RESEARCH ARTICLE

New strong convergence method for the sum of two maximal monotone operators

Yekini Shehu[1,2] · Qiao-Li Dong[3] · Lu-Lu Liu[3] · Jen-Chih Yao[4]

Abstract

This paper aims to obtain a strong convergence result for a Douglas–Rachford splitting method with inertial extrapolation step for finding a zero of the sum of two set-valued maximal monotone operators without any further assumption of uniform monotonicity on any of the involved maximal monotone operators. Furthermore, our proposed method is easy to implement and the inertial factor in our proposed method is a natural choice. Our method of proof is of independent interest. Finally, some numerical implementations are given to confirm the theoretical analysis.

✉ Yekini Shehu
   yekini.shehu@unn.edu.ng

   Qiao-Li Dong
   dongql@lsec.cc.ac.cn

   Lu-Lu Liu
   lululiumath@163.com

   Jen-Chih Yao
   yaojc@mail.cmu.edu.tw

1   Institute of Science and Technology (IST), Am Campus 1, 3400 Klosterneuburg, Austria

2   Department of Mathematics, Zhejiang Normal University, Jinhua 321004,
    People's Republic of China

3   College of Science, Civil Aviation University of China, Tianjin 300300,
    People's Republic of China

4   Center for General Education, China Medical University, Taichung, Taiwan

🖄 Springer

1 Introduction

Let H be a real Hilbert space with scalar product $\langle . , . \rangle$ and induced norm $\| \cdot \|$. An operator $A : H \to 2^H$ with domain $D(A)$ is said to be monotone if

$$\langle u - v, x - y \rangle \geq 0 \ \ \forall x, y \in D(A), \ \ u \in Ax, v \in Ay.$$

A is maximal monotone if its graph

$$G(A) := \{(x, y) : x \in D(A), y \in Ax\}$$

is not properly contained in the graph of any other monotone operators.

Let us consider the inclusion problem of the form

$$0 \in A(u) + B(u), \tag{1}$$

where A and B are set-valued maximal monotone operators in H. Throughout this paper, we assume that the set of solution, denoted by S, of (1) is nonempty.

The proximal point algorithm (PPA) is the well-known method for solving inclusion problem (1) (see, Lions and Mercier 1979; Martinet 1970; Moreau 1965; Rockafellar 1976). The PPA for solving (1) is expressed as

$$0 \in A(u_{k+1}) + B(u_{k+1}) + \frac{1}{\lambda}(u_{k+1} - u_k), \tag{2}$$

where $\lambda > 0$ is the proximal parameter. Now, implementing PPA (2) to solve (1) requires computing the resolvent operator of the sum $A + B$ exactly. This is very difficult to implement and could be as hard as the original inclusion problem (1). This difficulty has led many authors to consider the operator splitting approach to solve (1). The aim of operator splitting method is to circumvent the computation of J_{A+B}^λ when implementing (2) but rather consider the computation of J_A^λ and J_B^λ (Eckstein and Bertsekas 1992; Glowinski and Le Tallec 1989; Lions and Mercier 1979).

When both A and B are single-valued linear operators in (1), Douglas and Rachford (1956) proposed the following method for solving heat conduction problems:

$$\begin{cases} \frac{1}{\lambda}\left(u_{k+\frac{1}{2}} - u_k\right) + A\left(u_{k+\frac{1}{2}}\right) + B(u_k) = 0, \\ \frac{1}{\lambda}\left(u_{k+1} - u_{k+\frac{1}{2}}\right) + B\left(u_{k+1}\right) - B(u_k) = 0. \end{cases} \tag{3}$$

We can eliminate $u_{k+\frac{1}{2}}$ in (3) above and obtain

$$\left(J_B^\lambda\right)^{-1} u_{k+1} = \left(J_A^\lambda(2J_B^\lambda - I) + (I - J_B^\lambda)\right)\left(J_B^\lambda\right)^{-1} u_k. \tag{4}$$

Define $z_k := \left(J_B^\lambda\right)^{-1} u_k \Leftrightarrow u_k = J_B^\lambda(z_k)$. Then, (4) reduces to the following splitting method (known as Douglas–Rachford splitting method)

$$z_{k+1} = J_A^\lambda(2J_B^\lambda - I)z_k + (I - J_B^\lambda)z_k. \tag{5}$$

Lions and Mercier ([1979](#)) extended the Douglas–Rachford splitting method ([5](#)) to the generic case where both A and B are set-valued nonlinear operators as in our problem ([1](#)). The Douglas–Rachford splitting method ([5](#)) to the generic case is explained as follows in Lions and Mercier ([1979](#)): Starting from an arbitrary iterate u_1 in the domain of B, choosing $b_1 \in B(u_1)$ and setting $z_1 = u_1 + \lambda b_1$, then $u_1 = J_B^\lambda(z_1)$ (the existence of the pair (u_1, z_1) is unique by the Representation Lemma, see Eckstein and Bertsekas [1992](#), cor. 2.3). Thus a sequence $\{z_k\}$ is generated by the Douglas–Rachford scheme ([5](#)); and consequently a sequence $\{u_k := J_B^\lambda(z_k)\}$ converging to a solution point of ([1](#)) can be generated (see Eckstein [1989](#), Thm. 3.15). We refer to Combettes ([2004](#)) for the precise connection between ([5](#)) and the original Douglas–Rachford scheme in Douglas and Rachford ([1956](#)) for heat conduction problems. More details on Douglas–Rachford splitting method ([5](#)) can also be found in Fukushima ([1996](#)), Gabay and Mercier ([1976](#)) and Glowinski and Marrocco ([1975](#)).

1.1 Motivations and contributions

Boţ et al. ([2015](#)) gave the following method for solving ([1](#)): $z_0 = z_1$;

$$
\begin{cases}
u_k = J_B^\lambda(z_k + \alpha_k(z_k - z_{k-1})) \\[2mm]
w_k = J_A^\lambda(2u_k - z_k - \alpha_k(z_k - z_{k-1})) \\[2mm]
z_{k+1} := z_k + \alpha_k(z_k - z_{k-1}) + \beta_k(w_k - u_k),
\end{cases}
\tag{6}
$$

where $\{\alpha_k\}$ is a non-decreasing sequence with $0 \le \alpha_k \le \alpha < 1, \forall k \ge 1$ and $\lambda, \sigma, \delta > 0$ such that

(a) $\delta > \frac{\alpha^2(1+\alpha)+\alpha\sigma}{1-\alpha^2}$; and
(b) $0 < \lambda \le \beta_k \le \theta := 2\frac{\delta-\alpha[\alpha(1+\alpha)+\alpha\delta+\sigma]}{\delta[1+\alpha(1+\alpha)+\alpha\delta+\sigma]}$.

Boţ et al. ([2015](#)) obtained weak convergence analysis of algorithm ([6](#)) for finding common zeros of the sum of two maximal monotone operators and illustrate their results through some numerical experiments. The same conditions (a) and (b) above have been used in recent works in Dong et al. ([2018](#)), Shehu ([2018](#)) and other associated papers. When $\alpha_k = 0$, it was proved in Bauschke and Combettes ([2011](#), Thm. 25.6(vii)) that $\{z_k\}$ in ([6](#)) converges strongly to a solution of ([1](#)) if either A or B is uniformly monotone (A is uniformly monotone if $\langle x - y, u - v \rangle \ge \phi(\|x - y\|), \forall u \in Ax, v \in Ay$, where $\phi : [0, \infty) \to [0, \infty)$ is increasing and vanishes only at zero) on every nonempty bounded subset of its domain.

When $\beta_k = 1$ and $B \equiv 0$, then ([6](#)) reduces to the inertial proximal point method proposed by Alvarez and Attouch ([2001](#)). In this case, Alvarez and Attouch ([2001](#)) assumed that the inertial factor α_k satisfies the condition $0 \le \alpha_k \le \alpha_{k+1} \le \alpha < \frac{1}{3}$ in their convergence result. However, the assumption on the inertial factor α_k imposed in ([6](#)) does not appear as simple as condition $0 \le \alpha_k \le \alpha_{k+1} \le \alpha < \frac{1}{3}$, assumed by Alvarez and Attouch ([2001](#)).

Problems arise in infinite dimensional spaces in many disciplines like economics, image recovery, electromagnetics, quantum physics, and control theory. For such problems, strong convergence of sequence of iterates z_k of the proposed iterative procedure is often much more desirable than weak convergence. This is because strong convergence translates the physically tangible property that the energy $\|z_k - z\|$ of the error between the iterate z_k and a solution z eventually becomes arbitrarily small. Another importance of strong convergence is also underlined in the works of Güler (1991), where a convex function f is minimized through the proximal point algorithm. Güler (1991) showed that the rate of convergence of the value sequence $\{f(z_k)\}$ is better when $\{z_k\}$ converges strongly than when it converges weakly. For more details on importance of strong convergence, please see Bauschke and Combettes (2001).

Strong convergence methods for solving problem (1) when B is set-valued maximal monotone operator and A is a single-valued κ-inverse strongly monotone operator (i.e., $\langle Ax - Ay, x - y \rangle \geq \kappa \|Ax - Ay\|^2, \quad \forall x, y \in H$) have been studied extensively in the literature (see, for example, Boikanyo 2016; Chang et al. 2019; Cholamjiak 2016; Cholamjiak et al. 2018; Dong et al. 2017; Gibali and Thong 2018; López et al. 2012; Riahi et al. 2018; Shehu 2016, 2019; Shehu and Cai 2018; Thong and Cholamjiak 2019; Wang and Wang 2018). However, there are still few results on the strong convergence results concerning more general case of problem (1) when A and B are set-valued maximal monotone operators. This is the gap that this paper aims to fill in.

Our aim in this paper is to prove the strong convergence analysis of the inertial Douglas–Rachford splitting method with different conditions from the conditions (a) and (b) assumed in Boţ et al. (2015) without assuming uniform monotonicity on either maximal monotone operator A or B. Furthermore our assumptions on the inertial factor θ_k here in this paper are the same assumptions in the results of Alvarez and Attouch (2001) (which is a special case of our result). In summary,

- We prove strong convergence analysis of inertial Douglas–Rachford splitting method without using the conditions (a) and (b) assumed in Boţ et al. (2015). Our inertial conditions are the same as the ones assumed in Alvarez and Attouch (2001) for finding zero of a set-valued maximal monotone operator using inertial proximal method.
- We obtain strong convergence results without assuming that any of the involved maximal monotone operators is uniformly monotone on every nonempty bounded subset. Our strong convergence results are much more general than the current ones in Bauschke and Combettes (2011) and other associated works where strong convergence is obtained.
- Some numerical examples are given to confirm the importance of the presence of inertial term in our method.

The paper is therefore organized as follows: We first recall some basic explanations of Douglas–Rachford splitting method and introduce our inertial Douglas–Rachford splitting method alongside some results in Sect. 2. The analysis of strong convergence of our proposed method is then investigated in Sect. 3. We

give numerical implementations in Sect. 4 and conclude with some final remarks in Sect. 5.

2 Preliminaries

Let us first recall some basics that are required to derive and analyze the Douglas–Rachford splitting method; for the corresponding details, we refer, Eckstein and Bertsekas (1992), He and Yuan (2015), Svaiter (2011) and Zhang and Cheng (2013).

Let $\lambda > 0$ be a fixed parameter, and let us denote by

$$J_A^\lambda := (I + \lambda A)^{-1} \quad \text{and} \quad J_B^\lambda := (I + \lambda B)^{-1}$$

the *resolvents* of A and B, respectively, which are known to be firmly nonexpansive (operator T is firmly-nonexpansive if $\langle x - y, Tx - Ty \rangle \geq \|Tx - Ty\|^2, \ \forall x, y \in H$). Furthermore, let us write

$$R_A^\lambda := 2J_A^\lambda - I \quad \text{and} \quad R_B^\lambda := 2J_B^\lambda - I$$

for the corresponding *reflections* (also called *Cayley operators*), and note that the reflections are nonexpansive operators (T is nonexpansive if $\|Tx - Ty\| \leq \|x - y\|, \ \forall x, y \in H$).

In Eckstein and Bertsekas (1992) and He and Yuan (2015), the maximal monotone operator $S_{\lambda,A,B}$ is defined as

$$S_{\lambda,A,B} := \{(v + \lambda b, u - v) : (u, b) \in B, (v, a) \in A, v + \lambda a = u - \lambda b\}.$$

It was shown in Eckstein and Bertsekas (1992) that the Douglas–Rachford splitting method (5) can be converted to

$$z_{k+1} = \left(J_A^\lambda (2J_B^\lambda - I) + (I - J_B^\lambda) \right) z_k = (I + S_{\lambda,A,B})^{-1} z_k = J_{S_{\lambda,A,B}}(z_k).$$

By Eckstein and Bertsekas (1992, Thm. 5), for any given zero z^* of $S_{\lambda,A,B}$, $J_B^\lambda(z^*)$ is a zero of $A + B$. Therefore, $J_B^\lambda(z^*)$ is a solution of (1) whenever z^* satisfies

$$z^* = R_A^\lambda o R_B^\lambda(z^*). \tag{7}$$

Consequently, the Douglas–Rachford splitting method (5) can be rewritten as

$$\begin{aligned}
z_{k+1} &= J_A^\lambda (2J_B^\lambda - I) z_k + (I - J_B^\lambda) z_k \\
&= z_k + \frac{1}{2}(2J_A^\lambda (2J_B^\lambda(z_k) - z_k) - (2J_B^\lambda(z_k) - z_k) - z_k) \\
&= z_k + \frac{1}{2}(R_A^\lambda o R_B^\lambda(z_k) - z_k) \\
&= z_k - e(z_k, \lambda),
\end{aligned} \tag{8}$$

where $e(z_k, \lambda) := \frac{1}{2}(z_k - R_A^\lambda o R_B^\lambda(z_k))$.

In this paper, our convergence analysis will be conducted for an inertial generalized version of Douglas–Rachford splitting method (8): $z_0, z_1 \in H$,

$$\begin{cases} y_k = \alpha_k z_0 + (1 - \alpha_k) z_k + \theta_k (z_k - z_{k-1}) \\ z_{k+1} = y_k - \beta_k e(y_k, \lambda), \end{cases} \tag{9}$$

with $\alpha_k \in [0, 1), \beta_k \in (0, 1]$ and $\theta_k \in [0, 1)$. We get the original Douglas–Rachford method (8) when $\beta_k = 1, \theta_k = 0 = \alpha_k$ in (9).

We next recall some properties of the projection. For any point $u \in H$, there exists a unique point $P_C u \in C$ such that

$$\|u - P_C u\| \leq \|u - y\|, \quad \forall y \in C.$$

P_C is called the *metric projection* of H onto C. We know that P_C is a nonexpansive mapping of H onto C. It is also known that P_C satisfies

$$\langle x - y, P_C x - P_C y \rangle \geq \|P_C x - P_C y\|^2 \quad \forall x, y, \in H. \tag{10}$$

In particular, we get from (10) that

$$\langle x - y, x - P_C y \rangle \geq \|x - P_C y\|^2, \quad \forall x \in C, y \in H. \tag{11}$$

Furthermore, $P_C x$ is characterized by the properties

$$P_C x \in C \quad \text{and} \quad \langle x - P_C x, P_C x - y \rangle \geq 0, \, \forall y \in C. \tag{12}$$

This characterization implies that

$$\|x - y\|^2 \geq \|x - P_C x\|^2 + \|y - P_C x\|^2 \quad \forall x \in H, \forall y \in C. \tag{13}$$

The following result is obtained (Shehu et al. 2020) but we give the proof for the sake of completeness.

Lemma 2.1 *Let $S \subseteq H$ be a nonempty, closed, and convex subset of a real Hilbert space H. Let $u \in H$ be arbitrarily given, $z := P_S u$, and $\Omega := \{x \in H : \langle x - u, x - z \rangle \leq 0\}$. Then $\Omega \cap S = \{z\}$.*

Proof By definition, it follows immediately that $z \in \Omega \cap S$. Conversely, take an arbitrary $y \in \Omega \cap S$. Then, in particular, we have $y \in \Omega$, and it therefore follows that

$$\begin{align} \|y - z\|^2 &= \langle y - z, y - z \rangle \\ &= \langle y - z, y - u \rangle + \langle y - z, u - z \rangle \\ &\leq \langle y - z, u - z \rangle. \end{align} \tag{14}$$

Using $z = P_S u$ together with the characterization (12), we also have

$$\langle u - z, z - x \rangle \geq 0 \quad \forall x \in S.$$

In particular, since $y \in S$, we therefore have $\langle u - z, z - y \rangle \geq 0$. Hence (14) implies $\|y - z\|^2 \leq 0$, so that $y = z$. This completes the proof. □

Finally, we state some basic properties that will be used in our convergence theorems.

Lemma 2.2 *The following statements hold in H*:

(a) $\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2$ for all $x, y \in H$.
(b) $2\langle x - y, x - z \rangle = \|x - y\|^2 + \|x - z\|^2 - \|y - z\|^2$ for all $x, y, z \in H$.
(c) $\|tx + sy\|^2 = t(t + s)\|x\|^2 + s(t + s)\|y\|^2 - st\|x - y\|^2, \quad \forall x, y \in H, \forall s, t \in \mathbb{R}$.

Lemma 2.3 (Maingé 2008) *Assume that* $\varphi_k \in [0, \infty)$ *and* $\delta_k \in [0, \infty)$ *satisfy:*

(1) $\varphi_{k+1} - \varphi_n \leq \theta_k(\varphi_k - \varphi_{k-1}) + \delta_k$,
(2) $\sum_{k=1}^{\infty} \delta_k < \infty$,
(3) $\{\theta_k\} \subset [0, \theta]$, *where* $\theta \in (0, 1)$.

Then the sequence $\{\varphi_k\}$ *is convergent with* $\sum_{k=1}^{\infty}[\varphi_{k+1} - \varphi_k]_+ < \infty$, *where* $[t]_+ := \max\{t, 0\}$ *(for any* $t \in \mathbb{R}$).

3 Analysis of the convergence

For the rest of this paper, we assume that $S \neq \emptyset$, $\alpha_k \in (0, 1)$ with $\lim_{k \to \infty} \alpha_k = 0$ and $\sum_{k=1}^{\infty} \alpha_k = \infty$, $0 < \beta \leq \beta_k \leq 1$ and $0 \leq \theta_k \leq \theta_{k+1} \leq \theta < \frac{1}{3}$.

Lemma 3.1 *Let* $\{z_k\}$ *be the sequence generated by* (9). *For any z satisfying* (7), *we have*

$$\|z_{k+1} - z\|^2 \leq \|y_k - z\|^2 - \|z_{k+1} - y_k\|^2. \tag{15}$$

Proof By (9), we get

$$\begin{aligned}
\|z_{k+1} - z\|^2 &= \|y_k - z - \beta_k e(y_k, \lambda)\|^2 \\
&= \|y_k - z\|^2 - 2\beta_k \langle y_k - z, e(y_k, \lambda) \rangle + \beta_k^2 \|e(y_k, \lambda)\|^2.
\end{aligned} \tag{16}$$

We know that $e(y_k, \lambda) = \frac{1}{2}(y_k - R_A^\lambda o R_B^\lambda(y_k))$, where $\lambda > 0$ is the proximal parameter, is firmly-nonexpansive (see, He and Yuan 2015, lem. 2.2). Thus,

$$\langle x - y, e(x, \lambda) - e(y, \lambda) \rangle \geq \|e(x, \lambda) - e(y, \lambda)\|^2, \quad \forall x, y \in H, \lambda > 0.$$

In particular, for $z = R_A^\lambda o R_B^\lambda(z)$, we obtain

$$\langle y_k - z, e(y_k, \lambda) \rangle \geq \|e(y_k, \lambda)\|^2. \tag{17}$$

Putting (17) into (16), we have

$$\|z_{k+1} - z\|^2 \le \|y_k - z\|^2 - 2\beta_k \|e(y_k, \lambda)\|^2 + \beta_k^2 \|e(y_k, \lambda)\|^2$$
$$= \|y_k - z\|^2 - \beta_k(2 - \beta_k)\|e(y_k, \lambda)\|^2. \tag{18}$$

Recall that $\beta_k e(y_k, \lambda) = y_k - z_{k+1}$ implies that

$$e(y_k, \lambda) = \frac{1}{\beta_k}(y_k - z_{k+1}). \tag{19}$$

Using (19) in (18) and the condition that $0 < \beta \le \beta_k \le 1$, we have

$$\|z_{k+1} - z\|^2 \le \|y_k - z\|^2 - \beta_k(2 - \beta_k)\frac{1}{\beta_k^2}\|z_{k+1} - y_k\|^2$$
$$= \|y_k - z\|^2 - \frac{2 - \beta_k}{\beta_k}\|z_{k+1} - y_k\|^2$$
$$\le \|y_k - z\|^2 - \|z_{k+1} - y_k\|^2.$$

$\square$

Lemma 3.2 *Let $\{z_k\}$ be the sequence generated by (9). For any z satisfying (7), we have*

$$- 2\alpha_k\langle z_k - z, z_k - z_0\rangle$$
$$\ge \|z_{k+1} - z\|^2 - \|z_k - z\|^2 + 2\theta_{k+1}\|z_{k+1} - z_k\|^2 - 2\theta_k\|z_k - z_{k-1}\|^2$$
$$+ \alpha_{k+1}\|z_0 - z_{k+1}\|^2 - \alpha_k\|z_k - z_0\|^2 - \theta_k\|z_k - z\|^2 + \theta_{k-1}\|z_{k-1} - z\|^2 \tag{20}$$
$$+ (1 - 3\theta_{k+1} - \alpha_k)\|z_k - z_{k+1}\|^2.$$

Proof Moreover, from the definition of y_k, we obtain using Lemma 2.2 (a) that

$$\|y_k - z\|^2 = \|(z_k - z) + \theta_k(z_k - z_{k-1}) - \alpha_k(z_k - z_0)\|^2$$
$$= \|z_k - z\|^2 + \|\theta_k(z_k - z_{k-1}) - \alpha_k(z_k - z_0)\|^2$$
$$+ 2\langle z_k - z, \theta_k(z_k - z_{k-1}) - \alpha_k(z_k - z_0)\rangle \tag{21}$$
$$= \|z_k - z\|^2 + 2\theta_k\langle z_k - z, z_k - z_{k-1}\rangle - 2\alpha_k\langle z_k - z, z_k - z_0\rangle$$
$$+ \|\theta_k(z_k - z_{k-1}) - \alpha_k(z_k - z_0)\|^2,$$

and, similarly, with z replaced by z_{k+1} in the previous formula,

$$\|y_k - z_{k+1}\|^2$$
$$= \|z_k - z_{k+1}\|^2 + 2\theta_k\langle z_k - z_{k+1}, z_k - z_{k-1}\rangle \tag{22}$$
$$- 2\alpha_k\langle z_k - z_{k+1}, z_k - z_0\rangle + \|\theta_k(z_k - z_{k-1}) - \alpha_k(z_k - z_0)\|^2.$$

Substituting (21) and (22) into (15) and eliminating identical terms, we get

$$\|z_{k+1} - z\|^2$$
$$\leq \|z_k - z\|^2 + 2\theta_k \langle z_k - z, z_k - z_{k-1}\rangle$$
$$- 2\alpha_k \langle z_k - z, z_k - z_0\rangle - \|z_k - z_{k+1}\|^2$$
$$- 2\theta_k \langle z_k - z_{k+1}, z_k - z_{k-1}\rangle + 2\alpha_k \langle z_k - z_{k+1}, z_k - z_0\rangle$$
$$= \|z_k - z\|^2 + 2\theta_k \langle z_k - z, z_k - z_{k-1}\rangle$$
$$- 2\alpha_k \langle z_k - z, z_k - z_0\rangle - \|z_k - z_{k+1}\|^2 + \theta_k \|z_k - z_{k+1}\|^2 + \theta_k \|z_k - z_{k-1}\|^2$$
$$- \theta_k \|z_k - z_{k+1} + (z_k - z_{k-1})\|^2 + 2\alpha_k \langle z_k - z_{k+1}, z_k - z_0\rangle. \tag{23}$$

Therefore, we obtain

$$\|z_{k+1} - z\|^2 - \|z_k - z\|^2 - \theta_k \|z_k - z_{k-1}\|^2 + (1 - \theta_k)\|z_k - z_{k+1}\|^2$$
$$\leq -2\alpha_k \langle z_k - z, z_k - z_0\rangle + 2\theta_k \langle z_k - z, z_k - z_{k-1}\rangle + 2\alpha_k \langle z_k - z_{k+1}, z_k - z_0\rangle$$
$$= -2\alpha_k \langle z_k - z, z_k - z_0\rangle - \theta_k \|z_{k-1} - z\|^2 + \theta_k \|z_k - z\|^2 + \theta_k \|z_k - z_{k-1}\|^2$$
$$- \alpha_k \|z_0 - z_{k+1}\|^2 + \alpha_k \|z_{k+1} - z_k\|^2 + \alpha_k \|z_k - z_0\|^2, \tag{24}$$

where the last identity exploits Lemma 2.2 (a) twice. We therefore have

$$-2\alpha_k \langle z_k - z, z_k - z_0\rangle \tag{25}$$

$$\geq \|z_{k+1} - z\|^2 - \|z_k - z\|^2 + 2\theta_{k+1}\|z_{k+1} - z_k\|^2 - 2\theta_k \|z_k - z_{k-1}\|^2$$
$$+ \theta_k \big(\|z_{k-1} - z\|^2 - \|z_k - z\|^2\big) + \alpha_k \big(\|z_0 - z_{k+1}\|^2 - \|z_k - z_0\|^2\big) \tag{26}$$
$$+ (1 - \theta_k - 2\theta_{k+1} - \alpha_k)\|z_{k+1} - z_k\|^2.$$

Using the fact that $\{\theta_k\}$ is non-decreasing and $\{\alpha_k\}$ is non-increasing, we then obtain

$$-2\alpha_k \langle z_k - z, z_k - z_0\rangle$$
$$\geq \|z_{k+1} - z\|^2 - \|z_k - z\|^2 + 2\theta_{k+1}\|z_{k+1} - z_k\|^2 - 2\theta_k \|z_k - z_{k-1}\|^2$$
$$+ \alpha_{k+1}\|z_0 - z_{k+1}\|^2 - \alpha_k \|z_k - z_0\|^2 - \theta_k \|z_k - z\|^2 + \theta_{k-1}\|z_{k-1} - z\|^2$$
$$+ (1 - 3\theta_{k+1} - \alpha_k)\|z_k - z_{k+1}\|^2,$$

which is the desired inequality. □

Our first central result below shows that the sequence $\{z_k\}$ generated by (9) is bounded.

Lemma 3.3 *The sequence $\{z_k\}$ generated by (9) is bounded.*

Proof A simple re-ordering of (20) implies that

$$\|z_{k+1} - z\|^2 - \|z_k - z\|^2$$
$$\leq \theta_k\|z_k - z\|^2 - \theta_{k-1}\|z_{k-1} - z\|^2 - (1 - 3\theta_{k+1} - \alpha_k)\|z_k - z_{k+1}\|^2$$
$$\quad - 2\theta_{k+1}\|z_{k+1} - z_k\|^2 + 2\theta_k\|z_k - z_{k-1}\|^2 - \alpha_{k+1}\|z_0 - z_{k+1}\|^2$$
$$\quad + \alpha_k\|z_k - z_0\|^2 - 2\alpha_k\langle z_k - z_0, z_k - z\rangle \qquad (27)$$
$$= \theta_k\|z_k - z\|^2 - \theta_{k-1}\|z_{k-1} - z\|^2 - (1 - 3\theta_{k+1} - \alpha_k)\|z_k - z_{k+1}\|^2$$
$$\quad - 2\theta_{k+1}\|z_{k+1} - z_k\|^2 + 2\theta_k\|z_k - z_{k-1}\|^2 - \alpha_{k+1}\|z_0 - z_{k+1}\|^2$$
$$\quad + \alpha_k\|z_k - z_0\|^2 + \alpha_k\|z_0 - z\|^2 - \alpha_k\|z_k - z_0\|^2 - \alpha_k\|z_k - z\|^2,$$

where the equality uses once again Lemma 2.2 (a). Hence, by cancellation, re-ordering, and neglecting a non-positive term on the right-hand side, we obtain

$$\|z_{k+1} - z\|^2 - \|z_k - z\|^2 + \alpha_k\|z_k - z\|^2$$
$$\leq \theta_k\|z_k - z\|^2 - \theta_{k-1}\|z_{k-1} - z\|^2 - (1 - 3\theta_{k+1} - \alpha_k)\|z_k - z_{k+1}\|^2 \qquad (28)$$
$$\quad - 2\theta_{k+1}\|z_{k+1} - z_k\|^2 + 2\theta_k\|z_k - z_{k-1}\|^2 + \alpha_k\|z_0 - z\|^2.$$

Let $\mu_j := e^{\sum_{i=1}^j \alpha_i}, j \geq 1$. Using $1 - x \leq e^{-x}$ for all $x \in \mathbb{R}$ (or equivalently, $1 - e^{-x} \leq x, x \in \mathbb{R}$), we obtain

$$\frac{1}{\mu_{k+1}}(\mu_{k+1} - \mu_k) = 1 - \frac{\mu_k}{\mu_{k+1}}$$
$$= 1 - e^{(\sum_{i=1}^k \alpha_i - \sum_{i=1}^{k+1} \alpha_i)} \qquad (29)$$
$$= 1 - e^{-\alpha_{k+1}} \leq \alpha_{k+1}.$$

Then (29) consequently implies that

$$\frac{1}{\mu_{k+1}}\left(\mu_{k+1}\|z_{k+1} - z\|^2 - \mu_k\|z_k - z\|^2\right)$$
$$= \|z_{k+1} - z\|^2 - \|z_k - z\|^2 + \frac{1}{\mu_{k+1}}(\mu_{k+1} - \mu_k)\|z_k - z\|^2$$
$$\leq \|z_{k+1} - z\|^2 - \|z_k - z\|^2 + \alpha_{k+1}\|z_k - z\|^2.$$

Since $\{\alpha_k\}$ is non-increasing in (0,1), this implies

$$\frac{1}{\mu_{k+1}}\left(\mu_{k+1}\|z_{k+1} - z\|^2 - \mu_k\|z_k - z\|^2\right)$$
$$\leq \|z_{k+1} - z\|^2 - \|z_k - z\|^2 + \alpha_k\|z_k - z\|^2. \qquad (30)$$

It then follows from (28) and (30) that

$$\frac{1}{\mu_{k+1}} \left(\mu_{k+1} \|z_{k+1} - z\|^2 - \mu_k \|z_k - z\|^2 \right)$$
$$\leq \theta_k \|z_k - z\|^2 - \theta_{k-1} \|z_{k-1} - z\|^2$$
$$- (1 - 3\theta_{k+1} - \alpha_k) \|z_k - z_{k+1}\|^2 - 2\theta_{k+1} \|z_{k+1} - z_k\|^2$$
$$+ 2\theta_k \|z_k - z_{k-1}\|^2 + \alpha_k \|z_0 - z\|^2.$$

Since $\mu_k \leq \mu_{k+1}$, $\mu_{k+1} = \mu_k e^{\alpha_{k+1}}$ and $\{\alpha_k\}$ is non-increasing in $(0,1)$, we therefore get

$$\mu_{k+1} \|z_{k+1} - z\|^2 - \mu_k \|z_k - z\|^2$$
$$\leq \mu_{k+1} \theta_k \|z_k - z\|^2 - \mu_k \theta_{k-1} \|z_{k-1} - z\|^2 - \mu_{k+1}(1 - 3\theta_{k+1} - \alpha_k)\|z_{k+1} - z_k\|^2$$
$$- 2\mu_{k+1}\theta_{k+1}\|z_{k+1} - z_k\|^2 + 2\mu_k \theta_k e^{\alpha_{k+1}} \|z_k - z_{k-1}\|^2 + \mu_{k+1}\alpha_k \|z_0 - z\|^2,$$

which can be rewritten as (since $\{\alpha_k\}$ is non-increasing in $(0,1)$)

$$\mu_{k+1} \|z_{k+1} - z\|^2 - \mu_k \|z_k - z\|^2$$
$$\leq \mu_{k+1}\theta_k \|z_k - z\|^2 - \mu_k \theta_{k-1} \|z_{k-1} - z\|^2$$
$$- \mu_{k+1}\left[1 - \theta_{k+1}\left(3 + 2(e^{\alpha_{k+1}} - 1)\right) - \alpha_k\right]\|z_{k+1} - z_k\|^2$$
$$- 2\mu_{k+1}\theta_{k+1}e^{\alpha_{k+1}}\|z_{k+1} - z_k\|^2 + 2\mu_k \theta_k e^{\alpha_k}\|z_k - z_{k-1}\|^2 + \mu_{k+1}\alpha_k \|z_0 - z\|^2.$$

Since the sequence $\{\theta_k\}$ belongs to the interval $[0, \theta]$, we have

$$1 - \theta_{k+1}\left(3 + 2(e^{\alpha_{k+1}} - 1)\right) - \alpha_k \geq 1 - \theta\left(3 + 2(e^{\alpha_{k+1}} - 1)\right) - \alpha_k, \quad \forall k \in \mathbb{N}.$$

Using $\lim_{k \to \infty} \alpha_k = 0$ and $\theta \in [0, 1/3)$, it follows that the right-hand side is eventually bounded from below by a positive number, i.e., there is a constant $\gamma > 0$ such that $1 - \theta_{k+1}\left(3 + 2(e^{\alpha_{k+1}} - 1)\right) - \alpha_k \geq \gamma$ for all $k \in \mathbb{N}$ sufficiently large, say, for all $k \geq k_0$. Hence, we have

$$\mu_{k+1} \|z_{k+1} - z\|^2 - \mu_k \|z_k - z\|^2$$
$$\leq \mu_{k+1}\theta_k \|z_k - z\|^2 - \mu_k \theta_{k-1} \|z_{k-1} - z\|^2 - 2\mu_{k+1}\theta_{k+1}e^{\alpha_{k+1}}\|z_{k+1} - z_k\|^2$$
$$- \gamma \mu_{k+1}\|z_{k+1} - z_k\|^2 + 2\mu_k \theta_k e^{\alpha_k}\|z_k - z_{k-1}\|^2 + \mu_{k+1}\alpha_k \|z_0 - z\|^2.$$

This implies that for $k \geq k_0$,

$$\|z_0 - z\|^2 \sum_{j=k_0+1}^{k} \mu_{j+1}\alpha_j$$
$$\geq \mu_{k+1}\|z_{k+1} - z\|^2 + 2\mu_{k+1}\theta_{k+1}e^{\alpha_{k+1}}\|z_{k+1} - z_k\|^2 - \mu_{k+1}\theta_k \|z_k - z\|^2 \quad (31)$$
$$- \mu_{k_0+1}\|z_{k_0+1} - z\|^2 - 2\mu_{k_0+1}\theta_{k_0+1}e^{\alpha_{k_0+1}}\|z_{k_0+1} - z_{k_0}\|^2$$
$$+ \mu_{k_0+1}\theta_{k_0}\|z_{k_0} - z\|^2.$$

Thus, dividing by μ_{k+1} and omitting a non-positive term, we get

$$\|z_{k+1} - z\|^2 - \theta_k \|z_k - z\|^2$$
$$\leq e^{-t_{k+1}} \left[\mu_{k_0+1} \|z_{k_0+1} - z\|^2 + 2\mu_{k_0+1}\theta_{k_0+1} e^{\alpha_{k_0+1}} \|z_{k_0+1} - z_{k_0}\|^2 \right. \tag{32}$$
$$\left. - \mu_{k_0+1}\theta_{k_0} \|z_{k_0} - z\|^2 \right] + \|z_0 - z\|^2 e^{-t_{k+1}} \sum_{j=k_0+1}^{k} \alpha_j e^{t_j+1},$$

where $t_k := \sum_{i=1}^{k} \alpha_i$. Since $\alpha_k \in (0,1)$ for all $k \in \mathbb{N}$, it is easy to see that $\alpha_k e^{t_{k+1}} \leq e^2(e^{t_k} - e^{t_{k-1}})$ for all $k \geq 2$, so that

$$\sum_{j=k_0+1}^{k} \mu_{j+1}\alpha_j = \sum_{j=k_0+1}^{k} \alpha_j e^{t_j+1} \leq e^2 \left(e^{t_k} - e^{t_{k_0}} \right) \leq e^2 e^{t_k},$$

which, by (32), $e^{-t_{k+1}} \leq 1$, and the fact that $\{\theta_k\}$ belongs to the interval $[0, \theta] \subset [0, \frac{1}{3})$, yields

$$\|z_{k+1} - z\|^2$$
$$\leq \theta \|z_k - z\|^2 + \mu_{k_0+1} \|z_{k_0+1} - z\|^2 + 2\mu_{k_0+1}\theta_{k_0+1} e^{\alpha_{k_0+1}} \|z_{k_0+1} - z_{k_0}\|^2 \tag{33}$$
$$+ e^2 \|z_0 - z\|^2.$$

Using (33), $\theta \in [0,1)$, and the convergence of the geometric series, a simple calculation gives

$$\|z_{k+1} - z\|^2 \leq \theta^{k-k_0} \|z_{k_0+1} - z\|^2 + \frac{1}{1-\theta} \left[\mu_{k_0+1} \|z_{k_0+1} - z\|^2 \right.$$
$$\left. + 2\mu_{k_0+1}\theta_{k_0+1} e^{\alpha_{k_0+1}} \|z_{k_0+1} - z_{k_0}\|^2 + e^2 \|z_0 - z\|^2 \right].$$

Using once again that $\theta < 1$, this shows that $\{z_k\}$ is bounded. $\qquad \square$

Next, we formulate a simple lemma that turns out to be useful for proving the strong convergence result.

Lemma 3.4 *Let $\{z_k\}$ be the sequence generated by* (9). *Define*

$$u_k := \|z_k - z\|^2 - \theta_{k-1} \|z_{k-1} - z\|^2 + 2\theta_k \|z_k - z_{k-1}\|^2 + \alpha_k \|z_k - z_0\|^2$$

for all $k \in \mathbb{N}$. Then $u_k \geq 0$ for all $k \in \mathbb{N}$.

Proof Since $\{\theta_k\}$ is non-decreasing with $0 \leq \theta_k < \frac{1}{3}$, and by Lemma 2.2 (a), we have

$$
\begin{aligned}
u_k &= \|z_k - z\|^2 - \theta_{k-1}\|z_{k-1} - z_k + z_k - z\|^2 + 2\theta_k\|z_k - z_{k-1}\|^2 + \alpha_k\|z_k - z_0\|^2 \\
&= \|z_k - z\|^2 - \theta_{k-1}\big[\|z_{k-1} - z_k\|^2 + \|z_k - z\|^2 + 2\langle z_{k-1} - z_k, z_k - z\rangle\big] \\
&\quad + 2\theta_k\|z_k - z_{k-1}\|^2 + \alpha_k\|z_k - z_0\|^2 \\
&= \|z_k - z\|^2 - \theta_{k-1}\big[2\|z_{k-1} - z_k\|^2 + 2\|z_k - z\|^2 - \|z_{k-1} - 2z_k - z\|^2\big] \\
&\quad + 2\theta_k\|z_k - z_{k-1}\|^2 + \alpha_k\|z_k - z_0\|^2 \\
&= \|z_k - z\|^2 - 2\theta_{k-1}\|z_{k-1} - z_k\|^2 - 2\theta_{k-1}\|z_k - z\|^2 + \theta_{k-1}\|z_{k-1} - 2z_k - z\|^2 \\
&\quad + 2\theta_k\|z_k - z_{k-1}\|^2 + \alpha_k\|z_k - z_0\|^2 \\
&\geq \|z_k - z\|^2 - 2\theta_k\|z_{k-1} - z_k\|^2 - \frac{2}{3}\|z_k - z\|^2 + \theta_{k-1}\|z_{k-1} - 2z_k - z\|^2 \\
&\quad + 2\theta_k\|z_k - z_{k-1}\|^2 + \alpha_k\|z_k - z_0\|^2 \\
&\geq \frac{1}{3}\|z_k - z\|^2 + \alpha_k\|z_k - z_0\|^2 \\
&\geq 0,
\end{aligned}
$$

and this completes the proof. □

Before we prove our main strong convergence result, we state another preliminary result which provides sufficient conditions for the strong convergence of the sequence $\{z_k\}$ generated by our method (9). In our strong convergence result, we will then show that these sufficient conditions automatically hold.

Lemma 3.5 *Let $\{z_k\}$ be the sequence generated by* (9). *Assume that*

$$
\lim_{k\to\infty} \|z_{k+1} - z_k\| = 0
$$

and

$$
\lim_{k\to\infty}(\|z_{k+1} - z\|^2 - \theta_k\|z_k - z\|^2) = 0.
$$

Then the entire sequence $\{z_k\}$ converges strongly to the solution z.

Proof By assumption, we have

$$
\begin{aligned}
0 &= \lim_{k\to\infty}(\|z_{k+1} - z\|^2 - \theta_k\|z_k - z\|^2) \\
&= \lim_{k\to\infty}\Big[(\|z_{k+1} - z\| + \sqrt{\theta_k}\|z_k - z\|)(\|z_{k+1} - z\| - \sqrt{\theta_k}\|z_k - z\|)\Big].
\end{aligned}
\tag{34}
$$

We claim that this already implies

$$
\lim_{k\to\infty}(\|z_{k+1} - z\| + \sqrt{\theta_k}\|z_k - z\|) = 0,
$$

from which the strong convergence of the entire sequence $\{z_k\}$ to z follows immediately. Assume this limit does not hold. Then there is a subset $K \subseteq \mathbb{N}$ and a constant $\rho > 0$ such that

$$\|z_{k+1} - z\| + \sqrt{\theta_k}\|z_k - z\| \geq \rho, \forall k \in K. \tag{35}$$

Since $\lim_{k \to \infty} \|z_{k+1} - z_k\| = 0$ by the assumption and $0 \leq \theta < 1$, then (recall that if $\{a_k\}$ and $\{b_k\}$ are bounded sequences in $\mathbb{R}$ and one of either $\{a_k\}$ or $\{b_k\}$ converges, then $\lim \sup_{k \to \infty}(a_k + b_k) = \lim \sup_{k \to \infty} a_k + \lim \sup_{k \to \infty} b_k$)

$$\lim_{k \in K} \sup((1 - \sqrt{\theta})\|z_k - z\| - \|z_{k+1} - z_k\|) = \lim_{k \in K} \sup(1 - \sqrt{\theta})\|z_k - z\| - \lim_{k \in K}\|z_{k+1} - z_k\|$$

$$= (1 - \sqrt{\theta})\lim_{k \in K}\sup\|z_k - z\| - \lim_{k \in K}\|z_{k+1} - z_k\|.$$

Using (34) and $\theta_k \leq \theta < 1$, we get

$$0 = \lim_{k \in K}(\|z_{k+1} - z\| - \sqrt{\theta_k}\|z_k - z\|)$$

$$= \lim_{k \in K}\sup(\|z_{k+1} - z_k + z_k - z\| - \sqrt{\theta_k}\|z_k - z\|)$$

$$\geq \lim_{k \in K}\sup(\|z_k - z\| - \|z_{k+1} - z_k\| - \sqrt{\theta_k}\|z_k - z\|)$$

$$\geq \lim_{k \in K}\sup((1 - \sqrt{\theta})\|z_k - z\| - \|z_{k+1} - z_k\|)$$

$$= (1 - \sqrt{\theta})\lim_{k \in K}\sup\|z_k - z\| - \lim_{k \in K}\|z_{k+1} - z_k\|$$

$$= (1 - \sqrt{\theta})\lim_{k \in K}\sup\|z_k - z\|.$$

Consequently, we have $\lim \sup_{k \in K} \|z_k - z\| \leq 0$. Since $\lim \inf_{k \in K} \|z_k - z\| \geq 0$ obviously holds, it follows that $\lim_{k \in K} \|z_k - z\| = 0$. This implies [by (35)]

$$\|z_{k+1} - z_k\| \geq \|z_{k+1} - z\| - \|z_k - z\|$$

$$= \|z_{k+1} - z\| + \sqrt{\theta_k}\|z_k - z\| - (1 + \sqrt{\theta_k})\|z_k - z\|$$

$$\geq \frac{\rho}{2}$$

for all $k \in K$ sufficiently large, a contradiction to the assumption that $\lim_{k \to \infty} \|z_{k+1} - z_k\| = 0$. This completes the proof. $\qquad \square$

We are now ready to obtain strong convergence of the sequence $\{z_k\}$ generated by (9) to an element of S.

Theorem 3.6 *The sequence $\{z_k\}$ generated by (9) strongly converges to z, where $z = P_S z_0$.*

Proof Let u_k denote the nonnegative number defined in Lemma 3.4, and let us apply Lemma 3.2. We obtain from (20) that

$$
\begin{aligned}
u_{k+1} - u_k &+ (1 - 3\theta_{k+1} - \alpha_k)\|z_k - z_{k+1}\|^2 \\
&\leq -2\alpha_k \langle z_k - z, z_k - z_0 \rangle.
\end{aligned}
\tag{36}
$$

We now consider two cases.

Case 1 Suppose $\{u_k\}$ is eventually a monotonically decreasing sequence, i.e. for some $k_0 \in \mathbb{N}$ large enough, we have $u_{k+1} \leq u_k$ for all $k \geq k_0$. Then, since u_k is nonnegative for all $k \in \mathbb{N}$ by Lemma 3.4, we obviously get that $\{u_k\}$ is a convergent sequence. Consequently, it follows that $\lim_{k\to\infty} u_k = \lim_{k\to\infty} u_{k+1}$. Since $\{z_k\}$ is bounded by Theorem 3.3, there exists $M > 0$ such that $2|\langle z_k - z, z_k - z_0 \rangle| \leq M$. Moreover, it follows that there exist $N \in \mathbb{N}$ and $\gamma_1 > 0$ such that $1 - 3\theta_{k+1} - \alpha_k \geq \gamma_1$ for all $k \geq N$. Therefore, for $k \geq N$, we obtain from (36) that

$$
\begin{aligned}
\gamma_1 \|z_{k+1} - z_k\|^2 &\leq \alpha_k M + u_k - u_{k+1} \\
&\to 0 \quad \text{for } k \to \infty.
\end{aligned}
$$

Hence

$$
\lim_{k\to\infty} \|z_{k+1} - z_k\| \to 0.
$$

Together with $\alpha_k \to 0$, the boundedness of $\{z_k\}$, and the convergence of $\{u_k\}$, we therefore obtain from the definition of u_k that the limit

$$
\lambda := \lim_{k\to\infty} \left(\|z_{k+1} - z\|^2 - \theta_k \|z_k - z\|^2 \right)
\tag{37}
$$

exists and is equal to $\lim_{k\to\infty} u_{k+1}$. In particular, Lemma 3.4 therefore implies that $\lambda \geq 0$. We will show that $\lambda = 0$ holds; then (37) together with the fact that $\theta_k \leq \theta < 1$ for all $k \in \mathbb{N}$ yields the strong convergence of the sequence $\{z_k\}$ to the solution z.

By contradiction, assume that $\lambda > 0$. Since $\{z_k\}$ is bounded by Theorem 3.3, it is easy to see that we can choose a subsequence $\{z_{k_j}\}$ which converges weakly to an element $p \in H$ and such that

$$
\liminf_{k\to\infty} \langle z_k - z, z - z_0 \rangle = \lim_{j\to\infty} \langle z_{k_j} - z, z - z_0 \rangle = \langle p - z, z - z_0 \rangle.
$$

We show that $p \in S$. Observe that the updating rule for y_k implies

$$
\begin{aligned}
\|y_k - z_k\| = \|\alpha_k(z_0 - z_k) &+ \theta_k(z_k - z_{k-1})\| \\
&\leq \alpha_k \|z_0 - z_k\| + \theta_k \|z_k - z_{k-1}\| \to 0, \quad k \to \infty.
\end{aligned}
$$

This yields

$$
\|e(y_k, \lambda)\| \leq \frac{1}{\beta} \|z_{k+1} - y_k\| \leq \|z_k - y_k\| + \|z_{k+1} - z_k\| \to 0, \quad k \to \infty.
$$

Let $Ty := \frac{1}{2}y + \frac{1}{2}R_A^\lambda oR_B^\lambda(y)$, $y \in H$. Then it is clear that T is nonexpansive and $z \in F(T) := \{x \in H : x = Tx\}$ if and only if $z = R_A^\lambda oR_B^\lambda(z)$. Similarly, it is easy to see that $e(y_k, \lambda) = \frac{1}{2}(y_k - R_A^\lambda oR_B^\lambda(y_k)) = y_k - Ty_k$. Therefore,

$$\lim_{k \to \infty} \|y_k - Ty_k\| = \lim_{k \to \infty} \|e(y_k, \lambda)\| = 0.$$

Demiclosedness Principle of T implies that $p \in F(T)$. Hence, $p \in S$. This implies that

$$\liminf_{k \to \infty} \langle z_k - z, z - z_0 \rangle = \langle p - z, z - z_0 \rangle \geq 0, \tag{38}$$

where the inequality follows from the characterization (12) of a projection applied to $z = P_S z_0$ and $p \in S$. Since (37) yields

$$\liminf_{k \to \infty} \|z_{k+1} - z\|^2 \geq \lim_{k \to \infty} \left(\|z_{k+1} - z\|^2 - \theta_k \|z_k - z\|^2 \right) = \lambda,$$

and since $\lambda > 0$ by assumption, we have

$$\|z_{k+1} - z\|^2 \geq \frac{1}{2}\lambda \quad \forall k \geq k_1$$

for some sufficiently large $k_1 \in \mathbb{N}$. Using the identity

$$\langle z_k - z, z_k - z_0 \rangle = \|z_k - z\|^2 + \langle z_k - z, z - z_0 \rangle,$$

we therefore get

$$\begin{aligned}
\liminf_{k \to \infty} \langle z_k - z, z_k - z_0 \rangle &= \liminf_{k \to \infty} \left(\|z_k - z\|^2 + \langle z_k - z, z - z_0 \rangle \right) \\
&\geq \liminf_{k \to \infty} \left(\frac{1}{2}\lambda + \langle z_k - z, z - z_0 \rangle \right) \\
&= \frac{1}{2}\lambda + \liminf_{k \to \infty} \langle z_k - z, z - z_0 \rangle \\
&\geq \frac{1}{2}\lambda
\end{aligned}$$

from (38). Using once again the assumption that $\lambda > 0$, this implies

$$\langle z_k - z, z_k - z_0 \rangle \geq \frac{1}{4}\lambda \quad \forall k \geq k_2$$

for some sufficiently large $k_2 \in \mathbb{N}, k_2 \geq k_1$. From (36), we therefore obtain

$$u_{k+1} - u_k \leq -\frac{1}{2}\alpha_k \lambda \quad \forall k \geq k_2.$$

This implies

$$\frac{1}{2}\lambda \sum_{j=k_2}^{k} \alpha_j \leq u_{k_2} - u_{k+1} \leq u_{k_2} \quad \forall k \geq k_2,$$

where the second inequality follows from Lemma 3.4. Since $\lambda > 0$, this gives the summability of the sequence $\{\alpha_k\}$, a contradiction to our assumption. Hence we must have $\lambda = 0$, and this yields the strong convergence of the sequence $\{z_k\}$ to z.

Case 2 Assume that $\{u_k\}$ is not eventually monotonically decreasing. Then let $\tau : \mathbb{N} \to \mathbb{N}$ be the map defined for all $k \geq k_0$ (for some $k_0 \in \mathbb{N}$ large enough) by

$$\tau(k) := \max\{j \in \mathbb{N} : j \leq k, u_j \leq u_{j+1}\}. \tag{39}$$

Clearly, $\tau(k)$ is a non-decreasing sequence such that $\tau(k) \to \infty$ for $k \to \infty$ and $u_{\tau(k)} \leq u_{\tau(k)+1}$ for all $k \geq k_0$. Hence, similar to the proof of Case 1, we therefore obtain from (36) that

$$\gamma_1 \|x_{\tau(k)+1} - x_{\tau(k)}\|^2 \leq \alpha_{\tau(k)} M \to 0 \tag{40}$$

for some constant $M > 0$. Thus,

$$\|x_{\tau(k)+1} - x_{\tau(k)}\| \to 0, \quad k \to \infty. \tag{41}$$

Using the same technique of the proof as in Case 1, one can also derive the limits

$$\|x_{\tau(k)+1} - w_{\tau(k)}\| \to 0, \quad k \to \infty,$$
$$\|w_{\tau(k)} - x_{\tau(n)}\| \to 0, \quad k \to \infty, \tag{42}$$

$$\|x_{\tau(k)} - z_{\tau(k)}\| \to 0, \quad k \to \infty. \tag{43}$$

Again observe that for $j \geq 0$ by (36), we have $u_{j+1} < u_j$ when $x_j \notin \Omega := \{x \in H : \langle x - z_0, x - z \rangle \leq 0\}$ (note that this Ω is the same set as in Lemma 2.1). Hence $x_{\tau(k)} \in \Omega$ for all $k \geq k_0$ since $u_{\tau(k)} \leq u_{\tau(k)+1}$. Since $\{x_{\tau(k)}\}$ is bounded, we may choose a subsequence (which we again call $\{x_{\tau(k)}\}$) which converges weakly to some $x^* \in H$. As Ω is a closed and convex set, it is then weakly closed and so $x^* \in \Omega$. Using (43), one can see as in Case 1 that $z_{\tau(k)} \rightharpoonup x^*$ and $x^* \in S$. Consequently, we have $x^* \in \Omega \cap S$. In view of Lemma 2.1, however, the intersection $\Omega \cap S$ contains z as its only element. We therefore get $x^* = z$. Furthermore, we have

$$\|x_{\tau(k)} - z\|^2 = \langle x_{\tau(k)} - z_0, x_{\tau(k)} - z \rangle - \langle z - z_0, x_{\tau(k)} - z \rangle$$
$$\leq - \langle z - z_0, x_{\tau(k)} - z \rangle$$

since $x_{\tau(k)} \in \Omega$. Taking lim sup in this last inequality gives

$$\limsup_{k \to \infty} \|x_{\tau(k)} - z\| \leq 0.$$

Hence

$$\|x_{\tau(k)} - z\| \to 0, \quad k \to \infty. \tag{44}$$

We claim that this implies $\lim_{k \to \infty} u_{\tau(k)+1} = 0$. By definition, $u_{\tau(k)+1}$ is equal to

$$\|x_{\tau(k)+1} - z\|^2 - \theta_{\tau(k)}\|x_{\tau(k)} - z\|^2 + 2\theta_{\tau(k)+1}\|x_{\tau(k)+1} - x_{\tau(k)}\|^2 + \alpha_{\tau(k)+1}\|x_{\tau(k)+1} - z_0\|^2.$$

Adding and subtracting $x_{\tau(k)}$ inside the norm of the first term, and using (41), (44), we see that the first term goes to zero. The second term converges to zero also in view of (44), taking into account the boundedness of $\{\theta_k\}$. The third term vanishes in the limit because of (41) and noting once again that $\{\theta_k\}$ is a bounded sequence. Finally, the last term goes to zero since $\{\alpha_k\}$ converges to zero and the sequence $\{z_k\}$ is bounded by Theorem 3.3.

We next show that we actually have $\lim_{k\to\infty} u_k = 0$. To this end, first observe that, for $k \geq k_0$, one has $u_k \leq u_{\tau(k)+1}$ if $k \neq \tau(k)$ (that is, if $\tau(k) < k$) because we necessarily have $u_j > u_{j+1}$ for $\tau(k) + 1 \leq j \leq k - 1$. It follows that for all $k \geq k_0$, we have $u_k \leq \max\{u_{\tau(k)}, u_{\tau(k)+1}\} = u_{\tau(k)+1} \to 0$, hence $\limsup_{k\to\infty} u_k \leq 0$. On the other hand, Lemma 3.4 implies that $\liminf_{k\to\infty} u_k \geq 0$. Together we obtain $\lim_{k\to\infty} u_k = 0$.

Consequently, the boundedness of $\{z_k\}$, assumptions on our iterative parameters and (36) show that

$$\|z_k - z_{k+1}\| \to 0, \quad k \to \infty.$$

Hence the definition of u_k yields

$$\lim_{k\to\infty} \left(\|z_{k+1} - z\|^2 - \theta_k\|z_k - z\|^2 \right) = 0.$$

Using our assumption, it is not difficult to see that this implies the strong convergence of the entire sequence $\{z_k\}$ to the particular solution z. The statement therefore follows from Lemma 3.5. $\qquad\square$

In the special case when B is a set-valued maximal monotone operator and A is a single-valued κ-inverse strongly monotone operator in problem (1), iterative procedure (9) reduces to the following: $z_0, z_1 \in H$,

$$\begin{cases} y_k = \alpha_k z_0 + (1 - \alpha_k)z_k + \theta_k(z_k - z_{k-1}) \\ z_{k+1} = (1 - \beta_k)y_k + \beta_k(I + \lambda B)^{-1}(I - \lambda A)y_k, \end{cases} \tag{45}$$

with $0 < \lambda < 2\kappa$. Moreover, we obtain strong convergence for this special case of monotone inclusion for which its proof can be easily obtained by following line of arguments of previous lemmas and Theorem 3.6.

Corollary 3.7 *Suppose B is a set-valued maximal monotone operator and A is a single-valued κ-inverse strongly monotone operator. Assume that $S := \{x \in H : 0 \in Ax + Bx\} \neq \emptyset$. Let $\{z_k\}$ be the sequence generated by (45) with $0 < \beta \leq \beta_k \leq \frac{1}{2}$, $0 < \lambda < 2\kappa$ and $0 \leq \theta_k \leq \theta_{k+1} \leq \theta < \frac{1}{3}$. Then $\{z_k\}$ strongly converges to z, where $z = P_S z_0$.*

We next relate our results to some existing results from the literature.

Remark 3.8

(a) In the results of Thong and Vinh (see Thong and Vinh 2019, Thm. 3.5), strong convergence for monotone inclusion was obtained under some assumptions on the iterative sequence. The monotone inclusion studied in Thong and Vinh (2019) involves sum of a set-valued maximal monotone operator and single-valued inverse-strongly monotone operator. In this paper, our method is proposed such that no assumption is made on the iterative sequence even for a more general result considered here.

(b) The Algorithm (45) could be taken as the inertial strong convergence version of some recent results in Attouch and Cabot (2019), Boţ and Csetnek (2016), Lorenz and Pock (2015) and Villa et al. (2013). ◊

4 Numerical experiments

In all the examples in this section, we compare our proposed method (9) with the non-inertial version (when $\theta_n = 0$), Thong and Vinh results (see Thong and Vinh 2019, Thm. 3.5) and Shehu (2016). Our aim is to compare our method with other relevant strong convergence methods in the literature.

Example 4.1 Let $H = L^2([0, 1])$. Let $A := \partial \|.\|$ and $B = N_C$ in (1), where N_C is the normal cone of nonempty closed and convex subset C of H ($N_C(x) := \{x^* \in H : \langle y - x, x^* \rangle \leq 0, \forall y \in C\}$), . Then problem (1) reduces to the following minimization problem: find $x^* \in L^2([0, 1])$ such that

$$0 \in \partial \|x^*\| + N_C(x^*). \tag{46}$$

Note that $S \neq \emptyset$ since $0 \in S$. Furthermore, the resolvent $J_B^\lambda = (I + \lambda N_C)^{-1} = P_C$, and J_A^λ is given by the Moreau decomposition

$$
\begin{aligned}
J_A^\lambda(x) &= (I + \lambda \partial \|.\|)^{-1}(x) \\
&= \text{Prox}_{\lambda \|.\|}(x) = x - \lambda P_{B_{\|.\|_*}}\left(\frac{x}{\lambda}\right),
\end{aligned}
$$

where $\text{Prox}_{\lambda \|.\|}(x) := \text{argmin}_y \{\|y\| + \frac{1}{2}\|y - x\|^2\}$, $P_{B_{\|.\|_*}}$ is the projection operator and $B_{\|.\|_*}$ is the norm unit ball(of the dual norm). Note that in this case, $L^2([0, 1])$ is self dual. Moreover, the projection $P_{B_{\|.\|_*}}$ (see Bauschke and Combettes 2011; Cegielski 2012) is given by:

$$
P_{B_{\|.\|_*}}(x) = \begin{cases} \frac{x}{\|x\|}, & \|x\| > 1 \\ x, & \|x\| \leq 1. \end{cases}
$$

Therefore,

$$J_A^\lambda(x) = x - \lambda P_{B_{\|\cdot\|}}\left(\frac{x}{\lambda}\right) = \begin{cases} x - \lambda\frac{x}{\|x\|}, & \|\frac{x}{\lambda}\| > 1 \\ 0, & \|\frac{x}{\lambda}\| \leq 1. \end{cases}$$

We take C as the ball $C := \{x \in H : \|x - z\| \leq r\}$, then

$$P_C(x) = \begin{cases} x, & \|x - z\| \leq r \\ z + \frac{r(x-z)}{\|x-z\|}, & \|x - z\| > r \end{cases}$$

In particular, $C = \{x \in L^2([0,1]) : \int_0^1 |x(t) - \sin(\frac{t}{2\pi})|^2 dt \leq 16\}$.

Set $\lambda = 0.02$, $\beta_k = 0.6$ and $\alpha_k = 100/k$. Take $\|z_k - z_{k-1}\| \leq 10^{-3}$ as the stopping criterion (Fig. 1).

Example 4.2 Suppose $A : \mathbb{R}^3 \to \mathbb{R}^3$ and $B : \mathbb{R}^3 \to \mathbb{R}^3$ are given by

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 10 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad B\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

It can be shown that $\Omega = \{(0, 0, 0)\}$.

Let z_0 be randomly selected. In Algorithm 3 of Thong and Vinh (2019), we chose $\lambda = 0.1$, $\beta_k = 1/(k + 1)$ and $\tau_k = 1/(k + 1)^2$. In Algorithm (26) of Shehu (2016), we chose $\alpha_k = 1/k$, $\beta_k = k/(2k + 1)$ and $r_k = 0.1$. Take $\|z_k\| \leq 0.005$ as the stopping criterion.

For Examples 4.2 and 4.3, we take $\lambda = 0.2$, $\beta_k = 0.5$, $\alpha_k = \frac{1}{25k}$ in Algorithm (9) and $\beta = 0.2$ in Algorithm 3 of Thong and Vinh (2019).

We compared the algorithm (9), Algorithm 3 in Thong and Vinh (2019) and the algorithm (26) in Shehu (2016). From Fig. 2, we know that the performance of the algorithm (9) is better than that of the other two algorithms.

Example 4.3 Let us consider the following well known ℓ_1-regularized least squares problem, which consists of finding a sparse solution to an underdetermined linear system. Suppose that we solve the following problem:

$$\min \frac{1}{2}\|Dx - b\|_2^2 + \rho\|x\|_1, \tag{47}$$

where $D \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. In this case,

$$J_A^\lambda(x) = (D^t D + \lambda^{-1}I)^{-1}(D^t b + \lambda^{-1}x)$$

while

$$J_B^\lambda(x) = (sign(x_i).\max\{0, |x_i| - \lambda\rho\})_i, \quad i = 1, 2, \ldots, n.$$

We remark that there is a commercial software, based on the projected gradient method for solving problem (47), for example SPGL1 (van den Berg and Friedlander
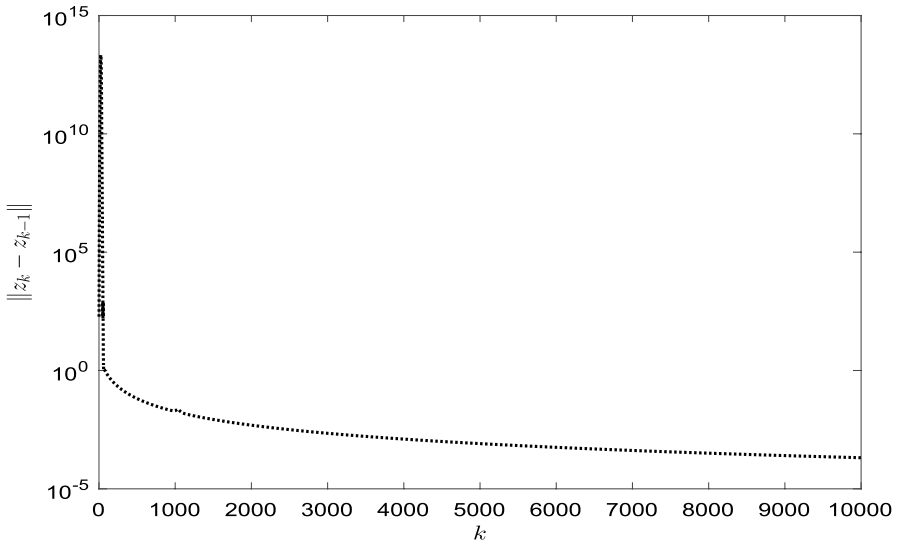
Fig. 1 Error attenuation trend of Algorithm (9)

2007; Lorenz 2013) and FISTA (Beck and Teboulle 2009), but this is beyond the scope of this paper. Our interest here is to demonstrate the efficiency of our proposed method (9) using problem (47).

We generate random problems using different choices of λ for $m = 100$ and $n = 1000$. In Algorithm 3 of Thong and Vinh (2019) and the algorithm (26) of Shehu (2016), we chose $\rho = 1$ and $\lambda = 1.9/(\max(eig(D^T D)))$, and in the algorithm (9), we chose $\rho = 0.5$. In addition, select $r_k = 0.2$ in algorithm (26) of Shehu (2016).

Table 3 shows that the algorithm (9) is better when $\theta_k = 0.33$. The numerical result is described in Fig. 3, it illustrates that the performance of Algorithm (9) is better than the other two algorithms.
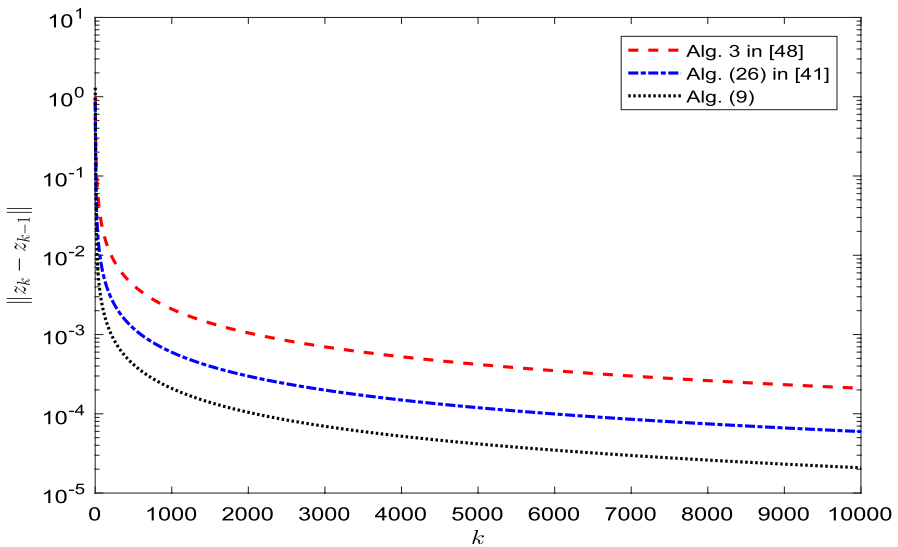
Remark 4.4

(a) It can be seen from the numerical examples that Algorithm (9) outperforms the methods in Shehu (2016) and Thong and Vinh (2019) (see Figs. 2, 3) for strong convergence of sum of maximal monotone operators. Furthermore, the additional of inertial term improves the acceleration of the proposed method as can be seen in the numerical examples that Algorithm (9) converges faster than the non-inertial case when $\theta_k = 0$ (please see, Tables 1, 2, 3). Also, the optimum choice of $\theta_k = 0$ should be close to the upper bound $\frac{1}{3}$ from our examples.

(b) Algorithm (9) is sensitive to the choice of the initial point z_0 as can be seen in our examples in Tables 1, 2 and 3.

Table 1 Comparing the change of θ_k under the same initial value for Example 4.1

z_0	θ_k	0	0.05	0.1	0.15	0.2	0.25	0.3	0.33
t	Iter.	76	70	66	64	64	67	67	71
	CPU time	0.0331	0.0516	0.0482	0.0490	0.0503	0.0473	0.0490	0.0493
$20 * e^t$	Iter.	4501	4501	4501	4501	4501	4501	4501	4501
	CPU time	1.9981	1.8217	1.8206	1.7393	1.7787	1.7789	1.7985	1.8972
$-1000 * \sin(\pi t)$	Iter.	21603	21603	21603	21603	21603	21603	21603	21603
	CPU time	9.4897	8.2634	8.1801	8.2655	8.2255	8.3546	8.2791	8.2253

Table 2 Comparing the change of θ_k under the same initial value for Example 4.2

| z_0	θ_k	0	0.05	0.1	0.15	0.2	0.25	0.3	0.33
$[1, 1, 1]^T$	Iter.	45	45	45	45	45	44	44	44
	CPU time	0.0261	0.0274	0.0255	0.0248	0.0254	0.0266	0.0249	0.0252
$[100, 100, 10]^T$	Iter.	3497	3497	3496	3496	3496	3496	3496	3496
	CPU time	0.3556	0.3515	0.3595	0.3807	0.3761	0.3802	0.3981	0.3782
$[-100, 20, 1000]^T$	Iter.	23005	23005	23004	23004	23004	23004	23004	23004
	CPU time	2.3476	2.3244	2.2244	2.2110	2.2698	2.2905	2.2292	2.2087

Fig. 2 Comparison of three algorithms

Remark 4.5

(a) We point out that there are different strategies in the current literature to enforce strong convergence on proximal-like algorithms (in particular, DR splitting); see, e.g., Solodov and Svaiter (2000) and Hirstoaga (2006). In this regard, the results

Table 3 Comparing the change of θ_k under the same initial value for Example 4.3

z_0	θ_k	0	0.05	0.1	0.15	0.2	0.25	0.3	0.33
$[1,...,1]^T$	Iter.	209	201	203	201	200	194	189	184
	CPU time	64.0070	61.1515	61.1118	62.5140	60.4273	58.5795	56.7839	55.8204
$[1,-20,...,1,-20]^T$	Iter.	589	617	592	570	545	523	500	485
	CPU time	184.3802	187.0893	178.6727	171.9601	165.8236	158.0672	151.0111	147.1679
$[20,100,...,20,100]^T$	Iter.	2324	2238	2134	2030	1922	1817	1715	1657
	CPU time	726.3275	670.4111	640.5669	608.4853	577.9857	546.6380	514.9161	498.6941
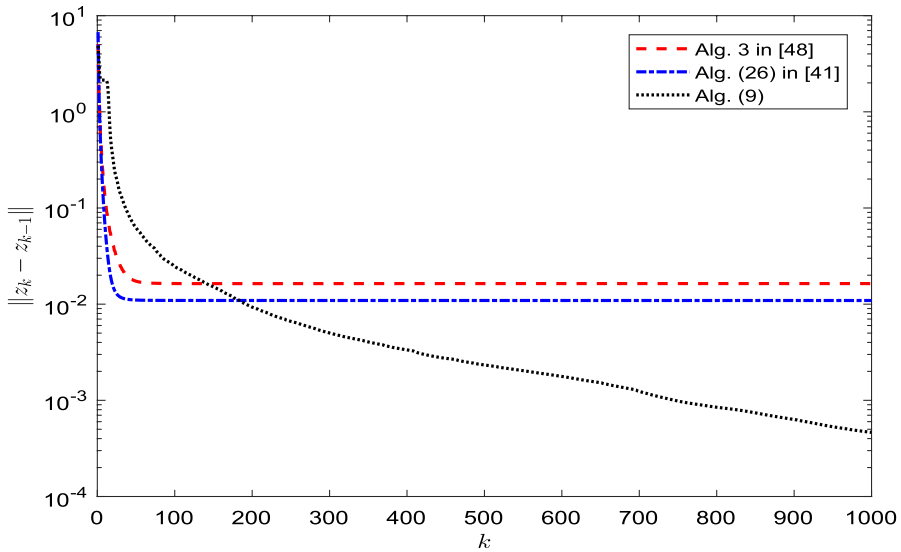
Fig. 3 Comparison of three algorithms

of Hirstoaga (2006) is concerned with "anchor-point" algorithms as employed in our proposed method (9). One can see in Algorithm 2.1 of Hirstoaga (2006), there is no presence of inertial extrapolation term, $\theta_k(z_k - z_{k-1})$, which has been proved in the literature to increase the speed of convergence of non-inertial counterpart in most optimization methods. From our proposed method (9), we see that when $\theta_k \neq 0$ (in this paper, we assume that $0 \leq \theta_k \leq \theta < \frac{1}{3}$), then our method (9) cannot be reduced to Algorithm 2.1 of Hirstoaga (2006) applied to the splitting operator of Eckstein and Bertsekas (1992). As confirmed in our numerical examples in Sect. 4, our method (9) outperforms Algorithm 2.1 of Hirstoaga (2006) when applied to the splitting operator of Eckstein and Bertsekas (1992). Also, our method of proof is different from the method of proof given in Hirstoaga (2006).

(b) The essence of our numerical examples in Sect. 4 is to drive home the implementations and effectiveness of our proposed method (9). As discussed in MacNamara and Strang (2016) and other related chapters in the book, applications of our method (9) to solve problems arising from wireless communications, imaging, networking, finance, hemodynamics, free-surface flows, and other science and engineering problems in infinite-dimensional Hilbert spaces would be discussed separately as a future project. ◊

5 Final remarks

In this paper we propose a Douglas–Rachford splitting method with inertial extrapolation step and give strong convergence analysis of the method. The method is much more applicable for a general class of maximal monotone operators and no uniform

monotonicity on any of the involved maximal monotone operators is assumed. Furthermore, the analysis of the algorithm is obtained under the natural condition of the inertial factor θ_k being monotone non-decreasing and bounded away from 1/3. Some numerical illustrations are given to test the efficiency and implemnetation of the proposed scheme. The results obtained in this paper could serve as the strong convergence counterpart of already obtained weak convergence methods for inertial Douglas–Rachford splitting methods (Bauschke and Combettes 2011; Beck and Teboulle 2009; Boţ et al. 2015; Lorenz and Pock 2015; Thong and Vinh 2019) in the literature.

Our future project include the following:

- to modify the proposed method (9) in this paper so that the bound of the inertial factor θ_k could exceed 1/3 and possibly lead to a faster convergence; and
- to obtain the rate of convergence of method (9). As far as we know, this has not been obtained before in the literature.

References

Alvarez F, Attouch H (2001) An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. Set-Valued Anal 9:3–11

Attouch H, Cabot A (2019) Convergence of a relaxed inertial forward–backward algorithm for structured monotone inclusions. Appl Math Optim. https://doi.org/10.1007/s00245-019-09584-z

Bauschke HH, Combettes PL (2001) A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces. Math Oper Res 26:248–264

Bauschke HH, Combettes PL (2011) Convex analysis and monotone operator theory in Hilbert spaces. CMS books in mathematics. Springer, New York

Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imaging Sci 2(1):183–202

Boikanyo OA (2016) The viscosity approximation forward–backward splitting method for zeros of the sum of monotone operators. Abstr Appl Anal. Article ID 2371857

Boţ RI, Csetnek ER (2016) An inertial forward-backward-forward primal-dual splitting algorithm for solving monotone inclusion problems. Numer. Algorithms 71:519–540

Boţ RI, Csetnek ER, Hendrich C (2015) Inertial Douglas–Rachford splitting for monotone inclusion problems. Appl Math Comput 256:472–487

Cegielski A (2012) Iterative methods for fixed point problems in Hilbert spaces. Lecture notes in mathematics 2057. Springer, Berlin

Chang S-S, Wen C-F, Yao J-C (2019) A generalized forward-backward splitting method for solving a system of quasi variational inclusions in Banach spaces. RACSAM 113:729–747

Cholamjiak P (2016) A generalized forward-backward splitting method for solving quasi inclusion problems in Banach spaces. Numer. Algorithms 71:915–932

Cholamjiak W, Cholamjiak P, Prasit SS (2018) An inertial forward-backward splitting method for solving inclusion problems in Hilbert spaces. J. Fixed Point Theory Appl. 20:42

Combettes PL (2004) Solving monotone inclusions via compositions of nonexpansive averaged operators. Optimization 53:475–504

Dong Q, Jiang D, Cholamjiak P, Shehu Y (2017) A strong convergence result involving an inertial forward–backward algorithm for monotone inclusions. J. Fixed Point Theory Appl. 19:3097–3118

Dong QL, Cho YJ, Zhong Zhong LL, Rassias TM (2018) Inertial projection and contraction algorithms for variational inequalities. J. Glob. Optim. 70(3):687–704

Douglas J, Rachford HH (1956) On the numerical solution of heat conduction problems in two or three space variables. Trans. Am. Math. Soc. 82:421–439

Eckstein J (1989) Splitting methods for monotone operators with applications to parallel optimization. Doctoral dissertation, Department of Civil Engineering, Massachusetts Institute of Technology. Available as Report LIDS-TH-1877, Laboratory for Information and Decision Sciences, MIT, Cambridge

Eckstein J, Bertsekas DP (1992) On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. Math. Program. 55(3):293–318

Fukushima M (1996) The primal Douglas–Rachford splitting algorithm for a class of monotone mappings with application to the traffic equilibrium problem. Math. Program. 72:1–15

Gabay D, Mercier B (1976) A dual algorithm for the solution of nonlinear variational problems via finite element approximations. Comput. Math. Appl. 2:17–40

Gibali A, Thong DV (2018) Tseng type methods for solving inclusion problems and its applications. Calcolo 55:49

Glowinski R, Le Tallec P (1989) Augmented Lagrangian and operator-splitting methods in nonlinear mechanics. SIAM Studies in Applied Mathematics, vol 9, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA. ISBN: 0-89871-230-0

Glowinski R, Marrocco A (1975) Approximation par é'ements finis d'ordre un et résolution par pénalisation-dualité d'une classe de problémes non linéaires. R.A.I.R.O. R2:41–76

Güler O (1991) On the convergence of the proximal point algorithm for convex minimization. SIAM J. Control Optim. 29:403–419

He B, Yuan X (2015) On the convergence rate of Douglas–Rachford operator splitting method. Math. Program. 153:715–722

Hirstoaga SA (2006) Iterative selection methods for common fixed point problems. J. Math. Anal. Appl. 324(2):1020–1035

Lions PL, Mercier B (1979) Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. 16:964–979

López G, Martín-Márquez V, Wang F, Xu H-K (2012) Forward–backward splitting methods for accretive operators in Banach spaces. Abstr. Appl. Anal. Article ID 109236

Lorenz DA (2013) Constructing test instances for Basis Pursuit Denoising. IEEE Trans. Signal Process. 61:1210–1214

Lorenz DA, Pock T (2015) An inertial forward–backward algorithm for monotone inclusions. J. Math. Imaging Vis. 51:311–325

MacNamara S, Strang G (2016) Operator splitting. In: Glowinski R, Osher S, Yin W (eds) Splitting methods in communication, imaging, science, and engineering. Springer, Berlin, pp 95–114

Maingé PE (2008) Convergence theorem for inertial KM-type algorithms. J. Comput. Appl. Math. 219(1):223–236

Martinet B (1970) Regularisation dequations variationnelles par approximations successives. Rev. Francaise Informat. Recherche. Operationnelle 4:154–158

Moreau JJ (1965) Proximité et dualit'e dans un espace Hilbertien. Bull. Soc. Math. Fr. 93:273–299

Riahi H, Chbani Z, Loumi M-T (2018) Weak and strong convergences of the generalized penalty Forward–Forward and Forward–Backward splitting algorithms for solving bilevel hierarchical pseudomonotone equilibrium problems. Optimization 67:1745–1767

Rockafellar RT (1976) Monotone operators and the proximal point algorithm. SIAM J. Control Optim. 14:877–898

Shehu Y (2016) Iterative approximations for zeros of sum of accretive operators in Banach spaces. J Funct Spaces. Article ID 5973468

Shehu Y (2018) Convergence rate analysis of inertial Krasnoselskii–Mann-type iteration with applications. Numer. Funct. Anal. Optim. 39:1077–1091

Shehu Y (2019) Convergence results of forward–backward algorithms for sum of monotone operators in Banach spaces. Results Math. 74:138

Shehu Y, Cai G (2018) Strong convergence result of forward–backward splitting methods for accretive operators in Banach spaces with applications. RACSAM 112:71–87

Shehu Y, Li X-H, Dong Q-L (2020) An efficient projection-type method for monotone variational inequalities in Hilbert spaces. Numer. Algorithms 84:365–388

Solodov MV, Svaiter BF (2000) Forcing strong convergence of proximal point iterations in a Hilbert space. Math. Program. Ser. A 87:189–202

Svaiter BF (2011) On weak convergence of the Douglas–Rachford method. SIAM J. Control Optim. 49:280–287

Thong DV, Cholamjiak P (2019) Strong convergence of a forward–backward splitting method with a new step size for solving monotone inclusions. Comput. Appl. Math. 38:94

Thong DV, Vinh NT (2019) Inertial methods for fixed point problems and zero point problems of the sum of two monotone mappings. Optimization 68:1037–1072

van den Berg E, Friedlander MP (2007) SPGL1: a solver for large-scale sparse reconstruction. http://www.cs.ubc.ca/labs/scl/spgl1. Version 1.9. Accessed 2015

Villa S, Salzo S, Baldassarres L, Verri A (2013) Accelerated and inexact forward–backward. SIAM J. Optim. 23:1607–1633

Wang Y, Wang F (2018) Strong convergence of the forward–backward splitting method with multiple parameters in Hilbert spaces. Optimization 67:493–505

Zhang H, Cheng L (2013) Projective splitting methods for sums of maximal monotone operators with applications. J. Math. Anal. Appl. 406:323–334