Optimization
A Journal of Mathematical Programming and Operations Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gopt20

Forward-reflected-backward splitting method without cocoercivity for the sum of maximal monotone operators in Banach spaces

A. U. Bello, C. C. Okeke, M. Isyaku & M. T. Omojola

To cite this article: A. U. Bello, C. C. Okeke, M. Isyaku & M. T. Omojola (2023) Forward-reflected-backward splitting method without cocoercivity for the sum of maximal monotone operators in Banach spaces, Optimization, 72:9, 2201-2222, DOI: 10.1080/02331934.2022.2057308

To link to this article: https://doi.org/10.1080/02331934.2022.2057308

Published online: 29 Mar 2022.

OPTIMIZATION
2023, VOL. 72, NO. 9, 2201–2222
https://doi.org/10.1080/02331934.2022.2057308

Forward-reflected-backward splitting method without cocoercivity for the sum of maximal monotone operators in Banach spaces

A. U. Bello [a,b], C. C. Okeke [c], M. Isyaku [b] and M. T. Omojola [a]

[a] Mathematics Institute African University of Science and Technology, Abuja, Nigeria; [b] Department of Mathematics, Federal University Dutsinma, Dutsin-Ma, Kastina State Nigeria; [c] School of Mathematics, University of the Witwatersrand, Johannesburg, South Africa

ABSTRACT
Let E be a real 2-uniformly convex Banach space with topological dual E^*. We prove weak convergence of the forward-reflected-backward splitting method to a solution of inclusion of sum of two monotone operators. Moreover, we give a variant of the method in which the step size does not depend on the Lipschitz constant of one of the operators. Finally, our results extend and complement several existing results in the literature. We also give some numerical illustrations of the proposed method in comparison with other method in the literature to further demonstrate the applicability and efficiency of our method.

ARTICLE HISTORY
Received 10 February 2021
Accepted 19 March 2022

KEYWORDS
Maximally monotone operators; Lipschitz-continuous operator; forward-reflected-backward splitting method; 2-uniformly convex spaces

2000 MATHEMATICS SUBJECT CLASSIFICATIONS
47H09; 47H10; 49J20; 49J40

1. Introduction

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

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

with $(A+B)^{-1}(0) \neq \emptyset$, where $A : E \to 2^{E^*}$ is a maximally monotone operator and $B : E \to E^*$ is a monotone and Lipschitz map. Interest in problem (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, [1–10] for more treatments of problem (1). Consider, for instance, the split feasibility problem, introduced by Censor and Elfvin [3], which

CONTACT C. C. Okeke chibueze.okeke87@yahoo.com

is to find

$$x \in C \quad \text{such that } Tx \in Q, \tag{2}$$

where $C \subset H_1, Q \subset H_2$ are non-empty closed and convex subsets of the Hilbert spaces H_1 and H_2, respectively, and $T : H_1 \to H_2$ is a bounded linear map. By setting

$$A(x) = N_C(x) \quad \text{and} \quad B(x) = \nabla\left(\frac{1}{2}\left\|Tx - P_Q Tx\right\|^2\right) = T^*(I - P_Q)Tx.$$

Then, (2) can be reformulated as (1).

Consider also the following problem in quantum mechanics which is formulated in a more general Banach space, see e.g. [11]. For $p > 1$, define the operator T by

$$Tu = a^2 \Delta u + (g(x) + b)u(x) + |u(x)|^{p-2}u(x)\int_{\mathbb{R}^3} \frac{|u(y)|^p}{\|x-y\|}\mathrm{d}y, \tag{3}$$

where $\Delta = \sum_{i=1}^{3} \frac{\delta^2}{\delta x_i^2}$ is the Laplace operator (Laplacian) in $\mathbb{R}^3$, a and b are real constants and $g = g_0 + g_1$, with $g_0 \in L^\infty(\mathbb{R}^3)$ and $g_1 \in L^p(\mathbb{R}^3)$. The operator T can be decomposed into the sum of two operators A and B with A as the linear part of T called the Schrödinger operator and B is the last term. It is well known that there exists $b \geq 0$ such that the operator A becomes positive. Also the operator B is the gradient of the proper convex and lower semi-continuous functional $\Phi : L^p(\mathbb{R}^3) \to L^q(\mathbb{R}^3)$ defined by

$$\Phi(u) = \frac{1}{2p}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{|u(x)|^p|u(y)|^p}{\|x-y\|}\mathrm{d}y, \tag{4}$$

where $\frac{1}{p} + \frac{1}{q} = 1$. Consequently, the operator B is also monotone. It is well known that for $p = 2$, then $g(x) = -2|x|^{-2}$ and $a = \frac{1}{2}$ correspond to the case of the coulomb potential in quantum mechanics. The operator T here describes the atom of helium in the situation when both electrons are in the same state. The zeros of the operator T corresponds to the zeros of the sum $A + B$ which is exactly problem (1). There are several methods of approximating solutions of (1), see, e.g. [1, 2, 6, 7, 9]. One of the simplest and efficient method is the *forward–backward* splitting method introduced by Passty [9], and Lions and Mercier [7]. The method generates a sequence $\{x_n\}$ iteratively defined by

$$x_{n+1} = J_{\lambda_n}^A(x_n - \lambda_n Bx_n), \quad n \geq 0. \tag{5}$$

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

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

and $\liminf \lambda_n > 0$ with $\limsup \lambda_n < 2\mu$, then the sequence $\{x_n\}$ generated by (5) converges weakly to a solution of (1). The coercivity 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 application where the forward–backward splitting method fails to converge due to the lack of cocoercivity of one of the operator. For instance, the first-order optimality condition for the saddle-points problems of the form

$$\min_{x\in H_1}\max_{y\in H_2}(f(x)+\Phi(x,y)-g(y)), \tag{6}$$

where $f : H_1 \to \mathbb{R}\cup\{+\infty\}$ and $g : H_2 \to \mathbb{R}\cup\{+\infty\}$ are proper convex and lower semi-continuous functions and $\Phi : H_1 \times H_2 \to \mathbb{R}$ is a smooth convex–concave function, is given by (1) with

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

Problem (6) 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 [12], Tseng introduced the *forward–backward–forward* splitting method for approximating solutions of (1). The method generates a sequence $\{x_n\}$ iteratively defined by

$$\begin{cases} y_n = J^A_{\lambda_n}(x_n-\lambda_n B(x_n))\\ x_{n+1}=y_n-\lambda_n B(y_n)+\lambda_n B(x_n) \quad \forall\, n\in\mathbb{N}\end{cases} \tag{7}$$

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. Very recently, Malitsky and Tam [13], introduced the *forward-reflected-backward* splitting method generated iteratively by

$$x_{n+1}=J^A_{\lambda_n}(x_n-\lambda_n B(x_n)-\lambda_{n-1}(B(x_n)-B(x_{n-1}))\quad \forall\, n\in\mathbb{N}, \tag{8}$$

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 forward–backward–forward method which requires two evaluations of the operator B per iteration.

It is worth mentioning that all the results mentioned above are done in the settings of Hilbert spaces. There are few results regarding the forward–backward method and its variants in Banach spaces, see, e.g. [14, 15]. As far as we know, there is no any result concerning forward-reflected-backward splitting in the setting of Banach spaces. One of the difficulties, perhaps, is the fact that the operators A and B go from the Banach space E to its dual E^*. The tools available in Hilbert spaces are not readily available in general Banach spaces. In this paper, we prove weak convergence of the forward-reflected-backward splitting method in real 2-uniformly convex Banach spaces. 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 would be used in the proof of our main results. Let E be a real normed linear space. Let S_E and B_E denote the unit sphere and the closed unit ball of E, respectively. The modulus of smoothness of $E \rho_E : [0,\infty) \to [0,\infty)$ is defined by

$$\rho_E(t) := \sup\left\{\frac{\|x+y\|+\|x-y\|}{2} - 1 : x \in S_E, \|y\| = t\right\}.$$

A real Banach space E is said to be 2-uniformly smooth, if there exists a fixed constant $c > 0$ such that $\rho_E(t) \le ct^2$. The space E is said to be smooth if

$$\lim_{t\to 0} \frac{\|x+ty\| - \|x\|}{t} \tag{9}$$

exists for all $x, y \in S_E$. The space E is also said to be uniformly smooth if (9) converges uniformly in $x, y \in S_E$. It is well known that every 2-uniformly smooth space is uniformly smooth. A real normed space E is said to be strictly convex if

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

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

$$\delta_E(\epsilon) := \inf\left\{1 - \left\|\frac{x+y}{2}\right\| \;:\; x, y \in B_E,\ \|x-y\| \ge \epsilon\right\} \tag{10}$$

for all $\epsilon \in [0, 2]$. The space E is said to be 2-uniformly convex if there exists $c > 0$ such that $\delta_E(\epsilon) \ge c\epsilon^2$ 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 $1 < p \le \infty$, and 2-uniformly convex whenever $1 < p \le 2$ (see [16]).

Let E be a real normed space. The normalized duality mapping of E into E^* is defined by

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

for all $x \in E$. The normalized duality mapping J has the following properties (see, e.g. [17]):

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

- If E is uniformly smooth, then J is uniformly norm-to-norm continues on each bounded subset of E.

Recall that a map T is said to be weakly sequential continuous if and only if for any sequence $x_n \in E$ such that x_n converges weakly to x we have Jx_n converges to Jx. The following example provides an infinite-dimensional non-Hilbert space for which the normalized duality mapping is weakly sequential continuous.

Example 2.1 (see [18]): Consider the space

$$X = H + V$$

endowed with the norm

$$\|x\| = \sqrt{\|h\|^2 + \|v\|^2}, \quad x = (h, v), \quad h \in H, \quad v \in V,$$

where H is a Hilbert space and V is a finite-dimensional smooth normed space with a norm not induced by an inner product. Then X is non-Hilbert; however, its normalized duality map is weakly continuous.

For more information on normalized duality mapping, please see ([18] and the references therein).

Definition 2.2: Let E be a real normed space with dual E^*. A map $A : E \to 2^{E^*}$ is called *monotone* if for each $x, y \in E$,

$$\langle \eta - v, x - y \rangle \geq 0, \quad \forall\, \eta \in Ax, \quad v \in Ay. \tag{11}$$

If A is single-valued, the map $A : E \to E$ is called *monotone* if

$$\langle Ax - Ay, x - y \rangle \geq 0, \quad \forall\, x, y \in E. \tag{12}$$

Then the mapping $A : E \to E^*$ is called

- γ-strongly monotone if there exists $\gamma \in (0, 1)$ such that

$$\langle Ax - Ay, x - y \rangle \geq \gamma \|x - y\|^2, \quad \forall\, x, y \in E.$$

- Cocoercive if there exists $\beta > 0$ such that if

$$\langle Ax - Ay, x - y \rangle \geq \beta \|Ax - Ay\|^2 \quad \forall\, x, y \in X;$$

- Lipschitz continuous if there exists a constant $L > 0$ such that

$$\|Ax - Ay\| \leq L\|x - y\| \quad \forall\, x, y \in E.$$

It can be easily seen that every cocoercive operator is Lipschitz.

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

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

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

where J is the normalized duality mapping on 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 E be a smooth real Banach space with dual E^*. The functional, $\phi : E \times E \to \mathbb{R}$, defined by

$$\phi(x, y) := \|x\|^2 - 2\langle x, Jy\rangle + \|y\|^2, \quad \forall\, x, y \in E, \tag{13}$$

where J is the normalized duality mapping on E will play a central role in the sequel. It was introduced by Alber and has been studied by Alber [19], Alber and Guerre-Delabriere [20], Kamimura and Takahashi [21], Reich [22], Chidume et al. [23, 24] and a host of other authors.

Lemma 2.3 ([19, 25]): *Let E be a real uniformly convex, smooth Banach space. Then, the following identities hold:*

(i) $\phi(x, y) = \phi(x, z) + \phi(z, y) + 2\langle x - z, Jz - Jy\rangle, \ \forall\, x, y, z \in E.$
(ii) $\phi(x, y) + \phi(y, x) = 2\langle x - y, Jx - Jy\rangle, \ \forall\, x, y \in E.$

Lemma 2.4 ([25]): *Let E be a real 2-uniformly convex Banach space. Then, there exists $\mu \geq 1$ such that*

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

Lemma 2.5 ([26]): *Let $A : E \to 2^{E^*}$ be a maximally monotone mapping and $B : E \to E^*$ be a Lipschitz continuous and monotone mapping. Then the mapping $A + B$ is a maximally monotone.*

Lemma 2.6 ([11]): *Let 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 $x, y \in E$ with $\|x\| \leq R$ and $\|y\| \leq R$ there exists a real non-negative and continuous function $\psi_R : [0, \infty) \to [0, \infty)$ such that $\psi_R(t) > 0$ for $t > 0$, $\psi_R(0) = 0$ and*

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

Lemma 2.7 (**[21]**)**:** *Let E be a uniformly convex and smooth Banach space, $\{x_n\}$ and $\{y_n\}$ be two sequences of E. If $\lim_{n\to\infty} \phi(x_n, y_n) = 0$ 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 a Forward–Backward Splitting Method in real 2-uniformly convex Banach space.

Theorem 3.1: *Let E be a real 2-uniformly convex and uniformly smooth Banach space. Let $A : E \to 2^{E^*}$ be a maximally monotone operator and $B : E \to E^*$ be monotone and Lipschitz. For $x_0, x_{-1} \in E$ define the sequence x_n iteratively by*

$$x_{n+1} = J^A_{\lambda_n} \circ J^{-1}(Jx_n - \lambda_n Bx_n - \lambda_{n-1}(Bx_n - Bx_{n-1})), \quad n \geq 0 \tag{14}$$

where $\lambda_n \subseteq [\epsilon, \frac{1-2\epsilon}{2\mu L}]$ for some $\epsilon > 0$ and $\mu \geq 1$. Suppose $(A + B)^{-1}(0) \neq \emptyset$, then the sequence $\{x_n\}$ is bounded.

Proof: Let $x^* \in (A + B)^{-1}(0)$. So that

$$-Bx^* \in Ax^*. \tag{15}$$

From (14) we have that

$$\frac{1}{\lambda_n}(Jx_n - \lambda_n Bx_n - \lambda_{n-1}(Bx_n - Bx_{n-1}) - Jx_{n+1}) \in Ax_{n+1} \tag{16}$$

Using (15), (16) and the monotonicity of A, we obtain

$$\langle Jx_{n+1} - Jx_n + \lambda_n(Bx_n - Bx^*) + \lambda_{n-1}(Bx_n - Bx_{n-1}), x^* - x_{n+1}\rangle \geq 0 \tag{17}$$

By Lemma 2.3(i), we have

$$2\langle Jx_{n+1} - Jx_n, x^* - x_{n+1}\rangle = \phi(x^*, x_n) - \phi(x^*, x_{n+1}) - \phi(x_{n+1}, x_n) \tag{18}$$

Also

$$\begin{aligned}&\langle Bx_n - Bx^*, x^* - x_{n+1}\rangle \\ &= \langle Bx_{n+1} - Bx^*, x^* - x_{n+1}\rangle + \langle Bx_n - Bx_{n+1}, x^* - x_{n+1}\rangle\end{aligned} \tag{19}$$

and

$$\begin{aligned}&\langle Bx_n - Bx_{n-1}, x^* - x_{n+1}\rangle \\ &= \langle Bx_n - Bx_{n-1}, x^* - x_n\rangle + \langle Bx_n - Bx_{n-1}, x_n - x_{n+1}\rangle\end{aligned} \tag{20}$$

Substituting (18), (19) and (20) in (17) we have:

$$\begin{aligned}
&\phi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle \\
&\leq \phi(x^*, x_n) + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle \\
&\quad + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x_n - x_{n+1}\rangle - \phi(x_{n+1}, x_n) \\
&\quad + 2\lambda_n\langle Bx_{n+1} - Bx^*, x^* - x_{n+1}\rangle
\end{aligned} \tag{21}$$

Using the monotonicity of B on the last term of Equation (21) and rearranging the equation we have

$$\begin{aligned}
&\phi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle + \phi(x_{n+1}, x_n) \leq \phi(x^*, x_n) \\
&\quad + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle \\
&\quad + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x_n - x_{n+1}\rangle
\end{aligned} \tag{22}$$

Using the Lipschitz property of B and Lemma 2.4 we have

$$\begin{aligned}
&2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x_n - x_{n+1}\rangle \\
&\leq 2\lambda_{n-1}\|Bx_n - Bx_{n-1}\|\|x_n - x_{n+1}\| \\
&\leq 2\lambda_{n-1}\|x_n - x_{n-1}\|\|x_n - x_{n+1}\| \\
&\leq \lambda_{n-1}L(\|x_n - x_{n-1}\|^2 + \|x_n - x_{n+1}\|^2) \\
&\leq \lambda_{n-1}\mu L(\phi(x_n, x_{n-1}) + \phi(x_{n+1}, x_n))
\end{aligned} \tag{23}$$

Substituting (23) in (22) we have

$$\begin{aligned}
&\phi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle \\
&\quad + (1 - \lambda_{n-1}\mu L)\phi(x_{n+1}, x_n) \leq \phi(x^*, x_n) \\
&\quad + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle + \lambda_{n-1}\mu L\phi(x^*, x_{n-1}).
\end{aligned} \tag{24}$$

Now, $\lambda_n \in [\epsilon, \frac{1-2\epsilon}{2\mu L}]$ implies $\lambda_n\mu L < \frac{1}{2}$ and $\frac{1}{2} + \epsilon \leq 1 - \lambda_n\mu L$. So that (24) becomes

$$\begin{aligned}
&\phi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle \\
&\quad + \left(\frac{1}{2} + \epsilon\right)\phi(x_{n+1}, x_n) \leq \phi(x^*, x_n) \\
&\quad + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle + \frac{1}{2}\phi(x^*, x_{n-1})
\end{aligned} \tag{25}$$

Thus, we have

$$\begin{aligned}
&\epsilon\phi(x_{n+1}, x_n) \\
&\leq \left[\phi(x^*, x_n) - \phi(x^*, x_{n+1})\right] + \left[2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle\right. \\
&\quad \left. -2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle\right] + \frac{1}{2}\left[\phi(x_n, x_{n-1}) - \phi(x_{n+1}, x_n)\right]
\end{aligned} \tag{26}$$

Telescoping the terms in (26) we have

$$\begin{aligned}
&\epsilon \sum_{k=0}^{n} \phi(x_{k+1}, x_k) \\
&\le \left[\phi(x^*, x_0) - \phi(x^*, x_{n+1})\right] + \left[2\lambda_{n-1}\langle Bx_0 - Bx_{-1}, x^* - x_0\rangle\right. \\
&\quad \left. -2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle\right] + \frac{1}{2}\left[\phi(x_0, x_{-1}) - \phi(x_{n+1}, x_n)\right]. \qquad (27)
\end{aligned}$$

Hence we have

$$\begin{aligned}
&\phi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle \\
&\quad + \frac{1}{2}\phi(x_{n+1}, x_n) + \epsilon \sum_{k=0}^{n} \phi(x_{k+1}, x_k) \\
&\quad \le \phi(x^*, x_0) + 2\lambda_{n-1}\langle Bx_0 - Bx_{-1}, x^* - x_0\rangle + \frac{1}{2}\phi(x_0, x_{-1}) \qquad (28)
\end{aligned}$$

Again, Using the Lipschitz property of B and Lemma 2.4 we obtain

$$-\lambda_n\mu L(\phi(x_{n+1}, x_n) + \phi(x^*, x_{n+1})) \le 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}.\rangle \qquad (29)$$

Setting $M = \phi(x^*, x_0) + 2\lambda_{n-1}\langle Bx_0 - Bx_{-1}, x^* - x_0\rangle + \phi(x_0, x_{-1})$ and substituting (29) in (28), we obtain

$$(1 - \lambda_n\mu L)\phi(x^*, x_{n+1}) + \left(\frac{1}{2} - \lambda_n\mu L\right).\phi(x_{n+1}, x_n) + \epsilon \sum_{k=0}^{\infty} \phi(x_{k+1}, x_k) \le M.$$

Since

$$\frac{1}{2} < 1 - \lambda_n\mu L \quad \text{and} \quad \frac{1}{2} - \lambda_n\mu L > 0,$$

we have

$$\frac{1}{2}\phi(x^*, x_{n+1}) + \epsilon \sum_{k=0}^{n} \phi(x_{k+1}, x_k) \le M \quad \forall\, n \ge 0. \qquad (30)$$

Hence, the sequence $\{\phi(x^*, x_{n+1})\}$ is bounded and consequently, the sequence $\{x_n\}$ is bounded. ■

Theorem 3.2: *Let E be a real 2-uniformly convex and uniformly smooth Banach space. Let $A : E \to 2^{E^*}$ be a maximally monotone operator and $B : E \to E^*$ be monotone and Lipschitz. For $x_0, x_{-1} \in E$ defined sequence x_n by*

$$x_{n+1} = J^A_{\lambda_n} \circ J^{-1}(Jx_n - \lambda_n Bx_n - \lambda_{n-1}(Bx_n - Bx_{n-1})), \quad n \ge 0 \qquad (31)$$

where $\lambda_n \subseteq [\epsilon, \frac{1-2\epsilon}{2\mu L}]$ for some $\epsilon > 0$ and $\mu \ge 1$. Suppose $(A + B)^{-1}(0) \ne \emptyset$ and that the normalized duality mapping J is weakly sequentially continuous, then the sequence $\{x_n\}$ converges weakly to an element of $(A + B)^{-1}(0)$.

Proof: From Theorem 3.1, we have that the sequence $\{x_n\}$ is bounded. Let x^* be a weak sequential cluster point of $\{x_n\}$ i.e. there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \rightharpoonup x^*$. We show that $x^* \in (A+B)^{-1}(0)$.

Let $(u,v) \in \text{Graph}(A+B)$. That is $v - Bu \in Au$

Also, from (42) we have that

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

Monotonicity of A implies

$$\begin{aligned}
&\langle u - x_{n+1}, v\rangle \\
&\geq \left\langle u - x_{n+1}, Bu + \frac{1}{\lambda_n}(Jx_n - Jx_{n+1} - \lambda_n Bx_n - \lambda_{n-1}(Bx_n - Bx_{n-1}))\right\rangle \\
&= \frac{1}{\lambda_n}\langle u - x_{n+1}, Jx_n - Jx_{n+1}\rangle - \frac{\lambda_{n-1}}{\lambda_n}\langle u - x_{n+1}, Bx_n - Bx_{n-1}\rangle \\
&\quad + \langle u - x_{n+1}, Bu - Bx_n\rangle \\
&\geq \frac{1}{\lambda_n}\langle u - x_{n+1}, Jx_n - Jx_{n+1}\rangle - \frac{\lambda_{n-1}}{\lambda_n}\langle u - x_{n+1}, Bx_n - Bx_{n-1}\rangle \\
&\quad + \langle u - x_{n+1}, Bx_{n+1} - Bx_n\rangle
\end{aligned} \tag{32}$$

where in the last inequality we have used the fact that $\langle u - x_{n+1}, Bu - Bx_{n+1}\rangle \geq 0$. Now from (30), we have that $\phi(x_{n+1}, x_n) \to 0$ as $n \to \infty$. Thus by Lemma 2.7, we have $\|x_{n+1} - x_n\| \to 0$ as $n \to +\infty$, and consequently, Lipschitz of B implies that $\|Bx_{n+1} - Bx_n\| \to 0$ as $n \to \infty$. We also have, by uniform continuity of J on bounded sets that, $Jx_n - Jx_{n+1} \to 0$ as $n \to +\infty$. Hence we have

$$\underline{\lim}\langle u - x_{n+1}, v\rangle \geq 0.$$

So that $\langle u - x^*, v\rangle = \lim_{k\to+\infty}\langle u - x_{n_k}, v\rangle \geq \underline{\lim}\langle u - x_n, v\rangle \geq 0$. Since $A+B$, by Lemma 2.5, is maximally monotone, we have $x^* \in (A+B)^{-1}(0)$.

Now set

$$a_n = \phi(x^*, x_n) + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle + \frac{1}{2}\phi(x_n, x_{n-1}).$$

We see from (25) that

$$a_{n+1} \leq a_n \quad \forall\, n \geq 0.$$

Also, by Lipschitz property of B and Lemma 2.4 we have

$$\begin{aligned}
a_{n+1} &= \phi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle + \frac{1}{2}\phi(x_{n+1}, x_n) \\
&\geq (1 - \lambda_n \mu L)\phi(x^*, x_{n+1}) + (\frac{1}{2} - \lambda_n \mu L)\phi(x_{n+1}, x_n) \\
&\geq \frac{1}{2}\phi(x^*, x_{n+1}) \geq 0.
\end{aligned}$$

So the limit $\lim_{n\to\infty} a_n$ exists. Now, since the limits: $\lim_{n\to\infty} \phi(x_{n+1}, x_n)$ and $\lim_{n\to\infty}\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle$ exist and equal 0. It implies that $\lim_{n\to\infty} \phi(x^*, x_n)$ exists.

Now Let x^* and y^* be two weak cluster points of $\{x_n\}$ Then

$$\phi(x^*, x_n) = \|x^*\|^2 - 2\langle x^*, Jx_n\rangle + \|x_n\|^2$$

and

$$\phi(y^*, x_n) = \|y^*\|^2 - 2\langle y^*, Jx_n\rangle + \|x_n\|^2.$$

So that

$$2\langle x^* - y^*, Jx_n\rangle = \phi(y^*, x_n) - \phi(x^*, x_n) + \|x^*\|^2 - \|y^*\|^2$$

So the limit $\lim_{n\to\infty}\langle x^* - y^*, Jx_n\rangle$ exists.

Using the fact that J is weakly sequentially continuous, we have

$$\langle x^* - y^*, Jx^*\rangle = \lim_{j\to\infty}\langle x^* - y^*, Jx_{n_j}\rangle = \lim_{k\to\infty}\langle x^* - y^*, Jx_{n_k}\rangle = \langle x^* - y^*, Jy^*\rangle$$

For some $\{x_{n_k}\}$ and $\{x_{n_j}\}$ such that $x_{n_k} \rightharpoonup y^*$ and $x_{n_j} \rightharpoonup x^*$.

$$\langle x^* - y^*, Jx^* - Jy^*\rangle = 0.$$

Using Lemma 2.6 we conclude that $x^* = y^*$. The proof is complete. Now, if we fix $\lambda_n = \lambda$ for all $n \geq 1$ in Theorem 3.1 we obtain the following corollary.

Corollary 3.3: *Let E be a real 2-uniformly convex and uniformly smooth Banach space. Let $A : E \to 2^{E^*}$ be a maximally monotone operator and $B : E \to E^*$ be monotone and Lipschitz. For $x_0, x_{-1} \in E$ defined sequence x_n by*

$$x_{n+1} = J_\lambda^A \circ J^{-1}(Jx_n - 2\lambda Bx_n + \lambda Bx_{n-1}), \quad n \geq 0 \tag{33}$$

where $\lambda \subseteq [\epsilon, \frac{1-2\epsilon}{2\mu L}]$ for some $\epsilon > 0$ and $\mu \geq 1$. Suppose $(A + B)^{-1}(0) \neq \emptyset$ and that the normalized duality mapping J is weakly sequentially continuous, then the sequence $\{x_n\}$ converges weakly to an element of $(A + B)^{-1}(0)$.

We now state Theorem 3.1 in a real Hilbert space.

Corollary 3.4: *Let H be a real Hilbert space. Let $A : H \to 2^H$ be a maximally monotone operator and $B : H \to H$ be monotone and Lipschitz. For $x_0, x_{-1} \in H$ defined sequence $\{x_n\}$ by*

$$x_{n+1} = J_{\lambda_n}^A(x_n - \lambda_n Bx_n - \lambda_{n-1}(Bx_n - Bx_{n-1})), \quad n \geq 0 \tag{34}$$

where $\lambda_n \subseteq [\epsilon, \frac{1-2\epsilon}{2L}]$ for some $\epsilon > 0$. Suppose $(A + B)^{-1}(0) \neq \emptyset$. Then $\{x_n\}$ converges weakly to an element of $(A + B)^{-1}(0)$.

Remark 3.5: In the implementation of Algorithm (42), we assume that the resolvent $(J+\lambda_n A)^{-1}J$ is easy to compute and the duality mapping J is easily computable as well. On the other hand, one has to obtain the Lipschitz constant, L, of the monotone mapping B (or an estimate of it). In a case when the Lipschitz constant cannot be accurately estimated or overestimated, this might result in too small step-sizes λ_n. This can be a drawback of Algorithm (42) One way to overcome this problem is to introduce linesearch rule in Algorithm (42).

Next we give an alternative method in which the step size $\{\lambda_n\}$ does not depend on the Lipschitz constant. Let A and B be as in Theorem 3.2. For arbitrary $x_0, x_{-1} \in X$, define the sequence $\{x_n\}$ by

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

where $\{\lambda_n\}$ is the largest $\lambda \in \{\gamma, \gamma l, \gamma l^2, \dots\}$ satisfying

$$\lambda_n \|Bx_{n+1} - Bx_n\| \leq \theta \|x_{n+1} - x_n\| \tag{35}$$

with $\theta \in (0, \frac{1-2\epsilon}{2\mu})$, $l \in (0,1)$ $\gamma > 0$ and some $\epsilon > 0$.

We first observe that $\{x_n\}$ defined above in (35) is well defined and $\min\{\gamma, \frac{\theta l}{L}\} \leq \lambda_n \leq \gamma$ where L is the Lipschitz constant of B. Indeed, By Lipschitz property of B we have

$$\|Bx_{n+1} - Bx_n\| \leq L\|x_{n+1} - x_n\|.$$

so that

$$\frac{\theta}{L}\|Bx_{n+1} - Bx_n\| \leq \theta\|x_{n+1} - x_n\|.$$

Hence, (35) will be satisfied whenever $\lambda_n \leq \frac{\theta}{L}$. So it is well defined. Also, from definition of λ_n we have that $\lambda_n \leq \gamma$. Now, Suppose $\lambda_n < \gamma$. Then $\frac{\lambda_n}{L}$ will not satisfy (35), thus,

$$\frac{\lambda_n}{L}\|Bx_{n+1} - Bx_n\| \geq \theta\|x_{n+1} - x_n\|.$$

That is

$$L\|x_{n+1} - x_n\| \geq \|Bx_{n+1} - Bx_n\| \geq \frac{\theta}{\lambda_n/L}\|x_{n+1} - x_n\|.$$

Hence, $\lambda_n \geq \frac{\theta}{L}$, and thus λ_n is bounded and below if non zero.

We now establish the weak convergence of $\{x_n\}$ with condition (35). From (22), we have that

$$\begin{aligned}&\phi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle + \phi(x_{n+1}, x_n)\\&\le \phi(x^*, x_n) + 2\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle\\&\quad + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x_n - x_{n-1}\rangle \end{aligned} \tag{36}$$

Using (35) and Lemma 2.4 we have

$$2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x_n - x_{n+1}\rangle \le \mu\theta(\phi(x_n, x_{n-1}) + \phi(x_{n+1}, x_n)) \tag{37}$$

Substituting (37) in (36) we have:

$$\begin{aligned}&\phi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle + (1-\mu\theta)\phi(x_{n+1}, x_n)\\&\le \phi(x^*, x_n) + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle + \mu\theta\phi(x_n, x_{n-1})\end{aligned}$$

Using the fact that $\theta \in (0, \frac{1-2\epsilon}{2\mu})$ we have

$$\begin{aligned}&\phi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle + \left(\frac{1}{2}+\epsilon\right)\phi(x_{n+1}, x_n)\\&\le \phi(x^*, x_n) + 2\lambda_{n-1}\langle Bx_n - Bx_{n-1}, x^* - x_n\rangle + \frac{1}{2}\phi(x_n, x_{n-1})\end{aligned}$$

Telescoping as in Theorem 3.2 we have

$$\begin{aligned}&\phi(x^*, x_{n+1}) + 2\lambda_n\langle Bx_{n+1} - Bx_n, Bx^* - x_{n+1}\rangle\\&\quad + \frac{1}{2}\phi(x_{n+1}, x_n) + \epsilon\sum_{k=0}^{n}\phi(x_{k+1}, x_k) \le M,\end{aligned} \tag{38}$$

where $M := \phi(x^*, x_0) + 2\lambda_{n-1}\langle Bx_0 - Bx_{-1}, x^* - x_0\rangle + \frac{1}{2}\phi(x_0, x_{-1})$.

Using condition (35) and Lemma 2.4, we have

$$\mu\theta\left(\phi(x_{n+1}, x_n) + \phi(x^*, x_{n+1})\right) \le 2\lambda_{n-1}\langle Bx_{n+1} - Bx_n, x^* - x_{n+1}\rangle$$

So that

$$(1-\mu\theta)\phi(x^*, x_{n+1}) + \left(\frac{1}{2} - \mu\theta\right).\phi(x_{n+1}, x_n) + \epsilon\sum_{k=0}^{n}\phi(x_{k+1}, x_k) \le M^* \tag{39}$$

Observe that $1 - \mu\theta > \frac{1}{2}$ and $\frac{1}{2} - \mu\theta > 0$ So that $\{\phi(x^*, x_{n+1})\}$ is bounded and $\phi(x_{n+1}, x_n) \to 0$ as $n \to +\infty$. Consequently the sequence $\{x_n\}$ is bounded and $\|x_{n+1} - x_n\| \to 0$ as $n \to +\infty$. The rest of the proof follow exactly as that of Theorem 3.2. ■

4. Application

In this section, we apply our results to the minimization of composite objective function of the type

$$\min_{x\in E} f(x) + g(x), \tag{40}$$

where $f : E \to \mathbb{R}$ is proper, convex and lower semi-continuous functional and $g : E \to \mathbb{R}$ is convex functional.

Many optimization problems from image processing [27], statistical regression, machine learning (see e.g. [28] and the references contained therein), etc., can be adopted into the form of (40). In this setting, we assume that g represents the 'smooth part' of the functional where f is assumed to be non-smooth. Specifically, if we assume that is Gâteaux differentiable with derivatives ∇g which is Lipschitz-continuous with constant L. Then by [29, Theorem 3.13], we have

$$\langle \nabla g(x) - \nabla g(y), x - y\rangle \geq \frac{1}{L}\|\nabla g(x) - \nabla g(y)\|^2, \quad \forall\, x, y \in E.$$

Therefore, ∇g is monotone and Lipschitz continuous with Lipschitz constant L. Observe that problem (40) is equivalent to find $x \in E$ such that

$$0 \in \partial f(x) + \nabla g(x). \tag{41}$$

Then problem (41) is a special case of the inclusion problem (1) with $A := \partial f$ and $B := \nabla g$.

Next, we obtain the resolvent of ∂f. Let us fix $r > 0$ and $z \in E$. Suppose $J_r^{\partial f}$ is the resolvent of ∂f. Then

$$Jz \in J(J_r^{\partial f}) + r\partial f(J_r^{\partial f}).$$

Hence we obtain

$$0 \in \partial f(J_r^{\partial f}) + \frac{1}{r}J(J_r^{\partial f}) - \frac{1}{r}Jz = \partial\left(f + \frac{1}{2r}\|\cdot\|^2 - \frac{1}{r}Jz\right) J_r^{\partial f}.$$

Therefore,

$$J_r^{\partial f}(z) = \arg\min_{y\in E}\left\{f(y) + \frac{1}{2r}\|y\|^2 - \frac{1}{r}\langle y, Jz\rangle\right\}.$$

We obtain the following weak convergence result for problem (40).

Theorem 4.1: *Let E be a real 2-uniformly convex and uniformly smooth Banach space and the solution set Υ of problem (40) be non-empty. For $x_0, x_{-1} \in E$ defined sequence x_n by*

$$x_{n+1} = J_{\lambda_n}^{\partial f} \circ J^{-1}(Jx_n - \lambda_n \nabla g(x_n) - \lambda_{n-1}(\nabla g(x_n) - \nabla g(x_{n-1}))), \quad n \geq 0 \tag{42}$$

where $\lambda_n \subseteq [\epsilon, \frac{1-2\epsilon}{2\mu L}]$ for some $\epsilon > 0$ and $\mu \geq 1$. Suppose $\Upsilon \neq \emptyset$ and that the normalized duality mapping J is weakly sequentially continuous, then the sequence $\{x_n\}$ converges weakly to an element of Υ.

Next, we now give an application in the framework of Hilbert spaces. We assume that C is a non-empty closed and convex subset of a Hilbert space H and P_C stands for metric projection from H onto C.

4.1. Variational inequality problems

A monotone variational inequality problem (VIP) is formulated as the problem of finding a point $x \in C$ with the property:

$$\langle Ax, y - x\rangle \geq 0, \quad \forall\, y \in C, \tag{43}$$

where $A : C \to H$ is nonlinear monotone operator. By Rockafellar [30], VIP (43) is equivalent to finding a point x so that

$$0 \in (A + B)x,$$

where B is the normal cone operator of C. In other words, VIPs are special case of the inclusion problems.

We shall denote by S, the set of solution of (43) and assume S is non-empty.

If we denote the normal cone of C as N_C. Let $A = N_C$ in Corollary 3.4, we find that $J_{\lambda_n} = J^A_{\lambda_n} = (I + \lambda_n N_C)^{-1} = P_C$. The following result can be obtained immediately

Theorem 4.2: *Let H be a real Hilbert space. Let $B : H \to H$ be monotone. Define a sequence by the iterative scheme,*

$$x_{n+1} = P_C(x_n - \lambda_n Bx_n - \lambda_{n-1}(Bx_n - Bx_{n-1})), \tag{44}$$

where $\lambda_n \subseteq [\epsilon, \frac{1-2\epsilon}{2L}]$ for some $\epsilon > 0$. Then the sequence $\{x_n\}$ generated by (44) converges to a solution of problem (43).

5. Numerical experiment

To show that the proposed method in Section 3 is efficient and easily implementable, we consider two examples in $\mathcal{L}_2([0, 1])$ and give numerical example of the proposed Algorithm (3.1). Numerical experiments were carried out on Matlab R2020a version. All programs were run on a 64-bit OS PC with Intel(R) Core(TM) i5-1035G1 CPU @ 1.00GHz 1.19 GHz and 3GB RAM. All figures were plotted using the loglog plot command, while integrals were approximated using the *trapz* command on MATLAB over the interval $[0, 1]$, being partitioned into 1000 subintervals.

Example 5.1: Let $H = \mathcal{L}_2([0,1])$, with norm and inner product function defined by

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

Observe that the normalized duality map, for $1 < p \leq 2$, $1/p + 1/q = 1$, $J : L_p([0,1]) \longrightarrow \mathcal{L}_q([0,1])$ and its inverse $J^{-1} : \mathcal{L}_q([0,1]) \longrightarrow \mathcal{L}_p([0,1])$ can be defined by (see, for instance, Alber [31])

$$J(f) = ||f||_p^{2-p} f|f|^{p-2} \quad \text{for all } f \in \mathcal{L}_p([0,1]), \tag{45}$$

$$J^{-1}(g) = ||g||_q^{2-q} g|g|^{q-2} \quad \text{for all } g \in \mathcal{L}_q([0,1]), \tag{46}$$

respectively. Setting $p, q = 2$ implies that J and J^{-1}, defined in (45) and (46), are identities. Now, define the operator $A, B : \mathcal{L}_2([0,1]) \to \mathcal{L}_2([0,1])$ by

$$Bx(t) = \int_0^1 \left(x(s) - \left(\frac{2tse^{t+s}}{e\sqrt{e^2-1}} \right) \cos x(s) \right) \mathrm{d}s + \frac{2te^t}{e\sqrt{e^2-1}}, \quad x \in \mathcal{L}_2([0,1]),$$
$$Ax(t) = \max\{0, x(t)\}, \quad t \in [0,1].$$

Then B is monotone and Lipschitz continuous with a Lipschitz constant, $L = 2$, while A is maximally monotone on $\mathcal{L}_2([0,1])$ (see e.g. [32]). Given the definition of A given above, the resolvent operator, for $r > 0$, $J_r^A : \mathcal{L}_2([0,1]) \to \mathcal{L}_2([0,1])$, becomes

$$J_r^A x(t) = \begin{cases} x(t), & Ax(t) = 0, \\ \dfrac{1}{1+r} x(t), & Ax(t) = x(t). \end{cases}$$

Let $\psi \in H$ be a continuous linear functional. By Riesz representation theorem, there exists a unique $f_\psi \in H$ such that $\langle \psi, x \rangle = \langle f_\psi, x \rangle$ forall $x \in H$. Since $(A + B)^{-1}(0) = 0$, we show that $x_n \rightharpoonup 0$ as $n \to \infty$. This implies showing that

$$\langle f_\psi, x_n \rangle = \langle \psi, x_n \rangle = \int_0^1 \psi(t) x_n(t) \mathrm{d}t \to 0 \quad \text{as} \quad n \to \infty.$$

In this example, ψ is chosen arbitrarily in $\mathcal{L}_2([0,1])$ as seen in Table 1. From this table and Figures 1–3, it can be seen that the number of function evaluations ($\mathbb{NFE}$) reduces as the fixed power, $j \in \mathbb{N}$, of $\psi(t) = t^j$, $t \in [0,1]$, increases. As illustrated in the aforementioned figures, the present algorithm generates a sequence that converges to zero using a loglog plot. It can also be inferred from Table 1 that it takes a less computational time to compute ψ as a monomial of higher degree j.

Example 5.2: Let $H = \mathcal{L}_2([0,1])$, with norm and inner product function defined

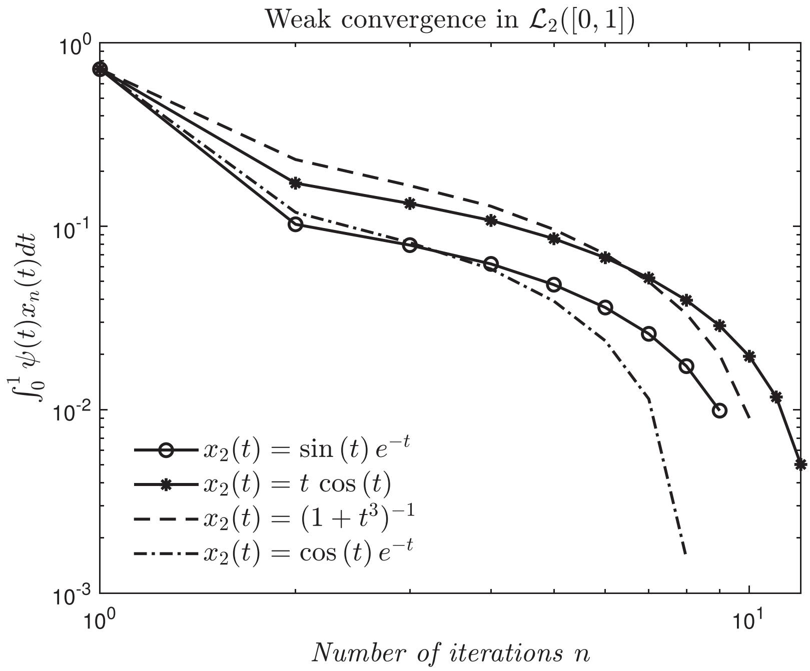

Figure 1. Example 5.1 when $x_1(t) = e^t$ and $\psi(t) = t$.

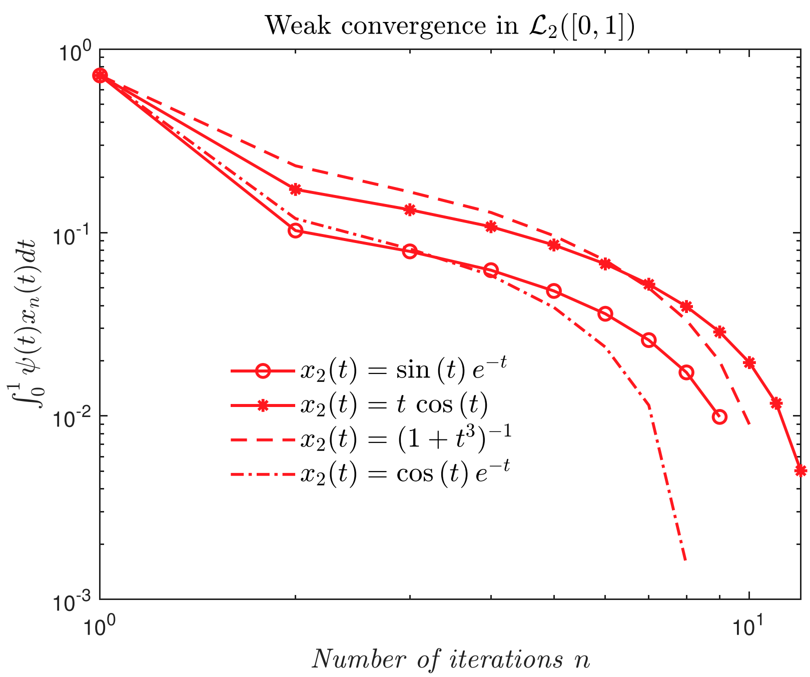

Figure 2. Example 5.1 when $x_1(t) = e^t$ and $\psi(t) = t^2$.

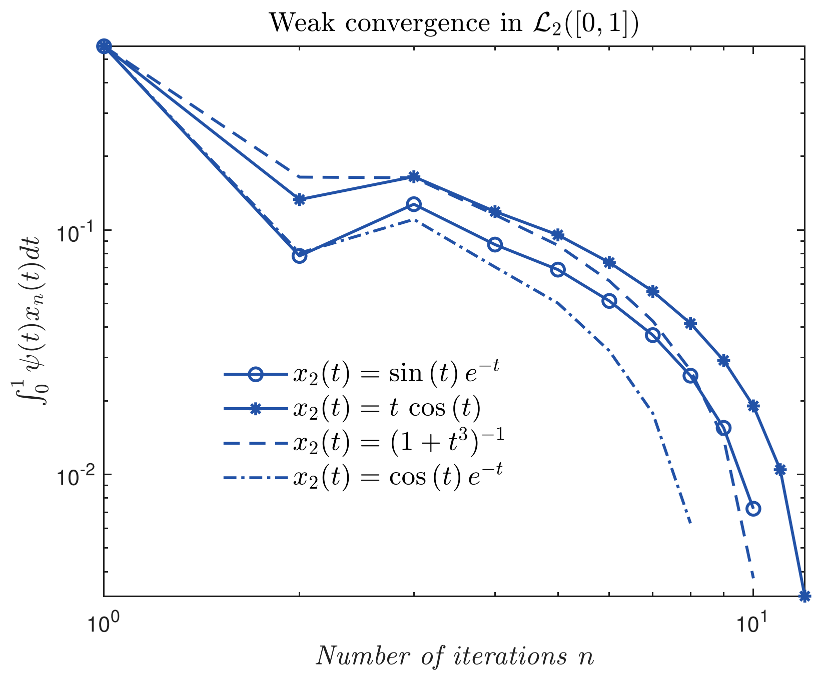

Figure 3. Example 5.1 when $x_1(t) = e^t$ and $\psi(t) = t^3$.

Table 1. Computational results for Example 5.1 when $\lambda_1 = 9/64$, $\epsilon = 1/8$, and $\mu = 1.1$.

TOL			$x_2(t)$	NFE	Time (secs)
1e−−2	$x_1(t) = e^t$	$\psi(t) = t$	$\sin(t)\, e^{-t}$	15	1.2890e−002
			$t \cos(t)$	17	3.1139e−003
			$(1+t^3)^{-1}$	15	1.5174e−003
			$\cos(t)\, e^{-t}$	13	1.3567e−003
		$\psi(t) = t^2$	$\sin(t)\, e^{-t}$	12	7.7586e−003
			$t \cos(t)$	13	2.8121e−003
			$(1+t^3)^{-1}$	11	1.1553e−003
			$\cos(t)\, e^{-t}$	10	1.0490e−003
		$\psi(t) = t^3$	$\sin(t)\, e^{-t}$	10	1.0675e−002
			$t \cos(t)$	12	3.5905e−003
			$(1+t^3)^{-1}$	10	1.6049e−003
			$\cos(t)\, e^{-t}$	8	1.3017e−003

by

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

The feasibility set in consideration is given by the hyperslab $C := \{v \in H : c \leq \langle a, v\rangle \leq b\}$, where $\mathbf{0} \neq a \in H$ and $b \in \mathbb{R}$. Suppose that N_C is the normal cone of C and let $A = N_C$. Then, for $r > 0$, $J_r^A = (I + rN_C)^{-1} = P_C$. Moreover, the

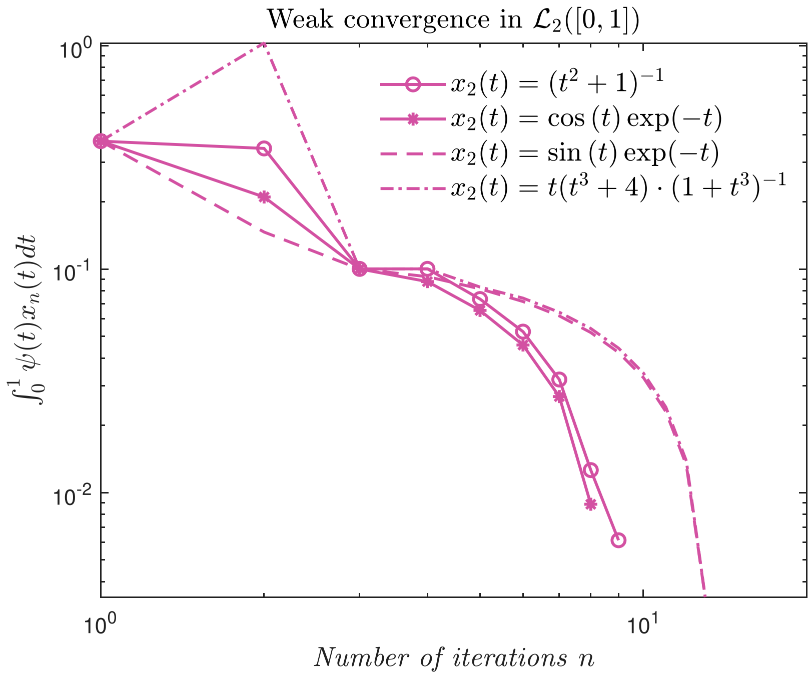

Figure 4. Example 5.2 when $x_1(t) = (1+t^3)^{-1}$ and $\psi(t) = t$.

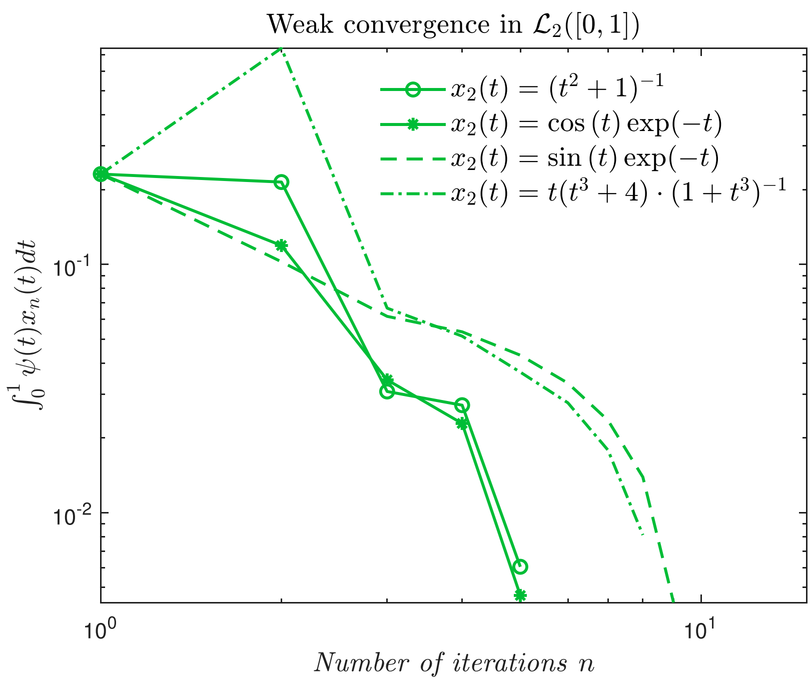

Figure 5. Example 5.2 when $x_1(t) = (1+t^3)^{-1}$ and $\psi(t) = t^2$.

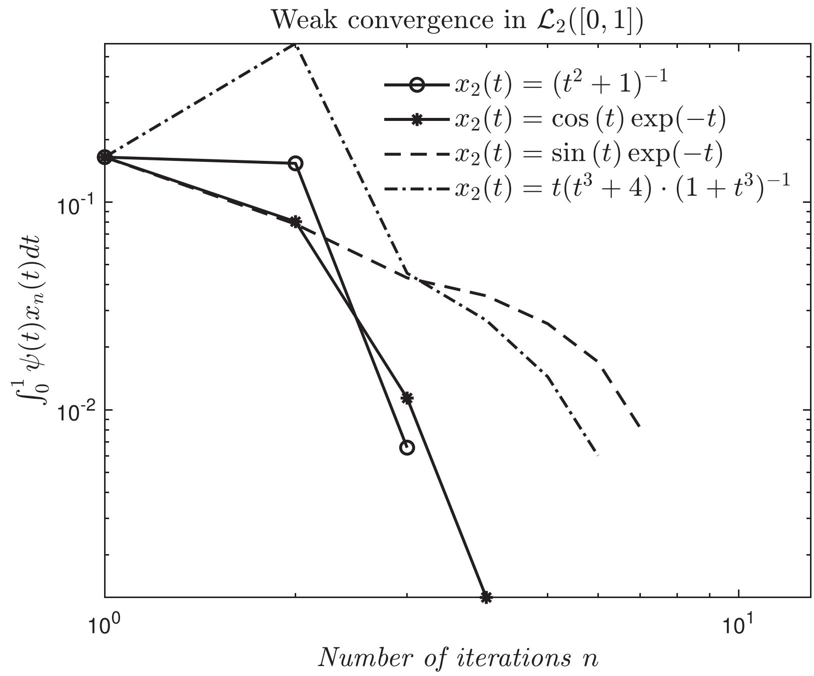

Figure 6. Example 5.2 when $x_1(t) = (1+t^3)^{-1}$ and $\psi(t) = t^3$.

Table 2. Computational results for Example 5.2 when $\lambda_1 = 9/64$, $\epsilon = 1/8$, and $\mu = 1.1$.

TOL			$x_2(t)$	NFE	Time (secs)
1e−2	$x_1(t) = (1+t^3)^{-1}$	$\psi(t) = t$	$(1+t^3)^{-1}$	9	1.1987e−002
			$\cos(t)\,e^{-t}$	8	3.9848e−003
			$\sin(t)\,e^{-t}$	13	1.4970e−003
			$t(t^3+4)(1+t^3)^{-1}$	13	3.2559e−003
	$x_1(t) = (1+t^3)^{-1}$	$\psi(t) = t^2$	$(1+t^3)^{-1}$	5	1.2753e−003
			$\cos(t)\,e^{-t}$	5	5.8930e−004
			$\sin(t)\,e^{-t}$	9	8.2250e−004
			$t(t^3+4)(1+t^3)^{-1}$	8	7.7780e−004
	$x_1(t) = (1+t^3)^{-1}$	$\psi(t) = t^3$	$(1+t^3)^{-1}$	3	3.4430e−004
			$\cos(t)\,e^{-t}$	4	4.7810e−004
			$\sin(t)\,e^{-t}$	7	8.6690e−004
			$t(t^3+4)(1+t^3)^{-1}$	6	9.0960e−004

projection map $P_C : H \longrightarrow C$ is defined by

$$P_C(v) = \begin{cases} v + \left(\dfrac{c - \langle a, v\rangle}{\|a\|^2}\right) a, & \text{if } \langle a, v\rangle < c, \\ v & \text{if } c \le \langle a, v\rangle \le b, \\ v + \left(\dfrac{b - \langle a, v\rangle}{\|a\|^2}\right) a, & \text{if } \langle a, v\rangle > b, \end{cases} \tag{47}$$

see, for example, [33, p. 419]. Setting $c = -100$ and $b = 100$, let $a : [0,1] \to \mathbb{R} : t \mapsto t$. All other maps except the projection map have been inherited from the previous example.

From Table 2 and Figures 4–6, it can be observed, as well, that when ψ is a monomial, i.e. for a fixed power, $j \in \mathbb{N}$, of $\psi(t) = t^j$, $t \in [0, 1]$, the algorithm converges faster. As illustrated in the table and figures, the method works for even rational functions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

A. U. Bello http://orcid.org/0000-0001-6965-8239

References

[1] Attouch H, Peypouquet J, Redont P. Backward–forward algorithms for structured monotone inclusions in Hilbert spaces. J Math Anal Appl. 2018;457(2):1095–1117.
[2] 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. 1977;61:159–164.
[3] Censor Y, Elfving T. A multi-projection algorithm using Bregman projections in a product space. Numer Algorithms. 1994;8:221–239.
[4] Chen GH-G, Rockafellar RT. Convergence rates in forward–backward splitting. SIAM J Optim. 1997;7(2):421–444.
[5] Combettes P, Wajs VR. Signal recovery by proximal forward–backward splitting. Multiscale Model Simul. 2005;4(4):1168–1200.
[6] Davis D, Yin WT. A three-operator splitting scheme and its optimization applications. Set-Valued Var Anal. 2017;25:829–858.
[7] Lions PL, Mercier B. Splitting algorithms for the sum of two nonlinear operators. SIAM J Numer Anal. 1979;16:964–979.
[8] Moudafi A, Thera M. Finding a zero of the sum of two maximal monotone operators. J Optim Theory Appl. 1997;94:425–448.
[9] Passty GB. Ergodic convergence to a zero of the sum of monotone operators in Hilbert spaces. J Math Anal Appl. 1979;72:383–390.
[10] Peaceman DH, Rachford HH. The numerical solutions of parabolic and elliptic differential equations. J Soc Ind Appl Math. 1955;3:28–41.
[11] Alber Y, Ryazantseva I. Nonlinear ill posed problems of monotone type. London: Springer; 2006.
[12] Tseng P. A modified forward–backward splitting method for maximal monotone mappings. SIAM J Control Optim. 2000;38:431–446.
[13] Malitsky Y, Tam MK. A forward–Backward spliting method for monotone inclusion without cocoercivity. SIAM J Optim. 2020;30(2):1451–1472.
[14] Shehu Y. Convergence results of forward–backward algorithm for sum of monotone operator in Banach spaces. Results Math. 2019;74:95.
[15] Tuyen TM, Promkam R, Sunthrayuth P. Strong convergence of a generalized forward–backward splitting method in reflexive Banach spaces. Optimization. 2020; DOI:10.1080/02331934.2020.1812607.

[16] Beauzamy B. Introduction to Banach spaces and their geometry. 2nd ed. Amsterdam: North-Holland Publishing Co.; 1985. (North-Holland mathematics studies; vol. 68. Notas de Matematica [Mathematical Notes], 86). xv+338 pp.
[17] Takahashi W. Nonlinear functional analysis. Yokohama: Yokohama Publishers; 2000.
[18] Xu H-K, Kim TH, Yin X. Weak continuity of the normalized duality map. J Nonlinear Convex Anal. 2014;15(3):595–604.
[19] 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. New York: Dekker; 1996. p. 15–50. (Lecture notes in pure and applied mathematics; vol. 178). .
[20] Alber Y, Guerre-Delabriere S. On the projection methods for fixed point problems. Analysis. 2001;21(1):17–40. (Munich).
[21] Kamimura S, Takahashi W. Strong convergence of a proximal-type algorithm in a Banach space. SIAM J Optim. 2003;13(3):938–945.
[22] Reich S. A weak convergence theorem for the alternating method with Bregman distances. In: Kartsatos AG, editor. Theory and applications of nonlinear operators of accretive and monotone type. New York: Dekker; 1996. p. 313–318. (Lecture notes pure applied mathematics; vol. 178).
[23] Chidume CE, Bello AU, Usman B. Krasnoselskii-type algorithm for zeros of strongly monotone Lipschitz maps in classical banach spaces. SpringerPlus. 2015;4:875. DOI:10.1186/s40064-015-1044-1.
[24] Chidume CE, Chidume CO, Bello AU. An algorithm for zeros of generalized phi-strongly monotone and bounded maps in classical Banach spaces. Optimization. 2016;65(4):827–839. DOI:10.1080/02331934.2015.1074686.
[25] Aoyama K, Kohsaka F. Strongly relatively nonexpansive sequences generated by firmly nonexpansive-like mappings. Fixed Point Theory Appl. 2014;95:13.
[26] Barbu V. Nonlinear semigroups and differential equations in Banach spaces. Bucharest: Editura Academiei R.S.R; 1976.
[27] Bredies K. A forward–backward splitting algorithm for the minimization of non-smooth convex functionals in Banach spaces. Inverse Probl. 2009;25(1):Article ID 015005.
[28] Wang Y, Xu H-K. Strong convergence for the proximal-gradient method. J Nonlinear Convex Anal. 2014;15(3):581–593.
[29] Peypouquet J. Convex optimization in normed spaces: theory, methods and examples. Cham: Springer; 2015. (With a foreword by Hedy Attouch. Springer briefs in optimization). xiv+124 pp.
[30] Rockafellar RT. On the maximality of sums of nonlinear monotone operators. Trans Am Math Soc. 1970;149:75–88.
[31] Alber Y. Generalized projection operators in banach space: properties and applications. Funct Differ Equ. 1993;1993:1–21. 2.12, 3.2.
[32] Hieu DV, Anh PK, Muu LD. Modified hybrid projection methods for finding common solutions to variational inequality problems. Comput Optim Appl. 2017;66:75–96.
[33] Bauschke HH, Combettes PL. Convex analysis and monotone operator theory in Hilbert spaces. 2011. (CMS books in mathematics). DOI:10.1007/978-1-4419-9467-7.