approach.

Another possible approach how to compute direction vectors is trust-region implementation [5]. There the direction vector is not determined as a solution to the system of linear equations solved by the indefinitely preconditioned conjugate gradient method, but as a solution to the trust-region subproblem whose optimality conditions lead to the same system of equations. However, it may happen that a trust-region constraint $\|y\| \le \Delta$ is incompatible with equality constraints contained in the trust-region subproblem, so the direction vector is decomposed into two parts, normal and tangential [22]. This a bit more difficult implementation is also included in the UFO system, so we have used it in our problems for comparison as well. The results using the line-search method described in the previous two paragraphs were quite good, see Tables 1–13, but using the trust-region method they were worse in terms of accuracy. So we decided to present only the results using the line-search implementation.

BFO [24] is a 'Brute-Force Optimizer', written in Matlab, for unconstrained or bound 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 [25] is a bound 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. In contrast to BFO, BOBYQA has in principle no restriction on the number of variables. Problems with hundreds of variables can be solved using BOBYQA. The only condition is that n must be at least two.

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 in 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 contains a very accurate and efficient system of updating the approximation models and 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 a 'good' set of interpolation points is too high or the approximation is poor, see [6, 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 suitably the final trust-region radius.