Journal of Computational and Applied Mathematics 460 (2025) 116405

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Convergence of an inertial reflected–forward–backward splitting algorithm for solving monotone inclusion problems with application to image recovery

Chinedu Izuchukwu [a], Simeon Reich [b], Yekini Shehu [c,*]

[a] School of Mathematics, University of the Witwatersrand, Private Bag 3, Johannesburg, 2050, South Africa
[b] Department of Mathematics, The Technion – Israel Institute of Technology, 32000 Haifa, Israel
[c] School of Mathematical Sciences, Zhejiang Normal University, Jinhua 321004, People's Republic of China

ARTICLE INFO

MSC:
47H09
47H10
49J20
49J40

Keywords:
Image restoration problem
Inertial method
Monotone inclusion
Monotone operator
Optimal control
Reflected–forward–backward algorithm

ABSTRACT

We first propose a reflected–forward–backward splitting algorithm with two inertial effects for solving monotone inclusions and then establish that the sequence of iterates it generates converges weakly in a real Hilbert space to a zero of the sum of a set-valued maximal monotone operator and a single-valued monotone Lipschitz continuous operator. The proposed algorithm involves only one forward evaluation of the single-valued operator and one backward evaluation of the set-valued operator at each iteration. One inertial parameter is non-negative while the other is non-positive. These features are absent in many other available inertial splitting algorithms in the literature. Finally, we discuss some problems in image restoration in connection with the implementation of our algorithm and compare it with some known related algorithms in the literature.

1. Introduction

Let $\mathcal{H}$ be a real Hilbert space, and let $A : \mathcal{H} \to 2^{\mathcal{H}}$ and $B : \mathcal{H} \to \mathcal{H}$ be two monotone operators. The problem of finding a zero of the sum of A and B, which is also known as the monotone inclusion problem (MIP), is defined as follows:

$$\text{find } \hat{x} \in \mathcal{H} \text{ such that } \mathbf{0} \in (A+B)\hat{x}. \tag{1.1}$$

We denote the solution set of this problem by $(A+B)^{-1}(\mathbf{0})$.

The monotone inclusion is an important problem in optimization as well as in signal processing, image recovery, and machine learning. It is often solved by splitting algorithms that involve tackling each of the two associated operators (A and B) separately (by means of forward evaluation of the single-valued operator and backward evaluation of the set-valued operator) rather than their sum. These algorithms have undergone tremendous research which has been motivated by the desire to devise faster, computationally inexpensive, and much more applicable methods.

Among these splitting methods, we now mention the following forward–backward splitting algorithm [1,2]:

$$x_{n+1} = (I_{\mathcal{H}} + \gamma A)^{-1}(x_n - \gamma B x_n), \ \ n \geq 1, \tag{1.2}$$

* Corresponding author.
E-mail addresses: chinedu.izuchukwu@wits.ac.za (C. Izuchukwu), sreich@technion.ac.il (S. Reich), yekini.shehu@zjnu.edu.cn (Y. Shehu).

https://doi.org/10.1016/j.cam.2024.116405
Received 12 December 2023; Received in revised form 22 November 2024
Available online 29 November 2024

where $I_{\mathcal{H}}$ is the identity operator on $\mathcal{H}$ and γ is a positive constant. This algorithm involves only one forward evaluation of B and one backward evaluation of A per iteration, and the sequences it generates are known to converge weakly to a solution of the inclusion problem (1.1) when B is L^{-1}-cocoercive (which is a strict condition), $\gamma \in (0, 2L^{-1})$, A is maximal monotone and $(A+B)^{-1}(\mathbf{0})$ is nonempty.

The strict cocoercivity assumption in algorithm (1.2) was removed in [3], where the author proposed the following forward–backward–forward (FBF) splitting algorithm (also known as Tseng's splitting method):

$$\begin{cases} y_n = (I_{\mathcal{H}} + \gamma A)^{-1}(x_n - \gamma B x_n), \\ x_{n+1} = y_n - \gamma B y_n + \gamma B x_n, \ n \geq 1. \end{cases} \tag{1.3}$$

An inexact variant, a variable metric variant and a stochastic variant of algorithm (1.3) were studied in [4,5], and [6], respectively. The main advantage of algorithm (1.3) is that the sequences it generates converge weakly to a solution of (1.1) under the much weaker assumptions that B is monotone and L-Lipschitz continuous, $\gamma \in (0, L^{-1})$, A is maximal monotone and $(A + B)^{-1}(\mathbf{0}) \neq \emptyset$. However, the main disadvantage of this algorithm is that it involves two forward evaluations of B. This might affect its efficiency especially when it is applied to solving large-scale optimization problems and optimization problems emanating from optimal control theory, where computations involving pertinent operators are often very expensive (see [7]).

This disadvantage was recently resolved in [8], where the authors proposed the following forward-reflected–backward (FRB) splitting algorithm:

$$x_{n+1} = (I_{\mathcal{H}} + \gamma A)^{-1}(x_n - 2\gamma B x_n + \gamma B x_{n-1}), \ n \geq 1, \tag{1.4}$$

where $\gamma \in (0, \frac{1}{2}L^{-1})$. The main advantages of algorithm (1.4) are that it requires only one forward evaluation of B and that B is monotone and Lipschitz continuous. See [9] for the reflexive Banach space variant of algorithm (1.4).

If B is linear, algorithm (1.4) has the same structure as the reflected–forward–backward (RFB) splitting algorithm proposed and studied in [10]:

$$\begin{cases} y_n = 2x_n - x_{n-1}, \\ x_{n+1} = (I_{\mathcal{H}} + \gamma A)^{-1}(x_n - \gamma B y_n), \ n \geq 1. \end{cases} \tag{1.5}$$

The weak convergence of the sequences generated by algorithm (1.5) to a solution of (1.1) was investigated in [10] under the assumption that B is monotone and L-Lipschitz continuous, $\gamma \in (0, L^{-1}(\sqrt{2} - 1))$, A is maximal monotone and $(A + B)^{-1}(\mathbf{0}) \neq \emptyset$. The weak convergence of the sequences generated by algorithm (1.5) was also investigated in [11] in the particular case where A in the inclusion problem (1.1) is the normal cone of a nonempty, closed, and convex subset of $\mathcal{H}$. Also, when A is the subdifferential of a proper convex and lower semicontinuous function, the line search variant of algorithm (1.5) was investigated in [12].

Recently, fast convergent algorithms for solving optimization problems have become of great interest to many researchers. The inertial extrapolation technique is one of the most important methods of improving the convergence speed of optimization algorithms. This technique is based upon a discrete analog of a second-order dissipative dynamical system, which is known for its efficiency in improving the convergence speed of optimization algorithms (see [13–17]). For notable results on inertial algorithms with one inertial parameter for solving the MIP (1.1), see [18–20] and references therein.

Advantages of an inertial extrapolation step with two inertial parameters. Poon and Liang [21,22] presented some limitations of algorithms with one inertial parameter for solving feasibility problems in $\mathbb{R}^2$, where we define $T_k = T := \frac{1}{2}\Big(I + (2P_{H_1} - I)(2P_{H_2} - I)\Big)$.

Let $H_1, H_2 \subset \mathbb{R}^2$ be two subspaces such that $H_1 \cap H_2 \neq \emptyset$. Find $x \in \mathbb{R}^2$ such that $x \in H_1 \cap H_2$.

It was shown in [22, Section 4] that the two-step inertial fixed point iteration

$$x_{n+1} = T(x_n + \theta(x_n - x_{n-1}) + \delta(x_{n-1} - x_{n-2}))$$

converges faster, in terms of both the number of iterations and the CPU time, than the one-step inertial fixed point iteration

$$x_{n+1} = T(x_n + \theta(x_n - x_{n-1})). \tag{1.6}$$

Furthermore, it was shown that the sequences generated by the one-step inertial fixed point iteration (1.6) converge more slowly than those generated by its non-inertial version (that is, when $\theta = 0$ in (1.6)). This example, therefore, shows that the one-step inertial fixed point iteration (1.6) may fail to provide acceleration in the case where T is the Douglas–Rachford splitting operator. Therefore, for certain cases, the use of the inertia of more than two points could be of numerical advantage. It was noted in [23, Chapter 4] that the use of more than two points x_n, x_{n-1} could provide acceleration. For example, the two-step inertial extrapolation

$$w_n = x_n + \theta(x_n - x_{n-1}) + \delta(x_{n-1} - x_{n-2})$$

with $\theta > 0$ and $\delta < 0$ can provide acceleration. The failure of one-step inertial acceleration of ADMM was also discussed in [21, Section 3] and adaptive acceleration for ADMM (Alternating direction method of multipliers) was proposed instead. Polyak [24] also observed that the multi-step inertial methods can boost the speed of optimization methods even though neither the convergence nor the rate result of such multi-step inertial methods is established in [24]. Some results on multi-step inertial methods have recently been presented in [25].

In this paper we investigate the convergence of the sequences generated by an inertial variant of the RFB splitting algorithm (1.5), that is, we propose an RFB splitting algorithm with two inertial parameters in which one is non-negative while the other is non-positive and prove that the sequences generated by our proposed algorithm converge weakly to a solution of Problem (1.1). The proposed algorithm involves only one forward evaluation of the monotone Lipschitz continuous operator B and one backward evaluation of the maximal monotone operator A at each iteration; a feature that is absent in many other inertial splitting algorithms which are available in the literature. Our weak convergence results are obtained without imposing summability conditions on the inertial parameters and the sequence of iterates; such conditions are assumed in recent results on multi-step inertial methods in the literature (see, for instance, [25]). Moreover, some previously known related methods in the literature are recovered. Furthermore, we conducted numerical experiments for problems arising from image restoration. Results from these experiments show that our algorithm is efficient and faster than both the RFB splitting algorithm (1.5) and the FRB splitting algorithm (1.4).

2. Preliminaries

Throughout this work, $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ are the inner product and norm, respectively, of a real Hilbert space $\mathcal{H}$.

The operator $B:\mathcal{H}\to\mathcal{H}$ is called L-cocoercive (or inverse strongly monotone) if there exists a number $L>0$ such that

$$\langle Ba-Bb, a-b\rangle \geq L\|Ba-Bb\|^2 \ \forall a,b\in\mathcal{H},$$

and monotone if

$$\langle Ba-Bb, a-b\rangle \geq 0 \ \ \forall a,b\in\mathcal{H}.$$

The operator B is called L-Lipschitz continuous if there exists a number $L>0$ such that

$$\|Ba-Bb\| \leq L\|a-b\| \ \forall a,b\in\mathcal{H}.$$

If A is a set-valued operator $A:\mathcal{H}\to 2^{\mathcal{H}}$, then A is said to be monotone if

$$\langle \hat{a}-\hat{b}, a-b\rangle \geq 0 \ \ \forall a,b\in\mathcal{H},\ \hat{a}\in Aa,\ \hat{b}\in Ab.$$

The monotone operator A is called maximal if the graph $\mathcal{G}(A)$ of A, defined by

$$\mathcal{G}(A) := \{(a,\hat{a})\in\mathcal{H}\times\mathcal{H} : \hat{a}\in Aa\},$$

is not properly contained in the graph of any other monotone operator. In other words, A is called a maximal monotone operator if for $(a,\hat{a})\in\mathcal{H}\times\mathcal{H}$, we have that $\langle\hat{a}-\hat{b}, a-b\rangle\geq 0$ for all $(b,\hat{b})\in\mathcal{G}(A)$ implies that $\hat{a}\in Aa$.

For a set-valued operator A, the resolvent associated with it is the mapping $J^A_{\bar{\delta}}:\mathcal{H}\to 2^{\mathcal{H}}$ defined by

$$J^A_{\bar{\delta}}(a) := (I_{\mathcal{H}}+\bar{\delta}A)^{-1}(a), \quad a\in\mathcal{H},\ \bar{\delta}>0.$$

If A is maximal monotone and B is single-valued, then both $J^A_{\bar{\delta}}$ and $J^A_{\bar{\delta}}(I_{\mathcal{H}}-\bar{\delta}B)$ are single-valued and everywhere defined on $\mathcal{H}$.

The following identities hold in a real Hilbert space $\mathcal{H}$:

$$2\langle\hat{a},\hat{b}\rangle = \|\hat{a}\|^2+\|\hat{b}\|^2-\|\hat{a}-\hat{b}\|^2 = \|\hat{a}+\hat{b}\|^2-\|\hat{a}\|^2-\|\hat{b}\|^2 \quad \forall\ \hat{a},\hat{b}\in\mathcal{H}. \tag{2.1}$$

Lemma 2.1. *Let $x,y,z\in\mathcal{H}$ and $a,b\in\mathbb{R}$. Then*

$$\begin{aligned}\|(1+a)x-(a-b)y-bz\|^2 \;=\;& (1+a)\|x\|^2-(a-b)\|y\|^2-b\|z\|^2+(1+a)(a-b)\|x-y\|^2\\ &+b(1+a)\|x-z\|^2-b(a-b)\|y-z\|^2.\end{aligned}$$

3. Proposed algorithm and weak convergence

In this section we present our proposed algorithm and establish a weak convergence theorem for it. We begin with the following assumptions.

Assumption 3.1.

(a) A is maximal monotone,
(b) B is monotone and Lipschitz continuous with Lipschitz constant $L>0$,
(c) $(A+B)^{-1}(\mathbf{0})$ is nonempty.

Algorithm 1. *Inertial RFB Splitting Algorithm*

Let $\gamma>0$, $\theta\in[0,1)$ and $\delta\in(-\infty,0]$ be given. For arbitrary points $x_{-1},x_0,x_1\in\mathcal{H}$, let the sequence $\{x_n\}$ be generated by

$$\begin{cases} w_n = x_n+\theta(x_n-x_{n-1})+\delta(x_{n-1}-x_{n-2}),\\ y_n = 2x_n-x_{n-1},\\ x_{n+1} = J^A_{\gamma}(w_n-\gamma By_n),\ \ n\geq 1.\end{cases}$$

Remark 3.1. We recover the RFB splitting algorithm studied in [10] (please see (1.5) above) when we take $\theta = \delta = 0$ in our Algorithm 1. We also recover the projected reflected gradient method [11, Algorithm 3.1] when A in Algorithm 1 is a normal cone (in this case, J_γ^A is a projection operator onto a closed and convex set C) and $\theta = \delta = 0$.

Lemma 3.1. *Let $\{x_n\}$ be generated by Algorithm 1 and assume that Assumption 3.1 holds. For all $x \in (A+B)^{-1}(\mathbf{0})$, define*

$$\begin{aligned} a_n &:= \|x_n - x\|^2 - \theta\|x_{n-1} - x\|^2 - \delta\|x_{n-2} - x\|^2 + (1 - |\delta| + 2\theta)\|x_n - x_{n-1}\|^2 \\ &\quad + (\gamma L + \theta)\|x_n - y_{n-1}\|^2 + \|p_n + \gamma Bx\|^2 - \|p_n + x_n - x_{n-1} + \gamma Bx\|^2, \end{aligned} \tag{3.1}$$

where

$$p_n := w_{n-1} - \gamma B y_{n-1} - x_n. \tag{3.2}$$

Then

$$\begin{aligned} a_{n+1} &\le a_n + 2\theta\|x_{n+1} - x_n\|^2 + (|\delta| + |\delta|\theta - \delta\theta)\|x_{n-1} - x_{n-2}\|^2 \\ &\quad - (1 - 3\theta - \gamma L(1+\sqrt{2}))\left[\|x_n - y_n\|^2 + \|x_{n+1} - y_n\|^2\right] \\ &\quad - 2\delta(1+\theta)\|x_n - x_{n-1}\|^2 - 2\delta\langle x_{n-2} - x_{n-3}, x_{n+1} - y_n\rangle. \end{aligned}$$

Proof. Let $x \in (A+B)^{-1}(\mathbf{0})$. From $x_n = J_\gamma^A(w_{n-1} - \gamma B y_{n-1})$, the definition of J_γ^A and (3.2), we get $p_n \in \gamma A x_n$.
Similarly, $p_{n+1} \in \gamma A x_{n+1}$. Also, we have

$$2\langle x_n - w_{n-1} + p_n + \gamma B y_{n-1}, x_{n+1} - y_n\rangle = 0 \tag{3.3}$$

and

$$2\langle x_{n+1} - w_n + p_{n+1} + \gamma B y_n, x - x_{n+1}\rangle = 0. \tag{3.4}$$

Adding (3.3) and (3.4), we get

$$\begin{aligned} 0 &= 2\langle x_n - w_{n-1}, x_{n+1} - y_n\rangle + 2\langle p_n + \gamma B y_{n-1}, x_{n+1} - y_n\rangle \\ &\quad + 2\langle x_{n+1} - w_n, x - x_{n+1}\rangle + 2\langle p_{n+1} + \gamma B y_n, x - x_{n+1}\rangle \\ &= 2\langle x_n - w_{n-1}, x_{n+1} - y_n\rangle + 2\langle p_n + \gamma B y_n, x_{n+1} - y_n\rangle + 2\gamma\langle B y_{n-1} - B y_n, x_{n+1} - y_n\rangle \\ &\quad + 2\langle x_{n+1} - w_n, x - x_{n+1}\rangle + 2\langle p_{n+1} + \gamma B y_n, x - x_{n+1}\rangle. \end{aligned} \tag{3.5}$$

Also, from $y_{n-1} = 2x_{n-1} - x_{n-2}$ and $w_{n-1} = x_{n-1} + \theta(x_{n-1} - x_{n-2}) + \delta(x_{n-2} - x_{n-3})$, we obtain

$$\begin{aligned} x_n - w_{n-1} &= x_n - x_{n-1} - \theta(x_{n-1} - x_{n-2}) \\ &\quad - \delta(x_{n-2} - x_{n-3}) \\ &= x_n - x_{n-1} - \theta(y_{n-1} - x_{n-1}) \\ &\quad - \delta(x_{n-2} - x_{n-3}) \\ &= (1-\theta)(x_n - x_{n-1}) + \theta(x_n - y_{n-1}) \\ &\quad - \delta(x_{n-2} - x_{n-3}) \\ &= (1-\theta)(y_n - x_n) + \theta(x_n - y_{n-1}) \\ &\quad - \delta(x_{n-2} - x_{n-3}). \end{aligned} \tag{3.6}$$

Using (3.6) in (3.5), we get

$$\begin{aligned} 0 &= 2(1-\theta)\langle y_n - x_n, x_{n+1} - y_n\rangle + 2\theta\langle x_n - y_{n-1}, x_{n+1} - y_n\rangle + 2\langle p_n + \gamma B y_n, x_{n+1} - y_n\rangle \\ &\quad + 2\gamma\langle B y_{n-1} - B y_n, x_{n+1} - y_n\rangle + 2\langle x_{n+1} - w_n, x - x_{n+1}\rangle + 2\langle p_{n+1} + \gamma B y_n, x - x_{n+1}\rangle \\ &\quad - 2\delta\langle x_{n-2} - x_{n-3}, x_{n+1} - y_n\rangle \\ &= 2(1-\theta)\langle y_n - x_n, x_{n+1} - y_n\rangle + 2\langle x_{n+1} - w_n, x - x_{n+1}\rangle + 2\gamma\langle B y_{n-1} - B y_n, x_{n+1} - y_n\rangle \\ &\quad + 2\langle p_n, x_{n+1} - y_n\rangle + 2\gamma\langle B y_n, x - y_n\rangle + 2\langle p_{n+1}, x - x_{n+1}\rangle + 2\theta\langle x_n - y_{n-1}, x_{n+1} - y_n\rangle \\ &\quad - 2\delta\langle x_{n-2} - x_{n-3}, x_{n+1} - y_n\rangle. \end{aligned} \tag{3.7}$$

Now, by (2.1), we have

$$2\langle y_n - x_n, x_{n+1} - y_n\rangle = \|x_{n+1} - x_n\|^2 - \|x_{n+1} - y_n\|^2 - \|x_n - y_n\|^2 \tag{3.8}$$

and

$$2\langle x_{n+1} - w_n, x - x_{n+1}\rangle = \|w_n - x\|^2 - \|x_{n+1} - x\|^2 - \|x_{n+1} - w_n\|^2. \tag{3.9}$$

Using the Cauchy–Schwarz inequality, we obtain

$$-2\theta\langle x_{n+1} - x_n, x_n - x_{n-1}\rangle \geq -2\theta\|x_{n+1} - x_n\|\|x_n - x_{n-1}\|, \tag{3.10}$$

$$-2\delta\langle x_{n+1} - x_n, x_{n-1} - x_{n-2}\rangle \geq -2|\delta|\|x_{n+1} - x_n\|\|x_{n-1} - x_{n-2}\|, \tag{3.11}$$

and

$$\begin{aligned} 2\delta\theta\langle x_n - x_{n-1}, x_{n-1} - x_{n-2}\rangle &= -2\delta\theta\langle x_{n-1} - x_n, x_{n-1} - x_{n-2}\rangle \\ &\geq -2|\delta|\theta\|x_n - x_{n-1}\|\|x_{n-1} - x_{n-2}\|. \end{aligned} \tag{3.12}$$

Using (3.10), (3.11) and (3.12), we obtain

$$\begin{aligned} \|x_{n+1} - w_n\|^2 &= \|x_{n+1} - (x_n + \theta(x_n - x_{n-1}) + \delta(x_{n-1} - x_{n-2}))\|^2 \\ &= \|x_{n+1} - x_n - \theta(x_n - x_{n-1}) - \delta(x_{n-1} - x_{n-2})\|^2 \\ &= \|x_{n+1} - x_n\|^2 - 2\theta\langle x_{n+1} - x_n, x_n - x_{n-1}\rangle \\ &\quad - 2\delta\langle x_{n+1} - x_n, x_{n-1} - x_{n-2}\rangle + \theta^2\|x_n - x_{n-1}\|^2 \\ &\quad + 2\delta\theta\langle x_n - x_{n-1}, x_{n-1} - x_{n-2}\rangle + \delta^2\|x_{n-1} - x_{n-2}\|^2 \\ &\geq \|x_{n+1} - x_n\|^2 - 2\theta\|x_{n+1} - x_n\|\|x_n - x_{n-1}\| \\ &\quad - 2|\delta|\|x_{n+1} - x_n\|\|x_{n-1} - x_{n-2}\| + \theta^2\|x_n - x_{n-1}\|^2 \\ &\quad - 2|\delta|\theta\|x_n - x_{n-1}\|\|x_{n-1} - x_{n-2}\| + \delta^2\|x_{n-1} - x_{n-2}\|^2 \\ &\geq \|x_{n+1} - x_n\|^2 - \theta\|x_{n+1} - x_n\|^2 - \theta\|x_n - x_{n-1}\|^2 \\ &\quad - |\delta|\|x_{n+1} - x_n\|^2 - |\delta|\|x_{n-1} - x_{n-2}\|^2 + \theta^2\|x_n - x_{n-1}\|^2 \\ &\quad - |\delta|\theta\|x_n - x_{n-1}\|^2 - |\delta|\theta\|x_{n-1} - x_{n-2}\|^2 + \delta^2\|x_{n-1} - x_{n-2}\|^2 \\ &= (1 - |\delta| - \theta)\|x_{n+1} - x_n\|^2 + (\theta^2 - \theta - |\delta|\theta)\|x_n - x_{n-1}\|^2 \\ &\quad + (\delta^2 - |\delta| - |\delta|\theta)\|x_{n-1} - x_{n-2}\|^2. \end{aligned} \tag{3.13}$$

Employing Lemma 2.1, we find that

$$\begin{aligned} \|w_n - x\|^2 &= \|(1+\theta)(x_n - x) - (\theta - \delta)(x_{n-1} - x) - \delta(x_{n-2} - x)\|^2 \\ &= (1+\theta)\|x_n - x\|^2 - (\theta - \delta)\|x_{n-1} - x\|^2 - \delta\|x_{n-2} - x\|^2 \\ &\quad + (1+\theta)(\theta - \delta)\|x_n - x_{n-1}\|^2 + \delta(1+\theta)\|x_n - x_{n-2}\|^2 \\ &\quad - \delta(\theta - \delta)\|x_{n-1} - x_{n-2}\|^2. \end{aligned} \tag{3.14}$$

Combining (3.13) and (3.14) with (3.9), we see that

$$\begin{aligned} 2\langle x_{n+1} - w_n, x - x_{n+1}\rangle &\leq (1+\theta)\|x_n - x\|^2 - (\theta - \delta)\|x_{n-1} - x\|^2 - \delta\|x_{n-2} - x\|^2 \\ &\quad + (1+\theta)(\theta - \delta)\|x_n - x_{n-1}\|^2 + \delta(1+\theta)\|x_n - x_{n-2}\|^2 \\ &\quad - \delta(\theta - \delta)\|x_{n-1} - x_{n-2}\|^2 - \|x_{n+1} - x\|^2 - (1 - |\delta| - \theta)\|x_{n+1} - x_n\|^2 \\ &\quad - (\theta^2 - \theta - |\delta|\theta)\|x_n - x_{n-1}\|^2 - (\delta^2 - |\delta| - |\delta|\theta)\|x_{n-1} - x_{n-2}\|^2 \\ &= (1+\theta)\|x_n - x\|^2 - (\theta - \delta)\|x_{n-1} - x\|^2 - \delta\|x_{n-2} - x\|^2 \\ &\quad + \Big((1+\theta)(\theta - \delta) - (\theta^2 - \theta - |\delta|\theta)\Big)\|x_n - x_{n-1}\|^2 \\ &\quad + \delta(1+\theta)\|x_n - x_{n-2}\|^2 - (1 - |\delta| - \theta)\|x_{n+1} - x_n\|^2 \\ &\quad - \Big(\delta(\theta - \delta) + (\delta^2 - |\delta| - |\delta|\theta)\Big)\|x_{n-1} - x_{n-2}\|^2 \\ &\quad - \|x_{n+1} - x\|^2 \\ &= (1+\theta)\|x_n - x\|^2 - (\theta - \delta)\|x_{n-1} - x\|^2 - \delta\|x_{n-2} - x\|^2 \\ &\quad + (2\theta - \delta - \delta\theta + |\delta|\theta)\|x_n - x_{n-1}\|^2 + \delta(1+\theta)\|x_n - x_{n-2}\|^2 \\ &\quad - (1 - |\delta| - \theta)\|x_{n+1} - x_n\|^2 + (|\delta| + |\delta|\theta - \delta\theta)\|x_{n-1} - x_{n-2}\|^2 \\ &\quad - \|x_{n+1} - x\|^2 \\ &\leq (1+\theta)\|x_n - x\|^2 - (\theta - \delta)\|x_{n-1} - x\|^2 - \delta\|x_{n-2} - x\|^2 \\ &\quad + (2\theta - \delta - \delta\theta + |\delta|\theta)\|x_n - x_{n-1}\|^2 - (1 - |\delta| - \theta)\|x_{n+1} - x_n\|^2 \\ &\quad + (|\delta| + |\delta|\theta - \delta\theta)\|x_{n-1} - x_{n-2}\|^2 - \|x_{n+1} - x\|^2. \end{aligned} \tag{3.15}$$

Next, using the L-Lipschitz continuity of B, we find that

$$
\begin{aligned}
2\gamma\langle By_{n-1}-By_n, x_{n+1}-y_n\rangle &\le 2\gamma L\|y_n-y_{n-1}\|\|x_{n+1}-y_n\| \\
&\le \gamma L\left(\frac{1}{\sqrt{2}}\|y_n-y_{n-1}\|^2+\sqrt{2}\|x_{n+1}-y_n\|^2\right) \\
&\le \gamma L\frac{1}{\sqrt{2}}((2+\sqrt{2})\|y_n-x_n\|^2+\sqrt{2}\|x_n-y_{n-1}\|^2)+\sqrt{2}\gamma L\|x_{n+1}-y_n\|^2 \\
&= \gamma L(1+\sqrt{2})\|y_n-x_n\|^2+\gamma L\|x_n-y_{n-1}\|^2+\sqrt{2}\gamma L\|x_{n+1}-y_n\|^2.
\end{aligned}
\tag{3.16}
$$

Now, since B is monotone, we have

$$
\gamma\langle By_n, x-y_n\rangle \le \gamma\langle Bx, x-y_n\rangle. \tag{3.17}
$$

Also, since $p_{n+1}\in\gamma Ax_{n+1}$, $-\gamma Bx\in\gamma Ax$ and A is monotone, we find that $0\le\langle -\gamma Bx-p_{n+1}, x-x_{n+1}\rangle$.

Hence,

$$
\begin{aligned}
\langle p_{n+1}, x-x_{n+1}\rangle &\le \langle p_{n+1}, x-x_{n+1}\rangle+\langle -\gamma Bx-p_{n+1}, x-x_{n+1}\rangle \\
&= \gamma\langle Bx, x_{n+1}-x\rangle.
\end{aligned}
\tag{3.18}
$$

Again, since $p_n\in\gamma Ax_n$, $p_{n+1}\in\gamma Ax_{n+1}$ and A is monotone, we see that

$$
\langle p_n, x_{n+1}-x_n\rangle \le \langle p_{n+1}, x_{n+1}-x_n\rangle. \tag{3.19}
$$

From (3.17) and (3.18), it follows that

$$
\begin{aligned}
&2\langle p_n, x_{n+1}-y_n\rangle+2\gamma\langle By_n, x-y_n\rangle+2\langle p_{n+1}, x-x_{n+1}\rangle \\
\le\ & 2\langle p_n, x_{n+1}-y_n\rangle+2\gamma\langle Bx, x-y_n\rangle+2\gamma\langle Bx, x_{n+1}-x\rangle \\
=\ & 2\langle p_n+\gamma Bx, x_{n+1}-y_n\rangle \\
=\ & 2\langle p_n+\gamma Bx, x_{n+1}-x_n\rangle-2\langle p_n+\gamma Bx, x_n-x_{n-1}\rangle \\
\le\ & 2\langle p_{n+1}+\gamma Bx, x_{n+1}-x_n\rangle-2\langle p_n+\gamma Bx, x_n-x_{n-1}\rangle,
\end{aligned}
\tag{3.20}
$$

where the last equality and inequality follow from $y_n=2x_n-x_{n-1}$ and (3.19), respectively. Now, putting (3.8), (3.15), (3.16) and (3.20) in (3.7), we obtain

$$
\begin{aligned}
0 \le\ & (1-\theta)\|x_{n+1}-x_n\|^2-(1-\theta)\|x_{n+1}-y_n\|^2-(1-\theta)\|x_n-y_n\|^2 \\
&+(1+\theta)\|x_n-x\|^2-(\theta-\delta)\|x_{n-1}-x\|^2-\delta\|x_{n-2}-x\|^2 \\
&+(2\theta-\delta-\delta\theta+|\delta|\theta)\|x_n-x_{n-1}\|^2-(1-|\delta|-\theta)\|x_{n+1}-x_n\|^2 \\
&+(|\delta|+|\delta|\theta-\delta\theta)\|x_{n-1}-x_{n-2}\|^2-\|x_{n+1}-x\|^2 \\
&+\gamma L(1+\sqrt{2})\|y_n-x_n\|^2+\gamma L\|x_n-y_{n-1}\|^2+\sqrt{2}\gamma L\|x_{n+1}-y_n\|^2 \\
&+2\langle p_{n+1}+\gamma Bx, x_{n+1}-x_n\rangle-2\langle p_n+\gamma Bx, x_n-x_{n-1}\rangle+2\theta\langle x_n-y_{n-1}, x_{n+1}-y_n\rangle \\
&-2\delta\langle x_{n-2}-x_{n-3}, x_{n+1}-y_n\rangle \\
\le\ & (1-\theta)\|x_{n+1}-x_n\|^2-(1-\theta)\|x_{n+1}-y_n\|^2-(1-\theta)\|x_n-y_n\|^2 \\
&+(1+\theta)\|x_n-x\|^2-(\theta-\delta)\|x_{n-1}-x\|^2-\delta\|x_{n-2}-x\|^2 \\
&+(2\theta-\delta-\delta\theta+|\delta|\theta)\|x_n-x_{n-1}\|^2-(1-|\delta|-\theta)\|x_{n+1}-x_n\|^2 \\
&+(|\delta|+|\delta|\theta-\delta\theta)\|x_{n-1}-x_{n-2}\|^2-\|x_{n+1}-x\|^2 \\
&+\gamma L(1+\sqrt{2})\|y_n-x_n\|^2+\gamma L\|x_n-y_{n-1}\|^2+\sqrt{2}\gamma L\|x_{n+1}-y_n\|^2 \\
&-\|p_{n+1}+\gamma Bx\|^2-\|x_{n+1}-x_n\|^2+\|p_{n+1}+\gamma Bx+x_{n+1}-x_n\|^2 \\
&+\|p_n+\gamma Bx\|^2+\|x_n-x_{n-1}\|^2-\|p_n+\gamma Bx+x_n-x_{n-1}\|^2 \\
&+\theta\|x_n-y_{n-1}\|^2+\theta\|x_{n+1}-y_n\|^2-2\delta\langle x_{n-2}-x_{n-3}, x_{n+1}-y_n\rangle.
\end{aligned}
\tag{3.21}
$$

Rearranging (3.21), we arrive at

$$
\begin{aligned}
&\|x_{n+1}-x\|^2-\theta\|x_n-x\|^2-\delta\|x_{n-1}-x\|^2+(1-|\delta|+2\theta)\|x_{n+1}-x_n\|^2 \\
&+(\gamma L+\theta)\|x_{n+1}-y_n\|^2+\|p_{n+1}+\gamma Bx\|^2-\|p_{n+1}+x_{n+1}-x_n+\gamma Bx\|^2 \\
\le\ & \|x_n-x\|^2-\theta\|x_{n-1}-x\|^2-\delta\|x_{n-2}-x\|^2+(1+2\theta-\delta-\delta\theta+|\delta|\theta)\|x_n-x_{n-1}\|^2 \\
&+(\gamma L+\theta)\|x_n-y_{n-1}\|^2+\|p_n+\gamma Bx\|^2-\|p_n+x_n-x_{n-1}+\gamma Bx\|^2 \\
&+2\theta\|x_{n+1}-x_n\|^2-(1-\theta-\gamma L(1+\sqrt{2}))\|x_n-y_n\|^2 \\
&-(1-3\theta-\gamma L(1+\sqrt{2}))\|x_{n+1}-y_n\|^2+(|\delta|+|\delta|\theta-\delta\theta)\|x_{n-1}-x_{n-2}\|^2 \\
&-2\delta\langle x_{n-2}-x_{n-3}, x_{n+1}-y_n\rangle
\end{aligned}
$$

$$\begin{aligned}
&\le \|x_n - x\|^2 - \theta\|x_{n-1} - x\|^2 - \delta\|x_{n-2} - x\|^2 + (1 + 2\theta - \delta - \delta\theta + |\delta|\theta)\|x_n - x_{n-1}\|^2 \\
&\quad + (\gamma L + \theta)\|x_n - y_{n-1}\|^2 + \|p_n + \gamma Bx\|^2 - \|p_n + x_n - x_{n-1} + \gamma Bx\|^2 \\
&\quad + 2\theta\|x_{n+1} - x_n\|^2 + (|\delta| + |\delta|\theta - \delta\theta)\|x_{n-1} - x_{n-2}\|^2 \\
&\quad - (1 - 3\theta - \gamma L(1 + \sqrt{2}))\left[\|x_n - y_n\|^2 + \|x_{n+1} - y_n\|^2\right] \\
&\quad - 2\delta\langle x_{n-2} - x_{n-3}, x_{n+1} - y_n\rangle \\
&= \|x_n - x\|^2 - \theta\|x_{n-1} - x\|^2 - \delta\|x_{n-2} - x\|^2 + (1 - |\delta| + 2\theta)\|x_n - x_{n-1}\|^2 \\
&\quad + (\gamma L + \theta)\|x_n - y_{n-1}\|^2 + \|p_n + \gamma Bx\|^2 - \|p_n + x_n - x_{n-1} + \gamma Bx\|^2 \\
&\quad + 2\theta\|x_{n+1} - x_n\|^2 + (|\delta| + |\delta|\theta - \delta\theta)\|x_{n-1} - x_{n-2}\|^2 \\
&\quad - (1 - 3\theta - \gamma L(1 + \sqrt{2}))\left[\|x_n - y_n\|^2 + \|x_{n+1} - y_n\|^2\right] \\
&\quad - 2\delta(1 + \theta)\|x_n - x_{n-1}\|^2 - 2\delta\langle x_{n-2} - x_{n-3}, x_{n+1} - y_n\rangle,
\end{aligned}$$

where the last equality follows from $-\delta = |\delta|$ (since $\delta \le 0$). Hence, the required conclusion. □

Next, we present the following conditions on the inertial parameters and the step size. They are employed in our main convergence result.

Assumption 3.2.

(i) Choose $\delta \in \left(-\frac{1}{9}, -\frac{5}{54}\right] \cup \{0\}$ and $\theta \in \left[0, \frac{9\delta+1}{7-8\delta}\right)$;

(ii) Pick the step size γ such that $0 < \gamma < \min\left\{\frac{1+\sqrt{1+16\theta(1-4\theta+36\delta\theta)}}{8L}, \frac{8\delta(1+\theta)+1+\delta-7\theta}{L(1+\sqrt{2})}\right\}$.

Theorem 3.1. *Let $\{x_n\}$ be generated by Algorithm 1 when Assumptions 3.1 and 3.2 hold. Then $\{x_n\}$ converges weakly to an element of $(A+B)^{-1}(\mathbf{0})$.*

Proof. Let $x \in (A+B)^{-1}(\mathbf{0})$. It follows from Lemma 3.1 that (note that $|\delta| = -\delta$)

$$\begin{aligned}
a_{n+1} &\le a_n + 2\theta\|x_{n+1} - x_n\|^2 - \delta(1 + 2\theta)\|x_{n-1} - x_{n-2}\|^2 \\
&\quad - (1 - 3\theta - \gamma L(1 + \sqrt{2}))\Big(\|x_{n+1} - y_n\|^2 + \|y_n - x_n\|^2\Big) \\
&\quad - 2\delta(1 + \theta)\|x_n - x_{n-1}\|^2 - 2\delta\langle x_{n-2} - x_{n-3}, x_{n+1} - y_n\rangle \\
&\le a_n + 2\theta\|x_{n+1} - x_n\|^2 - \delta(1 + 2\theta)\|x_{n-1} - x_{n-2}\|^2 \\
&\quad - (1 - 3\theta - \gamma L(1 + \sqrt{2}))\Big(\|x_{n+1} - y_n\|^2 + \|y_n - x_n\|^2\Big) \\
&\quad - 2\delta(1 + \theta)\|x_n - x_{n-1}\|^2 - \delta\|x_{n-2} - x_{n-3}\|^2 - \delta\|x_{n+1} - y_n\|^2 \\
&= a_n + 2\theta\|x_{n+1} - x_n\|^2 - \delta(1 + 2\theta)\|x_{n-1} - x_{n-2}\|^2 \\
&\quad - (1 - 3\theta + \delta - \gamma L(1 + \sqrt{2}))\|x_{n+1} - y_n\|^2 \\
&\quad - (1 - 3\theta - \gamma L(1 + \sqrt{2}))\|y_n - x_n\|^2 \\
&\quad - 2\delta(1 + \theta)\|x_n - x_{n-1}\|^2 - \delta\|x_{n-2} - x_{n-3}\|^2 \\
&\le a_n + 2\theta\|x_{n+1} - x_n\|^2 - \delta(1 + 2\theta)\|x_{n-1} - x_{n-2}\|^2 \\
&\quad - (1 - 3\theta + \delta - \gamma L(1 + \sqrt{2}))\Big(\|x_{n+1} - y_n\|^2 + \|y_n - x_n\|^2\Big) \\
&\quad - 2\delta(1 + \theta)\|x_n - x_{n-1}\|^2 - \delta\|x_{n-2} - x_{n-3}\|^2.
\end{aligned} \tag{3.22}$$

Since $\|x_{n+1} - x_n\|^2 \le 2\left(\|x_{n+1} - y_n\|^2 + \|y_n - x_n\|^2\right)$ and $1 - 3\theta + \delta - \gamma L(1 + \sqrt{2}) > 0$ (by Assumption 3.2 (ii)), we get from (3.22) that

$$\begin{aligned}
a_{n+1} &\le a_n + 2\theta\|x_{n+1} - x_n\|^2 - \frac{1}{2}(1 - 3\theta + \delta - \gamma L(1 + \sqrt{2}))\|x_{n+1} - x_n\|^2 \\
&\quad - \delta(1 + 2\theta)\|x_{n-1} - x_{n-2}\|^2 - 2\delta(1 + \theta)\|x_n - x_{n-1}\|^2 \\
&\quad - \delta\|x_{n-2} - x_{n-3}\|^2 \\
&= a_n - \frac{1}{2}(1 - 7\theta + \delta - \gamma L(1 + \sqrt{2}))\|x_{n+1} - x_n\|^2 \\
&\quad - \delta(1 + 2\theta)\|x_{n-1} - x_{n-2}\|^2 - 2\delta(1 + \theta)\|x_n - x_{n-1}\|^2 \\
&\quad - \delta\|x_{n-2} - x_{n-3}\|^2.
\end{aligned} \tag{3.23}$$

Next, we show that $a_n \ge 0$ for all $n \ge 1$. To this end, observe that

$$\begin{aligned}
\|p_n + x_n - x_{n-1} + \gamma Bx\|^2 &= \|w_{n-1} - \gamma By_{n-1} - x_{n-1} + \gamma Bx\|^2 \\
&= \|\theta(x_{n-1} - x_{n-2}) + \delta(x_{n-2} - x_{n-3}) - \gamma(By_{n-1} - Bx)\|^2
\end{aligned}$$

$$
\begin{aligned}
&\le 2\|\theta(x_{n-1}-x_{n-2})+\delta(x_{n-2}-x_{n-3})\|^2+2\gamma^2\|By_{n-1}-Bx\|^2\\
&\le 2\|\theta(x_{n-1}-x_{n-2})+\delta(x_{n-2}-x_{n-3})\|^2+2\gamma^2L^2\|y_{n-1}-x\|^2\\
&= 2\|\theta(y_{n-1}-x_{n-1})+\delta(y_{n-2}-x_{n-2})\|^2+2\gamma^2L^2\|y_{n-1}-x\|^2\\
&= 2\|\theta(y_{n-1}-x_n)+\theta(x_n-x_{n-1})+\delta(y_{n-2}-x_{n-2})\|^2\\
&\quad +2\gamma^2L^2\|y_{n-1}-x\|^2\\
&\le 2\|\theta(y_{n-1}-x_n)+\theta(x_n-x_{n-1})+\delta(y_{n-2}-x_{n-2})\|^2\\
&\quad +2\gamma^2L^2\Big[2\|y_{n-1}-x_n\|^2+2\|x_n-x\|^2\Big]\\
&= 2\|\theta(y_{n-1}-x_n)+\theta(x_n-x_{n-1})+\delta(y_{n-2}-x_{n-2})\|^2\\
&\quad +4\gamma^2L^2\|y_{n-1}-x_n\|^2+4\gamma^2L^2\|x_n-x\|^2\\
&\le 2\Big[\|\theta(y_{n-1}-x_n)+\theta(x_n-x_{n-1})\|^2\\
&\quad +2\langle\delta(y_{n-2}-x_{n-2}),\theta(y_{n-1}-x_n)+\theta(x_n-x_{n-1})+\delta(y_{n-2}-x_{n-2})\rangle\Big]\\
&\quad +4\gamma^2L^2\|y_{n-1}-x_n\|^2+4\gamma^2L^2\|x_n-x\|^2.
\end{aligned}
\tag{3.24}
$$

We obtain by the Cauchy–Schwarz inequality that

$$
\begin{aligned}
&2\langle\delta(y_{n-2}-x_{n-2}),\theta(y_{n-1}-x_n)+\theta(x_n-x_{n-1})+\delta(y_{n-2}-x_{n-2})\rangle\\
&= 2\delta\langle y_{n-2}-x_{n-2},\theta(y_{n-1}-x_n)+\theta(x_n-x_{n-1})+\delta(y_{n-2}-x_{n-2})\rangle\\
&\le |\delta|\Big[\frac{1}{18}\|y_{n-2}-x_{n-2}\|^2+18\theta^2\|y_{n-1}-x_n\|^2+\frac{1}{27}\|y_{n-2}-x_{n-2}\|^2\\
&\quad +27\theta^2\|x_n-x_{n-1}\|^2+\delta\|y_{n-2}-x_{n-2}\|^2\Big]\\
&= |\delta|\Big[\Big(\frac{1}{18}+\delta\Big)\|y_{n-2}-x_{n-2}\|^2+18\theta^2\|y_{n-1}-x_n\|^2\\
&\quad +\frac{1}{27}\|y_{n-2}-x_{n-2}\|^2+27\theta^2\|x_n-x_{n-1}\|^2\Big].
\end{aligned}
\tag{3.25}
$$

Plugging (3.25) into (3.24), we obtain

$$
\begin{aligned}
&\|p_n+x_n-x_{n-1}+\gamma Bx\|^2\le 2\|\theta(y_{n-1}-x_n)+\theta(x_n-x_{n-1})\|^2\\
&\quad +2|\delta|\Big[\Big(\frac{1}{18}+\delta\Big)\|y_{n-2}-x_{n-2}\|^2+18\theta^2\|y_{n-1}-x_n\|^2\\
&\quad +\frac{1}{27}\|y_{n-2}-x_{n-2}\|^2+27\theta^2\|x_n-x_{n-1}\|^2\Big]+4\gamma^2L^2\|y_{n-1}-x_n\|^2\\
&\quad +4\gamma^2L^2\|x_n-x\|^2\\
&\le 2\Big[2\theta^2\|y_{n-1}-x_n\|^2+2\theta^2\|x_n-x_{n-1}\|^2\Big]\\
&\quad +2|\delta|\Big[\Big(\frac{1}{18}+\frac{1}{27}+\delta\Big)\|y_{n-2}-x_{n-2}\|^2+18\theta^2\|y_{n-1}-x_n\|^2\\
&\quad +27\theta^2\|x_n-x_{n-1}\|^2\Big]+4\gamma^2L^2\|y_{n-1}-x_n\|^2+4\gamma^2L^2\|x_n-x\|^2.
\end{aligned}
\tag{3.26}
$$

We also have

$$
\|x_{n-1}-x\|^2\le 2\|x_{n-1}-x_n\|^2+2\|x_n-x\|^2. \tag{3.27}
$$

Now, combining (3.26) and (3.27) with (3.1), we obtain

$$
\begin{aligned}
a_n &\ge \|x_n-x\|^2-2\theta\|x_{n-1}-x_n\|^2-2\theta\|x_n-x\|^2+(1-|\delta|+2\theta)\|x_n-x_{n-1}\|^2+(\gamma L+\theta)\|x_n-y_{n-1}\|^2\\
&\quad +\|p_n+\gamma Bx\|^2-\Big(4\theta^2+36|\delta|\theta^2+4\gamma^2L^2\Big)\|y_{n-1}-x_n\|^2\\
&\quad -\Big(4\theta^2+54|\delta|\theta^2\Big)\|x_n-x_{n-1}\|^2-4\gamma^2L^2\|x_n-x\|^2-2|\delta|\Big(\frac{5}{54}+\delta\Big)\|y_{n-2}-x_{n-2}\|^2\\
&= \Big(1-2\theta-4\gamma^2L^2\Big)\|x_n-x\|^2+\Big(1-|\delta|-4\theta^2-54|\delta|\theta^2\Big)\|x_n-x_{n-1}\|^2\\
&\quad +\Big(\gamma L+\theta-4\theta^2-36|\delta|\theta^2-4\gamma^2L^2\Big)\|x_n-y_{n-1}\|^2+\|p_n+\gamma Bx\|^2\\
&\quad -2|\delta|\Big(\frac{5}{54}+\delta\Big)\|y_{n-2}-x_{n-2}\|^2.
\end{aligned}
\tag{3.28}
$$

Thus, using the conditions imposed on θ and γ (i.e., Assumption 3.2), we see that $a_n\ge 0$ for all $n\ge 1$, as asserted.

Now, from (3.23), we see that

$$
\begin{aligned}
a_{n+1}-\delta\|x_{n-1}-x_{n-2}\|^2 &\le a_n-\delta\|x_{n-2}-x_{n-3}\|^2\\
&\quad -2\delta(1+\theta)\|x_{n-1}-x_{n-2}\|^2-2\delta(1+\theta)\|x_n-x_{n-1}\|^2\\
&\quad -\frac{1}{2}(1-7\theta+\delta-\gamma L(1+\sqrt{2}))\|x_{n+1}-x_n\|^2.
\end{aligned}
$$

Hence,

$$\begin{aligned} a_{n+1} - \delta\|x_{n-1} - x_{n-2}\|^2 - 2\delta(1+\theta)\|x_n - x_{n-1}\|^2 \le\ & a_n - \delta\|x_{n-2} - x_{n-3}\|^2 \\ & - 2\delta(1+\theta)\|x_{n-1} - x_{n-2}\|^2 \\ & - \frac{1}{2}(1 - 7\theta + \delta - \gamma L(1+\sqrt{2}))\|x_{n+1} - x_n\|^2 \\ & - 4\delta(1+\theta)\|x_n - x_{n-1}\|^2. \end{aligned}$$

Consequently,

$$b_{n+1} \le b_n - \frac{1}{2}\Big(8\delta(1+\theta) + 1 - 7\theta + \delta - \gamma L(1+\sqrt{2})\Big)\|x_n - x_{n-1}\|^2,$$

where

$$b_n := a_n - \delta\|x_{n-2} - x_{n-3}\|^2 - 2\delta(1+\theta)\|x_{n-1} - x_{n-2}\|^2 + \frac{1}{2}(1 - 7\theta + \delta - \gamma L(1+\sqrt{2}))\|x_n - x_{n-1}\|^2.$$

Since $a_n \ge 0$, we also have $b_n \ge 0$. In addition, the conditions imposed on θ and γ imply that $\frac{1}{2}\Big(8\delta(1+\theta)+1-7\theta+\delta-\gamma L(1+\sqrt{2})\Big) > 0$. Therefore, we obtain that $\lim_{n\to\infty} b_n$ exists and

$$\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0. \tag{3.29}$$

Since $\lim_{n\to\infty} b_n$ exists, it follows from (3.29) that $\lim_{n\to\infty} a_n$ exists and thus, $\{a_n\}$ is bounded. Using (3.28), we obtain that $\{x_n\}$ is also bounded. Hence, $\{y_n\}$, $\{w_n\}$ and $\{By_n - Bx\}$ are bounded too. Therefore, by (3.29), we obtain

$$\lim_{n\to\infty} \langle x_{n-1} - x_n, By_{n-1} - Bx\rangle = 0. \tag{3.30}$$

Now, since $y_n = 2x_n - x_{n-1}$ and $w_n = x_n + \theta(x_n - x_{n-1}) + \delta(x_{n-1} - x_{n-2})$, it follows from (3.29) that

$$\lim_{n\to\infty} \|y_n - x_n\| = \lim_{n\to\infty} \|x_n - x_{n-1}\| = 0 = \lim_{n\to\infty} \|w_n - x_n\|. \tag{3.31}$$

We also have

$$\|x_{n+1} - y_n\| \le \|x_{n+1} - x_n\| + \|y_n - x_n\| \to 0,\ n \to \infty \tag{3.32}$$

and

$$\|x_{n+1} - w_n\| \le \|x_{n+1} - x_n\| + \|x_n - w_n\| \to 0,\ n \to \infty. \tag{3.33}$$

Combining (3.2) and (2.1), we find that

$$\begin{aligned} & \|p_n + \gamma Bx\|^2 - \|p_n + x_n - x_{n-1} + \gamma Bx\|^2 \\ =\ & \|w_{n-1} - x_n + \gamma(Bx - By_{n-1})\|^2 - \|w_{n-1} - x_{n-1} + \gamma(Bx - By_{n-1})\|^2 \\ =\ & \|w_{n-1} - x_n\|^2 - \|w_{n-1} - x_{n-1}\|^2 + 2\gamma\langle x_{n-1} - x_n, By_{n-1} - Bx\rangle \\ & \to 0,\ n \to \infty \text{ (by (3.33), (3.31) and (3.30)).} \end{aligned} \tag{3.34}$$

Now, since $\lim_{n\to\infty} a_n$ exists, it follows from (3.1), (3.29), (3.32) and (3.34) that

$$\lim_{n\to\infty} \left[\|x_n - x\|^2 - \theta\|x_{n-1} - x\|^2 - \delta\|x_{n-2} - x\|^2\right] \quad \text{exists.} \tag{3.35}$$

Since the sequence $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_k}\} \subset \{x_n\}$ such that $x_{n_k} \rightharpoonup p \in \mathcal{H}$.

Once again, using (3.33) and (3.32), we see that

$$w_{n_k} - x_{n_k+1} - \gamma By_{n_k} + \gamma Bx_{n_k+1} \to 0,\ n \to \infty. \tag{3.36}$$

Also,

$$w_{n_k} - x_{n_k+1} - \gamma By_{n_k} + \gamma Bx_{n_k+1} = p_{n_k+1} + \gamma Bx_{n_k+1} \in \gamma(A+B)x_{n_k+1}. \tag{3.37}$$

Since $A+B$ is maximal monotone, its graph is closed in $H^{\text{strong}} \times H^{\text{weak}}$. Thus, by (3.36) and (3.37), we find that $p \in (A+B)^{-1}(\mathbf{0})$.

We now show that $x_n \rightharpoonup p \in (A+B)^{-1}(\mathbf{0})$. To this end, assume that there exist two subsequences $\{x_{n_k}\} \subset \{x_n\}$ and $\{x_{m_k}\} \subset \{x_n\}$ such that $x_{n_k} \rightharpoonup p$ as $k \to \infty$ and $x_{m_k} \rightharpoonup p^*$ as $k \to \infty$. We claim that $p = p^*$.

To see this, observe first that

$$2\langle x_n, p^* - p\rangle = \|x_n - p\|^2 - \|x_n - p^*\|^2 - \|p\|^2 + \|p^*\|^2, \tag{3.38}$$

$$2\langle x_{n-1}, p^* - p\rangle = \|x_{n-1} - p\|^2 - \|x_{n-1} - p^*\|^2 - \|p\|^2 + \|p^*\|^2, \tag{3.39}$$

and

$$2\langle x_{n-2}, p^* - p\rangle = \|x_{n-2} - p\|^2 - \|x_{n-2} - p^*\|^2 - \|p\|^2 + \|p^*\|^2. \tag{3.40}$$

Hence, (3.39) and (3.40) imply

$$\begin{aligned}2\langle -\theta x_{n-1}, p^* - p\rangle &= -\theta\|x_{n-1} - p\|^2 + \theta\|x_{n-1} - p^*\|^2 \\ &\quad + \theta\|p\|^2 - \theta\|p^*\|^2\end{aligned} \tag{3.41}$$

and

$$\begin{aligned}2\langle -\delta x_{n-2}, p^* - p\rangle &= -\delta\|x_{n-2} - p\|^2 + \delta\|x_{n-2} - p^*\|^2 \\ &\quad + \delta\|p\|^2 - \delta\|p^*\|^2,\end{aligned} \tag{3.42}$$

respectively. Adding (3.38), (3.41) and (3.42), we arrive at

$$\begin{aligned}2\langle x_n - \theta x_{n-1} - \delta x_{n-2}, p^* - p\rangle &= \left(\|x_n - p\|^2 - \theta\|x_{n-1} - p\|^2 - \delta\|x_{n-2} - p\|^2\right) \\ &\quad - \left(\|x_n - p^*\|^2 - \theta\|x_{n-1} - p^*\|^2 - \delta\|x_{n-2} - p^*\|^2\right) \\ &\quad + (1 - \theta - \delta)(\|p^*\|^2 - \|p\|^2).\end{aligned}$$

In view of (3.35), both

$$\lim_{n\to\infty}\left[\|x_n - p^*\|^2 - \theta\|x_{n-1} - p^*\|^2 - \delta\|x_{n-2} - p^*\|^2\right]$$

and

$$\lim_{n\to\infty}\left[\|x_n - p\|^2 - \theta\|x_{n-1} - p\|^2 - \delta\|x_{n-2} - p\|^2\right]$$

exist. Therefore,

$$\lim_{n\to\infty}\langle x_n - \theta x_{n-1} - \delta x_{n-2}, p^* - p\rangle$$

exists. Now,

$$\begin{aligned}\langle p - \theta p - \delta p, p^* - p\rangle &= \lim_{k\to\infty}\langle x_{n_k} - \theta x_{n_k-1} - \delta x_{n_k-2}, p^* - p\rangle \\ &= \lim_{n\to\infty}\langle x_n - \theta x_{n-1} - \delta x_{n-2}, p^* - p\rangle \\ &= \lim_{k\to\infty}\langle x_{m_k} - \theta x_{m_k-1} - \delta x_{m_k-2}, p^* - p\rangle \\ &= \langle p^* - \theta p^* - \delta p^*, p^* - p\rangle,\end{aligned}$$

and this yields

$$(1 - \theta - \delta)\|p^* - p\|^2 = 0.$$

Since $\delta \leq 0 < 1 - \theta$, it follows that $p^* = p$, as claimed. Hence, $\{x_n\}$ converges weakly to a point in $(A + B)^{-1}(\mathbf{0})$, as asserted. This completes the proof. □

Remark 3.2. Our proposed Algorithm 1 is a bit different from the general context obtained in [26, Algorithm 1.2] and [26, Example 4.4]. This is because it is not yet clear if the operator $T(x_n, x_{n-1}) := J_\gamma^A(x_n - \gamma B(2x_n - x_{n-1}))$ in Algorithm 1 is an averaged quasi-nonexpansive (or quasi-nonexpansive) operator in order for it to properly fit into the framework of [26].

Remark 3.3. Fang and Huang [27] introduced and studied a class of generalized monotone operators called the class of relaxed η-α-monotone operators, which they defined as follows: An operator $M : \mathcal{H} \to \mathcal{H}$ is said to be relaxed $\eta - \alpha$- monotone, if there exist a mapping $\eta : \mathcal{H} \times \mathcal{H} \to \mathcal{H}$ and a function $\alpha : \mathcal{H} \to \mathbb{R}$ positively homogeneous of degree p (i.e., $\alpha(tz) = t^p\alpha(z)$ for all $t > 0$ and $z \in \mathcal{H}$, where $p > 1$) such that

$$\langle Mx - My, \eta(x, y)\rangle \geq \alpha(x - y)\ \forall x, y \in \mathcal{H}. \tag{3.43}$$

If M is a set-valued operator $M : \mathcal{H} \to 2^{\mathcal{H}}$, then inequality (3.43) becomes:

$$\langle u - v, \eta(x, y)\rangle \geq \alpha(x - y)\ \forall x, y \in \mathcal{H},\ u \in M(x), v \in M(y).$$

Observe that if $\eta(x, y) = x - y, \quad \forall x, y \in H$ and $\alpha(z) = \mu\|z\|^p$, for $p = 2$ and $\mu > 0$, then M is called μ-strongly monotone. Also, if $\eta(x, y) = x - y\ \forall x, y \in H$ and $\alpha = 0$, then the relaxed η-α monotone operator M reduces to a monotone mapping. Thus, the class of monotone operators is a subclass of the class of relaxed η-α monotone mappings.

Furthermore, this inclusion is proper. Indeed, let $H = \mathbb{R}$ and define $M : \mathbb{R} \to \mathbb{R}$ by $M(x) = -2x$. Then, M is a relaxed η-α-monotone operator but not a monotone operator. To see this, let $\eta : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined by $\eta(x, y) = 3(x - y)$ and $\alpha : \mathbb{R} \to \mathbb{R}$ by $\alpha(x) = -9x^2$. Then, for each $t > 0$ and $z \in \mathbb{R}$, we see that α is positively homogeneous of degree 2. That is,

$$\alpha(tz) = -9(tz)^2 = t^2(-9z^2) = t^2\alpha(z).$$

Moreover, for each $x, y \in \mathbb{R}$, we have

$$\langle Mx - My, \eta(x, y)\rangle = \langle -2(x - y), 3(x - y)\rangle$$

$$= -6(x-y)^2$$
$$\geq -9(x-y)^2 = \alpha(x-y).$$

Hence, M is relaxed η-α-monotone. However, it is not difficult to see that M is not monotone. Therefore, the relaxed η-α monotonicity is a more relaxed operator condition than the monotonicity assumption.

We can replace A with M (where M is a set-valued relaxed $\eta-\alpha$- monotone operator) or B with M (where M is a single-valued monotone operator) to consider the MIP (1.1) under more relaxed operator conditions. However, in doing so, we do not know if $A+M$ (or $B+M$) is still maximal monotone which is crucial in our convergence analysis. In particular, we have used the fact that $A+B$ is maximal monotone to obtain that $p \in (A+B)^{-1}(\mathbf{0})$ in our convergence analysis (immediately after (3.37)). Furthermore, we have used the fact that A is monotone to estimate the inequalities in (3.18) and (3.19), and the fact that B is monotone in (3.17). Since $\alpha(z)$ is in $\mathbb{R}$, not necessarily zero, and $\eta(x,y)$ is not necessarily equal to $x-y$, it might be difficult to estimate these inequalities. Lastly, we do not know if the resolvent $J^M_{\bar{\delta}}$ of M is single-valued, which is a property we used in the last equation of Algorithm 1. If $J^M_{\bar{\delta}}$ is not single-valued, then the last equation of Algorithm 1 becomes $x_{n+1} \in J^A_\gamma(w_n - \gamma By_n)$.

4. Numerical experiments

In this section we compare numerically, Algorithm 1 with the FRB splitting algorithm (1.4), the RFB splitting algorithm (1.5) (RFB), the two-step inertial FBF (2IFBF) algorithm [28] and the double inertial FBF (DIFBF) algorithm [29].

Example 1 (*Image Restoration Problem*).

$$\min_{w\in\mathcal{R}^l} \{\|\mathcal{P}w-e\|_2^2 + \bar{\delta}\|w\|_1\}, \tag{4.1}$$

where $\bar{\delta} > 0$ (in particular, we take $\bar{\delta} = 1$), $w \in \mathcal{R}^l$ is the original image that we intend to recover, $e \in \mathcal{R}^d$ is the observed image and $\mathcal{P} : \mathcal{R}^l \to \mathcal{R}^d$ is the blurring operator. Thus, we can apply Algorithm 1 to solve (4.1) by setting $A = \partial g$ the subdifferential of g, where $g(x) = \bar{\delta}\|w\|_1$, and setting $B = \nabla f$ the gradient of f, where $f(x) = \|\mathcal{P}w-e\|_2^2$. In this case, A is maximal monotone and B is $\|\mathcal{P}\|^2$-Lipschitz continuous and monotone (see, for example, [30, Section 5]). In order to measure the quality of the restored image, we use the signal-to-noise ratio (SNR) which is defined by

$$\mathrm{SNR} = 20 \times \log_{10}\left(\frac{\|w\|_2}{\|w-w^*\|_2}\right),$$

where w^* is the restored image. Note that the larger the SNR, the better the quality of the restored image.

For the implementation, we take $x_0 = \mathbf{0} \in \mathbb{R}^l$ and $x_{-1} = x_1 = \mathbf{1} \in \mathbb{R}^l$, and use the following image found in the MATLAB Image Processing Toolbox:

(a) Tire Image of size 205 × 232. To create the blurred and noisy image (observed image), we use the Gaussian blur of size 9 × 9 and standard deviation $\sigma = 4$, followed by an additive zero-mean white Gaussian noise with standard deviation (or noise level) 10^{-3} (see also [30, Subsection 5.1]).
(b) Cameraman Image of size 256 × 256. We use the Gaussian blur of size 7 × 7 and standard deviation $\sigma = 4$, followed by an additive zero-mean white Gaussian noise with standard deviation (or noise level) 10^{-3}.
(c) Medical Resonance Imaging (MRI) of size 128 × 128. We use the Gaussian blur of size 7 × 7 and standard deviation $\sigma = 4$, followed by an additive zero-mean white Gaussian noise with standard deviation (or noise level) 10^{-3}.

Example 2. Let $\mathcal{H} = l_2(\mathbb{R}) := \{x = (x_1, x_2, \dots, x_i, \dots),\ \ x_i \in \mathbb{R} : \sum_{i=1}^\infty |x_i|^2 < \infty\}$ and $\|x\|_{l_2} := \left(\sum_{i=1}^\infty |x_i|^2\right)^{\frac{1}{2}} \quad \forall x \in l_2(\mathbb{R})$.

Define $A, B : l_2 \to l_2$ by

$$Ax := (2x_1, 2x_2, \dots, 2x_i, \dots),\ \forall x \in l_2$$

and

$$Bx := \left(\frac{x_1+|x_1|}{2}, \frac{x_2+|x_2|}{2}, \dots, \frac{x_i+|x_i|}{2}, \dots\right),\ \forall x \in l_2.$$

Then A is maximal monotone and B is Lipschitz continuous and monotone with Lipschitz constant $L = 1$. For this example, we can find the exact solution to the MIP (1.1). Indeed, we have

$$(2x_1, 2x_2, \dots, 2x_i, \dots) + \left(\frac{x_1+|x_1|}{2}, \frac{x_2+|x_2|}{2}, \dots, \frac{x_i+|x_i|}{2}, \dots\right) = 0.$$

For each component i, we have

$$2x_i + \frac{x_i+|x_i|}{2} = 0.$$

If $x_i \geq 0$, then

$$2x_i + \frac{x_i+x_i}{2} = 0 \Longrightarrow x_i = 0.$$

Table 1
Numerical results for Example 1.

Algorithms	Tire		Cameraman		MRI	
	CPU	SNR	CPU	SNR	CPU	SNR
Algorithm 1($\delta = 0$)	1.0091	21.3905	1.1940	24.2937	0.3719	20.5504
Algorithm 1($\delta < 0$)	0.8084	24.0167	1.0284	27.2516	0.2810	22.2648
FRB	1.2183	21.3110	1.6017	24.7699	0.7217	20.7094
RFB	1.1101	20.7653	1.2044	22.0834	0.3160	20.0002
2IFBF	0.9901	22.0056	1.1041	25.1852	0.3009	20.8524
DIFBF	1.001	21.7181	1.1872	24.7729	0.3587	20.7678

Table 2
Numerical results for Example 2 with $\epsilon = 10^{-7}$.

Algorithms	Case 1		Case 2		Case 3		Case 4	
	CPU	Iter	CPU	Iter	CPU	Iter	CPU	Iter
Algorithm 1($\delta = 0$)	0.0520	261	0.0540	288	0.0410	275	0.0510	248
Algorithm 1($\delta < 0$)	0.0113	155	0.0161	172	0.0131	163	0.0181	148
FRB	1.0131	266	1.0202	294	1.0140	280	1.0101	253
RFB	0.1169	295	0.1318	326	0.1119	310	0.1092	280
2IFBF	0.0411	191	0.0401	212	0.0390	202	0.0389	184
DIFBF	0.0498	222	0.0472	258	0.0401	246	0.0499	224

Otherwise, we have

$$2x_i + \frac{x_i - x_i}{2} = 0 \Longrightarrow x_i = 0.$$

Thus, the solution set of MIP(1.1) for this example is $(A+B)^{-1}(\mathbf{0}) = \{(0, 0, \dots, 0, \dots)\}$.

For the implementation, we take the starting points:
Case 1: $x_0 = (\frac{2}{3}, \frac{4}{9}, \frac{8}{27}, \dots)$, $x_{-1} = x_1 = (\frac{2}{3}, \frac{4}{9}, \frac{8}{27}, \dots)$.
Case 2: $x_0 = (\frac{2}{3}, \frac{4}{9}, \frac{8}{27}, \dots)$, $x_{-1} = x_1 = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots)$.
Case 3: $x_0 = (1, \frac{1}{2}, \frac{1}{4}, \dots)$, $x_{-1} = x_1 = (\frac{4}{5}, \frac{16}{25}, \frac{64}{125}, \dots)$.
Case 4: $x_0 = (1, \frac{1}{4}, \frac{1}{9}, \dots)$, $x_{-1} = x_1 = (\frac{3}{4}, \frac{9}{16}, \frac{27}{64}, \dots)$.

During the implementation, we take $\theta = 0.012$ and $\delta = \{0, -0.001\}$ for Algorithm 1, and $\gamma = 0.001$ for Algorithm 1, the FRB splitting algorithm (1.4) and the RFB splitting algorithm (1.5). We also take $\mu = 0.9, \alpha = 0.012, \beta = -0.001, \lambda_1 = 0.001$ and $\tau_n = \frac{1}{n^2}$ for the 2IFBF [28], and $\mu = 0.9, \alpha_n = 1 - \frac{1}{10^n}, \beta_n = 0.1 - \frac{1}{1000+n}, \theta_n = 0.45 - \frac{1}{1000+n}, \rho_n = \frac{1}{n^2}, \mu_n = \frac{1}{n^2}$ and $\lambda_1 = 0.001$ for the DIFBF [29]. We then use the stopping criterion; $\text{TOL}_n := 0.5\|x_n - J^A(x_n - Bx_n)\|^2 < \epsilon$ for all algorithms, where ϵ is the predetermined error.

Note that the inertial parameters θ and δ are chosen randomly such that Assumption 3.2 is satisfied. This choice of parameters may not be the optimum choice but from the results of the numerical experiments, we can see that these random inertial parameters provide acceleration to our algorithm. Furthermore, the choice of $\gamma = 0.001$ does not give the optimal convergence performance of Algorithm 1, the FRB and the RFB. We only used this choice in our numerical experiments because it satisfies all the conditions in all the papers.

All the computations are performed using Matlab 2016 (b) which is running on a personal computer with an Intel(R) Core(TM) i5-2600 CPU at 2.30 GHz and 8.00 Gb-RAM.

In Tables 1–2, "Iter" and "CPU" mean the CPU time in seconds and the number of iterations, respectively (see Figs. 1–4).

5. Conclusion and future research

We have proposed an inertial RFB splitting algorithm for solving the MIP (1.1) in a real Hilbert space. We have also proved that the sequence generated by this algorithm converges weakly to a solution of the MIP (1.1). This algorithm inherits the attractive features of the FRB splitting algorithm (1.4) and the RFB splitting algorithm (1.5), namely, it involves only one forward evaluation of B (for which B is not required to be cocoercive) and one backward evaluation of A. However, numerical experiments show that our proposed inertial RFB splitting algorithm converges faster than both the FRB splitting algorithm and the RFB splitting algorithm. Part of our future research is to study the rate of convergence of the inertial RFB splitting algorithm. It would be interesting to study both the linear and strong convergence of the inertial RFB splitting algorithm. We also intend to develop a more general (relaxed) condition on the parameters θ, δ and γ in our future work.

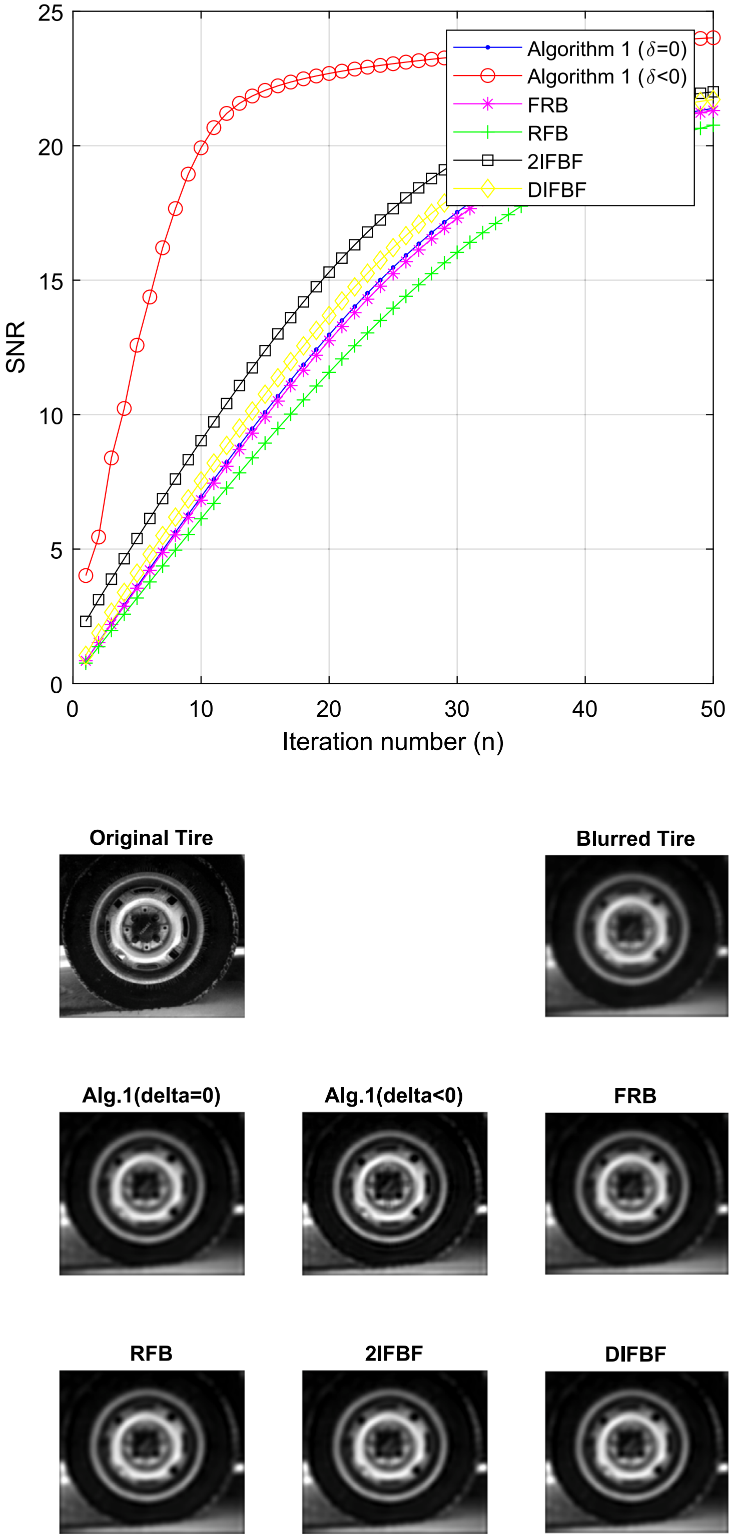

Fig. 1. Numerical results for Tire.

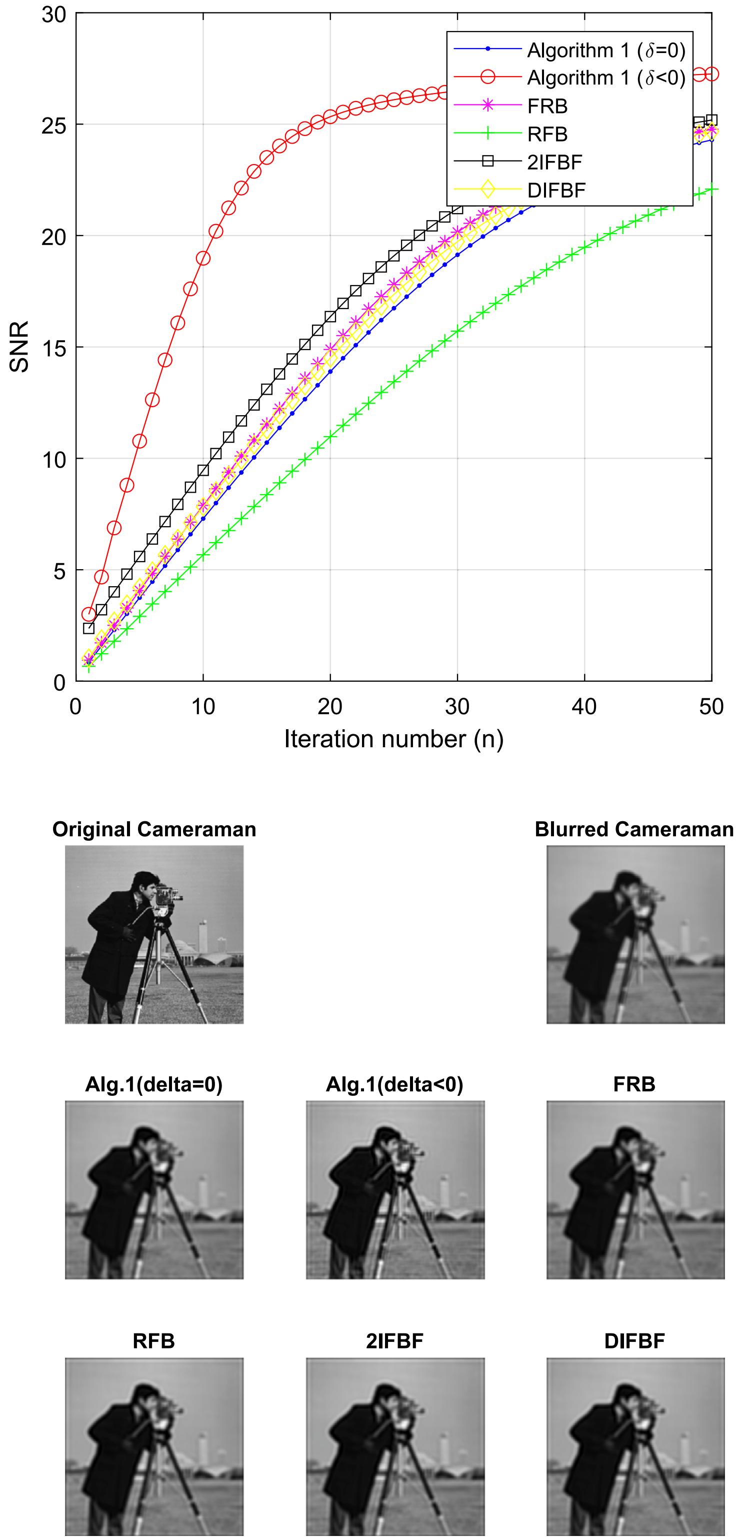

Fig. 2. Numerical results for Cameraman.

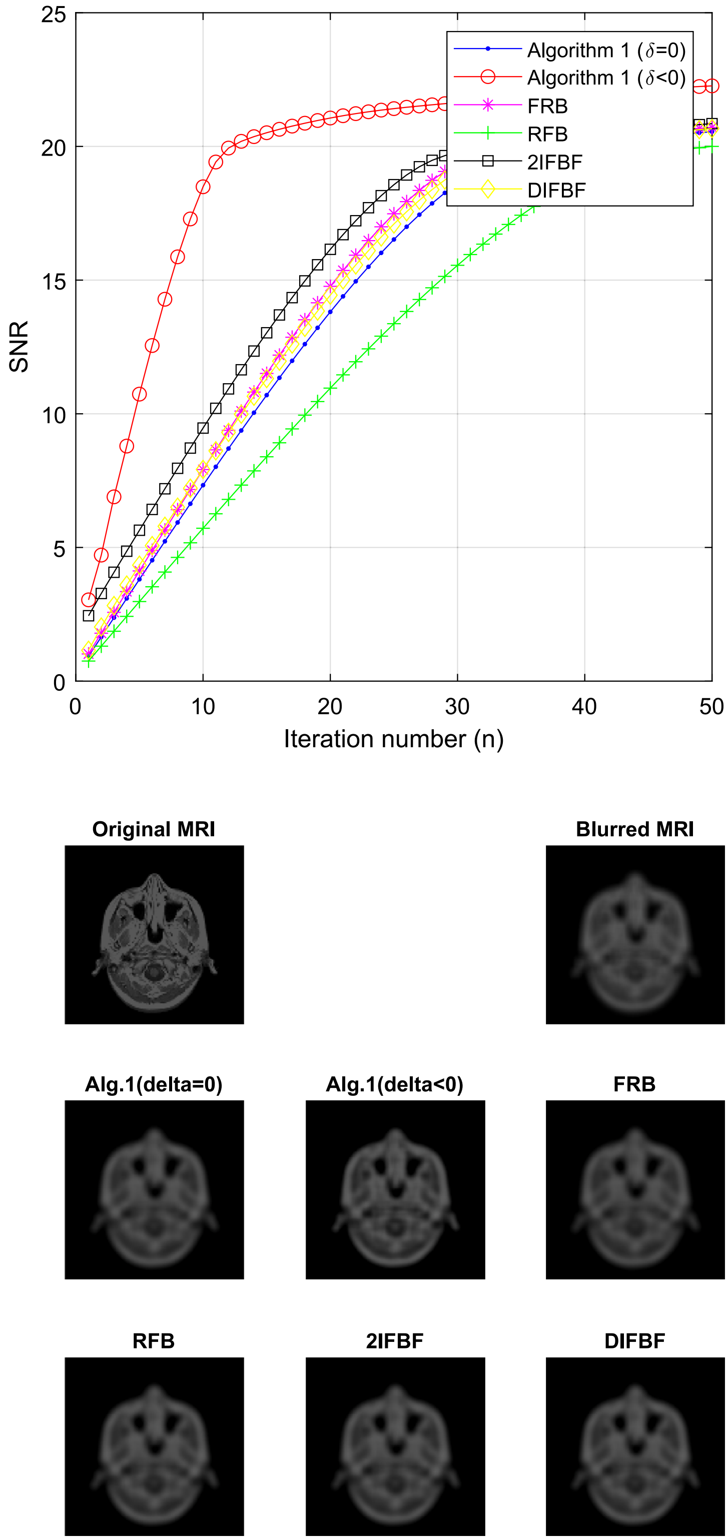

Fig. 3. Numerical results for MRI.

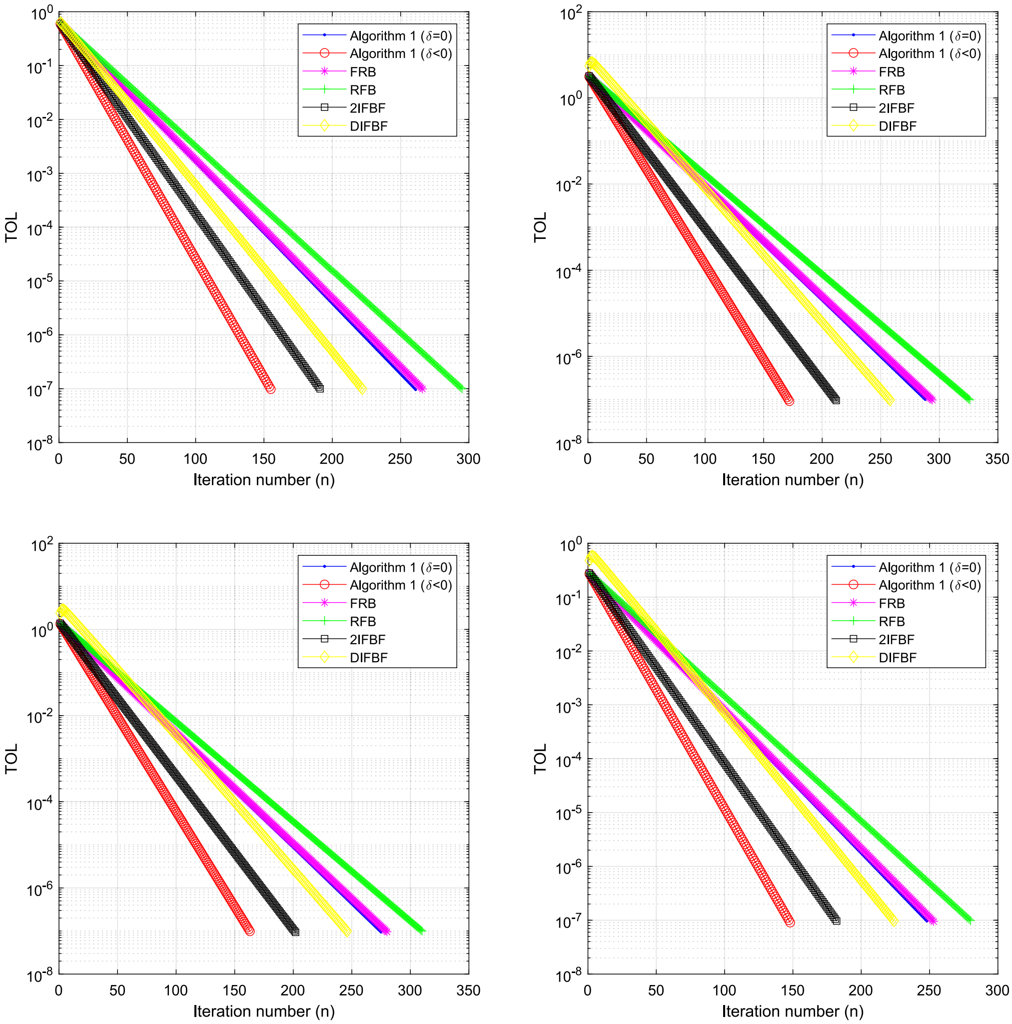

Fig. 4. The behavior of TOL_n for Example 2 with $\epsilon = 10^{-7}$: Top Left: **Case 1**; Top Right: **Case 2**; Bottom left: **Case 3**; Bottom Right: **Case 4**.

Acronyms

ADMM	Alternating direction method of multipliers (algorithm)
FBF	Forward-backward-forward (algorithm)
DIFBF	Double inertial FBF (algorithm)
2IFBF	Two-step inertial FBF (algorithm)
FRB	Forward-reflected-backward (algorithm)
MIP	Monotone inclusion problem
MRI	Medical resonance imaging
RFB	Reflected-forward–backward (algorithm)
SNR	Signal-to-noise ratio

Acknowledgments

The second author was partially supported by the Israel Science Foundation (Grant 820/17), the Fund for the Promotion of Research at the Technion, and by the Technion General Research Fund. The authors thank the reviewers and the principal editor for providing valuable comments and useful suggestions that helped improve the earlier version of this paper.

Data availability

No data was used for the research described in the article.

References

[1] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numer. Anal. 16 (1979) 964–979.
[2] G.B. Passty, Ergodic convergence to a zero of the sum of monotone operators in Hilbert spaces, J. Math. Anal. Appl. 72 (1979) 383–390.
[3] P. Tseng, A modified forward–backward splitting method for maximal monotone mappings, SIAM J. Control Optim. 38 (2000) 431–446.
[4] L.M. Briceño-Arias, P.L. Combettes, A monotone skew splitting model for composite monotone inclusions in duality, SIAM J. Optim. 21 (2011) 1230–1250.
[5] B.C. Vũ, A variable metric extension of the forward–backward–forward algorithm for monotone operators, Numer. Funct. Anal. Optim. 34 (2013) 1050–1065.
[6] B.C. Vũ, Almost sure convergence of the forward–backward-forward splitting algorithm, Optim. Lett. 10 (2016) 781–803.
[7] J.L. Lions, Optimal Control of Systems Governed By Partial Differential Equations, Springer, Berlin, 1971.
[8] Y. Malitsky, M.K. Tam, A forward–backward splitting method for monotone inclusions without cocoercivity, SIAM J. Optim. 30 (2020) 1451–1472.
[9] C. Izuchukwu, S. Reich, Y. Shehu, Convergence of two simple methods for solving monotone inclusion problems in reflexive Banach spaces, Results Math. 77 (2022) 143, http://dx.doi.org/10.1007/s00025-022-01694-5.
[10] V. Cevher, B.C. Vũ, A reflected forward–backward splitting method for monotone inclusions involving Lipschitzian operators, Set-Valued Var. Anal. 29 (2021) 163–174.
[11] Y.V. Malitsky, Projected reflected gradient methods for monotone variational inequalities, SIAM J. Optim. 25 (2015) 502–520.
[12] Y. Malitsky, Proximal extrapolated gradient methods for variational inequalities, Optim. Methods Softw. 33 (2018) 140–164.
[13] C. Izuchukwu, S. Reich, Y. Shehu, Relaxed inertial methods for solving the split monotone variational inclusion problem beyond co-coerciveness, Optimization 72 (2023) 607–646.
[14] Y. Nesterov, A method of solving a convex programming problem with convergence rate $O(1/k^2)$, Sov. Math. Doklady 27 (1983) 372–376.
[15] P. Peeyada, W. Cholamjiak, A new projection algorithm for variational inclusion problems and its application to cervical cancer disease prediction, J. Comput. Appl. Math. 441 (2024) 115702, http://dx.doi.org/10.1016/j.cam.2023.115702.
[16] P. Peeyada, H. Dutta, K. Shiangjen, W. Cholamjiak, A modified forward–backward splitting methods for the sum of two monotone operators with applications to breast cancer prediction, Math. Methods Appl. Sci. 46 (2023) 1251–1265.
[17] B.T. Polyak, Some methods of speeding up the convergence of iterates methods, USSR Comput. Math. Phys. 4 (1964) 1–17.
[18] W. Cholamjiak, P. Cholamjiak, S. Suantai, An inertial forward–backward splitting method for solving inclusion problems in Hilbert spaces, J. Fixed Point Theory Appl. 20 (2018) 42, http://dx.doi.org/10.1007/s11784-018-0526-5.
[19] D.A. Lorenz, T. Pock, An inertial forward–backward algorithm for monotone inclusions, J. Math. Imaging Vision 51 (2015) 311–325.
[20] A. Moudafi, M. Oliny, Convergence of a splitting inertial proximal method for monotone operators, J. Comput. Appl. Math. 155 (2003) 447–454.
[21] C. Poon, J. Liang, Trajectory of alternating direction method of multipliers and adaptive acceleration, in: 33rd Conference on Neural Information Processing Systems, Vancouver, Canada, 2019.
[22] C. Poon, J. Liang, Geometry of first-order methods and adaptive acceleration, arXiv:2003.03910.
[23] J. Liang, Convergence Rates of First-Order Operator Splitting Methods (Ph.D. thesis), Normandie Université, GREYC, CNRS UMR 6072, 2016.
[24] B.T. Polyak, Introduction to Optimization, Optimization Software, Publications Division, New York, 1987.
[25] Q.L. Dong, J.Z. Huang, X.H. Li, Y.J. Cho, Th.M. Rassias, MiKM: multi-step inertial Krasnosel'skii–Mann algorithm and its applications, J. Global Optim. 73 (2019) 801–824.
[26] P.L. Combettes, L.E. Glaudin, Quasi-nonexpansive iterations on the affine hull of orbits: from Mann's mean value algorithm to inertial methods, SIAM J. Optim. 27 (2017) 2356–2380.
[27] Y.P. Fang, N.J. Huang, Variational-like inequalities with generalized monotone mappings in Banach spaces, J. Optim. Theory Appl. 118 (2003) 327–338.
[28] V.T. Dung, P.K. Anh, D.V. Thong, Convergence of two-step inertial Tseng's extragradient methods for quasimonotone variational inequality problems, Commun. Nonlinear Sci. Numer. Simul. 136 (2024) 108110, http://dx.doi.org/10.1016/j.cnsns.2024.108110.
[29] V.T. Dung, P.K. Anh, D.V. Thong, A modified Tseng splitting method with double inertial steps for solving monotone inclusion problems, J. Sci. Comput. 96 (2023) 92, http://dx.doi.org/10.1007/s10915-023-02311-5.
[30] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imaging Sci. 2 (2009) 183–202.