A variant of Forward-Backward splitting method for the sum of two monotone operators with a new search strategy

J. Y. Bello Cruz* R. Díaz Millán†

May 8, 2013

Abstract

In this paper, we propose variants of Forward-Backward splitting method for finding a zero of the sum of two operators. A classical modification of Forward-Backward method was proposed by Tseng, which is known to converge when the forward and the backward operators are monotone and with Lipschitz continuity of the backward operator. The conceptual algorithm proposed here improves Tseng's method in some instances. The first and main part of our approach, contains an explicit Armijo-type search in the spirit of the extragradient-like methods for variational inequalities. During the iteration process the search performs only one calculation of the forward-backward operator, in each tentative of the step. This achieves a considerable computational saving when the forward-backward operator is computationally expensive. The second part of the scheme consists in special projection steps. The convergence analysis of the proposed scheme is given assuming monotonicity on both operators, without Lipschitz continuity assumption on the backward operator.

Keywords: Armijo-type search, Maximal monotone operators, Splitting methods
Mathematical Subject Classification (2008): 90C47, 49J35.

1 Introduction

First of all, we introduce the notation. The inner product in $\mathbb{R}^n$ is denoted by $\langle\cdot,\cdot\rangle$ and the norm induced by the inner product by $\|\cdot\|$. For X a nonempty, convex and closed subset of $\mathbb{R}^n$, we define the orthogonal projection of x onto X by $\mathcal{P}_X(x)$, as the unique point in X, such that $\|\mathcal{P}_X(x)-y\| \le \|x-y\|$ for all $y\in X$. Let $N_X(x)$ be the normal cone to X at $x\in X$, i.e., $N_X(x)=\{d\in\mathbb{R}^n : \langle d, x-y\rangle \ge 0 \ \ \forall y\in X\}$. Recall that an operator $T:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is monotone if, for all $(x,u),(y,v)\in Gr(T)$, we have $\langle x-y, u-v\rangle\ge 0$, and it is maximal if T has no proper monotone extension in the graph inclusion sense.

In this paper, we present a modified method for solving monotone inclusion problems for the sum of two operators. Given the monotone operators, $A: dom(A)\subseteq\mathbb{R}^n\to\mathbb{R}^n$ point-to-point and $B: dom(B)\subseteq\mathbb{R}^n\rightrightarrows\mathbb{R}^n$ point-to-set, the inclusion problem consists in:

$$\text{Find } \ x\in\mathbb{R}^n \ \text{ such that } \ 0\in(A+B)(x). \tag{1}$$

*Federal University of Goiás, Goiânia, Brazil, e-mail: yunier@impa.br
†Instituto de Matemática e Estatística. Federal University of Goiás, CEP 74001-970, Goiânia, Brazil, e-mail: rdm@impa.br

The solution set is denoted by $S^* := \{x \in \mathbb{R}^n : 0 \in (A+B)(x)\}$. This problem has recently received a lot attention due to the fact that many nonlinear problems, arising within applied areas, are mathematically modeled as nonlinear operator equations and/or inclusions, which are decomposed as the sum of two operators. We focus our attention in the called splitting method, which is an iterative method, for which each iteration involves only the individual operators, A or B, but not the sum, $A+B$; see [8, 19, 9].

A classical splitting method for solving problem (1) is the so called Forward-Backward splitting method as proposed in [15]. Assuming that $dom(B) \subseteq dom(A)$, the scheme is given as follows:

$$x^{k+1} = (I + \beta_k B)^{-1}(I - \beta_k A)(x^k), \tag{2}$$

where $\beta_k > 0$ for all k. The iteration defined by (2) converges when the inverse of the forward mapping is strongly monotone as well as over other undesired assumptions on the stepsize β_k and the operator B; see, for instance, [15] and [18]. An important and promising modification of Scheme (2) was presented by Tseng in [19]. It consists in:

$$\begin{aligned} J(x^k,\beta_k) &= (I+\beta_k B)^{-1}(I-\beta_k A)(x^k) && (3)\\ x^{k+1} &= \mathcal{P}_X\Big(J(x^k,\beta_k) - \beta_k\big[A(J(x^k,\beta_k)) - A(x^k)\big]\Big), && (4)\end{aligned}$$

where X is a suitable nonempty, closed and convex set, belonging to $dom(A)$. The stepsize β_k is chosen to be the largest $\beta \in \{\sigma, \sigma\theta, \sigma\theta^2, \cdots\}$, satisfying:

$$\beta \left\| A\big(J(x^k,\beta)\big) - A(x^k)\right\| \le \delta \left\| J(x^k,\beta) - x^k \right\|, \tag{5}$$

with $\theta, \delta \in (0,1)$ and $\sigma > 0$. Note that there exists various choices for the set X. If $dom(B)$ is closed, then the result of Minty in [16], implies that $dom(B)$ is convex, hence we may choose $X = dom(B)$; see [19].

The convergence of (3)-(5) was established assuming maximal monotonicity of A and B, as well as Lipschitz continuity of A. It is important to say that, in the above scheme, in order to compute β_k satisfying (5), the forward-backward operator (3) must be calculated, in each tentative of the step. From a computational point of view, this represents a considerable drawback.

In order to overcome these two serious limitations a conceptual algorithm has been proposed, containing three variants, which are denominated Algorithm 1, 2 and 3. We show the convergence to a solution of problem (1), assuming only monotonicity of both operators however without demands Lipschitz continuity of A. Our approach contains two parts. The first being a separating halfspace, containing the solution set of the problem, is found. This procedure employs a new Armijo-type search which performs only one calculation of the forward-backward operator instead Tseng's algorithm. In the remaining part, the current point is projected onto a suitable set. The main difference between the three proposed algorithms, is determined by the way of the projection steps are carried out.

When $B = N_X$, problem (1) may be written as $0 \in A(x) + N_X(x)$, or equivalently $\langle A(x), y - x\rangle \ge 0$ for all $y \in X$. This problem is the well studied variational inequality problem with numerous applications in optimization theory; see [11, 14]. An excellent survey of projection methods for variational inequality problems can be found in [10]. In this setting the variants of the proposed conceptual algorithm are related to the algorithms, presented in [12, 20, 3].

This work is organized as follows. The next section provides some preliminary results that will be used in the remainder of this paper. The conceptual algorithm is presented in Section 3, where its three variants, called Algorithm 1, 2 and 3, are proposed. Section 4 contains the convergence analysis of the algorithms. Finally, Section 5 gives some concluding remarks.

2 Preliminaries

In this section, we present some definitions and results needed for the convergence analysis of the proposed methods. First, we state two well-known facts on orthogonal projections.

Proposition 2.1. *Let X be any nonempty, closed and convex set in $\mathbb{R}^n$, and $\mathcal{P}_X$ the orthogonal projection onto X. For all $x, y \in \mathbb{R}^n$ and all $z \in X$ the following hold:*

i) $\|\mathcal{P}_X(x) - \mathcal{P}_X(y)\|^2 \leq \|x - y\|^2 - \|(\mathcal{P}_X(x) - x) - \big(\mathcal{P}_X(y) - y\big)\|^2$.

ii) $\langle x - \mathcal{P}_X(x), z - \mathcal{P}_X(x)\rangle \leq 0$.

iii) $\mathcal{P}_X = (I + N_X)^{-1}$.

Proof. See Proposition 2.3 in [2]. □

In the following we state some useful results on maximal monotone operators.

Lemma 2.1. *Let $T : dom(T) \subseteq \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximal monotone operator. Then,*

i) $G(T)$ is closed.

ii) T is bounded on bounded subsets of the interior of its domain.

Proof. i) See Proposition 4.2.1(ii) in [7].

ii) See Lemma 5(iii) in [4].

□

Proposition 2.2. *Let $T : dom(T) \subseteq \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a point-to-set and maximal monotone operator. Given $\beta > 0$ then the operator $(I + \beta T)^{-1} : \mathbb{R}^n \to dom(T)$ is single valued and maximal monotone.*

Proof. See Theorem 4 in [17]. □

Proposition 2.3. *Given $\beta > 0$ and $S : dom(S) \subseteq \mathbb{R}^n \to \mathbb{R}^n$ and $T : dom(T) \subseteq \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ two maximal monotone operators, then*

$$x = (I + \beta T)^{-1}(I - \beta S)(x),$$

if and only if, $0 \in (S + T)(x)$.

Proof. See Proposition 3.13 in [8]. □

Now, we define the so called Fejér convergence.

Definition 2.1. *Let S be a nonempty subset of $\mathbb{R}^n$. A sequence $\{x^k\} \subset \mathbb{R}^n$ is said to be Fejér convergent to S, if and only if, for all $x \in S$ there exists $k_0 \geq 0$, such that $\|x^{k+1} - x\| \leq \|x^k - x\|$ for all $k \geq k_0$.*

This definition was introduced in [5] and has been elaborated further in [13] and [1]. A useful result on Fejér sequences is the following.

Proposition 2.4. *If $\{x^k\}$ is Fejér convergent to S, then:*

i) the sequence $\{x^k\}$ is bounded;

ii) the sequence $\{\|x^k - x\|\}$ is convergent for all $x \in S$;

iii) if a cluster point x^* belongs to S, then the sequence $\{x^k\}$ converges to x^*.

Proof. See Theorem 1 in [6]. □

3 The Conceptual Algorithm

Let $A : \mathbb{R}^n \to \mathbb{R}^n$ and $B : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be two maximal monotone operators, with A point-to-point and B point-to-set. Assume that $dom(B) \subseteq dom(A)$. Choose any nonempty, closed and convex set, $X \subseteq dom(B)$, satisfying $X \cap S^* \neq \emptyset$. Thus, from now on, the solution set, S^*, is nonempty. Also we assume that the operator B satisfies, that for each bounded subset V of $dom(B)$ there exists $R > 0$, such that $B(x) \cap B[0, R] \neq \emptyset$, for all $x \in V$. We emphasize that this assumption holds trivially if $dom(B) = \mathbb{R}^n$ or $V \subset int(dom(B))$ or B is the normal cone in any subset of $dom(B)$; see Lemma 2.1(ii).

Let $\{\beta_k\}_{k=0}^{\infty}$ be a sequence such that $\{\beta_k\} \subseteq [\check{\beta}, \hat{\beta}]$ with $0 < \check{\beta} \leq \hat{\beta} < \infty$, and be $\theta, \delta \in (0, 1)$. The algorithm is defined as follows:

Conceptual Algorithm
Initialization Step 1: Take

$$x^0 \in X.$$

Iterative Step 1: Given x^k and β_k, compute the forward-backward operator at x^k,

$$J(x^k, \beta_k) := (I + \beta_k B)^{-1}(I - \beta_k A)(x^k). \tag{6}$$

Stop Criteria 1: If $x^k = J(x^k, \beta_k)$ stop.
Inner Loop: Otherwise, begin the inner loop over j.
Put $j = 0$ and chose any $u_j^k \in B\big(\theta^j J(x^k, \beta_k) + (1 - \theta^j)x^k\big) \cap B[0, R]$. If

$$\Big\langle A\big(\theta^j J(x^k, \beta_k) + (1 - \theta^j)x^k\big) + u_j^k, x^k - J(x^k, \beta_k)\Big\rangle \geq \frac{\delta}{\beta_k}\|x^k - J(x^k, \beta_k)\|^2, \tag{7}$$

then $j(k) = j$ and stop. Else, $j = j + 1$.
Iterative Step 2: Set

$$\alpha_k := \theta^{j(k)}, \tag{8}$$

$$\bar{u}^k = u_{j(k)}^k, \tag{9}$$

$$\bar{x}^k := \alpha_k J(x^k, \beta_k) + (1 - \alpha_k)x^k \tag{10}$$

and

$$x^{k+1} := \mathcal{F}(x^k). \tag{11}$$

Stop Criteria 2: If $x^{k+1} = x^k$ then stop.

Now we consider three variants on this conceptual algorithm. The difference is given by the definition of the procedure $\mathcal{F}$ in (11).

$$\mathcal{F}_1(x^k) = \mathcal{P}_X\big(\mathcal{P}_{H(\bar{x}^k, \bar{u}^k)}(x^k)\big); \tag{12}$$

$$\mathcal{F}_2(x^k) = \mathcal{P}_{X \cap H(\bar{x}^k, \bar{u}^k)}(x^k); \tag{13}$$

$$\mathcal{F}_3(x^k) = \mathcal{P}_{X \cap H(\bar{x}^k, \bar{u}^k) \cap W(x^k)}(x^0); \tag{14}$$

where

$$H(x, u) := \big\{y \in \mathbb{R}^n : \langle A(x) + u, y - x\rangle \leq 0\big\} \tag{15}$$

and

$$W(x) := \big\{y \in \mathbb{R}^n : \langle y - x, x^0 - x\rangle \leq 0\big\}. \tag{16}$$

The variants of the conceptual algorithm given in (12), (13) and (14) are called Algorithm 1, 2 and 3, respectively.

4 Convergence Analysis

In this section we analyze the convergence of the algorithms presented in the previous section. First, we present some general properties as well as prove the well-definition of the conceptual algorithm.

Lemma 4.1. *For all* $(x, u) \in Gr(B)$, $S^* \subseteq H(x, u)$.

Proof. Take $x^* \in S^*$. Using the definition of the solution, there exists $v^* \in B(x^*)$, such that $0 = A(x^*) + v^*$. By the monotonicity of $A + B$, we have

$$\langle A(x) + u - (A(x^*) + v^*), x - x^* \rangle \geq 0,$$

for all $(x, u) \in Gr(B)$. Hence,

$$\langle A(x) + u, x^* - x \rangle \leq 0$$

and by (15), $x^* \in H(x, u)$. □

From now on, $\{x^k\}$ is the sequence generated by the conceptual algorithm.

Proposition 4.1. *The conceptual algorithm is well-defined.*

Proof. By Proposition 2.3, Stop Criteria 1 is well-defined. The proof of the well-definition of $j(k)$ is by contradiction. Assume that for all $j \geq 0$ having chosen $u_j^k \in B\big(\theta^j J(x^k, \beta_k) + (1 - \theta^j) x^k\big) \cap B[0, R]$,

$$\Big\langle A\big(\theta^j J(x^k, \beta_k) + (1 - \theta^j) x^k\big) + u_j^k, x^k - J(x^k, \beta_k) \Big\rangle < \frac{\delta}{\beta_k} \|x^k - J(x^k, \beta_k)\|^2.$$

Since the sequence $\{u_j^k\}_{j=0}^{\infty}$ is bounded, there exists a subsequence $\{u_{\ell_j}^k\}_{j=0}^{\infty}$ of $\{u_j^k\}_{j=0}^{\infty}$, which converges to an element u^k belonging to $B(x^k)$ by maximality. Taking the limit over the subsequence $\{\ell_j\}$, we get

$$\big\langle \beta_k A(x^k) + \beta_k u^k, x^k - J(x^k, \beta_k) \big\rangle \leq \delta \|x^k - J(x^k, \beta_k)\|^2. \tag{17}$$

It follows from (6) that

$$\beta_k A(x^k) = x^k - J(x^k, \beta_k) - \beta_k v^k,$$

for some $v^k \in B(J(x^k, \beta_k))$.
Now, the above equality together with (17), lead to

$$\|x^k - J(x^k, \beta_k)\|^2 \leq \Big\langle x^k - J(x^k, \beta_k) - \beta_k v^k + \beta_k u^k, x^k - J(x^k, \beta_k) \Big\rangle \leq \delta \|x^k - J(x^k, \beta_k)\|^2,$$

using the monotonicity of B for the first inequality. So,

$$(1 - \delta) \|x^k - J(x^k, \beta_k)\|^2 \leq 0,$$

which contradicts Stop Criteria 1. Thus, the conceptual algorithm is well-defined. □

Proposition 4.2. $x^k \in H(\bar{x}^k, \bar{u}^k)$ *for* $\bar{x}^k$ *and* $\bar{u}^k$ *as in* (10) *and* (9)*, respectively, if and only if,* $x^k \in S^*$.

Proof. Since $x^k \in H(\bar{x}^k, \bar{u}^k)$, $\langle A(\bar{x}^k) + \bar{u}^k, x^k - \bar{x}^k \rangle \leq 0$. Using the Armijo-type search, given in (7), and (10), we obtain

$$0 \geq \langle A(\bar{x}^k) + \bar{u}^k, x^k - \bar{x}^k \rangle = \alpha_k \langle A(\bar{x}^k) + \bar{u}^k, x^k - J(x^k, \beta_k) \rangle \geq \frac{\alpha_k}{\beta_k} \delta \| x^k - J(x^k, \beta_k) \|^2 \geq 0,$$

which implies that $x^k = J(x^k, \beta_k)$. So, by Proposition 2.3, $x^k \in S^*$. Conversely, if $x^k \in S^*$ using Lemma 4.1, $x^k \in H(\bar{x}^k, \bar{u}^k)$. □

From now on, denote $H_k := H(\bar{x}^k, \bar{u}^k)$ as (15) and $W_k := W(x^k)$ as (16), for $\bar{x}^k$ and $\bar{u}^k$ as in (10) and (9).

Finally, a useful algebraic property on the sequence generated by the conceptual algorithm, which is a direct consequence of the inner loop and (10).

Corollary 4.1. *Let $\{x^k\}$ and $\{\alpha_k\}$ be sequences generated by the conceptual algorithm, $\{\beta_k\}$. With δ and $\hat{\beta}$ as in the algorithm. Then,*

$$\langle A(\bar{x}^k) + \bar{u}^k, x^k - \bar{x}^k \rangle \geq \frac{\alpha_k \delta}{\hat{\beta}} \| x^k - J(x^k, \beta_k) \|^2 \geq 0, \tag{18}$$

for all k.

4.1 Convergence of Algorithm 1

In this subsection all results are referent to Algorithm 1, i.e., with Iterative Step 2 as

$$x^{k+1} = \mathcal{F}_1(x^k) = \mathcal{P}_X\big(\mathcal{P}_{H_k}(x^k)\big).$$

Proposition 4.3. *If Algorithm 1 stops, then $x^k \in S^*$.*

Proof. If Stop Criteria 2 is satisfied, $x^{k+1} = \mathcal{P}_X\big(\mathcal{P}_{H_k}(x^k)\big) = x^k$. Using Proposition 2.1(ii), we have

$$\langle \mathcal{P}_{H_k}(x^k) - x^k, z - x^k \rangle \leq 0, \tag{19}$$

for all $z \in X$. Now using Proposition 2.1(ii),

$$\langle \mathcal{P}_{H_k}(x^k) - x^k, \mathcal{P}_{H_k}(x^k) - z \rangle \leq 0, \tag{20}$$

for all $z \in H_k$. Since $X \cap H_k \neq \emptyset$ summing (19) and (20), with $z \in X \cap H_k$, we get

$$\| x^k - \mathcal{P}_{H_k}(x^k) \|^2 = 0.$$

Hence, $x^k = \mathcal{P}_{H_k}(x^k)$, implying that $x^k \in H_k$ and by Proposition 4.2, $x^k \in S^*$. □

From now on, assume that Algorithm 1 does not stop. Note that by Lemma 4.1 H_k is nonempty for all k. Then, the projection step, given in (12), is well-defined, i.e., if Algorithm 1 does not stop, it generates an infinite sequence $\{x^k\}$.

Proposition 4.4. *i) The sequence $\{x^k\}$ is Fejér convergente to $S^* \cap X$.*

ii) The sequence $\{x^k\}$ is bounded.

iii) $\lim_{k\to\infty}\langle A(\bar{x}^k)+\bar{u}^k, x^k-\bar{x}^k\rangle = 0.$

Proof.

(i) Take $x^*\in S^*\cap X$. Using (12), Proposition 2.1(i) and Lemma 4.1, we have

$$\begin{aligned}\|x^{k+1}-x^*\|^2 &= \|P_X(P_{H_k}(x^k))-P_X(P_{H_k}(x^*))\|^2 \le \|P_{H_k}(x^k)-P_{H_k}(x^*)\|^2\\ &\le \|x^k-x^*\|^2-\|\mathcal{P}_{H_k}(x^k)-x^k\|^2. \end{aligned} \tag{21}$$

So, $\|x^{k+1}-x^*\|\le\|x^k-x^*\|$.

(ii) Follows inmedately from item (i) and Proposition 1(i).

(iii) Take $x^*\in S^*\cap X$. Using (9) and

$$\mathcal{P}_{H_k}(x^k)=x^k-\frac{\big\langle A(\bar{x}^k)+\bar{u}^k, x^k-\bar{x}^k\big\rangle}{\|A(\bar{x}^k)+\bar{u}^k\|^2}\big(A(\bar{x}^k)+\bar{u}^k\big), \tag{22}$$

combining with (21), yields

$$\begin{aligned}\|x^{k+1}-x^*\|^2 &\le \|x^k-x^*\|^2-\left\|x^k-\frac{\big\langle A(\bar{x}^k)+\bar{u}^k, x^k-\bar{x}^k\big\rangle}{\|A(\bar{x}^k)+\bar{u}^k\|^2}\big(A(\bar{x}^k)+\bar{u}^k\big)-x^k\right\|^2\\ &= \|x^k-x^*\|^2-\frac{(\langle A(\bar{x}^k)+\bar{u}^k, x^k-\bar{x}^k\rangle)^2}{\|A(\bar{x}^k)+\bar{u}^k\|^2}.\end{aligned}$$

Reordering the above inequality, we get

$$\frac{(\langle A(\bar{x}^k)+\bar{u}^k, x^k-\bar{x}^k\rangle)^2}{\|A(\bar{x}^k)+\bar{u}^k\|^2}\le\|x^k-x^*\|^2-\|x^{k+1}-x^*\|^2. \tag{23}$$

By Proposition 2.2 and the continuity of A, $\{J(x^k,\beta_k)\}$ is bounded. Since $\{x^k\}$ and $\{\beta_k\}$ are bounded, $\{\bar{x}^k\}$ is bounded, implying the boundedness of $\{\|A(\bar{x}^k)+\bar{u}^k\|\}$.
Using Proposition 2.4(ii), the right side of (23) goes to 0, when k goes to ∞, establishing the result. □

Next we establish our main convergence result on Algorithm 1.

Theorem 4.1. *The sequence $\{x^k\}$ converges to some element belonging to $S^*\cap X$.*

Proof. We claim that there exists a cluster point of $\{x^k\}$ belonging to S^*. The existence of the cluster points follows from Proposition 4.4(ii). Let $\{x^{i_k}\}$ be a convergent subsequence of $\{x^k\}$, which converges to $\tilde{x}$. Using Proposition 4.4(iii) and taking limits in (18) over the subsequence $\{i_k\}$, we have

$$0=\lim_{k\to\infty}\langle A(\bar{x}^{i_k})+\bar{u}^{i_k}, x^{i_k}-\bar{x}^{i_k}\rangle\ge\lim_{k\to\infty}\frac{\alpha_{i_k}\delta}{\hat{\beta}}\|x^{i_k}-J(x^{i_k},\beta_{i_k})\|^2\ge 0. \tag{24}$$

Therefore,

$$\lim_{k\to\infty}\alpha_{i_k}\|x^{i_k}-J(x^{i_k},\beta_{i_k})\|=0.$$

Now consider the two possibles cases.

(a) First, assume that $\lim_{k\to\infty}\alpha_{i_k}\neq 0$, i.e., $\alpha_{i_k}\geq\bar{\alpha}$ for all k and some $\bar{\alpha}>0$. In view of (24),

$$\lim_{k\to\infty}\|x^{i_k}-J(x^{i_k},\beta_{i_k})\|=0. \tag{25}$$

Taking a subsequence, if necessary, we may assume that $\lim_{k\to\infty}\beta_{i_k}=\tilde{\beta}$ such that $\tilde{\beta}\geq\check{\beta}>0$ and since J is continuous, by the continuity of A and $(I+\beta_k B)^{-1}$ and by Proposition 2.2, (25) becomes

$$\tilde{x}=J(\tilde{x},\tilde{\beta}),$$

which implies that $\tilde{x}\in S^*$. Establishing the claim.

(b) On the other hand, if $\lim_{k\to\infty}\alpha_{i_k}=0$. We have

$$\lim_{k\to\infty}\frac{\alpha_{i_k}}{\theta}=0.$$

Define

$$y^{i_k}:=\frac{\alpha_{i_k}}{\theta}J(x^{i_k},\beta_{i_k})+\Big(1-\frac{\alpha_{i_k}}{\theta}\Big)x^{i_k}.$$

Then,

$$\lim_{k\to\infty}y^{i_k}=\tilde{x}. \tag{26}$$

Using the definition of $j(k)$ and (8), y^{i_k} does not satisfy (7) implying

$$\Big\langle A(y^{i_k})+u^{i_k}_{j(i_k)-1},x^{i_k}-J(x^{i_k},\beta_{i_k})\Big\rangle<\frac{\delta}{\beta_{i_k}}\|x^{i_k}-J(x^{i_k},\beta_{i_k})\|^2, \tag{27}$$

for $u^{i_k}_{j(i_k)-1}\in B(y^{i_k})$ and all k.

Redefining the subsequence $\{i_k\}$, if necessary, we may assume that $\{\beta_{i_k}\}$ converges to some $\tilde{\beta}$ such that $\tilde{\beta}\geq\check{\beta}>0$ and $\{u^{i_k}_{j(i_k)-1}\}_{k=0}^{\infty}$ converges to $\tilde{u}$. By the maximality of B, $\tilde{u}$ belongs to $B(\tilde{x})$. Using the continuity of J, $\{J(x^{i_k},\beta_{i_k})\}$ converges to $J(\tilde{x},\tilde{\beta})$. Using (26) and taking limit in (27) over the subsequence $\{i_k\}$, we have

$$\Big\langle A(\tilde{x})+\tilde{u},\tilde{x}-J(\tilde{x},\tilde{\beta})\Big\rangle\leq\frac{\delta}{\tilde{\beta}}\|\tilde{x}-J(\tilde{x},\tilde{\beta})\|^2. \tag{28}$$

Using (6) and multiplying by $\tilde{\beta}$ on both sides of (28), we get

$$\langle\tilde{x}-J(\tilde{x},\tilde{\beta})-\tilde{\beta}\tilde{v}+\tilde{\beta}\tilde{u},\tilde{x}-J(\tilde{x},\tilde{\beta})\rangle\leq\delta\|\tilde{x}-J(\tilde{x},\tilde{\beta})\|^2,$$

where $\tilde{v}\in B(J(\tilde{x},\tilde{\beta}))$. Applying the monotonicity of B, we obtain

$$\|\tilde{x}-J(\tilde{x},\tilde{\beta})\|^2\leq\delta\|\tilde{x}-J(\tilde{x},\tilde{\beta})\|^2,$$

implying that $\|\tilde{x}-J(\tilde{x},\tilde{\beta})\|\leq 0$. Thus, $\tilde{x}=J(\tilde{x},\tilde{\beta})$ and hence, $\tilde{x}\in S^*$. □

4.2 Convergence of Algorithm 2

In this subsection all results are referent to Algorithm 2, i.e., with Iterative Step 2 as

$$x^{k+1} = \mathcal{F}_2(x^k) = \mathcal{P}_{X\cap H_k}(x^k).$$

Proposition 4.5. *If Algorithm 2 stops, then $x^k \in S^*$.*

Proof. If $x^{k+1} = \mathcal{P}_{X\cap H_k}(x^k) = x^k$ then $x^k \in X \cap H_k$ and by Proposition 4.2, $x^k \in S^* \cap X$. □

From now on, assume that Algorithm 2 does not stop.

Proposition 4.6. *The sequence $\{x^k\}$ is Féjer convergent to $S^* \cap X$. Moreover, it is bounded and*

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

Proof. Take $x^* \in S^* \cap X$. By Lemma 4.1, $x^* \in H_k \cap X$, for all k. Then

$$\|x^{k+1} - x^*\|^2 - \|x^k - x^*\|^2 + \|x^{k+1} - x^k\|^2 = 2\langle x^* - x^{k+1}, x^k - x^{k+1}\rangle \le 0,$$

using Proposition 2.1(ii) and (13) in the last inequality,

$$\|x^{k+1} - x^*\|^2 \le \|x^k - x^*\|^2 - \|x^{k+1} - x^k\|^2. \tag{29}$$

The above inequality implies that $\{x^k\}$ is Féjer convergent to $S^* \cap X$. Hence by Proposition 2.4(i) and (ii), $\{x^k\}$ is bounded and thus $\{\|x^k - x^*\|\}$ is a convergent sequence. Taking limits in (29), we get

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

□

The next proposition shows a relation between the projection steps in Algorithm 1 and 2. This fact has a geometry interpretation, since the projection of Algorithm 2 is done over a small set, improving the convergence of Algorithm 1. Note that this can be reduce the number of iterations, avoiding possible zigzagging of Algorithm 1.

Proposition 4.7. *Let $\{x^k\}$ the sequence generated by Algorithm 2. Then,*

i) $x^{k+1} = \mathcal{P}_{X\cap H_k}(\mathcal{P}_{H_k}(x^k))$.

ii) $\lim_{k\to\infty}\langle A(\bar{x}^k) + \bar{u}^k, x^k - \bar{x}^k\rangle = 0$.

Proof. (i) Fix any $y \in X \cap H_k$. Since $x^k \in X$ but $x^k \notin H_k$ by Proposition 4.2, there exists $\gamma \in [0,1]$, such that $\tilde{x} = \gamma x^k + (1-\gamma)y \in X \cap \partial H_k$, where $\partial H_k := \big\{x \in \mathbb{R}^n : \langle A(\bar{x}^k) + \bar{u}^k, x - \bar{x}^k\rangle = 0\big\}$. Hence,

$$\begin{aligned}
\|y - \mathcal{P}_{H_k}(x^k)\|^2 &\ge (1-\gamma)^2\|y - \mathcal{P}_{H_k}(x^k)\|^2\\
&= \|\tilde{x} - \gamma x^k - (1-\gamma)\mathcal{P}_{H_k}(x^k)\|^2\\
&= \|\tilde{x} - \mathcal{P}_{H_k}(x^k)\|^2 + \gamma^2\|x^k - \mathcal{P}_{H_k}(x^k)\|^2 - 2\gamma\langle \tilde{x} - \mathcal{P}_{H_k}(x^k), x^k - \mathcal{P}_{H_k}(x^k)\rangle\\
&\ge \|\tilde{x} - \mathcal{P}_{H_k}(x^k)\|^2,
\end{aligned} \tag{30}$$

where the last inequality follows from Proposition 2.1(ii), applied with $X = H_k$, $x = x^k$ and $z = \tilde{x} \in H_k$. Furthermore, we have

$$\begin{aligned}\|\tilde{x} - \mathcal{P}_{H_k}(x^k)\|^2 &= \|\tilde{x} - x^k\|^2 - \|x^k - \mathcal{P}_{H_k}(x^k)\|^2 \\ &\geq \|x^{k+1} - x^k\|^2 - \|x^k - \mathcal{P}_{H_k}(x^k)\|^2 \\ &= \|x^{k+1} - \mathcal{P}_{H_k}(x^k)\|^2, \end{aligned} \tag{31}$$

where the first equality follows by $\mathcal{P}_{H_k}(x^k) = \mathcal{P}_{\partial H_k}(x^k)$, $\tilde{x} \in \partial H_k$ and Pythagoras's Theorem, using the fact that $\tilde{x} \in X \cap H_k$ and $x^{k+1} = \mathcal{P}_{X \cap H_k}(x^k)$ in the first inequality, and Pythagoras's Theorem again in the last equality. Combining (30) and (31), we obtain

$$\|y - \mathcal{P}_{H_k}(x^k)\| \geq \|x^{k+1} - \mathcal{P}_{H_k}(x^k)\|,$$

for all $y \in X \cap H_k$. Hence, $x^{k+1} = \mathcal{P}_{X \cap H_k}(\mathcal{P}_{H_k}(x^k))$.

(ii) Take $x^* \in X \cap S^*$. By item (i), (13) and Proposition 2.1(i), we have

$$\begin{aligned}\|x^{k+1} - x^*\|^2 &\leq \|\mathcal{P}_{H_k}(x^k) - \mathcal{P}_{H_k}(x^*)\|^2 = \|\mathcal{P}_{H_k}(x^k) - \mathcal{P}_{H_k}(x^*)\|^2 \\ &\leq \|x^k - x^*\|^2 - \|\mathcal{P}_{H_k}(x^k) - x^k\|^2. \end{aligned}$$

The proof is similar to the proof of Proposition 4.4(iii).

□

Finally we present the convergence result for Algorithm 2.

Theorem 4.2. *The sequence $\{x^k\}$ converges to some point belonging to $S^* \cap X$.*

Proof. Repeat the proof of Theorem 4.1. □

4.3 Convergence of Algorithm 3

In this subsection all results are referent to Algorithm 3, i.e., with Iterative Step 2 as

$$x^{k+1} = \mathcal{F}_3(x^k) = \mathcal{P}_{X \cap H_k \cap W_k}(x^0).$$

Proposition 4.8. *If Algorithm 3 stops, then $x^k \in S^* \cap X$.*

Proof. If Stop Criteria 2 is satisfied then, $x^{k+1} = \mathcal{P}_{X \cap H_k \cap W_k}(x^0) = x^k$. So, $x^k \in X \cap H_k \cap W_k \subset X \cap H_k$ and finally using Proposition 4.2, $x^k \in S^* \cap X$. □

From now on we assume that Algorithm 3 does not stop. Observe that, in virtue of their definitions, W_k and H_k are convex and closed halfspaces, for each k. Therefore $X \cap H_k \cap W_k$ is a convex and closed set. So, if $X \cap H_k \cap W_k$ is nonempty, then the next iterate, x^{k+1}, is well-defined. The following lemma guarantees this fact.

Lemma 4.2. *$S^* \cap X \subset H_k \cap W_k$, for all k.*

Proof. We proceed by induction. By definition, $S^* \cap X \neq \emptyset$. By Lemma 4.1, $S^* \cap X \subset H_k$, for all k. For $k = 0$, as $W_0 = \mathbb{R}^n$, $S^* \cap X \subset H_0 \cap W_0$.

Assume that $S^* \cap X \subset H_\ell \cap W_\ell$, for $\ell \leq k$. Henceforth, $x^{k+1} = P_{X \cap H_k \cap W_k}(x^0)$ is well-defined. Then, by Lemma 4.1, we have

$$\langle x^* - x^{k+1}\, , \, x^0 - x^{k+1} \rangle = \langle x^* - P_{C \cap H_k \cap W_k}(x^0)\, , \, x^0 - P_{X \cap H_k \cap W_k}(x^0) \rangle \leq 0, \tag{32}$$

for all $x^* \in S^* \cap X$. The inequality follows by the induction hypothesis. Now, (32) implies that $x^* \in W_{k+1}$ and hence, $S^* \cap X \subset H_{k+1} \cap W_{k+1}$. □

The above lemma shows that the set $X \cap H_k \cap W_k$ is nonempty and in consequence the projection step, given in (14), is well-defined.

Corollary 4.2. *Algorithm 3 is well-defined.*

Proof. By Lemma 4.2 , $S^* \cap X \subset H_k \cap W_k$, for all k. Then, given x^0, the sequence $\{x^k\}$ is computable. □

Before proving the convergence of the sequence, we study its boundedness.The next lemma shows that the sequence remains in a ball determined by the initial point.

Lemma 4.3. *The sequence $\{x^k\}$ is bounded. Furthermore,*

$$\{x^k\} \subset B\left[\frac{1}{2}(x^0 + \bar{x}), \frac{1}{2}\rho\right] \cap X,$$

where $\bar{x} = P_{S^* \cap X}(x^0)$ and $\rho = \operatorname{dist}(x^0, S^* \cap X)$.

Proof. $S^* \cap X \subset H_k \cap W_k$ follows from Lemma 4.2. Moreover, from (14), we obtain that

$$\|x^{k+1} - x^0\| \leq \|z - x^0\|, \tag{33}$$

for all k and all $z \in S^* \cap X$. Henceforth, taking $z = \bar{x}$ in (33),

$$\|x^{k+1} - x^0\| \leq \|\bar{x} - x^0\| = \rho, \tag{34}$$

for all k. Thus, $\{x^k\}$ is bounded. Define $z^k = x^k - \frac{1}{2}(x^0 + \bar{x})$ and $\bar{z} = \bar{x} - \frac{1}{2}(x^0 + \bar{x})$. It follows from the fact $\bar{x} \in W_{k+1}$, that

$$\begin{aligned}
0 &\geq 2\langle \bar{x} - x^{k+1}, x^0 - x^{k+1} \rangle \\
&= 2\left\langle \bar{z} + \frac{1}{2}(x^0 + \bar{x}) - z^{k+1} - \frac{1}{2}(x^0 + \bar{x}), z^0 + \frac{1}{2}(x^0 + \bar{x}) - z^{k+1} - \frac{1}{2}(x^0 + \bar{x}) \right\rangle \\
&= 2\left\langle \bar{z} - z^{k+1}, z^0 - z^{k+1} \right\rangle = \left\langle \bar{z} - z^{k+1}, -\bar{z} - z^{k+1} \right\rangle = \|z^{k+1}\|^2 - \|\bar{z}\|^2,
\end{aligned}$$

where we have used that $\bar{z} = -z^0$ in the third equality. So,

$$\left\| x^{k+1} - \frac{x^0 + \bar{x}}{2} \right\| \leq \left\| \bar{x} - \frac{x^0 + \bar{x}}{2} \right\| = \frac{\rho}{2},$$

for all k. Now, the result follows from the feasibility of $\{x^k\}$, which, in turn, is a consequence of (14). □

Now, we focus on the properties of the accumulation points.

Lemma 4.4. *All accumulation points of $\{x^k\}$ belong to $S^* \cap X$.*

Proof. Since $x^{k+1} \in W_k$,

$$0 \geq 2\langle x^{k+1} - x^k, x^0 - x^k\rangle = \|x^{k+1} - x^k\|^2 - \|x^{k+1} - x^0\|^2 + \|x^k - x^0\|^2.$$

Equivalently

$$0 \leq \|x^{k+1} - x^k\|^2 \leq \|x^{k+1} - x^0\|^2 - \|x^k - x^0\|^2,$$

establishing that the sequence $\{\|x^k - x^0\|\}$ is monotone and nondecreasing. From Lemma 4.3, we get that $\{\|x^k - x^0\|\}$ is bounded, and thus, convergent. Therefore,

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0. \tag{35}$$

Since $x^{k+1} \in H_k$, we get

$$\langle A(\bar{x}^k) + \bar{u}^k, x^{k+1} - \bar{x}^k\rangle \leq 0, \tag{36}$$

with $\bar{u}^k$ and $\bar{x}^k$ as (9) and (10).
Using (10) and (36), we have

$$\langle A(\bar{x}^k) + \bar{u}^k, x^{k+1} - x^k\rangle + \alpha_k\big\langle A(\bar{x}^k) + \bar{u}^k, x^k - J(x^k, \beta_k)\big\rangle \leq 0.$$

Combining the above inequality with Corollary 4.1, we get

$$\langle A(\bar{x}^k) + \bar{u}^k, x^{k+1} - x^k\rangle + \frac{\alpha_k\delta}{\hat{\beta}}\|x^k - J(x^k, \beta_k)\|^2 \leq 0. \tag{37}$$

Choosing a subsequence $\{i_k\}$ such that the subsequences $\{x^{i_k}\}$, $\{\beta_{i_k}\}$ and $\{\bar{u}^{i_k}\}$ converge to $\tilde{x}$, $\tilde{\beta}$ and $\tilde{u}$ respectively. This is possible by the boundedness of $\{\bar{u}^k\}$, by hypothesis on B, bounded of $\{x^k\}$ and $\{\beta_k\}$. Taking limits in (37), we have

$$\lim_{k\to\infty} \alpha_{i_k}\|x^{i_k} - J(x^{i_k}, \beta_{i_k})\|^2 = 0. \tag{38}$$

Now we consider two cases, $\lim_{k\to\infty} \alpha_{i_k} = 0$ or $\lim_{k\to\infty} \alpha_{i_k} \neq 0$.

a) $\lim_{k\to\infty} \alpha_{i_k} \neq 0$, i.e., $x^{i_k} \geq \tilde{\alpha}$ for all k and some $\tilde{\alpha} > 0$. By (38),

$$\lim_{k\to\infty} \|x^{i_k} - J(x^{i_k}, \beta_{i_k})\|^2 = 0.$$

By continuity of J, we have $\tilde{x} = J(\tilde{x}, \tilde{\beta})$ and hence by Proposition 2.3, $\tilde{x} \in S^*$.

b) $\lim_{k\to\infty} \alpha_{i_k} = 0$, then $\lim_{k\to\infty} \frac{\alpha_{i_k}}{\theta} = 0$. It follows in the same, manner as in the proof of Theorem 4.1(b). □

Finally, we are ready to prove the convergence of the sequence $\{x^k\}$ generated by Algorithm 3, to the solution closest to x^0.

Theorem 4.3. *Define $\bar{x} = P_{S^*\cap X}(x^0)$. Then, $\{x^k\}$ converges to $\bar{x}$.*

Proof. By Lemma 4.3, $\{x^k\} \subset B\left[\frac{1}{2}(x^0 + \bar{x}), \frac{1}{2}\rho\right] \cap X$, so it is bounded. Let $\{x^{i_k}\}$ be a convergent subsequence of $\{x^k\}$, and let $\hat{x}$ be its limit. Evidently $\hat{x} \in B\left[\frac{1}{2}(x^0 + \bar{x}), \frac{1}{2}\rho\right] \cap X$. Furthermore, by Lemma 4.4, $\hat{x} \in S^* \cap X$. Then,

$$\hat{x} \in S^* \cap X \cap B\left[\frac{1}{2}(x^0 + \bar{x}), \frac{1}{2}\rho\right] = \{\bar{x}\},$$

implying $\hat{x} = \bar{x}$, i.e., $\bar{x}$ is the unique limit point of $\{x^k\}$. Hence, $\{x^k\}$ converges to $\bar{x} \in S^* \cap X$. □

5 Final Remarks

When $B = N_X$ problem (1) becomes the well-study variational inequality problem. In this case the proposed algorithms (1, 2 and 3) are related with the algorithms in [12, 20, 3]; see Proposition 2.1(iii). In this section, we present an example showing that there exist advantage to take, inside to the inner loop, a non-zero element, u_j^k, belonging to N_X in the application of Algorithm 1.

Example 5.1. *Consider $A : \mathbb{R}^2 \to \mathbb{R}^2$ defined as $A(x, y) = (-y, x)$ and $B : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ as $B = N_C$ where C is the ball centered in $(0, 0)$ and radius 1, i.e.,*

$$N_C(x, y) = \begin{cases} 0 & , \quad x^2 + y^2 < 1 \\ \mathbb{R}_+(x, y) & , \quad x^2 + y^2 = 1. \end{cases}$$

Clearly, A and B are monotone and the unique solution of the problem (1) is $x^* = (0, 0)$. Set $\beta_k = 1$ for all k, $\delta = \frac{1}{2}$ and $X = C$. We begin the Algorithm 1 with $x^0 = (a, b)$ such that $a^2 + b^2 = 1$. Then,

$$J(x^0) = \frac{\sqrt{2}}{2}\Big(a + b, b - a\Big).$$

Beginning the inner loop with $j = 0$, and take$u_0^0 \in N_C\big(J(x^0)\big)$, i.e,

$$u_0^0 = \frac{\sqrt{2}\,r}{2}\Big(a + b, b - a\Big),$$

where $r \geq 0$. For all $r \leq \sqrt{2}$, $J(x^0)$ and u_0^0 satisfies the Armijo condition in (7). Then, $j(0) = 0$, $\bar{x}^0 = J(x^0)$ and

$$\bar{u}_0^0 = \frac{\sqrt{2}\,r}{2}\Big(a + b, b - a\Big).$$

Thus,

$$x^1 = \mathcal{P}_C\Big((a, b) - \frac{1 + (1 - \sqrt{2})r}{2(r^2 + 1)}\big(a - b + r(a + b), a + b + r(b - a)\big)\Big) \tag{39}$$

$$= (a, b) - \frac{1 + (1 - \sqrt{2})r}{2(r^2 + 1)}\big(a - b + r(a + b), a + b + r(b - a)\big), \tag{40}$$

for all $0 \leq r \leq \sqrt{2}$. Therefore,

$$D(r) = \operatorname{dist}^2(x^*; x^1) = \|x^1\|^2 = \frac{3r^2 - 2r + 1}{2(r^2 + 1)},$$

for all $0 \leq r \leq \sqrt{2}$, which attains the unique minimum in $r = \sqrt{2} - 1$. Concluding that it is better to take $u_0^0 = \frac{\sqrt{2}-1}{\sqrt{2}}\Big(a + b, b - a\Big)$ in order to obtain the next point, x^1, nearest to the unique solution, x^*, of problem (1).

ACKNOWLEDGMENTS

Research for this paper was partially supported by PROCAD-nf - UFG/UnB/IMPA and by project CAPES-MES-CUBA 226/2012 “Modelos de Otimização e Aplicações”. The authors would like to thank to CNPq. We also are very grateful to Prof. Ole Peter Smith for correcting the manuscript. The second author was supported by his scholarship for his doctoral studies, granted by CAPES.

References

[1] Bauschke, H.H., Borwein, J.M. On projection algorithms for solving convex feasibility problems. *SIAM Review* **38** (1996) 367-426.

[2] Bauschke, H. H., Burke, J. V., Deutsch, F. R., Hundal, H. S., Vanderwerff, J. D. A new proximal point iteration that converges weakly but not in norm. *Proceedings of the American Mathematical Society* **133** (2005) 1829-1835.

[3] Bello Cruz, J.Y., Iusem, A.N. A strongly convergent direct method for monotone variational inequalities in Hilbert spaces. *Numerical Functional Analysis and Optimization* **30** (2009) 23-36.

[4] Bello Cruz, J.Y., Iusem, A.N. Convergence of direct methods for paramonotone variational inequalities. *Computation Optimization and Applications* **46** (2010) 247-263.

[5] Browder, F.E. Convergence theorems for sequences of nonlinear operators in Banach spaces. *Mathematische Zeitschrift* **100** (1967) 201-225.

[6] Burachik, R., Graña Drummond, L.M., Iusem, A.N., Svaiter, B.F. Full convergence of the steepest descent method with inexact line searches. *Optimization* **32** (1995) 137-146.

[7] Burachik, R.S., Iusem, A.N. *Set-Valued Mappings and Enlargements of Monotone Operators*. Springer, Berlin (2008).

[8] Eckstein, J. *Splitting Methods for Monotone Operators, with Applications to Parallel Optimization.* PhD thesis, Massachusetts Institute of Techonology, Cambridge, MA, 1989. Report LIDS-TH-1877, Laboratory for Information and Decision Systems, M.I.T.

[9] Eckstein, J., Svaiter, B.F. A family of projective splitting methods for the sum of two maximal monotone operators. *Mathematical Programming* **111** (2008) 173-199.

[10] Facchinei, F., Pang, J.S. *Finite-dimensional Variational Inequalities and Complementarity Problems.* Springer, Berlin (2003).

[11] Hartman, P., Stampacchia, G. On some non-linear elliptic differential-functional equations. *Acta Mathematica* **115** (1966) 271-310.

[12] Iusem, A.N., Svaiter, B.F. A variant of Korpelevich's method for variational inequalities with a new search strategy. *Optimization* **42** (1997) 309–321.

[13] Iusem, A.N., Svaiter, B.F., Teboulle, M. Entropy-like proximal methods in convex programming. *Mathematics of Operations Research* **19** (1994) 790-814.

[14] Kinderlehrer, D., Stampacchia, G. *An Introduction to Variational Inequalities and Their Applications*. Academic Press, New York (1980).

[15] Lions, P.L., Mercier, B. Splitting algorithms for the sum of two nonlinear operators. *SIAM of Journal Numerical Analysis.* **16** (1979) 964-979.

[16] Minty, G. On the maximal domain of a "monotone" function. *Michigan Mathematical Journal* **8** (1961) 135-137.

[17] Minty, G. Monotone (nonlinear) operators in Hilbert Space. *Duke Mathetematical Journal* **29** (1962) 341-346.

[18] Passty, G.B. Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *Journal of Mathematical Analysis and Applications* **72** (1979) 383-390.

[19] Tseng, P. A modified forward-backward splitting method for maximal monotone mappings. *SIAM on Journal Control Optimization* **38** (2000) 431-446.

[20] Solodov, M.V., Svaiter, B.F. A new projection method for variational inequality problems. *SIAM Journal on Control and Optimization* **37** (1999) 765-776.