

## An inexact proximal algorithm for variational inequalities

O. Sarmiento, E. A. Papa Quiroz and P. R. Oliveira

Programa de Ingeniería de Sistemas y Computación - COPPE, Universidad Federal de Rio de Janeiro, 68511  
CEP: 21941-972, Rio de Janeiro, Brasil.

DOI: <https://doi.org/10.33017/RevECIPeru2015.0018/>

### Abstract

This paper presents a new inexact proximal method for solving monotone variational inequality problems with a given separable structure. The resulting method combines the recent proximal distances theory introduced by Auslender and Teboulle (2006) with a decomposition method given by Chen and Teboulle that was proposed to solve convex optimization problems. This method extends and generalizes proximal methods using Bregman, Phi-divergences and Quadratic logarithmic distances. Taking mild assumptions we prove that the primal-dual sequences produced by algorithm is well-defined and converge to optimal solution of the variational inequality problem. Furthermore, we show some numerical experiments, for the particular case to solve convex optimization problem, showing that the algorithm is perfectly implementable.

**Keywords:** *Inexact proximal method, variational inequality, separable structure, proximal distances.*

### Resumen

En este artículo presentamos un nuevo método proximal inexacto para resolver problemas de desigualdad variacional monótono con una estructura separable. El método resultante combina la reciente teoría de distancias proximales introducidas por Auslender y Teboulle (2006) con un método de descomposición proximal dado por Chen y Teboulle que fue propuesto para resolver problemas de optimización convexa.

Este método extiende y generaliza métodos proximales usando distancias de Bregman, Phi-divergencias y logaritmo cuadrático, Asumiendo hipótesis adecuadas probamos que la sucesión primal-dual generada por el algoritmo está bien definido y converge a la solución óptima de un problema de desigualdad variacional. Además presentamos algunos resultados computacionales para el caso particular de resolver problemas de optimización convexa, mostrando así que el algoritmo es perfectamente implementable.

**Descriptores:** *Método proximal inexacto, desigualdad variacional, estructura separable, distancias proximales.*

### 1. Introduction

Let  $T: \mathbb{R}^n \times \mathbb{R}^p \rightrightarrows \mathbb{R}^n \times \mathbb{R}^p$  be a maximal monotone operator and let

$$\Omega := \{(x, z) \in \mathcal{C} \times \mathcal{K} : Ax + Bz = b\} \quad (1)$$

where  $\mathcal{C} \subset \mathbb{R}^n$  and  $\mathcal{K} \subset \mathbb{R}^n$  are nonempty open convex sets,  $\mathcal{C}$  and  $\mathcal{K}$  denote the closure (in the euclidean topology) of  $\mathcal{C}$  and  $\mathcal{K}$  respectively,  $A \in \mathbb{R}^{m \times n}$ ,  $B \in \mathbb{R}^{m \times p}$  and  $b \in \mathbb{R}^m$ . This paper considers the variational inequality problem  $VI(\Omega, T)$  with separate structure:

Find a pair  $(x^*, z^*) \in \Omega$  and  $g^* := (g_1^*, g_2^*) \in T(x^*, z^*)$

such that

$$\langle x - x^*, g_1^* \rangle + \langle z - z^*, g_2^* \rangle \geq 0, \quad \forall (x, z) \in \Omega, \quad (2)$$

where  $\langle \cdot, \cdot \rangle$  denotes the inner product in the appropriate Euclidean space.

Then, it can be easily verified that  $(x^*, z^*)$  solves  $VI(\Omega, T)$  if and only if there exists  $y^* \in \mathbb{R}^m$ , playing the role of a dual multiplier for the constraint (1) such that  $(x^*, z^*, y^*)$  solves the following primal-dual formulation of:

Find  $(x^*, z^*, y^*) \in \bar{\mathcal{C}} \times \bar{\mathcal{K}} \times \mathbb{R}^m$  and  $g^* := (g_1^*, g_2^*) \in T(x^*, z^*)$  such that for all  $(x, z) \in \mathcal{C} \times \mathcal{K}$  we have

$$\langle x - x^*, g_1^* + A^T y^* \rangle + \langle x - x^*, g_2^* + B^T y^* \rangle \geq 0, \quad (3)$$

$$Ax^* + Bz^* = b. \quad (4)$$

Various decomposition methods that exploit the special structure of the problem have been proposed. Some examples of such methods are: alternating directions method of multipliers [4,13,9,10] partial inverse of Spingarn method [8,14,15,19,18]. See also the recent works [11,12] and references therein.

For solving the primal-dual formulation of  $VI(T, \Omega)$ , we propose a inexact proximal decomposition algorithm using proximal distances, which combine the recent proximal distances theory introduced by Auslender and Teboulle in [3] with the Entropic Proximal Decomposition Method proposed in [2]. This scheme is in fact an extension of Chen and Teboulle method's [7] (which was developed for solving convex programs with a particular separable structure) and the Entropic Proximal Decomposition Method.

The extension is in two directions:

Firstly we consider the more general framework of variational inequalities with convex constraints and secondly we use here the recent proximal distance theory introduced by Auslender and Teboulle [3] in place of the usual quadratic proximal theory (for the case of Chen and Teboulle method's) and Logarithmic Quadratic proximal theory (for the case of the Entropic Proximal Decomposition Method).

The rest of the paper is organized as follows. In Section 2 we recall basic notions and properties on

set-valued maps, proximal distances, and induced proximal distances. The inexact proximal decomposition method for variational inequalities is presented in Section 3. Then, the well-definition and convergence of the new method is proved in Section 4 and 5 respectively. In Section 6 we present some numerical experiments. Finally, some conclusions are made in Section 7.

## 2. Basic definitions

Given a subset  $\mathcal{C} \subset \mathbb{R}^n$ , we denote by  $int(\mathcal{C})$  its interior and  $\bar{\mathcal{C}}$  its closure. A point-to-set mapping (or multifunction)  $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  is an operator which associates with each point  $x \in \mathbb{R}^n$  a set  $A(x) \subseteq \mathbb{R}^n$ . The domain and the graph of a point-to-set valued map  $A$  are defined as  $D(A) := \{x \in \mathbb{R}^n : A(x) \neq \emptyset\}$  and  $Gr(A) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : y \in A(x)\}$  respectively. A point-to-set  $A$  is said to be monotone if for all  $x, x' \in dom(A)$ ,  $\langle y' - y, x' - x \rangle \geq 0$ ,  $\forall y' \in A(x'), \forall y \in A(x)$ .  $A$  is said strictly monotone is the inequality above is strict for all  $x, x' \in dom(A)$  with  $x \neq x'$ . A monotone operator is said to be maximal when its graph is not properly contained in the graph of any other monotone operator.

A function  $d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$  is called a proximal distance with respect to an open nonempty convex set  $\mathcal{C} \subset \mathbb{R}^n$  if for each  $y \in \mathcal{C}$ ;  $d(\cdot, y)$  is proper, closed, convex on  $\mathbb{R}^n$  and continuously differentiable on  $\mathcal{C}$ ;  $dom d(\cdot, y) \subset \mathcal{C}$  and  $dom \partial_1 d(\cdot, y) = \mathcal{C}$ , where  $\partial_1 d(\cdot, y)$  denotes the classical subgradient map of the function  $d(\cdot, y)$  with respect to the first variable;  $d(\cdot, y)$  is coercive on  $\mathbb{R}^n$  e  $\$d(y, y) = 0$ .

Given a proximal distance  $d$ , a function  $H : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+ \cup \{+\infty\}$  is called the induced proximal distance to  $d$  if there exists  $\gamma \in (0,1]$  with  $H$  a finitevalued on  $\mathcal{C} \times \mathcal{C}$  and such that for each  $a, b \in \mathcal{C}$ , we have that  $H(a, a) = 0$ ; for each  $c \in \mathcal{C}$ ,  $H(c, \cdot)$  has level bounded sets on  $\mathcal{C}$  and  $\langle c - b, \nabla_1 d(b, a) \rangle \leq H(c, a) - H(c, b) - \gamma H(b, a)$ .

Finally, we write  $(d, H) \in \mathcal{F}_+(\mathcal{C})$  if  $d$  and  $H$  are proximal distance and induced proximal distance respectively and

(a)  $\forall y \in \mathcal{C}$  and  $\forall \{y^k\} \subset \mathcal{C}$  bounded with  $\lim_{k \rightarrow +\infty} H(y, y^k) = 0$ , we have  $\lim_{k \rightarrow +\infty} y^k = y$ .

(b)  $\forall y \in \mathcal{C}$  and  $\forall \{y^k\} \subset \mathcal{C}$  such that  $\lim_{k \rightarrow +\infty} y^k = y$ , we obtain  $\lim_{k \rightarrow +\infty} H(y, y^k) = 0$ .

Some classes of proximal distances are: Bregman distances, proximal distances based on  $\varphi$  divergences, self-proximal distances, and distances based on second order homogeneous proximal distances and were given by Auslender and Teboulle in [3].

### 3. The proximal decomposition algorithm with proximal distances (PDAPD)

Throughout this section we make the following standing assumption for the variational inequality problem  $VI(T, \Omega)$ .

#### Assumptions $\mathcal{A}$

( $\mathcal{A}_1$ ) Problem  $VI(T, \Omega)$  has a solution.

( $\mathcal{A}_2$ )  $dom T \cap (\mathcal{C} \times \mathcal{K}) \neq \emptyset$ .

**Remark 1.** This assumptions were suggested by Auslender and Teboulle (see [2], page 35), to derive well-definition of the Entropic Proximal Decomposition Method and are classical conditions given in proximal algorithms for variational inequality problems, see [2,1,6,5].

Now, we propose an algorithm to solve the problem (3) - (4). In this algorithm we use the class of proximal distances  $(d_0, H_0) \in \mathcal{F}_+(\bar{\mathcal{C}})$ ,  $(d'_0, H'_0) \in \mathcal{F}_+(\mathcal{K})$  and given  $\mu > 0$ ,  $\mu' > 0$  we define the following functions:

$$d(x, y) = d_0(x, y) + \left(\frac{\mu}{2}\right) \|x - y\|^2, \tag{5}$$

$$H(x, y) = H_0(x, y) + \left(\frac{\mu}{2}\right) \|x - y\|^2, \tag{6}$$

$$d'(x, y) = d'_0(x, y) + \left(\frac{\mu'}{2}\right) \|x - y\|^2, \tag{7}$$

$$H'(x, y) = H'_0(x, y) + \left(\frac{\mu'}{2}\right) \|x - y\|^2, \tag{8}$$

It is easy to check that  $(d, H) \in \mathcal{F}_+(\bar{\mathcal{C}})$  and  $(d', H') \in \mathcal{F}_+(\mathcal{K})$  for the same value of  $\gamma$  and  $\gamma'$  respectively.

The algorithm, which will be called Proximal Decomposition Algorithm with Proximal Distances (PDAPD) is as follows:

#### (PDAPD) Algorithm

**Step 0.** Choose two pairs  $(d_0, H_0) \in \mathcal{F}_+(\bar{\mathcal{C}})$ ,  $(d'_0, H'_0) \in \mathcal{F}_+(\mathcal{K})$  and define  $(d, H)$ ,  $(d', H')$  given by (5)-(6) and (7)-(8) respectively. Take  $\{\lambda_k\}$  a sequence of positive scalars. Start with arbitrary points  $(x^0, z^0, y^0) \in \mathcal{C} \times \mathcal{K} \times \mathbb{R}^m$ ,  $(e_1^{k+1}, e_2^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^p$  and generate the sequences  $\{x^k, z^k, y^k\} \in \mathcal{C} \times \mathcal{K} \times \mathbb{R}^m$  and  $(e_1^{k+1}, e_2^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^p$  as follows:

**Step 1.** For  $k = 0, 1, 2, \dots$ , calculate  $p^{k+1} \in \mathbb{R}^m$  by

$$p^{k+1} = y^k + \lambda_k (Ax^k + Bz^k - b). \tag{9}$$

**Step 2.** Find  $(x^{k+1}, z^{k+1}) \in \mathcal{C} \times \mathcal{K}$ ,  $(e_1^{k+1}, e_2^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^p$  and  $g_{k+1} := (g_{1k+1}, g_{2k+1}) \in T(x_{k+1}, z_{k+1})$  such that

$$g_{1k+1} + Atp_{k+1} + \lambda_{-k1} \nabla_1 d(x_{k+1}, x_k) = e_{1k+1}, \tag{10}$$

$$g_{2k+1} + B^T p_{k+1} + \lambda_k^{-1} \nabla_1 d'(z^{k+1}, z^k) = e_{2k+1}, \tag{11}$$

where  $(e_1^{k+1}, e_2^{k+1})$  is an approximation error which satisfies some conditions given.

**Step 3.** Compute

$$y^{k+1} = y^k + \lambda_k (Ax^{k+1} + Bz^{k+1} - b). \tag{12}$$

**Stopping Criterion:** If  $x^{k+1} = x^k$ ,  $z^{k+1} = z^k$  and  $y^{k+1} = y^k$  then stop. Otherwise to do  $k := k + 1$ , and go to Step 1.

**Remark 2.** Moreover, as we are interested in the asymptotic convergence of the method, we assume in each iteration that  $x^{k+1} \neq x^k$ ,  $z^{k+1} \neq z^k$  and  $y^{k+1} \neq y^k$  for each  $k = 1, 2, \dots$ . Indeed, if

$(x^{k+1}, z^{k+1}, y^{k+1}) = (x^k, z^k, y^k)$  for some  $k$ , then  $\nabla_1 d(x^{k+1}, x^k) = 0$  and  $\nabla_1 d'(z^{k+1}, z^k) = 0$  then from (10)-(11) we have that (3)-(4) hold approximately, that is,  $(x^k, z^k, y^k)$  is an approximate solution of primal-dual formulation of  $VI(T, \Omega)$ .

**4. Well-definition of (PDAPD) algorithm**

Before that we prove our existence result. We show a more general result of independent interest. For this consider a class of functions  $h : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  satisfying the following properties

- i)  $h$  is a closed proper convex function with  $dom h$  open,
- ii)  $h$  is differentiable on  $dom h$ ,
- iii)  $h_{+\infty}(s) = +\infty, \forall s \neq \emptyset$ .

Here  $h_{+\infty}$  denotes the recession function of  $h$ , see [16] for definition. We denote by  $\vartheta$  the class of functions satisfying (i), (ii) and (iii).

**Remark 3.** Note that for fixed  $(x^k, z^k) \in \mathcal{C} \times \mathcal{K}$  the functions  $d(\cdot, x^k)$  and  $d'(\cdot, z^k)$  clearly satisfies properties i) - ii) when  $dom d$  and  $dom d'$  are open. Moreover, It has been proven by the authors in [17] (see proof of Theorem 4.1) that  $d$  and  $d'$  satisfy iii).

Our existence result will be a consequence of the following general result whose proof is similar at the proof given by Auslender et al. for Proposition 2 in [1].

**Lemma 4.1** If  $h \in \vartheta$  then the gradient mapping  $\nabla h$  is onto. Furthermore, let  $T$  be a maximal monotone map such that  $dom T \cap dom h \neq \emptyset$  and let

$$U(x) = \begin{cases} T(x) + \nabla h(x), & \forall x \in dom T \cap dom h \\ \emptyset, & otherwise \end{cases}$$

Then there exists at least a solution  $x$  of the generalized equation:  $0 \in U(x)$  which is unique if in addition  $h$  is supposed to be strictly convex on its domain.

Due to Remark 3 we have to assume the following assumption:

( $\mathcal{A}_3$ ) The proximal distances  $d$  and  $d'$  have open domains.

**Remark 4.** A large class of proximal distances satisfy this assumption. For example:

□ **Separable Bregman distances:**

$d_h(x, y) := h(x) - (h(y) + \langle \nabla h(y), x - y \rangle)$  induced by  $h(x) = \sum_{j=1}^n \theta(x_j)$  where  $\theta$  can be defined by  $\theta(t) = -\log t$  (Burg entropy) or  $\theta(t) = t^{-1}$ . It is clear that, in both cases,  $dom d$  is open.

□ **Proximal distances based on  $\varphi$  - divergences** defined by

$$d_\varphi(x, y) = \sum_{i=1}^n y_i^r \varphi\left(\frac{x_i}{y_i}\right) \text{ with } r = 1, 2.$$

When  $r = 1$ ,  $d_\varphi$  is called  $\varphi$  - divergence proximal distance and  $dom d_\varphi$  is open for  $\varphi(t) = -\log t + t - 1$ .

When  $r = 2$  and  $\varphi(t) = \mu p(t) + \nu/2(t - 1)^2$  with  $\nu > \mu > 0, p(t) = -\log t + t - 1$ ,  $d_\varphi$  is called second order homogeneous proximal distance (also known as log-quadratic proximal distance) and clearly  $dom d_\varphi$  is open.

**Theorem 4.2** Assuming the assumptions ( $\mathcal{A}_1$ )-( $\mathcal{A}_3$ ). For any  $\lambda_k > 0, (x^k, z^k, y^k) \in \mathcal{C} \times \mathcal{K} \times \mathbb{R}^m, \forall k \geq 0$ , there exists a unique point  $(x^{k+1}, z^{k+1}) \in \mathcal{C} \times \mathcal{K}$  satisfying (10)-(11) with  $g^{k+1} \in T(x^{k+1}, z^{k+1})$ .

*Proof.* Let  $P_k(\cdot) := \lambda_k^{-1} \nabla_1 d(\cdot, x^k)$  and  $Q_k(\cdot) := \lambda_k^{-1} \nabla_1 d'(\cdot, z^k)$ . Then  $P_k$  and  $Q_k$  are strictly monotone operators because  $d$  and  $d'$  are strictly convex functions. This implies strict monotonicity of  $T_k := T + (A^T p^{k+1}, B^T p^{k+1}) + (P_k, Q_k)$ . Since  $T + (A^T p^{k+1}, B^T p^{k+1})$  is maximal monotone and  $d, d' \in \vartheta$ , by Lemma 4.1 we have that  $T_k$  has a zero in  $D(T_k)$ , which is unique by strict monotonicity. We call this zero  $(x^{k+1}, z^{k+1})$ . Thus, it is clear that

(10)-(11) hold. Now, we have to show that  $(x^{k+1}, z^{k+1})$  belongs to  $\mathcal{C} \times \mathcal{K}$ . By Definition,  $D(T_k) = D(T) \cap (\mathcal{C} \times \mathcal{K})$ , and since  $(x^{k+1}, z^{k+1}) \in D(T_k)$  we obtain that  $(x^{k+1}, z^{k+1}) \in \mathcal{C} \times \mathcal{K}$ .

**5. Global convergence of (PDAPD) algorithm**

In this section, under appropriate assumptions, we establish the global convergence of the PDAPD.

**Assumptions B.**

(B<sub>1</sub>) Given the parameters  $\mu > 0, \mu' > 0$ , defined in (5) and (7) respectively, the sequence  $\{\lambda_k\}$  satisfies

$$\eta < \lambda_k < \bar{c} - \eta \quad (13)$$

where  $\eta \in (0, \bar{c}/2)$  with  $\bar{c} := \min\left\{\frac{\sqrt{\gamma\mu}}{2\|A\|}, \frac{\sqrt{\gamma'\mu'}}{2\|B\|}\right\}$  and

$\gamma, \gamma'$  are positive constants related to  $H$  and  $H'$  respectively.

(B<sub>2</sub>) Given the sequences  $\{(x^k, z^k)\}$  generated by (PDAPD) algorithm, assume additional condition on the sequences of errors  $\{e_1^k, e_2^k\}$ . We suppose that

$$\sum_{k=0}^{\infty} \langle x^k, e_1^k \rangle < +\infty, \quad \sum_{k=0}^{\infty} \langle x^k, e_2^k \rangle < +\infty, \quad (14)$$

$$\sum_{k=0}^{\infty} \|e_1^k\| < +\infty, \quad \sum_{k=0}^{\infty} \|e_2^k\| < +\infty, \quad (15)$$

**Remark 5.** Assumptions (B<sub>1</sub>)-(B<sub>2</sub>) will be used to ensures the convergence of the method. Observe that the interval  $(\eta, \bar{c} - \eta)$  depends of  $\mu$  and  $\mu'$  which are arbitrary. The condition (B<sub>2</sub>) is a kind of conditions for the errors sequences given in the literature. Condition (14) might appear somewhat unnatural since it involves the iterates  $x^k$  and  $z^k$  which are apriori unknown. However, as it was noticed for Auslender in [2] p. 10., (14) is easily enforceable in practice, and also implied by the more easily verified condition

$$\sum_{k=0}^{\infty} \|e_1^k\| \|x^k\| < +\infty, \quad \sum_{k=0}^{\infty} \|e_2^k\| \|z^k\| < +\infty,$$

which is satisfied in particular when  $e_1^k = 0, e_2^k = 0$  for each  $k$  when  $\mathcal{C}$  and  $\mathcal{K}$  or  $dom T$  is bounded (in addition, with (15)).

The following convergence analysis follows a line of argument similar to that given in [2].

**Theorem 5.1** (Global convergence). Consider the variational inequality problem  $VI(T, \Omega)$  and suppose

Assumptions (A<sub>1</sub>) – (A<sub>3</sub>) and (B<sub>1</sub>) – (B<sub>2</sub>) hold. Let  $(d_0, H_0) \in \mathcal{F}_+(\bar{\mathcal{C}}), (d'_0, H'_0) \in \mathcal{F}_+(\bar{\mathcal{K}})$  be a proximal and induced proximal distances and let  $\{(x^k, z^k, y^k)\}$  be the sequence generated by PDAPD, then the sequence  $\{(x^k, z^k, y^k)\}$  globally converges to  $(x^*, z^*, y^*)$  with  $(x^*, z^*)$  solution of  $VI(T, \Omega)$ .

**6. Numerical experiments**

In this section we show numerical experiments in the case that the operator  $T$  is defined like  $T = (\partial f, \partial g)$ . To this case the variational inequality problem  $VI(T, \Omega)$  becomes a minimization problem. Consider the problem following:

**Example 6.1**

$$\begin{aligned} & \min (x_1 - 1)^2 + (x_2 - 1)^2 + (z_1 - 1)^2 + (z_2 - 1)^2 \\ \{ & \quad s. to: \quad \quad \quad x_1 + 2x_2 + 2z_1 - z_2 = 4 \\ & \quad \quad \quad \quad \quad \quad -2x_1 + x_2 + z_1 + z_2 = 1 \\ & x_i \geq 0, z_i \geq 0, i = 1, 2. \end{aligned}$$

The function  $f + g: \mathbb{R}^4 \rightarrow \mathbb{R}$  defined by

$$\begin{aligned} f(x_1, x_2) + g(z_1, z_2) & \\ & = (x_1 - 1)^2 + (x_2 - 1)^2 + (z_1 - 1)^2 \\ & \quad + (z_2 - 1)^2 \end{aligned}$$

is proper continuous and convex. The optimal point is  $(x^*, z^*) = ((1,1), (1,1))$  with  $(f + g)(x^*, z^*) = 0$  and  $y^* = (0,0)$  is the optimum Lagrange multiplier associated to the equality constraints. In this case,

$$\begin{aligned} A &= \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \\ \mathcal{C} &= \mathbb{R}^2_+, \quad \mathcal{K} = \mathbb{R}^2_+. \end{aligned}$$

In this example, we take  $tol = 10^{-5}, a_k = 0, b_k = 0, \lambda_k = 0.125$  for each  $k$  and an initial point  $(x^0, z^0, y^0) = ((1,2), (3,2), (1,1))$ .

For examples 6.1 and 6.2, we choose three proximal distances given in the literature:

- 1) Kullback-Liebler Bregman distance:

$$d_0(u, v) := \sum_{i=1}^2 u_i \log\left(\frac{u_i}{v_i}\right) + v_i - u_i,$$

2) Proximal  $\varphi$ -divergence distance  $d_0(u, v) := \sum_{i=1}^2 v_i \varphi\left(\frac{u_i}{v_i}\right)$ , with  $\varphi(t) := t - \log t - 1$ , we have,

$$d_0(u, v) := \sum_{i=1}^2 v_i \log\left(\frac{v_i}{u_i}\right) + u_i - v_i.$$

3) Second-order homogeneous proximal distance  $d_0(u, v) := \sum_{i=1}^2 v_i^2 \varphi\left(\frac{u_i}{v_i}\right)$ , with  $\varphi(t) := \sigma(-\log t + t - 1) + \nu/2(t - 1)^2$ ,  $\nu > \sigma > 0$ , we obtain

$$d_0(u, v) := \sum_{i=1}^2 \frac{\nu}{2}(u_i - v_i)^2 + \sigma(v_i^2 \log(v_i/u_i) + u_i v_i - v_i^2).$$

Using the (PDAPD) algorithm with  $\sigma = 0.001$  and  $\nu = 0.01$ , we obtain the results shown in the Fig.(a).

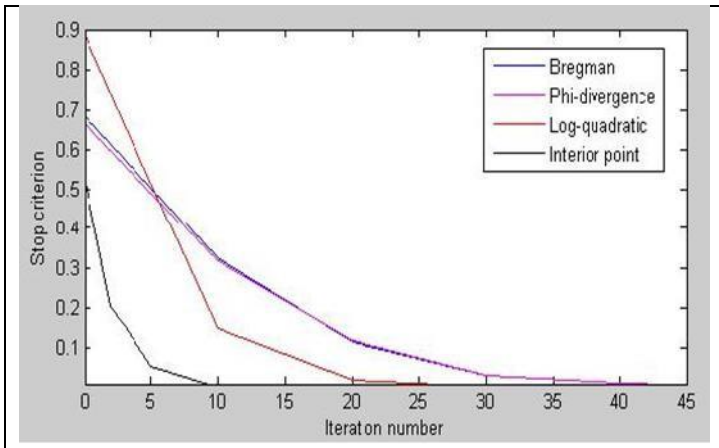


Fig.(a): Numerical results using proximal distances.

Furthermore, for this example we show also results given using the Interior point algorithm with logarithmic barrier.

Moreover, we can solve minimization problem for nondifferentiable functions, for example we consider this problem:

**Example 6.2**

$$\begin{cases} \min |x_1 - 1|^2 + |x_2 - 1|^2 + |z_1 - 1|^2 + |z_2 - 1|^2 \\ \text{s. to:} & x_1 + 2x_2 + 2z_1 + z_2 = 6 \\ & 4x_1 + 3x_2 + 5z_1 = 1 \\ & 0.5 \leq x_i \leq 2, \quad i = 1, 2 \\ & z_i \geq 0.5, \quad i = 1, 2. \end{cases}$$

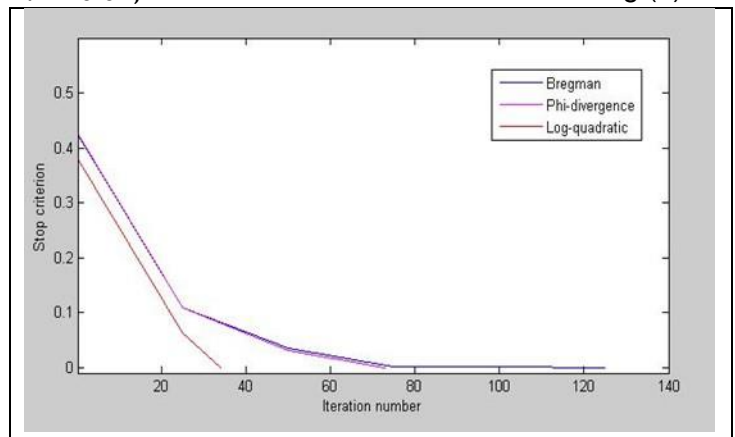
The function  $f + g: \mathbb{R}^4 \rightarrow \mathbb{R}$  defined by

$$\begin{aligned} f(x_1, x_2) + g(z_1, z_2) &= |x_1 - 1|^2 + |x_2 - 1|^2 + |z_1 - 1|^2 \\ &\quad + |z_2 - 1|^2 \end{aligned}$$

is proper continuous and convex. Furthermore, the optimal point is  $(x^*, z^*) = ((1,1), (1,1))$  with optimal

$$(x^0, z^0, y^0) = ((1,2), (3,2), (1,1)).$$

Using the (PDAPD) algorithm with the proximal distances given in 1), 2) and 3) with  $\sigma = 0.001$  and  $\nu = 0.01$ , we obtain the results shown in the Fig.(b).



value  $(f + g)(x^*, z^*) = 0$ .

Consider

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 13 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 5 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

$$\bar{C} = \{x \in \mathbb{R}^2 : 0.5 \leq x_1 \leq 2, 0.5 \leq x_2 \leq 2\}.$$

$$\bar{K} = \{z \in \mathbb{R}^2 : z_1 \geq 0.5, z_2 \geq 0.5\}.$$

In this example, we take  $tol = 10^{-5}$ ,  $a_k = 0$ ,  $b_k = 0$ ,  $\lambda_k = 0.0347$  for each  $k$  and the starting point

*Fig.(b): Numerical results using proximal distances*

In the figures above, Fig.(a) and (b), we can see that the method of interior point is more efficient than our (PADPD) algorithm and in the non-differentiable case the algorithm (PADPD) using Logarithmic Quadratic distance is more efficient than using the class of Bregman and  $\varphi$ -divergences distances.

## 7. Results and discussion

In Section 4, we introduce the (PADPD) Algorithm showing the well-definition of the algorithm, and in Section 5 we show its convergence. In Section 6 we show numerical experiments to solve minimization problems that is a particular case when the operator  $T$  is defined like a subdifferential operator of convex functions.

In the numerical experiments we note that the convergence of the method of Interior Point (in the differentiable case) is faster than our (PADPD) algorithm in the particular case given in example 6.1, and in the non-differentiable case (example 6.2) the convergence of the algorithm (PADPD) using Logarithmic Quadratic distance is faster than itself using the class of Bregman and  $\varphi$ -divergences distances. In future work we hope to show the convergence rate analysis of the algorithm and show computational results comparing with other existing algorithms in the literature.

## 7. Conclusions

The use of proximal distances in the (PDAPD) algorithm generalize the works of Chen and Teboulle method's and the Entropic Proximal Decomposition Method proposed in [2]. Assuming the assumptions  $(\mathcal{A}_1) - (\mathcal{A}_3)$  we prove in Section 4 that the iterations generated by the (PDAPD) algorithm are welldefined. In the literature there proximal algorithms using Bregman and Log-quadratic distances, see [2,7,17] to solve minimization problems, note that these distances are examples of proximal distances, see for example the work of O. Sarmiento et al. in [17]. In this paper we generalize proximal algorithms given in the

literature but now focusing on solving problems of variational inequalities covering a wider field to a minimization problem.

## Acknowledgements

The research of the first author and second author was supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and the PosDoctoral Scholarship CAPES-FAPERJ Edital PAPD-2011, respectively.

## References

- [1] A. Auslender, M. Teboulle, S. Ben-Tiba, *Comp. Optim. Appl.* **12** (1999) 31-40.
- [2] A. Auslender, M. Teboulle, *Math. Program.* **91** (2001) 33-47.
- [3] A. Auslender, M. Teboulle, *SIAM J. Optim.* **16** (2006) 697-725.
- [4] D.P. Bertsekas, J.N. Tsitsiklis, *Parallel and distributed computation*, Numerical Methods, Prentice Hall, Englewood Cliffs, NJ, 1989. [5] R.S. Burachik and A.N. Iusem, *SIAM J. Opt.* **8** (1998) 197-216.
- [6] R.S. Burachik and J. Dutta, *SIAM J. Opt.* **20** (1998) 2653-2678.
- [7] G. Chen, M. Teboulle, *Math. Program.* **64** (1994) 197-216.
- [8] J. Chiflet, P. Mahey, V. Reiner, *Telecommun. Syst.* **3** (1994) 1-10.
- [9] D. Gabay, Chapter ix applications of the method of multipliers to variational inequalities. *Studies in mathematics and its applications* **15** (1983) 299-331.
- [10] Gabay, B. Mercier, *Computers and Mathematics with Applications* **2** (1976) 17-40. [11] D. Han, X. Yuan, W. Zhang, *Mathematics of Computation* **83** (2014) 2263-2291. [12] B. He, M. Tao, X. Yuan, *IMA Journal of Numerical Analysis* **35(1)** (2015) 394-426. [13] P.L. Lions, B. Mercier, *SIAM Journal on Numerical Analysis* **16** (1979) 964-979.
- [14] P. Mahey, Oualibouch, D.T. Pham, *SIAM J. Optim.* **5** (1995) 454-466.
- [15] P. Mahey, A. Ourou, L. LeBlanc, J. Chiflet, *Networks* **31** (1997) 227-238.

- [16] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton NJ, 1970. [17] O. Sarmiento, E.A. Papa Quiroz, P.R.Oliveira, *Optimization* **64** (2015) In Press. [18] J.E.Spingarn, *Applied Mathematics and Optimization* **10** (1983) 247-265.
- [19] J.E. Spingarn, *Math. Program.*, **32** (1985) 199-223.