Applications of Mathematics

Prasit Cholamjiak; Yekini Shehu

Inertial forward-backward splitting method in Banach spaces with application to compressed sensing

*Applications of Mathematics*, Vol. 64 (2019), No. 4, 409–435

Persistent URL: http://dml.cz/dmlcz/147798

# INERTIAL FORWARD-BACKWARD SPLITTING METHOD IN BANACH SPACES WITH APPLICATION TO COMPRESSED SENSING

Prasit Cholamjiak, Phayao, Yekini Shehu, Nsukka

Received November 21, 2018. Published online May 5, 2019.

*Abstract.* We propose a Halpern-type forward-backward splitting with inertial extrapolation step for finding a zero of the sum of accretive operators in Banach spaces. Strong convergence of the sequence of iterates generated by the method proposed is obtained under mild assumptions. We give some numerical results in compressed sensing to validate the theoretical analysis results. Our result is one of the few available inertial-type methods for zeros of the sum of accretive operators in Banach spaces.

*Keywords*: inertial term; forward-backward splitting; inclusion problem; strong convergence; Banach space

*MSC 2010*: 47H05, 47J20, 47J25

## 1. Introduction

Suppose $X$ is a real Banach space. Assume that $A\colon X \to X$ is an operator and $B\colon X \to 2^X$ a set-valued operator. In this paper, we consider the following inclusion problem: find $\hat{x} \in X$ such that

$$0 \in A\hat{x} + B\hat{x}. \tag{1.1}$$

It is well known that this problem includes, as special cases, nonsmooth convex optimization problems, variational inequalities, and convex-concave saddle-point problem, which have applications in compressed sensing, image processing, computer vision, machine learning and signal processing to mention but a few.

P. Cholamjiak is supported by the Thailand Research Fund and the University of Phayao under grant RSA6180084. The research of the second author is supported by the Alexander von Humboldt-Foundation.

DOI: 10.21136/AM.2019.0323-18

A popular method for solving problem (1.1) is the forward-backward splitting method, which is defined in the following manner: $x_1 \in X$ and

$$x_{n+1} = J_r^B(x_n - rAx_n), \quad n \geqslant 1, \tag{1.2}$$

where $J_r^B := (I+rB)^{-1}$, $r > 0$. The forward-backward splitting method (1.2) (as the name implies) is based on an explicit forward step with respect to $A$ followed by an implicit backward step with respect to $B$. Furthermore, forward-backward splitting method (1.2) includes, in particular, the proximal point algorithm (see e.g. [8], [11], [21], [29], [35]) and the gradient method (see e.g. [5], [20]).

Forward-backward splitting method (1.2) has been studied by many authors in the literature, see, for example, [18], [25], [31], [40]. It has been established in these papers (see e.g. [31]) that forward-backward splitting method (1.2) converges weakly to a zero of (1.1) in general.

In [26], López et al. introduced the following Halpern-type forward-backward method: $x_1 \in X$ and

$$x_{n+1} = \alpha_n u + (1-\alpha_n)(J_{r_n}^B(x_n - r_n(Ax_n + a_n)) + b_n), \tag{1.3}$$

where $J_r^B$ is the resolvent of $B$, $\{r_n\} \subset (0,\infty)$, $\{\alpha_n\} \subset (0,1]$ and $\{a_n\}, \{b_n\}$ are error sequences in $X$. López et al. proved in [26] that the sequence $\{x_n\}$ generated by (1.3) strongly converges to a zero of (1.1) under some appropriate conditions. Several authors have obtained strong convergence results both in Hilbert and Banach spaces for finding a zero of (1.1), see, for example, [13], [14], [15], [17], [19], [36], [37], [39], [40].

Using the idea in [33], Alvarez and Attouch [1] introduced an inertial proximal point algorithm for finding a zero of (1.1) when $A = 0$ and $B$ is the maximal monotone operator in a real Hilbert space: $x_0, x_1 \in H$,

$$\begin{cases} y_n = x_n + \beta_n(x_n - x_{n-1}), \\ x_{n+1} = J_{r_n}^B(y_n), & n \geqslant 1. \end{cases} \tag{1.4}$$

Alvarez and Attouch [1] obtained weak convergence of (1.4) under appropriate conditions on $\{\beta_n\}$ and $\{r_n\}$. Using the ideas in [33] and [1], Lorenz and Pock in [27] introduced an accelerated iterative method which is a combination of the inertial extrapolation method and (1.2) for finding a zero of (1.1) in real Hilbert spaces. It was shown numerically in [27] that (1.2) with inertial extrapolation step (accelerated version) converges faster than the unaccelerated version. Several other modifications of (1.2) with inertial extrapolation step have been considered in Hilbert spaces by many authors (see, for example, [4], [6], [7], [10], [30], [32]).

**Contribution.** In this work, our main motivation are the results in [2], [15], [19], [26]. Our contribution is threefold:

- ▷ We extend the forward-backward splitting method with inertial extrapolation step for solving (1.1) from Hilbert spaces to Banach spaces. The inertial modification of the forward-backward splitting method has already been suggested in several papers such as [2], [15], [19]. However, the results presented in [2], [15], [19] are done in real Hilbert spaces. Furthermore, strong convergence results are presented in Hilbert spaces in [19] and [15] using Haugazeau approach [23] and Halpern regularization technique [22] respectively. The authors in [2] presented weak convergence analysis in real Hilbert spaces. In this paper, we present strong convergence analysis of inertial modification using the Halpern regularization approach in a uniformly convex and $q$-uniformly smooth Banach space (e.g., $L_p$ spaces with $1 < p < \infty$), which is more general than Hilbert space. Therefore, our results in this paper extend and complement the recent results in [2], [15], [19].
- ▷ We give strong convergence analysis of our proposed accelerated forward-backward splitting method in uniformly convex and $q$-uniformly smooth Banach space and give some applications to inverse problems in signal recovery and nonlinear integro-differential systems involving the generalized $p$-Laplacian. These complement the unaccelerated results of López et al. given in [26].
- ▷ We show, using the numerical implementations in compressed sensing and some constrained convex minimization problem, that our proposed accelerated forward-backward splitting method outperforms the unaccelerated method proposed in [26] by López et al.

## 2. Preliminaries

Let $X$ be a real Banach space. The *modulus of convexity* of $X$ is defined as the function $\delta\colon (0,2] \to [0,1]$,

$$\delta(\varepsilon) = \inf\Big\{1 - \Big\|\frac{x+y}{2}\Big\|\colon x, y \in X,\ \|x\| = \|y\| = 1,\ \|x-y\| \geqslant \varepsilon\Big\}. \tag{2.1}$$

Here $X$ is said to be *uniformly convex* if $\delta(\varepsilon) > 0$ for all $\varepsilon \in (0,2]$.

The *modulus of smoothness* of $X$ is the function $\varrho\colon \mathbb{R}^+ \to \mathbb{R}^+$ defined by

$$\varrho(t) = \sup\Big\{\frac{\|x+ty\| + \|x-ty\|}{2} - 1\colon x, y \in X,\ \|x\| = \|y\| = 1\Big\}. \tag{2.2}$$

We say $X$ is *uniformly smooth* if $\lim\limits_{t\to 0} \varrho(t)/t = 0$ and $X$ is said to be *$q$-uniformly smooth* with $1 < q \leqslant 2$, if there exists a constant $k_q > 0$ such that $\varrho(t) \leqslant k_q t^q$

for $t > 0$. If $X$ is $q$-uniformly smooth, then it is uniformly smooth (see e.g. [16]). Suppose that $X^*$ is the dual space of $X$. The generalized duality mapping $J_q$ $(q > 1)$ of $X$ is defined by $J_q(x) := \{j_q(x) \in X^*\colon \langle x, j_q(x)\rangle = \|x\|^q,\ \|j_q(x)\| = \|x\|^{q-1}\}$ for all $x \in X$, where $\langle \cdot, \cdot\rangle$ denotes the duality pairing between $X$ and $X^*$. In particular, we call $J_2 := J$, the normalized duality mapping on $X$. Furthermore, (see e.g. [42], pp. 1128)

$$J_q(x) = \|x\|^{q-2} J(x), \quad x \neq 0. \tag{2.3}$$

It is well known (see, for example, [16]) that $X$ is uniformly smooth if and only if the duality mapping $J_q$ is single-valued and norm-to-norm uniformly continuous on bounded subsets of $X$.

Let $B\colon X \to 2^X$. We denote the domain of $B$ by $D(B) = \{x \in X\colon Bx \neq \emptyset\}$ and its range by $R(B) = \bigcup\{Bz\colon z \in D(B)\}$. We say that $B$ is *accretive* if, for each $x, y \in D(A)$, there exists $j(x-y) \in J(x-y)$ such that

$$\big\langle u - v, j(x-y)\big\rangle \geqslant 0, \quad u \in Bx,\ v \in By. \tag{2.4}$$

The operator $B$ is said to be *m-accretive* if $R(I + rB) = X$ for all $r > 0$. Given $\alpha > 0$ and $q \in (1, \infty)$, we say that a single-valued accretive operator $A$ is $\alpha$-inverse strongly accretive ($\alpha$-isa, for short) of order $q$ if, for each $x, y \in D(A)$, there exists $j_q(x-y) \in J_q(x-y)$ such that

$$\langle Ax - Ay, j_q(x-y)\rangle \geqslant \alpha\|u - v\|^q. \tag{2.5}$$

Let $\emptyset \neq C \subset X$ and let $T\colon C \to C$ be a nonlinear mapping. The set of fixed points of $T$ is defined by $\mathrm{Fix}(T) := \{x \in C\colon x = Tx\}$.

Let $C$ be a nonempty, closed and convex subset of $X$ and let $D \subset C$. A *retraction* from $C$ to $D$ is a mapping $Q\colon C \to D$ such that $Qx = x$ for all $x \in D$. Furthermore, the retraction $Q$ is *nonexpansive* if $\|Qx - Qy\| \leqslant \|x - y\|$ for all $x, y \in C$ and *sunny* if, for each $x \in C$ and $t \geqslant 0$, we have

$$Q(tx + (1-t)Qx) = Qx, \tag{2.6}$$

whenever $tx + (1-t)Qx \in C$. The following result gives the information on how sunny nonexpansive retraction can be constructed.

**Theorem 2.1** ([34], Corollary 1)**.** *Let $X$ be a uniformly smooth Banach space and let $T\colon C \to C$ be a nonexpansive mapping with a fixed point. For each fixed $u \in C$ and every $t \in (0,1)$, the unique fixed point $x_t \in C$ of the contraction $C \ni x \mapsto tu + (1-t)Tx$ converges strongly as $t \to 0$ to a fixed point of $T$. Define $Q\colon C \to D$ by $Qu = \lim_{t\to 0} x_t$. Then $Q$ is the unique sunny nonexpansive retract from $C$ to $D$.*

For the rest of this paper, we will adopt the notation

$$T_r^{A,B} = J_r^B(I - rA) = (I + rB)^{-1}(I - rA), \quad r > 0. \tag{2.7}$$

The following lemmas will be used in the convergence analysis of this paper.

**Lemma 2.1** ([24], page 82). *If $x > y > 0$ and $r > 1$, then*

$$\frac{x^r - y^r}{x - y} < rx^{r-1}. \tag{2.8}$$

**Lemma 2.2** ([12], page 33). *Let $q > 1$ and let $X$ be a real normed space with the generalized duality mapping $J_q$. Then, for any $x, y \in X$, we have*

$$\|x + y\|^q \leqslant \|x\|^q + q\langle y, j_q(x + y)\rangle \tag{2.9}$$

*for all $j_q(x + y) \in J_q(x + y)$.*

**Lemma 2.3** ([42], Corollary 1′). *Let $1 < q \leqslant 2$ and let $X$ be a smooth Banach space. Then the following statements are equivalent:*

(i) *$X$ is $q$-uniformly smooth.*
(ii) *There is a constant $k_q > 0$ such that for all $x, y \in X$*

$$\|x + y\|^q \leqslant \|x\|^q + q\langle y, j_q(x)\rangle + k_q\|y\|^q. \tag{2.10}$$

The best constant $k_q$ will be called the $q$-uniform smoothness coefficient of $X$.

**Lemma 2.4** ([26], Lemmas 3.1, 3.2). *Let $X$ be a Banach space. Let $A\colon X \to X$ be an $\alpha$-isa of order $q$ and $B\colon X \to 2^X$ an $m$-accretive operator. Then:*

(i) *For $r > 0$, $\mathrm{Fix}(T_r^{A,B}) = (A + B)^{-1}(0)$.*
(ii) *For $0 < s \leqslant r$ and $x \in X$, $\|x - T_s^{A,B}x\| \leqslant 2\|x - T_r^{A,B}x\|$.*

**Lemma 2.5** ([26], Lemma 3.3). *Let $X$ be a uniformly convex and $q$-uniformly smooth Banach space for some $q \in (1, 2]$. Assume that $A$ is a single-valued $\alpha$-isa of order $q$ in $X$. Then, given $r > 0$, there exists a continuous, strictly increasing and convex function $\varphi_q\colon \mathbb{R}^+ \to \mathbb{R}^+$ with $\varphi_q(0) = 0$ such that, for all $x, y \in B_r$,*

$$\begin{aligned} \|T_r^{A,B}x - T_r^{A,B}y\|^q \leqslant{}& \|x - y\|^q - r(\alpha q - r^{q-1}k_q)\|Ax - Ay\|^q \\ &- \varphi_q(\|(I - J_r^B)(I - rA)x - (I - J_r^B)(I - rA)y\|), \end{aligned} \tag{2.11}$$

*where $k_q$ is the $q$-uniform smoothness coefficient of $X$.*

**Lemma 2.6** ([28], Lemma 3.1). *Let $\{a_n\}$ and $\{c_n\}$ be sequences of nonnegative real numbers such that*

$$a_{n+1} \leqslant (1-\delta_n)a_n + b_n + c_n, \quad n \geqslant 1, \tag{2.12}$$

*where $\{\delta_n\}$ is a sequence in $(0,1)$ and $\{b_n\}$ is a real sequence. Assume $\sum\limits_{n=1}^{\infty} c_n < \infty$. Then the following results hold:*

(i) *If $b_n \leqslant \delta_n M$ for some $M \geqslant 0$, then $\{a_n\}$ is a bounded sequence.*
(ii) *If $\sum\limits_{n=1}^{\infty} \delta_n = \infty$ and $\limsup\limits_{n\to\infty} b_n/\delta_n \leqslant 0$, then $\lim\limits_{n\to\infty} a_n = 0$.*

We will adopt the following notation in this paper:

▷ $x_n \to x$ means that $x_n \to x$ strongly.
▷ $x_n \rightharpoonup x$ means that $x_n \to x$ weakly.

## 3. Approximation method

In this section, we propose our method and state certain conditions under which we obtain the desired convergence for our proposed method. First, we give the conditions governing the cost function and the sequence of parameters below.

**Assumption 3.1.**

(a) Let $X$ be a uniformly convex and $q$-uniformly smooth Banach space.
(b) Let $A\colon X \to X$ be an $\alpha$-isa of order $q$ and $B\colon X \to 2^X$ an $m$-accretive operator.
(c) Assume that the solution set satisfies $S = (A+B)^{-1}(0) \neq \emptyset$.

**Assumption 3.2.** Choose sequences $\{\alpha_n\}_{n=1}^{\infty} \subset (0,1)$, $\{a_n\}_{n=1}^{\infty}, \{b_n\}_{n=1}^{\infty} \subset X$, and $\{\varepsilon_n\}_{n=1}^{\infty}, \{r_n\}_{n=1}^{\infty} \subset (0,\infty)$ such that the following conditions are satisfied:

(i) $\lim\limits_{n\to\infty} \|a_n\|/\alpha_n = 0$, $\lim\limits_{n\to\infty} \|b_n\|/\alpha_n = 0$,
(ii) $\lim\limits_{n\to\infty} \alpha_n = 0$, $\sum\limits_{n=1}^{\infty} \alpha_n = \infty$,
(iii) $0 < \liminf\limits_{n\to\infty} r_n \leqslant \limsup\limits_{n\to\infty} r_n < (\alpha q/k_q)^{1/(q-1)}$,
(iv) $\varepsilon_n = o(\alpha_n)$, which means $\lim\limits_{n\to\infty} \varepsilon_n/\alpha_n = 0$.

We now give our proposed method below.

**Algorithm 3.1**

*Step 0*: Let Assumptions 3.1 and 3.2 hold. Let $\beta \in [0,1)$ and $x_0, x_1 \in X$ be given starting points. Set $n := 1$.

*Step 1*: Given the iterates $x_{n-1}$ and $x_n$, $n \geqslant 1$, choose $\beta_n$ such that $0 \leqslant \beta_n \leqslant \overline{\beta}_n$, where

$$\overline{\beta}_n = \begin{cases} \min\Big\{\beta, \dfrac{\varepsilon_n}{\|x_n - x_{n-1}\|}\Big\}, & x_n \neq x_{n-1}, \\ \beta, & \text{otherwise.} \end{cases}$$

*Step 2*: Compute

$$\begin{cases} y_n = x_n + \beta_n(x_n - x_{n-1}), \\ x_{n+1} = \alpha_n x_0 + (1-\alpha_n)(J^B_{r_n}(y_n - r_n(Ay_n + a_n)) + b_n), \quad n \geqslant 1, \end{cases} \tag{3.1}$$

where $J^B_{r_n} = (I + r_n B)^{-1}$.

*Step 3*: Set $n \leftarrow n+1$, and go to *Step 1*.

Remark 3.1. (a) We remark that Step 1 in our Algorithm 3.1 can be easily implemented in numerical computation since it involves only the two previous iterates $x_{n-1}$ and $x_n$. Hence, the value of $\|x_n - x_{n-1}\|$ is a priori known before choosing $\beta_n$. (See [38].)

(b) Observe that Assumption 3.2 and Algorithm 3.1 imply

$$\lim_{n\to\infty} \beta_n \|x_n - x_{n-1}\| = 0 \quad \text{and} \quad \lim_{n\to\infty} \frac{\beta_n}{\alpha_n}\|x_n - x_{n-1}\| = 0.$$

**3.1. Convergence analysis.** In this section, we obtain the strong convergence analysis of our proposed Algorithm 3.1 to a zero of (1.1). To do this, we assume that Assumptions 3.1 and 3.2 hold for the rest of this paper. We first show that the generated sequences $\{x_n\}$ and $\{y_n\}$ in Algorithm 3.1 are bounded in the following lemma.

**Lemma 3.1.** *The sequences* $\{x_n\}$ *and* $\{y_n\}$ *generated by Algorithm* 3.1 *are bounded.*

Proof. Let $z = Q(x_0)$, where $Q$ is the sunny nonexpansive retraction of $X$ onto $S$. Then $z \in S$. Let $T^{A,B}_{r_n} := J^B_{r_n}(I - r_n A)$. Then we can write $J^B_{r_n}(y_n - r_n \times (Ay_n + a_n)) + b_n = T^{A,B}_{r_n} y_n + e_n$, where $e_n = J^B_{r_n}(y_n - r_n(Ay_n + a_n)) + b_n - T^{A,B}_{r_n} y_n$. Hence $x_{n+1} = \alpha_n x_0 + (1-\alpha_n)(T^{A,B}_{r_n} y_n + e_n)$. By Lemma 2.5, $T^{A,B}_{r_n}$ is nonexpansive and by Lemma 2.4 (i), $\mathrm{Fix}(T_{r_n}^{A,B}) = S$. It follows that

$$\begin{aligned}
(3.2)\qquad & \|x_{n+1} - z\| \\
&\leqslant \alpha_n\|x_0 - z\| + (1-\alpha_n)\|T_{r_n}^{A,B}y_n - T_{r_n}^{A,B}z\| + (1-\alpha_n)\|e_n\| \\
&\leqslant \alpha_n\|x_0 - z\| + (1-\alpha_n)\|y_n - z\| + (1-\alpha_n)\|e_n\| \\
&= \alpha_n\|x_0 - z\| + (1-\alpha_n)\|x_n - z + \beta_n(x_n - x_{n-1})\| + (1-\alpha_n)\|e_n\| \\
&\leqslant \alpha_n\|x_0 - z\| + (1-\alpha_n)[\|x_n - z\| + \beta_n\|x_n - x_{n-1}\|] + (1-\alpha_n)\|e_n\| \\
&= \alpha_n\|x_0 - z\| + (1-\alpha_n)\|x_n - z\| \\
&\quad + \alpha_n\Big[\frac{(1-\alpha_n)\beta_n\|x_n - x_{n-1}\|}{\alpha_n} + \frac{(1-\alpha_n)\|e_n\|}{\alpha_n}\Big].
\end{aligned}$$

Since $J_{r_n}^B$ is nonexpansive, we obtain

$$\begin{aligned}
(3.3)\qquad \|e_n\| &= \|J_{r_n}^B(y_n - r_n(Ay_n + a_n)) + b_n - T_{r_n}^{A,B}y_n\| \\
&\leqslant \|J_{r_n}^B(y_n - r_n(Ay_n + a_n)) - J_{r_n}^B(y_n - r_nAy_n)\| + \|b_n\| \\
&\leqslant \|(y_n - r_n(Ay_n + a_n)) - (y_n - r_nAy_n)\| + \|b_n\| = r_n\|a_n\| + \|b_n\| \\
&\leqslant \Big(\frac{\alpha q}{k_q}\Big)^{1/(q-1)}\|a_n\| + \|b_n\|,
\end{aligned}$$

which gives $\|e_n\|/\alpha_n \to 0$ as $n \to \infty$ by condition (i). So from condition (iv) we get that

$$t_n = \frac{(1-\alpha_n)\beta_n\|x_n - x_{n-1}\|}{\alpha_n} + \frac{(1-\alpha_n)\|e_n\|}{\alpha_n} \to 0$$

as $n \to \infty$. Hence, it is bounded. Put

$$M = \max\Big\{\|x_0 - z\|,\ \sup_{n\geqslant 1} t_n\Big\}.$$

Then (3.2) becomes

$$(3.4)\qquad \|x_{n+1} - z\| \leqslant (1-\alpha_n)\|x_n - z\| + \alpha_n M.$$

Applying Lemma 2.6 (i) in (3.4), we can conclude that $\{x_n\}$ is bounded and so is $\{y_n\}$. □

We now give the strong convergence theorem of Algorithm 3.1.

**Theorem 3.1.** *The sequence $\{x_n\}$ generated by Algorithm 3.1 converges in norm to $z = Q(x_0)$, where $Q$ is the sunny nonexpansive retraction of $X$ onto $S$.*

P r o o f. Using Lemma 2.3 (ii) in (3.1), we get

$$\|y_n - z\|^q = \|x_n - z + \beta_n(x_n - x_{n-1})\|^q \leqslant \|x_n - z\|^q + q\beta_n\langle x_n - x_{n-1}, j_q(x_n - z)\rangle + k_q\beta_n^q\|x_n - x_{n-1}\|^q. \tag{3.5}$$

By Lemma 2.3 (ii), we have $\langle y, j_q(x)\rangle \leqslant \frac{1}{q}[\|x\|^q + k_q\|y\|^q - \|x-y\|^q]$ for all $x, y \in X$, and

$$\langle x_n - x_{n-1}, j_q(x_n - z)\rangle \leqslant \frac{1}{q}(\|x_n - z\|^q + k_q\|x_n - x_{n-1}\|^q - \|x_{n-1} - z\|^q). \tag{3.6}$$

Combining (3.5) and (3.6), we get

$$\|y_n - z\|^q \leqslant \|x_n - z\|^q + \beta_n(\|x_n - z\|^q - \|x_{n-1} - z\|^q) + k_q\beta_n(\beta_n^{q-1} + 1)\|x_n - x_{n-1}\|^q. \tag{3.7}$$

Using Lemma 2.2 and Lemma 2.5, we get for some $M^* > 0$,

$$\begin{aligned}\|x_{n+1} - z\|^q &= \|\alpha_n(x_0 - z) + (1-\alpha_n)(T_{r_n}^{A,B}y_n + e_n - z)\|^q \\ &\leqslant \|\alpha_n(x_0 - z) + (1-\alpha_n)(T_{r_n}^{A,B}y_n - z)\|^q \\ &\quad + q(1-\alpha_n)\langle e_n, j_q(x_{n+1} - z)\rangle \\ &\leqslant \|\alpha_n(x_0 - z) + (1-\alpha_n)(T_{r_n}^{A,B}y_n - z)\|^q \\ &\quad + q(1-\alpha_n)\|e_n\|\|j_q(x_{n+1} - z)\| \\ &\leqslant \|\alpha_n(x_0 - z) + (1-\alpha_n)(T_{r_n}^{A,B}y_n - z)\|^q + q(1-\alpha_n)M^*\|e_n\| \\ &\leqslant (1-\alpha_n)^q\|T_{r_n}^{A,B}y_n - z\|^q + q\alpha_n\langle x_0 - z, j_q(x_{n+1} - z - (1-\alpha_n)e_n)\rangle \\ &\quad + q(1-\alpha_n)M^*\|e_n\| \\ &\leqslant (1-\alpha_n)\|y_n - z\|^q - r_n(1-\alpha_n)(\alpha q - r_n^{q-1}k_q)\|Ay_n - Az\|^q \\ &\quad - (1-\alpha_n)\varphi_q(\|(I - J_{r_n}^B)(I - r_nA)y_n - (I - J_{r_n}^B)(I - r_nA)z\|) \\ &\quad + \alpha_n q\langle x_0 - z, j_q(x_{n+1} - z - (1-\alpha_n)e_n)\rangle + qM^*\|e_n\|.\end{aligned} \tag{3.8}$$

Combining (3.7) and (3.8), we get

$$\begin{aligned}\|x_{n+1} - z\|^q &\leqslant (1-\alpha_n)\|x_n - z\|^q + (1-\alpha_n)\beta_n(\|x_n - z\|^q - \|x_{n-1} - z\|^q) \\ &\quad + (1-\alpha_n)k_q\beta_n(\beta_n^{q-1} + 1)\|x_n - x_{n-1}\|^q \\ &\quad - (1-\alpha_n)r_n(\alpha q - r_n^{q-1}k_q)\|Ay_n - Az\|^q \\ &\quad - (1-\alpha_n)\varphi_q(\|(I - J_{r_n}^B)(I - r_nA)y_n - (I - J_{r_n}^B)(I - r_nA)z\|) \\ &\quad + \alpha_n q\langle x_0 - z, j_q(x_{n+1} - z - (1-\alpha_n)e_n)\rangle + qM^*\|e_n\|.\end{aligned} \tag{3.9}$$

By condition (iii) of Assumption 3.2, there is $\delta > 0$ such that $r_n(\alpha q - r_n^{q-1}k_q) \geqslant \delta > 0$ for all $n \in \mathbb{N}$. Set $\Gamma_n = \|x_n - z\|^q$ for all $n \in \mathbb{N}$. Then we obtain from (3.9) that

$$\begin{aligned}\Gamma_{n+1} \leqslant{} & (1-\alpha_n)\Gamma_n + (1-\alpha_n)\beta_n(\Gamma_n - \Gamma_{n-1}) \\ & + (1-\alpha_n)k_q\beta_n(\beta_n^{q-1}+1)\|x_n - x_{n-1}\|^q - (1-\alpha_n)\delta\|Ay_n - Az\|^q \\ & - (1-\alpha_n)\varphi_q(\|(I-J_{r_n}^B)(I-r_nA)y_n - (I-J_{r_n}^B)(I-r_nA)z\|) \\ & + \alpha_n q\langle x_0 - z, j_q(x_{n+1} - z - (1-\alpha_n)e_n)\rangle + qM^*\|e_n\|.\end{aligned} \tag{3.10}$$

We next consider the following two cases:

*Case 1*: Suppose there exists $N \in \mathbb{N}$ such that $\Gamma_{n+1} \leqslant \Gamma_n$ for all $n \geqslant N$. In this case, $\lim \Gamma_n$ exists and (3.10) implies that

$$\begin{aligned}&(1-\alpha_n)\delta\|Ay_n - Az\|^q \\ &\quad\leqslant (\Gamma_n - \Gamma_{n+1}) + (1-\alpha_n)\beta_n(\Gamma_n - \Gamma_{n-1}) \\ &\qquad + (1-\alpha_n)k_q\beta_n(\beta_n^{q-1}+1)\|x_n - x_{n-1}\|^q \\ &\qquad + \alpha_n q\langle x_0 - z, j_q(x_{n+1} - z - (1-\alpha_n)e_n)\rangle + qM^*\|e_n\|\end{aligned} \tag{3.11}$$

and

$$\begin{aligned}&(1-\alpha_n)\varphi_q(\|(I-J_{r_n}^B)(I-r_nA)y_n - (I-J_{r_n}^B)(I-r_nA)z\|) \\ &\quad\leqslant (\Gamma_n - \Gamma_{n+1}) + (1-\alpha_n)\beta_n(\Gamma_n - \Gamma_{n-1}) \\ &\qquad + (1-\alpha_n)k_q\beta_n(\beta_n^{q-1}+1)\|x_n - x_{n-1}\|^q \\ &\qquad + \alpha_n q\langle x_0 - z, j_q(x_{n+1} - z - (1-\alpha_n)e_n)\rangle + qM^*\|e_n\|.\end{aligned} \tag{3.12}$$

Note that Assumption 3.2 (i) implies $\lim\limits_{n\to\infty}\|e_n\| = 0$. So from Assumption 3.2 (ii), the boundedness of $\{x_n\}$, and $\beta_n\|x_n - x_{n-1}\| \to 0$ in (3.11), we get

$$\lim_{n\to\infty}(1-\alpha_n)\delta\|Ay_n - Az\|^q = 0.$$

Hence,

$$\|Ay_n - Az\| \to 0, \quad n \to \infty. \tag{3.13}$$

Also from (3.12), we get

$$\varphi_q(\|(I-J_{r_n}^B)(I-r_nA)y_n - (I-J_{r_n}^B)(I-r_nA)z\|) \to 0, \quad n \to \infty. \tag{3.14}$$

By the continuity of $\varphi_q$, we obtain from (3.14) that

$$\|(I-J_{r_n}^B)(I-r_nA)y_n - (I-J_{r_n}^B)(I-r_nA)z\| \to 0, \quad n \to \infty.$$

Expanding and noting that $z = J_{r_n}^B(I - r_n A)z$, we have

$$\lim_{n\to\infty} \|y_n - r_n A y_n - T_{r_n}^{A,B} y_n + r_n A z\| = 0. \tag{3.15}$$

Relations (3.13) and (3.15) imply that

$$\lim_{n\to\infty} [\|Ay_n - Az\| + \|y_n - r_n A y_n - T_{r_n}^{A,B} y_n + r_n A z\|] = 0. \tag{3.16}$$

Consequently,

$$\lim_{n\to\infty} \|T_{r_n}^{A,B} y_n - y_n\| = 0. \tag{3.17}$$

Since $\liminf_{n\to\infty} r_n > 0$, there exists $\varepsilon > 0$ such that $r_n \geqslant \varepsilon$ for all $n \geqslant 1$. Then by Lemma 2.4 (ii), we get

$$\lim_{n\to\infty} \|T_{\varepsilon}^{A,B} y_n - y_n\| \leqslant 2 \lim_{n\to\infty} \|T_{r_n}^{A,B} y_n - y_n\| = 0. \tag{3.18}$$

Let $z_t$ be the unique fixed point of $z \mapsto tx_0 + (1-t)T_{\varepsilon}^{A,B} z$, $t \in (0,1)$. By Reich's theorem in [34], we get $z_t \to Q_S(x_0) = z$, $t \to 0$. By the subdifferential inequality, we obtain (noting that $T_{\varepsilon}^{A,B}$ is nonexpansive)

$$\begin{aligned}
\|z_t - y_n\|^2 &= \|t(x_0 - y_n) + (1-t)(T_{\varepsilon}^{A,B} z_t - y_n)\|^2 \\
&\leqslant (1-t)^2 \|T_{\varepsilon}^{A,B} z_t - y_n\|^2 + 2t\langle x_0 - y_n, j(z_t - y_n)\rangle \\
&\leqslant (1-t)^2 (\|T_{\varepsilon}^{A,B} z_t - T_{\varepsilon}^{A,B} y_n\| + \|T_{\varepsilon}^{A,B} y_n - y_n\|)^2 + 2t\|z_t - y_n\|^2 \\
&\quad + 2t\langle x_0 - z_t, j(z_t - y_n)\rangle \\
&\leqslant (1-t)^2 (\|z_t - y_n\| + \|T_{\varepsilon}^{A,B} y_n - y_n\|)^2 + 2t\|z_t - y_n\|^2 \\
&\quad + 2t\langle x_0 - z_t, j(z_t - y_n)\rangle \\
&= (1+t^2)\|z_t - y_n\|^2 + (1 + t^2 - 2t)(2\|z_t - y_n\| \|T_{\varepsilon}^{A,B} y_n - y_n\| \\
&\quad + \|T_{\varepsilon}^{A,B} y_n - y_n\|^2) + 2t\langle x_0 - z_t, j(z_t - y_n)\rangle \\
&\leqslant (1+t^2)\|z_t - y_n\|^2 + (2\|z_t - y_n\| + \|T_{\varepsilon}^{A,B} y_n - y_n\|)\|T_{\varepsilon}^{A,B} y_n - y_n\| \\
&\quad + 2t\langle x_0 - z_t, j(z_t - y_n)\rangle \\
&\leqslant (1+t^2)\|z_t - y_n\|^2 + M\|T_{\varepsilon}^{A,B} y_n - y_n\| + 2t\langle x_0 - z_t, j(z_t - y_n)\rangle,
\end{aligned} \tag{3.19}$$

where $M > 0$ is a constant such that

$$M > \max\{\|z_t - y_n\|^2, 2\|z_t - y_n\| + \|T_{\varepsilon}^{A,B} y_n - y_n\|\}, \quad t \in (0,1),\ n \in \mathbb{N}.$$

It follows from (3.19) that

$$\langle x_0 - z_t, j(y_n - z_t)\rangle \leqslant \frac{M}{2} t + \frac{M}{2t} \|T_{\varepsilon}^{A,B} y_n - y_n\|.$$

Taking lim sup yields

$$\limsup_{n\to\infty}\langle x_0 - z_t, j(y_n - z_t)\rangle \leqslant \frac{M}{2}t.$$

Then, letting $t \to 0$ and noting that the duality map $j$ is norm-to-norm uniformly continuous on bounded sets, we get that $\limsup_{n\to\infty}\langle x_0 - z, j(y_n - z)\rangle \leqslant 0$. On the other hand, we see that

$$\|y_n - x_n\| = \beta_n\|x_n - x_{n-1}\| \to 0, \tag{3.20}$$

which together with

$$\langle x_0 - z, j(x_n - z)\rangle = \langle x_0 - z, j(x_n - z) - j(y_n - z)\rangle + \langle x_0 - z, j(y_n - z)\rangle \tag{3.21}$$

and the fact that $j$ is norm-to-norm uniformly continuous on bounded sets, implies that $\limsup_{n\to\infty}\langle x_0 - z, j(x_n - z)\rangle \leqslant 0$. Equivalently, we have $\limsup_{n\to\infty}\langle x_0 - z, j_q(x_n - z)\rangle \leqslant 0$. Again from (3.10), we get (since $\Gamma_n \leqslant \Gamma_{n-1}$ )

$$\begin{aligned}\Gamma_{n+1} &\leqslant (1-\alpha_n)\Gamma_n + (1-\alpha_n)\beta_n(\Gamma_n - \Gamma_{n-1})\\ &\quad + (1-\alpha_n)k_q\beta_n(\beta_n^{q-1}+1)\|x_n - x_{n-1}\|^q\\ &\quad + \alpha_n q\langle x_0 - z, j_q(x_{n+1} - z)\rangle + q(1-\alpha_n)M^*\|e_n\|\\ &\leqslant (1-\alpha_n)\Gamma_n + (1-\alpha_n)k_q\beta_n(\beta_n^{q-1}+1)\|x_n - x_{n-1}\|^q\\ &\quad + \alpha_n q\langle x_0 - z, j_q(x_{n+1} - (1-\alpha_n)e_n - z)\rangle + qM^*\|e_n\|.\end{aligned} \tag{3.22}$$

From (3.17) and (3.20), we get

$$\|x_n - T_{r_n}^{A,B}y_n\| \leqslant \|y_n - x_n\| + \|T_{r_n}^{A,B}y_n - y_n\| \to 0, \tag{3.23}$$

as $n \to \infty$. Also from Assumption 3.2 (i), (ii), (3.1), and (3.3), we get

$$\|x_{n+1} - T_{r_n}^{A,B}y_n\| \leqslant \alpha_n\|x_0 - T_{r_n}^{A,B}y_n\| + \|e_n\| \to 0, \tag{3.24}$$

as $n \to \infty$. Using (3.23) and (3.24), we get

$$\|x_{n+1} - x_n\| \leqslant \|x_n - T_{r_n}^{A,B}y_n\| + \|x_{n+1} - T_{r_n}^{A,B}y_n\| \to 0, \tag{3.25}$$

as $n \to \infty$. By (3.25) and the norm-to-norm uniform continuity of the duality mapping, we get

$$\limsup_{n\to\infty}\langle x_0 - z, j_q(x_{n+1} - z)\rangle \leqslant 0. \tag{3.26}$$

Let $w_n := x_{n+1} - (1-\alpha_n)e_n$. Then

$$\|w_n - x_{n+1}\| = (1-\alpha_n)\|e_n\| \to 0, \quad n \to \infty.$$

Using the norm-to-norm uniform continuity of duality mapping again, we get

$$\limsup_{n\to\infty}\langle x_0 - z, j_q(x_{n+1} - (1-\alpha_n)e_n - z)\rangle = \limsup_{n\to\infty}\langle x_0 - z, j_q(w_n - z)\rangle \leqslant 0. \tag{3.27}$$

Using Lemma 2.6 (ii) and (3.27) in (3.22), we get $x_n \to z$.

*Case 2*: Assume that there is no $N \in \mathbb{N}$ such that $\{\Gamma_n\}_{n=N}^{\infty}$ is decreasing. Let $\tau\colon \mathbb{N} \to \mathbb{N}$ be a mapping defined for all $n \geqslant N$ (for some $N$ large enough) by $\tau(n) := \max\{k \in \mathbb{N}\colon k \leqslant n,\ \Gamma_k \leqslant \Gamma_{k+1}\}$. In other words, $\tau(n)$ is the largest number $k$ in $\{1, 2, \ldots, n\}$ such that $\Gamma_k$ increases at $k = \tau(n)$. Observe that in view of Case 2, $\tau(n)$ is well-defined for all sufficiently large $n$. Also, it is easy to see that $\tau$ is a nondecreasing sequence such that $\tau(n) \to \infty$ as $n \to \infty$ and $0 \leqslant \Gamma_{\tau(n)} \leqslant \Gamma_{\tau(n)+1}$ for all $n \geqslant N$.

By ideas similar to (3.11) and (3.12) (noting that $\{x_n\}$ is bounded, $\beta_{\tau(n)}\|x_{\tau(n)} - x_{\tau(n)-1}\| \to 0$, $\alpha_{\tau(n)} \to 0$ and $\|e_{\tau(n)}\| \to 0$), we can show that $\lim_{n\to\infty}\|Ay_{\tau(n)} - Az\| = 0$ and

$$\lim_{n\to\infty}\varphi_q(\|(I - J^B_{r_{\tau(n)}})(I - r_{\tau(n)}A)y_{\tau(n)} - (I - J^B_{r_{\tau(n)}})(I - r_{\tau(n)})z\|) = 0,$$

which consequently shows that $\lim_{n\to\infty}\|T^{A,B}_{\varepsilon}y_{\tau(n)} - y_{\tau(n)}\| = 0$, by ideas in (3.18). Furthermore, as in Case 1, we can obtain $\limsup_{n\to\infty}\langle u - z, j(y_{\tau(n)} - z)\rangle \leqslant 0$, $\lim_{n\to\infty}\|y_{\tau(n)} - x_{\tau(n)}\| = 0$, $\lim_{n\to\infty}\|x_{\tau(n)+1} - x_{\tau(n)}\| = 0$ and $\limsup_{n\to\infty}\langle u - z,\ j_q(x_{\tau(n)} - z)\rangle \leqslant 0$. By exploiting the arguments when obtaining (3.25) and (3.27), we can show that

$$\limsup_{n\to\infty}\langle x_0 - z, j_q(x_{\tau(n)+1} - z)\rangle \leqslant 0 \tag{3.28}$$

and

$$\limsup_{n\to\infty}\langle x_0 - z, j_q(x_{\tau(n)+1} - (1-\alpha_{\tau(n)})e_{\tau(n)} - z)\rangle \leqslant 0. \tag{3.29}$$

From (3.10), we get

$$\begin{aligned}\Gamma_{\tau(n)+1} \leqslant{}& (1-\alpha_{\tau(n)})\Gamma_{\tau(n)} + (1-\alpha_{\tau(n)})\beta_{\tau(n)}(\Gamma_{\tau(n)} - \Gamma_{\tau(n)-1})\\ &+ (1-\alpha_{\tau(n)})k_q\beta_{\tau(n)}(\beta^{q-1}_{\tau(n)} + 1)\|x_{\tau(n)} - x_{\tau(n)-1}\|^q\\ &+ \alpha_{\tau(n)}q\langle x_0 - z, j_q(x_{\tau(n)+1} - (1-\alpha_{\tau(n)})e_{\tau(n)} - z)\rangle\\ &+ qM^*\|e_{\tau(n)}\|.\end{aligned} \tag{3.30}$$

This implies that (since $\Gamma_{\tau(n)} \leqslant \Gamma_{\tau(n)+1}$)

$$
\begin{aligned}
(3.31)\quad \alpha_{\tau(n)}\Gamma_{\tau(n)} &\leqslant \Gamma_{\tau(n)} - \Gamma_{\tau(n)+1} + (1-\alpha_{\tau(n)})\beta_{\tau(n)}(\Gamma_{\tau(n)} - \Gamma_{\tau(n)-1}) \\
&\quad + (1-\alpha_{\tau(n)})k_q\beta_{\tau(n)}(\beta_{\tau(n)}^{q-1}+1)\|x_{\tau(n)} - x_{\tau(n)-1}\|^q \\
&\quad + \alpha_{\tau(n)}q\langle x_0 - z, j_q(x_{\tau(n)+1} - (1-\alpha_{\tau(n)})e_{\tau(n)} - z)\rangle \\
&\quad + qM^*\|e_{\tau(n)}\| \\
&\leqslant \beta_{\tau(n)}(\Gamma_{\tau(n)} - \Gamma_{\tau(n)-1}) + k_q\beta_{\tau(n)}(\beta_{\tau(n)}^{q-1}+1)\|x_{\tau(n)} - x_{\tau(n)-1}\|^q \\
&\quad + \alpha_{\tau(n)}q\langle x_0 - z, j_q(x_{\tau(n)+1} - (1-\alpha_{\tau(n)})e_{\tau(n)} - z)\rangle \\
&\quad + qM^*\|e_{\tau(n)}\|.
\end{aligned}
$$

Using Lemma 2.1, we have

$$
\begin{aligned}
(3.32)\qquad \Gamma_{\tau(n)} - \Gamma_{\tau(n)-1} &< q\|x_{\tau(n)} - z\|^{q-1}(\|x_{\tau(n)} - z\| - \|x_{\tau(n)-1} - z\|) \\
&\leqslant q\|x_{\tau(n)} - z\|^{q-1}\|x_{\tau(n)} - x_{\tau(n)-1}\| \\
&\leqslant \|x_{\tau(n)} - x_{\tau(n)-1}\|M_2,
\end{aligned}
$$

where $M_2 := q\sup\limits_{n\geqslant 1}\|x_{\tau(n)} - z\|^{q-1}$. Using (3.32) in (3.31), we get

$$
\begin{aligned}
(3.33)\qquad \alpha_{\tau(n)}\Gamma_{\tau(n)} &\leqslant \beta_{\tau(n)}\|x_{\tau(n)} - x_{\tau(n)-1}\|M_2 \\
&\quad + k_q\beta_{\tau(n)}(\beta_{\tau(n)}^{q-1}+1)\|x_{\tau(n)} - x_{\tau(n)-1}\|^q \\
&\quad + \alpha_{\tau(n)}q\langle x_0 - z, j_q(x_{\tau(n)+1} - (1-\alpha_{\tau(n)})e_{\tau(n)} - z)\rangle \\
&\quad + qM^*\|e_{\tau(n)}\|.
\end{aligned}
$$

Since $\alpha_{\tau(n)} \in (0,1)$, we have

$$
\begin{aligned}
\Gamma_{\tau(n)} &\leqslant \frac{\beta_{\tau(n)}}{\alpha_{\tau(n)}}\|x_{\tau(n)} - x_{\tau(n)-1}\|M_2 + k_q\frac{\beta_{\tau(n)}}{\alpha_{\tau(n)}}(\beta_{\tau(n)}^{q-1}+1)\|x_{\tau(n)} - x_{\tau(n)-1}\|^q \\
&\quad + q\langle x_0 - z, j_q(x_{\tau(n)+1} - (1-\alpha_{\tau(n)})e_{\tau(n)} - z)\rangle + qM^*\frac{\|e_{\tau(n)}\|}{\alpha_{\tau(n)}}.
\end{aligned}
$$

Thus,

$$\limsup_{n\to\infty}\Gamma_{\tau(n)} \leqslant 0.$$

Hence, $\lim\limits_{n\to\infty}\Gamma_{\tau(n)} = 0$. It follows that $\lim\limits_{n\to\infty}\|x_{\tau(n)} - z\| = 0$. Subsequently, we get $\lim\limits_{n\to\infty}\|x_{\tau(n)+1} - z\| = 0$. This means $\lim\limits_{n\to\infty}\Gamma_{\tau(n)+1} = 0$. For all $n \geqslant N$, it is easy to see that $\Gamma_n \leqslant \Gamma_{\tau(n)+1}$. Therefore, we obtain for all sufficiently large $n$ that $0 \leqslant \Gamma_n \leqslant \Gamma_{\tau(n)+1}$ and this implies that $\lim\limits_{n\to\infty}\Gamma_n = 0$. Hence, $\{x_n\}$ converges strongly to $z$. □

## 4. Applications

**4.1. Application to signal recovery.** In this subsection, we give some applications of our results to inverse problems in signal recovery in real Hilbert spaces $H$. These inverse problems are formulated as the problem of minimizing the sum of two convex functions.

Let $f\colon H \to (-\infty, \infty]$ and $g\colon H \to \mathbb{R}$ be two proper lower semicontinuous convex functions such that $g$ is differentiable on $H$ with a $\frac{1}{L}$-Lipschitz continuous gradient for some $L \in (0, \infty)$. Let us consider the following minimization problem:

$$\min_{x\in H} f(x) + g(x). \tag{4.1}$$

We denote the set of solutions to (4.1) by $S$. It was established in Proposition 3.1 (iii)(b) of [18] that

$$x \in H \text{ solves (4.1) } \Leftrightarrow 0 \in \partial f(x) + \nabla g(x) \Leftrightarrow x = \operatorname{prox}_{\gamma f}(x - \gamma\nabla g(x)), \quad \gamma \in (0, \infty),$$

where

$$\operatorname{prox}_{\gamma f}(x) := \arg\min_{u\in H} \Big\{ f(u) + \frac{1}{2\lambda}\|u - x\|^2 \Big\}.$$

Combettes and Wajs in [18] proved the following strong convergence result for problem (4.1).

**Theorem 4.1.** *Suppose that $S \neq \emptyset$. Let $\{\gamma_n\}$ be a sequence in $(0, \infty)$ such that $0 < \inf\limits_{n\geqslant 1} \gamma_n \leqslant \sup\limits_{n\geqslant 1} \gamma_n < 2L$, let $\{\lambda_n\}$ be a sequence in $(0, 1]$ such that $\inf\limits_{n\geqslant 1} \lambda_n > 0$, and let $\{a_n\}$ and $\{b_n\}$ be sequences in $H$ such that $\sum\limits_{n=1}^{\infty} \|a_n\| < \infty$ and $\sum\limits_{n=1}^{\infty} \|b_n\| < \infty$. Fix $x_1 \in H$ and, for every $n \geqslant 1$, set*

$$x_{n+1} = x_n + \lambda_n(\operatorname{prox}_{\gamma_n f}(x_n - \gamma_n(\nabla g(x_n) + a_n)) + b_n - x_n).$$

*Then $\{x_n\}$ converges strongly to $x \in S$ if and only if $\liminf\limits_{n\to\infty} d_S(x_n) = 0$, where $d_S(x_n) := \inf\limits_{s\in S} \|x_n - s\|$. In particular, strong convergence occurs if $\operatorname{int} S \neq \emptyset$.*

It is known from the Baillon-Haddad theorem [3] that $\nabla g$ is inverse strongly monotone and $\partial f$ is a maximal monotone operator (see [35]). Therefore, problem (4.1) is a special case of problem (1.1) when $A := \nabla g$ and $B := \partial f$. Hence, our Theorem 3.1 can be applied to solve problem (4.1). Modifying Algorithm 3.1 and applying Theorem 3.1, we obtain the following result for solving problem (4.1).

**Theorem 4.2.** *Suppose that $S \neq \emptyset$. Let $\{\gamma_n\}$ be a sequence in $(0,\infty)$ such that $0 < \inf\limits_{n\geqslant 1} \gamma_n \leqslant \sup\limits_{n\geqslant 1} \gamma_n < 2L$, let $\{\alpha_n\}$ be a sequence in $(0,1)$ such that $\lim \alpha_n = 0$, $\sum\limits_{n=1}^{\infty} \alpha_n = \infty$, and let $\{a_n\}$ and $\{b_n\}$ be sequences in $H$ such that $\lim\limits_{n\to\infty} \|a_n\|/\alpha_n = 0$ and $\lim\limits_{n\to\infty} \|b_n\|/\alpha_n = 0$. Fix $x_0, x_1 \in H$ and, for every $n \geqslant 1$, set*

$$\begin{cases} y_n = x_n + \beta_n(x_n - x_{n-1}), \\ x_{n+1} = \alpha_n x_0 + (1-\alpha_n)(\operatorname{prox}_{\gamma_n f}(y_n - \gamma_n(\nabla g(y_n) + a_n)) + b_n), \quad n \geqslant 1, \end{cases} \tag{4.2}$$

*where $\beta \in [0,1)$ and $\beta_n$ is chosen such that $0 \leqslant \beta_n \leqslant \overline{\beta}_n$,*

$$\overline{\beta}_n = \begin{cases} \min\Bigl\{\beta, \dfrac{\varepsilon_n}{\|x_n - x_{n-1}\|}\Bigr\}, & x_n \neq x_{n-1}, \\ \beta, & \text{otherwise.} \end{cases}$$

*Then $\{x_n\}$ converges strongly to $z = P_S(x_0)$, where $P_S$ is called the metric projection of $H$ onto $S$, which is the unique point $P_S(x_0) \in S$ such that*

$$\|x_0 - P_S(x_0)\| \leqslant \|x_0 - y\| \quad \forall\, y \in S.$$

Remark 4.1. We remark here that Theorem 4.1 proved in [18] cannot be applied to solve problem (4.1) in the case where $\liminf\limits_{n\to\infty} d_S(x_n) \neq 0$. Our Theorem 4.2 (even with inertial extrapolation step) can be applied without any restriction on $S$, as long as $S \neq \emptyset$.

As a particular case of problem (4.1), we consider the standard linear data formation model in signal and image restoration, in which signal $z \in H_2$ is related to signal $\bar{x} \in H_1$ via the model

$$z = A_2\bar{x} + w,$$

where $A_2\colon H_1 \to H_2$ is a linear operator and $w \in H_2$ stands for an additive noise perturbation. The problem is described as

$$\min_{x\in H_1} h(A_1x) + \frac{1}{2}\|A_2x - z\|^2, \tag{4.3}$$

where

(i) $A_2\colon H_1 \to H_2$ is a nonzero bounded linear operator;
(ii) $A_1\colon H_1 \to H_3$ is a bijective bounded linear operator such that $A_1^{-1} = A_1^*$ ($A_1^*$ is the dual of $A_1$) and $H_3$ is a real Hilbert space;
(iii) $h\colon H_3 \to (-\infty,\infty]$ is a proper lower semicontinuous convex function.

The term $\frac{1}{2}\|A_2x - z\|^2$ is the so-called data fidelity term which attempts to reflect the contribution of the data formation model, while the term $h(A_1x)$ promotes prior knowledge about the original signal.

Taking $f(x) = h(A_1x)$, $g(x) = \frac{1}{2}\|A_2x - z\|^2$ and $L = 1/\|A_2\|^2$, we see that problem (4.3) reduces to (4.1). Using Lemma 2.8 of [18], we obtain that

$$\operatorname{prox}_f = A_1^* o \operatorname{prox}_h o A_1.$$

Furthermore, $\nabla g(x) = A_2^*(A_2x - z)$ and we have the following result regarding the accelerated inexact, relaxed proximal Landweber method for solving problem (4.3).

**Theorem 4.3.** *Suppose that* $S \neq \emptyset$. *Let* $\{\gamma_n\}$ *be a sequence in* $(0, \infty)$ *such that* $0 < \inf_{n \geqslant 1} \gamma_n \leqslant \sup_{n \geqslant 1} \gamma_n < 2/\|A_2\|^2$, *let* $\{\alpha_n\}$ *be a sequence in* $(0, 1)$ *such that* $\lim \alpha_n = 0$, $\sum_{n=1}^{\infty} \alpha_n = \infty$, *and let* $\{a_n\}$ *and* $\{b_n\}$ *be sequences in* $H_1$ *such that* $\lim_{n\to\infty} \|a_n\|/\alpha_n = 0$ *and* $\lim_{n\to\infty} \|b_n\|/\alpha_n = 0$. *Fix* $x_0, x_1 \in H_1$ *and, for every* $n \geqslant 1$, *set*

$$\begin{cases} y_n = x_n + \beta_n(x_n - x_{n-1}), \\ x_{n+1} = \alpha_n x_0 + (1 - \alpha_n)((A_1^* o \operatorname{prox}_{\gamma_n h} o A_1)(y_n - \gamma_n(A_2^*(A_2y_n - z) + a_n)) + b_n). \end{cases}$$

*Here* $\beta_n$ *is chosen such that* $0 \leqslant \beta_n \leqslant \overline{\beta}_n$, *where*

$$\overline{\beta}_n = \begin{cases} \min\Big\{\beta, \dfrac{\varepsilon_n}{\|x_n - x_{n-1}\|}\Big\}, & x_n \neq x_{n-1}, \\ \beta, & \text{otherwise}, \end{cases}$$

*and* $\beta \in [0, 1)$, $\varepsilon_n = o(\alpha_n)$. *Then* $\{x_n\}$ *converges strongly to* $z = P_S(x_0)$.

**4.2. Integro-differential equation.** We will give an application of (1.1) to solving nonlinear integro-differential equations involving the generalized $p$-Laplacian, which have been studied in [41]. Consider the nonlinear integro-differential equation

$$(4.4)\quad \begin{cases} \dfrac{\partial u(x,t)}{\partial t} - \operatorname{div}[(C(x,t) + |\nabla u|^2)^{(p-2)/2}\nabla u] + \varepsilon|u|^{r-2}u \\ \qquad + g(x, u, \nabla u) + a\dfrac{\partial}{\partial t}\displaystyle\int_\Omega u \,\mathrm{d}x = f(x,t), & (x,t) \in \Omega \times (0,T), \\ -\langle \vartheta, (C(x,t) + |\nabla u|^2)^{(p-2)/2}\nabla u\rangle \in \beta_x(u), & (x,t) \in \Gamma \times (0,T), \\ u(x,0) = u(x,T), & x \in \Omega, \end{cases}$$

where $\Omega$ is a bounded conical domain of the Euclidean space $\mathbb{R}^N$ ($N \geqslant 1$), $\Gamma$ is the boundary of $\Omega \in C^1$ and $\vartheta$ denotes the exterior normal derivative to $\Gamma$. The symbols $\langle\cdot,\cdot\rangle$ and $|\cdot|$ denote the Euclidean inner-product and the Euclidean norm in $\mathbb{R}^N$, respectively, $T$ is a positive constant,

$$\nabla u = \Big(\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \ldots, \frac{\partial u}{\partial x_N}\Big)$$

and $x = (x_1, x_2, \ldots, x_N) \in \Omega$. Furthermore $\beta_x$ is the subdifferential of $\varphi_x$, where $\varphi_x = \varphi(x,\cdot)\colon \mathbb{R} \to \mathbb{R}$ for $x \in \Gamma$, $a$ and $\varepsilon$ are nonnegative constants, $0 \leqslant C(x,t) \in V := L^p(0,T;W^{1,p}(\Omega))$, $f(x,t) \in W := L^{\max\{p,p'\}}(0,T;L^{\max\{p,p'\}}(\Omega))$ and $g\colon \Omega \times \mathbb{R}^{N+1} \to \mathbb{R}$ are given functions.

Just like in [41], we need the following assumptions to discuss (4.4).

**Assumption 4.1.** $p \in \mathbb{R}$ with $2N/(N+1) < p < \infty$, $\alpha \in (0,1]$ and $r \in \mathbb{R}$ satisfies $2N/(N+1) < r < \min\{p,p'\} < \infty$, where $1/p + 1/p' = 1$ and $1/r + 1/r' = 1$.

**Assumption 4.2.** Green's formula is available.

**Assumption 4.3.** For each $x \in \Gamma$, $\varphi_x = \varphi(x,\cdot)\colon \mathbb{R} \to \mathbb{R}$ is a proper, convex and lower-semicontinuous function and $\varphi_x(0) = 0$.

**Assumption 4.4.** $0 \in \beta_x(0)$ and for each $t \in \mathbb{R}$, the function $x \in \Gamma \to (I + \lambda\beta_x)^{-1}(t) \in \mathbb{R}$ is measurable for $\lambda > 0$.

**Assumption 4.5.** Suppose that $g\colon \Omega \times \mathbb{R}^{N+1} \to \mathbb{R}$ satisfies the following conditions:

(a) Carathéodory's conditions;

(b) Growth condition:

$$|g(x, r_1, \ldots, r_{N+1})|^{\max\{p,p'\}} \leqslant |h(x,t)|^p + b|r_1|^p,$$

where $(r_1, r_2, \ldots, r_{N+1}) \in \mathbb{R}^{N+1}$, $h(x,t) \in W$ and $b$ is a positive constant;

(c) Monotone condition: $g$ is monotone in the following sense:

$$(g(x, r_1, \ldots, r_{N+1}) - g(x, t_1, \ldots, t_{N+1})) \geqslant (r_1 - t_1)$$

for all $x \in \Omega$ and $(r_1, \ldots, r_{N+1}), (t_1, \ldots, t_{N+1}) \in \mathbb{R}^{N+1}$.

**Assumption 4.6.** Let $V^*$ denote the dual space of $V$. Then the norm in $V$, $\|\cdot\|_V$, is defined by

$$\|u(x,t)\|_V = \Big(\int_0^T \|u(x,t)\|^p_{W^{1,p}(\Omega)}\,\mathrm{d}t\Big)^{1/p}, \quad u(x,t) \in V.$$

**Definition 4.1** ([41]). Define an operator $K\colon V \to V^*$ by

$$\langle w, Ku\rangle = \int_0^T\int_\Omega \langle (C(x,t) + |\nabla u|^2)^{(p-2)/2}\nabla u, \nabla w\rangle \,\mathrm{d}x\,\mathrm{d}t + \varepsilon\int_0^T\int_\Omega |u|^{r-2}uw\,\mathrm{d}x\,\mathrm{d}t$$

for $u, w \in V$.

**Definition 4.2** ([41]). Define a function $\Phi\colon V \to \mathbb{R}$ by

$$\Phi(u) = \int_0^T\int_\Gamma \varphi_x(u|_\Gamma(x,t))\,\mathrm{d}\Gamma(x)\,\mathrm{d}t$$

for $u(x,t) \in V$.

**Definition 4.3** ([41]). Define $S\colon D(S) = \{u(x,t) \in V\colon \partial u/\partial t \in V^*,\ u(x,0) = u(x,T)\} \to V^*$ by

$$Su = \frac{\partial u}{\partial t} + a\frac{\partial}{\partial t}\int_\Omega u\,\mathrm{d}x.$$

**Lemma 4.1** ([41]). *Define a mapping $B\colon W \to 2^W$ as follows:*

$$D(B) = \{u \in W;\ \text{there exists an } f \in W \text{ such that } f \in Ku + \partial\Phi(u) + Su\},$$

*where $\partial\Phi\colon V \to V^*$ is the subdifferential of $\Phi$. For $u \in D(B)$, we set $Bu = \{f \in W;\ f \in Ku + \partial\Phi(u) + Su\}$. Then $B\colon W \to 2^W$ is $m$-accretive.*

**Lemma 4.2** ([41]). *Define*

$$A\colon D(A) = L^{\max\{p,p'\}}(0,T; W^{1,\max\{p,p'\}}(\Omega)) \subset W \to W$$

*by*

$$(Au)(x,t) = g(x,u,\nabla u) - f(x,t)$$

*for all $u(x,t) \in D(A)$ and $f(x,t)$ is the same as that in (4.4). Then $A\colon D(A) \subset W \to W$ is continuous and strongly accretive. If we further assume that $g(x, r_1, \dots, r_{N+1}) \equiv r_1$, then $A$ is $\alpha$-inverse strongly accretive of order $p$.*

**Lemma 4.3** ([41]). *For $f(x,t) \in W$, the integro-differential equation (4.4) has a unique solution $u(x,t) \in W$.*

**Lemma 4.4** ([41]). *If $\varepsilon \equiv 0$, $g(x, r_1, \dots, r_{N+1}) \equiv r_1$ and $f(x,t) \equiv k$, where $k$ is a constant, then $u(x,t) \equiv k$ is the unique solution of the integro-differential equation (4.4). Moreover, $\{u(x,t) \in W;\ u(x,t) \equiv k \text{ satisfies (4.4)}\} = (A+B)^{-1}(0)$.*

Let operators $B$ and $A$ be as in Lemma 4.1 and Lemma 4.2 respectively. Then we can apply our proposed Algorithm 3.1 to solve the nonlinear integro-differential equation (4.4) in the following theorem:

**Theorem 4.4.** *Suppose Assumptions* 4.1–4.6 *hold. Let* $\{\alpha_n\}$ *be a sequence in* $(0,1)$ *such that* $\lim \alpha_n = 0$, $\sum\limits_{n=1}^{\infty} \alpha_n = \infty$, *and let* $\{a_n\}$ *and* $\{b_n\}$ *be sequences in* $D$ *such that* $\lim\limits_{n\to\infty} \|a_n\|/\alpha_n = 0$ *and* $\lim\limits_{n\to\infty} \|b_n\|/\alpha_n = 0$. *Assume that*

$$0 < \liminf_{n\to\infty} r_n \leqslant \limsup_{n\to\infty} r_n < \Big(\frac{\alpha p}{k_p}\Big)^{1/(p-1)}.$$

*Given* $u_0(x,t), u_1(x,t) \in D$, *for every* $n \geqslant 1$, *compute*

$$\begin{cases} y_n(x,t) = u_n(x,t) + \beta_n[u_n(x,t) - u_{n-1}(x,t)], \\ u_{n+1}(x,t) = \alpha_n u_0(x,t) + (1-\alpha_n)(J_{r_n}^B(y_n(x,t) - r_n(Ay_n(x,t) + a_n)) + b_n), \end{cases}$$

*where* $\beta \in [0,1)$ *and* $\beta_n$ *is chosen such that* $0 \leqslant \beta_n \leqslant \overline{\beta}_n$

$$\overline{\beta}_n = \begin{cases} \min\Big\{\beta, \dfrac{\varepsilon_n}{\|u_n(x,t) - u_{n-1}(x,t)\|_D}\Big\}, & u_n(x,t) \neq u_{n-1}(x,t) \text{ a.e.} \\ \beta, & \text{otherwise,} \end{cases}$$

*where* $\varepsilon_n = o(\alpha_n)$. *Suppose in the nonlinear integro-differential equation* (4.4), $\varepsilon \equiv 0$, $g(x, r_1, \dots, r_{N+1}) \equiv r_1$ *and* $f(x,t) \equiv k$, *where* $k$ *is a constant. Then* $\{u_n(x,t)\}$ *converges strongly to the unique solution* $u(x,t)$ *of* (4.4), *where*

$$u(x,t) = Q_{(A+B)^{-1}(0)}(u_0(x,t)).$$

## 5. Numerical example

In this section, we give some numerical examples to the signal recovery in compressed sensing. We aim at providing a comparison between our Algorithm 3.1 with and without inertial terms which is the algorithm (1.3) of López et al. [26]. Compressed sensing can be modeled as the underdeterminated linear equation

$$y = Cx + \varepsilon, \tag{5.1}$$

where $x \in \mathbb{R}^N$ is a vector with $m$ nonzero components to be recovered, $y \in \mathbb{R}^M$ is the observed or measured data with noise $\varepsilon$, and $C\colon \mathbb{R}^N \to \mathbb{R}^M$ $(M < N)$ is a bounded linear operator. It is known that to solve (5.1) can be seen as solving the LASSO problem

$$\min_{x\in\mathbb{R}^N} \frac{1}{2}\|y - Cx\|_2^2 + \lambda\|x\|_1, \tag{5.2}$$

where $\lambda > 0$. Hence we can apply our method for solving (5.2). In this case, we set $A = \nabla f$ the gradient of $f$ where $f(x) = \frac{1}{2}\|y - Cx\|_2^2$ and $B = \partial g$ the subdifferential of $g$ where $g(x) = \lambda\|x\|_1$. It is well-known that $\nabla f(x) = C^t(Cx - y)$ and it is $1/\|C\|^2$-isa [9]. Moreover, $\partial g$ is maximal monotone [35].

In our experiment, the sparse vector $x \in \mathbb{R}^N$ is generated from the uniform distribution in the interval $[-2, 2]$ with $m$ nonzero elements. The matrix $C \in \mathbb{R}^{M\times N}$ is generated from a normal distribution with mean zero and variance one. The observation $y$ is generated by white Gaussian noise with signal-to-noise ratio SNR = 40. The restoration accuracy is measured by the error

$$E_n = \|x_n - x\|_2 < \varepsilon,$$

where $\varepsilon$ is a given tolerance and $x_n$ is an estimated signal of $x$.

In what follows, let $r_n = 0.5/\|C\|^2$, $\alpha_n = 10^{-2}/n$, $\varepsilon_n = 1/n^{1.1}$ and $\beta_n = \overline{\beta}_n$ with $\beta = 0.5$. The error sequences $\{a_n\}$ and $\{b_n\}$ are null sequences in $\mathbb{R}^N$. The initial points are given by $x_0 = \text{ones}([N, 1])$ and $x_1 = \text{zeros}([N, 1])$. We denote by "CPU" the CPU time and by "Iter" the number of iterations. The stopping criterion is given by $\varepsilon = 10^{-5}$. The numerical results are reported as follows:

| $m$-sparse signal | Algorithm 3.1 | $N = 512$, $M = 256$ | | $N = 1024$, $M = 512$ | |
|---|---|---|---|---|---|
| | | CPU | Iter | CPU | Iter |
| $m = 10$ | $\beta = 0$ | 3.4125 | 1020 | 13.6435 | 1523 |
| | $\beta = 0.5$ | 1.4217 | 623 | 6.3521 | 824 |
| $m = 20$ | $\beta = 0$ | 8.1457 | 1736 | 20.4127 | 1834 |
| | $\beta = 0.5$ | 4.3251 | 1136 | 10.1458 | 1214 |
| $m = 30$ | $\beta = 0$ | 17.6321 | 2412 | 47.2568 | 2835 |
| | $\beta = 0.5$ | 9.1025 | 1732 | 22.3215 | 1911 |
| $m = 40$ | $\beta = 0$ | 31.3258 | 3214 | 75.3968 | 3536 |
| | $\beta = 0.5$ | 17.5032 | 2313 | 43.8457 | 2712 |
| $m = 50$ | $\beta = 0$ | 47.3625 | 3982 | 117.6321 | 4498 |
| | $\beta = 0.5$ | 17.6512 | 2996 | 49.7121 | 3485 |

Table 1. Computational results for solving the LASSO problem.

From Table 1 we observe that iterations increase as $m$ increases and it takes time to recover the signal. Also, for a given tolerance, our algorithms can be used to solve the LASSO problem in compressed sensing as well. However, it was revealed that our Algorithm 3.1 with inertial extrapolation takes significantly smaller number of iterations and less CPU time compared to Algorithm 3.1 without inertial extrapolation.

We next discuss the optimal choice of the parameter $\beta$ on the convergence behavior of the proposed Algorithm 3.1. In this case, all assumptions are given as above with $m = 50$ and then numerical results which are averaged 10 times in terms of CPU and Iter are obtained as follows:

| | $N = 512$, $M = 256$ | | $N = 1024$, $M = 512$ | |
|---|---|---|---|---|
| | CPU | Iter | CPU | Iter |
| $\beta = 0$ | 0.095124 | 1825.2 | 0.356312 | 2566.7 |
| $\beta = 0.1$ | 0.080124 | 1651.8 | 0.321257 | 2013.5 |
| $\beta = 0.2$ | 0.077256 | 1456.2 | 0.301452 | 1947.4 |
| $\beta = 0.3$ | 0.074965 | 1372.7 | 0.282541 | 1825.3 |
| $\beta = 0.4$ | 0.071968 | 1235.5 | 0.263692 | 1636.4 |
| $\beta = 0.5$ | 0.069124 | 1138.7 | 0.245869 | 1457.2 |
| $\beta = 0.6$ | 0.067369 | 1059.5 | 0.223687 | 1236.2 |
| $\beta = 0.7$ | 0.065125 | 936.1 | 0.212574 | 1120.6 |
| $\beta = 0.8$ | 0.062358 | 847.2 | 0.201247 | 1021.5 |
| $\beta = 0.9$ | 0.059135 | 796.3 | 0.200147 | 978.5 |

Table 2. Computational results for choices of $\beta$.

From Table 2 it is observed that the choice of $\beta$ affects the number of iterations and the CPU time of our algorithm. To be more precise, Iter and CPU have a small number when the values of $\beta$ tend to 1 and the worst case occurs when $\beta = 0$, i.e., without the inertial term.

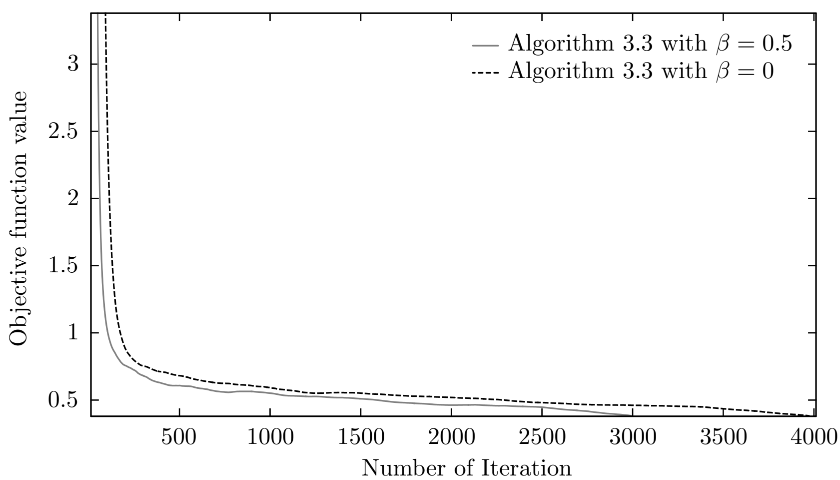


Figure 1. The objective function value versus the number of iterations in the case $N = 512$, $M = 256$.

We next provide some numerical experiments to illustrate the convergence behavior of all algorithms in comparison. We plot the number of iterations versus the objective function value and errors.

Figures 1 and 3 show the objective function values of Algorithm 3.1 with $\beta = 0$ and $\beta = 0.5$. From Table 1 we can see that for different choices of $m = 10, 20, 30, 40, 50$; $N = 512, M = 256$ and $N = 1024, M = 512$; the objective function values decrease faster when $\beta = 0.5$ than in the case when $\beta = 0$ (see the values of CPU and Iter in Table 1). Figures 2 and 4 compare the performance of the two versions of the algorithm in terms of errors.

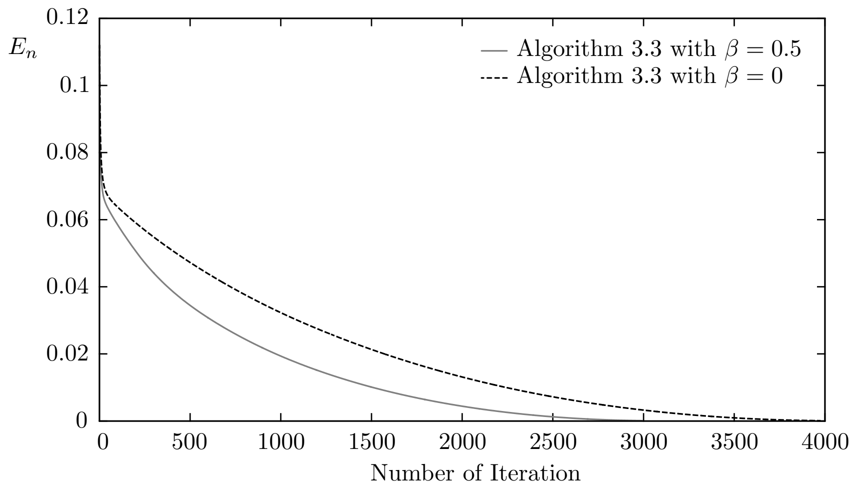


Figure 2. The errors versus the number of iterations in the case $N = 512$, $M = 256$.

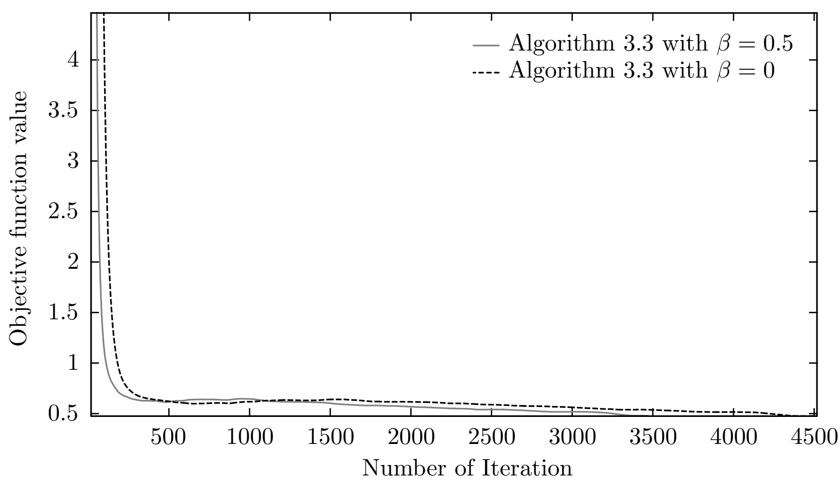


Figure 3. The objective function value versus the number of iterations in the case $N = 1024$, $M = 512$.

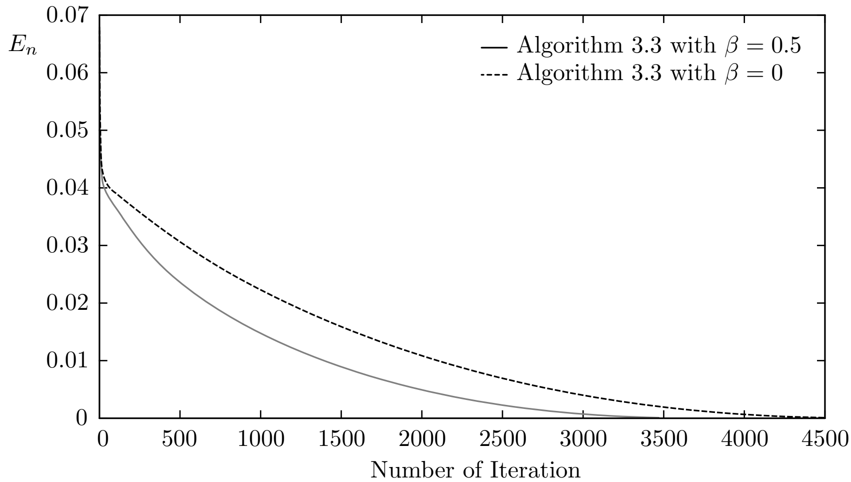


Figure 4. The errors versus the number of iterations in the case $N = 1024$, $M = 512$.

Next, we give another example in $L^2[0, 2\pi]$ which is an infinite dimensional space with the norm $\|x\| = (\int_0^{2\pi} x(t)^2\,\mathrm{d}t)^{1/2}$ and inner product $\langle x, y\rangle = \int_0^{2\pi} x(t)y(t)\,\mathrm{d}t$ for all $x, y \in L^2[0, 2\pi]$. Let $C = \{x \in L^2[0, 2\pi]\colon \int_0^{2\pi} \mathrm{e}^t x(t)\,\mathrm{d}t \leqslant 1\}$. Define $A\colon L^2[0, 2\pi] \to L^2[0, 2\pi]$ by $Ax(t) = x(t)/2$. In this case, we aim at minimizing the objective function $f + g$ where $g(x) = \frac{1}{2}\|Ax - b\|^2$, $b(t) = x(t)$ and $f(x) = \iota_C(x)$ is the indicator function of $C$. Take $a_n = 0 = b_n$. The iterations are terminated when $\|x_{n+1} - x_n\| < 10^{-5}$, where $10^{-5}$ is the tolerance. Then, using Algorithm 3.1, we obtain the following numerical results in Table 3:

| | $x_0 = 11t^2$, $x_1 = 7t^3$ | | $x_0 = t^2$, $x_1 = 2t^3 + t$ | |
|---|---|---|---|---|
| | CPU | Iter | CPU | Iter |
| $\beta = 0$ | 113.730573 | 199 | 139.126471 | 162 |
| $\beta = 0.2$ | 59.799427 | 160 | 85.415347 | 134 |
| $\beta = 0.4$ | 41.389116 | 119 | 65.401940 | 104 |
| $\beta = 0.6$ | 22.841785 | 69 | 38.475059 | 63 |
| $\beta = 0.8$ | 22.761412 | 67 | 36.607044 | 53 |
| | $x_0 = t^2$, $x_1 = 3\sin(t)$ | | $x_0 = t^3$, $x_1 = 2\mathrm{e}^t$ | |
| | CPU | Iter | CPU | Iter |
| $\beta = 0$ | 182.406959 | 130 | 658.656476 | 195 |
| $\beta = 0.2$ | 173.177190 | 118 | 571.352958 | 163 |
| $\beta = 0.4$ | 168.465893 | 96 | 442.208329 | 125 |
| $\beta = 0.6$ | 121.344645 | 61 | 271.401659 | 75 |
| $\beta = 0.8$ | 117.308429 | 55 | 232.989213 | 63 |

Table 3. Computational results in $L_2$-space.

From Table 3 we see that our proposed Algorithm 3.1 still works in this example and its convergence behavior becomes better when the value of $\beta$ approaches 1 as in Table 2.

R e m a r k 5.1. Our numerical examples on LASSO and constrained convex minimization problems show that our proposed Algorithm 3.1 can be implemented. In these numerical experiments, it is shown that Algorithm 3.1 outperforms its unaccelerated version. From Tables 2 and 3, it is reported that the number of iterations and the CPU time depend on the choice of the inertial factor $\beta$. In fact, Iter and CPU decrease as $\beta$ is close to 1.

A c k n o w l e d g m e n t s. We are sincerely grateful to the Editor and the anonymous reviewer for comments and suggestions which have improved the original manuscript greatly.

*References*

[1] *F. Alvarez, H. Attouch*: An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. Set-Valued Anal. *9* (2001), 3–11. zbl MR doi
[2] *H. Attouch, A. Cabot*: Convergence of a relaxed inertial forward-backward algorithm for structured monotone inclusions. Available at `https://hal.archives-ouvertes.fr/hal-01782016` (2018), Hal ID: 01782016, 35 pages.
[3] *J.-B. Baillon, G. Haddad*: Quelques propriétés des opérateurs angle-bornés et $n$-cycliquement monotones. Isr. J. Math. *26* (1977), 137–150. (In French.) zbl MR doi
[4] *A. Beck, M. Teboulle*: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. *2* (2009), 183–202. zbl MR doi
[5] *D. P. Bertsekas, J. N. Tsitsiklis*: Parallel and Distributed Computation: Numerical Methods. Athena Scientific, Belmont, 2014. zbl MR
[6] *R. I. Boţ, E. R. Csetnek*: An inertial alternating direction method of multipliers. Minimax Theory Appl. *1* (2016), 29–49. zbl MR
[7] *R. I. Boţ, E. R. Csetnek, C. Hendrich*: Inertial Douglas-Rachford splitting for monotone inclusion problems. Appl. Math. Comput. *256* (2015), 472–487. zbl MR doi
[8] *H. Brézis, P.-L. Lions*: Produits infinis de résolvantes. Isr. J. Math. *29* (1978), 329–345. (In French.) zbl MR doi
[9] *C. Byrne*: A unified treatment of some iterative algorithms in signal processing and image reconstruction. Inverse Probl. *20* (2004), 103–120. zbl MR doi
[10] *C. Chen, R. H. Chan, S. Ma, J. Yang*: Inertial proximal ADMM for linearly constrained separable convex optimization. SIAM J. Imaging Sci. *8* (2015), 2239–2267. zbl MR doi
[11] *G. H.-G. Chen, R. T. Rockafellar*: Convergence rates in forward-backward splitting. SIAM J. Optim. *7* (1997), 421–444. zbl MR doi
[12] *C. Chidume*: Geometric Properties of Banach Spaces and Nonlinear Iterations. Lecture Notes in Mathematics 1965, Springer, Berlin, 2009. zbl MR doi
[13] *P. Cholamjiak*: A generalized forward-backward splitting method for solving quasi inclusion problems in Banach spaces. Numer. Algorithms *71* (2016), 915–932. zbl MR doi
[14] *P. Cholamjiak, W. Cholamjiak, S. Suantai*: A modified regularization method for finding zeros of monotone operators in Hilbert spaces. J. Inequal. Appl. *2015* (2015), Article ID 220, 10 pages. zbl MR doi

[15] *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), Article ID 42, 17 pages. zbl MR doi
[16] *I. Cioranescu*: Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems. Mathematics and Its Applications 62, Kluwer Academic Publishers, Dordrecht, 1990. zbl MR doi
[17] *P. L. Combettes*: Iterative construction of the resolvent of a sum of maximal monotone operators. J. Convex Anal. *16* (2009), 727–748. zbl MR
[18] *P. L. Combettes, V. R. Wajs*: Signal recovery by proximal forward-backward splitting. Multiscale Model. Simul. *4* (2005), 1168–1200. zbl MR doi
[19] *Q. Dong, D. Jiang, P. Cholamjiak, Y. Shehu*: A strong convergence result involving an inertial forward-backward algorithm for monotone inclusions. J. Fixed Point Theory Appl. *19* (2017), 3097–3118. zbl MR doi
[20] *J. C. Dunn*: Convexity, monotonicity, and gradient processes in Hilbert space. J. Math. Anal. Appl. *53* (1976), 145–158. zbl MR doi
[21] *O. Güler*: On the convergence of the proximal point algorithm for convex minimization. SIAM J. Control Optim. *29* (1991), 403–419. zbl MR doi
[22] *B. Halpern*: Fixed points of nonexpanding maps. Bull. Am. Math. Soc. *73* (1967), 957–961. zbl MR doi
[23] *Y. Haugazeau*: Sur la minimisation des formes quadratiques avec contraintes. C. R. Acad. Sci., Paris, Sér. A *264* (1967), 322–324. (In French.) zbl MR
[24] *N. D. Kazarinoff*: Analytic Inequalities. Holt, Rinehart and Winston, New York, 1961. zbl MR
[25] *P.-L. Lions, B. Mercier*: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. *16* (1979), 964–979. zbl MR doi
[26] *G. López, V. Martín-Márquez, F. Wang, H.-K. Xu*: Forward-backward splitting methods for accretive operators in Banach spaces. Abstr. Appl. Anal. *2012* (2012), Article ID 109236, 25 pages. zbl MR doi
[27] *D. A. Lorenz, T. Pock*: An inertial forward-backward algorithm for monotone inclusions. J. Math. Imaging Vis. *51* (2015), 311–325. zbl MR doi
[28] *P.-E. Maingé*: Approximation methods for common fixed points of nonexpansive mappings in Hilbert spaces. J. Math. Anal. Appl. *325* (2007), 469–479. zbl MR doi
[29] *B. Martinet*: Régularisation d'inéquations variationnelles par approximations successives. Rev. Franç. Inform. Rech. Opér. *4* (1970), 154–158. (In French.) zbl MR
[30] *A. Moudafi, M. Oliny*: Convergence of a splitting inertial proximal method for monotone operators. J. Comput. Appl. Math. *155* (2003), 447–454. zbl MR doi
[31] *G. B. Passty*: Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. J. Math. Anal. Appl. *72* (1979), 383–390. zbl MR doi
[32] *J.-C. Pesquet, N. Pustelnik*: A parallel inertial proximal optimization method. Pac. J. Optim. *8* (2012), 273–306. zbl MR
[33] *B. T. Polyak*: Some methods of speeding up the convergence of iterative methods. U.S.S.R. Comput. Math. Math. Phys. *4* (1967), 1–17; translation from Zh. Vychisl. Mat. Mat. Fiz. *4* (1964), 791–803. zbl MR doi
[34] *S. Reich*: Strong convergence theorems for resolvents of accretive operators in Banach spaces. J. Math. Anal. Appl. *75* (1980), 287–292. zbl MR doi
[35] *R. T. Rockafellar*: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. *14* (1976), 877–898. zbl MR doi
[36] *Y. Shehu*: Iterative approximations for zeros of sum of accretive operators in Banach spaces. J. Funct. Spaces *2016* (2016), Article ID 5973468, 9 pages. zbl MR doi
[37] *Y. Shehu, G. Cai*: Strong convergence result of forward-backward splitting methods for accretive operators in Banach spaces with applications. Rev. R. Acad. Cienc. Exactas Fís. Nat., Ser. A Mat., RACSAM *112* (2018), 71–87. zbl MR doi

[38] *S. Suantai, N. Pholasa, P. Cholamjiak*: The modified inertial relaxed CQ algorithm for solving the split feasibility problems. J. Ind. Manag. Optim. *14* (2018), 1595–1615. doi
[39] *P. Sunthrayuth, P. Cholamjiak*: Iterative methods for solving quasi-variational inclusion and fixed point problem in $q$-uniformly smooth Banach spaces. Numer. Algorithms *78* (2018), 1019–1044. zbl MR doi
[40] *P. Tseng*: A modified forward-backward splitting method for maximal monotone mappings. SIAM J. Control Optim. *38* (2000), 431–446. zbl MR doi
[41] *L. Wei, R. P. Agarwal*: A new iterative algorithm for the sum of infinite $m$-accretive mappings and infinite $\mu_i$-inversely strongly accretive mappings and its applications to integro-differential systems. Fixed Point Theory Appl. *2016* (2016), Article ID 7, 22 pages. zbl MR doi
[42] *H.-K. Xu*: Inequalities in Banach spaces with applications. Nonlinear Anal., Theory Methods Appl. *16* (1991), 1127–1138. zbl MR doi

*Authors' addresses*: *Prasit Cholamjiak*, School of Science, University of Phayao, 19 Moo 2 Tambon Maeka Amphur Muang, Phayao 56000, Thailand, e-mail: `prasitch2008@yahoo.com`; *Yekini Shehu* (corresponding author), University of Nigeria, Department of Mathematics, Nsukka, Nigeria, e-mail: `yekini.shehu@unn.edu.ng`.