Fixed Point Theory, 23(2022), No. 1, 311-330
DOI: 10.24193/fpt-ro.2022.1.20
http://www.math.ubbcluj.ro/~nodeacj/sfptcj.html

LINEAR AND SUPERLINEAR CONVERGENCE OF AN INEXACT ALGORITHM WITH PROXIMAL DISTANCES FOR VARIATIONAL INEQUALITY PROBLEMS

E.A. PAPA QUIROZ* AND S. CRUZADO ACUÑA**

*Universidad Nacional Mayor de San Marcos and Universidad Privada del Norte
E-mail: erikpapa@gmail.com

**Universidad Nacional Mayor de San Marcos
E-mal: scruzadoa@gmail.com

Abstract. This paper introduces an inexact proximal point algorithm using proximal distances with linear and superlinear rate of convergence for solving variational inequality problems when the mapping is pseudomonotone or quasimonotone. This algorithm is new even for the monotone case and from the theoretical point of view the error criteria used improves recent works in the literature.
Key Words and Phrases: Variational inequalities, proximal distances, proximal point algorithms, quasimonotone and pseudomonotone mappings.
2020 Mathematics Subject Classification: 90C33, 47H10.

1. Introduction

In this paper we are interested in solving the Variational Inequality Problem (VIP) defined in a closed convex set on the Euclidean space. Observe that the (VIP) covers, as particular cases, minimization problems, urban traffic problems, complementarity problems, economic problems, nonlinear equations, among others. We cite, for example, Harker and Pang [14] and Vol I of Facchinei and Pang [11].

There are several algorithms for solving the (VIP), for example, algorithms based on merit functions, interior point methods, projective methods, proximal point methods, splitting methods, see for example, Vol II of Facchinei and Pang [11].

In this paper we are interested in the Proximal Point Method, (PPM) for now on, for the following reasons:

- The convergence of the sequence generated by the (PPM) has become the theoretical basis to justify the convergence of multiplier methods to solve optimization problems, in particular, it was proved that behind each multiplier method there is a proximal method that generates it, see Rockafellar [27].
- It has been proved that a large class of decomposition methods of convex problems are particular cases of the (PPM) to find a zero of a maximal monotone mapping, specifically the Douglas-Rachford method studied by Lions and

Mercier (which covers a large class of optimization methods) is a particular version of the (PPM), see Eckstein and Bertsekas, [10].

- Several variants of the (PPM) are useful for solving optimization problems, see for example Liu et al. [22], Han [13] and Cai [6] and references therein.

It is well known that, to solve the (VIP) with monotone or pseudomonotone mappings, the (PPM) method and its variants converge to a solution of the problem, see Theorem 12.3.7 of Facchinei and Pang [11] or Langenberg [19]. Furthermore, it was established by some authors, for example Solodov [33] and Tseng [37], that the rate of convergence of the (PPM) to solve the (VIP) is generally linear or sublinear. Some recent works related with proximal point methods, Bregman distances and variational inequality problems are the following [4, 7, 12, 15, 31, 38].

Some researches studied the convergence of the (PPM) when the mapping involved in the (VIP) is quasimonotone: Brito et al. [5], using a class of second order homogeneous distances which includes the logarithmic quadratic distance, proved a weak convergence property (if the intersection of the set of cluster points and a certain solution set is nonempty then the sequence converges) of an exact (PPM). Langenberg [20], under some appropriate assumptions, proved the convergence of an inexact (PPM) using a class of Bregman distances. Papa Quiroz et al. [24] proved the convergence of an inexact (PPM) using proximal distances which includes a class of Bregman distances, $\varphi-$divergence distances and the well known logarithmic quadratic distance.

We should observe that none of the above papers have obtained the linear or superlinear rate of convergence of their algorithms. It is the main motivation of the present work. Motivated from the previous works, [25], [23] and [24], we introduce an inexact proximal point algorithm using proximal distances to solve the (VIP) where the involved mapping is Pseudomonotone or Quasimonotone. We prove, under some natural assumptions, that the sequence generated by the algorithm converges to a solution point of the (VIP) and the rate of convergence is linear or superlinear. These results are very important to establish efficient proximal point methods to solve this class of problems.

This paper is organized as follows: Section 2, gives some basic results used throughout the paper, define proximal distances, rate of convergence based on induced proximal distance and give some definitions on point-to-set mappings. In Section 3, we present the (VIP) and we introduce an inexact proximal point algorithm using proximal distances called **IPP** algorithm. In Section 4, we analyze the convergence of the sequence generated by the **IPP** algorithm for the pseudomonotone and quasimonotone cases. In Section 5, we analyze the rate of convergence of the **IPP** algorithm obtaining superlinear and linear convergence properties. In Section 6, we present our conclusions, discussions and future researches.

2. Preliminaries

In this paper $\mathbb{R}^n$ 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 $B \in \mathbb{R}^{n\times n}$ 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 $B(x,\epsilon)=\{y\in\mathbb{R}^n : \|y-x\|<\epsilon\}$. The interior, closure and boundary of a subset $X\subset\mathbb{R}^n$ is denoted by $int(X)$, $\overline{X}$ and $bd(X)$, respectively.

Lemma 2.1. [26, Lemma 2, pp. 44.] *Let $\{v_k\}$, $\{\gamma_k\}$ and $\{\beta_k\}$ be nonnegative sequences of real numbers satisfying $v_{k+1}\le(1+\gamma_k)\,v_k+\beta_k$ such that*

$$\sum_{k=1}^{\infty}\beta_k<\infty,\quad \sum_{k=1}^{\infty}\gamma_k<\infty.$$

Then, the sequence $\{v_k\}$ converges.

Definition 2.1. 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.** dom $(d(\cdot,y))\subset\overline{C}$ and 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.

Definition 2.2. 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 there exists $\gamma\in(0,1]$ with H a finite-valued function on $C\times C$ and for each $a,b\in C$ we have:

- **(Ii)** $H(a,a)=0$.
- **(Iii)** $\langle c-b,\nabla_1 d(b,a)\rangle\le H(c,a)-H(c,b)-\gamma H(b,a),\quad \forall\, c\in C.$

Let us denoted by $(d,H)\in\mathcal{F}(C)$ to the proximal distance that satisfies the conditions of Definition 2.2.

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

- **(Iiii)** H is finite valued on $\overline{C}\times C$ satisfying (**Ii**) and (**Iii**), for each $c\in\overline{C}$.
- **(Iiv)** For each $c\in\overline{C}$, $H(c,\cdot)$ has level bounded sets on C.

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

- **(Iv)** $(d,H)\in\mathcal{F}(\overline{C})$.
- **(Ivi)** $\forall\, y\in\overline{C}$ y $\forall\,\{y^k\}\subset C$ bounded with $\lim\limits_{k\to+\infty}H(y,y^k)=0$, then $\lim\limits_{k\to+\infty}y^k=y$.
- **(Ivii)** $\forall\, y\in\overline{C}$, $\forall\,\{y^k\}\subset C$ such that $\lim\limits_{k\to+\infty}y^k=y$, then $\lim\limits_{k\to+\infty}H(y,y^k)=0$.

Remark 2.1. Examples of proximal distances which satisfy the above definitions may be seen in Auslender and Teboulle [2], Section 3. The condition (**Iii**) is satisfied by the class of Bregman distances, second order homogeneous distances and for some φ-divergences distances. However, it is well known that all φ-divergences distance satisfy the above inequalities when $\gamma=0$, see for example [36, Lemma 4.1 - (ii)] and [24, Definition 2.5 - (Iii)].

Remark 2.2. In this paper conditions (**Ivi**) and (**Ivii**) will ensure the global convergence of the sequence generated by the proposed algorithm. As we will see later in some results the condition (**Ivii**) will be substituted by the following condition:

(Ivii)' $H(.,.)$ is continuous in $C\times C$ and if $\{y^k\}\subset C$ such that $\lim\limits_{k\to+\infty} y^k = y \in bd(C)$ and $\overline{y}\neq y$ is another point in $bd(C)$ then $\lim_{k\to+\infty} H(\overline{y}, y^k) = +\infty$.

According to Langenberg, [21, page 643] (which is based on the papers of Kaplan and Tichatschke [16], [17]), the above condition for induced Bregman distances holds when nonlinear constraints are active at $y = \lim_{k\to+\infty} y^k$ while condition (**Ivii**) holds only when all the active constraints are affine.

Definition 2.3. 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

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

In the following definition we introduce a concept about the rate of convergence of a sequence related with proximal distance and induced distances.

Definition 2.4. Let $(d, H) \in \mathcal{F}(C)$ and $\{x^k\} \subset \mathbb{R}^n$ be a sequence such that $\{x^k\}$ converges to a point $\overline{x} \in \mathbb{R}^n$. Then, the convergence is said to be:

1. $H-$linear, if there exist a constant $0 < \theta < 1$ and $n_0 \in \mathbb{N}$ such that
$$H(x^k, \overline{x}) \leq \theta H(x^{k-1}, \overline{x}), \quad \forall\, k \geq n_0; \tag{2.1}$$
2. $H-$superlinear, if there exist a sequence $\{\beta_k\}$ converging to zero and $\overline{n} \in \mathbb{N}$ such that
$$H(x^k, \overline{x}) \leq \beta_k H(x^{k-1}, \overline{x}), \quad \forall\, k \geq \overline{n}. \tag{2.2}$$

Remark 2.3. In the particular case when the induced proximal distance H is given by $H(x, y) = \bar{\eta}\|x - y\|^2$, for some $\bar{\eta} > 0$, we obtain the usual rate of convergence definition.

Definition 2.5. Let $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a point-to-set mapping. The domain and the graph of T are defined, respectively, as $D(T) = \{x \in \mathbb{R}^n : T(x) \neq \emptyset\}$ and

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

Definition 2.6. A point-to-set mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is closed at $\bar{x}$ if for any sequence $\left\{x^k\right\} \subset \mathbb{R}^n$ and any sequence $\left\{v^k\right\} \subset \mathbb{R}^n$ such that $(x^k, v^k) \in G(T)$ and $(x^k, v^k) \to (\bar{x}, \bar{v})$, we have that $\bar{v} \in T(\bar{x})$.

Proposition 2.1. *A point-to-set mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ 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 $\{x^k\} \subset \mathbb{R}^n$ is bounded, then the sequence $\{v^k\}$ is bounded.*

Proof. See Proposition 5.15 of Rockafellar and Wets, [28]. □

Definition 2.7. A mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is:

i. Strongly monotone, if there exists $\alpha > 0$ such that
$$\langle u - v, x - y\rangle \geq \alpha \|x - y\|^2, \tag{2.3}$$
for all (x, u), $(y, v) \in G(T)$.

ii. Monotone if

$$\langle u - v, x - y\rangle \geq 0, \tag{2.4}$$

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

iii. Pseudomonotone if

$$\langle v, x - y\rangle \geq 0 \Rightarrow \langle u, x - y\rangle \geq 0, \tag{2.5}$$

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

iv. Quasimonotone if

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

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

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

$$\langle u - v, x - y\rangle \geq -\rho \left\|x - y\right\|^2, \tag{2.7}$$

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

3. Inexact proximal method

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

$$\langle y^*, x - x^*\rangle \geq 0, \forall x \in \overline{C}, \tag{3.8}$$

where $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a point-to-set mapping (not necessarily monotone), C is a nonempty open convex set, $\overline{C}$ is the closure of C in $\mathbb{R}^n$ and $D(T) \cap C \neq \emptyset$. Now, we propose an extension of the (PPM) with a proximal distance, called Inexact Proximal Point (IPP) algorithm, to solve the problem (3.8).

(IPP) Algorithm

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

$$x^0 \in C. \tag{3.9}$$

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

$$u^k + \lambda_k \nabla_1 d(x^k, x^{k-1}) = e^k, \tag{3.10}$$

where d is a proximal distance such that $(d, H) \in \mathcal{F}_+(\overline{C})$ and e^k is an approximation error which satisfies the following hypotheses:

$$\frac{\|e^k\|}{\lambda_k} \leq \eta_k \sqrt{H(x^k, x^{k-1})}, \tag{3.11}$$

$$\sum_{k=1}^{+\infty} \eta_k < +\infty. \tag{3.12}$$

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.

Remark 3.1. In practice, the sequence $\{\eta_k\}$ satisfying (3.12) should be given in the initialization step. For example, we can choose $\eta_k = 1/k^2$, all k.

Remark 3.2. Due that the equation (3.10) will be solved by an iterative method, the error criterion (3.11) is well reasonable. In fact, for simplicity consider that T is a point to point mapping and given a candidate point y to solve the equation $T_k(.) = 0$, where $T_k(.) = T(.) + \lambda_k \nabla_1 d(., x^{k-1})$; we have that $H(y, x^{k-1})$ is explicit and we only should verified the condition

$$\|T_k(y)\| \leq \lambda_k \eta_k \sqrt{H(y, x^{k-1})}.$$

If the above condition is satisfied, then we define $x^k = y$. Otherwise, we would search another candidate point and repeat the same process.

Remark 3.3. If T is weakly monotone with constant $\bar{\rho}$ and d is strongly convex in C with respect to the first variable and with respect to the norm $||.||_B$, then taking

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

where $\lambda_{min}(B)$ denotes the smallest eigenvalue of B, we obtain that

$$T(.) + \lambda_k \nabla_1 d(., x^{k-1})$$

is strongly monotone in C, see Lemma 2.2 of Papa Quiroz et al. [24]. Thus, the subproblems $0 \in T_k(x)$ are well conditioned and we may use an efficient algorithm to obtain (3.10).

Throughout this paper, we assume the following assumptions:

(H1) For each $k \in \mathbb{N}$, there exists $x^k \in C$.
(H2) The solution set $SOL(T, \overline{C})$ of (VIP) is nonempty.

Remark 3.4. Some sufficient conditions to ensure assumption **(H1)** were presented in Theorem 5.1 and Theorem 5.2 of Brito et al. [5] and Theorems 1 and 2 of Langenberg [19].

4. Global convergence

In this section, under some assumptions, we prove that the sequence generated by the proposed **IPP** algorithm converges. We divide the analysis in two cases: the pseudomonotone case and the quasimonotone ones. Moreover, as we are interested in the asymptotic convergence of the algorithm, we assume in each iteration that $x^k \neq x^{k-1}$ for each $k = 1, 2, ...$ In fact, if $x^k = x^{k-1}$, for some k, then $\nabla_1 d(x^k, x^{k-1}) = 0$ and from (3.10) we have that $e^k \in T(x^k)$, that is, x^k is an approximate solution of the (VIP).

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

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

Proof. As $\bar{x} \in SOL(T, \bar{C})$, there exists $\bar{u} \in T(\bar{x})$ such that $\langle \overline{u}, x^k - \overline{x} \rangle \geq 0$. Using the pseudomonotonicity condition of T, we obtain that $\langle u^k, x^k - \overline{x} \rangle \geq 0$, for all $u^k \in T(x^k)$. Then, from (3.10) we have:

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

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 2.2, (**Iii**), it follows that

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

Then, the result is obtained. □

We introduce the following extra condition on the induced proximal distance:

(Iviii) There exists $\theta > 0$ such that:

$$\|x - y\|^2 \leq \theta H(x, y),$$

for all $x \in \overline{C}$ and for all $y \in C$.

Remark 4.1. Some examples of proximal distances which satisfies the above condition are the following:

- **A $\Phi-$divergence proximal distance:** Define

$$d_\varphi(x, y) = \sum_{j=1}^{n} y_j \varphi\left(\frac{x_j}{y_j}\right) + \frac{\sigma}{2}||x - y||^2,$$

  where $\varphi(t) = t - \ln t - 1$. We obtain

$$d_\varphi(x, y) := \sum_{j=1}^{n} x_j - y_j - y_j \ln \frac{x_j}{y_j} + (\sigma/2) \|x - y\|^2.$$

  Defining,

$$H(x, y) = \sum_{j=1}^{n} x_j \ln\left(\frac{x_j}{y_j}\right) + y_j - x_j + \frac{\sigma}{2}||x - y||^2.$$

  It may be verified, see Auslender and Teboulle [1] (Section 2.3, case a2), that

$$\langle c - b, \nabla_1 d_\varphi(b, a) \rangle \leq H(c, a) - H(c, b) - \gamma H(b, a),$$

  $(d_\varphi, H) \in \mathcal{F}_+(\mathbb{R}^n_+)$ and may be verified easily that the condition **(Iviii)** is satisfied with $\theta = \frac{2}{\sigma}$.
- **Bregman distances generated by strongly convex functions:** Let $S \subseteq \mathbb{R}^n$ be a nonempty open convex set, and let $\bar{S}$ be its closure. Let $h : \bar{S} \to \mathbb{R} \cup \{+\infty\}$ be a proper, closed and convex function with $dom\, \nabla h = S$, strictly convex and continuous on $dom\, h$, and continuously differentiable on S. Define

$$H(x, y) := D_h(x, y) = h(x) - [h(y) + \langle \nabla h(y), x - y \rangle], \ \ \forall\, x \in \bar{S}, \forall\, y \in S.$$

The function D_h enjoys a remarkable three point identity (see [8], Lemma 3.1),

$H(c,a) = H(c,b) + H(b,a) + \langle c-b, \nabla_1 H(b,a)\rangle \ \ \forall\, a,b \in S, \forall\, c \in dom\, h.$

The function h is called a Bregman function with zone S, if it satisfies the following conditions:

(B_1) $dom\, h = \bar{S}$;

(B_2) (i) $\forall\, x \in \bar{S}$, $D_h(x,\cdot)$ is level bounded on $int(dom\, h)$;
(ii) $\forall\, y \in S$, $D_h(\cdot, y)$ is level bounded;

(B_3) $\forall\, y \in dom\, h$ and $\forall\, \{y^k\} \subset int(dom\, h)$ with $\lim_{k\to+\infty} y^k = y$, one has $\lim_{k\to+\infty} D_h(y,y^k) = 0$;

(B_4) If $\{y^k\}$ is a bounded sequence in $int(dom\, h)$ and $y \in dom\, h$ such that $\lim_{k\to+\infty} D_h(y,y^k) = 0$, then $\lim_{k\to+\infty} y^k = y$.

Note that (B_4) is a direct consequence of the first three properties, a fact proved by Kiwiel in ([18], Lemma 2.16). For some more properties we recommend the paper of Solodov and Svaiter [34].

Taking a Bregman function h and defining $d(x,y) = H(x,y) := D_h(x,y)$ we obtain that $(d,H) \in \mathcal{F}_+(\overline{C})$. Furthermore, if h is strongly convex then, there exists $\alpha > 0$ such that

$$H(x,y) = h(x) - h(y) - \langle \nabla h(y), x-y\rangle \geq \frac{\alpha}{2}\|x-y\|^2.$$

Defining $\theta = \frac{2}{\alpha}$ we obtain that the condition **(Iviii)** is satisfied.

- **Induced proximal distances by the Second-order homogeneous distances.** Let $\varphi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a closed proper convex function such that $dom\varphi \subset \mathbb{R}_+$ and $dom\partial\varphi = \mathbb{R}_{++}$. We suppose in addition that φ is $C^2(\mathbb{R}_{++})$, strictly convex, and nonnegative on $\mathbb{R}_{++}$ with $\varphi(1) = \varphi'(1) = 0$. We denote by Φ the class of such kernels and by

$$\bar{\Phi} = \left\{\varphi \in \Phi : \ \varphi''(1)\left(1 - \frac{1}{t}\right) \leq \varphi'(t) \leq \varphi''(1)(t-1) \quad \forall t > 0\right\}$$

the subclass of these kernels.

Let $\varphi(t) = \mu p(t) + \frac{\nu}{2}(t-1)^2$ with $\nu \geq \mu p''(1) > 0,\ p \in \bar{\Phi}$ and let the associated proximal distance be defined by

$$d_\varphi(x,y) = \sum_{j=1}^{n} y_j^2 \varphi\left(\frac{x_j}{y_j}\right).$$

The use of φ-divergence proximal distances is particularly suitable for handling polyhedral constraints. Let $C = \{x \in \mathbb{R}^n \ : \ Ax < b\}$, where A is an (m,n) matrix of full rank m $(m \geq n)$. Particularly important cases include $C = \mathbb{R}^n_{++}$ or $C = \{x \in \mathbb{R}^n_{++} \ : \ a_i < x_i < b_i \ \forall i = 1,\ldots,n\}$, with $a_i, b_i \in \mathbb{R}$. In [29], example (c) of the appendix section, was showed that for $H(x,y) = \bar{\eta}\,\|x-y\|^2$ with $\bar{\eta} = 2^{-1}(\nu + \mu p''(1))$, we have $(d_\varphi, H) \in \mathcal{F}_+(\mathbb{R}^n_+)$.

Proposition 4.2. *Let T be a pseudomonotone mapping, $(d,H) \in \mathcal{F}(\overline{C})$, and suppose that assumptions* (**H1**) *and* (**H2**) *are satisfied. If the proximal distance $H(.,.)$ satisfies the additional condition **(Iviii)**, then*

i) *there exists an integer $k_0 \geq 0$ such that for all $k \geq k_0$ and for all $\overline{x} \in SOL(T,\overline{C})$, we have*

$$H(\overline{x},x^k) \leq \left(1+\frac{\theta\eta_k}{1-\theta\eta_k}\right) H(\overline{x},x^{k-1}) + \left(\frac{\eta_k}{4}-\gamma\right) H(x^k,x^{k-1}); \tag{4.14}$$

ii) *$\{H(\overline{x},x^k)\}$ converges for all $\overline{x} \in SOL(T,\overline{C})$;*

iii) *$\{x^k\}$ is bounded;*

iv) *$\lim_{k\to+\infty} H(x^k,x^{k-1}) = 0$.*

Proof. **i)** Let $\overline{x} \in SOL(T,\overline{C})$, then

$$0 \leq \left\| \frac{e^k}{\sqrt{2\lambda_k\eta_k}} + \sqrt{2\lambda_k\eta_k}(\overline{x}-x^k)\right\|^2 = \frac{\|e^k\|^2}{2\lambda_k\eta_k} + 2\lambda_k\eta_k\|\overline{x}-x^k\|^2 + 2\langle e^k,\overline{x}-x^k\rangle,$$

thus,

$$-\frac{1}{\lambda_k}\langle e^k,\overline{x}-x^k\rangle \leq \frac{\|e^k\|^2}{4\lambda_k^2\eta_k} + \eta_k\|\overline{x}-x^k\|^2.$$

Replacing the previous expression in (4.13) we have

$$H(\overline{x},x^k) \leq H(\overline{x},x^{k-1}) - \gamma H(x^k,x^{k-1}) + \frac{\|e^k\|^2}{4\lambda_k^2\eta_k} + \eta_k\|\overline{x}-x^k\|^2.$$

Taking into account the hypothesis (3.11) and the condition **(Iviii)**, we have

$$H(\overline{x},x^k) \leq H(\overline{x},x^{k-1}) - \gamma H(x^k,x^{k-1}) + \frac{\eta_k}{4}H(x^k,x^{k-1}) + \theta\eta_k H(\overline{x},x^k),$$

thus,

$$(1-\theta\eta_k)H(\overline{x},x^k) \leq H(\overline{x},x^{k-1}) + \left(\frac{\eta_k}{4}-\gamma\right) H(x^k,x^{k-1}).$$

As $\eta_k \to 0^+$ (this is true from (3.12)) and $\theta > 0$, then there exists $k_0 \geq 0$ such that $0 < 1-\theta\eta_k \leq 1$ and $\frac{\eta_k}{4}-\gamma < 0$, for all $k \geq k_0$. So applying this fact in the previous expression we have

$$H(\overline{x},x^k) \leq \left(\frac{1}{1-\theta\eta_k}\right) H(\overline{x},x^{k-1}) + \left(\frac{1}{1-\theta\eta_k}\right)\left(\frac{\eta_k}{4}-\gamma\right) H(x^k,x^{k-1}).$$

As $1-\theta\eta_k \leq 1$, then the previous expression becomes

$$H(\overline{x},x^k) \leq \left(1+\frac{\theta\eta_k}{1-\theta\eta_k}\right) H(\overline{x},x^{k-1}) + \left(\frac{\eta_k}{4}-\gamma\right) H(x^k,x^{k-1}).$$

ii) From (4.14), it is clear that

$$H(\overline{x},x^k) \leq \left(1+\frac{\theta\eta_k}{1-\theta\eta_k}\right) H(\overline{x},x^{k-1}), \qquad \forall k \geq k_0. \tag{4.15}$$

As $\eta_k \to 0^+$ and $\theta > 0$, then for all $0 < \epsilon < 1$ there exists $\widetilde{k}_0 \in \mathbb{N}$, such that $\theta\eta_k < \epsilon$, for all $k \geq \widetilde{k}_0$, then $1 - \epsilon < 1 - \theta\eta_k \leq 1$. So

$$\frac{\theta\eta_k}{1-\theta\eta_k} < \frac{\theta\eta_k}{1-\epsilon}, \qquad \forall\, k \geq \widetilde{k}_0.$$

Applying summations, and taking into account (3.12), we have

$$\sum_{k=1}^{+\infty} \frac{\theta\eta_k}{1-\theta\eta_k} < +\infty. \tag{4.16}$$

Finally, taking

$$v_{k+1} = H(\overline{x}, x^k),\ v_k = H(\overline{x}, x^{k-1}),\ \gamma_k = \frac{\theta\eta_k}{1-\theta\eta_k} \text{ and } \beta_k = 0$$

in Lemma 2.1 and considering that $\sum_{k=1}^{+\infty}\gamma_k < +\infty$ we obtain that the sequence $\{H(\overline{x}, x^k\}$ converges.

iii) It is immediate from **(ii)** and Definition 2.2-(Iiv).

iv) From relation (4.14) we have for all $k \geq k_0$:

$$\left(\gamma - \frac{\eta_k}{4}\right) H(x^k, x^{k-1}) \leq H(\overline{x}, x^{k-1}) - H(\overline{x}, x^k) + \left(\frac{\theta\eta_k}{1-\theta\eta_k}\right) H(\overline{x}, x^{k-1}).$$

Take $0 < \alpha < \gamma$, then there exists a positive integer k_0' such that $\frac{\eta_k}{4} < \alpha$, for all $k \geq k_0'$. Taking $k \geq \bar{k} := \max\{k_0, k_0'\}$ we have

$$(\gamma - \alpha)\, H(x^k, x^{k-1}) \leq H(\overline{x}, x^{k-1}) - H(\overline{x}, x^k) + \left(\frac{\theta\eta_k}{1-\theta\eta_k}\right) H(\overline{x}, x^{k-1}).$$

Applying summations we have

$$(\gamma-\alpha)\sum_{k=\bar{k}}^{m} H(x^k, x^{k-1}) \leq H(\overline{x}, x^{\bar{k}-1}) - H(\overline{x}, x^m) + \max_{\bar{k}\leq k\leq m}\{H(\overline{x}, x^{k-1})\}\sum_{k=\bar{k}}^{m}\frac{\theta\eta_k}{1-\theta\eta_k}.$$

Since $\{H(\overline{x}, x^{k-1})\}$ is bounded, taking limit when $m \to +\infty$, and considering (4.16), we have

$$\sum_{k=\bar{k}}^{+\infty} H(x^k, x^{k-1}) < +\infty,$$

and therefore we obtain that $\lim_{k\to+\infty} H(x^k, x^{k-1}) = 0$. □

We will show that the sequence generated by the proposed algorithm converges to a solution of the (VIP) when T is a pseudomonotone mapping. So that, we need the following additional assumption:

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

Theorem 4.1. *Let T be a pseudomonotone mapping, $(d, H) \in \mathcal{F}_+(\overline{C})$, and suppose that both the assumptions* (**H1**), (**H2**) *and* (**H3**), *and* (**Iviii**) *are satisfied and* $0 < \lambda_k < \overline{\lambda}$, *then* $\left\{x^k\right\}$ *converges to a point of* $SOL(T, \overline{C})$.

Proof. From Propositions 4.2, we have that $\{x^k\}$ is bounded, so there exist a subsequence $\{x^{k_j}\} \subseteq \{x^k\}$ and a point x^* such that $x^{k_j} \to x^*$. Define

$$L := \{k_1, k_2, ..., k_j, ...\},$$

then $\{x^l\}_{l\in L} \to x^*$. We will prove that $x^* \in SOL(T, \overline{C})$.
From (3.10) we have that $\forall\, l \in L$ and $\forall\, x \in \overline{C}$:

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

Using Definition 2.2-**(Iii)** in the above equality, we obtain

$$\left\langle u^l, x - x^l \right\rangle \geq \left\langle e^l, x - x^l \right\rangle + \lambda_l \left[H(x, x^l) - H(x, x^{l-1}) + \gamma H(x^l, x^{l-1}) \right]. \tag{4.17}$$

Observe that, from (3.11) and due that $\{\lambda_k\}$ and $\{x^k\}$ are bounded, $\{\eta_k\}$ and $\{H(x^k, x^{k-1})\}$ converge to zero, then

$$\lim_{l\to\infty} \langle e^l, x - x^l \rangle = 0. \tag{4.18}$$

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

a) If $\{H(x, x^l)\}$ converges, then from Proposition 4.2-(**iv**), and the fact that $\{\lambda_l\}$ is bounded, we have $\lambda_l \left[H(x, x^l) - H(x, x^{l-1}) + \gamma H(x^l, x^{l-1}) \right] \to 0$. Applying this result and (4.18) in (4.17) we obtain

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

b) 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 $\{l_j\} \subset L$, for all $j \in \mathbb{N}$, such that $H(x, x^{l_j}) \geq H(x, x^{l_j-1})$, then

$$H(x, x^{l_j}) - H(x, x^{l_j-1}) + \gamma H(x^{l_j}, x^{l_j-1}) \geq \gamma H(x^{l_j}, x^{l_j-1}).$$

Taking into account this last result, Proposition 4.2 -(**iv**) and (4.18), in (4.17) we have:

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

and so

$$\liminf_{j\to\infty} \left\langle u^{l_j}, x - x^{l_j} \right\rangle \geq 0, \text{ with } u^{l_j} \in T(x^{l_j}).$$

Since T is locally bounded (see (**H3**)) and $\{x^l\}$ is bounded (see Propositions 4.2), then from Proposition 2.1, the sequence $\{u^l\} \subset T(x^l)$ is also bounded, so without loss of generality, there exists a subsequence $\{u^{l_j}\} \subseteq \{u^l\}$ and a point u^* such that $u^{l_j} \to u^*$; and since $G(T)$ is closed (see (**H3**)), $u^* \in T(x^*)$. Consequently we have $\langle u^*, x - x^* \rangle \geq 0$ for all $x \in \overline{C}$ and $u^* \in T(x^*)$, so $x^* \in SOL(T, \overline{C})$.
Suppose now that $\overline{x}$ is another cluster point of $\{x^k\} \subset C$ where $x^{k_l} \to \overline{x}$, then $\overline{x} \in SOL(T, \overline{C})$, so by Definition 2.2-**(Ivii)**, $H(\overline{x}, x^{k_l}) \to 0$. As $\{H(\overline{x}, x^k)\}$ is convergent (see Proposition 4.2-(**ii**)), and $H(\overline{x}, x^{k_l}) \to 0$, we obtain that $H(\overline{x}, x^{k_j}) \to 0$; so applying the Definition 2.2-**(Ivi)** we obtain that $x^{k_j} \to \overline{x}$, and due to the uniqueness of the limit we have $x^* = \overline{x}$. Thus $\{x^k\}$ converges to x^*. □

Theorem 4.2. *Suppose that T is a pseudomonotone mapping and that the assumptions* (**H1**), (**H2**) *and* (**H3**)*, and the condition* (**Iviii**) *are satisfied and* $0 < \lambda_k < \overline{\lambda}$. *If* $(d, H) \in \mathcal{F}_+(\overline{C})$ *satisfying the condition* $(\mathbf{Ivii})'$ *instead of* (**Ivii**)*, then* $\left\{x^k\right\}$ *converges to a point of* $SOL(T, \overline{C})$.

Proof. From Proposition 4.2-(**iii**), $\{x^k\}$ is bounded, then mimicking the proof of Theorem 4.1 any cluster point belongs to $SOL(T, \overline{C})$. Let $\overline{x}$ and x^* be two cluster point of $\{x^k\}$ with $x^{k_j} \to \overline{x}$ and $x^{k_l} \to x^*$, as $\overline{x}, x^* \in SOL(T, \overline{C})$, for the Proposition 4.2-(**ii**), both $\{H(\overline{x}, x^k)\}$ and $\{H(x^*, x^k)\}$ converge. We analyze three possibilities.

- If $x^*, \overline{x} \in bd(C)$, with $\overline{x} \neq x^*$, then from assumption $(\mathbf{Ivii})'$, $H(x^*, x^{k_j}) \to +\infty$, which contradict the convergence of $\{H(x^*, x^k)\}$, then we should have $\overline{x} = x^*$.
- If $x^*, \overline{x} \in C$; given that by condition $(\mathbf{Ivii})'$, $H(.,.)$ is continuous in C, then $H(x^*, x^{k_l}) \to H(x^*, x^*) = 0$. 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 $\overline{x} = x^*$.
- Without lost of generality we can suppose that $x^* \in C$ and $\overline{x} \in bd(C)$. Then, using the same argument as the last case we have that $\overline{x} = x^*$, which is a contradiction, so this case is not possible. □

Assume now that T is a quasimonotone mapping and consider the following subset of $SOL(T, \overline{C})$:

$$SOL^*(T, \overline{C}) = \left\{x^* \in SOL(T, \overline{C}) : \exists u^* \neq 0, u^* \in T(x^*)\right\}.$$

In this subsection we consider $SOL(T, \overline{C}) \cap bd(C) \neq \emptyset$ and we will use the following assumption:
(H2)' $SOL^*(T, \overline{C}) \neq \emptyset$.

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

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

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

Proof. See Papa Quiroz et al. [24], Lemma 4.1, page 13. □

Proposition 4.3. *If T is a quasimonotone mapping,* $(d, H) \in \mathcal{F}(\overline{C})$*, and also the assumptions* (**H1**) *and* **(H2)'** *are satisfied, then for all* $\overline{x} \in SOL^*(T, \overline{C})$ *and for each* $k \in \mathbb{N}$*, we have*

$$H(\overline{x}, x^k) \leq H(\overline{x}, x^{k-1}) - \gamma H(x^k, x^{k-1}) - \frac{1}{\lambda_k} \left\langle e^k, \overline{x} - x^k \right\rangle. \tag{4.19}$$

Proof. The steps are the same as the proof of Proposition 4.4 of Papa Quiroz et al. [24]. □

Proposition 4.4. *Let T be a quasimonotone mapping, $(d,H) \in \mathcal{F}(\overline{C})$, and suppose that the assumptions* (**H1**), **(H2)'**, ***(Iviii)*** *are satisfied, then*

i) *exists an integer $k_0 \geq 0$ such that for all $k \geq k_0$ and for all $\overline{x} \in SOL^*(T,\overline{C})$, we have*

$$H(\overline{x} - x^k) \leq \left(1 + \frac{\theta\eta_k}{1-\theta\eta_k}\right) H(\overline{x}, x^{k-1}) + \left(\frac{\eta_k}{4} - \gamma\right) H(x^k, x^{k-1});$$

ii) *$\{H(\overline{x}, x^k)\}$ converges for all $\overline{x} \in SOL^*(T,\overline{C})$;*
iii) *$\{x^k\}$ is a bounded sequence;*
iv) $\lim_{k\to+\infty} H(x^k, x^{k-1}) = 0$.

Proof. Similar to the proof of Proposition 4.2 but using (4.19) instead of (4.13) and $SOL^*(T,\overline{C})$ instead of $SOL(T,\overline{C})$. ☐

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

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

Theorem 4.3. *Let T be a quasimonotone mapping, $(d,H) \in \mathcal{F}_+(\overline{C})$, and suppose that the assumptions* (**H1**), **(H2)'**, (**H3**), *and* (**Iviii**) *are satisfied and $0 < \lambda_k < \overline{\lambda}$, then*

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

Proof.(**a**) From Proposition 4.4-(iii), $\{x^k\}$ is bounded, then there exist a subsequence $\{x^{k_j}\} \subseteq \{x^k\}$ and a point x^* such that $x^{k_j} \to x^*$, so $Acc(x^k) \neq \emptyset$ (since $x^* \in Acc(x^k)$). Finally mimicking the proof of Theorem 4.1, where in the proof we substitute Proposition 4.2-(i) by Proposition 4.4-(i), we obtain that $\overline{x} \in SOL(T,\overline{C})$.
(**b**) One more time, mimicking the last five line of the proof of Theorem 4.1 substituting Proposition 4.2-(i), by Proposition 4.4-(i), we obtain the result.

Theorem 4.4. *Suppose that T be a quasimonotone mapping, and that the assumptions* (**H1**), **(H2)'**, (**H3**), *and* (**Iviii**) *are satisfied and $0 < \lambda_k < \overline{\lambda}$. If $(d,H) \in \mathcal{F}_+(\overline{C})$, satisfying the condition* **(Ivii)'** *instead of* (**Ivii**), *then*

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

Proof.(**a**) From Proposition 4.4-(**iii**), $\{x^k\}$ is bounded, so $Acc\left(x^k\right) \neq \emptyset$. Take a subsequence $\{x^{k_j}\}$, such that $x^{k_j} \to \overline{x}$. Mimicking the proof of Theorem 4.1 we obtain that $\overline{x} \in SOL(T,\overline{C})$. (**b**) Consider that $Acc\left(x^k\right) \subset SOL^*(T,\overline{C})$, and that $\overline{x}$ and x^* are two cluster points of $\{x^k\}$ with $x^{k_j} \to \overline{x}$ and $x^{k_l} \to x^*$; as $\overline{x}, x^* \in SOL^*(T,\overline{C})$, then from Proposition 4.4-(ii) both $\{H(\overline{x}, x^k)\}$ and $\{H(x^*, x^k)\}$ converge. Finally mimicking the three possibilities analyzed in the proof of Theorem 4.2 we obtain the desired result. □

5. Rate of convergence

In this section we prove the linear or superlinear rate of convergence of the **IPP** algorithm. Without lost of generality we only analyze the pseudomonotone case because the quasimonotone ones is similar replacing in the assumption below the set $SOL(T,\overline{C})$ by $SOL^*(T,\overline{C})$.

Consider the following additional assumption:

(H4) For $\overline{x} \in SOL(T,\overline{C})$ such that $x^k \to \overline{x}$, there exist $\delta = \delta(\overline{x}) > 0$ and $\tau_k = \tau_k(\overline{x}) > 0$, such that for all $w \in B(0,\delta) \subset \mathbb{R}^n$ and for all x^k with $w \in T(x^k)$, we have

$$H(\overline{x}, x^k) \leq \tau_k \|w\|^2. \tag{5.20}$$

Another assumption that we also assume for the proximal distance $(d,H) \in \mathcal{F}_+(\overline{C})$ is the following:

(H5) The function $\bigtriangledown_1 d(.,u)$ satisfies the following condition: For any $x_0 \in \overline{C}$ there exist $L > 0$ and $r > 0$ such that

$$\| \bigtriangledown_1 d(x,u) - \bigtriangledown_1 d(\overline{x},u)\| \leq L\|x - \overline{x}\|, \qquad \forall x, \overline{x} \in B(x_0, r) \cap C,\ \forall u \in C.$$

Remark 5.1. With respect to previous conditions we make the following comments:

(1) The assumption (**H4**), also called a growth condition at the point of convergence $\overline{x} \in \mathbb{R}^n$; has been motivated from Baygorrea et al. [3] and Tang and Huang [35]. From our point of view, is the first time that this condition includes the induced proximal distance $H(.,.)$.

(2) A broad class of φ-divergence proximal distances, second order homogeneous proximal distances and Bregman distances satisfy the assumption (**H5**) for any $x_0 \in C$, with $C = \mathbb{R}^n_{++}$. Indeed, from Definition

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

with $\varphi \in C^2(C)$, $\varphi(t) = uh(t) + \frac{v}{2}(t-1)^2$, $v \geq uh''(1) > 0$, and $h \in C^2(C)$, we note that $\bigtriangledown_1 d_\varphi \in C^1(C)$, so that $\bigtriangledown_1 d_\varphi(.,y)$ is locally Lipschitz continuous on C. Therefore (**H5**) holds for any $x_0 \in C$.

Also if we consider h as a Bregman function such that $h \in C^2(C)$, then by Definition of Bregman distance, $\bigtriangledown_1 D_p(.,y) = \bigtriangledown h(.) - \bigtriangledown h(y)$ is continuously differentiable, then $\bigtriangledown_1 D_p(.,y)$ is locally Lipschitz on C. Therefore, again (**H5**) holds for any $x_0 \in C$.

Lemma 5.1. *Let T be a pseudomonotone mapping, $(d,H) \in \mathcal{F}_+(\overline{C})$ and suppose that both assumptions **(H1)**, **(H2)**, **(H3)**, **(H4)**, **(H5)** and condition **(Iviii)** are satisfied and $0 < \lambda_k < \overline{\lambda}$. Then*

i) *there exists* $\widetilde{k} \in \mathbb{N}$ *such that*

$$\|u^k\| < \delta, \quad \forall\, k \geq \widetilde{k}, \tag{5.21}$$

where u^k *is given by (3.10);*

ii) *it holds that*

$$H(\overline{x}, x^k) \leq \tau_k \lambda_k^2 (\eta_k + L\sqrt{\theta})^2 H(x^k, x^{k-1}), \quad \forall\, k \geq \widetilde{k}. \tag{5.22}$$

Proof. **i)** Let $\overline{x} = \lim_{k\to+\infty} x^k$, then $\overline{x} \in SOL(T, \overline{C})$, and thus from assumption (**H5**), there exist $L > 0$ and $r > 0$ such that

$$\| \bigtriangledown_1 d(x, x^{k-1}) - \bigtriangledown_1 d(y, x^{k-1})\| \leq L\|x - y\|, \qquad \forall x, y \in B(\overline{x}, r) \cap C.$$

As $\overline{x} = \lim_{k\to+\infty} x^k$, then there exists $l_0 \in I\!N$ such that $x^l \in B(\overline{x}, r)$, for all $l \geq l_0$. Taking $k \geq l_0 + 1$ and from the above inequality we have

$$\|\bigtriangledown_1 d(x^k, x^{k-1})\| = \|\bigtriangledown_1 d(x^k, x^{k-1}) - \bigtriangledown_1 d(x^{k-1}, x^{k-1})\| \leq L\|x^k - x^{k-1}\|. \tag{5.23}$$

From (3.10) we obtain

$$\|u^k\| = \|e^k - \lambda_k \bigtriangledown_1 d(x^k, x^{k-1})\| \leq \|e^k\| + \lambda_k\| \bigtriangledown_1 d(x^k, x^{k-1})\|, \tag{5.24}$$

so, taking into account (3.11), (5.23), the condition (**Iviii**), and the fact that $\lambda_k \leq \overline{\lambda}$, we have that for $k \geq l_0 + 1$ the inequality (5.24) implies

$$\begin{aligned} \|u^k\| &\leq \lambda_k \eta_k \sqrt{H(x^k, x^{k-1})} + \lambda_k L\sqrt{\theta}\sqrt{H(x^k, x^{k-1})} \\ &= \lambda_k(\eta_k + L\sqrt{\theta})\sqrt{H(x^k, x^{k-1})} \qquad (5.25) \\ &\leq \overline{\lambda}(\eta_k + L\sqrt{\theta})\sqrt{H(x^k, x^{k-1})}. \qquad (5.26) \end{aligned}$$

Since $\eta_k \to 0$ and $H(x^k, x^{k-1}) \to 0$ (see Proposition 4.2-(**iv**)), taking $\delta > 0$, there exists $\widetilde{k} \in \mathbb{N}$ with $\widetilde{k} \geq l_0 + 1$ such that $\|u^k\| < \delta$ for all $k \geq \widetilde{k}$.

ii) In (5.20) taking $w = u^k$ for all $k \geq \widetilde{k}$, we have

$$H(\overline{x}, x^k) \leq \tau_k \|u^k\|^2. \tag{5.27}$$

Therefore, the relation (5.22) follows from the last inequality combined with (5.25). □

Below we present a theorem related to the convergence rate of the inexact algorithm, thus completing the convergence result given by Theorem 4.1 and Theorem 4.2.

Theorem 5.1. *Let* T *be a pseudomonotone mapping,* $(d, H) \in \mathcal{F}_+(\overline{C})$ *and suppose that both the assumptions (**H1**), (**H2**), (**H3**), (**H4**), (**H5**) and condition (**Iviii**) are satisfied and* $0 < \lambda_k \leq \overline{\lambda}$. *Then,*

$$H(\overline{x}, x^k) \leq r_k H(\overline{x}, x^{k-1}), \tag{5.28}$$

for k *sufficiently large, where*

$$r_k = \left(\frac{4\tau_k(\eta_k + L\sqrt{\theta})^2}{4\tau_k(\eta_k + L\sqrt{\theta})^2 + \frac{4\gamma - \eta_k}{\overline{\lambda}^2}} \right) \left(\frac{1}{1 - \theta\eta_k} \right).$$

(1) *If* $\tau_k = \tau > 0$ *then,* $\{x^k\}$ *converges* H*-linearly to* $\overline{x} \in SOL(T, \overline{C})$.

(2) *If* $\{\tau_k\}$ *converges to zero then,* $\{x^k\}$ *converges* H*-superlinearly to* $\overline{x} \in SOL(T, \overline{C})$.
(3) *If* $\tau_k = \tau > 0$ *and* $\lambda_k \searrow 0$ *then,* $\{x^k\}$ *converges* H*-superlinearly to* $\overline{x} \in SOL(T, \overline{C})$.

Proof. Let $\overline{x} \in SOL(T, \overline{C})$ be the limit point of the sequence $\{x^k\}$ and $u^k \in T(x^k)$ given by (3.10). Due to the relationship (5.21) we have to $\|u^k\| < \delta$ for all $k \geq \widetilde{k}$. So $u^k \in B(0, \delta)$, for all $k \geq \widetilde{k}$.
Considering the inequality (5.22) in (4.14) for all $k \geq \max\{k_0, \widetilde{k}\}$, it follows that

$$H(\overline{x}, x^k) \leq \left(1 + \frac{\theta\eta_k}{1 - \theta\eta_k}\right) H(\overline{x}, x^{k-1}) - \left(\gamma - \frac{\eta_k}{4}\right)\left(\frac{1}{\tau_k\lambda_k^2(\eta_k + L\sqrt{\theta})^2}\right) H(\overline{x}, x^k),$$

Thus, we obtain for all $k \geq \max\{k_0, \widetilde{k}\}$:

$$\left(1 + \frac{4\gamma - \eta_k}{4\tau_k\lambda_k^2(\eta_k + L\sqrt{\theta})^2}\right) H(\overline{x}, x^k) \leq \left(\frac{1}{1 - \theta\eta_k}\right) H(\overline{x}, x^{k-1}).$$

As $\tau_k > 0$ and also $(4\gamma - \eta_k) > 0$, then for all $k \geq \max\{k_0, \widetilde{k}\}$ we have

$$H(\overline{x}, x^k) \leq \beta_k H(\overline{x}, x^{k-1}), \tag{5.29}$$

where

$$\beta_k = \left(\frac{4\tau_k(\eta_k + L\sqrt{\theta})^2}{4\tau_k(\eta_k + L\sqrt{\theta})^2 + \frac{4\gamma - \eta_k}{\lambda_k^2}}\right)\left(\frac{1}{1 - \theta\eta_k}\right). \tag{5.30}$$

Since that $\lambda_k \leq \overline{\lambda}$ for all $k \in \mathbb{N}$, we obtain

$$\beta_k \leq r_k, \tag{5.31}$$

where

$$r_k = \left(\frac{4\tau_k(\eta_k + L\sqrt{\theta})^2}{4\tau_k(\eta_k + L\sqrt{\theta})^2 + \frac{4\gamma - \eta_k}{\overline{\lambda}^2}}\right)\left(\frac{1}{1 - \theta\eta_k}\right).$$

Thus we obtain (5.28).

(1) Let $\tau_k = \tau > 0$, then taking into account that $\eta_k \to 0$, then

$$r_k \to \left(\frac{4\tau L^2\theta}{4\tau L^2\theta + \frac{4\gamma}{\overline{\lambda}^2}}\right).$$

Thus, there exists a positive number $k_1 \in \mathbb{N}$ with $k \geq k_1$, such that

$$\beta_k \leq r_k < \frac{1}{2}\left(1 + \frac{4\tau L^2\theta}{4\tau L^2\theta + \frac{4\gamma}{\overline{\lambda}^2}}\right) < 1 \quad \forall\, k \geq k_1.$$

Then, in (5.29) we have for all $k \geq \max\{k_0, \widetilde{k}, k_1\}$:

$$H(\overline{x}, x^k) \leq \bar{\theta} H(\overline{x}, x^{k-1}),$$

where

$$\bar{\theta} = \left(\frac{4\tau L^2 \theta}{4\tau L^2 \theta + \frac{4\gamma}{\bar{\lambda}^2}} \right).$$

Thus, the sequence $\{x^k\}$ converges $H-$linearly to $\overline{x}$.

(2) If $\{\tau_k\}$ converges to zero then from (5.28) we have that $\{r_k\}$ converges to zero and thus we obtain that the sequence $\{x^k\}$ converges $H-$superlinearly to $\overline{x}$.

(3) Let $\tau_k = \tau > 0$, we have from (5.29) and (5.30)

$$H(\overline{x}, x^k) \leq \beta_k H(\overline{x}, x^{k-1}), \tag{5.32}$$

where

$$\beta_k = \left(\frac{4\tau(\eta_k + L\sqrt{\theta})^2}{4\tau(\eta_k + L\sqrt{\theta})^2 + \frac{4\gamma - \eta_k}{\lambda_k^2}} \right) \left(\frac{1}{1 - \theta\eta_k} \right). \tag{5.33}$$

As $\lambda_k \searrow 0$ and $\eta_k \to 0$, then sequence $\{x^k\}$ converges $H-$superlinearly to $\overline{x}$.
□

The following result shows that we obtain the genuine linear and superlinear rate of convergence for a class of proximal distances which includes the class of logarithmic quadratic distances.

Corollary 5.1. *Let T be a pseudomonotone mapping, $(d, H) \in \mathcal{F}_+(\overline{C})$, and suppose that both assumptions (**H1**), (**H2**), (**H3**), (**H4**),(**H5**) and the condition*

$$H(x, y) = \theta\|x - y\|^2 \tag{5.34}$$

for some $\theta > 0$, are satisfied and $0 < \lambda_k \leq \overline{\lambda}$. Then

(1) *If $\tau_k = \tau > 0$, then $\{x^k\}$ converges linearly to $\overline{x} \in SOL(T, \overline{C})$*
(2) *If $\{\tau_k\}$ converges to zero then $\{x^k\}$ converges superlinearly to $\overline{x} \in SOL(T, \overline{C})$*
(3) *If $\tau_k = \tau > 0$, and $\lambda_k \searrow 0$ then $\{x^k\}$ converges superlinearly to $\overline{x} \in SOL(T, \overline{C})$.*

Remark 5.2. A class of proximal distances which satisfies the above condition (5.34) is the proximal distance with second order homogeneous distances, see Remark 4.1. Other class of proximal distances which satisfy the condition (5.34) are the induced proximal distances by double regularization, introduced by Silva and Eckstein [32], see Sarmiento et al. [30].

6. Conclusions, discussions and future researches

In this paper we obtain the global convergence of an inexact proximal point algorithm using a class of proximal distance, called **IPP** algorithm, to solve pseudomonotone and quasimonotone (VIP) defined on convex sets with interior nonempty. We also prove a general linear and superlinear rate of convergence of this algorithm.

For the class of second order proximal distance, in particular for the logarithmic quadratic proximal distance, and the induced proximal distances by double regularization we obtain genuine linear and superlinear convergence of the **IPP** algorithm. For a class of $\phi-$divergence distance and Bregman distance defined by strongly convex functions we obtain $H-$linear and $H-$superlinear rate of convergence.

The results of this paper are new even for monotone (VIP) and improve, for the quasimonotone case, the results obtained by Langenberg [20], Brito et al. [5] and Papa Quiroz et al. [24], in the following sense:

- In Langenberg [20] was proved the global convergence of the following inexact proximal algorithm: find $x^{k+1} \in C$, $y^k \in C$ and $t^k \in T(y^k)$ such that
$$t^k + \lambda_k \nabla_1 D_h(x^{k+1}, x^k) = e^k, \tag{6.35}$$
where $||e^k|| \leq \delta_k$ and
$$D_h(y^k, x^{k+1}) \leq \sigma^2 D_h(y^k, x^k)$$
with $\sigma \in [0, 1)$ and D_h is a Bregman distance generated by a strongly convex Bregman function h. The error criteria should satisfy the following condition
$$\sum_{k=1}^{\infty} \delta_k < +\infty. \tag{6.36}$$
However, the paper of Langenberg [20] does not present rate of convergence results. Observe that for the case $y^k = x^{k+1}$ the equation (6.35) is a particular case of the equation (3.10) of the **IPP** algorithm when the proximal distance is the Bregman distance. However, the condition (6.36) is different of the conditions (3.11)-(3.12) of **IPP** algorithm. In this sense our error criteria permit to obtain rate of convergence results.
- Brito et al. [5] presented an exact proximal algorithm using the logarithmic quadratic proximal distance. In this case the **IPP** algorithm is a inexact version of the paper [5] using the the logarithmic quadratic proximal distance.
- Papa Quiroz et al. [24] presented an inexact proximal algorithm using the following error criteria (introduced by Eckstein [9]):
$$\sum_{k=1}^{+\infty} \frac{\|e^k\|}{\lambda_k} < +\infty \tag{6.37}$$
$$\sum_{k=1}^{+\infty} \frac{|\langle e^k, x^k \rangle|}{\lambda_k} < +\infty. \tag{6.38}$$
They observed that it is possible to get rid the condition (6.38) for some proximal distances. Observe that if $\{\lambda_k\}$ is bounded then from (6.37) we obtain that
$$\sum_{k=1}^{+\infty} \|e^k\| < +\infty,$$
and therefore for the case $y^k = x^{k+1}$ the algorithm introduced by Langenberg in [20] is a particular case of the paper of Papa Quiroz et al. [24] applied

to Bregman distances with strongly convex Bregman functions. However, in that paper the rate of convergence does not studied.

A future research may be the introduction of a (PPM), based in the **IPP** algorithm and the paper of Auslender and Teboulle [1], to solve the following variational inequality problem: find $x^* \in \overline{C}$ and $y^* \in T(x^*)$, such that

$$\langle y^*, x - x^* \rangle \geq 0, \tag{6.39}$$

$\forall x \in \overline{C} \cap \{x : Ax = b\}$ where, A is an $m \times n$ matrix, $b \in \mathbb{R}^n$, $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a point-to-set mapping not necessarily monotone, C is a nonempty open convex set, $\overline{C}$ is the closure of C in $\mathbb{R}^n$.

References

[1] A. Auslender, M. Teboulle, *Interior projection-like methods for monotone variational inequalities*, Math. Program, Serie A, (2005), no. 104, 39-68.

[2] A. Auslender, M. Teboulle, *Interior gradient and proximal methods for convex and conic optimization*, SIAM Journal of Optimization, **16**(2006), no. 3, 697-725.

[3] N. Baygorrea, E.A. Papa Quiroz, N. Maculan, *Inexact proximal point methods for quasiconvex minimization on Hadamard manifolds*, J. Oper. Res. Soc. China, **4**(2016), no. 4, 397-424.

[4] A. Bnouhachem, X. Qin, *An inertial proximal Peaceman-Rachford splitting method with SQP regularization for convex programming*, J. Nonlinear Funct. Anal., **50**(2020), 1-17.

[5] S.A. Brito, J.X. da Cruz Neto, J.O. Lopes, P.R. Oliveira, *Interior algorithm for quasiconvex programming problems and variational inequalities with linear constraints*, J. Optim Theory Appl., **154**(2012), no. 1, 217-234.

[6] X. Cai, *A proximal point algorithm with asymmetric linear term*, Optimization Letters, **13**(2019), no. 4, 777-793.

[7] L.C. Ceng, *Two inertial linesearch extragradient algorithms for the bilevel split pseudomonotone variational inequality with constraints*, J. Appl. Numer. Optim., **2**(2020), 213-233.

[8] G. Chen, M. Teboulle, *Convergence analysis of a proximal-like minimization algorithm using Bregman functions*, SIAM J. Optim., (1993), no. 3, 538-543.

[9] J. Eckstein, *Approximate iterations in Bregman-function-based proximal algorithms*, Mathematical Programming, **83**(1998), 113-123.

[10] J. Eckstein, D.P. Bertsekas, *On the Douglas-Rachford splitting method and the proximal algorithm for maximal monotone operators*, Mathematical Programming, **55**(1992), no. 3, 293-318.

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

[12] A. Gibali, *A new Bregman projection method for solving variational inequalities in Hilbert*, Pure Appl. Funct. Anal., **3**(2018), 403-415.

[13] D. Han, *A generalized proximal-point-based prediction-correction method for variational inequality problems*, Journal of Computational and Applied Mathematics, **221**(2008), 183-193.

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

[15] D.V. Hieu, P. Cholambjiak, *Modified extragradient method with Bregman distance for variational inequalities*, Applicable Analysis, doi.org/10.1080/00036811.2020.1757078

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

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

[18] K.C. Kiwiel, *Proximal minimization methods with generalized Bregman functions*, SIAM J. Control Optim., **35**(1997), 1142-1168.

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

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

[21] N. Langenberg, R. Tichatschke, *Interior proximal methods for quasiconvex optimization*, J. Global Optim., **52**(2012), 641-661.

[22] Y. Liu, D. Sun, K. Toh, *An implementable proximal point algorithmic framework for nuclear norm minimization*, Mathematical Programming, **133**(2012), no. 1-2, 399-436.

[23] E.A. Papa Quiroz, L. Mallma Ramirez, P.R. Oliveira, *An inexact proximal method for quasiconvex minimization*, European J. Operational Research, **246**(2015), 721-729.

[24] E.A. Papa Quiroz, L. Mallma Ramirez, P.R. Oliveira, *An inexact algorithm with proximal distances for variational inequalities*, RAIRO Oper. Res., **52**(2018), no. 1, 159-176.

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

[26] B.T. Polyak, *Introduction to Optimization*, Optimization Software, New York, 1987.

[27] R.T. Rockafellar, *Monotone operations and the proximal point method*, SIAM J. Control Optimization, **14**(1976), 877-898.

[28] R.T Rockafellar, R. Wets, *Variational Analysis*, Grundlehren der Mathematischen, Wissenschaften 317, Springer 1998.

[29] O. Sarmiento, E.A. Papa Quiroz, P.R. Oliveira, *A proximal multiplier method for separable convex minimization*, Optimization, **65**(2016), no. 2, 501-537.

[30] O. Sarmiento, E.A. Papa Quiroz, P.R. Oliveira, *Convergence rate of a proximal multiplier algorithm for separable convex minimization*, Optimization, **66**(2017), no. 2, 251-270.

[31] K. Shimizu, K. Hishinuma, H. Iiduka, *Parallel computing proximal method for non-smooth convex optimization with fixed point constraints of quasi-nonexpansive mappings*, Appl. Set-Valued Anal. Optim., (2020), no. 2, 1-17.

[32] P.J.S. Silva, J. Eckstein, *Double-regularization proximal methods, with complementarity applications*, Comput. Optim. Appl., **33**(2006), 115-156.

[33] M.V. Solodov, *Convergence rate analysis of iteractive algorithms for solving variational inequality problems*, Mathematical Programming, Ser. A., **96**(2003), 513-528.

[34] M.V. Solodov, B.F. Svaiter, *An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions*, Mathematics of Operations Research, **25**(2000), 214-230.

[35] G.J. Tang, N.J. Huang, *An inexact proximal point algorithm for maximal monotone vector fields on Hadamard manifolds*, Operations Research Letters, **41**(2013), no. 6, 586-591.

[36] M. Teboulle, *Convergence of proximal-like algorithms*, SIAM J. Optimization, **7**(1997), no. 1, 1069-1083.

[37] P. Tseng, *On linear convergence of iterative methods for the variational inequality problem*, Journal of Computational and Applied Mathematics, **60**(1995), 237-252.

[38] X. Wang, X. Ou, T. Zhang, J. Chen, *An alternate minimization method beyond positive definite proximal regularization: convergence and complexity*, J. Nonlinear Var. Anal., (2019), no. 3, 333-355.

Received: September 6, 2020; Accepted: March 22, 2021.