A convergent inexact solution method for equilibrium problems*

David Di Lorenzo† Mauro Passacantando‡ Marco Sciandrone†

Abstract. We consider equilibrium problems with differentiable bifunctions. We adopt the well-known approach based on the reformulation of the equilibrium problem as a global optimization problem through an appropriate gap function. We propose a solution method based on the inexact (and hence, less expensive) evaluation of the gap function, and on the employment of a nonmonotone line search. We prove global convergence properties of the proposed inexact method under standard assumptions. Some preliminary numerical results show the potential computational advantages of the inexact method compared with a standard exact descent method.

Keywords: Equilibrium problem, gap function, inexact method, nonmonotone line search.

AMS Subject classification: 90C33, 90C30.

1 Introduction

Given a closed and convex set $C \subseteq \mathbb{R}^n$ and a bifunction $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, we consider the following *equilibrium problem*:

$$\text{find } x^* \in C \text{ s.t. } f(x^*, y) \geq 0, \qquad \forall\, y \in C. \tag{EP}$$

It is well-known (see e.g. [6, 3]) that (EP) provides a general setting which includes several problems such as scalar and vector optimization, complementarity, variational inequality, fixed point, saddle point, inverse optimization, and Nash equilibrium problems in noncooperative games.

Several methods to solve equilibrium problems have been proposed in the literature (see the recent survey paper [3]), often extending those originally conceived for optimization problems or variational inequalities (see, for instance, [12, 15]) to the framework of more general equilibrium problems.

*This paper has been published in Optimization Methods and Software, vol. 29 (5), pp. 979-991, see: `https://www.tandfonline.com/doi/full/10.1080/10556788.2013.796376`.

†Dipartimento di Sistemi e Informatica, Università di Firenze, Via di Santa Marta 3, 50139 Firenze, Italy

‡Dipartimento di Informatica, Università di Pisa, Largo B. Pontecorvo 3, 56127 Pisa, Italy

A well-known class of solution methods is given by the so-called descent methods, which are based on the reformulation of the equilibrium problem as a global optimization problem through appropriate gap or D-gap functions [2, 4, 5, 8, 9, 18, 20, 22, 26, 27].

In all these approaches the evaluation of the gap function at a point and the computation of the search direction consist in finding the exact solution of a convex optimization problem. However, these evaluations could become computationally expensive when the number of variables and constraints increases. For this reason, in this paper we propose an inexact solution method for solving (EP) where at each iteration both the search direction and the values of the gap function are computed by finding approximate solutions of convex optimization problems. The search direction computed by inexactly solving the convex optimization problem may be a non descent direction, so the adoption of a nonmonotone line search is necessary for computing the stepsize along it. In particular, we employ a nonmonotone line search previously introduced in the context of derivative-free optimization [10, 21]. The global convergence of the proposed algorithm is proved under not restrictive assumptions. Summarizing, we present a convergent nonmonotone line search-based method, where both the search directions and the function values are computed by inexactly solving suitable optimization problems in order to reduce the computational cost.

Other inexact methods for equilibrium problems, which are based on different approaches, have been proposed in the literature: for instance, in [14, 16, 23] proximal point methods are proposed, in [17, 19] Tikhonov-Browder methods are studied, and projection-type methods are described in [24, 25]. However, to the best of our knowledge, no inexact approach based on gap functions has been proposed so far.

The rest of the paper is organized as follows. In section 2 we recall the concept of gap function and describe the standard descent methods for (EP). Then, we show in section 3 our inexact solution method and prove its global convergence under assumptions which are either the same or weaker than those needed in other well-known methods. Finally, section 4 provides the results of some preliminary numerical tests. The obtained results show the potential computational advantages of the proposed inexact method compared with a standard exact descent method.

Throughout all the paper we will assume that the set C is bounded, the bifunction f is continuously differentiable, $f(x, \cdot)$ is convex and $f(x, x) = 0$ for all $x \in C$. It is well-known (see e.g. [13]) that these assumptions guarantee the existence of at least one solution to (EP).

2 Preliminary background

A function $g : C \to \mathbb{R}$ is said to be a *gap function* for (EP) if g is non-negative on C and x^* solves (EP) if and only if $x^* \in C$ and $g(x^*) = 0$. Thus, gap functions are tools to reformulate an equilibrium problem as a global optimization problem, whose optimal value is known a priori. In order to build gap functions with good regularity properties, it is usual to consider an auxiliary bifunction $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ which is continuously differentiable and satisfies the following conditions:

$$h(x, y) \geq 0 \text{ for all } x, y \in C \text{ and } h(z, z) = 0 \text{ for all } z \in C, \tag{1}$$

$$h(x, \cdot) \text{ is strongly convex for all } x \in C, \tag{2}$$

$$\nabla_y h(z, z) = 0 \text{ for all } z \in C, \tag{3}$$

$$\langle \nabla_x h(x, y) + \nabla_y h(x, y), y - x \rangle \geq 0 \text{ for all } x, y \in C. \tag{4}$$

We remark that the most used regularizing bifunction is $h(x, y) = ||y - x||^2$.

Given any $\alpha > 0$, the function defined as

$$\varphi_\alpha(x) = - \min_{y \in C} \ \{f(x, y) + \alpha\, h(x, y)\} \tag{5}$$

is a gap function (see e.g. [22]). Since the objective function $f(x, \cdot) + \alpha\, h(x, \cdot)$ is strongly convex, there exists a unique optimal solution $y_\alpha(x)$ of the problem which defines the gap function (5). Moreover, it is well known [22] that φ_α is continuously differentiable with

$$\nabla \varphi_\alpha(x) = -\nabla_x f(x, y_\alpha(x)) - \alpha\, \nabla_x h(x, y_\alpha(x)),$$

and x^* solves (EP) if and only if $x^* = y_\alpha(x^*)$.

We note that condition (4) is not needed for proving that φ_α is a gap function. However, assumptions of this kind are the key tools to devise descent methods.

Definition 2.1 ([2]). *A differentiable bifunction f is called*

- ∇–monotone *on C if $\langle \nabla_x f(x, y) + \nabla_y f(x, y), y - x \rangle \geq 0$ for all $x, y \in C$;*
- strictly ∇–monotone *on C if $\langle \nabla_x f(x, y) + \nabla_y f(x, y), y - x \rangle > 0$ for all $x, y \in C$ with $x \neq y$;*
- strongly ∇–monotone *on C with constant $\tau > 0$ if*

$$\langle \nabla_x f(x, y) + \nabla_y f(x, y), y - x \rangle \geq \tau \|y - x\|^2, \qquad \textit{for all } x, y \in C.$$

For some special classes of equilibrium problems, the above defined concepts reduce to well–known ones. For instance, when $f(x, y) = \phi(y) - \phi(x)$, with $\phi : \mathbb{R}^n \to \mathbb{R}$, (EP) coincides with the problem of minimizing ϕ over C: in this case the (strict, strong) ∇–monotonicity of f is equivalent to the (strict, strong) convexity of ϕ. When $f(x, y) = \langle F(x), y - x \rangle$ for some mapping $F : \mathbb{R}^n \to \mathbb{R}^n$, (EP) is actually a variational inequality problem: the (strong) ∇–monotonicity of f is equivalent to the (strong) monotonicity of F and the strict ∇–monotonicity of f implies the strict monotonicity of F.

However, it is important to notice that for general nonlinear bifunctions the well-known monotonicity condition on f (see e.g. [1]), i.e.

$$f(x, y) + f(y, x) \leq 0, \qquad \forall\ x, y \in C,$$

is not related to the ∇-monotonicity condition defined above, as the following examples show (see [2]). In fact, the bifunction $f(x, y) = -x^2 - xy + 2y^2$ is strongly ∇–monotone on $\mathbb{R}^2$, because

$$\langle \nabla_x f(x, y) + \nabla_y f(x, y), y - x \rangle = 3(y - x)^2, \qquad \forall\ x, y \in \mathbb{R},$$

but it is not monotone since

$$f(x,y)+f(y,x)=x^2-2xy+y^2=(x-y)^2>0, \qquad \forall\, x,y\in\mathbb{R},\ x\neq y.$$

Vice versa, the bifunction $f(x,y)=e^{x^2}(y^2-x^2)$ is monotone on $\mathbb{R}^2$ because

$$f(x,y)+f(y,x)=-(e^{y^2}-e^{x^2})(y^2-x^2)\leq 0, \qquad \forall\, x,y\in\mathbb{R},$$

but it is not ∇–monotone since

$$\langle \nabla_x f(x,y)+\nabla_y f(x,y), y-x\rangle = 2e^{x^2}(y-x)^2(x^2+xy+1)$$

is negative for instance when $x=1$ and $y=-3$.

On the other hand, we now show a special class of equilibrium problems where the ∇-monotonicity condition is weaker than the monotonicity condition. We consider a class of Nash equilibrium problems in noncooperative games with N players, where each player i has a set of possible strategies $K_i\subseteq \mathbb{R}^{n_i}$ and aims at minimizing a cost function $f_i: C\to\mathbb{R}$ with $C=K_1\times\cdots\times K_N$. It is well-known that finding a Nash equilibrium of the game amounts to solving (EP) with the bifunction f equal to the so-called Nikaido-Isoda function:

$$f(x,y)=\sum_{i=1}^{N}\left[f_i(y_i,x_{-i})-f_i(x)\right], \tag{6}$$

where the vector x_{-i} denotes the strategies of all the players different from i. We assume that the players cost functions are quadratic, so we are using

$$f_i(x)=\frac{1}{2}x_i^T A_{ii}x_i+\sum_{\substack{j=1\\ j\neq i}}^{N} x_i^T A_{ij}x_j + x_i^T b_i, \qquad i=1,\dots,N,$$

where the squared matrices $A_{11},\dots,A_{NN}$ are positive semidefinite. In this setting, it is easy to check that

$$f(x,y)+f(y,x)=-(y-x)^T\begin{pmatrix} 0 & A_{12} & \dots & A_{1n}\\ A_{21} & 0 & \dots & A_{2n}\\ \vdots & & \ddots & \vdots\\ A_{n1} & \dots & \dots & 0\end{pmatrix}(y-x) \tag{7}$$

and

$$\langle \nabla_x f(x,y)+\nabla_y f(x,y), y-x\rangle = (y-x)^T\begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n}\\ A_{21} & A_{22} & \dots & A_{2n}\\ \vdots & & \ddots & \vdots\\ A_{n1} & \dots & \dots & A_{nn}\end{pmatrix}(y-x). \tag{8}$$

Therefore, if f is monotone then it is also ∇-monotone; moreover, when the matrices $A_{11},\dots,A_{NN}$ are positive definite, the monotonicity of f implies the strong ∇-monotonicity of f. Finally, we note that it is easy to find matrices A_{ij} such that f is (strongly) ∇-monotone but not monotone (see also examples in section 4).

The strict ∇-monotonicity of f provides a natural descent direction for φ_α and guarantees that stationary points of φ_α coincide with its global minima, i.e. the solutions of (EP).

Theorem 2.1 ([2, 22]). *Suppose that f is strictly ∇–monotone on C.*

(a) If $x \in C$ does not solve (EP), then $y_\alpha(x) - x$ is a descent direction for φ_α at x, i.e. $\langle \nabla\varphi_\alpha(x), y_\alpha(x) - x\rangle < 0$.

(b) If x^* is a stationary point of φ_α over C, i.e. $\langle \nabla\varphi_\alpha(x^*), y - x^*\rangle \geq 0$ for all $y \in C$, then x^* solves (EP).

It follows from the above theorem that standard descent methods can be devised moving away from a non-stationary point x^k along the direction $d^k = y_\alpha(x^k) - x^k$ with a suitable stepsize $t_k \in (0, 1]$ to obtain the new iterate $x^{k+1} = x^k + t_k d^k$. In [22] it has been proved that this method with the exact line search

$$t_k \in \arg\min\{\ \varphi_\alpha(x^k + td^k)\ \ :\ \ t \in [0, 1]\ \}$$

converges to solutions of (EP) provided that f is strictly ∇–monotone. In [9, 20, 22] Armijo-type inexact line search has been considered choosing the stepsize $t_k = \gamma^s$ with $\gamma \in (0, 1)$ and s being the smallest nonnegative integer such that

$$\varphi_\alpha(x^k + \gamma^s\, d^k) \leq \varphi_\alpha(x^k) - \beta\, \gamma^s\, \|d^k\|^2.$$

Convergence of these methods is guaranteed provided that f is strongly ∇–monotone with constant $\tau > \beta$.

3 Inexact solution method

In this section we propose a solution method for (EP) which is based on the approximate evaluation of the gap function φ_α at each iteration. The main motivation lies in the fact that the computation of the search direction $y_\alpha(x) - x$ and the line search procedure require to exactly solve the optimization problem (5) defining φ_α, and this could be too expensive from a computational point of view. In this paper, we will consider approximate solutions of the optimization problem (5) within error ε. We recall that, whenever the exact solution $y_\alpha(x)$ of (5) is computed we have, by definition,

$$\varphi_\alpha(x) = -f(x, y_\alpha(x)) - \alpha\, h(x, y_\alpha(x)) \geq -f(x, y) - \alpha\, h(x, y) \quad \forall\ y \in C.$$

Given a positive scalar $\varepsilon > 0$, we say that a point $y^\varepsilon_\alpha(x) \in C$ is an inexact ε-solution of (5) if

$$\varphi^\varepsilon_\alpha(x) := -f(x, y^\varepsilon_\alpha(x)) - \alpha\, h(x, y^\varepsilon_\alpha(x)) \geq \varphi_\alpha(x) - \varepsilon. \tag{9}$$

Thus, by definition we have

$$\varphi_\alpha(x) - \varepsilon \leq \varphi^\varepsilon_\alpha(x) \leq \varphi_\alpha(x).$$

We remark that it is possible to determine an inexact ε-solution of (5) without knowing $\varphi_\alpha(x)$ (which would require the computation of the exact solution $y_\alpha(x)$). This will be shown in Section 3.1.

We now describe the proposed inexact solution method. At the $k-$th iteration, starting from the current point x^k, we compute an inexact ε_k-solution, i.e., we compute a point $y^k = y_\alpha^{\varepsilon_k}(x^k)$ such that

$$\varphi_\alpha(x^k) - \varepsilon_k \le \varphi_\alpha^{\varepsilon_k}(x^k) := -f(x^k, y^k) - \alpha\, h(x^k, y^k). \tag{10}$$

Then, we define the search direction

$$d^k = y^k - x^k,$$

and we perform an inexact line search along it for computing the stepsize t_k. Note that the search direction d^k may be a non descent direction, so that we adopt a nonmonotone line search (previosuly proposed in [10, 21]) based on an acceptance rule of the form

$$\varphi_\alpha^{\varepsilon_k}(x^k + t_k\, d^k) \le \varphi_\alpha^{\varepsilon_k}(x^k) - \beta\, t_k^2\, \|d^k\|^2 + \delta_k,$$

where δ_k is a positive scalar. The algorithm needs to manage two sequences of positive scalars, $\{\delta_k\}$ and $\{\varepsilon_k\}$, the first one related to the nonmonotone line search, the second one related to the degree of accuracy in inexactly solving the optimization problem (5). In order to guarantee global convergence properties of the generated sequence $\{x^k\}$, both $\sum_{k=0}^{\infty} \delta_k$ and $\sum_{k=0}^{\infty} \varepsilon_k$ must be convergent, furthermore, the following relation must hold

$$0 < \varepsilon_k < \delta_k, \qquad \forall\, k.$$

The algorithm is formally described below.

Inexact Algorithm (IA)

0. Choose $\alpha > 0$, $\beta, \gamma \in (0, 1)$, two sequences $\{\delta_k\}$ and $\{\varepsilon_k\}$ such that $0 < \varepsilon_k < \delta_k$ and $\sum_{k=0}^{\infty} \delta_k$ is convergent. Let $x^0 \in C$ and set $k = 0$.

1. Compute $y^k \in C$ such that

$$\varphi_\alpha^{\varepsilon_k}(x^k) := -f(x^k, y^k) - \alpha\, h(x^k, y^k) \ge \varphi_\alpha(x^k) - \varepsilon_k,$$

set $d^k = y^k - x^k$ and compute the smallest non-negative integer s such that

$$\varphi_\alpha^{\varepsilon_k}(x^k + \gamma^s\, d^k) \le \varphi_\alpha^{\varepsilon_k}(x^k) - \beta\, \gamma^{2s}\, \|d^k\|^2 + \delta_k.$$

2. Set $t_k = \gamma^s$, $x^{k+1} = x^k + t_k\, d^k$, $k = k + 1$ and go to Step 1.

In the following we state some lemmas which are necessary to prove the global convergence of the method.

Lemma 3.1. *The line search procedure is well-defined, i.e. it terminates in a finite number of steps.*

Proof. By contradiction, assume that

$$\varphi_\alpha^{\varepsilon_k}(x^k + \gamma^s\, d^k) - \varphi_\alpha^{\varepsilon_k}(x^k) > -\beta\, \gamma^{2s}\, \|d^k\|^2 + \delta_k \tag{11}$$

holds for all $s \in \mathbb{N}$. Since we have

$$\varphi_\alpha^{\varepsilon_k}(x^k + \gamma^s\, d^k) \le \varphi_\alpha(x^k + \gamma^s\, d^k) \quad \text{and} \quad \varphi_\alpha^{\varepsilon_k}(x^k) \ge \varphi_\alpha(x^k) - \varepsilon_k,$$

we obtain from (11) that

$$\varphi_\alpha(x^k + \gamma^s\, d^k) - \varphi_\alpha(x^k) > -\beta\, \gamma^{2s}\, \|d^k\|^2 + \delta_k - \varepsilon_k$$

holds for all $s \in \mathbb{N}$. For s sufficiently large we have

$$\frac{\delta_k - \varepsilon_k}{\gamma^s} \ge \varphi_\alpha'(x^k; d^k) + 1.$$

Therefore, we can write

$$\frac{\varphi_\alpha(x^k + \gamma^s\, d^k) - \varphi_\alpha(x^k)}{\gamma^s} \ge -\beta\, \gamma^s\, \|d^k\|^2 + \frac{\delta_k - \varepsilon_k}{\gamma^s} \ge -\beta\, \gamma^s\, \|d^k\|^2 + \varphi_\alpha'(x^k; d^k) + 1.$$

Taking the limits for $s \to \infty$ we obtain $0 \ge 1$. □

Lemma 3.2. *It holds*

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0. \tag{12}$$

Proof. We have

$$\varphi_\alpha(x^{k+1}) - \varepsilon_k \le \varphi_\alpha^{\varepsilon_k}(x^{k+1}) \le \varphi_\alpha^{\varepsilon_k}(x^k) - \beta\, \|t_k\, d^k\|^2 + \delta_k \le \varphi_\alpha(x^k) - \beta\, \|t_k\, d^k\|^2 + \delta_k.$$

Thus for all $k \in \mathbb{N}$ we obtain

$$\varphi_\alpha(x^{k+1}) \le \varphi_\alpha(x^k) - \beta\, \|t_k\, d^k\|^2 + \delta_k + \varepsilon_k.$$

We can write

$$\sum_{l=0}^{k} \beta \|t_l\, d^l\|^2 \le \varphi_\alpha(x^0) - \varphi_\alpha(x^{k+1}) + \sum_{l=0}^{k} (\delta_l + \varepsilon_l) \le \varphi_\alpha(x^0) + 2\, \sum_{l=0}^{k} \delta_l.$$

Since the series $\sum_{k=0}^{\infty} \delta_k$ is convergent, we obtain that also the series $\sum_{k=0}^{\infty} \|t_k\, d^k\|^2$ is convergent. Therefore we have that

$$\lim_{k\to\infty} \|t_k\, d^k\| = 0,$$

and hence (12) holds. □

Lemma 3.3. *For any $\alpha > 0$ we have*

$$\lim_{k\to\infty} \|y^k - y_\alpha(x^k)\| = 0.$$

Proof. At iteration k we consider the function $F_k(y) := f(x^k, y) + \alpha h(x^k, y)$. Thus

$$y_\alpha(x^k) = \arg\min_{y \in C} F_k(y).$$

We have

$$0 \le F_k(y^k) - F_k(y_\alpha(x^k)) \le \varepsilon_k,$$

thus

$$\lim_{k\to\infty} \Big[F_k(y^k) - F_k(y_\alpha(x^k)) \Big] = 0. \tag{13}$$

Since the function F_k is strongly convex, there must exist a positive scalar M such that

$$F_k(y^k) - F_k(y_\alpha(x^k)) \ge \langle \nabla F_k(y_\alpha(x^k)), y^k - y_\alpha(x^k) \rangle + M \|y^k - y_\alpha(x^k)\|^2.$$

It follows from the definition of $y_\alpha(x^k)$ that

$$\langle \nabla F_k(y_\alpha(x^k)), y^k - y_\alpha(x^k) \rangle \ge 0,$$

hence

$$F_k(y^k) - F_k(y_\alpha(x^k)) \ge M \, \|y^k - y_\alpha(x^k)\|^2. \tag{14}$$

Therefore the thesis follows from (13). □

Theorem 3.1. *If C is bounded and f is strictly ∇–monotone on C, then the Inexact Algorithm produces a sequence $\{x^k\}$ such that any of its cluster points solves (EP).*

Proof. The existence of cluster points of the sequence $\{x^k\}$ follows from the boundedness of C. Let x^* be any cluster point of $\{x^k\}$. Then there exists an infinite subset $K \subseteq \mathbb{N}$ such that

$$\lim_{k\in K, k\to\infty} x^k = x^*.$$

Let $d^* = y_\alpha(x^*) - x^*$. Since $d^k = y^k - y_\alpha(x^k) + y_\alpha(x^k) - x^k$, by continuity of the mapping y_α and by Lemma 3.3 we have

$$\lim_{k\in K, k\to\infty} d^k = d^*.$$

If $d^* = 0$, then x^* solves (EP). Otherwise, using (12), it follows

$$\lim_{k\in K, k\to\infty} t_k = 0. \tag{15}$$

The step size rule implies that for all $k \in K$ and k sufficiently large we have

$$\varphi_\alpha(x^k + t_k\gamma^{-1}d^k) > \varphi_\alpha(x^k) - \beta(t_k\gamma^{-1})^2 \|d^k\|^2.$$

In fact, we have

$$\begin{aligned}
\varphi_\alpha(x^k + t_k\gamma^{-1}d^k) &\ge \varphi_\alpha^{\varepsilon_k}(x^k + t_k\gamma^{-1}d^k) \\
&> \varphi_\alpha^{\varepsilon_k}(x^k) - \beta(t_k\gamma^{-1})^2 \|d^k\|^2 + \delta_k \\
&\ge \varphi_\alpha(x^k) - \beta(t_k\gamma^{-1})^2 \|d^k\|^2 + \delta_k - \varepsilon_k \\
&> \varphi_\alpha(x^k) - \beta(t_k\gamma^{-1})^2 \|d^k\|^2.
\end{aligned}$$

By the mean value theorem we can write

$$\langle \nabla \varphi_\alpha(x^k + \theta^k\, t_k\, \gamma^{-1} d^k), d^k \rangle > -\beta t_k\, \gamma^{-1} \|d^k\|^2,$$

where $\theta^k \in (0,1)$. Taking the limit as $k \in K, k \to +\infty$, using (15), the continuity of y_α and $\nabla \varphi_\alpha$, we obtain

$$\langle \nabla \varphi_\alpha(x^*), d^* \rangle \geq 0.$$

Then, by Theorem 2.1a, x^* solves (EP). ☐

We note that the inexact algorithm improves the classical descent methods given in [9, 20, 22] for two reasons: at each iteration they have to find exact solutions of convex optimization problems instead of inexact ones as described in our method; moreover the key assumption for the convergence of our method (i.e. the strict ∇-monotonicity of f) is either the same or weaker than those needed in the classical methods.

3.1 On the computation of an inexact ε-solution

A key issue of the proposed algorithm concerns the computation of an inexact solution of the optimization problem (5). More specifically, the algorithm requires to compute an ε-solution of the optimization problem (5), i.e., a point $y^\varepsilon_\alpha(x) \in C$ satisfying (9). Now we state a condition sufficient to guarantee that a point $\bar{y} \in C$ is an ε-solution. Given $\alpha > 0$ and a vector $x \in C$, to simplify the notation we set

$$F(y) = f(x, y) + \alpha\, h(x, y).$$

Proposition 3.1. *Let $\bar{y} \in C$, $\hat{y} = P[\bar{y} - \nabla F(\bar{y})]$ where P is the Euclidean projection on C, and let $\Omega = diam(C) + \|\nabla F(\bar{y})\|$. If*

$$\|\bar{y} - \hat{y}\| \leq \frac{\varepsilon}{\Omega}, \tag{16}$$

then $\bar{y}$ is an ε-solution, i.e.,

$$f(x, \bar{y}) + \alpha\, h(x, \bar{y}) \leq -\varphi_\alpha(x) + \varepsilon.$$

Proof. Letting

$$y^\star = \arg\min_{y \in C} F(y)$$

we have $F(y^\star) = -\varphi_\alpha(x)$. Then, in order to prove the thesis, we must show that

$$F(\bar{y}) \leq F(y^\star) + \varepsilon.$$

Using the properties of the projection mapping we have

$$\langle \bar{y} - \nabla F(\bar{y}) - \hat{y}, y^\star - \hat{y} \rangle \leq 0,$$

so that we obtain

$$\langle \nabla F(\bar{y}), y^\star - \hat{y} \rangle \geq \langle \bar{y} - \hat{y}, y^\star - \hat{y} \rangle,$$

from which we get

$$\langle \bar{y} - \hat{y}, y^\star - \hat{y} \rangle \leq \langle \nabla F(\bar{y}), y^\star - \bar{y} \rangle + \langle \nabla F(\bar{y}), \bar{y} - \hat{y} \rangle.$$

Thus, we can write

$$\begin{aligned} \langle \nabla F(\bar{y}), y^\star - \bar{y} \rangle \quad &\geq \langle \bar{y} - \hat{y}, y^\star - \hat{y} - \nabla F(\bar{y}) \rangle \\ &\geq -\|\bar{y} - \hat{y}\| \|y^\star - \hat{y} - \nabla F(\bar{y})\| \\ &\geq -\frac{\varepsilon}{\Omega} \left[\|y^\star - \hat{y}\| + \|\nabla F(\bar{y})\| \right] \geq -\varepsilon, \end{aligned}$$

so that, recalling the convexity of $F(y)$, we obtain

$$F(y^\star) \geq F(\bar{y}) + \langle \nabla F(\bar{y}), y^\star - \bar{y} \rangle \geq F(\bar{y}) - \varepsilon,$$

and this concludes the proof. □

On the basis of the preceding proposition, in order to compute an ε-solution any globally convergent method generating a sequence $\{y^\ell\}$ can be applied and stopped at iteration $\bar{\ell}$ whenever condition (16) holds with $\bar{y} = y^{\bar{\ell}}$. The global convergence properties of the method generating $\{y^\ell\}$ ensure that an ε-solution is determined in a finite number of iterations.

4 Preliminary computational experiments

In this section we show the results of some computational experiments performed on a class of Nash equilibrium problems. The aim of the numerical experiments is to evaluate possible computational advantages of the inexact algorithm compared with a standard exact algorithm.

Implementation details of Inexact Algorithm

The algorithm has been implemented in Python, using the scipy library v0.11. The sequences $\{\epsilon_k\}$ and $\{\delta_k\}$ are defined as follows:

$$\varepsilon_k = \max\{(0.8)^k, 10^{-12}\}, \qquad \delta_k = 1.1\, \varepsilon_k.$$

The line search parameters are $\beta = 10^{-4}$ and $\gamma = 0.5$. The used regularizing bifunction is $h(x, y) = \|y - x\|^2$. The parameter α is set equal to 0.1 and the starting point is randomly uniformly generated into C.

We use LBFGS-B algorithm [7] to approximately evaluate φ_α, namely to solve problem (5) with a given tolerance ε (the subroutine stops when the infinity norm of the projected gradient is less than or equal to ε).

Every 10 iterations we evaluate the (almost) exact value of $\varphi_\alpha(x^k)$ by setting equal to 10^{-12} the tolerance parameter in the stopping criterion of the optimization algorithm. We use such value to decide if the algorithm should stop: we terminate the algorithm when $\varphi_\alpha(x^k) \leq 10^{-8}$.

If, after 10000 iterations, the algorithm fails to converge, we stop the program.

Numerical results

We compare the proposed inexact algorithm (IA) with two different versions of the exact algorithm. The two versions have been derived from the implementation of IA setting, as tolerance parameter in the stopping criterion of the optimization algorithm, $\epsilon_k = 10^{-8}$ and $\epsilon_k = 10^{-12}$ respectively. In both the cases the line search scalar δ_k has been set equal to zero for all k, which corresponds to perform a monotone line search.

In order to perform extensive test cases, we consider the class of Nash equilibrium problems with quadratic payoff functions described in section 2, that is we assume that the cost function of each player is

$$f_i(x) = \frac{1}{2} x_i^T A_{ii} x_i + \sum_{\substack{j=1 \\ j \neq i}}^{N} x_i^T A_{ij} x_j + x_i^T b_i, \qquad i = 1, \dots, N.$$

In our experiments the number of players is n, each of them controls one variable and the feasible set C is defined as follows

$$C = \{x \in \mathbb{R}^n : \; -1 \leq x_i \leq 1, \quad i = 1, \dots, n\}.$$

Matrix A is generated as follows:

0. Randomly generate every element of A uniformly in the range $[-1, 1]$
1. Set $A = A^T A$. A is now symmetric and positive semidefinite.
2. Add 10^{-4} to every element in the diagonal of A, which now becomes symmetric positive definite.
3. Add to A a random antisymmetric matrix, with every element uniformly generated in the range $[-1, 1]$. A is now a positive definite matrix.

Every vector b_i is also randomly generated, with every element in the range $[-1, 1]$.

Since each generated matrix A is positive definite, it follows from relation (8) that the bifunction f is strongly ∇-monotone, hence the convergence of the Inexact algorithm is guaranteed by Theorem 3.1.

Moreover, we now prove that in this class of test problems the bifunction f is not monotone. In fact, if f was monotone, then by (7) the matrix

$$M = \begin{pmatrix} 0 & A_{12} & \dots & A_{1n} \\ A_{21} & 0 & \dots & A_{2n} \\ \vdots & & \ddots & \vdots \\ A_{n1} & \dots & \dots & 0 \end{pmatrix}$$

should be positive semidefinite. If we choose two indices $i, j \in \{1, \dots, n\}$ with $i \neq j$, and define two vectors $v, w \in \mathbb{R}^n$ as follows:

$$v_k = \begin{cases} 0 & \text{if } k \neq i, \, k \neq j \\ 1 & \text{if } k = i \\ 1 & \text{if } k = j \end{cases} \qquad w_k = \begin{cases} 0 & \text{if } k \neq i, \, k \neq j \\ 1 & \text{if } k = i \\ -1 & \text{if } k = j \end{cases}$$

then we obtain

$$v^T M v = A_{ij} + A_{ji} \geq 0, \qquad w^T M w = -(A_{ij} + A_{ji}) \geq 0,$$

i.e. $A_{ij} = -A_{ji}$. Hence the matrix M should be skew-symmetric. However, the procedure to generate matrix A guarantees that no matrix M is skew-symmetric, therefore the bifunction f can not be monotone.

We tested the algorithms on Nash equilibrium problems with sizes $n = 5$, 10, 20 and 50. For every different size, we randomly generated 100 instances of the problem. Figures 1–4 show the performance profiles [11] in terms of number of function evaluations.

Numerical results show that the proposed inexact algorithm clearly outperforms the two versions of the exact algorithm. The computational advantages of the inexact algorithm are clear in all cases, and particularly significant for $n = 20$ and $n = 50$. Summarizing, the computational experiments, although limited to a class of problems, show the validity of the proposed inexact nonmonotone approach.

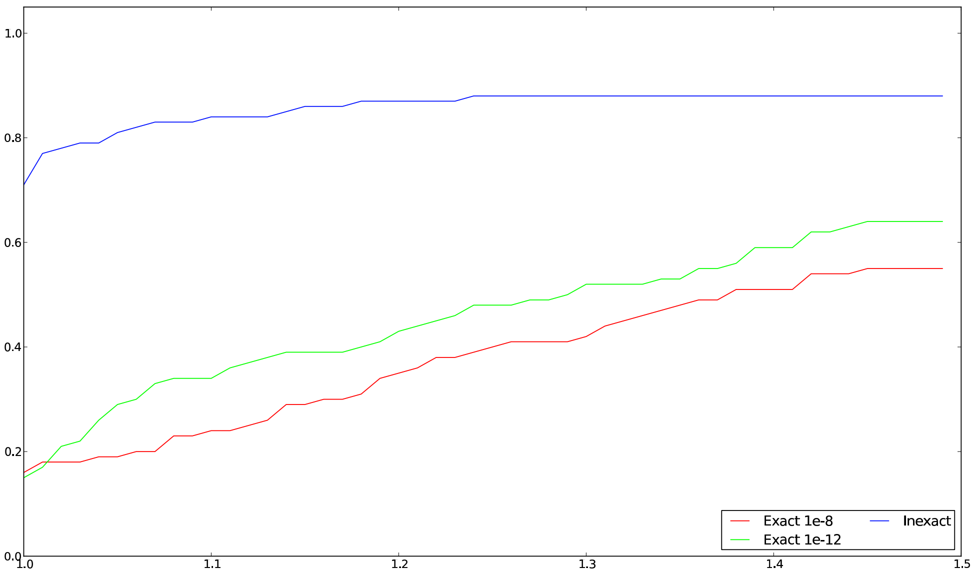

Figure 1: Function evaluations performance profile ($n = 5$)

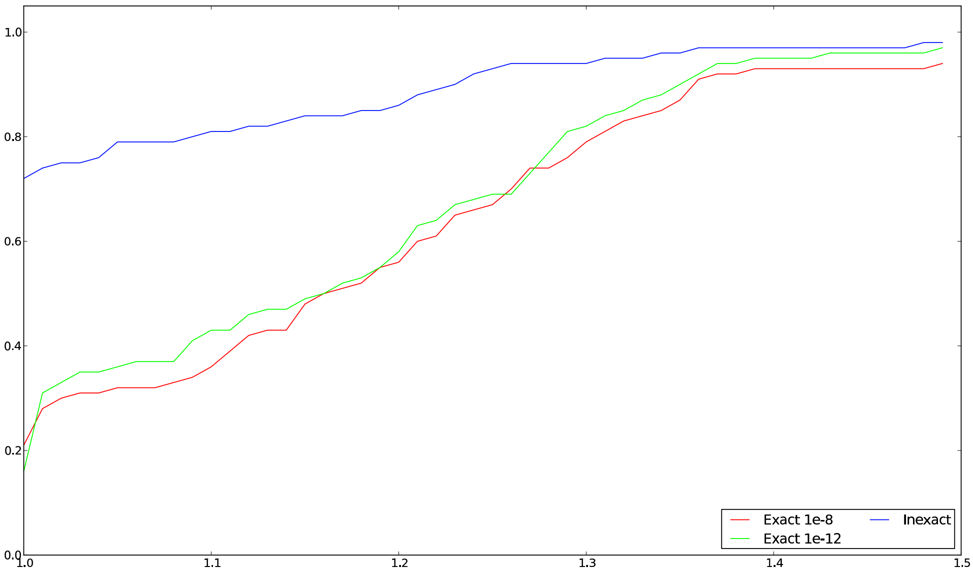

Figure 2: Function evaluations performance profile ($n = 10$)

References

[1] M. Bianchi and S. Schaible, Generalized monotone bifunctions and equilibrium problems, J. Optim. Theory Appl. 90 (1996), pp. 31–43.

[2] G. Bigi, M. Castellani, and M. Pappalardo, *A new solution method for equilibrium problems*, Optim. Methods Softw. 24 (2009), pp. 895–911.

[3] G. Bigi, M. Castellani, M. Pappalardo, and M. Passacantando, *Existence and solution methods for equilibria*, European J. Oper. Res. 227 (2013), pp. 1–11.

[4] G. Bigi and M. Passacantando, *Gap functions and penalization for solving equilibrium problems with nonlinear constraints*, Comput. Optim. Appl. 53 (2012), pp. 323–346.

[5] G. Bigi and M. Passacantando, *Descent and penalization techniques for equilibrium problems with nonlinear constraints*, Journal of Optimization Theory and Applications (2013), doi: 10.1007/s10957-013-0473-7.

[6] E. Blum and W. Oettli, *From optimization and variational inequalities to equilibrium problems*, Math. Student 63 (1994), pp. 123–145.

[7] R.H. Byrd, P. Lu, J. Nocedal, and C.Y. Zhu, *A limited memory algorithm for bound constrained optimization*, SIAM J. Sci. Comput. 16 (1995), pp. 1190–1208

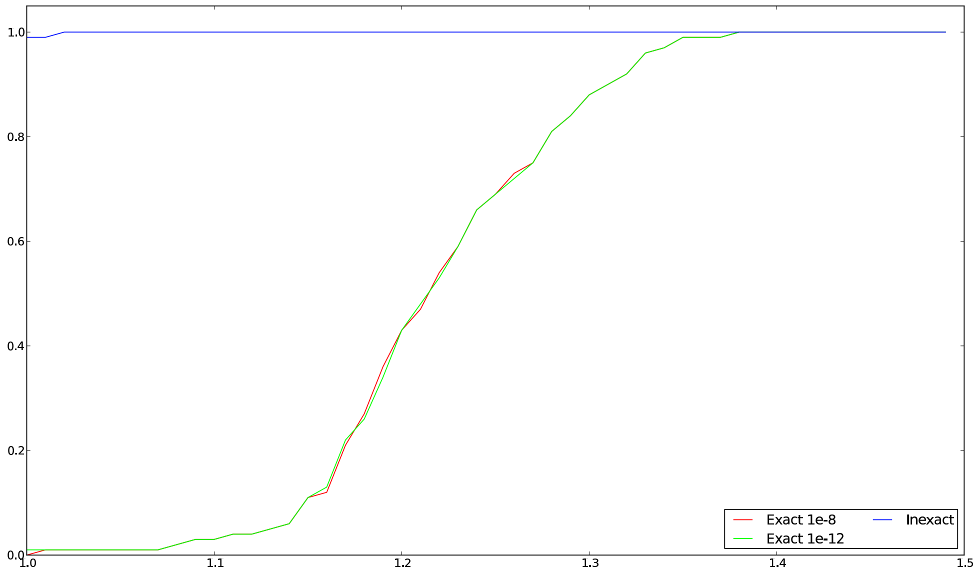

Figure 3: Function evaluations performance profile ($n = 20$)

[8] C. Cherugondi, *A note on D-gap functions for equilibrium problems*, Optimization 62 (2013), pp. 211–226.

[9] O. Chadli, I.V. Konnov, and J.C. Yao, *Descent methods for equilibrium problems in a Banach space*, Comput. Math. Appl. 48 (2004), pp. 609–616.

[10] M.A. Diniz-Ehrhardt, J.M. Martinez, and M. Raydan, *A derivative-free nonmonotone line search technique for unconstrained optimization*, J. Comput. Appl. Math. 219 (2008), pp. 383–397.

[11] E. D. Dolan and J. J. Moré, *Benchmarking Optimization Software with Performance Profiles*, Math. Program. Ser. A 91 (2002), pp. 201–213.

[12] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, Berlin, 2003.

[13] K. Fan, *A minimax inequality and applications*, in Inequalities, III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969; dedicated to the memory of Theodore S. Motzkin), Academic Press, New York, 1972, pp. 103–113.

[14] A.N. Iusem and M. Nasri, *Inexact proximal point methods for equilibrium problems in Banach spaces*, Numer. Funct. Anal. Optim. 28 (2007), pp. 1279–1308.

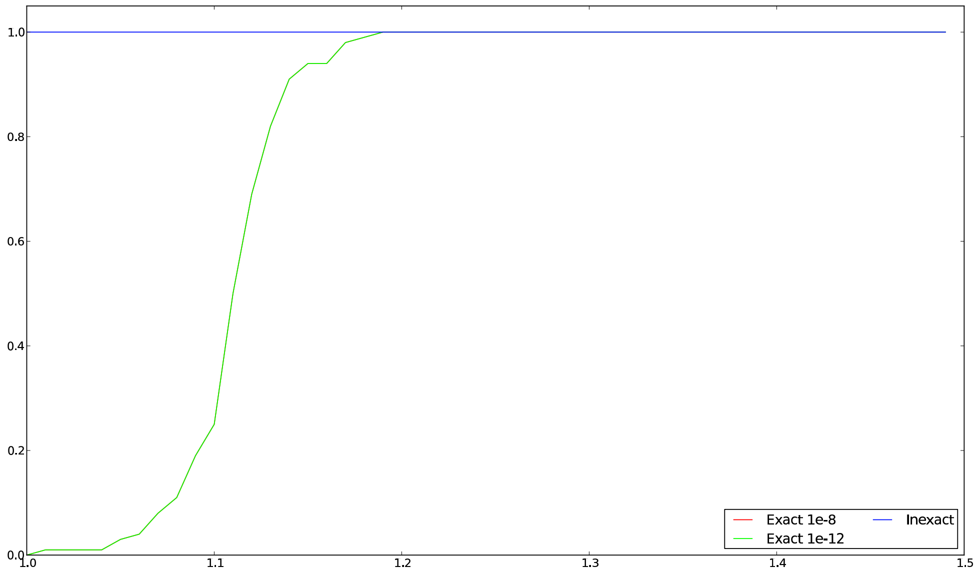

Figure 4: Function evaluations performance profile ($n = 50$)

[15] I.V. Konnov, *Combined Relaxation Methods for Variational Inequalities*, Springer, Berlin, 2001.

[16] I.V. Konnov, *Partial proximal point method for nonmonotone equilibrium problems*, Optim. Methods Softw. 21 (2006), pp. 373–384.

[17] I.V. Konnov and M.S.S. Ali, *Descent methods for monotone equilibrium problems in Banach spaces*, J. Comput. Appl. Math. 188 (2006), pp. 165–179.

[18] I.V. Konnov and O.V. Pinyagina, *D-gap functions for a class of equilibrium problems in Banach spaces*, Comput. Methods Appl. Math. 3 (2003), pp. 274–286.

[19] I.V. Konnov and O.V. Pinyagina, *D-gap functions and descent methods for a class of monotone equilibrium problems*, Lobachevsky J. Math. 13 (2003), pp. 57–65.

[20] I.V. Konnov and O.V. Pinyagina, *A descent method with inexact linear search for nonsmooth equilibrium problems*, Comput. Math. Math. Phys. 48 (2008), pp. 1777–1783.

[21] D.H. Li and M. Fukushima, *A derivative-free line search and global convergence of Broyden-like method for nonlinear equations*, Optim. Methods Softw. 13 (2000), pp. 181–201.

[22] G. Mastroeni, *Gap functions for equilibrium problems*, J. Global Optim. 27 (2003), pp. 411–426.

[23] A. Moudafi, *Proximal point methods extended to equilibrium problems*, J. Natural Geometry 15 (1999), pp. 91–100.

[24] P. Santos and S. Scheimberg, *An inexact subgradient algorithm for equilibrium problems*, Comput. Appl. Math. 30 (2011), pp. 91–107.

[25] S. Scheimberg and P.S.M. Santos, *A relaxed projection method for finite-dimensional equilibrium problems*, Optimization 60 (2011), pp. 1193–1208.

[26] L. Zhang and J.Y. Han, *Unconstrained optimization reformulations of equilibrium problems*, Acta Math. Sin. (Engl. Ser.) 25 (2009), pp. 343–354.

[27] L. Zhang and S.-Y. Wu, *An algorithm based on the generalized D-gap function for equilibrium problems*, J. Comput. Appl. Math. 231 (2009), pp. 403–411.