https://doi.org/10.1186/s13663-022-00732-9

Fixed Point Theory and Algorithms for Sciences and Engineering

RESEARCH Open Access

Self-adaptive forward–backward splitting algorithm for the sum of two monotone operators in Banach spaces

Abdulmalik U. Bello[1,2*], Charles E. Chidume[1] and Maryam Alka[1,3]

*Correspondence: uabdulmalik@aust.edu.ng
[1] African University of Science and Technology, Km 10 Airport Road, Abuja, Nigeria
[2] Federal University Dutsinma, Katsina, Nigeria
Full list of author information is available at the end of the article

Abstract
In this work, we prove the weak convergence of a one-step self-adaptive algorithm to a solution of the sum of two monotone operators in 2-uniformly convex and uniformly smooth real Banach spaces. We give numerical examples in infinite-dimensional spaces to compare our result with some existing algorithms. Finally, our results extend and complement several existing results in the literature.

MSC: 47H09; 47H10; 49J20; 49J40

Keywords: Maximal monotone operators; Lipschitz-continuous operator; Forward–reflected–backward splitting method; 2-uniformly convex spaces

1 Introduction

Let $\mathcal{E}$ be a real Banach space and $\mathcal{E}^*$ be its topological dual. A problem of significant interest in nonlinear analysis is to find

$$x \in (A+B)^{-1}(0) \tag{1.1}$$

with $(A+B)^{-1}(0) \neq \emptyset$, where $A : \mathcal{E} \to 2^{\mathcal{E}^*}$ is a maximal monotone operator and $B : \mathcal{E} \to \mathcal{E}^*$ is a monotone and Lipschitz map. Interest in problem (1.1) stems from its diverse application in different areas of nonlinear analysis such as optimization, variational inequality, split feasibility problems, and saddle-point problems with applications to signal and image processing and machine learning; see, for instance, Attouch *et al.* [6], Bruck [10], Censor and Elfvin [11], Chen and Rockafellar [12], Combettes and Wajs [15], Davis and Yin [16], Lions and Mercier [19], Moudafi and Thera [22], Passty [23], Peaceman and Rachford [24] for more treatments of problem (1.1). Consider, for instance, the split-feasibility problem, introduced by Censor and Elfvin [11], which is to find

$$x \in \mathcal{C}_1 \quad \text{such that } Tx \in \mathcal{C}_2, \tag{1.2}$$

where $\mathcal{C}_1 \subset \mathcal{H}_1$, $\mathcal{C}_2 \subset \mathcal{H}_2$ are nonempty, closed, and convex subsets of the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, and $T : \mathcal{H}_1 \to \mathcal{H}_2$ is a bounded linear map. Then, (1.2) can be transformed into the monotone inclusion

$$\text{find } x \in \mathcal{C}_1 \quad \text{such that}, 0 \in N_{\mathcal{C}_1}(x) + T^*(I - P_{\mathcal{C}_2})Tx.$$

By setting

$$A(x) = N_{\mathcal{C}_1}(x) \quad \text{and} \quad B(x) = \nabla\left(\frac{1}{2}\|Tx - P_{\mathcal{C}_2}Tx\|^2\right) = T^*(I - P_{\mathcal{C}_2})Tx,$$

where $N_{\mathcal{C}_1}(x)$ is the normal cone of $\mathcal{C}_1$ at x and T^* is the adjoint operator of T, (1.2) can be reformulated as (1.1).

There are several methods of approximating solutions of (1.1), see, e.g., [1, 6, 9, 10, 16, 19, 20, 23, 30]. One of the most efficient methods is the *forward–backward* splitting method introduced by Passty [23], and Lions and Mercier [19]. The method generates a sequence $\{x_n\}$ iteratively defined by

$$x_{n+1} = J^A_{\lambda_n}(x_n - \lambda_n Bx_n), \quad n \geq 0. \tag{1.3}$$

They proved that if the operator B is μ-cocoercive, that is, there exists $\mu > 0$ such that

$$\langle x - y, B(x) - B(y)\rangle \geq \mu \|B(x) - B(y)\|^2 \quad \forall x, y \in \mathcal{E},$$

and $\liminf \lambda_n > 0$ with $\limsup \lambda_n < 2\mu$, then the sequence $\{x_n\}$ generated by (1.3) converges weakly to a solution of (1.1). The cocoercivity requirement imposed on the operator B limits the class of operators for which the forward–backward splitting method is applicable. In fact, there are some important problems in applications where the forward–backward splitting method fails to converge due to the lack of coercivity of one of the operators. For instance, the first-order optimality condition for the saddle-point problems of the form

$$\min_{x \in \mathcal{H}_1} \max_{y \in \mathcal{H}_2} f_1(x) + \Phi(x, y) + f_2(y), \tag{1.4}$$

where $f_1 : \mathcal{H}_1 \to \mathbb{R} \cup \{+\infty\}$ and $f_2 : \mathcal{H}_2 \to \mathbb{R} \cup \{+\infty\}$ are proper convex and lower semicontinuous functions and $\Phi : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathbb{R}$ is a smooth convex–concave function. Then, (1.4) can be expressed as

$$\text{find } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathcal{H}_1 \times \mathcal{H}_2, \quad \text{such that } \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \begin{pmatrix} \partial f_1(x) \\ \partial f_2(y) \end{pmatrix} + \begin{pmatrix} \nabla_x \Phi(x, y) \\ -\nabla_y \Phi(x, y) \end{pmatrix}.$$

This can be seen as (1.1) with

$$A = \begin{pmatrix} \partial f_1(x) \\ \partial f_2(y) \end{pmatrix}, \quad \text{and} \quad B = \begin{pmatrix} \nabla_x \Phi(x, y) \\ -\nabla_y \Phi(x, y) \end{pmatrix}.$$

Problem (1.4) arises naturally in different areas of application such as statistics, machine learning, and optimization to mention but a few. Although the operator B, in this case, is Lipschitz whenever $\nabla\Phi$ is, B is never cocoercive even when Φ is bilinear. Thus, the development of an iterative method in which the cocoercivity of B is dispensed with is desirable.

In [28], Tseng introduced the *forward–backward–forward splitting method* (*FBFSM*) for approximating solutions of (1.1). The method generates a sequence $\{x_n\}$ iteratively defined by

$$\begin{cases} y_n = J_{\lambda_n}^A(x_n - \lambda_n B(x_n)), \\ x_{n+1} = y_n - \lambda_n B(y_n) + \lambda_n B(x_n), \quad \forall n \in \mathbb{N}, \end{cases} \tag{1.5}$$

with $\lambda_n \in (0, \frac{1}{L})$, where L is the Lipschitz constant of B. Tseng was able to dispense with the cocoercivity of the operator B at the expense of its evaluation twice per iteration. Recently, Malitsky and Tam [21], introduced the *forward–reflected–backward splitting method* (*FRBSM*) generated iteratively by

$$x_{n+1} = J_{\lambda_n}^A\big(x_n - \lambda_n B(x_n) - \lambda_{n-1}\big(B(x_n) - B(x_{n-1})\big)\big), \quad \forall n \in \mathbb{N}, \tag{1.6}$$

with $\lambda_n \in (\epsilon, \frac{1-2\epsilon}{2L})$ and $\epsilon > 0$. The forward–reflected–backward splitting method requires only one evaluation of the operator B per iteration. Thus, improving on the computational cost when compared to the forward–backward–forward method that requires two evaluations of the operator B per iteration. It is worth noting that the step sizes in each of the algorithms introduced by Tseng; and Tam and Malitski heavily depend on the prior knowledge of the Lipshitz constant of one of the operators that, some times, may be difficult to compute.

To overcome this difficulty, very recently, Hieu *et. al.* [17], introduced the *modified forward–reflected–backward splitting method* (*MFRBSM*) generated iteratively by

$$x_{n+1} = J_{\lambda_n}^A\big(x_n - \lambda_n B(x_n) - \lambda_{n-1}\big(B(x_n) - B(x_{n-1})\big)\big), \quad \forall n \in \mathbb{N}, \tag{1.7}$$

with

$$\lambda_{n+1} := \min\left\{\lambda_n, \frac{\theta\|x_{n+1} - x_n\|}{\|Bx_{n+1} - Bx_n\|}\right\}, \quad \theta \in \left(0, \frac{1}{2}\right).$$

They proved the weak convergence of Algorithm (1.7) to a solution of (1.1). It is worth noting that the variable step sizes here do not require prior knowlegde of the Lipschitz constant.

All the results mentioned above are obtained in the setting of Hilbert spaces. There are few results regarding the forward–backward method and its variants in Banach spaces, see, e.g., [26, 29]. One of the difficulties, perhaps, is the fact that the operators A and B go from the Banach space $\mathcal{E}$ to its dual $\mathcal{E}^*$. The tools available in Hilbert spaces are not readily available in general Banach spaces. Moreover, the Lipschitz constant is, in general, often unknown in practice. In fact, in nonlinear problems it may be difficult to approximate. In those cases an algorithm with a linesearch is often used (see, e.g., [26]). However, a linesearch algorithm needs an inner loop with some stopping criterion over iterations and this task may be time consuming. In this paper, we prove the weak convergence of the forward–reflected–backward splitting method in 2-uniformly convex uniformly smooth real Banach spaces with variable step sizes that do not depend on the Lipschitz constant and without any linesearch Procedure. Our results extend, unify, and complement many existing results in the literature.

2 Preliminaries

In this section, we give some basic definitions and lemmas that will be used in the proof of our main results. Let $\mathcal{E}$ be a real normed linear space. Let $S_{\mathcal{E}}$ and $B_{\mathcal{E}}$ denote the unit sphere and the closed unit ball of $\mathcal{E}$, respectively. The modulus of smoothness of $\mathcal{E}$, $\rho_{\mathcal{E}}: [0,\infty) \to [0,\infty)$ is defined by

$$\rho_{\mathcal{E}}(t) := \sup\left\{ \frac{\|x+y\| + \|x-y\|}{2} - 1 : x \in S_{\mathcal{E}}, \|y\| = t \right\}.$$

The space $\mathcal{E}$ is said to be smooth if

$$\lim_{t\to 0} \frac{\|x+ty\| - \|x\|}{t} \tag{2.1}$$

exists for all $x, y \in S_{\mathcal{E}}$. The space $\mathcal{E}$ is also said to be uniformly smooth if the limit in (2.1) converges uniformly for all $x, y \in S_{\mathcal{E}}$; and $\mathcal{E}$ is said to be 2-uniformly smooth, if there exists a fixed constant $c > 0$ such that $\rho_{\mathcal{E}}(t) \leq ct^2$. It is well known that every 2-uniformly smooth space is uniformly smooth. A real normed space $\mathcal{E}$ is said to be strictly convex if

$$\left\| \frac{(x+y)}{2} \right\| < 1 \quad \text{for all } x, y \in S_{\mathcal{E}} \text{ and } x \neq y.$$

$\mathcal{E}$ is said to be uniformly convex if $\delta_{\mathcal{E}}(\epsilon) > 0$ for all $\epsilon \in (0,2]$, where $\delta_{\mathcal{E}}$ is the modulus of convexity of $\mathcal{E}$ defined by

$$\delta_{\mathcal{E}}(\epsilon) := \inf\left\{ 1 - \left\| \frac{x+y}{2} \right\| : x, y \in B_{\mathcal{E}}, \|x-y\| \geq \epsilon \right\}, \tag{2.2}$$

for all $\epsilon \in (0,2]$. The space $\mathcal{E}$ is said to be 2-uniformly convex if there exists $c > 0$ such that $\delta_{\mathcal{E}}(\epsilon) \geq c\epsilon^2$ for all $\epsilon \in (0,2]$. It is obvious that every 2-uniformly convex Banach space is uniformly convex. It is known that all Hilbert spaces are uniformly smooth and 2-uniformly convex. It is also known that all the Lebesgue spaces L_p are uniformly smooth for $1 < p \leq \infty$, and 2-uniformly convex whenever $1 < p \leq 2$ (see [8]).

Let $\mathcal{E}$ be a real normed space. The normalized duality mapping of $\mathcal{E}$ into $\mathcal{E}^*$ is defined by

$$Jx := \left\{ x^* \in \mathcal{E}^* : \langle x^*, x \rangle = \|x^*\|^2 = \|x\|^2 \right\},$$

for all $x \in \mathcal{E}$. The normalized duality mapping J has the following properties (see, e.g., [27]):

- if $\mathcal{E}$ is reflexive and strictly convex with the strictly convex dual space $\mathcal{E}^*$, then J is a single-valued, one-to-one, and onto mapping. In this case, we can define the single-valued mapping $J^{-1} : \mathcal{E}^* \to \mathcal{E}$ and we have $J^{-1} = J^*$, where J^* is the normalized duality mapping on $\mathcal{E}^*$;

- if $\mathcal{E}$ is uniformly smooth, then J is norm-to-norm uniformly continuous on each bounded subset of $\mathcal{E}$.

Definition 2.1 Let $\mathcal{E}$ be a real normed space. A map $A : \mathcal{E} \to 2^{\mathcal{E}^*}$ is called *monotone* if for each $x, y \in \mathcal{E}$,

$$\langle \eta - \nu, x - y \rangle \geq 0, \quad \forall \eta \in Ax, \nu \in Ay. \tag{2.3}$$

If A is single valued, the map $A : \mathcal{E} \to \mathcal{E}^*$ is called *monotone* if

$$\langle Ax - Ay, x - y \rangle \geq 0, \quad \forall x, y \in \mathcal{E}. \tag{2.4}$$

A multivalued monotone operator $A : \mathcal{E} \to \mathcal{E}^*$ is said to be maximal monotone if $A = B$ whenever $B : \mathcal{E} \to 2^{\mathcal{E}^*}$ is monotone and $G(A) \subset G(B)$, where $G(A) = \{(x, x^*) : x^* \in Ax\}$ is the graph of A.

Let $\mathcal{E}$ be a real reflexive, strictly convex, and smooth Banach space and let $A : \mathcal{E} \to 2^{\mathcal{E}^*}$ be a maximal monotone operator. Then, for each $r > 0$ the resolvent of A, $J_r^A : \mathcal{E} \to \mathcal{E}$ is defined by

$$J_r^A(x) = (J + rA)^{-1} Jx,$$

where J is the normalized duality mapping on $\mathcal{E}$. It is easy to show that $A^{-1}0 = F(J_r^B)$ for all $r > 0$, where $F(J_r^A)$ denotes the set of fixed points of J_r^A. Let $\mathcal{E}$ be a smooth real Banach space with dual $\mathcal{E}^*$. The functional, $\psi : \mathcal{E} \times \mathcal{E} \to \mathbb{R}$, defined by

$$\psi(x, y) := \|x\|^2 - 2\langle x, Jy \rangle + \|y\|^2, \quad \forall x, y \in \mathcal{E}, \tag{2.5}$$

where J is the normalized duality mapping on $\mathcal{E}$ will play a central role in the following. It was introduced by Alber and has been studied by Alber [2], Alber and Guerre-Delabriere [3], Kamimura and Takahashi [18], Reich [25], Chidume *et al.* [13, 14], and a host of other authors.

Lemma 2.2 ([2, 5]) *Let $\mathcal{E}$ be a real uniformly convex, smooth Banach space. Then, the following identities hold*:

(i) $\psi(x, y) = \psi(x, z) + \psi(z, y) + 2\langle x - z, Jz - Jy \rangle, \forall x, y, z \in \mathcal{E}$.

(ii) $\psi(x, y) + \psi(y, x) = 2\langle x - y, Jx - Jy \rangle, \forall x, y \in \mathcal{E}$.

Lemma 2.3 ([5]) *Let $\mathcal{E}$ be a real 2-uniformly convex Banach space. Then, there exists $\mu \geq 1$ such that*

$$\frac{1}{\mu} \|x - y\|^2 \leq \psi(x, y) \quad \forall x, y \in X.$$

Lemma 2.4 ([7]) *Let $A : \mathcal{E} \to 2^{\mathcal{E}^*}$ be a maximal monotone mapping and $B : \mathcal{E} \to \mathcal{E}^*$ be a Lipschitz continuous and monotone mapping. Then, the mapping $A + B$ is a maximal monotone.*

Lemma 2.5 ([4]) *Let $\mathcal{E}$ be a uniformly convex Banach space. Then, the normalized duality mapping, J, is uniformly monotone on every bounded set. That is, for every $R > 0$ and arbitrary $x, y \in \mathcal{E}$ with $\|x\| \leq R$ and $\|y\| \leq R$ there exists a real nonnegative and continuous function $\psi_R : [0,\infty) \to [0,\infty)$ such that $\psi_R(t) > 0$ for $t > 0$, $\psi_R(0) = 0$ and*

$$\langle Jx - Jy, x - y\rangle \geq \psi_R\bigl(\|x - y\|\bigr).$$

Lemma 2.6 ([18]) *Let $\mathcal{E}$ be a uniformly convex and smooth Banach space, and $\{x_n\}$ and $\{y_n\}$ be two sequences of $\mathcal{E}$. If $\lim_{n\to\infty} \psi(x_n, y_n) = 0$ and either $\{x_n\}$ or $\{y_n\}$ is bounded, then $\lim_{n\to\infty} \|x_n - y_n\| = 0$.*

3 Main results

In this section, we state and prove a weak convergence result for the Modified Forward–Reflected–Backward Splitting Method in a 2-uniformly convex uniformly smooth real Banach space. The method does not require the prior knowlegde or an estimate of the Lipschitz constant. In the following, we assume that the solution set $(A + B)^{-1}(0)$ of problem (1.1) is nonempty.

Theorem 3.1 *Let $\mathcal{E}$ be a real 2-uniformly convex uniformly smooth Banach space. Let $A : \mathcal{E} \to 2^{\mathcal{E}^*}$ be a maximal monotone operator and $B : \mathcal{E} \to \mathcal{E}^*$ be monotone and Lipschitz. Let $x_{-1}, x_0 \in \mathcal{E}$ be arbitrary and $\lambda_{-1}, \lambda_0 > 0$. Define the sequence $\{x_n\}$ iteratively by*

$$x_{n+1} = J^A_{\lambda_n} \circ J^{-1}\bigl(Jx_n - \lambda_n Bx_n - \lambda_{n-1}(Bx_n - Bx_{n-1})\bigr), \quad n \geq 0, \tag{3.1}$$

with

$$\lambda_{n+1} := \min\left\{\lambda_n, \frac{\theta\|x_{n+1} - x_n\|}{\|Bx_{n+1} - Bx_n\|}\right\}, \quad \theta \in \left(0, \frac{1}{2\mu}\right), \mu \geq 1.$$

Suppose that $(A + B)^{-1} \neq \emptyset$ and that the duality mapping is weakly sequentially continuous, then the sequence $\{x_n\}$ generated by (3.1) converges weakly to a solution of (1.1).

Proof We first show that the sequence $\{x_n\}$ is bounded. Let $x^* \in (A + B)^{-1}(0)$, so that

$$-Bx^* \in Ax^*. \tag{3.2}$$

From (3.1), we have that

$$\frac{1}{\lambda_n}\bigl(Jx_n - \lambda_n Bx_n - \lambda_{n-1}(Bx_n - Bx_{n-1}) - Jx_{n+1}\bigr) \in Ax_{n+1}. \tag{3.3}$$

Using (3.2) and (3.3) and the monotonicity of A, we obtain

$$\bigl\langle Jx_{n+1} - Jx_n + \lambda_n\bigl(Bx_n - Bx^*\bigr) + \lambda_{n-1}(Bx_n - Bx_{n-1}), x^* - x_{n+1}\bigr\rangle \geq 0. \tag{3.4}$$

By Lemma 2.2(i), we have

$$2\bigl\langle Jx_{n+1} - Jx_n, x^* - x_{n+1}\bigr\rangle = \psi\bigl(x^*, x_n\bigr) - \psi\bigl(x^*, x_{n+1}\bigr) - \psi(x_{n+1}, x_n). \tag{3.5}$$

Also,

$$\langle Bx_n - Bx^*, x^* - x_{n+1}\rangle = \langle Bx_{n+1} - Bx^*, x^* - x_{n+1}\rangle + \langle Bx_n - Bx_{n+1}, x^* - x_{n+1}\rangle \tag{3.6}$$

and

$$\langle Bx_n - Bx_{n-1}, x^* - x_{n+1}\rangle = \langle Bx_n - Bx_{n-1}, x^* - x_n\rangle + \langle Bx_n - Bx_{n-1}, x_n - x_{n+1}\rangle. \tag{3.7}$$

Substituting (3.5), (3.6), and (3.7) into (3.4) we have:

$$\begin{aligned}
&\psi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle \\
&\quad\le \psi(x^*, x_n) + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle \\
&\qquad + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x_n - x_{n+1}\rangle - \psi(x_{n+1}, x_n) \\
&\qquad + 2\lambda_n\langle Bx_{n+1} - Bx^*, x^* - x_{n+1}\rangle.
\end{aligned} \tag{3.8}$$

Using the monotonicity of B on the last term of equation (3.8) and rearranging the equation we have

$$\begin{aligned}
&\psi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle + \psi(x_{n+1}, x_n) \\
&\quad\le \psi(x^*, x_n) + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle \\
&\qquad + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x_n - x_{n+1}\rangle.
\end{aligned} \tag{3.9}$$

Using the definition of λ_n and Lemma 2.3, we have

$$\begin{aligned}
2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x_n - x_{n+1}\rangle &\le 2\lambda_{n-1}\|Bx_n - Bx_{n-1}\|\|x_n - x_{n+1}\| \\
&\le 2\theta\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n-1}\|\|x_n - x_{n+1}\| \\
&\le \theta\frac{\lambda_{n-1}}{\lambda_n}\left(\|x_n - x_{n-1}\|^2 + \|x_n - x_{n+1}\|^2\right) \\
&\le \mu\theta\frac{\lambda_{n-1}}{\lambda_n}\left(\psi(x_n, x_{n-1}) + \psi(x_{n+1}, x_n)\right).
\end{aligned} \tag{3.10}$$

Substituting (3.10) into (3.9), we have

$$\begin{aligned}
&\psi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle + \psi(x_{n+1}, x_n) \\
&\quad\le \mu\theta\frac{\lambda_{n-1}}{\lambda_n}\left(\psi(x_n, x_{n-1}) + \psi(x_{n+1}, x_n)\right) + \psi(x^*, x_n) \\
&\qquad + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle.
\end{aligned} \tag{3.11}$$

Rearranging the above inequality we obtain,

$$\begin{aligned}
&\psi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle + \left(1 - \mu\theta\frac{\lambda_{n-1}}{\lambda_n}\right)\psi(x_{n+1}, x_n) \\
&\quad\le \mu\theta\frac{\lambda_{n-1}}{\lambda_n}\psi(x_n, x_{n-1}) + \psi(x^*, x_n) + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle.
\end{aligned} \tag{3.12}$$

Now, define

$$E_n(x^*) = \psi(x^*, x_n) + \mu\theta \frac{\lambda_{n-1}}{\lambda_n} \psi(x_n, x_{n-1}) + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle. \quad (3.13)$$

Using definition of $E_n(x^*)$ in (3.12), we have

$$E_{n+1}(x^*) \le E_n(x^*) - \left(1 - \mu\theta \frac{\lambda_{n-1}}{\lambda_n} - \mu\theta \frac{\lambda_n}{\lambda_{n+1}}\right) \psi(x_{n+1}, x_n). \quad (3.14)$$

Let $\delta \in (0, 1 - 2\mu\theta)$ be fixed, since $\lambda_n \to \lambda > 0$, we derive

$$\lim_{n\to\infty} \left(1 - \mu\theta \frac{\lambda_{n-1}}{\lambda_n} - \mu\theta \frac{\lambda_n}{\lambda_{n+1}}\right) = 1 - 2\mu\theta > \delta.$$

Thus, there exists $n_1 \ge 1$ such that

$$1 - \mu\theta \frac{\lambda_{n-1}}{\lambda_n} - \mu\theta \frac{\lambda_n}{\lambda_{n+1}} \ge \delta, \quad \forall n \ge n_1. \quad (3.15)$$

It follows from (3.14) and (3.15) that

$$E_{n+1}(x^*) \le E_n(x^*) - \delta\psi(x_{n+1}, x_n) \le E_n(x^*) \quad \forall n \ge n_1. \quad (3.16)$$

Therefore, the sequence $\{E_n\}_{n\ge n_1}$ is nonincreasing.

Now, from the definition of E_n and λ_n for each $n \ge n_1$ we see that

$$\begin{aligned}
E_n(x^*) &= \psi(x^*, x_n) + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle + \mu\theta \frac{\lambda_{n-1}}{\lambda_n} \psi(x_n, x_{n-1}) \\
&\ge \psi(x^*, x_n) - 2\lambda_{n-1}\|Bx_n - Bx_{n-1}\| \|x^* - x_n\| + \mu\theta \frac{\lambda_{n-1}}{\lambda_n} \psi(x_n, x_{n-1}) \\
&\ge \psi(x^*, x_n) - 2\theta \frac{\lambda_{n-1}}{\lambda_n} \|x_n - x_{n-1}\| \|x^* - x_n\| + \mu\theta \frac{\lambda_{n-1}}{\lambda_n} \psi(x_n, x_{n-1}) \\
&\ge \psi(x^*, x_n) - \theta \frac{\lambda_{n-1}}{\lambda_n} \left(\|x_n - x_{n-1}\|^2 + \|x^* - x_n\|^2\right) + \mu\theta \frac{\lambda_{n-1}}{\lambda_n} \psi(x_n, x_{n-1}) \\
&\ge \psi(x^*, x_n) - \mu\theta \frac{\lambda_{n-1}}{\lambda_n} \left(\psi(x_n, x_{n-1}) + \psi(x^*, x_n)\right) + \mu\theta \frac{\lambda_{n-1}}{\lambda_n} \psi(x_n, x_{n-1}) \\
&= \left(1 - \mu\theta \frac{\lambda_{n-1}}{\lambda_n}\right) \psi(x^*, x_n) \\
&\ge \left(1 - \mu\theta \frac{\lambda_{n-1}}{\lambda_n} - \mu\theta \frac{\lambda_n}{\lambda_{n+1}}\right) \psi(x^*, x_n) \\
&\ge \delta\psi(x^*, x_n) \ge 0.
\end{aligned}$$

Thus, the limit $\lim_{n\to\infty} E_n$ exists.

Also, the boundedness of $\{\psi(x^*, x_n)\}$ implies that $\{x_n\}$ is bounded. Moreover, from (3.16), we have by telescoping, that

$$E_{n+1}(x^*) \le E_{n_1}(x^*) - \delta \sum_{n=n_1}^{\infty} \psi(x_{n+1}, x_n). \quad (3.17)$$

That is,

$$\delta \sum_{n=n_1}^{\infty} \psi(x_{n+1},x_n) \leq E_{n_1}(x^*) - \lim_{n\to\infty} E_{n+1}(x^*) < +\infty.$$

Hence, the limit $\lim_{n\to\infty} \psi(x_{n+1},x_n)$ exists. Since B is Lipschitz continuous, $\{x_n\}$ is bounded, $\lambda_n \to \lambda > 0$, then from (3.17) and Lemma 2.6, we obtain that

$$\lim_{n\to\infty} \left(2\lambda\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle + \mu\theta \frac{\lambda_{n-1}}{\lambda_n} \psi(x_n,x_{n-1}) \right) = 0.$$

Using the definition of E_n, we have

$$\lim_{n\to\infty} E_n(x^*) = \lim_{n\to\infty} \psi(x^*,x_n).$$

That is, the limit of $\psi(x^*,x_n)$ exists for each $x^* \in (A+B)^{-1}(0)$.

We now prove that $\{x_n\}$ converges weakly to an element of $(A+B)^{-1}(0)$. Let ρ be a weak cluster point of $\{x_n\}$. Then, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \rightharpoonup \rho$. We show that $\rho \in (A+B)^{-1}(0)$.

From the definition of x_n in (3.1), we have

$$\frac{1}{\lambda_n}(Jx_n - Jx_{n+1}) + (Bx_{n+1} - Bx_n) - \frac{\lambda_{n-1}}{\lambda_n}(Bx_n - Bx_{n-1}) \in (A+B)x_{n+1.} \tag{3.18}$$

Since, by Lemma 2.4, A + B is maximal monotone, then we have that its graph is demi-closed. Now, passing the limit in (3.18) we obtain that

$$0 \in (A+B)(\rho).$$

Next, we show that the whole sequence $\{x_n\}$ converges weakly to ρ.

Suppose there exists ρ' such that $x_{n_j} \rightharpoonup \rho'$ for some subsequence $\{x_{n_j}\}$ of $\{x_n\}$ with $\rho' \neq \rho$. Then, we have

$$\psi(\rho,x_n) = \|\rho\|^2 - 2\langle\rho, Jx_n\rangle + \|x_n\|^2,$$

and

$$\psi(\rho',x_n) = \|\rho'\|^2 - 2\langle\rho', Jx_n\rangle + \|x_n\|^2.$$

Thus, we have

$$2\langle\rho' - \rho, Jx_n\rangle = \psi(\rho,x_n) - \psi(\rho',x_n) + \|\rho'\|^2 - \|\rho\|^2.$$

Hence, the limit $\lim_{n\to\infty}\langle\rho' - \rho, Jx_n\rangle$ exists. Since J is weakly sequentially continuous, we have

$$\langle\rho' - \rho, J\rho\rangle = \lim_{k\to\infty}\langle\rho' - \rho, Jx_{n_k}\rangle = \lim_{j\to\infty}\langle\rho' - \rho, Jx_{n_j}\rangle = \langle\rho' - \rho, J\rho'\rangle.$$

Using Lemma 2.5, we have that $\rho' = \rho$. Hence $\{x_n\}$ converges weakly to ρ. □

We now state Theorem 3.1 in Hilbert spaces.

Corollary 3.2 *Let $\mathcal{H}$ be a real Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximal monotone operator and $B : \mathcal{H} \to \mathcal{H}$ be monotone and Lipschitz. Choose $x_{-1}, x_0 \in \mathcal{H}, \lambda_{-1}, \lambda_0 > 0$. Let $\{x_n\}$ be the sequence defined by*

$$\begin{cases} x_{n+1} = J^A_{\lambda_n}(x_n - \lambda_n Bx_n - \lambda_{n-1}(Bx_n - Bx_{n-1})), & n \geq 0, \\ \lambda_{n+1} := \min\{\lambda_n, \frac{\theta\|x_{n+1}-x_n\|}{\|Bx_{n+1}-Bx_n\|}\}, & \theta \in (0, \frac{1}{2}). \end{cases}$$

Suppose $(A + B)^{-1}(0) \neq \emptyset$. Then, the sequence $\{x_n\}$ converges weakly to an element of $(A + B)^{-1}(0)$.

4 Numerical examples in infinite-dimensional spaces

In this section, we compare Algorithm (3.1) with FBFSM and FRBSM introduced in [28] and [21], respectively. For easy referencing, we term FBFSM and FRBSM as TSENG and TAM, respectively. Numerical experiments were carried out on MATLAB R2015a version. All programs were run on a 64-bit OS PC with an Intel(R) Core(TM) i7-3540M CPU @ 1.00 GHz, 1.19 GHz and 3 GB RAM. All figures were plotted using the log log plot command.

Example 1 Let $\mathcal{H} = L_2([0, 1])$, with the norm and inner product defined as

$$\|x\|_2 = \left(\int_0^1 |x(t)|^2 \, dt\right)^{\frac{1}{2}} \quad \text{and} \quad \langle x, y\rangle = \int_0^1 x(t)y(t)\, dt, \quad \text{respectively.}$$

Define the operator $B : \mathcal{H} \to \mathcal{H}$ by

$$Bx(t) = \int_0^1 \left[x(s) - \left(\frac{2tse^{t+s}}{e\sqrt{e^2 - 1}}\right)\cos x(s)\right] ds + \frac{2te^t}{e\sqrt{e^2 - 1}}, \quad x \in L_2\big([0, 1]\big),$$

then, B is monotone and Lipschitz with Lipschitz constant $L = 2$. Let $A : \mathcal{L}_2([0, 1]) \to \mathcal{L}_2([0, 1])$ be defined by

$$Ax(t) = \max\big\{x(t), 0\big\},$$

then, A is maximal monotone and for any $r > 0$, the resolvent, $J^A_r : \mathcal{L}_2([0, 1]) \to \mathcal{L}_2([0, 1])$, of A, is given by

$$J^A_r x(t) = \begin{cases} x(t), & Ax(t) = 0, \\ \frac{1}{1+r}x(t), & Ax(t) = x(t). \end{cases}$$

Clearly,

$$0 \in (A + B)^{-1}(0).$$

Table 1 Computational Results for Example 1

Algorithms	Tolerance (TOL)	$(\theta, \lambda_n, \lambda)$	No. of Iter.	Time (s)
Algorithm (3.1)	$TOL = 10^{-4}$	$\theta = 0.4$	40	16.9233
Tam		$\lambda_n = 0.125 + (0.01562)n^{-1}$	41	12.3243
Tseng		$\lambda = 0.1406$	400	134.3984
Algorithm (3.1)	$TOL = 10^{-4}$	$\theta = 0.4$	46	16.0526
Tam		$\lambda_n = 0.01 + (0.2)n^{-1}$	329	113.968
Tseng		$\lambda = 0.2100$	400	82.4403
Algorithm (3.1)	$TOL = 9 \times 10^{-3}$	$\theta = 0.01$	238	115.2452
Tam		$\lambda_n = 0.125 + (0.01562)n^{-1}$	21	9.2270
Tseng		$\lambda = 0.1406$	400	125.1663
Algorithm (3.1)	$TOL = 9 \times 10^{-3}$	$\theta = 0.01$	244	107.1992
Tam		$\lambda_n = 0.01 + (0.2)n^{-1}$	122	52.6718
Tseng		$\lambda = 0.2100$	400	94.7951

We show that $x_n \rightharpoonup 0$. We recall that the sequence $\{x_n\}$ converges weakly to 0 in $\mathcal{L}_2([0,1])$ if and only if

$$\langle \varphi, x_n \rangle = \int_0^1 \varphi(t) x_n(t)\, dt \to 0 \quad \text{as } n \to \infty$$

for any $\psi \in \mathcal{H}^*$. We conduct the experiment with various functions ψ in $\mathcal{L}_2([0,1])$. The integrals were approximated using the *trapz* and *int* command on MATLAB over the interval $[0,1]$. The results of the experiment are displayed in Table 1 and Figs. 1, 2, 3, and 4.

Example 2 Let $\mathcal{H} = L_2([0,1])$, with the norm and inner product defined as

$$\|x\|_2 = \left(\int_0^1 \left| x(t) \right|^2 dt \right)^{\frac{1}{2}} \quad \text{and} \quad \langle x, y \rangle = \int_0^1 x(t) y(t)\, dt, \quad \text{respectively.}$$

We inherit the map A from (1) above, while the map B is defined by

$$Bx(t) = \frac{x(t) + |x(t)|}{2}.$$

Clearly, B is monotone and Lipschitz and

$$0 \in (A + B)^{-1}(0).$$

We show that $x_n \rightharpoonup 0$ just as in Example 1 above. The results of the experiment are displayed in Table 2 and Figs. 5, 6, 7, and 8.

Remark 1 From the results displayed in Tables 1 and 2 it is clear that the speed of the convergence of Algorithm (3.1) heavily depends on the value of θ. For instance, Algorithm (3.1) converges faster as the value of θ moves closer to 0.5. Thus, if the value of θ is appropriately chosen Algorithm (3.1), seems to have cheaper computations compared to its counterparts. On the other hand, the Algorithm TAM depends on the step size $\{\lambda_n\}$, while that of TSENG depends on λ. The algorithm converges faster when the step sizes are chosen very close to the upper bound of the interval of choice. Finally, we note that the number

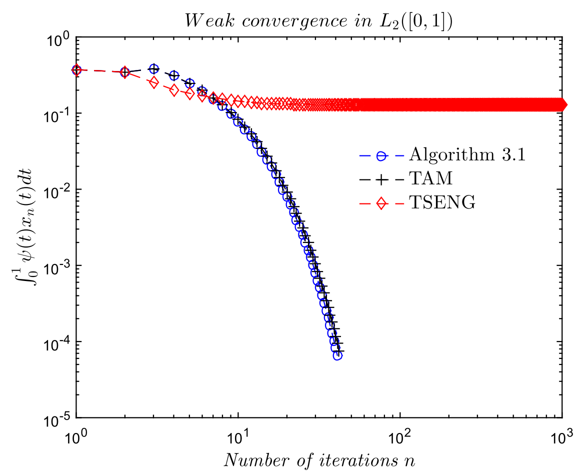

Figure 1 Example 1 with $\theta = 0.4$ and $\lambda_n = 0.1250 + (0.0156)n^{-1}$

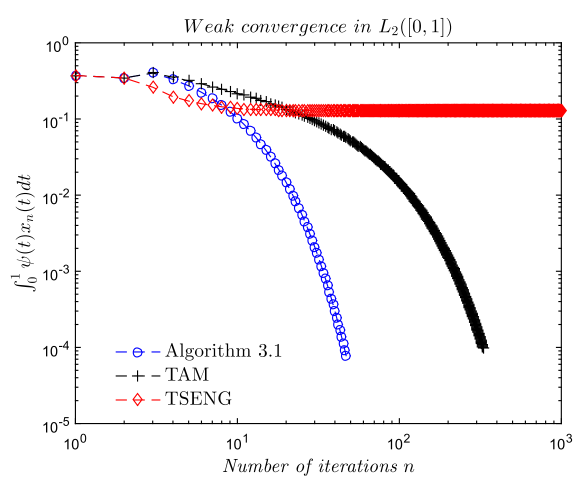

Figure 2 Example 1 with $\theta = 0.4$ and $\lambda_n = 0.01 + (0.2)n^{-1}$

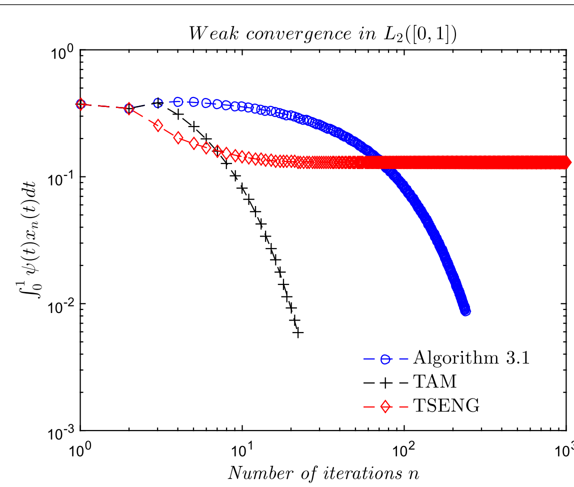

Figure 3 Example 1 with $\theta = 0.01$ and $\lambda_n = 0.1250 + (0.0156)n^{-1}$

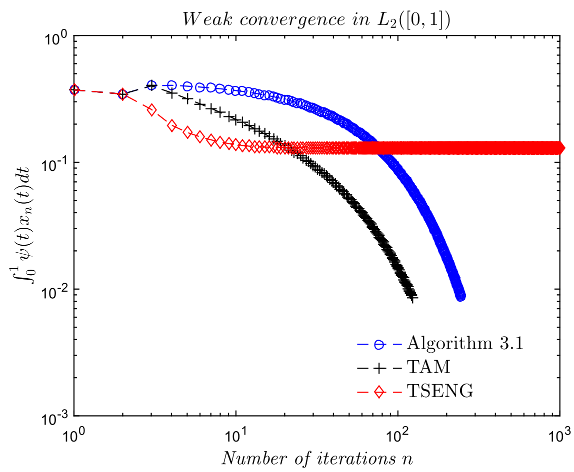

Figure 4 Example 1 with $\theta = 0.01$ and $\lambda_n = 0.01 + (0.2)n^{-1}$

Table 2 Computational Results for Example 2

Algorithms	Tolerance (TOL)	$(\theta, \lambda_n, \lambda)$	No. of Iter.	Time (s)
Algorithm (3.1)	TOL = 10^{-5}	$\theta = 0.4$	50	0.0693
Tam		$\lambda_n = 0.125 + (0.01562)n^{-1}$	52	0.0027
Tseng		$\lambda = 0.1406$	45	0.0332
Algorithm (3.1)	TOL = 10^{-5}	$\theta = 0.4$	58	0.0602
Tam		$\lambda_n = 0.01 + (0.2)n^{-1}$	440	0.0222
Tseng		$\lambda = 0.2100$	34	0.0328
Algorithm (3.1)	TOL = 10^{-5}	$\theta = 0.01$	538	0.1166
Tam		$\lambda_n = 0.125 + (0.01562)n^{-1}$	52	0.0043
Tseng		$\lambda = 0.1406$	45	0.0346
Algorithm (3.1)	TOL = 10^{-5}	$\theta = 0.01$	540	0.1130
Tam		$\lambda_n = 0.01 + (0.2)n^{-1}$	440	0.0226
Tseng		$\lambda = 0.2100$	34	0.0336

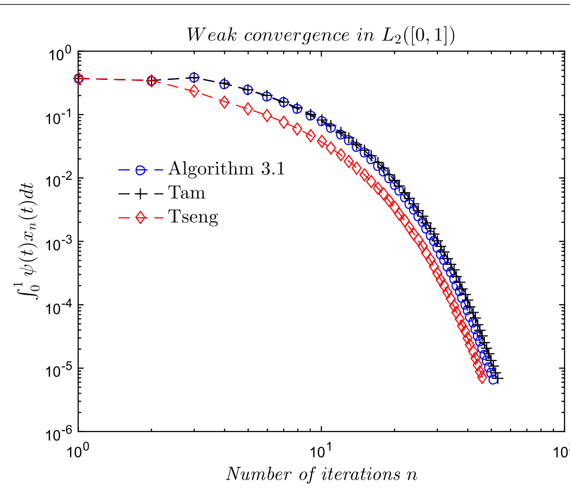

Figure 5 Example 2 with $\theta = 0.4$ and $\lambda_n = 0.1250 + (0.0156)n^{-1}$

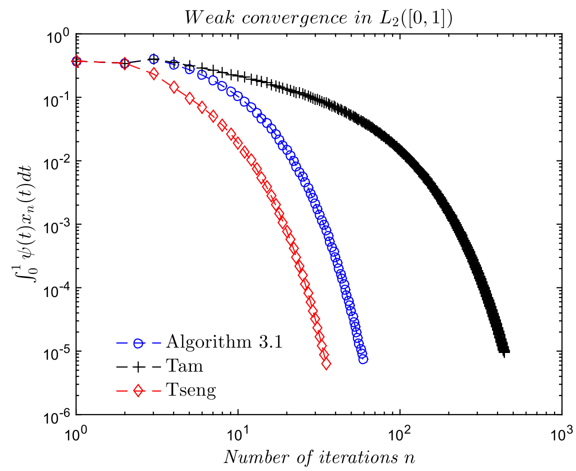

Figure 6 Example 2 with $\theta = 0.4$ and $\lambda_n = 0.01 + (0.2)n^{-1}$

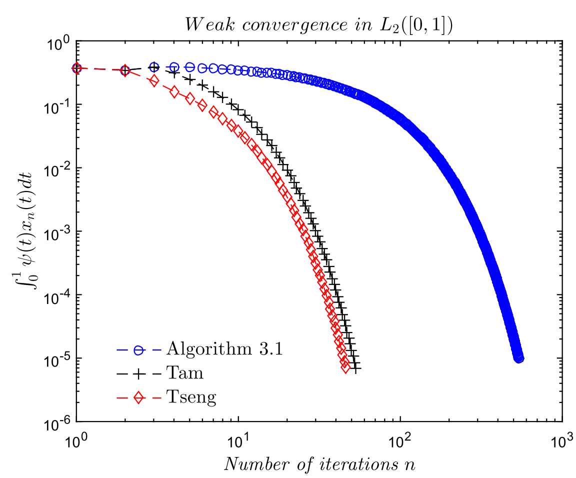

Figure 7 Example 2 with $\theta = 0.01$ and $\lambda_n = 0.1250 + (0.0156)n^{-1}$

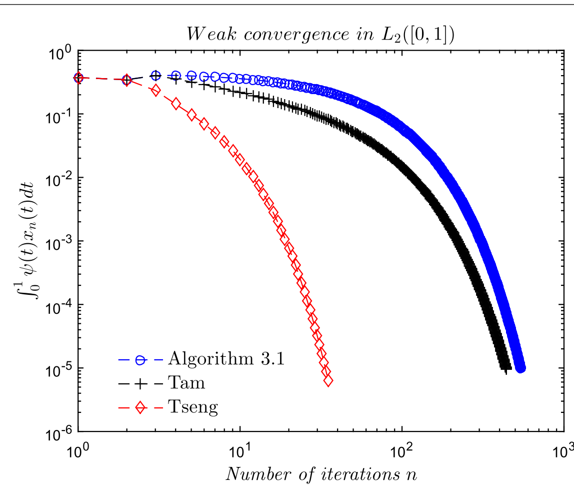

Figure 8 Example 2 with $\theta = 0.01$ and $\lambda_n = 0.01 + (0.2)n^{-1}$

of iterations for TSENG in Table 1 was cut short due to the large number iterations needed before the tolerance is reached.

5 Conclusion

In this work, we have proved the weak convergence of a one-step self-adaptive algorithm to a solution of the sum of two monotone operators in 2-uniformly convex and uniformly smooth Banach spaces. Numerical results were presented to illustrate how Algorithm (3.1) competes with some existing algorithms. Finally, our results generalize and complement some existing results in the literature.

Acknowledgements
The authors appreciate the support of their institution and AfDB.

Funding
This work is supported from AfDB Research Grant Funds to AUST.

Availability of data and materials
Data sharing is not applicable to this article.

Declarations

Competing interests
The authors declare no competing interests.

Author contribution
The problem was formulated by AUB and the computations and proofs were carried out jointly by CEC and MA. All authors have read and agreed the final manuscript.

Author details
[1] African University of Science and Technology, Km 10 Airport Road, Abuja, Nigeria. [2] Federal University Dutsinma, Katsina, Nigeria. [3] Planning Research and Statistics Department, Federal Inland Revenue Service, FCT, Abuja, Nigeria.

Publisher's Note

Received: 11 October 2021 Accepted: 21 September 2022 Published online: 06 December 2022

References

1. Abass, H.A., Aremu, K.O., Jolaoso, L.O., Mewomo, O.T.: An inertial forward–backward splitting method for approximating solutions of certain optimization problems. J. Nonlinear Funct. Anal. **2020**, Article ID 6 (2020)
2. Alber, Y.: Metric and generalized projection operators in Banach spaces: properties and applications. In: Theory and Applications of Nonlinear Operators of Accretive and Monotone Type, Lecture Notes in Pure and Appl. Math., vol. 178, pp. 15–50. Dekker, New York (1996)
3. Alber, Y., Guerre-Delabriere, S.: On the projections for fixed points problems. Analysis **21**(1), 17–39 (2001)
4. Alber, Y., Ryazantseva, I.: Nonlinear Ill Posed Problems of Monotone Type. Springer, London (2006)
5. Aoyama, K., Kohsaka, F.: Strongly relatively nonexpansive sequences generated by firmly nonexpansive-like mappings. Fixed Point Theory Appl. **2014**, Article ID 95 (2014)
6. Attouch, H., Peypouquet, J., Redont, P.: Backward–forward algorithms for structured monotone inclusions in Hilbert spaces. J. Math. Anal. Appl. **457**, 1095–1117 (2018)
7. Barbu, V.: Nonlinear Semigroups and Differential Equations in Banach Spaces. Editura Academiei R.S.R, Bucharest (1976)
8. Beauzamy, B.: Introduction to Banach Spaces and Their Geometry, 2nd edn. North-Holland Mathematics Studies, vol. 68, p. xv+338 North-Holland, Amsterdam (1985). ISBN 0-444-87878-5. Notas de Matematica [Mathematical Notes], 86
9. Bredies, K.: A forward–backward splitting algorithm for the minimization of non-smooth convex functionals in Banach space. Inverse Probl. **25**(1), 015005 (2009)
10. Bruck, R.: On the weak convergence of an ergodic iteration for the solution of variational inequalities for monotone operators in Hilbert space. J. Math. Anal. Appl. **61**, 159–164 (1977)
11. Censor, Y., Elfving, T.: A multi-projection algorithm using Bregman projections in a product space. Numer. Algorithms **8**, 221–239 (1994)
12. Chen, G.H.G., Rockafellar, R.T.: Convergence rates in forward–backward splitting. SIAM J. Optim. **7**(2), 421–444 (1997)
13. Chidume, C.E., Bello, A.U., Usman, B.: Iterative Algorithms for Zeros of Strongly Monotone Lipschitz Maps in Classical Banach Spaces p. 9. Springer, Berlin (2015). https://doi.org/10.1186/s40064-015-1044-1

14. Chidume, C.E., Chidume, C.O., Bello, A.U.: An algorithm for zeros of generalized phi-strongly monotone and bounded maps in classical Banach spaces. Optimization (2015). https://doi.org/10.1080/02331934.2015.1074686
15. Combettes, P., Wajs, V.R.: Signal recovery by proximal forward–backward splitting. Multiscale Model. Simul. **4**(4), 1168–1200 (2005)
16. Davis, D., Yin, W.T.: A three-operator splitting scheme and its optimization applications. Set-Valued Var. Anal. **25**, 829–858 (2017)
17. Hieu, D.V., Anh, P.K., Muu, L.D.: Modified Forward Reflected Backward Spliting Method for Variational Inclusions. Springer, Berlin (2020). https://doi.org/10.1007/s10288-020-00440-3
18. Kamimura, S., Takahashi, W.: Strong convergence of a proximal-type algorithm in a Banach space. SIAM J. Optim. **13**(3), 938–945 (2003)
19. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**, 964–979 (1979)
20. Liu, L.: Strong convergence of a modified inertial forward–backward splitting algorithm for a inclusion problem. J. Appl. Numer. Optim. **2**, 373–385 (2020)
21. Malitsky, Y., Tam, M.K.: A forward–backward spliting method for monotone inclusion without cocoercivity. SIAM J. Optim. **30**(2), 1451–1472 (2020)
22. Moudafi, A., Thera, M.: Finding a zero of the sum of two maximal monotone operators. J. Optim. Theory Appl. **94**, 425–448 (1997)
23. Passty, G.B.: Ergodic convergence to a zero of the sum of monotone operators in Hilbert spaces. J. Math. Anal. Appl. **72**, 383–390 (1979)
24. Peaceman, D.H., Rachford, H.H.: The numerical solutions of parabolic and elliptic differential equations. J. Soc. Ind. Appl. Math. **3**, 28–41 (1955)
25. Reich, S.: A weak convergence theorem for the alternating method with Bregman distances. In: Kartsatos, A.G. (ed.) Theory and Applications of Nonlinear Operators of Accretive and Monotone Type. Lecture Notes Pure Appl. Math., vol. 178, pp. 313–318. Dekker, New York (1996)
26. Shehu, Y.: Convergence results of forward–backward algorithm for sum of monotone operator in Banach spaces. Results Math. **74**, 138 (2019)
27. Takahashi, W.: Nonlinear Functional Analysis. Yokohama Publishers, Yokohama (2000)
28. Tseng, P.: A modified forward–backward splitting method for maximal monotone mappings. SIAM J. Control Optim. **38**, 431–446 (2000)
29. Tuyen, T.M., Promkam, R., Sunthrayuth, P.: Strong convergence of a generalized forward–backward splitting method in reflexive Banach spaces. Optimization **71**(6), 1483–1508(2020). https://doi.org/10.1080/02331934.2020.1812607
30. Wang, Y., Xu, H.K.: Strong convergence for the proximal-gradient method. J. Nonlinear Convex Anal. **15**(3), 581–593 (2014)