On the Weak and Strong Convergence of a Conceptual Algorithm for Solving Three Operator Monotone Inclusions

Yunier Bello-Cruz*   Oday Hazaimah†

April 11, 2021

Abstract

In this paper, a conceptual algorithm modifying the forward-backward-half-forward (FBHF) splitting method for solving three operator monotone inclusion problems is investigated. The FBHF splitting method adjusts and improves Tseng's forward-backward-forward (FBF) splitting method when the inclusion problem has a third-part operator that is cocoercive. The FBHF method recovers the FBF iteration (when this aforementioned part is zero), and it also works without using the widely used Lipschitz continuity assumption. The conceptual algorithm proposed in this paper also has those advantages, and it derives two variants (called Method 1 and Method 2) by choosing different projection (forward) steps. Both proposed methods also work efficiently without assuming the Lipschitz continuity and without directly using the cocoercive constant. Moreover, they have the following desired features: (i) very general iterations are derived for both methods, recovering the FBF and the FBHF iterations and allowing possibly larger stepsizes if the projection steps are over-relaxing; and (ii) strong convergence to the best approximation solution of the problem is proved for Method 2. To the best of our knowledge, this is the first time that an FBF-type method converges strongly for finding the best approximation solution of the three operator monotone inclusion.

2010 Mathematics Subject Classification: 47H05, 47J22, 49J52, 65K15, 90C25

Keywords: Best approximation solutions, Forward-backward-forward splitting method, Operator splitting algorithms, Separating hyperplanes, Strong convergence.

1   Introduction

In this work, we present a conceptual algorithm for solving monotone inclusion problems involving the sum of three maximal monotone operators in a real Hilbert space $\mathcal{H}$. The general formulation of the *inclusion problem* is as follows:

$$\text{Find} \quad x \in \mathcal{H} \quad \text{such that } 0 \in (A + B)x, \tag{1}$$

where $A : \mathcal{H} \to \mathcal{H}$ (single-valued) and $B : \text{dom}(B) \subseteq \mathcal{H} \rightrightarrows \mathcal{H}$ (set-valued) are maximal monotone operators. Throughout this paper we assume that the solution set of this problem, $\text{zer}(A + B)$,

---

*Department of Mathematical Sciences, Northern Illinois University. Watson Hall 366, DeKalb, IL, USA - 60115. E-mail: `yunierbello@niu.edu`.

†Department of Mathematical Sciences, Northern Illinois University. Watson Hall 366, DeKalb, IL, USA - 60115. E-mail: `oday@niu.edu`.

is non-empty. Problem (1) appears in different fields of applied mathematics and optimization including signal processing, numerous important structured optimization, composite convex optimization, saddle point, and inverse problems; see, for instance, [12,14]. One of the most relevant setting that can be represented by a particular case of the inclusion problem (1) is the broadly-studied *variational inequality problem* (VIP)

$$\langle Ax, y - x \rangle \geq 0 \quad \text{for all} \quad y \in C. \tag{2}$$

The set C is a convex and closed subset of $\mathcal{H}$ and $\mathcal{H}$ is equipped with the inner product $\langle \cdot, \cdot \rangle$ and the induced Euclidean norm $\| \cdot \|$. This problem is a particular case of problem (1) by taking $B = \mathcal{N}_C$ the normal cone of C, i.e., find $x \in C$ such that $0 \in Ax + \mathcal{N}_C(x)$. A popular strategy to solve problem (2) is the so-called cutting plane (a.k.a. localization) idea which is based on finding a suitable hyperplane that separates the solution of the problem from the current iterate and then performs a metric projection step. This kind of idea is used by the famous *Extragradient method* and its variants for solving problem (2); see, for instance, [3,10,14].

Here we apply the cutting plane idea to perform the first phase of the proposed iteration, described below when the single value operator A inside of problem (1) is the sum of two parts. The considered iteration solves problem (1) when $A = A_1 + A_2$ such that A_1 is cocoercive and A_2 is maximally monotone. Moreover, it uses a novel backtracking procedure that allows larger stepsizes in general. Furthermore, the *forward* steps are special projection steps onto suitable halfspaces which only evaluate A_2. It is worth noting that this kind of modification was presented in [5] to improve Tseng's scheme in finite dimension for solving problem (1) (without considering the cocoercive part). In general, the scheme in [5] fails to keep the splitting structure of Tseng's splitting method and requires finding a uniformly bounded sequence in the image of the set-valued operator. A similar approach (using normals vectors) for solving problem (2) was presented in [6].

1.1 Splitting Iterations Description

We focus our attention on a class of schemes, called splitting methods, which only use each operator individually rather than evaluating their sum directly. We refer to *forward step* when the single-valued operator is evaluated, and *backward step* when the resolvent operator of the set-valued operator is computed. Recalling that the *resolvent operator* of B is the full domain single-valued operator in $\mathcal{H}$ given by $J_B := (I + B)^{-1}$ where $I \colon \mathcal{H} \to \mathcal{H}$ denotes the identity operator.

One of the most important classical splitting methods to find a zero of the sum $A + B$ is the so-called *forward-backward* (FB) splitting method introduced in [9] which is given as follows:

$$x^{k+1} := J_{\alpha_k B}(x^k - \alpha_k A x^k), \tag{3}$$

where $\alpha_k > 0$ for all $k \in \mathbb{N}$. This iteration converges weakly to a point in the solution of problem (1), $\mathrm{zer}(A + B)$, when the inverse of A is β-strongly monotone (or equivalently A being β-cocoercive), i.e.,

$$\forall x, y \in \mathcal{H}, \qquad \langle Ax - Ay, x - y \rangle \geq \beta \| Ax - Ay \|^2,$$

where $\alpha_k \leq \beta$ for all $k \in \mathbb{N}$ and $\liminf_{k \to \infty} \alpha_k > 0$; see, for instance, [15, 17]. Unfortunately, there is no chance to relax the cocoercivity assumption on A to plain monotonicity and still prove convergence for the FB iteration given in (3). For example, if we set A as the $\pi/2$-rotation operator which is monotone and $B = 0$, iteration (3) moves away from zero (the unique solution) for any positive stepsize and starting at any point. Moreover, iteration (3) converges only weakly and the

strong convergence could fail in general; see [13]. It is worth emphasizing that the cocoercivity assumption of an operator is a strictly stronger property than Lipschitz continuity. Recalling that, for some $L \geq 0$, A is L-Lipschitz if

$$\forall x, y \in \mathcal{H}, \qquad \|Ax - Ay\| \leq L\|x - y\|.$$

Note that β-cocoercive operators are monotone and $1/\beta$-Lipschitz continuous, but the converse does not hold in general, i.e., There exist monotone and Lipschitz continuous operators that are not cocoercive. Although, for gradients of lower semicontinuous, proper and convex functions, the cocoercivity is equivalent to the global Lipschitz continuity assumption. This nice and surprising fact is strongly used in the convergence analysis of the FB iteration (3) for solving the sum of two convex function (problem (1) with $A = \nabla f$ and $B = \partial g$), is known as the *Baillon-Haddad Theorem*; see Corollary 18.16 of [2]. Another useful feature that *Baillon-Haddad Theorem* assures is that ∇f is firmly nonexpansive if and only if it is a nonexpansive map. Recalling that an operator A is non expansive if it is Lipschitz with constant 1, and A is said to be firmly nonexpansiveness if

$$\forall x, y \in \mathcal{H}, \qquad \|Ax - Ay\|^2 \leq \|x - y\|^2 - \|(x - Ax) - (y - Ay)\|^2.$$

In order to relax the cocoercivity assumption, Tseng [19] proposed a modification of the FB splitting method, known as the *forward-backward-forward* (FBF) splitting method, which usually requires the L-Lipschitz continuity assumption of A and an additional forward step. The FBF splitting iteration is:

$$\bar{x}^k := J_{\alpha_k B}(x^k - \alpha_k A x^k) \tag{4}$$
$$x^{k+1} := \bar{x}^k - \alpha_k [A\bar{x}^k - A x^k]. \tag{5}$$

This iteration converges weakly to a point in $\text{zer}(A + B)$, if:

(i) the operator A is monotone and L-Lipschitz and L is available by taking $\alpha_k \leq 1/L$ for all $k \in \mathbb{N}$ and $\liminf_{k \to \infty} \alpha_k > 0$; or

(ii) the operator A is locally uniformly continuous on $\text{dom}(B)$ and the function $x \mapsto \min_{w \in (A+B)x} \|w\|$ is locally bounded on $\text{dom}(B)$ by choosing α_k to be the largest $\alpha \in \{\sigma, \sigma\theta, \sigma\theta^2, \ldots\}$ with $\sigma > 0$ and $\theta, \delta \in]0, 1[$ satisfying

$$\alpha \|A\bar{x}^k - A x^k\| \leq \delta \|x^k - \bar{x}^k\|. \tag{6}$$

It is worth noting that there are relatively few effective alternatives to Tseng's FBF algorithm (4)-(5) for solving inclusions in the form of problem (1) [11, 15, 20].

In this paper, we assume that the single-valued operator A can be split as the sum of A_1 (β-cocoercive operator) and A_2 (monotone operator). Hence, problem (1) takes the following form:

$$\text{Find} \quad x \in \mathcal{H} \quad \text{such that} \quad 0 \in (A_1 + A_2 + B)x. \tag{7}$$

For convenience, we also denote $\text{zer}(A_1 + A_2 + B)$ the solution set of this problem, which from now on is assumed to be nonempty. Moreover, the FBF iteration (4)-(5) can be modified to the following *forward-backward-half-forward* (FBHF) splitting iteration proposed in [8] as follows:

$$\bar{x}^k := J_{\alpha_k B}(x^k - \alpha_k A x^k) \tag{8}$$
$$x^{k+1} := \bar{x}^k - \alpha_k [A_2\bar{x}^k - A_2 x^k]. \tag{9}$$

The weak convergence occurs when:

(i) the operator A_1 is β-cocoercive, A_2 is L-Lipschitz and β and L are available by taking $\alpha_k \in \,]\eta, \min\{\beta, 1/(2L)\}[$ with $\eta > 0$ for all $k \in \mathbb{N}$; or

(ii) the operator A_1 is β-cocoercive and β is available and A_2 is uniformly continuous in weakly compact subset of $\mathrm{dom}(B)$ by choosing α_k as the largest $\alpha \in \{2\beta\epsilon\theta, 2\beta\epsilon\theta^2, \ldots\}$ with $\epsilon, \theta \in \,]0,1[$ satisfying (6) with $\delta \in]0, \sqrt{1-\epsilon}[$ and $A = A_2$.

It is worth mentioning that this last backtracking strategy, described in (ii), to find α_k uses β and allows smaller stepsizes than β. The main difference between iterations (4)-(5) and (8)-(9) yields in the last forward steps (5) and (9), i.e., the operator A_1 is not evaluated in (9) (remind that $A = A_1 + A_2$). This is possible because of the cocoercive assumption, which allows the one-step FB splitting iteration (3) to be convergent for certain values of α_k related to β. Actually, if $A_2 = 0$, the FB splitting iteration (3) is recovered by the FBHF splitting iteration (8)-(9). In this context, with $A_1 = \nabla f$ and $B = \partial g$ with f, g convex functions, was proposed by using the backtracking procedure (6) a weakly convergent proximal gradient method in [7] without any boundedness of the image of ∂g or Lipschitz continuity assumption of ∇f. Note further that, in the particular case that $A_1 = 0$, the problem (7) (with $A = A_2$) becomes problem (1) and the FBHF iteration (8)-(9) coincides with the FBF iteration (4)-(5). Motivations and applications for such kind of splitting structure giving in problem (7) can be found in [2,8]. For example, Briceño-Arias and Davis in [8] applied the algorithm to primal-dual composite monotone inclusions with non-self-adjoint linear operators. In nonsmooth empirical risk minimizations in machine learning, one minimizes finite approximations of expected value for the loss function and constraints. For more, we refer the reader to [8] and the references therein.

The proposed conceptual algorithm here modifies and extends the FBHF iteration (8)-(9) by using two phases: (I) in the spirit of cutting plane methods, a backtracking search is performed to construct a suitable separating hyperplane; and (II) two special projection (forward) steps onto suitable separating hyperplanes deliver two different methods. Convergence analysis of both methods is presented without the Lipschitz continuity assumption and the knowledge of the cocoercive constant of A_1. In addition, if A_2 is L_2-Lipschitz, the proposed backtracking strategy may allow larger stepsizes than $1/L_2$. Furthermore, the first variant relies on a very general iteration which recovers the FBHF iteration (8)-(9) as a special case. The second variant has the following desirable properties: (i) the generated sequence is entirely contained in a ball with a diameter equal to the distance between the initial and the solution set; and (ii) the whole sequence converges strongly to the solution of the problem that lies closest to the initial point. Emphasizing that only weak convergence is known for the FBHF splitting method.

The presentation of this paper is as follows. In the next section, we provide some relevant background and useful facts that will be used throughout this paper. The proposed conceptual algorithm is presented in Section 3 and its two versions, called **Method 1** and **Method 2** are described. Section 4 contains the convergence analysis of these methods. Section 5 gives some concluding remarks.

2   Preliminaries

In this section, we present some definitions and conventional results needed for the convergence analysis of the proposed methods. The notation and results we discuss are standard, and interested readers can find further information in [2].

Throughout this paper, we write $p := q$ to indicate that p is defined to be equal to q. Let $\mathcal{H}$ be a real Hilbert space equipped with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \| := \sqrt{\langle \cdot, \cdot \rangle}$. We write $\mathbb{N}$ for the nonnegative integers $\{0, 1, 2, \ldots\}$. The closed ball centered at $x \in \mathcal{H}$ with radius $\gamma > 0$ will be denoted by $\mathbb{B}[x; \gamma] := \{y \in \mathcal{H} \mid \|y - x\| \leq \gamma\}$. Let $T : \mathcal{H} \rightrightarrows \mathcal{H}$ be a set-valued operator and its domain be denoted by $\operatorname{dom}(T) := \{x \in \mathcal{H} \mid T(x) \neq \varnothing\}$ and, for simplicity, we usually write $Tx := T(x)$. Define the graph of T by $\operatorname{Gph}(T) := \{(x, u) \in \operatorname{dom}(T) \times \mathcal{H} \mid u \in Tx\}$. We say that T is monotone if

$$\forall (x, u), (y, v) \in \operatorname{Gph}(T), \qquad \langle x - y, u - v \rangle \geq 0,$$

and it is maximally monotone if there exists no monotone operator T' such that $\operatorname{Gph}(T')$ properly contains $\operatorname{Gph}(T)$.

In the following, we state some important facts and preliminary results on set-valued mappings that are maximally monotone and addresses their graphs properties.

Lemma 2.1 (Proposition 20.31 and Proposition 20.33 of [2]) *Let $T : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximal monotone operator and let $x \in \mathcal{H}$. Then the following hold:*

(i) *Tx is closed and convex;*

(ii) *For every sequence $(x^k, u^k)_{k \in \mathbb{N}} \subset \operatorname{Gph}(T)$ and every point $(x, u) \in \operatorname{dom}(T) \times \mathcal{H}$, if $x^k \rightharpoonup x$ and $u^k \to u$, then $(x, u) \in \operatorname{Gph}(T)$, i.e. $\operatorname{Gph}(T)$ is sequentially closed in the weak-strong topology;*

(iii) *$\operatorname{Gph}(T)$ is closed in $\mathcal{H} \times \mathcal{H}$ in the strong topology.*

Note that the graph of a maximal monotone operator, in general, need not be sequentially closed in the weak topology of $\mathcal{H} \times \mathcal{H}$.

Proposition 2.2 (Theorem 4 of [16]) *Let $T : \operatorname{dom}(T) \subseteq \mathcal{H} \rightrightarrows \mathcal{H}$ be a set-valued and maximal monotone operator. If $\alpha > 0$ then the resolvent operator $J_{\alpha T} := (I + \alpha T)^{-1} : \mathcal{H} \to \operatorname{dom}(T)$ is a full domain, single-valued and firmly nonexpansive operator, i.e.,*

$$\forall x, y \in \mathcal{H}, \qquad \|J_{\alpha T}(x) - J_{\alpha T}(y)\|^2 + \|(I - J_{\alpha T})(x) - (I - J_{\alpha T})(y)\|^2 \leq \|x - y\|^2.$$

The inverse of T is the set-valued operator defined by $T^{-1} : u \mapsto \{x \in \mathcal{H} \mid u \in T(x)\}$. The zero set of T is $\operatorname{zer}(T) := T^{-1}(0)$. If $T = A + B$ then the solution of problem (1) is

$$\operatorname{zer}(A + B) = (A + B)^{-1}(0) = \{x \in \mathcal{H} \mid 0 \in (A + B)x\}.$$

The next result characterizes the above solution set as the fixed points of the forward-backward operator.

Proposition 2.3 (Proposition 23.28 of [2]) *Let $\alpha > 0$, and $A : \mathcal{H} \to \mathcal{H}$ and $B : \operatorname{dom}(B) \subseteq \mathcal{H} \rightrightarrows \mathcal{H}$ be two maximal monotone operators. Then,*

$$x = (I + \alpha B)^{-1}(x - \alpha Ax) \quad \text{if and only if} \quad x \in \operatorname{zer}(A + B).$$

Note further that, for all $x, y \in \mathcal{H}$,

$$y = (I + \alpha B)^{-1}(x - \alpha Ax) \quad \text{if and only if} \quad \frac{x - y}{\alpha} - Ax \in By. \tag{10}$$

Let C be a nonempty, convex and closed subset of $\mathcal{H}$, and define the normal cone operator with respect to a nonempty closed convex set $C \subseteq \mathcal{H}$ as

$$\mathcal{N}_C(x) := \begin{cases} \varnothing, & \text{if } x \notin C \\ \{y \in \mathcal{H} \mid \langle y, z - x \rangle \leq 0, \ \forall z \in C\}, & \text{if } x \in C. \end{cases}$$

Hence, the orthogonal projection of x onto C, $\Pi_C(x)$, is given by $\Pi_C(x) = J_{\mathcal{N}_C}(x) = (I + \mathcal{N}_C)^{-1}(x)$. Now, we state two well-known facts on orthogonal projections.

Proposition 2.4 (Theorem 3.16 and Proposition 4.16 of [2]) *Let C be nonempty closed convex subset of $\mathcal{H}$, and Π_C be the orthogonal projection onto C. For all $x, y \in \mathcal{H}$ and all $z \in C$ the following hold:*

(i) $\|\Pi_C(x) - \Pi_C(y)\|^2 \leq \|x - y\|^2 - \|(x - \Pi_C(x)) - (y - \Pi_C(y))\|^2$;

(ii) $\langle x - \Pi_C(x), z - \Pi_C(x) \rangle \leq 0$.

In the following, we present some useful formulae to describe the iterates of the proposed methods by introducing suitable hyperplanes and orthogonal projections onto these hyperplanes.

Proposition 2.5 (Proposition 28.19 of [2]) *Let $y, v \in \mathcal{H}$, $r \in \mathbb{R}$,*

$$T_{y,v}^r := \{x \in \mathcal{H} \mid \langle v, x - y \rangle \leq r\},$$

and

$$\Gamma_{z,x^0} := \{x \in \mathcal{H} \mid \langle x^0 - z, x - z \rangle \leq 0\}.$$

Then,

$$\Pi_{T_{y,v}^r}(w) = \begin{cases} w, & \text{if } w \in T_{y,v}^r \\ w - \dfrac{\langle v, w - y \rangle - r}{\|v\|^2} v, & \text{if } w \notin T_{y,v}^r. \end{cases}$$

Moreover,

(i) *if $x^0 - z$ is linearly dependent to v (or equivalently, $\|x^0 - z\|\|v\| = \langle x^0 - z, v \rangle$), $T_{y,v}^r \subset \Gamma_{z,x^0}$ and*

$$\Pi_{T_{y,v}^r \cap \Gamma_{z,x^0}}(x^0) = \Pi_{T_{y,v}^r}(x^0).$$

(ii) *if $x^0 - z$ is linearly independent to v (or equivalently, $\|x^0 - z\|\|v\| > \langle x^0 - z, v \rangle$),*

$$\Pi_{T_{y,v}^r \cap \Gamma_{z,x^0}}(x^0) = x^0 - \lambda_1 v - \lambda_2(x^0 - z),$$

where λ_1, λ_2 are solutions of the linear system:

$$\lambda_1 \|v\|^2 + \lambda_2 \langle v, x^0 - z \rangle = \langle v, x^0 - y \rangle - r$$
$$\lambda_1 \langle v, x^0 - z \rangle + \lambda_2 \|x^0 - z\|^2 = \langle x^0 - z, x^0 - z \rangle.$$

Now we define an important concept, the so-called Fejér monotonicity.

Definition 2.6 *Let S be a nonempty subset of $\mathcal{H}$. A sequence $(x^k)_{k \in \mathbb{N}} \subset \mathcal{H}$ is said to be Fejér monotone with respect to S, if and only if, for all $x \in S$ there exists $k_0 \in \mathbb{N}$, such that*

$$\|x^{k+1} - x\| \leq \|x^k - x\| \quad \text{for all} \quad k \geq k_0.$$

Useful properties on Fejér monotone sequences are the following.

Proposition 2.7 (Proposition 5.4 and Theorem 5.5 of [2]) *Let $(x^k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and let S be a non empty subset of $\mathcal{H}$. If $(x^k)_{k\in\mathbb{N}}$ is Fejér monotone with respect to S, then:*

(i) *The sequence $(x^k)_{k\in\mathbb{N}}$ is bounded;*

(ii) *The sequence $\left(\|x^k - x\|\right)_{k\in\mathbb{N}}$ is convergent for all $x \in S$;*

(iii) *If every weak accumulation point x^* of $(x^k)_{k\in\mathbb{N}}$ belongs to S, then $(x^k)_{k\in\mathbb{N}}$ converges weakly to x^*.*

3 The Conceptual Forward-Backward-Half-Forward Algorithm

The conceptual modification of the FBHF splitting algorithm uses the parameters $\theta, \delta \in (0,1)$. It is defined as follows:

Conceptual Algorithm.

Step 0. (Initialization): Take
$$x^0 \in \mathcal{H}, \quad \text{and} \quad \alpha_{-1} > 0.$$

Step 1. (Backtracking): Given x^k and α_{k-1} define
$$\bar{x}_j^k := J_{\alpha_{k-1}\theta^j B}(x^k - \alpha_{k-1}\theta^j A x^k). \tag{11}$$

Start the inner loop in j to compute $j(k)$ as the smallest $j \in \mathbb{N}$ such that
$$\alpha_k \langle A_2 x^k - A_2 \bar{x}_j^k, x^k - \bar{x}_j^k \rangle \leq \delta \|x^k - \bar{x}_j^k\|^2. \tag{12}$$

Step 2. (Forward step): Set $\alpha_k := \alpha_{k-1}\theta^{j(k)}$,
$$\bar{x}^k := \bar{x}_{j(k)}^k = J_{\alpha_k B}(x^k - \alpha_k A x^k) \tag{13}$$

and
$$x^{k+1} := \mathcal{F}(x^k, \bar{x}^k). \tag{14}$$

Stopping Criterion: If $x^{k+1} = x^k$ then stop.

We consider two projection variants of the **Conceptual Algorithm**, which are called **Method 1** and **Method 2** respectively. It will be used two different forward steps $\mathcal{F} = \mathcal{F}_1$ and $\mathcal{F} = \mathcal{F}_2$ on the projection steps in (14) as follows. Take any $\bar{\delta}$ such that $1 - \delta - \bar{\delta} > 0$ and define $r_k := \frac{\bar{\delta}}{\alpha_k}\|x^k - \bar{x}^k\|^2$,

$$T_k := \left\{ x \in \mathcal{H} \,\middle|\, \left\langle \frac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k), x - \bar{x}^k \right\rangle \leq r_k \right\} \tag{15}$$

and
$$\mathcal{F}_1(x^k, \bar{x}^k) := \Pi_{T_k}(x^k). \tag{16}$$

Moreover, set
$$\Gamma_k := \left\{ x \in \mathcal{H} \mid \langle x^0 - x^k, x - x^k \rangle \leq 0 \right\} \tag{17}$$

and

$$\mathcal{F}_2(x^k, \bar{x}^k) := \Pi_{T_k \cap \Gamma_k}(x^0). \tag{18}$$

Similar forward steps have been used in several papers for solving nonsmooth convex optimization problems [4], variational inequalities [3, 6], and nonsmooth inclusion problems [1, 18]. The existence of $j(k)$, satisfying (12), and the well-definition of (16) and (18) will be proved in the next section. It is worth mentioning that the projection steps defined in (16) and (18) do not increase the computational burden per iteration, that is, both steps have closed and inexpensive formulae. By using Proposition 2.5 with $y = \bar{x}^k$, $z = x^k$,

$$u = r^k := (\bar{\delta}/\alpha_k)\|x^k - \bar{x}^k\|^2$$

and

$$v = \bar{w}_2^k := \frac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k),$$

we have $T_k = T^{r_k}_{\bar{x}^k, \bar{w}^k}$ and $\Gamma_k = \Gamma_{x^k, x^0}$. Moreover, since $x^k \notin T_k$ which is proved below in Proposition 4.5,

$$\mathcal{F}_1(x^k, \bar{x}^k) = \Pi_{T_k}(x^k) = x^k - \frac{\langle \bar{w}_2^k, x^k - \bar{x}^k \rangle - r_k}{\|\bar{w}_2^k\|^2} \bar{w}_2^k$$

and also

$$\mathcal{F}_2(x^k, \bar{x}^k) = \Pi_{T_k \cap \Gamma_k}(x^0) = x^0 - \lambda_1^k \bar{w}_2^k - \lambda_2^k(x^0 - x^k),$$

where λ_1^k, λ_2^k are given by,

$$\lambda_1^k = \begin{cases} \dfrac{\left(\langle \bar{w}_2^k, x^0 - \bar{x}^k \rangle - r_k\right)\|x^0 - x^k\|^2 - \langle \bar{w}_2^k, x^0 - x^k \rangle \|x^0 - x^k\|^2}{\|\bar{w}_2^k\|^2 \|x^0 - x^k\|^2 - \langle \bar{w}_2^k, x^0 - x^k \rangle^2}, & \text{if } \dfrac{\langle \bar{w}_2^k, x^0 - x^k \rangle}{\|\bar{w}_2^k\|\|x^0 - x^k\|} < 1 \\[4mm] \dfrac{\langle \bar{w}_2^k, x^0 - \bar{x}^k \rangle - r_k}{\|\bar{w}_2^k\|^2}, & \text{if } \dfrac{\langle \bar{w}_2^k, x^0 - x^k \rangle}{\|\bar{w}_2^k\|\|x^0 - x^k\|} = 1 \end{cases} \tag{19}$$

and

$$\lambda_2^k = \begin{cases} \dfrac{\|\bar{w}_2^k\|^2 \|x^0 - x^k\|^2 - \langle \bar{w}_2^k, x^0 - x^k \rangle \left(\langle \bar{w}_2^k, x^0 - \bar{x}^k \rangle - r_k\right)}{\|\bar{w}_2^k\|^2 \|x^0 - x^k\|^2 - \langle \bar{w}_2^k, x^0 - x^k \rangle^2}, & \text{if } \dfrac{\langle \bar{w}_2^k, x^0 - x^k \rangle}{\|\bar{w}_2^k\|\|x^0 - x^k\|} < 1 \\[4mm] 0, & \text{if } \dfrac{\langle \bar{w}_2^k, x^0 - x^k \rangle}{\|\bar{w}_2^k\|\|x^0 - x^k\|} = 1. \end{cases} \tag{20}$$

4 Convergence Analysis

We start this section by presenting some technical results that are useful in analyzing the convergence properties of the two proposed methods. We also prove that **Conceptual Algorithm** is well-defined. We start proving that (12) is satisfied by j sufficiently large, hence α_k is well defined. From now on, we assume that $A_2 : \mathcal{H} \to \mathcal{H}$ is a uniformly continuous mapping. This assumption is standard to prove weak convergence of the FBHF splitting method without the Lipschitz continuity assumption. Moreover, we assume that the solution set of the inclusion problem (7), $\text{zer}(A_1 + A_2 + B)$, is nonempty.

Proposition 4.1 *The inequality* (12) *in the backtracking strategy holds after finitely many steps.*

Proof. If $x^k \in \text{zer}(A_1 + A_2 + B)$ then (12) automatically holds because Proposition 2.3. Thus, assume that x^k is not a solution, i.e., $x^k \notin \text{zer}(A_1 + A_2 + B)$. So, using (11) and Proposition 2.3, we have $x^k \neq \bar{x}_j^k = (I + \alpha_{k-1}\theta^j B)^{-1}(x^k - \alpha_{k-1}\theta^j A x^k)$ for all $j \in \mathbb{N}$. The proof of the well-definition of $j(k)$ goes by contradiction. Assume that (12) does not hold, i.e.,

$$\delta\|x^k - \bar{x}_j^k\|^2 < \alpha_{k-1}\theta^j \langle A_2 x^k - A_2\bar{x}_j^k, x^k - \bar{x}_j^k \rangle \leq \alpha_{k-1}\theta^j \|A_2 x^k - A_2\bar{x}_j^k\| \|x^k - \bar{x}_j^k\|,$$

using the Cauchy-Schwartz inequality in the last inequality. Dividing by $\|x^k - \bar{x}_j^k\| \neq 0$, we get

$$\delta\|x^k - \bar{x}_j^k\| < \alpha_{k-1}\theta^j \|A_2 x^k - A_2\bar{x}_j^k\|.$$

Since $\theta \in (0,1)$ and $\|A_2 x^k - A_2\bar{x}_j^k\|$ is bounded for all $j \in \mathbb{N}$, then by letting $j \to +\infty$ the right hand side of the last inequality goes to zero. Hence $\|x^k - \bar{x}_j^k\| \to 0$. Since A is uniformly continuous, we have

$$\|A_2 x^k - A_2\bar{x}_j^k\| \to 0. \tag{21}$$

Consequently,

$$\frac{\|x^k - \bar{x}_j^k\|}{\alpha_{k-1}\theta^j} \to 0. \tag{22}$$

Moreover, the β-cocoersivity of A_1 implies $1/\beta$-Lipschitz continuity. Then,

$$\lim_{j \to \infty} \|A_1 x^k - A_1\bar{x}_j^k\| \leq 1/\beta \lim_{j \to \infty} \|x^k - \bar{x}_j^k\| = 0,$$

which implies

$$\|A_1 x^k - A_1\bar{x}_j^k\| \to 0. \tag{23}$$

Define,

$$\bar{w}_j^k := \frac{x^k - \bar{x}_j^k}{\alpha_{k-1}\theta^j} - (A x^k - A\bar{x}_j^k).$$

It follows from (10) that $\bar{w}_j^k \in (A + B)\bar{x}_j^k$, or equivalently,

$$(\bar{x}_j^k, \bar{w}_j^k) \in \text{Gph}(A + B).$$

Observe that $A = A_1 + A_2$ and by the Cauchy-Schwarz inequality, we get

$$\|\bar{w}_j^k\| = \left\| \frac{x^k - \bar{x}_j^k}{\alpha_{k-1}\theta^j} - \left[(A_1 + A_2)x^k - (A_1 + A_2)\bar{x}_j^k \right] \right\|$$

$$\leq \frac{\|x^k - \bar{x}_j^k\|}{\alpha_{k-1}\theta^j} + \|A_1 x^k - A_1\bar{x}_j^k\| + \|A_2 x^k - A_2\bar{x}_j^k\|.$$

Hence, $\bar{w}_j^k$ converges to 0 by using (21), (22) and (23) above. Since $\bar{x}_j^k \rightharpoonup x^k$ and $\bar{w}_j^k \to 0$ and by Lemma 2.1(ii), $\text{Gph}(A + B)$ is closed in the weak-strong topology. So,

$$(x^k, 0) \in \text{Gph}(A + B),$$

or equivalently, $0 \in (A + B)x^k = (A_1 + A_2 + B)x^k$. Therefore, $x^k \in \text{zer}(A_1 + A_2 + B)$ which is a contradiction. ∎

Remark 4.2 We notice that if A_2 is L_2-Lipschitz continuous then any $\alpha \leq \dfrac{\delta}{L_2}$ satisfies the backtracking inequality (12). Actually, if we use the Cauchy-Schwartz inequality in the left hand side of (12), we get

$$\alpha \langle A_2 x^k - A_2 \bar{x}^k, x^k - \bar{x}^k \rangle \leq \frac{\delta}{L_2} \|A_2 x^k - A_2 \bar{x}^k\| \|x^k - \bar{x}^k\| \leq \delta \|x^k - \bar{x}^k\|^2.$$

Moreover, it is easy to prove that $(\alpha_k)_{k \in \mathbb{N}}$ the sequence generated by the backtracking strategy given by (12) satisfies

$$\alpha_k \geq \min \left\{ \alpha_{-1}, \frac{\delta}{L_2} \right\}$$

for all $k \in \mathbb{N}$.

From now on, we assume that $\alpha_{-1} \leq 4\beta\bar{\delta}$ where β is the cocoercive constant for A_1.

Lemma 4.3 Let $(x^k)_{k \in \mathbb{N}}$, $(\bar{x}^k)_{k \in \mathbb{N}}$ and $(\alpha_k)_{k \in \mathbb{N}}$ be the sequences generated by **Conceptual Algorithm**. Then, for all $k \in \mathbb{N}$,

(i) $\dfrac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k) \in (A_2 + B)\bar{x}^k + A_1 x^k$;

(ii) $\operatorname{zer}(A_1 + A_2 + B) \subseteq \mathrm{T}_k$.

Proof. By the definition of $\bar{x}^k$ given in (13) and (10), $\dfrac{x^k - \bar{x}^k}{\alpha_k} - Ax^k \in B\bar{x}^k$. So, (i) follows after add $A_2 \bar{x}^k + A_1 x^k$ in both sides and use that $A = A_1 + A_2$. To prove (ii) take any $x^* \in \operatorname{zer}(A_1 + A_2 + B)$. Then, there exists $v^* \in B(x^*)$, such that $0 = A_1 x^* + A_2 x^* + v^*$. Using (i) and also the monotonicity of $A_2 + B$ give us

$$0 \leq \left\langle \frac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k) - A_1 x^k - (A_2 x^* + v^*), \bar{x}^k - x^* \right\rangle$$

$$= \left\langle \frac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k) - A_1 x^k + A_1 x^*, \bar{x}^k - x^* \right\rangle.$$

Rearranging, we have

$$\left\langle \frac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k), x^* - \bar{x}^k \right\rangle \leq \left\langle A_1 x^* - A_1 x^k, \bar{x}^k - x^* \right\rangle.$$

Now, by using the β-cocoercivity of A_1, we get

$$\langle A_1 x^* - A_1 x^k, \bar{x}^k - x^* \rangle = \langle A_1 x^* - A_1 x^k, x^k - x^* \rangle + \langle A_1 x^* - A_1 x^k, \bar{x}^k - x^k \rangle$$

$$\leq -\beta \|A_1 x^* - A_1 x^k\|^2 + \frac{1}{2\alpha_k} \left[2\langle \alpha_k A_1 x^* - \alpha_k A_1 x^k, \bar{x}^k - x^k \rangle \right].$$

So, using the identity, for any $\alpha, \gamma > 0$ and $a, b \in \mathcal{H}$, $2\langle \alpha b, a \rangle = \gamma \|a\|^2 + \frac{\alpha^2}{\gamma} \|b\|^2 - \gamma \|a - \frac{\alpha}{\gamma} b\|^2$ in the right hand side of the last inequality, we get

$$\langle A_1 x^* - A_1 x^k, \bar{x}^k - x^* \rangle \leq \frac{\gamma}{2\alpha_k} \|\bar{x}^k - x^k\|^2 + \left(\frac{\alpha_k}{2\gamma} - \beta \right) \|A_1 x^* - A_1 x^k\|^2$$

$$- \frac{\gamma}{2\alpha_k} \|\bar{x}^k - x^k - \frac{\alpha_k}{\gamma}(A_1 x^* - A_1 x^k)\|^2,$$

for any $\gamma > 0$. Therefore, taking $\gamma = \frac{\alpha_{-1}}{2\beta} \geq \frac{\alpha_k}{2\beta}$ for all $k \in \mathbb{N}$ and using that $\alpha_{-1} \leq 4\beta\bar{\delta}$, we have

$$\left\langle \frac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k), x^* - \bar{x}^k \right\rangle \leq \frac{\bar{\delta}}{\alpha_k} \|x^k - \bar{x}^k\|^2 = r_k$$

and by (15), $x^* \in T_k$ as desired. ∎

Proposition 4.4 *Let $(x^k)_{k \in \mathbb{N}}$, $(\bar{x}^k)_{k \in \mathbb{N}}$ and $(\alpha_k)_{k \in \mathbb{N}}$ be the sequences generated by **Conceptual Algorithm**. Then, for all $k \in \mathbb{N}$, $\alpha_k \leq \alpha_{-1}$ and*

$$\left\langle \frac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k), x^k - \bar{x}^k \right\rangle \geq \frac{1 - \delta}{\alpha_k} \|x^k - \bar{x}^k\|^2 \geq 0. \tag{24}$$

Proof. The fact that $\alpha_k \leq \alpha_{-1}$ for all $k \in \mathbb{N}$ follows from the definition of the backtracking strategy inside of **Conceptual Algorithm**. Using the line search inequality, we have

$$\begin{aligned}
\left\langle \frac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k), x^k - \bar{x}^k \right\rangle &= \frac{\|x^k - \bar{x}^k\|^2}{\alpha_k} - \langle A_2 x^k - A_2 \bar{x}^k, x^k - \bar{x}^k \rangle \\
&\geq \frac{\|x^k - \bar{x}^k\|^2}{\alpha_k} - \frac{\delta}{\alpha_k} \|x^k - \bar{x}^k\|^2 \\
&= \frac{1 - \delta}{\alpha_k} \|x^k - \bar{x}^k\|^2.
\end{aligned}$$

So, the result follows. ∎

Proposition 4.5 *Let $(x^k)_{k \in \mathbb{N}}$, $(\bar{x}^k)_{k \in \mathbb{N}}$ and $(\alpha_k)_{k \in \mathbb{N}}$ be the sequences generated by **Conceptual Algorithm**. Then, $x^k \in T_k$ if and only if $x^k \in \mathrm{zer}(A_1 + A_2 + B)$.*

Proof. Clearly, from Lemma 4.3, if $x^k \in \mathrm{zer}(A_1 + A_2 + B)$ then $x^k \in T_k$. Conversely, if $x^k \in T_k$ then

$$\frac{\bar{\delta}}{\alpha_k} \|x^k - \bar{x}^k\|^2 \geq \left\langle \frac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k), x^k - \bar{x}^k \right\rangle \geq \frac{1 - \delta}{\alpha_k} \|x^k - \bar{x}^k\|^2,$$

using Proposition 4.4 in the second inequality. Hence, $\frac{(1 - \delta - \bar{\delta})}{\alpha_k} \|x^k - \bar{x}^k\|^2 \leq 0$, which implies that $x^k = \bar{x}^k$ and by Proposition 2.3, $x^k \in \mathrm{zer}(A_1 + A_2 + B)$. ∎

4.1 Convergence of Method 1

In this subsection all results are referred to **Method 1**, i.e., with **Conceptual Algorithm** with the iterative step (14) in **Step 2.** as

$$x^{k+1} = \mathcal{F}(x^k, \bar{x}^k) = \mathcal{F}_1(x^k, \bar{x}^k) = \Pi_{T_k}(x^k).$$

or equivalently,

$$x^{k+1} = \Pi_{T_k}(x^k) = x^k - \frac{\langle \bar{w}_2^k, x^k - \bar{x}^k \rangle - \frac{\bar{\delta}}{\alpha_k} \|x^k - \bar{x}^k\|^2}{\|\bar{w}_2^k\|^2} \bar{w}_2^k, \tag{25}$$

11

reminding that $\bar{w}_2^k = \dfrac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k)$. If we define

$$\lambda_k := \frac{\langle \bar{w}_2^k, x^k - \bar{x}^k \rangle - \dfrac{\bar{\delta}}{\alpha_k} \|x^k - \bar{x}^k\|^2}{\|\bar{w}_2^k\|^2} \tag{26}$$

then (25) becomes

$$x^{k+1} = \left(1 - \frac{\lambda_k}{\alpha_k}\right) x^k + \frac{\lambda_k}{\alpha_k} \bar{x}^k - \lambda_k (A_2 \bar{x}^k - A_2 x^k). \tag{27}$$

The above forward step is interesting on its own. It is possible to use (27) to allow over and under projections onto T_k, i.e., changing λ_k by $\gamma \lambda_k$ with $\gamma \in (0, 2)$. The analysis of convergence for the resulted new variants follows the same lines of convergence of **Method 1**. Note further that if $\lambda_k = \alpha_k$, the forward step of the FBHF splitting iteration is recovered from (27).

Proposition 4.6 *If **Method 1** stops then $x^k \in \mathrm{zer}(A_1 + A_2 + B)$.*

Proof. If Stopping Criterion is satisfied, then $x^{k+1} = \Pi_{T_k}(x^k) = x^k$, which implies that $x^k \in T_k$ and by Proposition 4.5, $x^k \in \mathrm{zer}(A_1 + A_2 + B)$. ∎

From now on, we assume that **Method 1** does not stop. Note that T_k is nonempty for all $k \in \mathbb{N}$ by Lemma 4.5. Then, the projection step (16) is well-defined, i.e., if **Method 1** does not stop, it generates an infinite sequence $(x^k)_{k \in \mathbb{N}}$.

Proposition 4.7 *Let $(x^k)_{k \in \mathbb{N}}$ be a generated sequence by **Method 1**. Then the following items are satisfied:*

(i) *The sequence $(x^k)_{k \in \mathbb{N}}$ is Fejér monotone with respect to $\mathrm{zer}(A_1 + A_2 + B)$;*

(ii) *The sequence $(x^k)_{k \in \mathbb{N}}$ is bounded;*

(iii) $\lim_{k \to \infty} \left\langle x^k - \bar{x}^k - \alpha_k (A_2 x^k - A_2 \bar{x}^k), x^k - \bar{x}^k \right\rangle - \bar{\delta} \|x^k - \bar{x}^k\|^2 = 0.$

Proof. (i) Take $x^* \in \mathrm{zer}(A_1 + A_2 + B)$. Using (16), Proposition 2.4(i) and Lemma 4.5, we have

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 &= \|\Pi_{T_k}(x^k) - \Pi_{T_k}(x^*)\|^2 \\
&\leq \|x^k - x^*\|^2 - \|\Pi_{H_k}(x^k) - x^k\|^2 \\
&\leq \|x^k - x^*\|.
\end{aligned} \tag{28}$$

Thus, the Fejér monotonicity applies. (ii) Using proposition 2.7 then the sequence is bounded. (iii) It follows from (25) that

$$x^{k+1} = \Pi_{T_k}(x^k) = x^k - \frac{\langle \bar{w}_2^k, x^k - \bar{x}^k \rangle - \dfrac{\bar{\delta}}{\alpha_k} \|x^k - \bar{x}^k\|^2}{\|\bar{w}_2^k\|^2} \bar{w}_2^k,$$

and combining this with the second line of (28), we have

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \left\| x^k - \frac{\langle \bar{w}_2^k, x^k - \bar{x}^k \rangle - \dfrac{\bar{\delta}}{\alpha_k} \|x^k - \bar{x}^k\|^2}{\|\bar{w}_2^k\|^2} \bar{w}_2^k - x^k \right\|^2$$

12

$$= \|x^k - x^*\|^2 - \frac{\left(\langle \bar{w}_2^k, x^k - \bar{x}^k \rangle - \frac{\bar{\delta}}{\alpha_k} \|x^k - \bar{x}^k\|^2 \right)^2}{\|\bar{w}_2^k\|^2}.$$

Reordering,

$$\frac{\left(\langle \bar{w}_2^k, x^k - \bar{x}^k \rangle - \frac{\bar{\delta}}{\alpha_k} \|x^k - \bar{x}^k\|^2 \right)^2}{\|\bar{w}_2^k\|^2} \leq \|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2.$$

It follows from Proposition 2.7(ii) and the definition of $\bar{w}_2^k$ that

$$0 = \lim_{k \to \infty} \frac{\left(\left\langle \frac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k), x^k - \bar{x}^k \right\rangle - \frac{\bar{\delta}}{\alpha_k} \|x^k - \bar{x}^k\|^2 \right)^2}{\left\| \frac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k) \right\|^2}$$

$$= \lim_{k \to \infty} \frac{\left(\left\langle x^k - \bar{x}^k - \alpha_k (A_2 x^k - A_2 \bar{x}^k), x^k - \bar{x}^k \right\rangle - \bar{\delta} \|x^k - \bar{x}^k\|^2 \right)^2}{\|x^k - \bar{x}^k - \alpha_k (Ax^k - A\bar{x}^k)\|^2}.$$

The sequence $\left(\|x^k - \bar{x}^k - \alpha_k (A_2 x^k - A_2 \bar{x}^k)\| \right)_{k \in \mathbb{N}}$ is bounded because the sequence $(x^k)_{k \in \mathbb{N}}$ and $(\bar{x}^k)_{k \in \mathbb{N}}$ are bounded and $\|x^k - \bar{x}^k - \alpha_k (A_2 x^k - A_2 \bar{x}^k)\| \leq \|x^k - \bar{x}^k\| + \alpha_k \|A_2 x^k - A_2 \bar{x}^k\|$, proving the desired result. ∎

Next we establish our main convergence result on **Method 1**.

Theorem 4.8 *All weak accumulation points of* $(x^k)_{k \in \mathbb{N}}$ *belong to* $\mathrm{zer}(A_1 + A_2 + B)$.

Proof. Using Proposition 4.7(iii) and taking limits in (24), we have

$$0 = \lim_{k \to \infty} \left\langle x^k - \bar{x}^k - \alpha_k (A_2 x^k - A_2 \bar{x}^k), x^k - \bar{x}^k \right\rangle - \bar{\delta} \|x^k - \bar{x}^k\|^2$$

$$\geq \lim_{k \to \infty} (1 - \delta - \bar{\delta}) \|x^k - \bar{x}^k\|^2 \geq 0.$$

Implying that,

$$\lim_{k \to \infty} \|x^k - \bar{x}^k\| = 0. \tag{29}$$

Proposition 4.7(ii) guarantees the existence of weak accumulation points of $(x^k)_{k \in \mathbb{N}}$. Let $\hat{x}$ be any weak accumulation point of $(x^k)_{k \in \mathbb{N}}$ and assume that $(x^{i_k})_{k \in \mathbb{N}}$ is any subsequence of $(x^k)_{k \in \mathbb{N}}$ that converges weakly to $\hat{x}$ and also, without loss of generality, assume that $\lim_{k \to \infty} \alpha_{i_k} = \bar{\alpha}$. Hence, it follows from (29) that $(\bar{x}^{i_k})_{k \in \mathbb{N}}$ converges weakly to $\hat{x}$ as well.

In the following, we can consider two cases.

(a) Assume that $\bar{\alpha} > 0$. Hence, (29) implies

$$\lim_{k \to \infty} \frac{\|x^{i_k} - \bar{x}^{i_k}\|}{\alpha_{i_k}} \leq \frac{1}{\bar{\alpha}} \cdot \lim_{k \to \infty} \|x^{i_k} - \bar{x}^{i_k}\| = 0, \tag{30}$$

because the sequence $(\alpha_k)_{k \in \mathbb{N}}$ is a nonincreasing sequence. Notice that,

$$\bar{w}^{i_k} := \frac{x^{i_k} - \bar{x}^{i_k}}{\alpha_{i_k}} - (A_1 x^{i_k} - A_1 \bar{x}^{i_k}) - (A_2 x^{i_k} - A_2 \bar{x}^{i_k}) \in (A_1 + A_2 + B)\bar{x}^{i_k},$$

13

which is equivalent to $(\bar{x}^{i_k}, \bar{w}^{i_k}) \in \text{Gph}(A_1 + A_2 + B)$. The fact that $\bar{w}^{i_k} \to 0$ follows by using Cauchy-Schwartz, (30), (29), the uniformly continuity assumption of A_2 and the cocoercivity of A_1. Since $\bar{w}^{i_k} \to 0$ and $\bar{x}^{i_k} \rightharpoonup \hat{x}$ then the closedness of $\text{Gph}(A_1 + A_2 + B)$ implies that $(\hat{x}, 0) \in \text{Gph}(A_1 + A_2 + B)$, and therefore $\hat{x} \in \text{zer}(A_1 + A_2 + B)$.

(b) Assume that $\bar{\alpha} = 0$, and the proof will be similar to the proof of Proposition 4.1. For simplicity, define $\hat{\alpha}_{i_k} := \dfrac{\alpha_{i_k}}{\theta}$. Then, because $\theta \in (0,1)$, we have that $\hat{\alpha}_{i_k} > \alpha_{i_k}$ and

$$\lim_{k \to \infty} \hat{\alpha}_{i_k} = 0. \tag{31}$$

Define

$$\hat{x}^{i_k} := (I + \hat{\alpha}_{i_k} B)^{-1}(x^{i_k} - \hat{\alpha}_{i_k} A x^{i_k}) \tag{32}$$

and for all $k \in \mathbb{N}$. Hence, $\hat{\alpha}_{i_k} \langle A_2 x^{i_k} - A_2 \hat{x}^{i_k}, x^{i_k} - \hat{x}^{i_k} \rangle > \delta \| x^{i_k} - \hat{x}^{i_k} \|^2$. Using the Cauchy-Schwartz inequality, we obtain

$$\delta \| x^{i_k} - \hat{x}^{i_k} \|^2 < \hat{\alpha}_{i_k} \| A_2 x^{i_k} - A_2 \hat{x}^{i_k} \| \| x^{i_k} - \hat{x}^{i_k} \|.$$

Dividing by $\| x^{i_k} - \hat{x}^{i_k} \|$, we get

$$\delta \| x^{i_k} - \hat{x}^{i_k} \| < \hat{\alpha}_{i_k} \| A_2 x^{i_k} - A_2 \hat{x}^{i_k} \|. \tag{33}$$

Letting $k \to +\infty$, and since $(\| A_2 x^{i_k} - A_2 \hat{x}^{i_k} \|)_{k \in \mathbb{N}}$ is bounded and (31), we have $\| x^{i_k} - \hat{x}^{i_k} \| \to 0$. Hence, $\hat{x}^{i_k} \rightharpoonup \hat{x}$. Since A_2 is uniformly continuous then

$$\| A_2 x^{i_k} - A_2 \hat{x}^{i_k} \| \to 0. \tag{34}$$

Using again (33),

$$\frac{\| x^{i_k} - \hat{x}^{i_k} \|}{\hat{\alpha}_{i_k}} \to 0. \tag{35}$$

The definition of $\hat{x}^{i_k}$ in (32) together with (10) imply

$$\hat{w}^{i_k} := \frac{x^k - \hat{x}^{i_k}}{\hat{\alpha}_{i_k}} - (A_1 x^{i_k} - A_1 \hat{x}^{i_k}) - (A_2 x^{i_k} - A_2 \hat{x}^{i_k}) \in (A_1 + A_2 + B)\hat{x}^{i_k},$$

or equivalently,

$$(\hat{x}^{i_k}, \hat{w}^{i_k}) \in \text{Gph}(A_1 + A_2 + B).$$

Observe that

$$\| \hat{w}^{i_k} \| = \left\| \frac{x^{i_k} - \hat{x}^{i_k}}{\hat{\alpha}_{i_k}} - (A_1 x^{i_k} - A_1 \hat{x}^{i_k}) - (A_2 x^{i_k} - A_2 \hat{x}^{i_k}) \right\|$$

$$\leq \frac{\| x^{i_k} - \bar{x}^{i_k} \|}{\hat{\alpha}_{i_k}} + \| A_1 x^{i_k} - A_1 \bar{x}^{i_k} \| + \| A_2 x^{i_k} - A_2 \bar{x}^{i_k} \|$$

$$\leq \frac{\| x^{i_k} - \bar{x}^{i_k} \|}{\hat{\alpha}_{i_k}} + \beta \| x^{i_k} - \bar{x}^{i_k} \| + \| A_2 x^{i_k} - A_2 \bar{x}^{i_k} \|$$

Hence, $\hat{w}^{i_k}$ converges to 0 by using (35), (29) and (34). Since $\hat{x}^{i_k} \rightharpoonup \hat{x}$ and $\hat{w}^{i_k} \to 0$ and by the sequentially closedness in the weak-strong topology of $\text{Gph}(A_1 + A_2 + B)$ (Lemma 2.1(ii)), we have

$$(\hat{x}, 0) \in \text{Gph}(A_1 + A_2 + B),$$

or equivalently, $0 \in (A_1 + A_2 + B)\hat{x}$. Then, $\hat{x} \in \text{zer}(A_1 + A_2 + B)$. Then, all weak accumulation points of $(x^k)_{k \in \mathbb{N}}$ belong to $\text{zer}(A_1 + A_2 + B)$. ∎

Theorem 4.9 *The generated sequence* $(x^k)_{k \in \mathbb{N}}$ *by* **Method 1** *converges weakly to some element belonging to the optimal solutions set* $\text{zer}(A_1 + A_2 + B)$.

Proof. The result follows from the Fejér monotonicity of $(x^k)_{k \in \mathbb{N}}$ to $\text{zer}(A_1 + A_2 + B)$ proved in Proposition 4.7(i), the optimality of the weak accumulation points showed in Theorem 4.8 and Proposition 2.7(iii). $\blacksquare$

4.2 Convergence of Method 2

In this subsection all results are referent to **Method 2**, i.e., with the iterative step (14) in **Step 2.** as

$$x^{k+1} = \mathcal{F}(x^k, \bar{x}^k) = \mathcal{F}_2(x^k, \bar{x}^k) = \Pi_{T_k \cap \Gamma_k}(x^0).$$

Provided that,

$$x^{k+1} = \Pi_{T_k \cap \Gamma_k}(x^0) = x^0 - \lambda_1^k \bar{w}^k - \lambda_2^k(x^0 - x^k)$$

$$= (1 - \lambda_2^k)(x^0 - x^k) + \left(1 - \frac{\lambda_1^k}{\alpha_k}\right)x^k + \frac{\lambda_1^k}{\alpha_k}\bar{x}^k - \lambda_1^k \left[A_2 \bar{x}^k - A_2 x^k\right].$$

where λ_1^k and λ_2^k are given by (19) and (20). This forward step is still more general than the forward step of **Method 1** and it is interesting on its own. Actually, if $\lambda_1^k = \lambda_k$ given in (26) and $\lambda_2^k = 1$ then it recovers the projection-forward step of **Method 1**. Moreover, the projection forward step for **Method 2** can be replaced by under and over projections (or even inexact projections) onto $T_k \cap \Gamma_k$ producing strongly convergent versions of **Method 2**.

Proposition 4.10 *If* **Method 2** *stops then* $x^k \in \text{zer}(A_1 + A_2 + B)$.

Proof. If **Method 2** stops then $x^{k+1} = \Pi_{T_k \cap \Gamma_k}(x^0) = x^k$. So, $x^k \in T_k \cap \Gamma_k \subset T_k$ and it implies that, by using Proposition 4.5, $x^k \in \text{zer}(A_1 + A_2 + B)$. $\blacksquare$

To avoid this case, we may assume that **Method 2** does not stop. Note that, Γ_k and T_k are convex and closed halfspaces, for all k. Therefore $T_k \cap \Gamma_k$ is a convex and closed set. So, if $T_k \cap \Gamma_k$ is nonempty, then x^{k+1} is well-defined. This notion is confirmed using the following lemma.

Lemma 4.11 $\text{zer}(A_1 + A_2 + B) \subseteq T_k \cap \Gamma_k$, *for all* $k \in \mathbb{N}$.

Proof. Note that, the set of optimal solutions $\text{zer}(A_1 + A_2 + B) \neq \emptyset$. By Proposition 4.3, $\text{zer}(A_1 + A_2 + B) \subseteq T_k$, for all $k \in \mathbb{N}$. When $k = 0$, it is the case $T_0 = \mathcal{H}$, we get $\text{zer}(A_1 + A_2 + B) \subset T_0 \cap \Gamma_0$. For all $\ell \in \mathbb{N}$ such that $\ell \leq k$, we have by induction hypothesis that $\text{zer}(A_1 + A_2 + B) \subset T_\ell \cap \Gamma_\ell$. Therefore, $x^{k+1} = \Pi_{T_k \cap \Gamma_k}(x^0)$ is well-defined. Then, the following inequality is consequence of the induction hypothesis and by Proposition 2.4(ii), i.e.,

$$\langle x^* - x^{k+1}, x^0 - x^{k+1}\rangle = \langle x^* - \Pi_{T_k \cap \Gamma_k}(x^0), x^0 - \Pi_{T_k \cap \Gamma_k}(x^0)\rangle \leq 0, \tag{36}$$

for all $x^* \in \text{zer}(A_1 + A_2 + B)$. Notice that (36) and (17) imply that $x^* \in \Gamma_{k+1}$ and hence, $\text{zer}(A_1 + A_2 + B) \subset T_{k+1} \cap \Gamma_{k+1}$. So, the result follows. $\blacksquare$

We have shown that the set $T_k \cap \Gamma_k$ is nonempty and therefore the projection step, given in (18), is well-defined.

Corollary 4.12 **Method 2** *is well-defined.*

Proof. By Lemma 4.11 , $\mathrm{zer}(A_1 + A_2 + B) \subset \mathrm{T}_k \cap \Gamma_k$, for all $k \in \mathbb{N}$. Then, for a given initial x^0, the sequence $(x^k)_{k \in \mathbb{N}}$ is attainable. ∎

To prove the convergence of the sequence, we need first to show conditions and bounds on the sequence. Next lemma tackles this issue by restricting the sequence in a ball defined by its initial point x^0.

Lemma 4.13 *The sequence $(x^k)_{k \in \mathbb{N}}$ is bounded. Moreover,*

$$(x^k)_{k \in \mathbb{N}} \subset \mathbb{B}\left[\frac{1}{2}(x^0 + \bar{x}); \frac{1}{2}\rho\right],$$

where $\bar{x} := \Pi_{\mathrm{zer}(A_1 + A_2 + B)}(x^0)$ and $\rho := \mathrm{dist}(x^0, \mathrm{zer}(A_1 + A_2 + B))$.

Proof. Using Lemma 4.11 we see that $\mathrm{zer}(A_1 + A_2 + B) \subseteq \mathrm{T}_k \cap \Gamma_k$, and by (18), we obtain that

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

for all $k \in \mathbb{N}$ and all $x^* \in \mathrm{zer}(A_1 + A_2 + B)$. Taking $x^* = \bar{x}$ in the above inequality, we have

$$\|x^{k+1} - x^0\| \leq \|\bar{x} - x^0\| = \rho, \quad \text{for all} \quad k \in \mathbb{N}.$$

Proving that $(x^k)_{k \in \mathbb{N}}$ is bounded.
Now, define $x_k^* := x^k - \frac{1}{2}(x^0 + \bar{x})$ and $\bar{x}^* := \bar{x} - \frac{1}{2}(x^0 + \bar{x})$. Since $\bar{x} \in \Gamma_{k+1}$, then we have

$$
\begin{aligned}
0 &\geq 2\langle \bar{x} - x^{k+1}, x^0 - x^{k+1} \rangle \\
&= 2\left\langle \bar{x}^* + \frac{1}{2}(x^0 + \bar{x}) - x_{k+1}^* - \frac{1}{2}(x^0 + \bar{x}), x_0^* + \frac{1}{2}(x^0 + \bar{x}) - x_{k+1}^* - \frac{1}{2}(x^0 + \bar{x}) \right\rangle \\
&= 2\left\langle \bar{x}^* - x_{k+1}^*, x_0^* - x_k^* \right\rangle = \left\langle \bar{x}^* - x_{k+1}^*, -\bar{x}^* - x_{k+1}^* \right\rangle = \|x_{k+1}^*\|^2 - \|\bar{x}^*\|^2,
\end{aligned}
$$

Thus, after manipulating, we obtain

$$\left\| x^{k+1} - \frac{x^0 + \bar{x}}{2} \right\| \leq \left\| \bar{x} - \frac{x^0 + \bar{x}}{2} \right\| = \frac{\rho}{2}, \quad \text{for all} \quad k \in \mathbb{N}.$$

Hence, the result follows and $(x^k)_{k \in \mathbb{N}}$ lies in the ball determined by the initial point x^0. ∎

An important question arisen here is, whether the set of optimal solutions contains any weak accumulation point of $(x^k)_{k \in \mathbb{N}}$ or not. Next lemma answers this question by showing that all weak accumulation points must lie in the optimal solutions set, $\mathrm{zer}(A_1 + A_2 + B)$.

Lemma 4.14 *Let $(x^k)_{k \in \mathbb{N}}$ be a sequence generated by **Method 2**. Then, all weak accumulation points of the sequence $(x^k)_{k \in \mathbb{N}}$ belong to $\mathrm{zer}(A_1 + A_2 + B)$.*

Proof. If $x^{k+1} \in \Gamma_k$, then

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

which implies $0 \leq \|x^{k+1} - x^k\|^2 \leq \|x^{k+1} - x^0\|^2 - \|x^k - x^0\|^2$. Hence, $(\|x^k - x^0\|)_{k \in \mathbb{N}}$ is a nondecreasing sequence. Since the sequence $(\|x^k - x^0\|)_{k \in \mathbb{N}}$ is bounded, by Lemma 4.13, it is therefore convergent. Thus,

$$\lim_{k \to \infty} \|x^{k+1} - x^k\| = 0. \tag{37}$$

Since the projection of x^k onto the halfspace T_k is denoted by $\Pi_{T_k}(x^k)$, then we have the inequality

$$0 \le \|x^k - \Pi_{T_k}(x^k)\| \le \|x^k - x\| \qquad \text{for all} \quad x \in T_k.$$

The fact that $x^{k+1} \in T_k$ implies $0 \le \|x^k - \Pi_{T_k}(x^k)\| \le \|x^k - x^{k+1}\|$. Since $\|x^k - x^{k+1}\| \to 0$ by (37), we have

$$0 = \lim_{k\to\infty} \|x^k - \Pi_{T_k}(x^k)\| = \lim_{k\to\infty} \frac{\left\langle \dfrac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k), x^k - \bar{x}^k \right\rangle - \dfrac{\bar{\delta}}{\alpha_k} \|x^k - \bar{x}^k\|^2}{\left\| \dfrac{x^k - \bar{x}^k}{\alpha_k} - (A_2 x^k - A_2 \bar{x}^k) \right\|}$$

$$= \lim_{k\to\infty} \frac{\left\langle x^k - \bar{x}^k - \alpha_k(A_2 x^k - A_2 \bar{x}^k), x^k - \bar{x}^k \right\rangle - \bar{\delta}\|x^k - \bar{x}^k\|^2}{\|x^k - \bar{x}^k - \alpha_k(A_2 x^k - A_2 \bar{x}^k)\|}.$$

The sequence $\left(\|x^k - \bar{x}^k - \alpha_k(A_2 x^k - A_2 \bar{x}^k)\| \right)_{k\in\mathbb{N}}$ is bounded because the sequence $(x^k)_{k\in\mathbb{N}}$ and $(\bar{x}^k)_{k\in\mathbb{N}}$ are bounded and $\|x^k - \bar{x}^k - \alpha_k(A_2 x^k - A_2 \bar{x}^k)\| \le \|x^k - \bar{x}^k\| + \alpha_k\|A_2 x^k - A_2 \bar{x}^k\|$. Hence,

$$0 = \lim_{k\to\infty} \langle x^k - \bar{x}^k - \alpha_k(A_2 x^k - A_2 \bar{x}^k), x^k - \bar{x}^k \rangle - \bar{\delta}\|x^k - \bar{x}^k\|^2 \ge (1 - \delta - \bar{\delta}) \lim_{k\to\infty} \|x^k - \bar{x}^k\|^2 \ge 0,$$

using (12) in the first inequality. Hence, $\lim_{k\to\infty} \|x^k - \bar{x}^k\| = 0$. Now the rest of the proof follows similarly to the proof of Theorem 4.8 after (29). ∎

Finally, we prove the main theorem of convergence of the sequence $(x^k)_{k\in\mathbb{N}}$ generated by **Method 2**, which converges to the nearest solution to x^0.

Theorem 4.15 *Let $(x^k)_{k\in\mathbb{N}}$ be a sequence generated by* **Method 2**. *Define $\bar{x} := \Pi_{\text{zer}(A_1+A_2+B)}(x^0)$. Then, $(x^k)_{k\in\mathbb{N}}$ converges strongly to $\bar{x}$.*

Proof. Since $\text{zer}(A_1 + A_2 + B)$ is nonempty closed convex by Lemma 2.1(i) then there exists a metric (orthogonal) projection of x^0 onto $\text{zer}(A_1 + A_2 + B)$, i.e. $\bar{x} = \Pi_{\text{zer}(A_1+A_2+B)}(x^0)$. Since $\bar{x} \in \text{zer}(A_1 + A_2 + B) \subseteq T_k \cap \Gamma_k$, and by using the projection definition of x^{k+1} onto the intersection of the hyperplanes $T_k \cap \Gamma_k$ as in (18), we then have the property

$$\|x^k - x^0\| \le \|\bar{x} - x^0\| \quad \text{for all} \quad k \in \mathbb{N}. \tag{38}$$

Since $(x^k)_{k\in\mathbb{N}}$ is bounded by Lemma 4.13, and every weak accumulation point of the sequence $(x^k)_{k\in\mathbb{N}}$ is in the set of optimal solutions $\text{zer}(A_1 + A_2 + B)$ by Lemma 4.14, then any subsequence $(x^{i_k})_{k\in\mathbb{N}}$ of $(x^k)_{k\in\mathbb{N}}$ must converge weakly to an accumulation point belonging to $\text{zer}(A_1 + A_2 + B)$ (say $x^{i_k} \rightharpoonup \hat{x} \in \text{zer}(A_1 + A_2 + B)$). Thus,

$$\|x^{i_k} - \bar{x}\|^2 = \|x^{i_k} - x^0 - (\bar{x} - x^0)\|^2$$
$$= \|x^{i_k} - x^0\|^2 + \|\bar{x} - x^0\|^2 - 2\langle x^{i_k} - x^0, \bar{x} - x^0 \rangle$$
$$\le 2\|\bar{x} - x^0\|^2 - 2\langle x^{i_k} - x^0, \bar{x} - x^0 \rangle,$$

using (38) in the last inequality. Since the subsequence $(x^{i_k})_{k\in\mathbb{N}}$ is weakly convergent to $\hat{x}$, we get

$$\limsup_{k\to\infty} \|x^{i_k} - \bar{x}\|^2 \le 2\|\bar{x} - x^0\|^2 - 2\langle \hat{x} - x^0, \bar{x} - x^0 \rangle$$

$$= 2\langle \bar{x} - \hat{x}, \bar{x} - x^0 \rangle + 2\langle \hat{x} - x^0, \bar{x} - x^0 \rangle - 2\langle \hat{x} - x^0, \bar{x} - x^0 \rangle$$
$$= 2\langle \bar{x} - \hat{x}, \bar{x} - x^0 \rangle$$
$$\leq 0,$$

where the last inequality follows from the definition of $\bar{x}$ and Proposition 2.4(ii). This implies that $x^{i_k} \to \bar{x}$. Hence, we have proved that every weakly convergent subsequence of the generated sequence $(x^k)_{k \in \mathbb{N}}$ converges strongly to $\bar{x} \in \mathrm{zer}(A_1 + A_2 + B)$. Hence, the whole sequence $(x^k)_{k \in \mathbb{N}}$ converges strongly to $\bar{x} = \Pi_{\mathrm{zer}(A_1+A_2+B)}(x^0)$ proving the desired result. ∎

5   Concluding Remarks

This paper dealt with the weak and strong convergence of a modification of the forward-backward-forward (FBF) splitting method for solving monotone inclusions. We propose a conceptual algorithm modifying the forward-backward-half-forward (FBHF) splitting algorithm for finding a zero to the sum of three monotone operators. The two proposed modified variants use a backtracking strategy to find a suitable separating hyperplane, which splits the space into two halfspaces, one containing the optimal solution set and one containing the current iterate. Two explicit projection steps onto suitable halfspaces produce two methods with desired features. It is worth emphasizing that one of the variants is strongly convergent to the best approximation solution, and both proposed methods produce very general iterations recovering the FBHF splitting method.

We finalize by mentioning that replacing the Euclidean norm in the projection (forward) steps by a more general norm or Bregman distance may produce variants of the proposed iterations with a certain interest in some applications. Therefore, extending the presented analysis for more general projection steps could be a promising subject for future research.

Data Availability Statement: Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

References

[1] Bauschke, H.H., Combettes, Patrick L. A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces. *Mathematics of Operations Research* **26** (2001) 248–264.

[2] Bauschke, H.H., Combettes, Patrick L. Convex Analysis and Monotone Operator Theory in Hilbert Spaces. *Springer, New York* (2011).

[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. A strongly convergent method for nonsmooth convex minimization in Hilbert spaces. *Numerical Functional Analysis and Optimization* **32** (2011) 1009–1018.

[5] Bello-Cruz, J.Y., Díaz Millán, R. A variant of Forward-Backward splitting method for the sum of two monotone operators with a new search strategy. *Optimization* **64** (2015) 1471–1486.

[6] Bello-Cruz, J.Y., Díaz Millán, R., Phan, M. Hung. Conditional extragradient algorithms for variational inequalities. *Pacific Journal of Optimization* **15** (2019) 331–357.

[7] Bello-Cruz, J.Y., Nghia, T.T.A. On the convergence of the proximal forward-backward splitting method with linesearches. *Optimization methods & software* **31** (2016) 1209–1238.

[8] Briceño-Arias L.M., Davis D. Forward-Backward-Half-Forward Algorithm for Solving Montone Inclusions. *SIAM Journal on Optimization* **28** (2018) 2839–2871.

[9] Douglas, J., Rachford, H.H. On the numerical solution of heat conduction problems in two or three space variables. *Transactions of the American Mathematical Society* **82** (1956) 421–439.

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

[11] Giselsson, P. Nonlinear Forward-Backward Splitting with Projection Correction. arXiv:1908.07449, (2021).

[12] Glowinski, R., Osher, S. J. Yin, W., Editors. *Splitting methods in communication, imaging, science, and engineering.* Scientific Computation. Springer, Cham, (2016).

[13] Güler, O. New proximal point algorithms for convex minimization. *SIAM Journal on Optimization* **2** (1992) 649–664.

[14] Korpelevich, G.M. The extragradient method for finding saddle points and other problems. *Ekonomika i Matematcheskie Metody* **12** (1976) 747–756.

[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. Monotone (nonlinear) operators in Hilbert Space. *Duke Mathetematical Journal* **29** (1962) 341–346.

[17] 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.

[18] Solodov, M.V., Svaiter, B.F. Forcing strong convergence of proximal point iterations in a Hilbert space. *Mathematical Programming* **87** (2000) 189–202.

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

[20] Zhu, D.L., Marcotte, P. Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities. *SIAM Journal on Optimization* **6** (1996) 714–726.