# The Interior Proximal Extragradient Method for Solving Equilibrium Problems

Nguyen Thi Thu Van,[*] Jean-Jacques Strodiot,[†] and Van Hien Nguyen[‡]

**Abstract** In this paper we present a new and efficient method for solving equilibrium problems on polyhedra. The method is based on an interior-quadratic proximal term which replaces the usual quadratic proximal term. This leads to an interior proximal type algorithm. Each iteration consists in a prediction step followed by a correction step as in the extragradient method. In a first algorithm each of these steps is obtained by solving an unconstrained minimization problem, while in a second algorithm the correction step is replaced by an Armijo-backtracking linesearch followed by an hyperplane projection step. We prove that our algorithms are convergent under mild assumptions: pseudomonotonicity for the two algorithms and a Lipschitz property for the first one. Finally we present some numerical experiments to illustrate the behavior of the proposed algorithms.

**Keywords** Interior proximal method, logarithmic-quadratic proximal method, extragradient method, Armijo-backtracking linesearch, equilibrium problems.

## 1 Introduction

An equilibrium problem in the sense of Blum and Oettli [13], denoted EP, is to find a point $x^* \in C$ such that

$$f(x^*, y) \geq 0 \text{ for all } y \in C,$$

where $C$ is a closed convex subset of $\mathbb{R}^n$ and $f : C \times C \to \mathbb{R}$ satisfies $f(x, x) = 0$ for all $x \in C$. In this paper, we assume that $C$ is a polyhedral set with a nonempty interior given

[*]Department of Mathematics, University of Namur (FUNDP), Namur, Belgium, and Faculty of Mathematics and Informatics, University of Natural Sciences, Vietnam National University, Ho Chi Minh City, Vietnam. E-mail: tvnguyen@fundp.ac.be

[†]Department of Mathematics, University of Namur (FUNDP), Namur, Belgium.
E-mail: jean-jacques.strodiot@fundp.ac.be

[‡]Department of Mathematics, University of Namur (FUNDP), Namur, Belgium.
E-mail: vhnguyen@fundp.ac.be.

by

$$C = \{\, x \mid Ax \leq b \,\},$$

with $A$ an $m \times n$ $(m \geq n)$ matrix of full rank with rows $a_i$, and $b$ is a vector in $\mathbb{R}^m$ with rows $b_i$. An important example of such a $C$ is the nonnegative orthant of $\mathbb{R}^n$. We also assume that $f$ is continuous on $C \times C$ and that $f(x, \cdot)$ is convex and subdifferentiable on $C$ for all $x \in C$. We denote by $S^*$ the solution set of EP and we assume that there exists at least one solution to problem EP. Existence results for this problem can be found, for instance, in [13].

This problem is very general in the sense that it includes, as particular cases, the optimization problem, the variational inequality problem, the Nash equilibrium problem in noncooperative games, the fixed point problem, the nonlinear complementarity problem and the vector optimization problem (see, for instance, Blum and Oettli [13] and the references quoted therein). The interest of this problem is that it unifies all these particular problems in a convenient way. Moreover, many methods devoted to solving one of these problems can be extended, with suitable modifications, to solving the general equilibrium problem.

A typical method to solve the equilibrium problem is based on a fixed-point formulation of EP which starts with a point $x^0 \in C$ and generates a sequence $\{x^k\}$ defined, for all $k \in \mathbb{N}$, by

$$x^{k+1} = \arg\min_{x \in C} f(x^k, x).$$

However, this problem, in general, may not have a solution, and if it does, the solution may not be unique. To avoid this situation, it is more convenient to use the auxiliary problem principle (see Cohen [15]) which is based on the following fixed point property: $x^* \in C$ is a solution of problem EP if and only if it is a solution of the regularized problem

$$\min_{y \in C} \{c\, f(x^*, y) + \frac{1}{2}\|x^* - y\|^2\},$$

where $c > 0$. Observe that this problem has a unique solution. As a result, the corresponding fixed point iteration is: *Given $x^k \in C$, find $x^{k+1} \in C$ the solution of*

$$\min_{y \in C} \{c_k\, f(x^k, y) + \frac{1}{2}\|x^k - y\|^2\}. \tag{1}$$

This method has been proven to be convergent by Mastroeni [23] under the assumptions that $f$ is strongly monotone on $C \times C$, i.e., that there exists $\gamma > 0$ such that

$$f(x, y) + f(y, x) \leq -\gamma\|y - x\|^2, \quad \forall x, y \in C,$$

and that $f$ satisfies the following property: there exist $c_1, c_2 > 0$ such that

$$\forall x, y, z \in C \quad f(x, y) + f(y, z) \geq f(x, z) - c_1\|y - x\|^2 - c_2\|z - y\|^2. \tag{2}$$

This is a Lipschitz-type condition. Indeed, when $f(x,y) = \langle F(x), y-x\rangle$ with $F : I\!R^n \rightarrow I\!R^n$, problem EP amounts to the variational inequality problem: find $x^* \in C$ such that $\langle F(x^*), y-x^*\rangle \geq 0$ for all $y \in C$. In that case, $f(x,y)+f(y,z)-f(x,z) = \langle F(x)-F(y), y-z\rangle$ for all $x,y,z \in C$, and it is easy to see that if $F$ is Lipschitz continuous on $C$ (with constant $L > 0$), then for all $x,y,z \in C$,

$$|\langle F(x) - F(y), y - z\rangle| \leq L\|x-y\|\,\|y-z\| \leq \frac{L}{2}\,[\|x-y\|^2 + \|y-z\|^2],$$

and thus, $f$ satisfies condition (2). Furthermore, when $z = x$, this condition becomes

$$f(x,y) + f(y,x) \geq -(c_1 + c_2)\,\|y-x\|^2, \quad \forall\, x,y \in C.$$

This gives a lower bound on $f(x,y) + f(y,x)$ while the strong monotonicity gives an upper bound on $f(x,y) + f(y,x)$.

In order to avoid the strong monotonicity assumption on $f$, Antipin [1], [17] proposed to add at each iteration an extrapolation step after solving (1). This strategy has also been recently incorporated in Mastroeni's algorithm by Tran et al. [30] to obtain the convergence under the weaker assumption that $f$ is pseudomonotone on $C \times C$, i.e.,

$$f(x,y) \geq 0 \Rightarrow f(y,x) \leq 0, \quad \forall x,y \in C,$$

and that $f$ satisfies condition (2). More precisely, they denote by $y^k$ the unique solution of (1), and they take for $x^{k+1}$ the solution of the following problem

$$\min_{y \in C}\{c_k\, f(y^k, y) + \frac{1}{2}\|y - x^k\|^2\}.$$

When $f(x,y) = \langle F(x), y-x\rangle$ where $F : I\!R^n \rightarrow I\!R^n$, i.e., when problem EP is a variational inequality problem, it is easy to see that the iterates $y^k$ and $x^{k+1}$ are given by

$$y^k = P_C(x^k - c_k F(x^k)) \quad \text{and} \quad x^{k+1} = P_C(x^k - c_k F(y^k)),$$

where $P_C(y)$ is the orthogonal projection of $y$ over $C$.

This is the classical extragradient method for solving the single-valued variational inequality problem. This method has been introduced by Korpelevich in [21]. However, condition (2) is rather strong. So a well-known strategy to avoid it, is to use an Armijo-backtracking linesearch along the direction $y^k - x^k$ to get a point $z^k$ which is used to construct an hyperplane separating $x^k$ from the solution set. Then the new iterate $x^{k+1}$ is the projection of $x^k$ onto this hyperplane (for more details concerning variational inequality problems, see [16] and [19]). This strategy has been adapted by Konnov [18], [20] for solving differentiable monotone equilibrium problems. Recently, using the same strategy but for subdifferentiable equilibrium problems, Tran et al. [30] proved the convergence of the resulting method under the sole assumption that $f$ is pseudomonotone. Nevertheless, all these methods always assume that solving constrained subproblems can be done efficiently. But it is well known that the boundary of constraints can destroy some of the nice properties of unconstrained methods (see a discussion about this in [8]). So it seems interesting to consider unconstrained subproblems instead of constrained ones.

In this context and when $\operatorname{int} C \neq \emptyset$, Auslender et al. have proposed in [4] a new type of interior proximal method for solving convex programs by replacing in subproblems (1) the quadratic term $\frac{1}{2}\|x^k - y\|^2$ by some nonlinear function $D(y, x^k)$ composed of two parts: the first part is based on entropic proximal terms and will play a role of barrier function forcing the iterates $\{x^k\}$ to remain in the interior of $C$. The second part is a quadratic convex regularization based on the set $C$ to preserve the nice properties of the Auxiliary Problem Principle. So the classical difficulties associated with the boundary of the constraints are automatically eliminated. This way to transform a constrained problem into an unconstrained one has already been used by Antipin [17] but with a distance-like function $D(y, x^k)$ based on Bregman functions. However, as mentioned by Auslender (see Theorem 2.1 of [8]), the distance-like function based on the logarithmic-quadratic function enjoys several nice properties not shared by other nonquadratic functions. It is the reason why we concentrate our study on this class of functions.

In this paper, we assume that the interior of $C$ is nonempty. Then the distance-like function, denoted $D_\varphi(x, y)$, is constructed from a class of functions $\varphi : I\!R \to (-\infty, +\infty]$ of the form

$$\varphi(t) = \mu h(t) + \frac{\nu}{2}(t-1)^2, \tag{3}$$

where $\nu > \mu > 0$ and $h$ is a closed and proper convex function satisfying the following additional properties:

(a) $h$ is twice continuously differentiable on $(0, +\infty)$, the interior of its domain,

(b) $h$ is strictly convex on its domain,

(c) $\lim_{t \to 0^+} h^{'}(t) = -\infty$,

(d) $h(1) = h^{'}(1) = 0$ and $h^{''}(1) = 1$, and

(e) For $t > 0$, $1 - t^{-1} \leq h^{'}(t) \leq t - 1$.

Amongst all the functions $h$ satisfying properties $(a) - (e)$, let us mention the following one:

$$h(t) = \begin{cases} t - \log t - 1 & \text{if} \quad t > 0, \\ +\infty & \text{otherwise.} \end{cases}$$

The corresponding function $\varphi$ is called the logarithmic-quadratic function. It enjoys attractive properties for developing efficient algorithms (see [8] and [7] for the properties of this function).

Another function $h$ which is also often used in the literature (see, for example, [11] and

[26]) is

$$h(t) = \begin{cases} t\log t - t + 1 & \text{if} \quad t > 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Associated with $\varphi$, we consider the $\varphi-$divergence proximal distance

$$d_\varphi(x,y) = \sum_{j=1}^{n} y_j^2 \varphi\left(\frac{x_j}{y_j}\right), \quad \forall x, y \in I\!R^n_{++},$$

and for any $x, y \in \operatorname{int} C$, we define the distance-like function $D_\varphi$ by

$$D_\varphi(x,y) = d_\varphi(l(x), l(y)), \quad \forall x, y \in \operatorname{int} C,$$

where $l(x) = (l_1(x), \dots, l_n(x))$ and $l_j(x) = b_j - \langle a_j, x\rangle$, $j = 1, \dots, n$.
It is easy to see that

$$D_\varphi(x,y) = \mu D_h(x,y) + \frac{\nu}{2}\|A(x-y)\|^2, \quad \forall x, y \in \operatorname{int} C,$$

showing the barrier and regularization terms. Note that $A$ being of full rank, the function $(u,v) \to \langle A^T Au, v\rangle$ defines on $I\!R^n$ an inner product denoted $\langle u, v\rangle_A$ with $\|u\|_A := \|Au\| = \langle Au, Au\rangle^{\frac{1}{2}}$, so that we can write

$$D_\varphi(x,y) = \mu D_h(x,y) + \frac{\nu}{2}\|x-y\|_A^2, \quad \forall x, y \in \operatorname{int} C. \tag{4}$$

With this distance, the basic iteration of our method can be written as follows: *Given $x^k \in \operatorname{int} C$, find $x^{k+1} \in \operatorname{int} C$, the solution of the unconstrained problem*

$(P_k)$ $\qquad \min_y\{c_k f(x^k, y) + D_\varphi(y, x^k)\}$.

This method has been intensively studied by Auslender et al. for solving particular equilibrium problems as the convex optimization problems (see, for example, [4], [6], [8], [10]) and the variational inequality problems (see, for example, [3], [5], [6], [9]). See also ([11], [12], [14], [24], [31]).

Our aim in this paper is to study extragradient methods based on problem $(P_k)$ for solving the equilibrium problem EP where $C = \{x \mid Ax \le b\}$. In the next two sections, we assume that $\varphi$ is of the form (3) with $h$ a function satisfying properties $(a) - (e)$.

# 2 The interior proximal extragradient algorithm

Let us recall some preliminary results which will be used later in our analysis. First, for all $x, y, z \in I\!R^n$, it is easy to see that

$$\|x-y\|_A^2 + \|x-z\|_A^2 = \|y-z\|_A^2 + 2\langle x-z, x-y\rangle_A. \tag{5}$$

Next, let us introduce a lemma that plays a key role in the convergence analysis.

**Lemma 2.1.** *For all $x, y \in intC$ and $z \in C$, it holds that*
*(i) $D_\varphi(\cdot, y)$ is differentiable and strongly convex on $intC$ with modulus $\nu$, i.e.,*

$$\langle \nabla_1 D_\varphi(x,p) - \nabla_1 D_\varphi(y,p),\, x - y \rangle \geq \nu \, \|x-y\|_A^2 \quad \forall p \in \, intC,$$

*where $\nabla_1 D_\varphi(x,p)$ denotes the gradient of $D_\varphi(\cdot,p)$ at $x$.*
*(ii) $D_\varphi(x,y) = 0$ if and only if $x = y$,*
*(iii) $\nabla_1 D_\varphi(x,y) = 0$ if and only if $x = y$,*
*(iv)* $\langle \nabla_1 D_\varphi(x,y), x - z \rangle \geq \left(\dfrac{\nu+\mu}{2}\right)(\|x-z\|_A^2 - \|y-z\|_A^2) + \left(\dfrac{\nu-\mu}{2}\right)\|x-y\|_A^2.$

*Proof.* See Proposition 2.1. in [8] and Proposition 4.1 in [14]. ∎

The next result is crucial to establish the existence and the characterization of a solution to subproblem $(P_k)$.

**Theorem 2.1.** *Let $F : C \to I\!R \cup \{+\infty\}$ be a closed proper convex function such that $domF \cap intC \neq \emptyset$. Given $x \in intC$ and $c_k > 0$. Then there exists a unique $y \in intC$ such that*

$$y = \arg\min_z \{c_k F(z) + D_\varphi(z,x)\}$$

*and*

$$0 \in c_k \partial F(y) + \nabla_1 D_\varphi(y,x),$$

*where $\partial F(y)$ denotes the subdifferential of $F$ at $y$.*

*Proof.* See Lemma 3.3 in [4]. ∎

Now we present a first interior proximal extragradient algorithm for solving EP.

**The Interior Proximal Extragradient Algorithm (IPE).**

**Step 0.** Take $x^0 \in C$, choose $c_0 > 0$ and a couple of positive parameters $(\nu, \mu)$ such that $\nu > \mu$ and set $k = 0$. The corresponding distance function is denoted $D_\varphi$.

**Step 1.** Solve the interior proximal convex program

$$\min_y \{c_k f(x^k, y) + D_\varphi(y, x^k)\} \tag{6}$$

to obtain its unique solution $y^k$. If $y^k = x^k$, then stop: $x^k$ is a solution to EP.
Otherwise, go to Step 2.

**Step 2.** Solve the interior proximal convex program

$$\min_y \{c_k f(y^k, y) + D_\varphi(y, x^k)\} \tag{7}$$

to obtain its unique solution $x^{k+1}$.

**Step 3.** Increase $k$ by 1, choose $c_k > 0$ and return to Step 1.

First observe that the algorithm is well defined. Indeed, thanks to Theorem 2.1 with function $F$ defined by $f(x^k, \cdot)$ and $f(y^k, \cdot)$, respectively, the subproblems (6) and (7) have a unique solution and

$$0 \in c_k \partial_2 f(x^k, y^k) + \nabla_1 D_\varphi(y^k, x^k) \quad \text{and} \quad 0 \in c_k \partial_2 f(y^k, x^{k+1}) + \nabla_1 D_\varphi(x^{k+1}, x^k),$$

where $\partial_2 f(x, y)$ denotes the subdifferential of $f(x, \cdot)$ at $y$.
Consequently, using the definition of the subdifferential, we can write

$$c_k f(x^k, y) \geq c_k f(x^k, y^k) + \langle \nabla_1 D_\varphi(y^k, x^k), y^k - y \rangle, \quad \forall y \in C, \tag{8}$$

and

$$c_k f(y^k, y) \geq c_k f(y^k, x^{k+1}) + \langle \nabla_1 D_\varphi(x^{k+1}, x^k), x^{k+1} - y \rangle, \quad \forall y \in C. \tag{9}$$

In the next proposition, we justify the stopping criterion.

**Proposition 2.1.** *If $y^k = x^k$, then $x^k$ is a solution to EP.*

*Proof.* When $y^k = x^k$, the inequality (8) becomes

$$c_k f(x^k, y) \geq c_k f(x^k, x^k) + \langle \nabla_1 D_\varphi(x^k, x^k), x^k - y \rangle, \quad \forall y \in C.$$

Since $f(x^k, x^k) = 0$ and $\nabla_1 D_\varphi(x^k, x^k) = 0$ (by Lemma 2.1(iii)), it follows that

$$c_k f(x^k, y) \geq 0, \quad \forall y \in C,$$

i.e., that $x^k$ is a solution to EP. ∎

In order to prove the convergence of the algorithm, let us first consider the dual problem of EP, namely:

$$\text{Find} \quad x^* \in C \quad \text{such that} \quad f(y, x^*) \leq 0, \quad \forall y \in C.$$

We denote by $S_d^*$ the solution set of this problem and we recall that $S_d^* = S^*$ when $f$ is pseudomonotone (see Proposition 1.1.2, page 5 in [19]). Now we are in a position to prove the convergence of the IPE algorithm.

**Theorem 2.2.** *Assume that $\nu > 5\mu$ and that there exist two positive parameters $c_1$ and $c_2$ such that*

$$\forall x, y, z \in C \quad f(x, y) + f(y, z) \geq f(x, z) - c_1 \|y - x\|_A^2 - c_2 \|z - y\|_A^2. \tag{10}$$

*Then the following statements hold:*
*(i) If* $x^* \in S_d^*$*, then*

$$\Delta(x^k) - \Delta(x^{k+1}) \geq \left(\frac{1}{2} - \frac{2\mu + c_k\, c_1}{\nu - \mu}\right) \|y^k - x^k\|_A^2 + \left(\frac{1}{2} - \frac{\mu + c_k\, c_2}{\nu - \mu}\right) \|x^{k+1} - y^k\|_A^2, \quad (11)$$

*where* $\Delta(x) = \left(\frac{1}{2} + \dfrac{\mu}{\nu - \mu}\right) \|x - x^*\|_A^2$*;*
*(ii) If* $0 < c \leq c_k < \min\left\{\dfrac{\nu - 5\mu}{2c_1}, \dfrac{\nu - 3\mu}{2c_2}\right\}$*, then the sequence* $\{x^k\}$ *is bounded and every limit point of* $\{x^k\}$ *is a solution to EP. In addition, if* $S_d^* = S^*$*, then the whole sequence* $\{x^k\}$ *tends to a solution of EP.*

*Proof.* (i) Take any $x^* \in S_d^*$ and consider the inequality (8) with $y = x^{k+1}$. Then

$$c_k f(x^k, x^{k+1}) - c_k f(x^k, y^k) \geq \langle \nabla_1 D_\varphi(y^k, x^k), y^k - x^{k+1}\rangle.$$

Using first Lemma 2.1 $(iv)$ to the right hand side of this inequality and then the equality (5) with $x = y^k, y = x^{k+1}$ and $z = x^k$, we obtain successively

$$\begin{aligned} c_k f(x^k, x^{k+1}) - c_k f(x^k, y^k) &\geq \theta\,(\|y^k - x^{k+1}\|_A^2 - \|x^k - x^{k+1}\|_A^2) + (\nu - \theta)\|y^k - x^k\|_A^2 \\ &= \theta\,(-\|y^k - x^k\|_A^2 + 2\langle y^k - x^k, y^k - x^{k+1}\rangle_A) + (\nu - \theta)\|y^k - x^k\|_A^2 \\ &= -\mu\,\|y^k - x^k\|_A^2 + (\nu + \mu)\langle y^k - x^k, y^k - x^{k+1}\rangle_A, \qquad (12) \end{aligned}$$

where $\theta = \frac{\nu + \mu}{2}$.

On the other hand, considering the inequality (9) with $y = x^*$, we have

$$c_k f(y^k, x^*) - c_k f(y^k, x^{k+1}) \geq \langle \nabla_1 D_\varphi(x^{k+1}, x^k), x^{k+1} - x^*\rangle.$$

Using again Lemma 2.1 $(iv)$ and the equality (5) with $x = x^{k+1} \in \operatorname{int} C, y = x^k \in \operatorname{int} C$, $z = x^* \in C$, we obtain that

$$\begin{aligned} c_k f(y^k, x^*) - c_k f(y^k, x^{k+1}) &\geq \theta(\|x^{k+1} - x^*\|_A^2 - \|x^k - x^*\|_A^2) + (\nu - \theta)\|x^{k+1} - x^k\|_A^2 \\ &= \theta(\|x^{k+1} - x^*\|_A^2 - \|x^k - x^*\|_A^2) + (\nu - \theta)\|x^k - x^*\|_A^2 \\ &- (\nu - \theta)\|x^{k+1} - x^*\|_A^2 \\ &+ 2(\nu - \theta)\langle x^{k+1} - x^k, x^{k+1} - x^*\rangle_A \\ &= \mu\|x^{k+1} - x^*\|_A^2 - \mu\|x^k - x^*\|_A^2 \\ &+ (\nu - \mu)\langle x^{k+1} - x^k, x^{k+1} - x^*\rangle_A, \end{aligned}$$

Noting that $\nu - \mu > 0$ and $f(y^k, x^*) \leq 0$ because $x^* \in S_d^*$, we deduce from the above inequality that

$$\begin{aligned}
\langle x^{k+1} - x^k, x^* - x^{k+1}\rangle_A &\geq \frac{c_k}{\nu - \mu} f(y^k, x^{k+1}) \\
&+ \frac{\mu}{\nu - \mu}\|x^{k+1} - x^*\|_A^2 - \frac{\mu}{\nu - \mu}\|x^k - x^*\|_A^2 \\
&\geq \frac{c_k}{\nu - \mu}[f(x^k, x^{k+1}) - f(x^k, y^k)] \\
&- \frac{c_k c_1}{\nu - \mu}\|y^k - x^k\|_A^2 - \frac{c_k c_2}{\nu - \mu}\|x^{k+1} - y^k\|_A^2 \\
&+ \frac{\mu}{\nu - \mu}\|x^{k+1} - x^*\|_A^2 - \frac{\mu}{\nu - \mu}\|x^k - x^*\|_A^2 \qquad (13)
\end{aligned}$$

where the second inequality is obtained after using assumption (10) with $x = x^k, y = y^k$ and $z = x^{k+1}$.

On the other hand, from equality (5) with $x = x^{k+1}$, $y = x^*$, $z = x^k$ and then with $x = y^k$, $y = x^{k+1}$ and $z = x^k$, we deduce

$$\|x^k - x^*\|_A^2 - \|x^{k+1} - x^*\|_A^2 = \|x^{k+1} - x^k\|_A^2 + 2\langle x^{k+1} - x^k, x^* - x^{k+1}\rangle_A, \qquad (14)$$

and

$$\|x^{k+1} - x^k\|_A^2 = -2\langle y^k - x^k, y^k - x^{k+1}\rangle_A + \|x^{k+1} - y^k\|_A^2 + \|y^k - x^k\|_A^2. \qquad (15)$$

Finally, using successively (14), (13), (12), (15), and the inequality

$$\langle y^k - x^k, y^k - x^{k+1}\rangle_A \geq -\|y^k - x^k\|_A \|y^k - x^{k+1}\|_A \geq -\frac{1}{2}\|y^k - x^k\|_A^2 - \frac{1}{2}\|y^k - x^{k+1}\|_A^2,$$

we obtain the following equalities and inequalities

$$\begin{aligned}
\Delta(x^k) - \Delta(x^{k+1}) &= \tfrac{1}{2}\|x^{k+1} - x^k\|_A^2 + \langle x^{k+1} - x^k, x^* - x^{k+1}\rangle_A \\
&+ \tfrac{\mu}{\nu-\mu}\|x^k - x^*\|_A^2 - \tfrac{\mu}{\nu-\mu}\|x^{k+1} - x^*\|_A^2 \\
&\geq \tfrac{1}{2}\|x^{k+1} - x^k\|_A^2 + \tfrac{c_k}{\nu-\mu}[f(x^k, x^{k+1}) - f(x^k, y^k)] \\
&- \tfrac{c_k c_1}{\nu-\mu}\|y^k - x^k\|_A^2 - \tfrac{c_k c_2}{\nu-\mu}\|x^{k+1} - y^k\|_A^2 \\
&\geq \tfrac{1}{2}\|x^{k+1} - x^k\|_A^2 + \tfrac{\nu+\mu}{\nu-\mu}\langle y^k - x^k, y^k - x^{k+1}\rangle_A \\
&- \tfrac{\mu + c_k c_1}{\nu-\mu}\|y^k - x^k\|_A^2 - \tfrac{c_k c_2}{\nu-\mu}\|x^{k+1} - y^k\|_A^2 \\
&= (\tfrac{1}{2} - \tfrac{\mu + c_k c_1}{\nu-\mu})\|y^k - x^k\|_A^2 + (\tfrac{1}{2} - \tfrac{c_k c_2}{\nu-\mu})\|x^{k+1} - y^k\|_A^2 \\
&+ \tfrac{2\mu}{\nu-\mu}\langle y^k - x^k, y^k - x^{k+1}\rangle_A \\
&\geq (\tfrac{1}{2} - \tfrac{2\mu + c_k c_1}{\nu-\mu})\|y^k - x^k\|_A^2 + (\tfrac{1}{2} - \tfrac{\mu + c_k c_2}{\nu-\mu})\|x^{k+1} - y^k\|_A^2
\end{aligned}$$

(ii) Since $\nu > 5\mu$ and $0 < c_k < \min\{\frac{\nu - 5\mu}{2c_1}, \frac{\nu - 3\mu}{2c_2}\}$, we have $\frac{1}{2} - \frac{2\mu + c_k c_1}{\nu - \mu} > 0$ and $\frac{1}{2} - \frac{\mu + c_k c_2}{\nu - \mu} > 0$. Consequently, from part $(i)$, we obtain that

$$\Delta(x^k) - \Delta(x^{k+1}) \geq \left(\frac{1}{2} - \frac{2\mu + c_k c_1}{\nu - \mu}\right) \|y^k - x^k\|_A^2 \geq 0, \quad \forall k.$$

This implies that the positive sequence $\{\Delta(x^k)\}$ is nonincreasing. Hence this sequence converges in $I\!R$ and consequently, is bounded and such that

$$\lim_{k\to+\infty} \|y^k - x^k\|_A^2 = 0. \tag{16}$$

Let $\bar{x}$ be a limit point of $\{x^k\}$. Then $\bar{x} = \lim_{j\to+\infty} x^{k_j}$, and, by (16), $\bar{x} = \lim_{j\to+\infty} y^{k_j}$. Using (8) and Lemma 2.1$(iv)$, we have for all $y \in C$ and all $j$ that

$$\begin{aligned} c_{k_j} f(x^{k_j}, y) - c_{k_j} f(x^{k_j}, y^{k_j}) &\geq \langle \nabla_1 D_\varphi(y^{k_j}, x^{k_j}), y^{k_j} - y \rangle, \\ &\geq \frac{\nu + \mu}{2}(\|y^{k_j} - y\|_A^2 - \|x^{k_j} - y\|_A^2). \end{aligned} \tag{17}$$

Taking $j \to +\infty$ in (17) and noting that $f(\bar{x}, \bar{x}) = 0$ and $0 < c < c_k \leq \min\{\frac{\nu-5\mu}{2c_1}, \frac{\nu-3\mu}{2c_2}\}$, we obtain

$$f(\bar{x}, y) \geq 0, \quad \forall y \in C,$$

which means that $\bar{x}$ is a solution to EP.

Suppose now that $S_d^* = S^*$. Then the whole sequence $\{x^k\}$ converges to $\bar{x}$. Indeed, defining $\Delta(x^k)$ with $x^* = \bar{x} \in S_d^*$, we have $\Delta(x^{k_j}) \to 0$ because $x^{k_j} \to \bar{x}$. So the sequence $\Delta(x^k)$ being nonincreasing, the whole sequence $\{\Delta(x^k)\}$ also converges to 0 and thus $\|x^k - \bar{x}\|_A \to 0$, i.e., $x^k \to \bar{x}$. ■

When the function $f(x, \cdot)$ is nonsmooth, it can be difficult to solve subproblems (6) and (7). In that case, we can use a bundle strategy as in nonsmooth optimization [8] (see also [24], [25]). For subproblem (6), the idea is to approximate the function $f(x^k, \cdot)$ from below by a piecewise linear convex function $\psi^k$ and to take for the next iterate the solution $y^k$ of the following subproblem

$$\min_y \{c_k \psi^k(y) + D_\varphi(y, x^k)\}. \tag{18}$$

More precisely, $\psi^k$ is constructed, thanks to a sequence $\psi_i^k$, $i = 1, 2, \ldots$ as follows:

The starting data are $y_0^k = x^k$, $g_0^k \in \partial_2 f(x^k, y_0^k)$ and $\psi_1^k(y) = f(x^k, y_0^k) + \langle g_0^k, y - y_0^k \rangle$ for all $y \in I\!R^n$.

Suppose at iteration $i \geq 1$ that $\psi_i^k$ is known. Then $\psi_{i+1}^k$ is obtained by the following steps:

**Step 1.** Solve subproblem (18) with $\psi^k$ replaced by $\psi_i^k$ to get $y_i^k$; set $d_i^k = -\nabla_1 D_\varphi(y_i^k, x^k)$

and $l_i^k(y) = \psi_i^k(y_i^k) + \langle d_i^k, y - y_i^k \rangle$.

**Step 2.** Choose $\psi_{i+1}^k : I\!R^n \to I\!R$ as a piecewise linear convex function satisfying the conditions:
(C1) $l_i^k \le \psi_{i+1}^k \le f(x^k, \cdot)$
(C2) $f(x^k, y_i^k) + \langle g_i^k, y - y_i^k \rangle \le \psi_{i+1}^k(y) \quad \forall y \in I\!R^n \quad$ with $g_i^k \in \partial_2 f(x^k, y_i^k)$.

It can be proven (see Theorem 3.2 in [8]) that after finitely many steps $i$, this algorithm gives a point $y_i^k$ and a model $\psi_i^k$ such that

$$f(x^k, y_i^k) \le \eta\, \psi_i^k(y_i^k) \quad (0 < \eta < 1).$$

In that case we consider that the approximate function $\psi_i^k$ is appropriate and we set $\psi^k = \psi_i^k$ and $y^k = y_i^k$.
Next, in order to obtain an efficient algorithm, the functions $\psi_i^k$ must be chosen in such a way that the subproblems (18) (with $\psi^k$ replaced by $\psi_i^k$) can be easily solved. In [8] it is shown that for $C = I\!R_+^n$ and

$$\psi_{i+1}^k(y) = \max\{l_i^k(y), f(x^k, y_i^k) + \langle g_i^k, y - y_i^k \rangle\} \quad \text{for all } y \in I\!R^n,$$

the conditions (C1) and (C2) are satisfied and the subproblems (18) can be simplified and reduced to minimizing a function of a single variable (see also [25], for other examples and properties of the models $\psi_i^k$). Since a similar strategy can be developed for solving subproblem (7), we finally obtain an implementable algorithm whose convergence results can be proven exactly as in [25].

# 3 The interior proximal linesearch extragradient method

Convergence of the IPE algorithm requires that the function $f$ satisfies condition (10). This condition depends on two positive parameters $c_1$ and $c_2$ and in some cases, they are unknown or difficult to approximate. So in this section, we modify the second step of the algorithm using a linesearch and an hyperplane projection step in order to obtain the convergence without assuming that condition (10) is satisfied. When a quadratic regularization term is used, this strategy has been initiated by Konnov [18], [19], [20] in the particular case where $f$ is differentiable. The nondifferentiable convex case has been recently considered by Tran et al. [30]. In this section, we replace the usual quadratic proximal distance by the $\varphi$-divergence proximal distance $D_\varphi$ defined in (4), and as in [30], we suppose that

**(A1)** $C$ is contained in an open convex set $\Lambda \subset I\!R^n$,
**(A2)** $f : \Lambda \times \Lambda \to I\!R$ is a continuous function satisfying $f(x, x) = 0$ for each $x \in \Lambda$ and

$f(x, \cdot)$ is convex for each $x \in \Lambda$.

Before giving the algorithm and in order to obtain more flexibility in the choice of the steplength, we introduce a sequence $\{\gamma_k\}$ which satisfies the properties

$$\gamma_k \in (0, 2) \quad \forall k = 0, 1, \ldots \text{ and } \liminf_{k \to +\infty} \gamma_k (2 - \gamma_k) > 0. \tag{19}$$

Obviously, $\gamma_k = 1$ for all $k$ is an example of such a sequence.

**The Interior Proximal Linesearch Extragradient Algorithm (IPLE).**

**Step 0.** Take $x^0 \in \operatorname{int} C$, choose $\theta \in (0, 1)$, $\tau \in (0, 1)$, $\alpha \in (0, 1)$, $c > 0$, $c_0 \geq c > 0$ and choose positive parameters $\nu, \mu$ such that $\nu > \mu$. Set $k = 0$.

**Step 1.** Solve the convex program

$$\min_y \{c_k f(x^k, y) + D_\varphi(y, x^k)\} \tag{20}$$

to obtain its unique solution $y^k$. If $y^k = x^k$, then STOP: $x^k$ is a solution to EP. Otherwise, go to Step 2.

**Step 2.** Find the smallest nonnegative integer $m$ such that

$$f(z^{k,m}, x^k) - f(z^{k,m}, y^k) \geq \frac{\alpha}{c_k} D_\varphi(y^k, x^k), \tag{21}$$

where $z^{k,m} = (1 - \theta^m)x^k + \theta^m y^k$. Set $z^k = z^{k,m}$ and go to Step 3.

**Step 3.** Take $g^k \in \partial_2 f(z^k, x^k)$.
Compute $\sigma_k = \dfrac{f(z^k, x^k)}{\|g^k\|^2}$ and $x^{k+1} = (1 - \tau)\, x^k + \tau\, P_C(x^k - \gamma_k\, \sigma_k\, g^k)$, where $P_C(z)$ denotes the orthogonal projection of $z$ over $C$.

**Step 4.** Increase $k$ by 1, choose $c_k \geq c > 0$ and return to Step 1.

**Remark 3.1.** Algorithm (IPLE) is an extension of the combined relaxation method proposed by Konnov [20] for solving a differentiable monotone equilibrium problem. The Armijo-backtracking linesearch (Step 2) is slightly different from Konnov's one to take account of the $\varphi$-divergence proximal distance and of the fact that $f$ is nondifferentiable. The hyperplane projection step (Step 3) is similar, a subgradient $g_k$ of $f(z^k, \cdot)$ replacing the gradient of $f(z^k, \cdot)$.

In order to see that Algorithm (IPLE) is well defined, first observe that, by Theorem 2.1, the solution $y^k$ of problem (20) exists and is unique. Furthermore, if $x^k \in \operatorname{int} C$, then $x^{k+1}$ also belongs to $\operatorname{int} C$ because $\tau \in (0, 1)$. Finally to state that the linesearch is also well defined, we introduce the following lemma:

**Lemma 3.1.** *Assume that $y^k \neq x^k$ for some $k$. Then the next three properties hold:*
*(i) There exists a nonnegative integer $m$ satisfying (21);*
*(ii) $f(z^k, x^k) > 0$;*
*(iii) $0 \notin \partial_2 f(z^k, x^k)$.*

*Proof.* (i) By contradiction, we suppose that statement $(i)$ is not true, i.e., that for all nonnegative integer $m$, we have the inequality

$$f(z^{k,m}, x^k) - f(z^{k,m}, y^k) < \frac{\alpha}{c_k} D_\varphi(y^k, x^k).$$

Let $m \to +\infty$. Then $z^{k,m} \to x^k$ and because $f$ is continuous on $C \times C$ and $f(x, x) = 0$ for all $x \in C$, we obtain

$$c_k f(x^k, y^k) + \alpha D_\varphi(y^k, x^k) \geq 0. \tag{22}$$

On the other hand, because $y^k$ is a solution of (20), we have

$$c_k f(x^k, y^k) + D_\varphi(y^k, x^k) \leq c_k f(x^k, y) + D_\varphi(y, x^k), \quad \forall y \in \operatorname{int} C.$$

Taking $y = x^k$ in this inequality and noting that $f(x^k, x^k) = 0$ and $D_\varphi(x^k, x^k) = 0$, we deduce

$$c_k f(x^k, y^k) + D_\varphi(y^k, x^k) \leq 0.$$

Combining this inequality and (22) and noting that $D_\varphi(y^k, x^k) > 0$ because $y^k \neq x^k$, we obtain $\alpha \geq 1$. But this contradicts the assumption and thus there exists a nonnegative integer $m$ satisfying (21).
(ii) Because $f$ is convex with respect to the second argument, it follows from the definition of $z^k$ that

$$0 = f(z^k, z^k) \leq (1 - \theta^m) f(z^k, x^k) + \theta^m f(z^k, y^k). \tag{23}$$

Hence, using (21), we obtain

$$f(z^k, x^k) \geq \theta^m (f(z^k, x^k) - f(z^k, y^k)) \geq \frac{\alpha \theta^m}{c_k} D_\varphi(y^k, x^k) > 0.$$

(iii) By contradiction, let us suppose that $0 \in \partial_2 f(z^k, x^k)$, i.e., that

$$f(z^k, y) \geq f(z^k, x^k), \quad \forall y \in C.$$

Taking $y = z^k$, we obtain that $f(z^k, x^k) \leq 0$. This contradicts $(ii)$, and so $(iii)$ holds. ∎

The following lemmas are the key results in our analysis of the convergence of the algorithm (IPLE).

**Lemma 3.2.** *(i) The sequence $\{x^k\}$ is bounded and for every solution $x^* \in S_d^*$, the following inequality holds*

$$\|x^{k+1} - x^*\|^2 \le \|x^k - x^*\|^2 - \tau\gamma_k(2-\gamma_k)(\sigma_k\|g^k\|)^2.$$

*(ii)* $\displaystyle\sum_{k=0}^{+\infty} \gamma_k(2-\gamma_k)(\sigma_k\|g^k\|)^2 < +\infty.$

*Proof.* (i) Take $x^* \in S_d^*$. Using successively the definition of $x^{k+1}$, the convexity of $\|\cdot\|^2$ and the nonexpansiveness of the projection, we have

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 &= \|(1-\tau)x^k + \tau P_C(x^k - \gamma_k\sigma_k g^k) - x^*\|^2 \\
&= \|(1-\tau)(x^k - x^*) + \tau[P_C(x^k - \gamma_k\sigma_k g^k) - x^*]\|^2 \\
&\le (1-\tau)\|x^k - x^*\|^2 + \tau\|P_C(x^k - \gamma_k\sigma_k g^k) - x^*\|^2 \\
&\le (1-\tau)\|x^k - x^*\|^2 + \tau\|x^k - \gamma_k\sigma_k g^k - x^*\|^2 \\
&= \|x^k - x^*\|^2 + \tau\|\gamma_k\sigma_k g^k\|^2 - 2\tau\langle\gamma_k\sigma_k g^k, x^k - x^*\rangle. \qquad (24)
\end{aligned}$$

On the other hand, because $g^k \in \partial_2 f(z^k, x^k)$, it follows that

$$f(z^k, x^*) \ge f(z^k, x^k) + \langle g^k, x^* - x^k\rangle.$$

Furthermore, since $f(z^k, x^*) \le 0$ and $\sigma_k = \dfrac{f(z^k, x^k)}{\|g^k\|^2}$, we obtain from the previous inequality that

$$\langle g^k, x^k - x^*\rangle \ge \sigma_k\|g^k\|^2.$$

Using this inequality in (24), we deduce that

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 &\le \|x^k - x^*\|^2 + \tau\|\gamma_k\sigma_k g^k\|^2 - 2\tau\gamma_k\|\sigma_k g^k\|^2 \\
&= \|x^k - x^*\|^2 - \tau\gamma_k(2-\gamma_k)(\sigma_k\|g^k\|)^2.
\end{aligned}$$

In particular, this implies that the sequence $\{x^k\}$ is bounded.
(ii) We easily deduce from part $(i)$ that for all $m \in I\!N$, we have

$$0 \le \sum_{k=0}^{m} \tau\gamma_k(2-\gamma_k)(\sigma_k\|g^k\|)^2 \le \|x^0 - x^*\|^2 - \|x^{m+1} - x^*\|^2 \le \|x^0 - x^*\|^2.$$

So, taking $m \to +\infty$, we obtain

$$\sum_{k=0}^{+\infty} \gamma_k(2-\gamma_k)(\sigma_k\|g^k\|)^2 < +\infty.$$

∎

**Lemma 3.3.** *Let $\bar{x}$ be a limit point of $\{x^k\}$ and let $x^{k_j} \to \bar{x}$. Then the sequences $\{y^{k_j}\}$, $\{z^{k_j}\}$ and $\{g^{k_j}\}$ are bounded providing that $c_{k_j} \le \bar{c}$ for all $j$.*

*Proof.* Since the sequence $\{x^k\}$ is bounded, it suffices to prove that there exists $M$ such that $\|x^{k_j} - y^{k_j}\| \le M$ for $j$ large enough to obtain that the sequence $\{y^{k_j}\}$ is bounded. Without loss of generality, we suppose, that $y^{k_j} \neq x^{k_j}$ for all $j$, and we set $S(y) = c_{k_j} f(x^{k_j}, y) + \bar{D}_\varphi(y, x^{k_j})$.

Since $f(x^{k_j}, \cdot)$ is convex and since, by Lemma 2.1($i$), the function $D_\varphi(\cdot, x^{k_j})$ is strongly convex on $\operatorname{int} C$ with modulus $\nu > 0$, we have, for all $y_1, y_2 \in \operatorname{int} C$, $g_1 \in \partial S(y_1)$ and $g_2 \in \partial S(y_2)$ that

$$\langle g_1 - g_2, y_1 - y_2 \rangle \ge \nu \|y_1 - y_2\|_A^2 \ge \nu\, \lambda_{min}(A^T A) \|y_1 - y_2\|^2,$$

where $\lambda_{min}(A^T A)$ denotes the smallest eigenvalue of the matrix $A^T A$.

Taking $y_1 = x^{k_j}$ and $y_2 = y^{k_j}$ and noting that $0 \in \partial S(y^{k_j})$ by definition of $y^{k_j}$, we deduce from the previous inequality that

$$\forall g_j \in \partial S(x^{k_j}) \quad \nu\, \lambda_{min}(A^T A) \|x^{k_j} - y^{k_j}\|^2 \le \langle g_j, x^{k_j} - y^{k_j} \rangle \le \|g_j\| \, \|x^{k_j} - y^{k_j}\|.$$

Since $y^{k_j} \neq x^{k_j}$, and since, by Lemma 2.1($iii$), $\nabla_1 D_\varphi(x^{k_j}, x^{k_j}) = 0$, we can write

$$\forall g_j \in \partial_2 f(x^{k_j}, x^{k_j}) \quad \nu\, \lambda_{min}(A^T A) \|x^{k_j} - y^{k_j}\| \le \|g_j\|. \tag{25}$$

On the other hand, let the sequence $\{f_j\}_{j \in \mathbb{N}}$ be defined for all $j \in \mathbb{N}$ by $f_j = f(x^{k_j}, \cdot)$. By continuity of $f$, this sequence of convex functions converges pointwise to the convex function $f(\bar{x}, \cdot)$. Since $x^{k_j} \to \bar{x} \in \Lambda$ and since $f(\bar{x}, \cdot)$ is finite on $\Lambda$, it follows from Theorem 24.5 in [28] that there exists an index $j_0$ such that

$$\forall j \ge j_0, \quad \partial f(x^{k_j}, x^{k_j}) \subset \partial f(\bar{x}, \bar{x}) + B,$$

where $B$ is the closed Euclidean unit ball of $\mathbb{R}^n$. Since $g_j \in \partial_2 f(x^{k_j}, x^{k_j})$ for all $j$ and $\partial_2 f(\bar{x}, \bar{x})$ is bounded, this inclusion implies that the right-hand side of (25) is bounded. So there exists $M > 0$ such that $\|x^{k_j} - y^{k_j}\| \le M$ for all $j \ge j_0$, and the sequence $\{y^{k_j}\}$ is bounded.
The sequence $\{z^{k_j}\}$ being a convex combination of $x^{k_j}$ and $y^{k_j}$, it is very easy to see that the sequence $\{z^{k_j}\}$ is also bounded and that there exists a subsequence of $\{z^{k_j}\}$, again denoted $\{z^{k_j}\}$, that converges to $\bar{z} \in C$.

Finally, to prove that the sequence $\{g^{k_j}\}$ is bounded, we proceed exactly as for the sequence $\{g_j\}$ but this time with the sequence $\{f_j\}_{j \in \mathbb{N}}$ defined for all $j \in \mathbb{N}$ by $f_j = f(z^{k_j}, \cdot)$. ■

Thanks to Lemmas 3.2 and 3.3, we can deduce the following convergence result.

**Theorem 3.1.** *Assume that the properties (A1) and (A2) are satisfied and that $0 < c \leq c_k \leq \bar{c}$ for all $k$. Then the following statements hold:*
*(i) Every limit point of $\{x^k\}$ is a solution to EP.*
*(ii) If $S^* = S_d^*$ then the whole sequence $\{x^k\}$ converges to a solution of EP.*

*Proof.* (i) Let $\bar{x}$ be a limit point of $\{x^k\}$ and $x^{k_j} \to \bar{x}$. Applying Lemma 3.2 $(ii)$ and (19), we deduce that

$$\sigma_{k_j} \|g^{k_j}\| \to 0,$$

i.e., by using the definition of $\sigma_{k_j}$, that

$$\frac{f(z^{k_j}, x^{k_j})}{\|g^{k_j}\|} \to 0.$$

Since, by Lemma 3.3, the sequence $\{g^{k_j}\}$ is bounded, we obtain that $f(z^{k_j}, x^{k_j}) \to 0$ as $j \to +\infty$. Furthermore, it follows from (23) that for all $j$,

$$f(z^{k_j}, x^{k_j}) - f(z^{k_j}, y^{k_j}) \leq \frac{1}{\theta^m} f(z^{k_j}, x^{k_j}).$$

Combining this inequality with (21) and noting, from (4), that $D_\varphi(y^{k_j}, x^{k_j}) \geq \frac{\nu}{2}\|y^{k_j} - x^{k_j}\|_A^2$, we have

$$\frac{\alpha\, \nu}{2c_{k_j}} \|y^{k_j} - x^{k_j}\|_A^2 \leq \frac{1}{\theta^m} f(z^{k_j}, x^{k_j}).$$

Consequently, since $c_{k_j} \leq \bar{c}$ for all $j$ and $f(z^{k_j}, x^{k_j}) \to 0$ as $j \to +\infty$, we have

$$\lim_{j \to +\infty} \|y^{k_j} - x^{k_j}\|_A^2 = 0,$$

and $y^{k_j} \to \bar{x}$ because $x^{k_j} \to \bar{x}$. Finally, using Theorem 2.1 and Lemma 2.1, we obtain again inequality (17). Taking the limit $j \to +\infty$ in (17), using the continuity of $f$ and observing that $f(\bar{x}, \bar{x}) = 0$ and $0 < c \leq c_{k_j} \leq \bar{c}$ for all $j$, we deduce immediately that $f(\bar{x}, y) \geq 0$ for all $y \in C$, i.e., $\bar{x}$ is a solution to EP.
(ii) Let $\bar{x} \in S^*$ be a limit point of the sequence $\{x^k\}$. Because $S^* = S_d^*$, it follows that $\bar{x} \in S_d^*$. Applying Lemma 3.2 $(i)$, we have that the sequence $\{\|x^k - \bar{x}\|\}_k$ is nonincreasing and since it has a subsequence converging to 0, it converges to zero. Hence, the whole sequence $\{x^k\}$ converges to $\bar{x} \in S^*$. ∎

**Remark 3.2.** The (IPE) and (IPLE) algorithms can be interpreted as prediction-correction methods. Indeed, Step 1 gives a prediction step while Step 2 for (IPE) and Step 3 for (IPLE) bring a correction step. Recently, such strategies have been intensively used for solving nonlinear complementarity problems (NLC), i.e., problems where the constraint set and the equilibrium function are given by

$$C = I\!R^n_+ \quad \text{and} \quad f(x,y) = \langle F(x), y - x \rangle, \quad \forall x, y \in C, \tag{26}$$

with $F : \mathbb{R}^n \to \mathbb{R}^n$ a (pseudo)monotone and continuous mapping (see, for example, [11], [26], [31] and [32]).

In these papers, the proximal-point iteration is used in the prediction step and consists, given $x^k$, in finding a solution $\tilde{x}^k$ of the system in $x$:

$$c_k F(x) + x - (1-\mu)x^k - \mu X_k^2 x^{-1} = \xi^k$$

when $\varphi(t) = \frac{\nu}{2}(t-1)^2 + \mu(t - \log t - 1)$ and of the system

$$c_k F(x) + x - x^k + \mu X_k \log \frac{x}{x^k} = \xi^k$$

when $\varphi(t) = \frac{1}{2}(t-1)^2 + \mu(t \log t - t + 1)$.
Here $X_k$ =diag$(x^k)$ and $x^{-1}$ denotes the vector $(x_1^{-1}, \ldots, x_n^{-1})$. Furthermore, the error $\xi^k$ must satisfy the condition: $\|\xi^k\| \leq \eta \|x^k - \tilde{x}^k\|$, $0 < \eta < 1$.
For the NLC problem, a practical choice for $\xi^k$ is to take $\xi^k = c_k F(x) - c_k F(x^k)$ so that the two previous systems become

$$c_k F(x^k) + x - (1-\mu)x^k - \mu X_k^2 x^{-1} = 0 \quad \text{and} \quad c_k F(x^k) + x - x^k + \mu X_k \log \frac{x}{x^k} = 0,$$

following the choice of $\varphi$. These systems are in fact the optimality conditions associated with Step 1 in our two algorithms when $C$ and $f$ are as in (26). Let us also observe that the correction step in [26] is similar to Step 2 in (IPE) and the ones in [11], [31], [32] to Step 3 in (IPLE) when no Armijo-backtracking linesearch is done. In this sense, we can say that our algorithms can be considered as generalizations of the algorithms mentioned above for solving the NLC problem. To end this section, let us also notice that a comparison with Solodov and Svaiter's method [29] is developed in [11].

# 4 Numerical Results

The aim of this section is to illustrate the proposed algorithms on a class of equilibrium problems where $C = \mathbb{R}^n_+$ and the equilibrium function $f : C \times C \to \mathbb{R}$ is of the form

$$f(x, y) = \langle Px + Qy + q, y - x \rangle,$$

with $P$ and $Q$ two matrices of dimension $n$. The corresponding equilibrium problem is a generalized form of an equilibrium problem defined by the Nash-Cournot equilibrium model considered in [27]. Let us also notice that this problem, in general, is not a variational inequality problem.

In order to fulfill the assumptions imposed in the previous sections, we suppose that the matrices $P$ and $Q$ are chosen such that $Q$ is symmetric positive definite and $Q - P$ is negative semidefinite. Under these assumptions, it can be proven (see [30], p.23) that $f$ is continuous and monotone, that $f(x,\cdot)$ is differentiable and convex for all $x \in C$ and that condition (2) is satisfied with $c_1 = c_2 = \frac{1}{2}\|P - Q\|$.

With this choice of function $f$, solving subproblem (6) amounts to solving the subproblem

$$\min_y \; g(y) + D_\varphi(y, x^k) \tag{27}$$

where $g(y) = c_k y^T Q y + c_k b^T y$ and $b = (P - Q)x + q$. The domain of the objective function of this problem is $I\!R^n_+$. So it is advisable to first solve its Fenchel dual

$$\min_u \; g^*(u) + D_\varphi(\cdot, x^k)^*(-u)$$

for the reason that its objective function is finite everywhere. Indeed the domain of $\varphi^*$ is equal to $I\!R$ and for all $u \in I\!R^n$, the functions

$$g^*(u) = \frac{1}{4c_k}\,\langle u - c_k b, Q^{-1}(u - c_k b)\rangle \quad \text{and} \quad D_\varphi(\cdot, x^k)^*(u) = \sum_{j=1}^n (x_j^k)^2\, \varphi^*(\frac{u_j}{x_j^k})$$

are finite.
Furthermore, since $\varphi^*(t)$ and $(\varphi^*)'(t)$ can be explicitly computed [8], it is possible to solve the Fenchel dual by using an efficient unconstrained optimization method. Let $u^*$ denote the solution of this problem. Then the solution $y^k$ of the subproblem (27) can be recovered by using the formula

$$(y^k)_j = x_j^k (\varphi^*)'(-\frac{u_j^*}{x_j^k}) \quad j = 1, \ldots, n.$$

To illustrate our two algorithms, we introduce three academic numerical tests of small size. Our purpose is to compare the behavior of the two algorithms. The data are the following ones: for the first two examples, the matrix $Q$ and the vector $q$ are

$$Q = \begin{bmatrix} 1.6 & 1 & 0 & 0 & 0 \\ 1 & 1.6 & 0 & 0 & 0 \\ 0 & 0 & 1.5 & 1 & 0 \\ 0 & 0 & 1 & 1.5 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \quad \text{and} \quad q = \begin{bmatrix} -1 \\ -2 \\ -1 \\ 2 \\ -1 \end{bmatrix}$$

while for the third example, they are

$$Q = \begin{bmatrix} 2.3550 & 1.6364 & 1.8430 & 2.1540 & 0.7586 \\ 1.6364 & 1.6620 & 1.5323 & 1.4876 & 0.2901 \\ 1.8430 & 1.5323 & 2.4317 & 2.2961 & 1.0964 \\ 2.1540 & 1.4876 & 2.2961 & 2.8473 & 1.2273 \\ 0.7586 & 0.2901 & 1.0964 & 1.2273 & 0.8085 \end{bmatrix} \quad \text{and} \quad q = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The matrix $P$ is chosen successively equal to

$$\begin{bmatrix} 3.1 & 2.0 & 0 & 0 & 0 \\ 2.0 & 3.6 & 0 & 0 & 0 \\ 0 & 0 & 3.5 & 2 & 0 \\ 0 & 0 & 2 & 3.3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 3.1 & 2.0 & 0 & 0 & 0 \\ 2.0 & 3.6 & 0 & 0 & 0 \\ 0 & 0 & 3.5 & 2 & 0 \\ 0 & 0 & 2 & 3.3 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \quad \text{and} \quad 10I,$$

where $I$ denotes the identity matrix. The parameters are fixed to $\nu = 7$, $\mu = 1$, $c_k = 1/c_1$ for Algorithm IPE, and $\nu = 2$, $\mu = 1$, $\theta = 0.99$, $\alpha = 0.49$, $\tau = 0.999$ for Algorithm IPLE. For this algorithm, $c_k$ is equal to 0.7 for the first two examples and to 0.1 for the third one. Finally the starting point is $x_0 = (1, 3, 1, 1, 2)$ for all the tests. The results are reported in the table below:

| | Example 1 | | Example 2 | | Example 3 | |
|---|---|---|---|---|---|---|
| Algorithm | IPE | IPLE | IPE | IPLE | IPE | IPLE |
| it | 19 | 1305 | 20 | 1342 | 40 | 228 |
| cpu (second) | 1.078 | 26.89 | 1.296 | 27.64 | 10.875 | 13.25 |
| optimality | -0.00000 | -0.00257 | -0.00000 | -0.00237 | -0.00006 | -0.00152 |

where ‘it’ and ‘cpu’ stand for the number of iterations and the cpu time, respectively. The two algorithms give the same solution for each example. Two constraints are active at the solution for Examples 1 and 2. Three constraints are active for the third example. Furthermore for checking the quality of the solution $x$ obtained by each algorithm, we solve the minimization problem $\min_{y \geq 0} f(x, y)$ whose optimal value must be equal to zero when $x$ is the exact solution of (EP). This optimal value is denoted ‘optimality’ in the table.

¿From the preliminary numerical results reported in the table, the first algorithm seems to be the most efficient. For each example, the total number of iterations is much smaller for this algorithm than for the second one as well as the cpu time. Furthermore it is also the most robust in the sense that the quality of the solution is the best. But this could be due to the fact that for the second algorithm, an unconstrained minimization problem is replaced at each iteration by a gradient step which usually slows down the convergence.

**Acknowledgements.** We are very grateful to the referees for their really helpful and constructive comments; in particular to one of the referees for bringing to our attention references [1, 17, 18, 20, 21].
The first author wishes to thank the University of Namur for financial support in the form of a FUNDP Research Scholarship during the period September 2004 to August 2008.

## References


[1] Antipin, A.S.: The convergence of proximal methods to fixed points of extremal mappings and estimates of their rates of convergence. Comput. Maths Math. Phys. **35**, 539–551 (1995)

[2] Aubin, J.P. and Ekeland, I.: Applied Nonlinear Analysis. Wiley, New York (1984)

[3] Auslender, A., Teboulle, M., and Ben-Tiba, S.: A logarithmic-quadratic proximal method for variational inequalities. Comput. Optim. Appl. **12**, 31–40 (1999)

[4] Auslender, A., Teboulle, M., and Ben-Tiba, S.: Interior proximal and multiplier methods based on second order homogeous kernels. Math. Oper. Res. **24**, 645–668 (1999)

[5] Auslender, A. and Teboulle, M.: Lagrangian duality and related multiplier methods for variational inequality problems. SIAM J. Optim. **10**, 4, 1097–1115 (2000)

[6] Auslender, A. and Teboulle, M.: Entropic proximal decomposition methods for convex programs and variational inequalities. Math. Program. **91**, 33–47 (2001)

[7] Auslender, A. and Teboulle, M.: The log-quadratic proximal methodology in convex optimization algorithms and variational inequalities. In: P. Daniele, F. Giannessi and A. Maugeri, (eds.), Equilibrium Problems and Variational Models. Nonconvex Optimization and its Applications. Kluwer Academic Publishers, Dordrecht, The Netherlands. **68**, 19–52 (2003)

[8] Auslender, A. and Teboulle, M.: Interior gradient and epsilon-subgradient descent methods for constrained convex minimization. Math. Oper. Res. **29**, 1–26 (2004)

[9] Auslender, A., Teboulle, M.: Interior projection-like methods for monotone variational inequalities. Math. Program. **104**, 39–68 (2005)

[10] Auslender, A. and Teboulle, M.: Interior gradient and proximal methods for convex and conic optimization. SIAM J. Optim. **16**, 3, 697–725 (2006)

[11] Bnouhachem, A.: An LQP method for pseudomonotone variational inequalities. J. Global Optim. **36**, 351–363 (2006)

[12] Bnouhachem, A.: A new inexactness criterion for approximate logarithmic-quadratic proximal methods. Numer. Math. J. Chinese Univ. (English Ser.) **15**, 1, 74–81 (2006)

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

[14] Burachik, R. S. and Svaiter B.F.: A relative error tolerance for a family of generalized proximal point methods. Math. Method Oper. Res. **26**, 4, 816–831 (2001)

[15] Cohen, G.: Auxiliary problem principle extended to variational inequalities. J. Optim. Theory Appl. **59**, 325–333 (1988)

[16] Facchinei, F. and Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer-Verlag, New-York (2003)

[17] Flåm, S.D. and Antipin, A.S.: Equilibrium programming using proximal-like algorithms. Math. Program. **78**, 29–41 (1997)

[18] Konnov, I.: Combined relaxation methods for finding equilibrium points and solving related problems. Russ. Math. **37**, 44–51 (1993)

[19] Konnov, I.: Combined Relaxation Methods for Variational Inequalities. Springer (2001)

[20] Konnov, I.: Combined relaxation method for monotone equilibrium problems. J. Optim. Theory Appl. **111**, 327–340 (2001)

[21] Korpelevich, G.M.: Extragradient method for finding saddle points and other problems. Matecon **12**, 747–756 (1976)

[22] Lemaréchal, C., Strodiot, J.J., and Bihain, A.: On a Bundle Method for Nonsmooth Optimization. In: O.L Mangasarian, R.R. Meyer, and S.M Robinson, (eds.), Nonlinear Programming **4**, 245–282. Academic Press, New York (1981)

[23] Mastroeni, G.: On Auxiliary Principle for Equilibrium Problems. In: P. Daniele, F. Giannessi and A. Maugeri, (eds.), Equilibrium Problems and Variational Models **68**, Nonconvex Optimization and its Applications, 289–298. Kluwer Academic Publishers, Dordrecht, The Netherlands (2003)

[24] Nguyen, T.T.V., Nguyen, V.H., and Strodiot, J.J.: A bundle interior proximal method for solving convex minimization problems. J. Convex Anal. **12**, 95–111 (2005)

[25] Nguyen, T.T.V., Strodiot, J.J., and Nguyen, V.H.: A bundle method for solving equilibrium problems. Math. Program. (DOI 10.1007/s10107-007-0112-x)

[26] Noor, M.A. and Bnouhachem, A.: Modified proximal-point method for nonlinear complementarity problems. J. Comput. Appl. Math. **197**, 395–405 (2006)

[27] Muu, L.D., Nguyen, V.H., and Quy, N.V.: On Nash-Cournot oligopolistic market equilibrium models with concave cost functions. J. Global Optim. (DOI 10.1007/s10898-007-9243-0)

[28] Rockafellar, R.T., Convex Analysis. Princeton, New Jersey (1970)

[29] Solodov, M. and Svaiter, B.: Error bounds for proximal point subproblems and associated inexact proximal point algorithms. Math. Program. **88**, 371–389 (2000)

[30] Tran, D.Q, Muu, L.D., and Nguyen, V.H.: Extragradient algorithms extended to equilibrium problems. Optimization (DOI: 10.1080/02331930601122876)

[31] Xu, Y., He, B., and Yuan, X.: A hybrid inexact logarithmic-quadratic proximal method for nonlinear complementarity problems. J. Math. Anal. Appl. **322**, 276–287 (2006)

[32] Yuan, X.: The prediction-correction approach to nonlinear complementarity problems. Eur. J. Oper. Res. **176**, 1357–1370 (2007)

[33] Zhu, D. and Marcotte, P.: Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities. SIAM J. Optim. **6**, 714–726 (1996)