An Inexact Proximal Algorithm for Pseudomonotone and Quasimonotone Variational Inequalities

E.A. Papa Quiroz, * L. Mallma Ramirez   and   P.R. Oliveira

Federal University of Rio de Janeiro, Brazil
COPPE-PESC
erikpapa@gmail.com

Federal University of Rio de Janeiro, Brazil
COPPE-PESC
lenninmr@gmail.com

Federal University of Rio de Janeiro, Brazil
COPPE-PESC
poliveir@cos.ufrj.br

Abstract

In this paper we introduce an inexact proximal point algorithm using proximal distances for solving variational inequality problems when the mapping is pseudomonotone or quasimonotone. Under some natural assumptions we prove that the sequence generates by the algorithm is convergent for the pseudomonotone case and weakly convergent for the quasimonotone ones. This approach unifies the results obtained by Auslender, Teboulle and Ben-Tiba [1], Brito et al. [3] and extends the convergence properties for the class of $\varphi-$divergence distances and Bregman distances.

Keywords: Variational inequalities, proximal distance, proximal point algorithm, quasimonotone and pseudomonotone mapping.

1   Introduction

In this paper we consider the following Variational Inequality Problem (VIP): find $x^* \in \bar{C}$ and $y^* \in T(x^*)$, such that

$$\langle y^*, x - x^* \rangle \geq 0, \forall x \in \bar{C}, \tag{1.1}$$

where $T : I\!\!R^n \rightrightarrows I\!\!R^n$ is a (point-to-set) mapping, C is a nonempty open convex set in $I\!\!R^n$ and $\bar{C}$ is the euclidean closure of C. The above model covers as particular cases optimization problems, urban traffic equilibrium problems, linear and nonlinear complementarity problems, economic equilibrium problems, among others, see for example Harker and Pang [7] and Facchinei and Pang [6].

---

*Corresponding author: E.A. Papa Quiroz

There are several methods for solving the (VIP), for example, methods based on merit functions, interior point methods, projective methods, proximal point methods, splitting methods, among others, see Vol II of the book of Facchinei and Pang [6].

Given the theoretical importance of the Proximal Point Method (PPM) for very general problem classes and algorithms, in this paper we are interested in solving the problem (1.1), but when the mapping is not necessarily monotone, specifically, when T is pseudomonotone or quasimonotone. Recall that in the monotone case it is possible to obtain global convergence of the proximal method to a solution of (1.1), see for example Theorem 12.3.7 of Facchinei and Pang [6].

The (VIP) when the mapping T is pseudomonotone and quasimonotone, have been recently studied by some researches. Langenberg [12] studied the convergence properties of a inexact version of the (PPM) using Bregman like distances when the mapping T is pseudomonotone, this author has been proved the global convergence of the proposed method. Brito et al.[3] proved a weak convergence property of an exact (PPM) when the mapping is quasimonotone using the class of second order homogeneous distances which includes the logarithmic quadratic distance. Langenberg, [11], under some appropriate assumptions, proved the convergence of an inexact (PPM) using a class of Bregman distances.

On the other hand, Auslender and Teboulle [2] have been developed an unified analysis of the (PPM) using the so called proximal distances, which include Bregman distances, logarithmic quadratic and φ-divergence distances, to solve convex optimization problems.

To obtain convergence properties for a large class of distances, it would be interesting to extend the above approach now to solve (VIP) when the mapping is not necessarily monotone. This is the motivation of the present paper. Specifically, we study the inexact proximal iterations of the form: given $x^{k-1} \in C$ find $x^k \in C$ y $u^k \in T(x^k)$ such that

$$u^k + \lambda_k \nabla_1 d(x^k, x^{k-1}) = e^k,$$

where T is a pseudomonotone or quasimonotone mapping, λ_k is a positive parameter, d is a proximal distances, see Subsection 2.1, and $e^k \in I\!R^n$ is the approximation error given by:

$$\sum_{k=1}^{+\infty} \frac{\left\| e^k \right\|}{\lambda_k} < +\infty \tag{1.2}$$

$$\sum_{k=1}^{+\infty} \frac{\left| \left\langle e^k, x^k \right\rangle \right|}{\lambda_k} < +\infty \tag{1.3}$$

The main contributions of this paper are the following:

- The strong convergence of the proposed algorithm, using proximal distances, when T is a pseudomonotone mapping and the weak convergence when the mapping is quasimonotone. Our results unifies the convergence to the proximal methods introduced by Auslender, Teboulle and Ben-Tiba [1] and Brito et al. [3] and extends the convergence properties for the class of $\varphi-$divergence and Bregman distances.

- We analyze the substitution of a classical condition on the proximal distance, see condition (**Ivii**) in Definition 2.5, by another ones, see condition (**Iviii**) on the same definition, to work with nonlinear constraints in the (VIP). We obtain, in general, the same convergence properties of the proposed algorithm.

- Introducing an extra condition on the proximal distance, see condition (**Iix**), to get rid the condition (1.3) in the proposed algorithm. Furthermore, we give a new algorithm for constrained minimization problems with quasiconvex objective functions with strong convergence results.

This paper is organized as follows: Section 2 gives some basic results used throughout the paper. In the Section 3 we introduce the proposed method. In Section 4 we study the convergence of the sequence generated by the method, analyzing the monotone and quasimonotone cases respectively. In Section 5 we analyze the adaptation and a variant of the method to solve minimization problems with quasiconvex objective functions.

2 Basic Results

Throughout this paper $I\!R^n$ is the Euclidean space endowed with the canonical inner product $\langle\,,\,\rangle$ and the norm of x given by $\|x\| := \langle x,x\rangle^{1/2}$. Let $B \in I\!R^{n\times n}$ be a symmetric and positive definite matrix, we denote $\|x\|_B := \langle Bx,x\rangle^{1/2}$. We also denote the Euclidean ball centered at x with ratio ϵ as $B(x,\epsilon) = \{y \in I\!R^n : \|y - x\| < \epsilon\}$. The interior, closure and boundary of a subset $X \subset I\!R^n$ is denoted by $int(X)$, $\bar{X}$ and $bd(X)$, respectively.

Lemma 2.1 *Let $\{v_k\},\{\gamma_k\}$ and $\{\beta_k\}$ be nonnegative sequences of real numbers satisfying $v_{k+1} \leq (1 + \gamma_k)\,v_k + \beta_k$ such that $\sum_{k=1}^{\infty}\beta_k < \infty$, $\sum_{k=1}^{\infty}\gamma_k < \infty$. Then , the sequence $\{v_k\}$ converges.*

Proof. See Lemma 2, page. 44, of Polyak, [17]. ∎

Definition 2.1 *Let $T : I\!R^n \rightrightarrows I\!R^n$ be a mapping. The domain and the graph of T are defined as*

$$D(T) = \{x \in I\!R^n : T(x) \neq \emptyset\}.$$

$$G(T) = \{(x,v) \in I\!R^n \times I\!R^n : x \in D(T), v \in T(x)\}.$$

Definition 2.2 *A mapping $T : I\!R^n \rightrightarrows I\!R^n$ is closed at $\bar{x}$ if for any sequence $\left\{x^k\right\} \subset I\!R^n$ and any sequence $\left\{v^k\right\} \subset I\!R^n$ such that $(x^k, v^k) \in G(T)$ and $(x^k, v^k) \to (\bar{x}, \bar{v})$, implies that $\bar{v} \in T(\bar{x})$.*

Proposition 2.1 *A mapping $T : I\!R^n \rightrightarrows I\!R^n$ is locally bounded if and only if $T(B)$ is bounded for every bounded set B. This is equivalent to the property that whenever $v^k \in T(x^k)$ and the sequence $\{x^k\} \subset I\!R^n$ is bounded, then the sequence $\{v^k\}$ is bounded.*

Proof. See Proposition 5.15 of Rockafellar and Wets, [18]. ∎

Definition 2.3 *A mapping $T : I\!R^n \rightrightarrows I\!R^n$ is:*

  i. *Strongly monotone if there exists $\alpha > 0$ such that*

$$\langle u - v, x - y\rangle \geq \alpha \|x - y\|^2, \qquad (2.4)$$

  for all $(x, u), (y, v) \in G(T)$.

  ii. *Monotone if*

$$\langle u - v, x - y\rangle \geq 0, \qquad (2.5)$$

  for all $(x, u), (y, v) \in G(T)$.

iii. *Pseudomonotone if*

$$\langle v, x - y \rangle \geq 0 \Rightarrow \langle u, x - y \rangle \geq 0, \tag{2.6}$$

for all $(x, u), (y, v) \in G(T)$.

iv. *Quasimonotone if*

$$\langle v, x - y \rangle > 0 \Rightarrow \langle u, x - y \rangle \geq 0, \tag{2.7}$$

for all $(x, u), (y, v) \in G(T)$.

v. *Weakly monotone if exists* $\rho > 0$ *such that*

$$\langle u - v, x - y \rangle \geq -\rho \|x - y\|^2, \tag{2.8}$$

for all $(x, u), (y, v) \in G(T)$.

vi. *Locally weakly monotone if for each* $x \in int(D(T))$ *there exist* $\epsilon_x > 0$ *and* $\rho_x > 0$ *such that for all* $z, y \in B(x, \epsilon_x) \cap D(T)$ *we have*

$$\langle u - v, z - y \rangle \geq -\rho_x \|z - y\|^2, \tag{2.9}$$

for all $u \in T(z)$ *and all* $v \in T(y)$.

2.1 Proximal Distances

In this subsection we present the definitions of proximal distance and induced proximal distance, introduced by Auslender and Teboulle, [2]. This approach has been used in the works of Villacorta and Oliveira [19], Papa Quiroz and Oliveira [15], Papa Quiroz et al. [16].

Definition 2.4 *A function* $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \cup \{+\infty\}$ *is called a proximal distance with respect to an open nonempty convex set* C *if for each* $y \in C$ *it satisfies the following properties:*

i. $d(\cdot, y)$ *is proper, lower semicontinuous, strictly convex and continuously differentiable on* C;

ii. $\mathrm{dom}\,(d(\cdot, y)) \subset \bar{C}$ *and* $\mathrm{dom}(\partial_1 d(\cdot, y)) = C$, *where* $\partial_1 d(\cdot, y)$ *denotes the classical subgradient map of the function* $d(\cdot, y)$ *with respect to the first variable;*

iii. $d(\cdot, y)$ *is coercive on* $\mathbb{R}^n$ *(i.e.,* $\lim_{\|u\| \to \infty} d(u, y) = +\infty$*);*

iv. $d(y, y) = 0$.

We denote by $D(C)$ *the family of functions satisfying the above definition.*

Property **i.** is needed to preserve convexity of $d(\cdot, y)$, property **ii** will force the iteration of the proximal method to stay in C, and the property **iii** is useful to guarantee the existence of the proximal iterations. For each $y \in C$, let $\nabla_1 d(\cdot, y)$ denote the gradient map of the function $d(\cdot, y)$ with respect to the first variable. Note that by definition $d(\cdot, \cdot) \geq 0$ and from **iv.** the global minimum of $d(\cdot, y)$ is obtained at y, which shows that $\nabla_1 d(y, y) = 0$.

Definition 2.5 *Given* $d \in D(C)$, *a function* $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \cup \{+\infty\}$ *is called the induced proximal distance to* d *if* H *is a finite-valued function on* $C \times C$ *and for each* $a, b \in C$ *we have:*

(Ii) $H(a, a) = 0$.

(Iii) $\langle c - b, \nabla_1 d(b, a) \rangle \le H(c, a) - H(c, b), \quad \forall c \in C.$

Let us denoted by $(d, H) \in \mathcal{F}(C)$ to the proximal distance that satisfies the conditions of Definition 2.5.

We also denote $(d, H) \in \mathcal{F}(\bar{C})$ if there exists H such that:

(Iiii) H is finite valued on $\bar{C} \times C$ satisfying **(Ii)** and **(Iii)**, for each $c \in \bar{C}$.

(Iiv) For each $c \in \bar{C}$, $H(c, \cdot)$ has level bounded sets on C.

Finally, denote $(d, H) \in \mathcal{F}_+(\bar{C})$ if

(Iv) $(d, H) \in \mathcal{F}(\bar{C})$.

(Ivi) $\forall y \in \bar{C}$ y $\forall \{y^k\} \subset C$ bounded with $\lim_{k \to +\infty} H(y, y^k) = 0$, then $\lim_{k \to +\infty} y^k = y$.

(Ivii) $\forall y \in \bar{C}$, y $\forall \{y^k\} \subset C$ such that $\lim_{k \to +\infty} y^k = y$, then $\lim_{k \to +\infty} H(y, y^k) = 0$.

The main result on proximal point method will be when $(d, H) \in \mathcal{F}_+(\bar{C})$. Several examples of proximal distances which satisfy the above definitions, for example Bregman distances, distances based on $\varphi-$ divergences and distances based on second order homogeneous proximal distances, were given by Auslender and Teboulle, [2], Section 3.

Remark 2.1 *The conditions* **(Ivi)** *and* **(Ivii)** *will ensure the global convergence of the sequence generates by the proposed algorithm in this paper. As we will see in Proposition 2.3, the condition* **(Ivii)** *may be substituted by the following:*

(Iviii) $H(., .)$ *is continuous in* $C \times C$ *and if* $\{y^k\} \subset C$ *such that* $\lim_{k \to +\infty} y^k = y \in bd(C)$ *and* $\bar{y} \ne y$ *is another point in* $bd(C)$ *then* $\lim_{k \to +\infty} H(\bar{y}, y^k) = +\infty.$

According Langenberg and Tichatschke, [11], page 643, which is based on the papers of Kaplan and Tichatschke [9] and Kaplan and Tichatschke [10], the above condition for induced Bregman distances holds when nonlinear constraints are active at $y = \lim_{k \to +\infty} y^k$ *while the condition* **(Ivii)** *holds when only affine constraints are active at* y.

Definition 2.6 *Let* $(d, H) \in \mathcal{F}(\bar{C})$. *We say that the sequence* $\{z^l\} \subset C$ *is H-quasi-Fejér convergent to a set* $U \subset \bar{C}$ *if for each* $u \in U$ *there exists a sequence* $\{\epsilon_l\}$, *with* $\epsilon_l \ge 0$ *and* $\sum_{l=1}^{+\infty} \epsilon_l < +\infty$ *such that*

$$H(u, z^l) \le H(u, z^{l-1}) + \epsilon_l.$$

Proposition 2.2 *Let* $(d, H) \in \mathcal{F}_+(\bar{C})$ *and* $\{z^l\} \subset C$ *be a sequence H-quasi-Fejér convergent to* $U \subset \bar{C}$ *then* $\{z^l\}$ *is bounded. If, furthermore, there exists a cluster point* $\bar{z}$ *of* $\{z^l\}$, *belongs to* U, *then the whole sequence* $\{z^l\}$ *converges to* $\bar{z}$.

Proof. Let $u \in U$, from the H-quasi-Fejér convergent assumption we have

$$H(u, z^l) \le H(u, z^0) + \sum_{l=1}^{+\infty} \epsilon_l.$$

Thus, $z^l \in L_H(u, \alpha) := \{y \in C : H(u, y) \le \alpha\}$, where $\alpha = H(u, z^0) + \sum_{l=1}^{+\infty} \epsilon_l$. From Definition 2.5, **(Iiv)**, $L_H(u, \alpha)$ is bounded and therefore $\{z^l\}$ is a bounded sequence.

Let $\bar{z}$ and z^* two cluster points of $\{z^l\}$ where $z^{l_j} \to \bar{z}$ and $z^{l_k} \to z^*$ with $\bar{z} \in U$, then from Definition 2.5 (**Ivii**), $H(\bar{z}, z^{l_j}) \to 0$ and $H(z^*, z^{l_k}) \to 0$. As $H(\bar{z}, z^l)$ is convergent, see Lemma 2.1, and the sequence $H(\bar{z}, z^{l_j})$ converges to zero we obtain that $H(\bar{z}, z^l) \to 0$ and particularly $H(\bar{z}, z^{l_k}) \to 0$. From Definition 2.5, (**Ivi**), we obtain that $z^{l_k} \to \bar{z}$ and due to the uniqueness of the limit we have $z^* = \bar{z}$. Thus, $\{z^l\}$ converges to $\bar{z}$. ∎

The following proposition weaken the above result and it will be important to stabilize the global convergence of the proposed algorithm, when we substitute the condition (**Iviii**) instead of (**Ivii**) in Definition 2.5.

Proposition 2.3 *Let $(d, H) \in \mathcal{F}_+(\bar{C})$ satisfy the condition (**Iviii**) instead of (**Ivii**) and $\{z^l\} \subset C$ be a sequence H-quasi-Fejér convergent to $U \subset \bar{C}$ then $\{z^l\}$ is bounded. If, furthermore, any cluster point of $\{z^l\}$, belongs to U, then the whole sequence $\{z^l\}$ converges.*

Proof. The boundedness of $\{z^l\}$ is immediate. Let $\bar{z}$ and z^* two cluster points of $\{z^l\}$ where $z^{l_j} \to \bar{z}$ and $z^{l_k} \to z^*$, then, from the assumption, $\bar{z}, z^* \in U$ and so both $\{H(\bar{z}, z^l)\}$ and $\{H(z^*, z^l)\}$ converge. We analyze three possibilities.
i. If z^* and $\bar{z}$ belong to $bd(C)$ and suppose that $\bar{z} \neq z^*$, then from assumption (**Iviii**), $H(z^*, z^{l_j}) \to +\infty$, which contradict the convergence of $\{H(z^*, z^l)\}$, then we should have $\bar{z} = z^*$.
ii. If z^* and $\bar{z}$ belong to C, from continuity of $H(.,.)$ in C we have $H(z^*, z^{l_k}) \to 0$. As $\{H(z^*, z^l)\}$ converges then $H(z^*, z^{l_j}) \to 0$. Using the condition (**Ivi**) we have $z^{l_j} \to z^*$, thus $\bar{z} = z^*$.
iii. Without lost of generality we can suppose that $z^* \in C$ and $\bar{z} \in bd(C)$. Then, using the same argument as the case *ii*, we have that $\bar{z} = z^*$, which is a contradiction, so this case is not possible.

Definition 2.7 *Given a symmetric and positive definite matrix $B \in \mathbb{R}^{n \times n}$ and $d \in \mathcal{D}(C)$. We say that d is strongly convex in C with respect to the first variable and with respect to the norm $\|.\|_B$, if for each $y \in C$ there exists $\alpha > 0$ such that*

$$\langle \nabla_1 d(x_1, y) - \nabla_1 d(x_2, y), x_1 - x_2 \rangle \geq \alpha \|x_1 - x_2\|_B^2, \forall x_1, x_2 \in C.$$

Definition 2.8 *Given a symmetric and positive definite matrix $B \in \mathbb{R}^{n \times n}$ and $d \in \mathcal{D}(C)$. We say that d is locally strongly convex in C with respect to the first variable and with respect to the norm $\|.\|_B$, if for each $y \in C$ and each $x \in C$ there exist $\epsilon_x > 0$ and $\alpha_x > 0$ such that*

$$\langle \nabla_1 d(x_1, y) - \nabla_1 d(x_2, y), x_1 - x_2 \rangle \geq \alpha_x \|x_1 - x_2\|_B^2, \forall x_1, x_2 \in B(x, \epsilon_x).$$

Lemma 2.2 *Let $d \in D(C)$, $\text{dom}(T) \cap C \neq \emptyset$ and $B \in \mathbb{R}^{n \times n}$ a symmetric and positive definite matrix. Given an arbitrary point $y \in C$, if T is a locally weakly monotone mapping with constant $\rho > 0$, $d(., y)$ is locally strongly convex with respect to the norm $\|.\|_B$, with constant α and $\{\beta_k\}$ is a sequence of positive numbers satisfying*

$$\beta_k \geq \beta > \frac{\rho}{\alpha \lambda_{min}(B)},$$

where $\lambda_{min}(B)$ denotes the smallest eigenvalue of B, then $F(.) := T + \beta_k \nabla_1 d(., y)$ is locally strongly monotone with constant $\beta \alpha \lambda_{min}(B) - \rho$.

Proof. The proof follows the same steps from Lemma 5.1 of Brito, et al., [3], considering local information, T a point to set mapping and substituting in that lemma $\nabla_1 D_\phi$ by $\nabla_1 d$ and $A^T A$ by B respectively. ∎

3 Inexact Proximal Method

We are interested in solving the (VIP): find $x^* \in \bar{C}$ and $y^* \in T(x^*)$, such that

$$\langle y^*, x - x^* \rangle \geq 0, \forall x \in \bar{C}, \tag{3.10}$$

where $T : I\!\!R^n \Rightarrow I\!\!R^n$ is a mapping not necessarily monotone, C is a nonempty open convex set, $\bar{C}$ is the closure of C in $I\!\!R^n$ and

$$D(T) \cap C \neq \emptyset.$$

Now, we propose an extension of the proximal point method with a proximal distance to solve the problem (3.10).

Inexact Algorithm

Initialization: Let $\{\lambda_k\}$ be a sequence of positive parameters and a starting point:

$$x^0 \in C. \tag{3.11}$$

Main Steps: For $k = 1, 2, \ldots$, and $x^{k-1} \in C$, find $x^k \in C$ and $u^k \in T(x^k)$, such that:

$$u^k + \lambda_k \nabla_1 d(x^k, x^{k-1}) = e^k, \tag{3.12}$$

where d is a proximal distance such that $(d, H) \in \mathcal{F}_+(\bar{C})$ and e^k is an approximation error which satisfies some conditions to be specific later.

Stop Criterion: If $x^k = x^{k-1}$ or $0 \in T(x^k)$, then finish. Otherwise, to do $k - 1 \leftarrow k$ and return to Main Steps.

Throughout this paper, we assume the following assumptions:

(H1) For each $k \in I\!\!N$, there exists $x^k \in C$.

(H2) The solution set $SOL(T, \bar{C})$ of (VIP) is nonempty.

Remark 3.1 *Some sufficient conditions to ensure the assumption* **(H1)** *were presented in Theorems 5.1 and 5.2 of Brito et al. [3] and Theorems 1 and 2 of Langenberg [12].*

Remark 3.2 *Suppose that T is locally weakly monotone where each locally weakly monotone constants ρ is bounded from above by $m > 0$, that is, for each $x \in C$, the constant ρ_x from Definition 2.3, **vi**, satisfies $\rho_x \leq m$. If $d(., y)$ is strongly convex for each $y \in C$ with constant α then, taking*

$$\lambda_k \geq \frac{m}{\alpha \lambda_{\min}(B)},$$

for each $k \in I\!\!N$, the mapping $F := T(.) + \lambda_k \nabla_1 d(., x^{k-1})$ is always locally strongly monotone in C. Furthermore, if T is (globally) weakly monotone with constant $\bar{\rho}$ then taking

$$\lambda_k \geq \frac{\bar{\rho}}{\alpha \lambda_{\min}(B)},$$

we obtain that $T(.) + \lambda_k \nabla_1 d(., x^{k-1})$ is always (globally) strongly monotone in C.
In the above cases, the subproblem (3.12) is well conditioning. Moreover, assuming for simplicity that F is sufficiently smooth and given $x^{k-1} \in C$, to find the point $x^k \in C$ in (3.12) we can apply efficiently, for example, some of the following methods:

- *The so-called damped Newton method:*

$$z^{l+1} = z^l - \alpha_l (JF(z^l))^{-1} F(z^l), \quad l = 0, 1, 2, \ldots$$

  where z^0 is an arbitrary point of C, JF is the Jacobian of F and the step size α_l, $0 < \alpha_l \leq 1$ is chosen so that $F(z)$ decrease monotonicaly, that is, $||F(z^{l+1})|| < ||F(z^l)||$.

- *The Levenberg-Marquardt method:*

$$z^{l+1} = z^l - \left(\alpha_l I + JF(z^l) \right)^{-1} F(z^l), \quad l = 0, 1, 2, \ldots$$

  where $\alpha_l > 0$ is a certain parameter for each l. Observe that there exist different strategies for adjusting α_l, see for example Ortega and Rheinbolt [14].

4   Convergence Results

In this section, under some natural conditions, we prove that the proposed method converges. We divide the analysis in two cases, the pseudomonotone case and the quasimonotone one. Moreover, as we are interested in the asymptotic convergence of the method, we assume in each iteration that $x^k \neq x^{k-1}$ for each $k = 1, 2, \ldots$ In fact, if $x^k = x^{k-1}$, for some k, then $\nabla_1 d(x^k, x^{k-1}) = 0$ and from (3.11)-(3.12) we have that $e^k \in T(x^k)$, that is, x^k is an approximate solution of (VIP).

4.1   Pseudomonotone Case

Proposition 4.1 *Let T be a pseudomonotone mapping, $(d, H) \in \mathcal{F}(\bar{C})$, and suppose that the assumptions (**H1**) and (**H2**) are satisfied. Then, for each $k \in \mathbb{N}$, we have*

$$H(\bar{x}, x^k) \leq H(\bar{x}, x^{k-1}) - \frac{1}{\lambda_k} \left\langle e^k, \bar{x} - x^k \right\rangle, \tag{4.13}$$

for all $\bar{x} \in SOL(T, \bar{C})$.

Proof. Since $\bar{x} \in SOL(T, \bar{C})$, there exists $\bar{u} \in T(\bar{x})$ such that $\langle \bar{u}, z - \bar{x} \rangle \geq 0$ for all $z \in \bar{C}$, in particular $\left\langle \bar{u}, x^k - \bar{x} \right\rangle \geq 0$. Using the pseudomonotonicity of T, we have that for all $u^k \in T(x^k), k \in \mathbb{N}$, it has $\left\langle u^k, x^k - \bar{x} \right\rangle \geq 0$. Now, from (3.12) we have:

$$0 \leq \left\langle u^k, x^k - \bar{x} \right\rangle = \left\langle e^k - \lambda_k \nabla_1 d(x^k, x^{k-1}), x^k - \bar{x} \right\rangle = \left\langle e^k, x^k - \bar{x} \right\rangle + \lambda_k \left\langle \nabla_1 d(x^k, x^{k-1}), \bar{x} - x^k \right\rangle,$$

thus,

$$0 \leq \left\langle e^k, x^k - \bar{x} \right\rangle + \lambda_k \left\langle \nabla_1 d(x^k, x^{k-1}), \bar{x} - x^k \right\rangle.$$

Since $(d, H) \in \mathcal{F}(\bar{C})$ and from Definition 2.5, (**Iii**), it follows that

$$0 \leq \left\langle e^k, x^k - \bar{x} \right\rangle + \lambda_k \left[H(\bar{x}, x^{k-1}) - H(\bar{x}, x^k) \right].$$

Then,

$$H(\bar{x}, x^k) \leq H(\bar{x}, x^{k-1}) - \frac{1}{\lambda_k} \left\langle e^k, \bar{x} - x^k \right\rangle.$$

$\blacksquare$

Proposition 4.2 *Let T be a pseudomonotone mapping, $(d, H) \in \mathcal{F}(\bar{C})$, and suppose that the assumptions ($\mathbf{H1}$) and ($\mathbf{H2}$) are satisfied. If the following additional conditions are satisfied:*

$$\sum_{k=1}^{+\infty} \frac{\left\| e^k \right\|}{\lambda_k} < +\infty \tag{4.14}$$

$$\sum_{k=1}^{+\infty} \frac{\left| \langle e^k, x^k \rangle \right|}{\lambda_k} < +\infty \tag{4.15}$$

then

a). $\{x^k\}$ *is H-quasi-Fejér convergent to the set $SOL(T, \bar{C})$, it is,*

$$H\left(\bar{x}, x^k \right) \leq H\left(\bar{x}, x^{k-1} \right) + \epsilon^k,$$

for each $k \in \mathbb{N}$ and all $\bar{x} \in SOL(T, \bar{C})$, where $\epsilon^k = \frac{1}{\lambda_k} \left(\|e^k\| \|\bar{x}\| + |\langle e^k, x^k \rangle| \right)$ with $\sum_{k=1}^{+\infty} \epsilon^k < +\infty$.

b). $\{H(\bar{x}, x^k)\}$ *converges for all $\bar{x} \in SOL(T, \bar{C})$.*

c). $\{x^k\}$ *is bounded.*

Proof.

a). Using the Cauchy-Schwarz inequality in (4.13) have for all $\bar{x} \in SOL(T, \bar{C})$ that:

$$H(\bar{x}, x^k) \leq H(\bar{x}, x^{k-1}) + \frac{1}{\lambda_k} \left(\|e^k\| \|\bar{x}\| + |\langle e^k, x^k \rangle| \right) \tag{4.16}$$

Let $\epsilon^k = \frac{1}{\lambda_k} \left(\|e^k\| \|\bar{x}\| + |\langle e^k, x^k \rangle| \right)$, then $H\left(\bar{x}, x^k \right) \leq H\left(\bar{x}, x^{k-1} \right) + \epsilon^k$, and from (4.14) and (4.15) we have $\sum_{k=1}^{+\infty} \epsilon^k < \infty$.

b). It is immediately from **a)** and Lemma 2.1.

c). It is also immediately from **a)** and Proposition 2.2. ∎

It is possible to get rid the assumption (4.15) in Proposition 4.2, to obtain that $\{H(x, x^k)\}$ is convergent and $\{x^k\}$ is bounded, for a class of induced proximal distances which includes Bregman distances given by the standard entropy kernel and all strongly convex Bregman functions, see Kaplan and Tichatsche, [8]. We prove this fact in the following proposition.

Proposition 4.3 *Let T be a pseudomonotone mapping, $(d, H) \in \mathcal{F}(\bar{C})$, and suppose that the assumptions ($\mathbf{H1}$) and ($\mathbf{H2}$) are satisfied. If only the condition (4.14) is satisfied and suppose that the induced proximal distance $H(.,.)$ satisfies the following additional property:*

(Iix) *For each $x \in \bar{C}$ there exist $\alpha(x) > 0$ and $c(x) > 0$ such that:*

$$H(x, v) + c(x) \geq \alpha(x) \|x - v\|, \forall v \in C;$$

then

a. *For all* $\bar{x} \in SOL(T, \bar{C})$, *we have*

$$H(\bar{x}, x^k) \leq \left(1 + 2\frac{\left\|e^k\right\|}{\lambda_k \alpha(\bar{x})}\right) H(\bar{x}, x^{k-1}) + 2\frac{\left\|e^k\right\| c(\bar{x})}{\lambda_k \alpha(\bar{x})},$$

for k *sufficiently large and therefore* $\{H(\bar{x}, x^k)\}$ *converges.*

b. $\{x^k\}$ *is bounded.*

Proof. Let $\bar{x} \in SOL(T, \bar{C})$, from (4.13) we have

$$H(\bar{x}, x^k) \leq H(\bar{x}, x^{k-1}) + \frac{1}{\lambda_k}\left\|e^k\right\| \left\|x^k - \bar{x}\right\|. \tag{4.17}$$

Taking $x = \bar{x}$ and $v = x^k$ in (**Iix**) and using in (4.17) we obtain

$$\left(1 - \frac{\left\|e^k\right\|}{\lambda_k \alpha(\bar{x})}\right) H(\bar{x}, x^k) \leq H(\bar{x}, x^{k-1}) + \frac{c(\bar{x})\left\|e^k\right\|}{\alpha(\bar{x})\lambda_k}. \tag{4.18}$$

From (4.14), there exists $k_0 \equiv k_0(\bar{x})$ such that $\frac{\left\|e^k\right\|}{\lambda_k \alpha(\bar{x})} \leq \frac{1}{2}$, for all $k \geq k_0$. Then

$$1 \leq \left(1 - \frac{\left\|e^k\right\|}{\lambda_k \alpha(\bar{x})}\right)^{-1} \leq 1 + 2\frac{\left\|e^k\right\|}{\lambda_k \alpha(\bar{x})} \leq 2, \forall k \geq k_0. \tag{4.19}$$

From (4.18) and (4.19) we have

$$H(\bar{x}, x^k) \leq \left(1 + 2\frac{\left\|e^k\right\|}{\lambda_k \alpha(\bar{x})}\right) H(\bar{x}, x^{k-1}) + 2\frac{\left\|e^k\right\| c(\bar{x})}{\lambda_k \alpha(\bar{x})}, \forall k \geq k_0.$$

Thus, from Lemma 2.1 $\left\{H(\bar{x}, x^k)\right\}$ is convergent and from Definition 2.5, (**Iiv**), $\left\{x^k\right\}$ is bounded. ∎

We now show that the proposed algorithm solves the (VIP) when T is a pseudomonotone mapping. We need the following additional assumption:

(H3) T is a locally bounded mapping and $G(T)$ is closed.

The following result is motivated from Theorem 9 of Langenberg and Tichatschke, [11].

Theorem 4.1 *Suppose that* T *is a pseudomonotone mapping and that the assumptions* (**H1**)-(**H3**) *are satisfied. If* $(d, H) \in \mathcal{F}_+(\bar{C})$, $0 < \lambda_k < \bar{\lambda}$, *for some* $\bar{\lambda} > 0$, *and one of the following condition is satisfied:*

i). The conditions (4.14)-(4.15) are satisfied.

ii). (d, H) *satisfies* (**Iix**) *and only the condition (4.14) is satisfied.*

Then, $\left\{x^k\right\}$ *converges to a point of* $SOL(T, \bar{C})$.

Proof. Since Proposition 4.2 (for the condition i)) and Proposition 4.3 (for the condition ii)) assure that $\{x^k\}$ is bounded, let x^* be a cluster point of $\{x^k\}$ and $\{x^{k_j}\}$ be a subsequence which converges to x^*. Define $L := \{k_1, k_2, ..., k_j, ...\}$, then from above we obtain $\{x^l\}_{l \in L} \to x^*$. From (3.12) we have

$$\left\langle u^l, x - x^l \right\rangle = \left\langle e^l, x - x^l \right\rangle - \lambda_l \left\langle \nabla_1 d(x^l, x^{l-1}), x - x^l \right\rangle, \qquad (4.20)$$

for all $l \in L$ and for each $x \in \bar{C}$. In view of (4.14) and that $\{\lambda_l\}$ is bounded from above we have that $||e^l|| \to 0$. Then, as $\{x^l\}$ is bounded, we obtain $\left\langle e^l, x - x^l \right\rangle \to 0$. Thus, only is necessary to analyze the convergence of the sequence

$$-\lambda_l \left\langle \nabla_1 d(x^l, x^{l-1}), x - x^l \right\rangle.$$

From Definition 2.5 (**Iii**), we have

$$-\lambda_l \left\langle \nabla_1 d(x^l, x^{l-1}), x - x^l \right\rangle \geq \lambda_l \left[H(x, x^l) - H(x, x^{l-1}) \right].$$

Fix $x \in \bar{C}$, we analyze two cases:

If $\{H(x, x^l)\}$ converges, then $\lambda_l \left[H(x, x^l) - H(x, x^{l-1}) \right] \to 0$, since $\{\lambda_l\}$ is bounded, and from (4.20):

$$\liminf_{l \to \infty} \left\langle u^l, x - x^l \right\rangle \geq 0.$$

If $\{H(x, x^l)\}$ is not convergent, then the sequence is not monotonically decreasing and so there are infinite $l \in L$ such that $H(x, x^l) \geq H(x, x^{l-1})$. Let $\{l_j\} \subset L$ such that $H(x, x^{l_j}) \geq H(x, x^{l_j-1})$, for all $j \in \mathbb{N}$, then

$$\liminf_{j \to \infty} \left\langle u^{l_j}, x - x^{l_j} \right\rangle \geq \liminf_{j \to \infty} \lambda_{l_j} \left[H(x, x^{l_j}) - H(x, x^{l_j-1}) \right] \geq 0,$$

and so

$$\liminf_{j \to \infty} \left\langle u^{l_j}, x - x^{l_j} \right\rangle \geq 0,$$

with $u^{l_j} \in T(x^{l_j})$.

Since T is locally bounded and as $\{x^l\}$ is bounded then, from Proposition 2.1, the sequence $\{u^l\}$ is also bounded. Thus, without loss of generality, there exists a subsequence $\{u^{l_j}\}$ which converges to some u^* and since $G(T)$ is closed, $u^* \in T(x^*)$. Consequently we have $\langle u^*, x - x^* \rangle \geq 0$ for all $x \in \bar{C}$ and $u^* \in T(x^*)$, implying that $x^* \in SOL(T, \bar{C})$.

If the condition i) is satisfied, then (4.15) is true and using Proposition 4.2, **a)** and Proposition 2.2, we have $\{x^k\}$ converges to x^*.

Now, if the condition ii) is satisfied, then (d, H) holds the condition (**Iix**). Let $\bar{x}$ be another cluster point of $\{x^k\}$ where $x^{k_l} \to \bar{x}$, then $\bar{x} \in SOL(T, \bar{C})$, and from Definition 2.5 (**Ivii**), $H(\bar{x}, x^{k_l}) \to 0$. As $H(\bar{x}, x^k)$ is convergent, see Proposition 4.3, **a**, and the sequence $H(\bar{x}, x^{k_l})$ converges to zero we obtain that $H(\bar{x}, x^{k_j}) \to 0$. From Definition 2.5, (**Ivi**), we obtain that $x^{k_j} \to \bar{x}$ and due to the uniqueness of the limit we have $x^* = \bar{x}$. Thus, $\{x^k\}$ converges to x^*. ∎

Theorem 4.2 *Suppose that T is a pseudomonotone mapping and that the assumptions (**H1**)-(**H3**) are satisfied. If $(d, H) \in \mathcal{F}_+(\bar{C})$ satisfying the condition (**Iviii**) instead of (**Ivii**), $0 < \lambda_k < \bar{\lambda}$, for some $\bar{\lambda} > 0$, and one of the following condition is satisfied:*

11

i. the conditions (4.14)-(4.15) *are satisfied;*

ii (d, H) *satisfies* (**Iix**) *and the condition* (4.14) *is satisfied;*

then, $\{x^k\}$ *converges to a point of* $SOL(T, \bar{C})$.

Proof.
i. If the conditions (4.14)-(4.15) are satisfied then from Proposition 4.2, **a)**, $\{x^k\}$ is H-quasi-Fejér convergent to $SOL(T, \bar{C})$. As any cluster point of $\{x^k\}$ belongs to $SOL(T, \bar{C})$, see the first part of the proof of Theorem 4.1, then using Proposition 2.3 we obtain the result.
ii. From Proposition 4.3, **b**, $\{x^k\}$ is bounded, mimicking the proof of Theorem 4.1 any cluster point belongs to $SOL(T, \bar{C})$. Let $\bar{x}$ and x^* two cluster points of $\{x^k\}$ with $x^{k_j} \to \bar{x}$ and $x^{k_l} \to x^*$, as $\bar{x}, x^* \in SOL(T, \bar{C})$, from Proposition 4.3, **a**, both $\{H(\bar{x}, x^k)\}$ and $\{H(x^*, x^k)\}$ converge. We analyze the three possibilities.
If x^* and $\bar{x}$ belong to $bd(C)$ and suppose that $\bar{x} \neq x^*$, then from assumption (**Iviii**), $H(x^*, x^{k_j}) \to +\infty$, which contradict the convergence of $\{H(x^*, x^k)\}$, then we should have $\bar{x} = x^*$.
If x^* and $\bar{x}$ belong to C, from continuity of $H(.,,)$ in C we have $H(x^*, x^{k_l}) \to 0$. As $\{H(x^*, x^k)\}$ converges then $H(x^*, x^{k_j}) \to 0$. Using the condition (**Ivi**) we have $x^{k_j} \to x^*$, thus $\bar{x} = x^*$.
Without lost of generality we can suppose that $x^* \in C$ and $\bar{x} \in bd(C)$. Then, using the same argument as the last case we have that $\bar{x} = x^*$, which is a contradiction, so this case is not possible. ∎

4.2 Quasimonotone Case

Assume now that T is a quasimonotone mapping and consider the following subset of $SOL(T, \bar{C})$:
$$SOL^*(T, \bar{C}) = \{x^* \in SOL(T, \bar{C}) : \exists u^* \neq 0, u^* \in T(x^*)\}.$$

In this subsection we consider $SOL(T, \bar{C}) \cap bd(C) \neq \emptyset$.
We will use the following assumption:

(**H2**)' $SOL^*(T, \bar{C}) \neq \emptyset$

The following lemma was proved simultaneously by Brito et al. [3] and Langenberg [13], but for completeness we establish the proof.

Lemma 4.1 *Assume that the assumption* (**H2**)' *is satisfied. If* $x^* \in SOL^*(T, \bar{C})$, *then*

$$\langle u^*, w - x^* \rangle > 0, \forall w \in C,$$

where $u^* \neq 0, u^* \in T(x^*)$.

Proof. Given $x^* \in SOL^*(T, \bar{C})$, then $\langle u^*, x - x^* \rangle \geq 0$ for all $x \in \bar{C}$. Suppose that there exists $\bar{w} \in C$ such that $\langle u^*, \bar{w} - x^* \rangle = 0$. Since $\bar{w} \in C$ and C is an open set, there exists $B(\bar{w}, r) \subset C$, where $B(\bar{w}, r)$ is a ball centered on $\bar{w}$ and radio $r > 0$.
Now as $\|u^*\| > 0$ and $r > 0$, we obtain that there exists $\epsilon > 0$ $(\epsilon < r \|u^*\|)$ such that $\bar{x} = \bar{w} - \frac{\epsilon}{\|u^*\|^2} u^* \in B(\bar{w}, r)$ and therefore $\bar{x} \in C$. Consequently $\langle u^*, \bar{x} - x^* \rangle = -\epsilon < 0$, which is a contradiction. ∎

Proposition 4.4 *If T is a quasimonotone mapping, $(d, H) \in \mathcal{F}(\bar{C})$, and the assumptions* (**H1**) *and* (**H2**)' *are satisfied, then*

$$H(\bar{x}, x^k) \leq H(\bar{x}, x^{k-1}) - \frac{1}{\lambda_k} \left\langle e^k, \bar{x} - x^k \right\rangle, \tag{4.21}$$

for all $\bar{x} \in SOL^*(T, \bar{C})$.

Proof. Given $\bar{x} \in SOL^*(T, \bar{C})$, then there exists $\bar{u} \in T(\bar{x})$ with $\bar{u} \neq 0$ such that $\langle \bar{u}, x - \bar{x} \rangle \geq 0$, for all $x \in \bar{C}$. By assumption (**H1**), we have that $x^k \in C$ and using Lemma 4.1 we have that $\left\langle \bar{u}, x^k - \bar{x} \right\rangle > 0$. Using the property that T is quasimonotone, see Definition 2.3, **iv**, we have that $\left\langle u^k, x^k - \bar{x} \right\rangle \geq 0$, for all $u^k \in T(x^k)$. From here, the steps are the same as the proof of Proposition 4.1. ∎

Proposition 4.5 *Let T be a quasimonotone mapping, $(d, H) \in \mathcal{F}(\bar{C})$, and suppose that the assumptions* (**H1**) *and* (**H2**)' *are satisfied. If the conditions (4.14) and (4.15) are satisfied, then*

   a. $\{x^k\}$ *is H-quasi-Fejér convergent to $SOL^*(T, \bar{C})$.*

   b. $\{H(\bar{x}, x^k)\}$ *convergent for all $\bar{x} \in SOL^*(T, \bar{C})$.*

   c. $\{x^k\}$ *is bounded.*

Proof. Similar to the proof of Proposition 4.2 but using (4.21) instead of (4.13) and $SOL^*(T, \bar{C})$ instead of $SOL(T, \bar{C})$. ∎

Proposition 4.6 *Let T be a quasimomonotone mapping, $(d, H) \in \mathcal{F}(\bar{C})$, and suppose that the assumptions* (**H1**) *and* (**H2**)' *are satisfied. If only the condition (4.14) is satisfied and suppose that the induced proximal distance H with $(d, H) \in \mathcal{F}(\bar{C})$ satisfies* (**Iix**), *then*

   a. *For all $\bar{x} \in SOL^*(T, \bar{C})$, we have*

$$H(\bar{x}, x^k) \leq \left(1 + 2\frac{\left\|e^k\right\|}{\lambda_k \alpha(\bar{x})}\right) H(\bar{x}, x^{k-1}) + 2\frac{\left\|e^k\right\| c(\bar{x})}{\lambda_k \alpha(\bar{x})},$$

   for k sufficiently large and therefore $\{H(\bar{x}, x^k)\}$ converges.

   b. $\{x^k\}$ *is bounded.*

Proof. The proof follows the same steps of the proof of Proposition 4.3 but using (4.21) instead of (4.13) and $SOL^*(T, \bar{C})$ instead of $SOL(T, \bar{C})$.. ∎

Denote $Acc\left(x^k\right)$ as the set of all accumulation points of $\{x^k\}$, that is,

$$Acc\left(x^k\right) = \{z \in \bar{C}: \text{ there exists a subsequence } \{x^{k_l}\} \text{ of } \{x^k\}: x^{k_l} \to z\}.$$

Theorem 4.3 *Let T be a quasimonotone mapping, $(d, H) \in \mathcal{F}_+(\bar{C})$, and suppose that the assumptions* (**H1**), (**H2**)' *and* (**H3**) *are true, $0 < \lambda_k < \bar{\lambda}$, for some $\bar{\lambda} > 0$. If one of the following conditions holds:*

13

i) the conditions (4.14)-(4.15) *are satisfied;*

ii) (d, H) *satisfies* (**Iix**) *and only the condition* (4.14) *holds;*

then,

(a) $\left\{x^k\right\}$ *converges weakly to an element of* $SOL(T, \bar{C})$, *that is,* $Acc\left(x^k\right) \neq \emptyset$ *and every element of* $Acc\left(x^k\right)$ *is a point of* $SOL(T, \bar{C})$.

(b) *If* $Acc\left(x^k\right) \cap SOL^*(T, \bar{C}) \neq \emptyset$ *then* $\left\{x^k\right\}$ *Converges to an element of* $SOL^*(T, \bar{C})$.

Proof. Consider true the first case *i*).
From Proposition 4.5, **c** , $\{x^k\}$ is bounded, and therefore there exists a convergent subsequence and thus $Acc\left(x^k\right) \neq \emptyset$. Take a subsequence $\{x^{k_j}\}$, such that $x^{k_j} \to \bar{x}$. Mimicking the proof of Theorem 4.1 we obtain that $\bar{x} \in SOL(T, \bar{C})$.
Assume that $\bar{x} \in SOL^*(T, \bar{C})$, then from Proposition 4.5, **a**, and Proposition 2.2 we have the result.
Now we consider true the second case *ii*).
From Proposition 4.6, **b** , $\{x^k\}$ is bounded, so $Acc\left(x^k\right) \neq \emptyset$. Take a subsequence $\{x^{k_j}\}$, such that $x^{k_j} \to \bar{x}$. Mimicking the proof of Theorem 4.1 we obtain that $\bar{x} \in SOL(T, \bar{C})$. One more time, mimicking the last five line of the proof of Theorem 4.1 substituting Proposition 4.3, **a**, by Proposition 4.6. ∎

Theorem 4.4 *Let* T *be a quasimonotone mapping,* $(d, H) \in \mathcal{F}_+(\bar{C})$, *satisfying the condition* (**Iviii**) *instead of* (**Ivii**), *and suppose that the assumptions* (**H1**), (**H2**)' *and* (**H3**) *are true,* $0 < \lambda_k < \bar{\lambda}$, *for some* $\bar{\lambda} > 0$. *If one of the following conditions holds:*

i) the conditions (4.14)-(4.15) *are satisfied*

ii) (d, H) *satisfies* (**Iix**) *and only the condition* (4.14) *holds;*

then

(a) $\left\{x^k\right\}$ *converges weakly to an element of* $SOL(T, \bar{C})$, *that is,* $Acc\left(x^k\right) \neq \emptyset$ *and every element of* $Acc\left(x^k\right)$ *is a point of* $SOL(T, \bar{C})$.

(b) *If* $Acc\left(x^k\right) \subset SOL^*(T, \bar{C})$ *then* $\left\{x^k\right\}$ *converges to an element of* $SOL^*(T, \bar{C})$.

Proof. Consider true the first case *i*).
From Proposition 4.5, **c**, $\{x^k\}$ is bounded, and therefore $Acc\left(x^k\right) \neq \emptyset$. Take a subsequence $\{x^{k_j}\}$, such that $x^{k_j} \to \bar{x}$. Mimicking the proof of Theorem 4.1 we obtain that $\bar{x} \in SOL(T, \bar{C})$. From Proposition 4.5, **a**, $\{x^k\}$ is H-quasi-Fejér convergent to $SOL^*(T, \bar{C})$, and if we suppose $Acc\left(x^k\right) \subset SOL^*(T, \bar{C})$ then from Proposition 2.3 we have the aimed result.
Now we consider true the second case *ii*).
From Proposition 4.6, **b**, $\{x^k\}$ is bounded, so $Acc\left(x^k\right) \neq \emptyset$. Take a subsequence $\{x^{k_j}\}$, such that $x^{k_j} \to \bar{x}$. Mimicking the proof of Theorem 4.1 we obtain that $\bar{x} \in SOL(T, \bar{C})$.
If $Acc\left(x^k\right) \subset SOL^*(T, \bar{C})$, let $\bar{x}$ and x^* two cluster points of $\{x^k\}$ with $x^{k_j} \to \bar{x}$ and $x^{k_l} \to x^*$, as $\bar{x}, x^* \in SOL^*(T, \bar{C})$, $\{H(\bar{x}, x^k)\}$ and $\{H(x^*, x^k)\}$ converge. We analyze the three possibilities

If x^* and $\bar{x}$ belong to $bd(C)$ and suppose that $\bar{x} \neq x^*$, then from assumption (**Iviii**), $H(x^*, x^{k_j}) \to +\infty$, which contradict the convergence of $\{H(x^*, x^k)\}$, then we should have $\bar{x} = x^*$.

If x^* and $\bar{x}$ belong to C, from continuity of $H(.,)$ in C we have $H(x^*, x^{k_l}) \to 0$. As $\{H(x^*, x^k)\}$ converges then $H(x^*, x^{k_j}) \to 0$. Using the condition (**Ivi**) we have $x^{k_j} \to x^*$, thus $\bar{x} = x^*$.

Without lost of generality we can suppose that $x^* \in C$ and $\bar{x} \in bd(C)$. Then, using the same argument as the last case we have that $\bar{x} = x^*$, which is a contradiction, so this case is not possible. ∎

5 Quasiconvex Minimization

Now consider the minimization problem

$$\min\left\{ f(x) : x \in \bar{C} \right\} \tag{5.22}$$

where $f : {I\!\!R}^n \to {I\!\!R} \cup \{\pm\infty\}$ is a function and C is a open convex set and $\bar{C}$ denotes the Euclidean closure of C. We assume the following assumptions:

Assumption A. f is a proper lower semicontinuous quasiconvex function.

Assumption B. f is locally lipschitzian and bounded from below.

Assumption C. $\mathrm{dom}(f) \cap \bar{C} \neq \emptyset$.

We use the Clarke subdifferential, which will be denoted by ∂°, see Clarke [4] or [5] for details.

5.1 Algorithm 1 for Minimization

The Inexact Algorithm, introduced in Section 3, but now adapted to solve the problem (5.22) is the following:

Inexact Minimization Algorithm 1 (IMA1)

Initialization: Let $x^0 \in C$.

Main Steps: For $k = 1, 2, \ldots$, and $x^{k-1} \in C$, find $x^k \in C$ and $u^k \in \partial^\circ f(x^k)$, such that:

$$u^k + \lambda_k \nabla_1 d(x^k, x^{k-1}) = e^k,$$

where d is a proximal distance such that $(d, H) \in F_+(\bar{C})$ and e^k is an approximation error satisfying the following conditions:

$$\sum_{k=1}^{+\infty} \frac{\left\| e^k \right\|}{\lambda_k} < +\infty \tag{5.23}$$

$$\sum_{k=1}^{+\infty} \frac{\left| \langle e^k, x^k \rangle \right|}{\lambda_k} < +\infty \tag{5.24}$$

Stop Criterium: If $x^k = x^{k-1}$ or $0 \in \partial^\circ f(x^k)$, then finish. Otherwise, to do $k - 1 \leftarrow k$ and return to Main Steps.

It is easy to prove that if f is a proper lower semicontinuous function and locally Lipschitz function and lower bounded on $\text{dom}(f) \cap \bar{C}$ and $d \in D(C)$, then the assumption (**H1**) is satisfied, see Theorem 4.1 of Papa Quiroz et al. [16]. Moreover, from the properties of the Clarke subdifferential we have that $T = \partial^\circ f$ satisfies the assumption (**H3**). Thus, the particular case of Theorem 4.3 for the minimization problem (5.22) is the following

Corollary 5.1 *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a proper, lower semicontinuous quasiconvex and locally Lipschitz function and lower bounded on $\text{dom}(f) \cap \bar{C}$. If the assumption (**H2**)' is satisfied, $0 < \lambda_k < \bar{\lambda}$, for some $\bar{\lambda} > 0$, the error criterion (5.23) and (5.24) are satisfied and $(d, H) \in \mathcal{F}_+(\bar{C})$, then under the assumption $Acc\left(x^k\right) \cap SOL^*(\partial^\circ f, \bar{C}) \neq \emptyset$, we obtain that $\left\{x^k\right\}$ converges to an element of $SOL^*(\partial^\circ f, \bar{C})$. Furthermore, if (d, H) satisfies the additional condition (**Iix**), then we obtain convergence of $\left\{x^k\right\}$ using only the error criterion (5.23).*

The particular case of Theorem 4.4 is the following result

Corollary 5.2 *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a proper, lower semicontinuous quasiconvex and locally Lipschitz function and lower bounded on $\text{dom}(f) \cap \bar{C}$. If the assumption (**H2**)' is satisfied, $0 < \lambda_k < \bar{\lambda}$, for some $\bar{\lambda} > 0$, the error criterion (5.23) and (5.24) are satisfied and $(d, H) \in \mathcal{F}_+(\bar{C})$ satisfyes the condition (**Iviii**) instead of (**Ivii**), then under the assumption $Acc\left(x^k\right) \subset SOL^*(\partial^\circ f, \bar{C}) \neq \emptyset$, we obtain that $\left\{x^k\right\}$ converges to an element of $SOL^*(\partial^\circ f, \bar{C})$. Furthermore, if (d, H) satisfies the additional condition (**Iix**), then we obtain convergence of $\left\{x^k\right\}$ using only the error criterion (5.23).*

The above results complement, when $Acc\left(x^k\right) \cap SOL^*(T, C) \neq \emptyset$, the convergence results of the proximal point method for quasiconvex functions obtained in Papa et al., [16] where, in that paper, we used simultaneously (5.23)-(5.24) and

$$\left\|x^{k-1} - x^k - e^k\right\| \leq max\left\{\left\|e^k\right\|, \left\|x^k - x^{k-1}\right\|\right\}.$$

5.2 Algorithm 2 for Minimization

In the first bullet point of Conclusion and Future Works section of our recently published paper, Papa Quiroz et al.[16], we wrote that it might be possible to get rid the condition

$$\sum_{k=1}^{+\infty} \frac{\left|\langle e^k, x^k \rangle\right|}{\lambda_k} < +\infty$$

in that algorithm for a class of induced proximal distances following the paper of Kaplan and Tichatschke, [8], but it has not been proved. In this subsection we prove this affirmation. For that, to solve (5.22) consider the following algorithm .

Inexact Minimization Algorithm 2 (IMA2)

Initialization: Let $\{\lambda_k\}$ be a sequence of real positive numbers and an initial point

$$x^0 \in C$$

Main Steps: For $k = 1, 2, ...$, and given $x^{k-1} \in C$, find $x^k \in C$ and $g^k \in \partial^\circ f(x^k)$ such that

$$e^k = g^k + \lambda_k \nabla_1 d(x^k, x^{k-1})$$

where e^k is an approximation error satisfying the following conditions:

$$\left\| x^{k-1} - x^k - e^k \right\| \leq max \left\{ \left\| e^k \right\|, \left\| x^k - x^{k-1} \right\| \right\}$$

$$\sum_{k=1}^{+\infty} \frac{\left\| e^k \right\|}{\lambda_k} < +\infty$$

and d is a proximal distance such that $(d, H) \in \mathcal{F}(\bar{C})$ and satisfies **(Iix)**.
Stop Criterium: If $x^k = x^{k-1}$, or $0 \in \partial^\circ f(x^k)$ then stop. Otherwise to do $k - 1 \leftarrow k$ and return to Main Steps.

Define

$$U_+ := \left\{ x \in \bar{C} : f(x) \leq \inf_{j \in I\!N} f(x^j) \right\}.$$

Suppose that U_+ is nonempty, then under assumptions A, B and C and from Proposition 4.2 of Papa Quiroz et al., [16], we have

$$H(x, x^k) \leq H(x, x^{k-1}) - \frac{1}{\lambda_k} \left\langle e^k, x - x^k \right\rangle, \forall x \in U_+.$$

Proposition 5.1 *Let $(d, H) \in \mathcal{F}(\bar{C})$ satisfying* **(Iix)** *and suppose that the assumptions A, B and C are satisfied. The sequence $\{x^k\}$ generates by (IMA2) satisfies:*

    a. *For all $\bar{x} \in U_+$, we have*

$$H(\bar{x}, x^k) \leq \left(1 + 2 \frac{\left\| e^k \right\|}{\lambda_k \alpha(\bar{x})} \right) H(\bar{x}, x^{k-1}) + 2 \frac{\left\| e^k \right\| c(\bar{x})}{\lambda_k \alpha(\bar{x})}$$

    for k sufficiently large and therefore $\{H(\bar{x}, x^k)\}$ converges.

    b. *$\{x^k\}$ is bounded.*

Proof. The proof is very simmilar to Proposition 4.3 substituting U_+ by $SOL(T, \bar{C})$.

Theorem 5.1 *Let $(d, H) \in \mathcal{F}_+(\bar{C})$ satisfying* **(Iix)** *and suppose that the assumptions A, B and C are satisfied, then the sequence $\{x^k\}$ converges to some point of U_+.*

Proof. From previous Proposition 5.1 $\{x^k\}$ is bounded, then there exists a subsequence $\{x^{k_j}\}$ which converges to $\bar{x}$, that is, $\lim_{j \to +\infty} x^{k_j} = \bar{x}$. As f is lower semicontinuous, $\{f(x^k)\}$ is nonincreasing, see Proposition 4.1 of Papa Quiroz et al., [16], and converges then we have $\bar{x} \in U_+$. Suppose that there exists another sequence $\{x^{k_l}\}$ such that $\lim_{l \to +\infty} x^{k_l} = z \in U_+$. Using property **(Ivii)**, from Definition 2.5, we obtain $\lim_{l \to +\infty} H(z, x^{k_l}) = 0$, and from the convergence of $\{H(z, x^k)\}$, $\lim_{k \to +\infty} H(z, x^k) = 0$. Thus $\lim_{j \to +\infty} H(z, x^{k_j}) = 0$. Using property **(Ivi)**, from Definition 2.5, we obtain that $\lim_{j \to +\infty} x^{k_j} = z$, that is, $\bar{x} = z$. ∎

Remark 5.1 *The above result also is true if the condition* (**Ivii**) *on* (d, H) *is substitute by* (**Iviii**). *In fact, let* $\bar{x}$ *and* x^* *two cluster points of* $\{x^k\}$ *with* $x^{k_j} \to \bar{x}$ *and* $x^{k_l} \to x^*$, *as* $\bar{x}, x^* \in U_+$, $\{H(\bar{x}, x^k)\}$ *and* $\{H(x^*, x^k)\}$ *converge. We analyze the three possibilities*
If x^* *and* $\bar{x}$ *belong to* $bd(C)$ *and suppose that* $\bar{x} \neq x^*$, *then from assumption* (**Iviii**), $H(x^*, x^{k_j}) \to +\infty$, *which contradict the convergence of* $\{H(x^*, x^k)\}$, *then we should have* $\bar{x} = x^*$.
If x^* *and* $\bar{x}$ *belong to* C, *from continuity of* $H(., ,)$ *in* C *we have* $H(x^*, x^{k_l}) \to 0$. *As* $\{H(x^*, x^k)\}$ *converges then* $H(x^*, x^{k_j}) \to 0$. *Using the condition* (**Ivi**) *we have* $x^{k_j} \to x^*$, *thus* $\bar{x} = x^*$.
Without lost of generality we can suppose that $x^* \in C$ *and* $\bar{x} \in bd(C)$. *Then, using the same argument as the last case we have that* $\bar{x} = x^*$, *which is a contradiction, so this case is not possible. From the three possibility we obtain* $\bar{x} = x^*$ *and so* $\{x^k\}$ *converges.*

Finally, we give the global convergence of $\{x^k\}$ to a stationary point of the problem.

Theorem 5.2 *Let* $(d, H) \in \mathcal{F}(\bar{C})$ *satisfying* (**Iix**) *and suppose that the assumptions* A, B *and* C *are satisfied and furthermore* $\{\lambda_k\}$ *is bounded from above, then sequence* $\{x^k\}$ *generates by* *(IMA2) converges to a stationary point* $\bar{x} \in C$ *of* f, *i.e. exists* $\bar{g} \in \partial^\circ f(x)$ *such that* $\forall x \in \bar{C}$ *we have*

$$\langle \bar{g}, x - \bar{x} \rangle \geq 0.$$

The above result also is true if the condition (**Ivii**) *on* (d, H) *is substitute by* (**Iviii**).

Proof. See Theorem 4.3 of Papa Quiroz et al., [16]. ∎

6 Conclusions and Future Work

- This article shows significant progress in the construction of inexact proximal methods for solving variational inequalities with quasimonotone mapping. Specifically we introduce an inexact proximal point algorithm and we prove two type of convergence, global for the pseudomonotone case and weak for the quasimonotone ones. Some relational works are the paper of Langenberg, [13] and Brito et al., [3]. Langenberg in that paper introduced a proximal point method using Bregman distances but with other criterion error and Brito et al. proposed a proximal exact method using the class of second order homogeneous distance. Observe that our work includes as a particular case the class of $\varphi-$divergence distances.

- In this paper we assume, for the variational inequality problem, the existence of the proximal iterations in the interior of the convex set model, we believe that a future work should be to find some sufficient conditions to guarantee the existence of these iterations.

- This work motivates to investigate the following question: Is it feasible to develop a Forward-Backward algorithm with inexact proximal distances to solve the (VIP) with quasimonote mapping?

7 Acknowledgements

References

[1] Auslender, A., Teboulle, M., Ben-Tiba, S.: Interior proximal and multiplier methods based on second order homogeneous functionals. Math Oper. Research. 24, 3, 645-668 (1999)

[2] Auslender, A., Teboulle, M.: Interior gradient and proximal methods for convex and conic optimization. SIAM Journal of Optimization. 16, 3, 697-725 (2006)

[3] Brito, S.A., da Cruz Neto, J.X., Lopes, J. O., Oliveira, P.R.: Interior algorithm for quasiconvex programming problems and variational inequalities with linear constraints. J. Optim Theory Appl. 154, 1, 217-234 (2012)

[4] Clarke, F.H.: Generalized Gradient and Applications. Transaction of the American Mathematical Society (1975)

[5] Clarke, F.H.: Optimization and Nonsmooth Analysis. New York, Wiley (1990)

[6] Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, Vol I and II, Springer Series in Operations Research. Springer-Verlag, New York (2003)

[7] Harker, P.T., Pang, J.S.: Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. Mathematical Programming, 48, 161-220 (1990)

[8] Kaplan, A., Tichatschke, R.: On inexact generalized proximal methods with a weakened error tolerance criterion. Optimization. 53, 3-17 (2004)

[9] Kaplan, A., Tichatschke, R.: Interior proximal method for variational inequalities on non-polyhedral sets. Discuss. Math. Diff. Inclusions Control Optim. 30, 51-59 (2007)

[10] Kaplan, A., Tichatschke, R.: Note on the paper: interior proximal method for variational inequalities on non-polyhedral sets. Discuss. Math. Diff. Inclusions Control Optim. 30, 51-59 (2010)

[11] Langenberg, N., Tichatschke, R.: Interior proximal methods for quasiconvex optimization. J. Glob Optim. 52, 641-661 (2012)

[12] Langenberg, N.: Pseudomnonotone operators and the Bregman proximal point algorithm. J. Glob Optim. 47, 537-555 (2010)

[13] Langenberg, N.: An interior proximal method for a class of quasimonotone variational inequalities. J. Optim Theory Appl. 155, 902-922 (2012)

[14] Ortega J.M., Rheinboldt W.C.: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York-London (1970)

[15] Papa Quiroz, E.A., Oliveira, P.R.: An extension of proximal methods for quasiconvex minimization on the nonnegative orthant. European Journal of Operational Research. 216, 26-32 (2012)

[16] Papa Quiroz, E.A., Mallma Ramirez, L., Oliveira, P.R.: An inexact proximal method for quasiconvex minimization. Accepted for Publication in European Journal of Operational Research (2015)

[17] Polyak, B.T.: Introduction to Optimization, Optimization Software, New York (1987)

[18] Rockafellar, R.T., Wets, R.: Variational Analysis. Grundlehren der Mathematischen, Wissenschaften 317, Springer (1998).

[19] Villacorta, K.D.V., Oliveira, P.R.: An interior proximal method in vector optimization. European Journal of Operational Research. 214, 485-492 (2011)