J. Appl. Numer. Optim. 2 (2020), No. 3, pp. 279-295
Available online at http://jano.biemdas.com
https://doi.org/10.23952/jano.2.2020.3.02

AN INERTIAL SPLITTING ALGORITHM FOR SOLVING INCLUSION PROBLEMS AND ITS APPLICATIONS TO COMPRESSED SENSING

YINGLIN LUO

Institute of Fundamental and Frontier Sciences,
University of Electronic Science and Technology of China, Chengdu 611731, China

Abstract. This paper investigates the problem of a variational inclusion problem via an inertial self-adaptive splitting algorithm. A solution theorem of the strong convergence is obtained in the framework of real Hilbert spaces. The effectiveness of our algorithm is validated via numerical examples.

Keywords. Forward-backward splitting algorithm; Inertial algorithm; Monotone operator; Inclusion problem; Signal processing.

1. Introduction

Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$. The purpose of this paper is to study the following known inclusion problem: find $x^* \in H$ such that

$$0 \in (A+B)x^*, \tag{1.1}$$

where $A : H \to H$ is a single-valued operator and $B : H \to 2^H$ is a set-valued operator. The solution set of (1.1) is denoted by Ω in this paper. This problem stands at the core of many mathematical problems, such as, convex programming problems, inverse problems, split feasibility problems, and minimization problems, see, e.g., [1, 7, 8, 14, 17, 24, 25]. Due to its wide and efficient applications in machine learning, signal processing, statistical regression, image restoration, etc., the inclusion problem has received much attention from many mathematicians and engineerings; see, e.g., [9, 15, 18, 23, 26]. For solving inclusion problem (1.1), many authors developed various iterative methods. One of the most popular methods is the forward-backward splitting method, which was proposed by Lions and Mercier, see [13]. The iterative scheme of the forward-backward splitting method reads as follows:

$$x_{n+1} = (I+\lambda_n B)^{-1}(I-\lambda_n A)x_n,$$

where $\lambda_n > 0$. One sees that this algorithm involves only with A as the forward step and B as the backward step, but nor the sum of A and B. It is known that if $A = \nabla f$ and $B = \partial g$, then the forward-backward algorithm reduces to the proximal gradient algorithm:

$$x_{n+1} = (I+\lambda_n \partial g)^{-1}(I-\lambda_n \nabla f)x_n,$$

*Corresponding author.
E-mail address: luoyinglink@163.com.
Received March 9, 2020; Accepted September 14, 2020.

where f and g are two proper lower semicontinuous convex functions. The proximal gradient algorithm converges weakly if ∇f is Lipschitz continuous. In [4], Bello and Nghia established some complexity results of the proximal gradient algorithm without using the Lipschitz continuity assumptions. In recent years, many authors improved this method in various ways, see, e.g., [2, 5, 6, 13, 16, 19, 20]. In [27], Tseng introduced a modified forward-backward splitting method, which is known as the Tseng's splitting algorithm. Since the Tseng's splitting algorithm converges in real Hilbert spaces in weak topology, some authors anticipated to obtain a modification which converges in norm and can be easily implemented. In [22], Shehu, Iyiola and Ogbuisi introduced a Halpern-type algorithm with both inertial terms and errors for approximating fixed points of a nonexpansive mapping. In [28], Wang and Wang proposed a modification of the forward-backward algorithm. They established two strong convergence theorems under some general conditions. In [11], based on viscosity ideas, Aviv and Thong proposed the following modification of the forward-backward splitting method with a new step size rule in real Hilbert spaces, see Algorithm 1.

Algorithm 1

Initialization: Give $\lambda_0 > 0$, $\mu \in (0,1)$. Let $x_0 \in H$ be arbitrary.
Iterative Steps: Given the current iterate x_n, calculate the next iterate as follows:
Step 1. Compute

$$y_n = (I + \lambda_n B)^{-1}(I - \lambda_n A)x_n.$$

If $x_n = y_n$, stop and y_n is a solution of inclusion problem (1.1). Otherwise,
Step 2. Compute

$$z_n = y_n - \lambda_n(Ay_n - Ax_n).$$

and

$$x_{n+1} = (1 - \alpha_n - \gamma_n)x_n + \gamma_n z_n.$$

Update

$$\lambda_{n+1} = \begin{cases} \min\left\{ \dfrac{\mu\|x_n - y_n\|}{\|Ax_n - Ay_n\|}, \lambda_n \right\}, & Ax_n - Ay_n \neq 0; \\ \lambda_n, & otherwise. \end{cases} \tag{1.2}$$

Set $n \leftarrow n+1$ and go to Step 1.

In this paper, motivated by above research results, we establish solution theorems of the strong convergence for variational inclusion problem (1.1) under some mild assumptions on iterative parameters. Our proposed method is a combination of self-adaptive inertial extrapolation steps, the Halpern's method and the Tseng's extragradient method. Moreover, we give some numerical experiments to support our convergence results. We also give a comparison between our method and some existing methods to demonstrate the efficiency of the our algorithm.

This paper is organized as follows: In Section 2, we recall some definitions and preliminary lemmas. In Section 3, we propose an inertial self-adaptive algorithm for solving inclusion problem (1.1) and give strong convergence analysis of our method. In Section 4, some numerical comparison examples are given. Finally, we give conclusions to summarize our paper in Section 5.

2. PRELIMINARIES

In this section, we give some known lemmas and definitions which will be used in our paper.

Let H be a real Hilbert space. The symbols $x_n \to x$ and $x_n \rightharpoonup x$ are employed to denote the strong and weak convergence of the sequence $\{x_n\}$, respectively. Let T be an operator defined on H. The set of fixed points of T is denoted by $\mathrm{Fix}(T) := \{x \in H | Tx = x\}$.

It is well known that, for each $x, y, z \in H$ and $\alpha, \beta, \gamma \in [0,1]$ with $\alpha + \beta + \gamma = 1$, the following properties hold.

(1) $\|x+y\|^2 - \|x\|^2 \leq 2\langle y, x+y\rangle$;
(2) $\|\alpha x + (1-\alpha)y\|^2 + \alpha(1-\alpha)\|x-y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2$;
(3) $\|\alpha x + \beta y + \gamma z\|^2 + \alpha\beta\|x-y\|^2 + \alpha\gamma\|x-z\|^2 + \beta\gamma\|y-z\|^2 = \alpha\|x\|^2 + \beta\|y\|^2 + \gamma\|z\|^2$.

Definition 2.1. Let $T : H \to H$ be an operator. For all $x, y \in H$,

(1) T is said to be L-Lipschitz continuous with $L > 0$ if
$$\|Tx - Ty\| \leq L\|x-y\|.$$
If $L = 1$, one says that T is nonexpansive.

(2) T is said to be monotone if
$$\langle Tx - Ty, x-y\rangle \geq 0.$$

(3) T is said to be firmly monotone if
$$\|Tx - Ty\|^2 \leq \langle Tx - Ty, x-y\rangle,$$
or equivalently,
$$\|Tx - Ty\|^2 + \|(I-T)x - (I-T)y\|^2 \leq \|x-y\|^2.$$

A set-valued operator $B : H \to 2^H$ is said to be monotone if, for all $x, y \in H$, $u \in Bx$ and $v \in By$, the inequality $\langle x-y, u-v\rangle \geq 0$ holds. If B is monotone and there exists no monotone operator $T : H \to 2^H$ such that Graph(T) properly contains Graph(B), i.e., for each $(x,u) \in H \times H$,
$$\langle x-y, u-v\rangle \geq 0, \forall (y,v) \in \mathrm{Graph}(B),$$
implies that $u \in Bx$. Then one says that B is maximal monotone. Let $A : H \to 2^H$ be a set-valued operator. $D(A) := \{x \in H : Ax \neq \emptyset\}$ denotes the domain of A and $R(A) := \bigcup\{Az : z \in D(A)\}$ denotes the range of A, respectively. One supposes that A is a maximal monotone operator. For each $r > 0$, one defines the resolvent $J_r^A : R(I+rA) \to D(A)$ of A by $J_r^A := (I+rA)^{-1}$. The zero set of A is denoted by $A^{-1}(0)$, that is, $A^{-1}(0) := \{x \in D(A) : 0 \in Ax\}$. It is known that if A is a maximal monotone set-valued operator, then J_r^A is single-valued and firmly nonexpansive. Moreover, $\mathrm{Fix}(J_r^A) := A^{-1}(0)$ for each $x \in R(I+rA)$, where $\mathrm{Fix}(J_r^A) := \{x \in R(I+rA) : J_r^A(x) = x\}$. In addition, if $A : H \to H$ is a Lipschitz continuous monotone mapping, which is maximal monotone, and $B : H \to 2^H$ is a maximal monotone operator, then $A+B$ is a maximal monotone operator, see [3] and the references therein.

Lemma 2.1. [11] *Let H be a real Hilbert space. Let $A : H \to H$ be a mapping and let $B : H \to 2^H$ be a maximal monotone operator. Then* $\mathrm{Fix}((I+rB)^{-1}(I-rA)) = (A+B)^{-1}(0)$ $(\forall r > 0)$.

Lemma 2.2. [21] *Let $\{a_n\}$ be a sequence of nonnegative real numbers and let $\{b_n\}$ be a sequence of real numbers. Let $\{\alpha_n\}$ be a sequence of real numbers in $(0,1)$ such that $\sum_{n=1}^{\infty}\alpha_n=\infty$. Assume that $a_{n+1}\le\alpha_n b_n+(1-\alpha_n)a_n, \forall n\ge 1$. If for every subsequence $\{a_{n_k}\}$ of $\{a_n\}$ satisfying* $\liminf_{k\to\infty}(a_{n_k+1}-a_{n_k})\ge 0$, *implies* $\limsup_{k\to\infty} b_{n_k}\le 0$, *then* $\lim_{n\to\infty}a_n=0$.

3. MAIN RESULTS

In this section, we give a precise statement of our methods and analyze the strong convergence of the considered methods. We give some assumptions that will be used next.

Assumption 3.1. (1) The solution set of inclusion problem (1.1) is nonempty, that is, $\Omega:=(A+B)^{-1}(0)\neq\emptyset$.

(2) The mapping $A:H\to H$ is L-Lipschitz continuous and monotone and the mapping $B:H\to H$ is a maximal monotone operator.

Assumption 3.2. Suppose that $\{\alpha_n\},\{\beta_n\}$ and $\{\gamma_n\}$ are three real sequences in $(0,1)$ which satisfy the following conditions:

(1) $\sum_{n=1}^{\infty}\alpha_n=\infty$ and $\lim_{n\to\infty}\alpha_n=0$;
(2) $\delta_n=o(\alpha_n)$, i.e., $\lim_{n\to\infty}\frac{\delta_n}{\alpha_n}=0$;
(3) $\alpha_n+\beta_n+\gamma_n=1$ and $\liminf_{n\to\infty}\gamma_n>0$.

Assumption 3.3. Let $f:H\to H$ be a contraction with constant $\rho\in[0,1)$.

Algorithm 2

Initialization: Give $\theta,\ \ \mu\in(0,1)$ and $\lambda_0>0$. Let $x_0,\ \ x_1\in H$ be arbitrary.
Iterative Steps: Given the current iterate x_n, calculate the next iterate as follows:
Step 1. Compute θ_n, such that $0\le\theta_n\le\theta_n^*$, where

$$\theta_n^*=\begin{cases}\min\left\{\theta,\dfrac{\|\delta_n\|}{\|x_n-x_{n-1}\|}\right\}, & x_n\neq x_{n-1};\\ \theta, & otherwise.\end{cases}\tag{3.1}$$

Step 2. Compute

$$w_n=x_n+\theta_n(x_n-x_{n-1}),$$

and

$$y_n=(I+\lambda_n B)^{-1}(I-\lambda_n A)w_n.$$

If $w_n=y_n$, then stop and y_n is a solution of inclusion problem (1.1). Otherwise,
Step 3. Compute

$$z_n=y_n-\lambda_n(Ay_n-Aw_n).$$

and

$$x_{n+1}=\alpha_n f(x_n)+\beta_n w_n+\gamma_n z_n.$$

Update

$$\lambda_{n+1}=\begin{cases}\min\left\{\dfrac{\mu\|w_n-y_n\|}{\|Aw_n-Ay_n\|},\lambda_n\right\}, & Aw_n\neq Ay_n;\\ \lambda_n, & otherwise.\end{cases}\tag{3.2}$$

Set $n\leftarrow n+1$ and go to Step 1.

Remark 3.1. Form Assumption 3.2 and (3.1), we see that $\lim_{n\to\infty}\theta_n\|x_n-x_{n-1}\|=0$ and $\lim_{n\to\infty}\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\|=0$.

Lemma 3.1. *The sequence $\{\lambda_n\}$ generated by (3.2) is nonincreasing and*

$$\lim_{n\to\infty}\lambda_n=\lambda\geq\min\left\{\lambda_0,\frac{\mu}{L}\right\}.$$

Proof. From (3.2), it is obvious that $\{\lambda_n\}$ is nonincreasing. On the other hand, if $Aw_n-Ay_n\neq 0$, then the L-lipschitz continuity of A yields

$$\frac{\mu\|w_n-y_n\|}{\|Aw_n-Ay_n\|}\geq\frac{\mu}{L}.$$

Therefore, $\{\lambda_n\}$ has the lower bound $\min\left\{\lambda_0,\frac{\mu}{L}\right\}$. This completes the proof. □

Lemma 3.2. *Suppose that Assumption 3.1, Assumption 3.2 and Assumption 3.3 hold. Let $\{z_n\}$ be any sequence generated by Algorithm (2). Then*

$$\|z_n-p\|^2\leq\|w_n-p\|^2-(1-\mu^2\frac{\lambda_n^2}{\lambda_{n+1}^2})\|w_n-y_n\|^2,\ \ \forall p\in\Omega. \tag{3.3}$$

Proof. First, one claims

$$\|Aw_n-Ay_n\|\leq\frac{\mu}{\lambda_{n+1}}\|w_n-y_n\|. \tag{3.4}$$

If $Aw_n=Ay_n$, then (3.4) holds immediately. If $Aw_n\neq Ay_n$, then

$$\lambda_{n+1}=\min\left\{\frac{\mu\|w_n-y_n\|}{\|Aw_n-Ay_n\|},\lambda_n\right\}\leq\frac{\mu\|w_n-y_n\|}{\|Aw_n-Ay_n\|}.$$

This implies that (3.4) holds.

Next, one claims that (3.3) holds.

$$\begin{aligned}\|z_n-p\|^2&=\lambda_n^2\|Ay_n-Aw_n\|^2+\|y_n-p\|^2-2\lambda_n\langle y_n-p,Ay_n-Aw_n\rangle\\&=\|w_n-p\|^2+\|w_n-y_n\|^2+2\langle w_n-p,y_n-w_n\rangle+\lambda_n^2\|Ay_n-Aw_n\|^2\\&\quad-2\lambda_n\langle y_n-p,Ay_n-Aw_n\rangle\\&=\|w_n-p\|^2-\|w_n-y_n\|^2-2\langle y_n-p,w_n-y_n-\lambda_n(Aw_n-Ay_n)\rangle\\&\quad+\lambda_n^2\|Ay_n-Aw_n\|^2.\end{aligned} \tag{3.5}$$

It follows from (3.4) and (3.5) that

$$\|z_n-p\|^2\leq\|w_n-p\|^2-(1-\mu^2\frac{\lambda_n^2}{\lambda_{n+1}^2})\|w_n-y_n\|^2-2\langle w_n-y_n-\lambda_n(Aw_n-Ay_n),y_n-p\rangle. \tag{3.6}$$

Since $y_n=(I+\lambda_nB)^{-1}(I-\lambda_nA)w_n$, one has $(I-\lambda_nA)w_n\in(I+\lambda_nB)y_n$. Since B is a maximal monotone mapping, there exists $v_n\in By_n$ such that

$$(I-\lambda_nA)w_n=y_n+\lambda_nv_n.$$

This is equivalent to

$$v_n=\frac{1}{\lambda_n}(w_n-y_n-\lambda_nAw_n). \tag{3.7}$$

On the other hand, we see that $0 \in (A+B)p$ and $Ay_n + v_n \in (A+B)y_n$. From the fact that $A+B$ is maximal monotone, one arrives at

$$\langle Ay_n + v_n, y_n - p\rangle \geq 0. \tag{3.8}$$

Combining (3.7) and (3.8), one has

$$\frac{1}{\lambda_n}\langle w_n - y_n - \lambda_n(Aw_n - Ay_n), y_n - p\rangle \geq 0.$$

Hence,

$$\langle w_n - y_n - \lambda_n(Aw_n - Ay_n), y_n - p\rangle \geq 0. \tag{3.9}$$

It follows from (3.6) and (3.9) that

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - (1 - \mu^2\frac{\lambda_n^2}{\lambda_{n+1}^2})\|w_n - y_n\|^2.$$

This completes the proof. □

Lemma 3.3. *Assume that $\{x_n\}, \{y_n\}$ and $\{w_n\}$ are three sequences generated by Algorithm (2). If $\|x_n - w_n\| \to 0$ and $\|y_n - w_n\| \to 0$, then $\{x_n\}$ converges weakly to $z \in \Omega$.*

Proof. Suppose that $(\eta, \zeta) \in \mathrm{Graph}(A+B)$. This implies that $\zeta - A\eta \in B\eta$. On the other hand, there exists a subsequence $\{w_{n_k}\}$ of $\{w_n\}$ such that $y_{n_k} = (I + \lambda_{n_k}B)^{-1}(I - \lambda_{n_k}A)w_{n_k}$. Then one has

$$(I - \lambda_{n_k}A)w_{n_k} \in (I + \lambda_{n_k}B)y_{n_k},$$

that is,

$$\frac{1}{\lambda_{n_k}}(w_{n_k} - y_{n_k} - \lambda_{n_k}Aw_{n_k}) \in By_{n_k}.$$

Since B is maximal monotone, one concludes

$$\langle \eta - y_{n_k}, \zeta - A\eta - \frac{1}{\lambda_{n_k}}(w_{n_k} - y_{n_k} - \lambda_{n_k}Aw_{n_k})\rangle \geq 0. \tag{3.10}$$

It follows from the maximal monotone of A and (3.10) that

$$\begin{aligned}
\langle \eta - y_{n_k}, \zeta\rangle &\geq \langle \eta - y_{n_k}, A\eta + \frac{1}{\lambda_{n_k}}(w_{n_k} - y_{n_k} - \lambda_{n_k}Aw_{n_k})\rangle \\
&= \langle \eta - y_{n_k}, A\eta - Ay_{n_k}\rangle + \langle \eta - y_{n_k}, Ay_{n_k} - Aw_{n_k}\rangle \\
&\quad + \langle \eta - y_{n_k}, \frac{1}{\lambda_{n_k}}(w_{n_k} - y_{n_k})\rangle \\
&\geq \langle \eta - y_{n_k}, Ay_{n_k} - Aw_{n_k}\rangle + \langle \eta - y_{n_k}, \frac{1}{\lambda_{n_k}}(w_{n_k} - y_{n_k})\rangle.
\end{aligned} \tag{3.11}$$

Since $\|y_n - w_n\| \to 0$ and A is Lipschitz continuous, one has $\lim_{k\to\infty}\|Ay_{n_k} - Aw_{n_k}\| = 0$. Letting $k \to \infty$ in (3.11), one obtains from Lemma 3.1 that

$$\langle \eta - z, \zeta\rangle = \lim_{k\to\infty}\langle \eta - y_{n_k}, \zeta\rangle \geq 0.$$

By the maximal monotonicity of $(A+B)$, one asserts that $0 \in (A+B)z$, that is, $z \in \Omega$. This completes the proof. □

Theorem 3.1. *Suppose that Assumption 3.1, Assumption 3.2 and Assumption 3.3 hold. Let $\{x_n\}$ be a sequence generated by Algorithm 2. Then $\{x_n\}$ converges strongly to $p = P_\Omega \circ f(p)$.*

Proof. The proof is split into three steps.
Step 1. One claims that $\{x_n\}$ is bounded. In the view of Lemma 3.1, one sees that

$$\lim_{n\to\infty}(1-\mu^2\frac{\lambda_n^2}{\lambda_{n+1}^2}) = 1-\mu^2 > 0.$$

Then, there exists $n_0 \in N$ such that

$$1-\mu^2\frac{\lambda_n^2}{\lambda_{n+1}^2} > 0, \forall n \geq n_0.$$

Combining this with (3.3), one finds

$$\|z_n - p\| \leq \|w_n - p\|, \forall n \geq n_0. \tag{3.12}$$

Since $w_n = x_n + \theta_n(x_n - x_{n-1})$, one concludes

$$\|w_n - p\| \leq \alpha_n \frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| + \|x_n - p\|. \tag{3.13}$$

From Remark 3.1, one sees that $\lim_{n\to\infty} \dfrac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| = 0$. Thus, there exists a constant $M > 0$ such that

$$\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| \leq M, \ \ \forall n \geq 1.$$

Combining this with (3.13), one arrives at

$$\|z_n - p\| \leq \|w_n - p\| \leq \|x_n - p\| + \alpha_n M, \ \ \forall n \geq n_0. \tag{3.14}$$

Hence, for all $n \geq n_0$,

$$\begin{aligned}
\|x_{n+1} - p\| &\leq \beta_n\|w_n - p\| + \alpha_n\|p - f(x_n)\| + \gamma_n\|z_n - p\| \\
&\leq \alpha_n(\|f(x_n) - f(p)\| + \|p - f(p)\|) + (1-\alpha_n)\|x_n - p\| + \alpha_n(1-\alpha_n)M \\
&\leq \alpha_n\rho\|x_n - p\| + \alpha_n\|f(p) - p\| + (1-\alpha_n)\|x_n - p\| + \alpha_n M \\
&\leq (1-\alpha_n(1-\rho))\|x_n - p\| + \alpha_n(1-\rho)\frac{\|f(p)-p\| + M}{1-\rho} \\
&\leq \max\{\frac{\|f(p)-p\|+M}{1-\rho}, \|x_n - p\|\} \\
&\ \ \vdots \\
&\leq \max\{\frac{\|f(p)-p\|+M}{1-\rho}, \|x_0 - p\|\}.
\end{aligned}$$

This implies that $\{x_n\}$ is bounded. From (3.14), one sees that $\{w_n\}$ and $\{z_n\}$ are also bounded.
Step 2. One claims that, for all $n \geq n_0$,

$$\begin{aligned}
\|x_{n+1} - p\|^2 \leq &(1-\alpha_n(1-\rho^2))\|x_n - p\|^2 + \frac{2}{(1-\rho^2)}\langle f(p) - p, x_{n+1} - p\rangle \\
&+ \alpha_n(1-\rho^2)(\frac{2(1-\alpha_n)}{1-\rho^2}\cdot\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\|\|w_n - p\|).
\end{aligned} \tag{3.15}$$

By Lemma 3.2, one obtains

$$\begin{aligned}\|x_{n+1}-p\|^2 &\le \gamma_n\|z_n-p\|^2+\beta_n\|w_n-p\|^2+\alpha_n\|f(x_n)-p\|^2+\gamma_n\|w_n-p\|^2\\ &\le \beta_n\|w_n-p\|^2+\alpha_n(\rho\|x_n-p\|+\|f(p)-p\|)^2+\gamma_n\|w_n-p\|^2\\ &\quad -\gamma_n(1-\mu^2\frac{\lambda_n^2}{\lambda_{n+1}^2})\|w_n-y_n\|^2\\ &\le \alpha_n\rho^2\|x_n-p\|^2+2\alpha_n\rho\|x_n-p\|\|f(p)-p\|+\alpha_n\|p-f(p)\|^2\\ &\quad +(1-\alpha_n)\|w_n-p\|^2-\gamma_n(1-\mu^2\frac{\lambda_n^2}{\lambda_{n+1}^2})\|w_n-y_n\|^2.\end{aligned} \tag{3.16}$$

Since $w_n=x_n+\theta_n(x_n-x_{n-1})$, one asserts

$$\begin{aligned}\|w_n-p\|^2 &\le \|x_n-p\|^2+2\langle\theta_n(x_n-x_{n-1}),w_n-p\rangle\\ &\le \|x_n-p\|^2+2\theta_n\|x_n-x_{n-1}\|\|w_n-p\|.\end{aligned} \tag{3.17}$$

Combining (3.16) and (3.17) yields that

$$\begin{aligned}\|x_{n+1}-p\|^2 &\le (1-\alpha_n+\alpha_n\rho^2)\|x_n-p\|^2+2(1-\alpha_n)\theta_n\|x_n-x_{n-1}\|\|w_n-p\|\\ &\quad +2\alpha_n\rho\|x_n-p\|\|f(p)-p\|+\alpha_n\|f(p)-p\|^2-\gamma_n(1-\mu^2\frac{\lambda_n^2}{\lambda_{n+1}^2})\|w_n-y_n\|^2.\end{aligned}$$

This implies that

$$\begin{aligned}\gamma_n(1-\mu^2\frac{\lambda_n^2}{\lambda_{n+1}^2})\|w_n-y_n\|^2 &\le \|x_n-p\|^2-\|x_{n+1}-p\|^2+\alpha_n(2\rho\|x_n-p\|\|f(p)-p\|\\ &\quad +2(1-\alpha_n)\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\|\|w_n-p\|+\|f(p)-p\|^2).\end{aligned} \tag{3.18}$$

On the other hand, for all $n\ge n_0$, it follows from (3.12) and (3.17) that

$$\begin{aligned}\|x_{n+1}-p\|^2 &\le \|\alpha_n(f(x_n)-f(p))+\beta_n(w_n-p)+\gamma_n(z_n-p)\|^2+2\alpha_n\langle f(p)-p,x_{n+1}-p\rangle\\ &\le \alpha_n\rho^2\|x_n-p\|^2+\beta_n\|w_n-p\|^2+\gamma_n\|z_n-p\|^2+2\alpha_n\langle f(p)-p,x_{n+1}-p\rangle\\ &\le \alpha_n\rho^2\|x_n-p\|^2+(1-\alpha_n)\|w_n-p\|^2+2\alpha_n\langle f(p)-p,x_{n+1}-p\rangle\\ &\le (1-\alpha_n(1-\rho^2))\|x_n-p\|^2+\alpha_n(1-\rho^2)(\frac{2(1-\alpha_n)}{1-\rho^2}\cdot\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\|\|w_n-p\|\\ &\quad +\frac{2}{(1-\rho^2)}\langle f(p)-p,x_{n+1}-p\rangle).\end{aligned} \tag{3.19}$$

Step 3. One claims that $\{\|x_n-p\|^2\}$ converges to zero. We see from Remark 3.1 that

$$\lim_{n\to\infty}\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\|=0.$$

Since $\{w_n\}$ is bounded, in the view of Lemma 2.2 and (3.19), we only need to show that every subsequence $\{\|x_{n_k}-p\|\}$ of $\{\|x_n-p\|\}$ satisfying

$$\liminf_{k\to\infty}(\|x_{n_k+1}-p\|-\|x_{n_k}-p\|)\ge 0,$$

implies that $\limsup_{k\to\infty}\langle f(p)-p, x_{n_k+1}-p\rangle \le 0$. For this, one assumes that $\{\|x_{n_k}-p\|\}$ is a subsequence of $\{\|x_n-p\|\}$ such that

$$\liminf_{k\to\infty}(\|x_{n_k+1}-p\|-\|x_{n_k}-p\|)\ge 0.$$

This implies that

$$\begin{aligned}&\liminf_{k\to\infty}(\|x_{n_k+1}-p\|^2-\|x_{n_k}-p\|^2)\\&=\liminf_{k\to\infty}((\|x_{n_k+1}-p\|-\|x_{n_k}-p\|)\cdot(\|x_{n_k+1}+p\|+\|x_{n_k}-p\|))\ge 0.\end{aligned}\tag{3.20}$$

Since $\{w_n\}$ and $\{x_n\}$ are bounded, it follows from (3.18), (3.20), Assumption 3.2 and Remark 3.1 that

$$\lim_{k\to\infty}\|w_{n_k}-y_{n_k}\|=0.\tag{3.21}$$

Since $z_n=y_n-\lambda_n(Ay_n-Aw_n)$, we have

$$\begin{aligned}\|z_{n_k}-w_{n_k}\|&=\|y_{n_k}-\lambda_{n_k}(Ay_{n_k}-Aw_{n_k})-w_{n_k}\|\\&\le\|y_{n_k}-w_{n_k}\|+\lambda_n\|Ay_{n_k}-Aw_{n_k}\|\\&\le(1+\lambda_nL)\|y_{n_k}-w_{n_k}\|.\end{aligned}$$

Thus,

$$\lim_{k\to\infty}\|z_{n_k}-w_{n_k}\|=0.\tag{3.22}$$

By virtue of the definition of x_n, we obtain

$$\begin{aligned}\|x_{n_k+1}-w_{n_k}\|&\le\|\alpha_{n_k}(f(x_{n_k})-w_{n_k})+\gamma_{n_k}(z_{n_k}-w_{n_k})\|\\&\le\alpha_{n_k}\|f(x_{n_k})-w_{n_k}\|+\gamma_{n_k}\|z_{n_k}-w_{n_k}\|.\end{aligned}\tag{3.23}$$

Moreover, we see that

$$\begin{aligned}\|f(x_{n_k})-w_{n_k}\|&\le\|f(x_{n_k})-f(p)\|+\|f(p)-w_{n_k}\|\\&\le\rho\|x_{n_k}-p\|+\|f(p)-w_{n_k}\|.\end{aligned}$$

Since $\{x_n\}$ and $\{w_n\}$ are bounded, we see from Assumption 3.2 and (3.23) that

$$\lim_{n\to\infty}\|x_{n_k+1}-w_{n_k}\|=0.\tag{3.24}$$

By Remark 3.1, we see that

$$\|x_{n_k}-w_{n_k}\|=\theta_{n_k}\|x_{n_k}-x_{n_k-1}\|\to 0.\tag{3.25}$$

From (3.24) and (3.25), we obtain

$$\|x_{n_k+1}-x_{n_k}\|\le\|x_{n_k+1}-w_{n_k}\|+\|w_{n_k}-x_{n_k}\|\to 0.\tag{3.26}$$

In view of Step 1, we see that $\{x_{n_k}\}$ is bounded. Hence, there exists a subsequence $\{x_{n_{k_j}}\}$ of $\{x_{n_k}\}$, which converges weakly to some $z\in H$. This implies that

$$\begin{aligned}&\limsup_{k\to\infty}\langle f(p)-p,x_{n_k}-p\rangle=\limsup_{j\to\infty}\langle f(p)-p,x_{n_{k_j}}-p\rangle\\&=\langle f(p)-p,z-p\rangle.\end{aligned}$$

From (3.21), (3.25) and Lemma 3.3, we have $z\in\Omega$. Since $p=P_\Omega\circ f(p)$, we arrive at

$$\limsup_{k\to\infty}\langle f(p)-p,x_{n_k}-p\rangle=\langle f(p)-p,z-p\rangle\le 0.$$

By using (3.26), we obtain

$$\begin{aligned}
&\limsup_{k\to\infty}\langle f(p)-p,x_{n_k+1}-p\rangle \\
&=\limsup_{k\to\infty}\langle f(p)-p,x_{n_k}-p\rangle+\limsup_{k\to\infty}\langle f(p)-p,x_{n_k+1}-x_{n_k}\rangle \\
&=\langle f(p)-p,z-p\rangle\le 0.
\end{aligned} \tag{3.27}$$

Therefore, we have $\|x_n-p\|\to 0$. This completes the proof. □

By Theorem 3.1, we obtain the following conclusions immediately.

Algorithm 3

Initialization: Give $l,\theta,\mu\in(0,1)$ and $\gamma>0$. Let $x_0,x_1\in H$ be arbitrary.
Iterative Steps: Given the current iterator x_n, calculate the next iterator as follows:
Step 1. Compute θ_n, such that $0\le\theta_n\le\theta_n^*$, where

$$\theta_n^*=\begin{cases}\min\left\{\theta,\dfrac{\|\delta_n\|}{\|x_n-x_{n-1}\|}\right\}, & x_n-x_{n-1}\neq 0;\\ \theta, & otherwise.\end{cases} \tag{3.28}$$

Step 2. Compute

$$w_n=x_n+\theta_n(x_n-x_{n-1}),$$

and

$$y_n=(I+\lambda_nB)^{-1}(I-\lambda_nA)w_n,$$

where λ_n is chosen to be the largest $\lambda\in\{\gamma,\gamma l,\gamma l^2,\cdots\}$ satisfying

$$\lambda\|Aw_n-Ay_n\|\le\mu\|w_n-y_n\|. \tag{3.29}$$

If $w_n=y_n$, stop and y_n is a solution of inclusion problem (1.1). Otherwise,
Step 3. Compute

$$z_n=y_n-\lambda_n(Ay_n-Aw_n).$$

and

$$x_{n+1}=\gamma_nz_n+\beta_nw_n+\alpha_nf(x_n).$$

Set $n\leftarrow n+1$ and go to Step 1.

Theorem 3.2. *Suppose that Assumption 3.1, Assumption 3.2 and Assumption 3.3 hold. Let $\{x_n\}$ be a sequence generated by Algorithm 3. Then $\{x_n\}$ converges strongly to $p=P_\Omega\circ f(p)$.*

Proof. Since the proof is similar to that of Theorem 3.1, we only show that, in each iteration of algorithm 3, search rule (3.29) is well defined and

$$\min\{\gamma,\frac{\mu l}{L}\}\le\lambda_n\le\gamma.$$

Since A is $L-$Lipschitz continuous on H, one has

$$\|Aw_n-A(J^B_{\lambda_n}(I-\lambda_nA)w_n)\|\le L\|w_n-J^B_{\lambda_n}(I-\lambda_nA)w_n\|.$$

This is equivalent to

$$\frac{\mu}{L}\|Aw_n-A(J^B_{\lambda_n}(w_n-\lambda_nAw_n))\|\le\mu\|w_n-J^B_{\lambda_n}(w_n-\lambda_nAw_n)\|.$$

This implies that (3.29) holds immediately provided $\lambda \leq \dfrac{\mu}{L}$. Hence, λ_n is well defined.

It is easy to see that $\lambda_n \leq \gamma$. If $\lambda_n = \gamma$, then this lemma is proved. Suppose $\lambda_n < \gamma$. Since λ_n is the largest $\lambda \in \{\gamma, \gamma l, \gamma l^2, \cdots\}$ satisfying (3.29) and $l \in (0,1)$, we see that $\dfrac{\lambda_n}{l}$ must contradict inequality (3.29), i.e.,

$$\frac{\lambda_n}{l}\|Aw_n - A(J^B_{\lambda_n}(w_n - \frac{\lambda_n}{l}w_n))\| \geq \mu\|w_n - J^B_{\lambda_n}(w_n - \frac{\lambda_n}{l}w_n)\|. \tag{3.30}$$

Since A is $L-$Lipschitz continuous on H, one concludes

$$\|Aw_n - A(J^B_{\lambda_n}(w_n - \frac{\lambda_n}{l}w_n))\| \leq L\|w_n - J^B_{\lambda_n}(w_n - \frac{\lambda_n}{l}w_n)\|. \tag{3.31}$$

Combining (3.30) and (3.31), one obtains

$$\lambda_n \geq \frac{\mu l}{L}.$$

Since the rest of the proof is similar to that of Theorem 3.1, we omit it here. This completes the proof. □

4. Numerical results

In this section, we present some numerical results to illustrate the convergence of Algorithm 2. Moreover, we give applications of our algorithm in signal processing. In order to demonstrate the efficiency of proposed algorithm, we also give some comparisons between our method and some existing methods. All the programs are performed on a computer with PC Desktop Intel (R) Core (TM) i5-8250U CPU@1.60GHz.

Example 4.1. Suppose that $A : R^N \to R^N$ is defined by $Ax = 2x + (1,1,1,\cdots,1)$. Let $B : R^N \to R^N$ be defined by $Bx = 5x$ and let $f(x) = 0.8x$. It is easy to verify that A and B satisfy Assumption 3.1 and $f(x)$ satisfies Assumption 3.3. Now, we test the convergence of Algorithm 2. One sees that

$$\begin{aligned} J^B_{\lambda_n}(x - \lambda_n Ax) &= (I + \lambda_n B)^{-1}(x - \lambda_n Ax) \\ &= \frac{1-2\lambda_n}{1+5\lambda_n}x - \frac{\lambda_n}{1+5\lambda_n}(1,1,\cdots,1). \end{aligned}$$

The selection of parameter sequences is as follows. Each component of x_0 and x_1 is generated randomly in (-1,1), $\lambda_0 = 0.1$, $\mu = 0.5$, $\theta_0 = 0.5$, $\alpha_n = \frac{1}{1000n+1}$, $\beta_n = \frac{1}{10n}$, $\gamma_n = 1 - \alpha_n - \beta_n$ and $\delta_n = \frac{1}{1000(n+1)^2}$. We take $Errors = \|x_n - x_{n-1}\|$ in this numerical example. Taking $N = 10$, we obtain the following numerical results in Fig.1 and Tab.1. Fig.1 shows the convergence behavior of each component of x_n, and also shows the relationship between the number of iterations and the errors. On the other hand, Tab.1 gives the value of iterators and the errors. From Fig.1 and Tab.1, we see that the solution is $(-0.142857033, -0.142857033, \cdots, -0.142857033) \in R^{10}$.

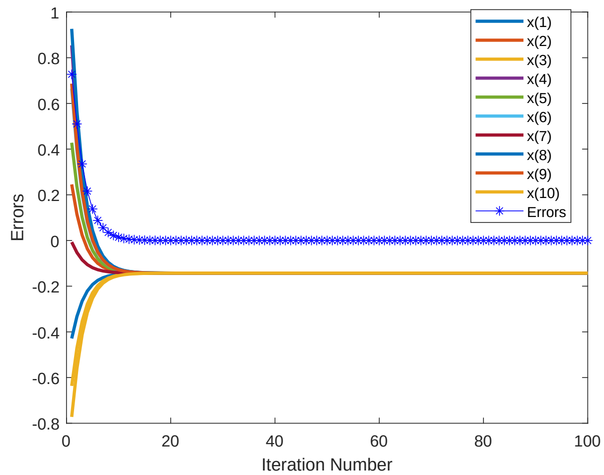

FIGURE 1. The convergence behavior of Algorithm 2 in Example 4.1

TABLE 1. Numerical results of Example 4.1

N	$x_n = (x_1, x_2, x_3, \cdots, x_{10}) \in R^{10}$	Errors
1	(-0.428806933,0.245548195,-0.636455027,$\cdots$)	0.727552254
10	(-0.147881123,-0.136006508,-0.151536564,$\cdots$)	0.014102404
30	(-0.142854779,-0.14285367,-0.14285512,$\cdots$)	1.51E-06
50	(-0.142855518,-0.142855518,-0.142855518,$\cdots$)	1.07E-07
80	(-0.142856151,-0.142856151,-0.142856151,$\cdots$)	4.02E-08
100	(-0.142856355,-0.142856355,-0.142856355,$\cdots$)	2.54E-08
$\vdots$	$\vdots$	$\vdots$
200	(-0.142856755,-0.142856755,-0.142856755,$\cdots$)	6.19E-09
300	(-0.142856885,-0.142856885,-0.142856885,$\cdots$)	2.73E-09
400	(-0.14285695,-0.14285695,-0.14285695,$\cdots$)	1 .53E-09
500	(-0.142856989,-0.142856989,-0.142856989,$\cdots$)	9.77E-10
600	(-0.142857015,-0.142857015,-0.142857015,$\cdots$)	6.77E-10
700	(-0.142857033,-0.142857033,-0.142857033,$\cdots$)	4.97E-10

Example 4.2. Let H be a Hilbert space. Let h and g be two convex, lower semi-continuous functions such that h is differentiable with L-Lipschitz continuous gradient, and the proximal map of g can be computed. The convex minimization problem is that find $x^* \in H$ such that

$$h(x^*) + g(x^*) \leq h(x) + g(x), \forall x \in H.$$

Take $A := \nabla h$ and $B := \partial g$. Then, the convex minimization problem can be reduced to inclusion problem (1.1), that is, find $x^* \in H$ such that

$$0 \in \nabla h(x^*) + \partial g(x^*),$$

where ∇h is a gradient of h and ∂g is a subdifferential of g. Now, we solve the following minimization problem:

$$\min_{x\in R^3} \|x\|_2^2+(3,5,-1)x+9+\|x\|_1,$$

where $x=(x_1,x_2,x_3)\in R^3$. Let $h(x)=\|x\|_2^2+(3,5,-1)x$, $g(x)=\|x\|_1$ and $\Phi(x)=h(x)+g(x)$, we obtain that $\nabla h(x)=2x+(3,5,-1)$. From [12], we have

$$(I+\partial g)^{-1}x=(\max\{|x_1|-r,0\}\text{sign}(x_1),\\ \max\{|x_2|-r,0\}\text{sign}(x_2),\max\{|x_3|-r,0\}\text{sign}(x_3)).$$

We solve this problem by Algorithm 1 and Algorithm 2.

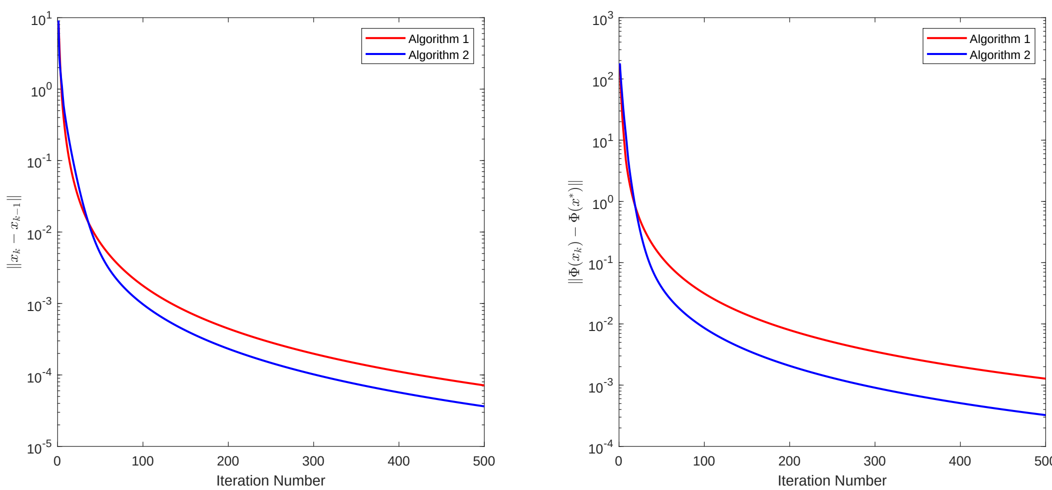

FIGURE 2. Comparison of Algorithms 1 and Algorithm 2 in Example 4.2

We choose $\lambda_0=2$, $\mu=0.5$, $\theta=0.5$. x_0 and x_1 are generated randomly in (0,1). The selection of parameter sequences is as follows: $\alpha_n=\frac{1}{n+1}$, $\beta_n=\gamma_n=\frac{n}{2(n+1)}$ and $\delta_n=\frac{1}{(n+1)^2}$. In Algorithm 2, we choose $f(x)=0.5x$. The numerical results are as follows. In Algorithm 1, $x_{500}=(-9.8403e-01,-1.9681e+00,5.5290e-17)$, $Error_1=7.1411e-05$ and $\|\Phi(x_k)-\Phi(x^*))\|=1.3e-3$. In Algorithm 2, $x_{500}=(-9.9195e-01,-1.9839e+00,5.5400e-17)$, $Error_2=3.6298e-05$ and $\|\Phi(x_k)-\Phi(x^*))\|=3.2396e-04$. Actually, the solutions of this problem are $x^*=(-1,-2,0)$ and $\Phi(x^*)=4$. From numerical results, we see that Algorithm 2 converges more efficiently.

Example 4.3. In this example, we consider the following linear inverse problem:

$$b=Ax_0+w\in\mathbb{R}^M,$$

where $x_0\in\mathbb{R}^N$ is the (unknown) signal to recover, $w\in\mathbb{R}^M$ is a noise vector and $A\in\mathbb{R}^{M\times N}$ models the acquisition device. To recover an approximation of the signal x_0, we use the Basis Pursuit denoising method:

$$\min_{x\in\mathbb{R}^N}\frac{1}{2}\|Ax-b\|^2+\lambda\|x\|_1,$$

where $\|x\|_1$ is a regularizer, the ℓ_1-norm is defined as $\|x\|_1 = \sum_i |x_i|$ and the parameter λ is related to the level of noise $\|w\|$. This problem can be reduced to the minimization of $A+B$, where

$$A(x) = \frac{1}{2}\|Ax-b\|^2 \quad \text{and} \quad B(x) = \lambda\|x\|_1.$$

We see that A is a smooth function satisfying $\nabla A(x) = A^*(Ax-b)$ and ∇A is L-Lipschitz continuous with $L = \|A^*A\|$. Since $B(x) = \lambda\|x\|_1$, the proximal operator of B is a soft thresholding

$$\operatorname{prox}_{\gamma B}(x)_k = \max\left\{0, 1-\frac{\lambda\gamma}{|x_k|}\right\} x_k.$$

A common linear model of signal processing is to consider a linear operator, that is, a filtering

$$Ax = \varphi * x,$$

where φ is the second derivative of Gaussian and $\varphi * x$ is the convolution of φ and x. In this example, our aim is to recover a sparse signal $x_0 \in \mathbb{R}^{400}$ with 16 non zero elements. The purpose of our model is to solve $b = Ax_0 + w$, where w is a realization of Gaussian white noise with the variance is 10^{-2}. The problem can be rewritten as

$$\min_{x\in\mathbb{R}^{400}} \frac{1}{2}||Ax-b||_2^2 + \lambda||x||_1.$$

We choose the regularization parameter $\lambda = 0.5$, the step size $\gamma = 1.9/L$ and the maximum number of iterations is 5×10^4. The selection of the corresponding parameters are the same as in Example 4.2. The recovery results are as follows.

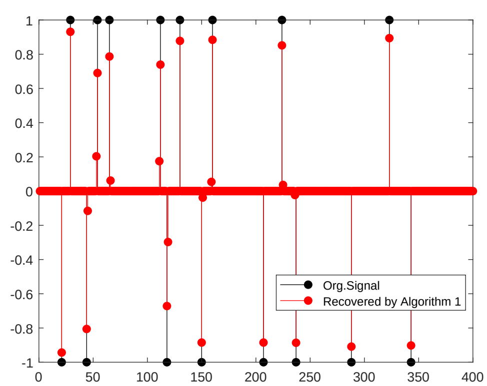

FIGURE 3. Algorithm 1 recovery of a sparse $k = 16$ signal

In Fig.3 and Fig.4, we present the recover sparse signal (with 16 nonzero elements) x_0 from noise observation vector b by Algorithm 1 and Algorithm 2, respectively. The true signal is plotted with the recovered one. Further more, we present the discrepancy of the term $\frac{1}{2}\|Ax-b\|^2$, see Fig.5 (left). The details of the discrepancy term also present in Fig.5 (right). In science and engineering, signal to noise ratio (SNR) is a measure that compares the level of a desired signal to the level of background noise. SNR is the ratio of signal power to the noise power, often expressed in decibels, which is defined as follows, see [10].

$$\text{SNR(dB)} = 10\log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right),$$

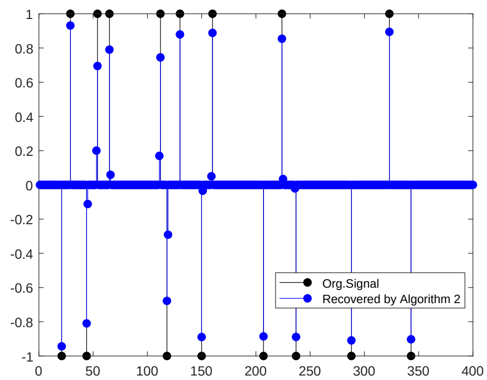

FIGURE 4. Algorithm 2 recovery of a sparse $k = 16$ signal

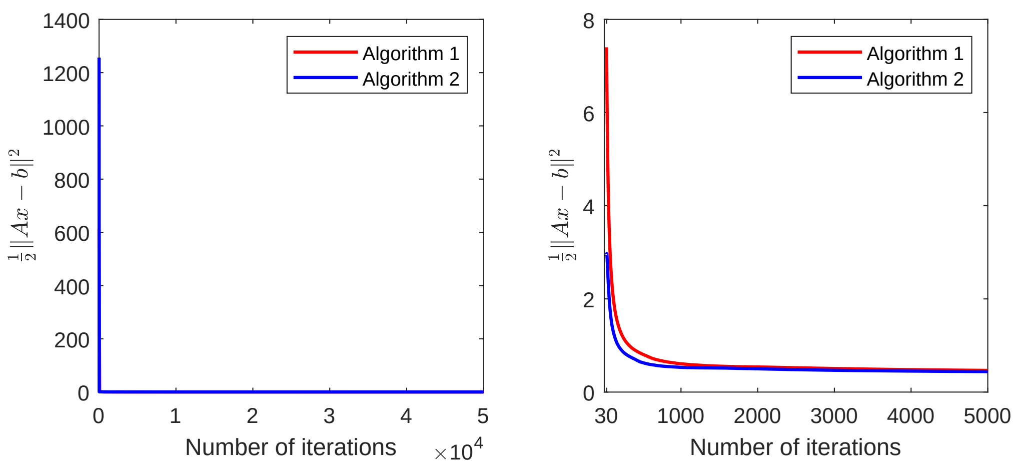

FIGURE 5. Comparison of Algorithm 1 and Algorithm 2 for the $\frac{1}{2}\|Ax-b\|^2$

where P_{signal} is Power of Signal, P_{noise} is Power of Noise. The numerical results show that SNR=13.8090(dB) in Algorithm 1 and SNR=13.9787(dB) in Algorithm 2. We see that both methods are effective in solving the problem, but Algorithm 2 is better.

5. THE CONCLUSION

In this paper, an self-adaptive forward-backward algorithm was proposed for solving the inclusion problem in real Hilbert spaces. The algorithm combines inertial extrapolation technique, the Halpern's method and the Tseng's extragradient method. Strong convergence theorems were obtained under some weak assumptions imposed on the sequence of parameters. Moreover, numerical examples and real applications were provided in the last section.

Acknowledgments
The author is grateful to the referees for their useful comments which improved the presentation of this paper.

REFERENCES

[1] F. Alvarez, Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space, SIAM J. Optim. 14 (2004), 773-782.

[2] H. Attouch, J. Peypouquet, P. Redont, Backward-forward algorithms for structured monotone inclusions in Hilbert spaces, J. Math Anal Appl. 457 (2018), 1095-1117.

[3] H.H. Bauschke, P.L. Combettes, Convex analysis and monotone operator theory in Hilbert spaces, New York, Springer, 2011.

[4] J.Y. Bello Cruz, T.T. Nghia, On the convergence of the forward-backward splitting method with linesearches, Optim Method Softw. 31 (2016), 1209-1238.

[5] R.I. Boţ, E.R. Csetnek, An inertial forward-backward-forward primal-dual splitting algorithm for solving monotone inclusion problems, Numer Algor. 71 (2016), 519-540.

[6] S.Y. Cho, Strong convergence analysis of a hybrid algorithm for nonlinear operators in a Banach space, J. Appl. Anal. Comput. 8 (2018), 19-31.

[7] S.Y. Cho, W. Li, S.M. Kang, Convergence analysis of an iterative algorithm for monotone operators, J. Inequal. Appl. 1 (2013), 199.

[8] S.Y. Cho, X. Qin, L. Wang, Strong convergence of a splitting algorithm for treating monotone operators, Fixed Point Theory Appl. 2014 (2014), Article ID 94.

[9] P.L. Combettes, V. Wajs, Signal recovery by proximal forward-backward splitting, SIAM Multiscale Model. Simul. 4 (2005), 1168-1200.

[10] D.L. Donoho, Compressed sensing, IEEE Trans. Information Theory, 52 (2006), 1289-1306.

[11] A. Gibali, D.V. Thong, Tseng type methods for solving inclusion problems and its applications, Calcolo. 55 (2018), 49.

[12] E.T. Hale, W. Yin, Y. Zhang, A fixed-point continuation method for l_1-regularized minimization with applications to compressed sensing, CAAM TR07-07, Rice University, 43 (2007), 44.

[13] P. L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numer. Anal. 16 (1979), 964-979.

[14] L. Liu, A hybrid steepest descent method for solving split feasibility problems involving nonexpansive mappings, J. Nonlinear Convex Anal., 20 (2019), 471-488.

[15] D.A. Lorenz, Constructing test instances for basis pursuit denoising, IEEE Trans. Signal Process. 61 (2013), 1210-1214.

[16] D.A. Lorenz, T. Pock, An inertial forward-backward algorithm for monotone inclusions, J. Math. Imaging Vis. 51 (2015), 311-325.

[17] P.E. Maingé, A hybrid extragradient-viscosity method for monotone operators and fixed point problems, SIAM J. Control Optim. 47 (2008), 1499-1515.

[18] 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.

[19] H. Raguet, J. Fadili, G.Peyré, A generalized forward-backward splitting, SIAM J. Imaging Sci. 6 (2013), 1199-1226.

[20] X. Qin, S.Y. Cho, L. Wang, A regularization method for treating zero points of the sum of two monotone operators, Fixed Point Theory Appl. 2014 (2014), Article ID 75.

[21] S. Saejung, P. Yotkaew, Approximation of zeros of inverse strongly monotone operators in Banach spaces, Nonlinear Anal. 75 (2012), 742-750.

[22] Y. Shehu, O.S. Iyiola, F.U. Ogbuisi, Iterative method with inertial terms for nonexpansive mappings: applications to compressed sensing, Numer Algor. 83 (2020), 1321-1347.

[23] S. Sra, S. Nowozin, S.J. Wright, Optimization for Machine Learning, Cambridge, MIT Pres. (2012).

[24] W. Takahashi, C.F. Wen, J.C. Yao, The shrinking projection method for a finite family of demimetric mappings with variational inequality problems in a Hilbert space, Fixed Point Theory, 19 (2018) , 407-419.

[25] D.V. Thong, N.T. Vinh, Y.J. Cho, A strong convergence theorem for Tseng's extragradient method for solving variational inequality problems, Opti. Lett. (2019), 1-19.

[26] R. Tibshirani, Regression shrinkage and selection via the lasso, J. Roy Stat Soc. Ser. B. 58 (1996), 267-288.

[27] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, SIAM J. Control Optim. 38 (2000), 431-446.
[28] Y. Wang, F. Wang, Strong convergence of the forward-backward splitting method with multiple parameters in Hilbert spaces, Optimization. 67 (2018), 493-505.