Thai Journal of Mathematics
Volume 17 (2019) Number 3 : 607–625
http://thaijmath.in.cmu.ac.th
ISSN 1686-0209

Hybrid Forward-Backward Algorithms Using Linesearch Rule for Minimization Problem

Kunrada Kankam, Nattawut Pholasa and Prasit Cholamjiak[1]

School of Science, University of Phayao, Phayao 56000, Thailand
e-mail : kunradazzz@gmail.com (K. Kankam)
nattawut_math@hotmail.com (N. Pholasa)
prasitch2008@yahoo.com (P. Cholamjiak)

Abstract : In this work, we investigate strong convergence of the sequences generated by the forward-backward algorithms using hybrid projection method and shrinking projection method for solving the minimization problem. The main advantage of our algorithms is that the Lipschitz constants of the gradient of functions do not require in computation. Finally, we present numerical experiments of our algorithms which are defined by two kinds of projection methods to show the efficiency and the implementation for LASSO problem in signal recovery.

Keywords : forward-backward algorithm; minimization problem; strong convergence; Hilbert space.
2010 Mathematics Subject Classification : 47H04; 47H10.

1 Introduction

This paper is interested in solving the convex minimization problem of the form:

$$\min_{x\in H} \; f(x) + g(x), \tag{1.1}$$

where $f : H \to \mathbb{R} \cup \{+\infty\}$ and $g : H \to \mathbb{R} \cup \{+\infty\}$ are two proper, lower-semicontinuous and convex functions. Problem (1.1) includes many optimization

[1]Corresponding author.

problems arising from applied areas such as signal processing, image recovery, system identification and machine learning [1–4].

Using the fixed point terminology, we know that the problem (1.1) is equivalent to solve the fixed point equation $x = \operatorname{prox}_{\alpha g}(x - \alpha\nabla f(x))$ where α is a positive real number and prox_g is the proximal operator of g defined by $\operatorname{prox}_g = (I + \partial g)^{-1}$. The forward-backward algorithm is a classical method for solving problem (1.1). It is generated by the following manner:

$$x^{k+1} = \underbrace{\operatorname{prox}_{\alpha_k g}}_{\text{backward step}} \underbrace{(x^k - \alpha_k \nabla f(x^k))}_{\text{forward step}}, \tag{1.2}$$

where α_k is a suitable stepsize. This method includes, in particular, the proximal point algorithm [4–7] and the gradient method [8–11]. Due to its wide applications, there have been modifications of (1.2) invented in the literature (see [12–17]).

In 2003, Nakajo and Takahashi [18] introduced the following hybrid projection method and prove its strong convergence for finding a fixed point of a nonexpansive mapping T. Let C be a nonempty closed convex subset of a real Hilbert spaces H. They investigated the sequence (x^k) generated by: $x^0 \in C$ and

$$\begin{cases} y^k = \alpha_k x^k + (1 - \alpha_k) T x^k, \\ C_k = \{z \in C : \|y^k - z\| \le \|x^k - z\|\}, \\ Q_k = \{z \in C : \langle z - x^k, x^0 - x^k \rangle \le 0\}, \\ x^{k+1} = P_{C_k \cap Q_k}(x^0), \end{cases} \tag{1.3}$$

for every $k \in \mathbb{N} \cup \{0\}$, where $(\alpha_k) \subset [0, a]$ for some $a \in [0, 1)$. They proved that (x^k) converges strongly to a fixed point of T. Furthermore, Takahashi et al. [19] proposed the shrinking projection method which is defined by: $x^0 \in C$, $C_1 = C$, $x^1 = P_{C_1}(x^0)$ and

$$\begin{cases} y^k = \alpha_k x^k + (1 - \alpha_k) T x^k, \\ C_{k+1} = \{z \in C_k : \|y^k - z\| \le \|x^k - z\|\}, \\ x^{k+1} = P_{C_{k+1}}(x^0), \end{cases} \tag{1.4}$$

where $0 \le (\alpha_k) < a < 1$ for all $k \in \mathbb{N}$. It was proved that the sequence (x^k) generated by (1.4) converges strongly to a fixed point of a nonexpansive mapping T.

In 2012, Lin and Takahashi [20] introduced the following forward-backward algorithm using the viscosity approximation method.

Algorithm 1.1. :
Initialization Step. *Take* $x^0 \in H$
Iterative Step. *Give* x^k *and set*

$$x^{k+1} = a^k h(x^k) + (1 - a^k)\operatorname{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)),$$

where $h : H \to H$ *is a* $\rho-$*contraction for some* $\rho \in [0,1)$, *i.e.* $\|h(x) - h(y)\| \leq \rho\|x - y\|$ *for all* $x, y \in H$ *and* ∇f *is a* $\nu-$*inverse strongly monotone with the following conditions:*

$$\lim_{n\to\infty} a^k = 0,\ \sum_{k=1}^{\infty} a^k = \infty,\ \sum_{k=1}^{\infty} |a^k - a^{k+1}| < \infty;$$

$$\sum_{k=1}^{\infty} |\alpha_k - \alpha_{k+1}| < \infty,\ 0 < b \leq \alpha_k \leq 2\nu$$

Stop Criteria. *If* $x^{k+1} = x^k$, *then stop.*

Recently, Cruz and Nghia [21] proposed the proximal gradient algorithm using linesearch technique for solving the convex minimization problem in Hilbert spaces. The main advantage of the proposed method is that the Lipschitz condition on the gradient of functions is dropped in computing. The linesearch is defined as follows:

Linesearch 1.1. *Given* $x,\ \sigma > 0,\ \theta \in (0,1)$ *and* $\delta \in (0, \frac{1}{2})$.
Input. *Set* $\alpha = \sigma$ *and* $J(x,\alpha) := \text{prox}_{\alpha g}(x - \alpha\nabla f(x))$ *with* $x \in \text{dom}g$
While $\alpha\|\nabla f(J(x,\alpha)) - \nabla f(x)\| > \delta\|J(x,\alpha) - x\|$
do $\alpha = \theta\alpha$.
End While
Output.α.

It was proved that Linesearch 1.1 is well-defined, *i.e.*, this linesearch stops after finitely many steps. They defined the following algorithm:

Algorithm 1.2. :
Initialization Step. *Take* $x^0 \in \text{dom}g,\ \sigma > 0,\ \theta \in (0,1)$ *and* $\delta \in (0, \frac{1}{2})$
Iterative Step. *Give* x^k *and set*

$$x^{k+1} = \text{prox}_{\alpha_k g}(x^k - \alpha_k\nabla f(x^k)),$$

with $\alpha_k :=$ *Linesearch* $1.1(x^k, \sigma, \theta, \delta)$.
Stop Criteria. *If* $x^{k+1} = x^k$, *then stop.*

It was shown that the sequence generated by Algorithm 1.2 converges weakly to minimizers of $f + g$. Moreover, if the gradient of f is globally Lipschitz continuous on dom$g$ with a constant $L > 0$, then $\alpha_k \geq \min\{\sigma, \delta\theta/L\}$ for all $k \in \mathbb{N}$. However, their algorithms have only weak convergence in real Hilbert spaces. As pointed out, for example, by Bauschke and Combettes [22], the weak convergence of an iterative scheme is an unsatisfactory property in an infinite dimensional setting. Moreover, it is our academic interests to analyze the strong convergence using the linesearch technique.

In this paper, based on Algorithm 1.2, hybrid projection method (1.3) and shrinking projection method (1.4), we introduce new hybrid projection algorithms for solving the convex minimization problem. We then prove the strong convergence theorems of the proposed methods using the linesearch technique. Finally, some numerical experiments in signal recovery are provided to show the efficiency and the implementation of our algorithms.

The rest of this paper is organized as follows: In Sect. 2, we give some definitions and lemmas used for our proof. In Sect. 3, we establish the strong convergence of the proposed algorithms. Finally, in Sect. 4, we show numerical examples to support the convergence of our algorithms. In Sect. 5, we give the conclusions of this paper.

2 Preliminaries

This section contains some definitions and lemmas that play an essential role in our analysis. Let C be a nonempty, closed and convex subset of a real Hilbert space H with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$. The strong (weak) convergence of a sequence $(x^k)_{k\in\mathbb{N}}$ to x is denoted by $x^k \to x$ ($x^k \rightharpoonup x$), respectively.

Definition 2.1. A mapping $T : C \to C$ is said to be nonexpansive if

$$\|Tx - Ty\| \leq \|x - y\|, \forall x, y \in C.$$

In a real Hilbert space, we know that for any point $x \in H$, there exists a unique point $P_C x \in C$ such that

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

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

$$\langle x - y, P_C x - P_C y\rangle \geq \|P_C x - P_C y\|^2, \tag{2.1}$$

for all $x, y \in H$. Furthermore, $P_C x$ is characterized by the property

$$\langle x - P_C x, P_C x - y\rangle \geq 0, \tag{2.2}$$

for all $y \in C$. Moreover, we know that

$$\|y - P_C x\|^2 + \|x - P_C x\|^2 \leq \|x - y\|^2, \ \forall x, y \in H. \tag{2.3}$$

We also know that all Hilbert space has the Kadec-Klee property, that is, (x^k) converges weakly to x and $\|x^k\| \to \|x\|$ imply (x^k) converges strongly to x.

Definition 2.2. The subdifferential of a function h at x is defined by

$$\partial h(x) = \{v \in H : \langle v, y - x\rangle \leq h(y) - h(x), y \in H\}.$$

Fact 2.1. [22, Proposition 17.2] *Let $h : H \to \mathbb{R} \cup \{+\infty\}$ be a proper, lower-semicontinuous and convex function. Then, for $x \in \mathrm{dom}h$ and $y \in H$,*

$$h'(x; y - x) + h(x) \leq h(y).$$

Lemma 2.3. [23] *The subdifferential operator ∂h is maximal monotone. Moreover, the graph of ∂h, $\mathrm{Gph}(\partial h) = \{(x,v) \in H \times H : v \in \partial h(x)\}$ is demiclosed, i.e., if the sequence $(x^k, v^k) \subset \mathrm{Gph}(\partial h)$ satisfies that $(x^k)_{k\in\mathbb{N}}$ converges weakly to x and $(v^k)_{k\in\mathbb{N}}$ converges strongly to v, then $(x,v) \in \mathrm{Gph}(\partial h)$.*

Let us recall the proximal operator $\mathrm{prox}_g : H \to \mathrm{dom}g$ with $\mathrm{prox}_g(z) = (\mathrm{I} + \partial g)^{-1}(z), z \in H$. Here I denotes the identity operator. It is well-known that the proximal operator is single-valued with full domain. It is also known that

$$\frac{z - \mathrm{prox}_{\alpha g}(z)}{\alpha} \in \partial g(\mathrm{prox}_{\alpha g}(z)) \text{ for all } z \in H,\ \alpha > 0. \tag{2.4}$$

3 Main Results

In this section, we propose the forward-backward splitting algorithm using the projection algorithm and prove the strong convergence theorem.

Following [21], we assume that two below conditions hold:
(A1) $f, g : H \to \mathbb{R} \cup \{+\infty\}$ are two proper, lower-semicontinuous and convex functions with $\mathrm{dom}g \subseteq \mathrm{dom}f$ and $\mathrm{dom}g$ is nonempty, closed and convex.
(A2) The function f is Fréchet differentiable on an open set containing $\mathrm{dom}g$. The gradient ∇f is uniformly continuous on any bounded subset of $\mathrm{dom}g$ and maps any bounded subset of $\mathrm{dom}g$ to a bounded set in H.

Algorithm 3.1. *(step 0) Choose $x^0 \in \mathrm{dom}g$, take $\delta \in (0, \frac{1}{2})$, $\sigma > 0$ and $\theta \in (0,1)$.*
(step 1) Set $\alpha_k = \sigma\theta^{m_k}$ and m_k is the smallest nonnegative integer such that

$$\begin{aligned} & \alpha_k \|\nabla f(\mathrm{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k))) - \nabla f(x^k)\| \\ \leq\ & \delta \|\mathrm{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)) - x^k\|. \end{aligned} \tag{3.1}$$

(step 2) Set

$$y^k = \mathrm{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)). \tag{3.2}$$

(step 3) Compute

$$C_k = \{x_* \in \mathrm{dom}g : \|y^k - x_*\| \leq \|x^k - x_*\|\}$$

and

$$Q_k = \{x_* \in \mathrm{dom}g : \langle x_* - x^k, x^0 - x^k \rangle \leq 0\}. \tag{3.3}$$

(step 4) Compute

$$x^{k+1} = P_{C_k \cap Q_k}(x^0). \tag{3.4}$$

(step 5) Set $k \leftarrow k+1$, and go to (step 1).

Throughout this paper, we denote Ω by the solution set of (1.1) and assume that Ω is nonempty.

Theorem 3.1. *Let H be a real Hilbert space. Assume that there exists $\alpha > 0$ such that $\alpha_k \geq \alpha > 0$. Then the sequence $(x^k)_{k=0}^{\infty}$ generated by Algorithm 3.1 converges strongly to $\bar{x} = P_{\Omega}(x^0)$.*

Proof. We divide our proof into four steps.
Step 1 Show that $(x_k)_{k=0}^{\infty}$ is well-defined and $\Omega \subset C_k \cap Q_k, \forall k \geq 0$. For each $x \in \text{dom}g$, we see that

$$\begin{aligned}\|y^k - x\| \leq \|x^k - x\| \quad &\leftrightarrow \quad \|y^k\|^2 - 2\langle x, y^k\rangle \leq \|x^k\|^2 - 2\langle x, x^k\rangle \\ &\leftrightarrow \quad 2\langle x, x^k - y^k\rangle \leq \|x^k\|^2 - \|y^k\|^2 \\ &\leftrightarrow \quad \langle x, x^k - y^k\rangle \leq \frac{1}{2}[\|x^k\|^2 - \|y^k\|^2].\end{aligned} \tag{3.5}$$

Therefore C_k is closed and convex for all $k \geq 0$. Moreover, it is easy to see that Q_k is closed and convex for all $k \geq 0$. Therefore, $C_k \cap Q_k$ is closed and convex for all $k \geq 0$. Using (2.4) and (3.2), we observe that

$$\frac{x^k - y^k}{\alpha_k} - \nabla f(x^k) = \frac{x^k - \text{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k))}{\alpha_k} - \nabla f(x^k) \in \partial g(y^k).$$

The convexity of g gives

$$g(x) - g(y^k) \geq \langle \frac{x^k - y^k}{\alpha_k} - \nabla f(x^k), x - y^k\rangle, \forall x \in \text{dom}g. \tag{3.6}$$

The convexity of f also implies

$$f(x) - f(y) \geq \langle \nabla f(y), x - y\rangle, \forall x \in \text{dom}f, \ y \in \text{dom}g. \tag{3.7}$$

From (3.6) and (3.7) with any $x \in \text{dom}g \subseteq \text{dom}f$ and $y = x^k \in \text{dom}g$, we have

$$\begin{aligned}(f+g)(x) \quad &\geq \quad f(x^k) + g(y^k) + \langle \frac{x^k - y^k}{\alpha_k} - \nabla f(x^k), x - y^k\rangle \\ &\quad\quad + \langle \nabla f(x^k), x - x^k\rangle \\ &= \quad f(x^k) + g(y^k) + \frac{1}{\alpha_k}\langle x^k - y^k, x - y^k\rangle \\ &\quad\quad + \langle \nabla f(x^k) - \nabla f(y^k), y^k - x^k\rangle + \langle \nabla f(y^k), y^k - x^k\rangle \\ &\geq \quad f(x^k) + g(y^k) + \frac{1}{\alpha_k}\langle x^k - y^k, x - y^k\rangle \\ &\quad\quad - \|\nabla f(x^k) - \nabla f(y^k)\|\|y^k - x^k\| + \langle \nabla f(y^k), y^k - x^k\rangle \\ &\geq \quad f(x^k) + g(y^k) + \frac{1}{\alpha_k}\langle x^k - y^k, x - y^k\rangle - \frac{\delta}{\alpha_k}\|x^k - y^k\|^2 \\ &\quad\quad + \langle \nabla f(y^k), y^k - x^k\rangle,\end{aligned}$$

where the last inequality follows from the linesearch (3.1). Hence we obtain

$$\begin{aligned}\langle x^k - y^k, y^k - x\rangle &\geq \alpha_k[f(x^k) + g(y^k) - (f+g)(x) + \langle \nabla f(y^k), y^k - x^k\rangle] \\ &\quad -\delta\|x^k - y^k\|^2. \end{aligned} \tag{3.8}$$

Replacing $x = x^k$ and $y = y^k$ in (3.7), we have $f(x^k) - f(y^k) \geq \langle \nabla f(y^k), x^k - y^k\rangle$. From (3.8), we get

$$\langle x^k - y^k, y^k - x\rangle \geq \alpha_k[(f+g)(y^k) - (f+g)(x)] - \delta\|x^k - y^k\|^2. \tag{3.9}$$

Since $2\langle x^k - y^k, y^k - x\rangle = \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - x\|^2$, by (3.9), it follows that

$$\begin{aligned}\|y^k - x\|^2 &\leq \|x^k - x\|^2 - 2\alpha_k[(f+g)(y^k) - (f+g)(x)] \\ &\quad -(1-2\delta)\|x^k - y^k\|^2. \end{aligned} \tag{3.10}$$

Let $x_* \in \Omega$ and set $x = x_*$ in (3.10). Hence we have

$$\|y^k - x_*\| \leq \|x^k - x_*\|. \tag{3.11}$$

Thus $x_* \in C_k, \forall k \geq 0$. Therefore, $\Omega \subset C_k, \forall k \geq 0$. For $k = 0$, we have that $x^0 \in \text{dom}g$ and $Q_0 = \text{dom}g$ and hence $\Omega \subset C_0 \cap Q_0$. Assume that x^n is given and $\Omega \subset C_n \cap Q_n$ for some $n \in \{0, 1, 2, ...\}$. Since Ω is nonempty, $C_n \cap Q_n$ is nonempty, closed and convex. So there exists a unique element $x^{n+1} \in C_n \cap Q_n$ such that $x^{n+1} = P_{C_n \cap Q_n}(x^0)$. This gives

$$\langle x_* - x^{n+1}, x^0 - x^{n+1}\rangle \leq 0, \ \forall x_* \in C_n \cap Q_n. \tag{3.12}$$

Since $\Omega \subset C_n \cap Q_n$, in particular, we obtain

$$\langle x_* - x^{n+1}, x^0 - x^{n+1}\rangle \leq 0, \ \forall x_* \in \Omega. \tag{3.13}$$

This implies that $\Omega \subset Q_{n+1}$. By induction we conclude that, $\Omega \subset C_k \cap Q_k, \forall k \geq 0$ and thus $(x^k)_{k=0}^{\infty}$ is well-defined.

Step 2 Show that $(x^k)_{k=0}^{\infty}$ is bounded. From (3.3), we see that

$$\langle x_* - x^k, x^0 - x^k\rangle \leq 0, \forall x_* \in Q_k.$$

This implies that $x^k = P_{Q_k}(x^0)$. Then we have

$$\|x^k - x^0\| \leq \|x^0 - x_*\|, \forall x_* \in Q_k.$$

Since $\Omega \subset Q_k$, it follows that

$$\|x^k - x^0\| \leq \|x^0 - x_*\|, \forall x_* \in \Omega. \tag{3.14}$$

In particular, since $x^{k+1} \in Q_k$,

$$\|x^k - x^0\| \leq \|x^{k+1} - x^0\|. \tag{3.15}$$

By (3.14) and (3.15), we obtain $\lim_{k\to\infty}\|x^k - x^0\|$ exists. Hence $(x^k)_{k=0}^{\infty}$ is bounded.

Step 3 Show that $\lim_{k\to\infty}\|x^{k+1}-x^k\| = 0$. By (2.3) and the fact that $x^k = P_{Q_k}(x^0)$, we see that

$$\|x^{k+1} - x^k\|^2 \leq \|x^{k+1} - x^0\|^2 - \|x^k - x^0\|^2.$$

Since $\lim_{k\to\infty}\|x^k - x^0\|$ exists, it follows that $\lim_{k\to\infty}\|x^{k+1} - x^k\| = 0$.

Step 4 Show that $\lim_{k\to\infty} x^k = \bar{x}$, where $\bar{x} = P_\Omega(x^0)$. From (3.3), $x^{k+1} \in C_k$ and Step 3, we see that

$$\|y^k - x^{k+1}\| \leq \|x^k - x^{k+1}\| \to 0,\ k \to \infty.$$

Hence we obtain

$$\begin{aligned}\|y^k - x^k\| &\leq \|y^k - x^{k+1}\| + \|x^{k+1} - x^k\| \\ &\to 0,\ k \to \infty. \end{aligned} \tag{3.16}$$

Since $(x^k)_{k=0}^{\infty}$ is bounded, the set of its weak accumulation point is nonempty. Take any weak accumulation point ω of (x^k). So there is a subsequence $(x^{k_n})_{n=0}^{\infty}$ of $(x^k)_{k=0}^{\infty}$ weakly converging to ω. We get from (3.16) and assumption (A2) that

$$\lim_{n\to\infty}\|\nabla f(y^{k_n}) - \nabla f(x^{k_n})\| = 0. \tag{3.17}$$

Since $y^{k_n} = \text{prox}_{\alpha_{k_n} g}(x^{k_n} - \alpha_{k_n}\nabla f(x^{k_n}))$, it follows from (2.4) that

$$\frac{x^{k_n} - \alpha_{k_n}\nabla f(x^{k_n}) - y^{k_n}}{\alpha_{k_n}} \in \partial g(y^{k_n})$$

which implies that

$$\frac{x^{k_n} - y^{k_n}}{\alpha_{k_n}} + \nabla f(y^{k_n}) - \nabla f(x^{k_n}) \in \nabla f(y^{k_n}) + \partial g(y^{k_n}) \subseteq \partial(f+g)(y^{k_n}). \tag{3.18}$$

From (3.16), (3.17) and (3.18), we conclude that $\omega \in \Omega$ by Lemma 2.3. If $\bar{x} = P_\Omega(x^0)$, it then follows from (3.14), the fact that $\omega \in \Omega$ and the lower semicontinuity of the norm that,

$$\begin{aligned}\|x^0 - \bar{x}\| &\leq \|x^0 - \omega\| \\ &\leq \liminf_{n\to\infty}\|x^0 - x^{k_n}\| \\ &\leq \limsup_{n\to\infty}\|x^0 - x^{k_n}\| \\ &\leq \|x^0 - \bar{x}\|. \end{aligned} \tag{3.19}$$

Hence we obtain $\lim_{n\to\infty}\|x^{k_n} - x^0\| = \|x^0 - \omega\| = \|x^0 - \bar{x}\|$. This yields $x^{k_n} \to \omega = \bar{x}, n \to \infty$. It follows that (x^k) converges weakly to $\bar{x}$. So we have

$$\begin{aligned}\|x^0 - \bar{x}\| &\leq \liminf_{n\to\infty}\|x^0 - x^k\| \\ &\leq \limsup_{n\to\infty}\|x^0 - x^k\| \\ &\leq \|x^0 - \bar{x}\|. \end{aligned} \tag{3.20}$$

This shows that $\lim_{n\to\infty}\|x^k - x^0\| = \|x^0 - \bar{x}\|$. From $x^k \rightharpoonup \bar{x}$, we also have $x^k - x^0 \rightharpoonup \bar{x} - x^0$. Since H satisfies the Kadec-Klee property, it follows that $x^k - x^0 \to \bar{x} - x^0$. Therefore $x^k \to \bar{x}$ as $k \to \infty$. This completes the proof. □

We next introduce another version of the forward-backward algorithm based on the shrinking projection method.

Algorithm 3.2. *(step 0) Set* $C_0 = \mathrm{dom} g$*, choose* $x^0 \in \mathrm{dom} g$*, take* $\delta \in (0, \frac{1}{2})$*,* $\sigma > 0$ *and* $\theta \in (0, 1)$*.*
(step 1) Set $\alpha_k = \sigma\theta^{m_k}$ *and* m_k *is the smallest nonnegative integer such that*

$$\alpha_k \|\nabla f(y^k) - \nabla f(x^k)\| \le \delta \|y^k - x^k\|. \tag{3.21}$$

(step 2) Set

$$y^k = \mathrm{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)). \tag{3.22}$$

(step 3) Compute

$$C_{k+1} \quad = \quad \{x_* \in C_k : \|y^k - x_*\| \le \|x^k - x_*\|\}. \tag{3.23}$$

(step 4) Compute

$$x^{k+1} = P_{C_{k+1}}(x^0). \tag{3.24}$$

(step 5) Set $k \leftarrow k+1$*, and go to (step 1).*

Theorem 3.2. *Let* H *be a real Hilbert space. Assume that there exists* $\alpha > 0$ *such that* $\alpha_k \ge \alpha > 0$*. Then the sequence* $(x^k)_{k=0}^{\infty}$ *generated by Algorithm* 3.2 *converges strongly to* $\bar{x} = P_\Omega(x^0)$*.*

Proof. We divide our proof into five steps.
Step 1 Show that $P_{C_{k+1}}(x^0)$ is well-defined and $\Omega \subseteq C_{k+1}, \forall k \ge 0$. Similar to Step 1 in Theorem 3.1, we can show that C_{k+1} is closed and convex, $\forall k \ge 0$. Also, we can show that

$$\|x^k - x^0\| \le \|x^0 - x_*\|, \forall x_* \in C_k.$$

Thus, if $x_* \in \Omega$, then we have $x_* \in C_{k+1}$. So $\Omega \subseteq C_{k+1}$ and $P_{C_{k+1}}(x^0)$ is well-defined.
Step 2 Show that $\lim_{k\to\infty}\|x^k - x^0\|$ exists. From $x^k = P_{C_k}x^0$, $C_{k+1} \subset C_k$ and $x^{k+1} \in C_k$, $\forall k \ge 1$, we get

$$\|x^k - x^0\| \le \|x^{k+1} - x^0\|, \forall k \ge 0. \tag{3.25}$$

On the other hand, since $\Omega \subset C_k$, we obtain

$$\|x^k - x^0\| \le \|x_* - x^0\|, \forall x_* \in \Omega. \tag{3.26}$$

It follows that the sequence (x^k) is bounded and nondecreasing. Therefore, $\lim_{k\to\infty} \|x^k - x^0\|$ exists.
Step 3 Show that $x^k \to \bar{x}$ as $k \to \infty$. For $l > k$, by the definition of C_k, we see that $x^l = P_{C_l}(x^0) \in C_l \subset C_k$. So we obtain

$$\|x^l - x^k\|^2 \leq \|x^l - x^0\|^2 - \|x^k - x^0\|^2.$$

From Step 2, we have $(x^k)_{k=0}^{\infty}$ is a Cauchy sequence. Hence, $x^k \to \bar{x}$ as $k \to \infty$.
Step 4 Show that $\bar{x} \in \Omega$. From Step 3, we see that

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0.$$

Since $x^{k+1} \in C_{k+1} \subset C_k$, we have

$$\|y^k - x^{k+1}\| \leq \|x^k - x^{k+1}\| \to 0, \; k \to \infty.$$

It follows that

$$\begin{aligned} \|y^k - x^k\| &\leq \|y^k - x^{k+1}\| + \|x^{k+1} - x^k\| \\ &\to 0, \; k \to \infty. \end{aligned} \tag{3.27}$$

We get from (3.27) and assumption (A2) that

$$\lim_{k\to\infty} \|\nabla f(y^k) - \nabla f(x^k)\| = 0. \tag{3.28}$$

Since $y^k = \text{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k))$, it follows from (2.4) that

$$\frac{x^k - \alpha_k \nabla f(x^k) - y^k}{\alpha_k} \in \partial g(y^k)$$

which implies that

$$\frac{x^k - y^k}{\alpha_k} + \nabla f(y^k) - \nabla f(x^k) \in \nabla f(y^k) + \partial g(y^k) \subseteq \partial(f+g)(y^k). \tag{3.29}$$

From (3.27), (3.28) and (3.29), we have $\bar{x} \in \Omega$. by Lemma 2.3
Step 5 Show that $\bar{x} = P_\Omega(x^0)$. Since $x^k = P_{C_k}(x^0)$ and $\Omega \subset C_k$, we obtain

$$\langle x^0 - x^k, x^k - x_* \rangle \geq 0, \forall x_* \in \Omega. \tag{3.30}$$

By taking the limit in (3.30), we obtain

$$\langle x^0 - \bar{x}, \bar{x} - x_* \rangle \geq 0, \forall x_* \in \Omega. \tag{3.31}$$

This shows that $\bar{x} = P_\Omega(x^0)$. We thus complete the proof. □

4 Numerical Experiments

In this section, we present some numerical examples to the signal recovery. We consider our first algorithm defined by projection method and provide a comparison among Algorithm 1.1, Algorithm 1.3 and Algorithm 3.1. In this case, we set $Tx^k = \text{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k))$. It is known that T is a nonexpansive mapping when $\alpha_k \in (0, \frac{2}{L})$ and L is the Lipschitz constant of ∇f. Compressed sensing can be modeled as the following underdeterminated linear equation system:

$$y = Ax + \epsilon, \tag{4.1}$$

where $x \in \mathbb{R}^N$ is a vector with k nonzero components to be recovered, $y \in \mathbb{R}^M$ is the observed or measured data with noisy ϵ, and $A : \mathbb{R}^N \to \mathbb{R}^M (M < N)$ is a bounded linear operator. It is known that to solve (4.1) can be seen as solving the LASSO problem:

$$\min_{x \in \mathbb{R}^N} \frac{1}{2}\|y - Ax\|_2^2 + \lambda \|x\|_1, \tag{4.2}$$

where $\lambda > 0$. So we can apply our method for solving (4.2) in case $f(x) = \frac{1}{2}\|y - Ax\|_2^2$ and $g(x) = \lambda\|x\|_1$. It is noted that $\nabla f(x) = A^T(Ax - y)$.

In our experiment, the sparse vector $x \in \mathbb{R}^N$ is generated from uniform distribution in the interval [-2,2] with k nonzero elements. The matrix $A \in \mathbb{R}^{M\times N}$ is generated from a normal distribution with mean zero and invariance one. The observation y is generated by with Gaussian noise white signal-to-noise ratio SNR=40. The initial point x^0 is picked randomly. The restoration accuracy is measured by the mean squared error as follows:

$$\text{MSE} = \frac{1}{N}\|x^k - x^*\|_2^2 < 10^{-5},$$

where x^* is an estimated signal of x.

In what follows, let $\sigma = 5$, $\theta = 0.4$, and $\delta = 0.4$ in both Algorithm 1.3 and Algorithm 3.1 and let the step size α_k in Algorithm 1.1 and Algorithm 1.3 be $\frac{1}{\|A\|^2}$. Let $h(x) = \frac{x}{5}$ be a contraction and choose $a^k = \frac{1}{100k}$ in all algorithms. We denote by CPU the time using in CPU and Iter the number of iterations. The numerical results are reported by Table 1.

The data in Table 1 shows that, for a given tolerance, all algorithms can be used to solve the LASSO problem in compressed sensing. To be more precise, Algorithm 3.1 with a linesearch take significantly less number of iterations and CPU time compared to Algorithm 1.1 of [20] and Algorithm 1.3 of [21].

Table 1: Computational results for solving the LASSO problem

m-sparse signal	Method	N=512, M=256		N=1024, M=512	
		CPU	Iter	CPU	Iter
m=20	Algorithm 3.1	4.3612	673	35.2779	1258
	Algorithm 1.3	41.5479	3645	265.4392	6851
	Algorithm 1.1	9.7712	1742	65.0949	3249
m=30	Algorithm 3.1	6.0680	793	32.7622	1335
	Algorithm 1.3	56.6697	4370	282.0070	7265
	Algorithm 1.1	13.0234	2109	64.8357	3457
m=40	Algorithm 3.1	5.5765	790	35.2468	1391
	Algorithm 1.3	57.1358	4495	324.6561	7639
	Algorithm 1.1	14.2279	2175	71.0742	3649
m=50	Algorithm 3.1	7.8385	1024	41.1793	1416
	Algorithm 1.3	96.3842	5901	357.4149	7818
	Algorithm 1.1	24.7290	2873	88.7461	3731
m=60	Algorithm 3.1	16.6168	1486	52.7078	1534
	Algorithm 1.3	155.5344	8933	287.6895	8668
	Algorithm 1.1	36.4613	4381	70.8262	4164

We next provide some numerical experiments for two cases to illustrate the convergence behavior of all algorithm in comparison. We plot the original signal, observation data, recovered signal, the number of iterations versus MSE.

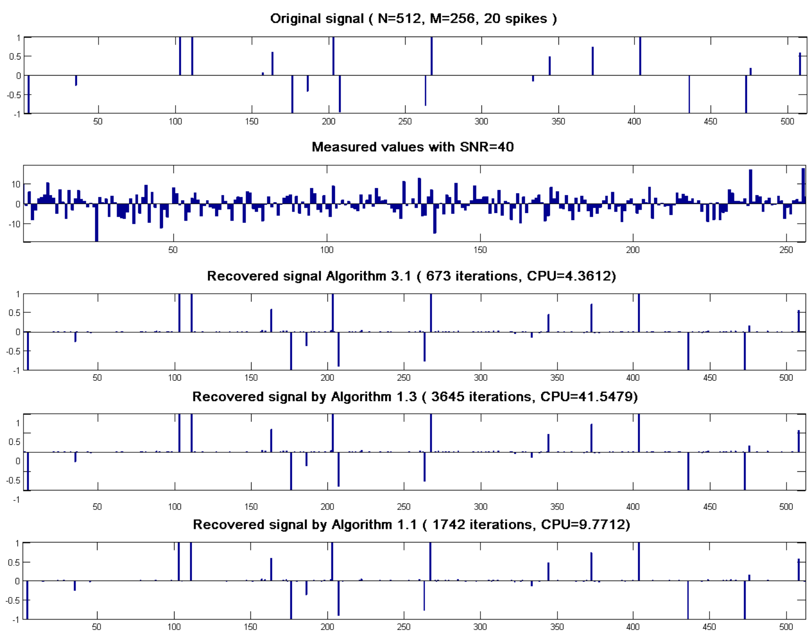

Figure 1: From top to bottom: original signal, observation data, recovered signal by Algorithm 3.1, Algorithm 1.3 and Algorithm 1.1 with $N = 512$ and $M = 256$, respectively.

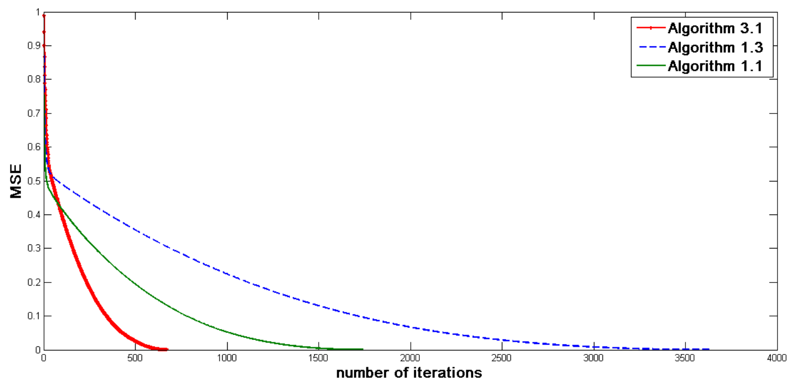

Figure 2: The objective function value versus number of iterations in case N=512, M=256.

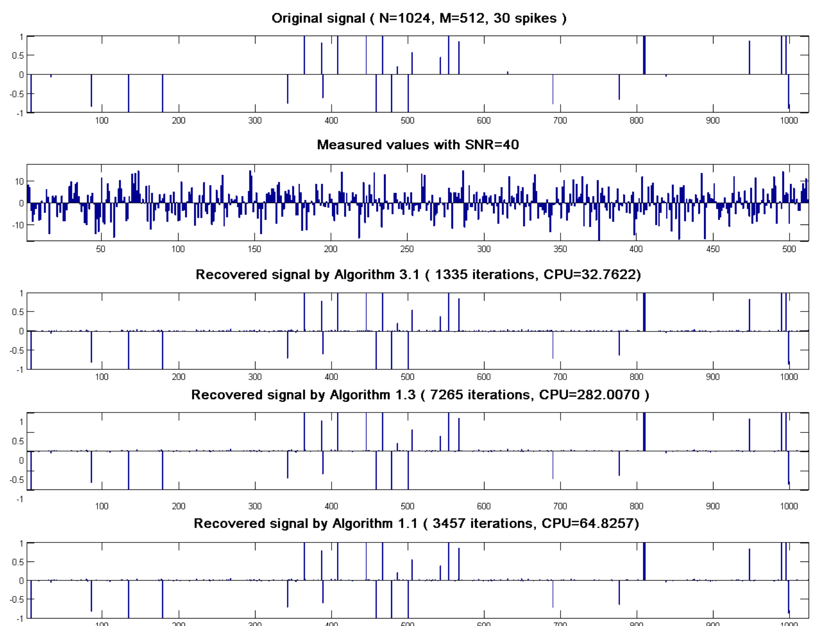

Figure 3: From top to bottom: original signal, observation data, recovered signal by Algorithm 3.1, Algorithm 1.3 and Algorithm 1.1 with $N = 512$ and $M = 256$, respectively.

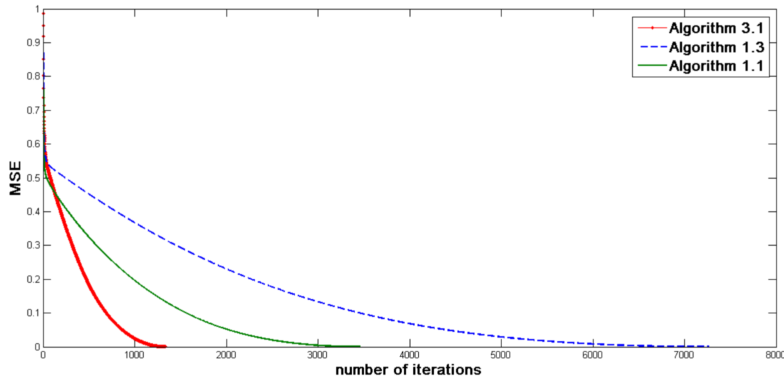

Figure 4: The objective function value versus number of iterations in case N=512, M=256.

Next, we discuss our forward-backward algorithm defined by the shrinking projection method. We provide a comparison among Algorithm 1.1, Algorithm 1.4 and Algorithm 3.2. For convenience, we set all condition as in the previous example.

Table 2: Computational results for solving the LASSO problem

m-sparse signal	Method	N=512, M=256		N=1024, M=512	
		CPU	Iter	CPU	Iter
m=20	Algorithm 3.2	5.2158	660	40.9654	1247
	Algorithm 1.4	29.0458	3548	183.4528	6789
	Algorithm 1.1	5.8716	1696	46.3793	3222
m=30	Algorithm 3.2	8.7975	865	45.0510	1369
	Algorithm 1.4	42.1279	4820	236.4308	7648
	Algorithm 1.1	9.8976	2325	62.7609	3645
m=40	Algorithm 3.2	7.4329	926	42.7707	1369
	Algorithm 1.4	53.7019	5079	224.0050	7551
	Algorithm 1.1	12.2709	2461	54.1403	3608
m=50	Algorithm 3.2	8.6868	1099	56.2084	1508
	Algorithm 1.4	107.2563	6309	308.9143	8439
	Algorithm 1.1	20.1471	3085	67.8876	4053
m=60	Algorithm 3.2	13.3035	1223	40.4446	1439
	Algorithm 1.4	90.2037	7025	233.2904	8016
	Algorithm 1.1	24.5405	3447	57.9466	3838

The data in Table 2 shows that, for a given tolerance, all algorithms can be used to solve the LASSO problem in compressed sensing. To be more precise, Algorithm 3.2 with a linesearch take significantly less number of iterations and CPU time compared to Algorithm 1.1 of [20] and Algorithm 1.4 of [18].

We plot the original signal, observation data, recovered signal, the number of iterations versus MSE.

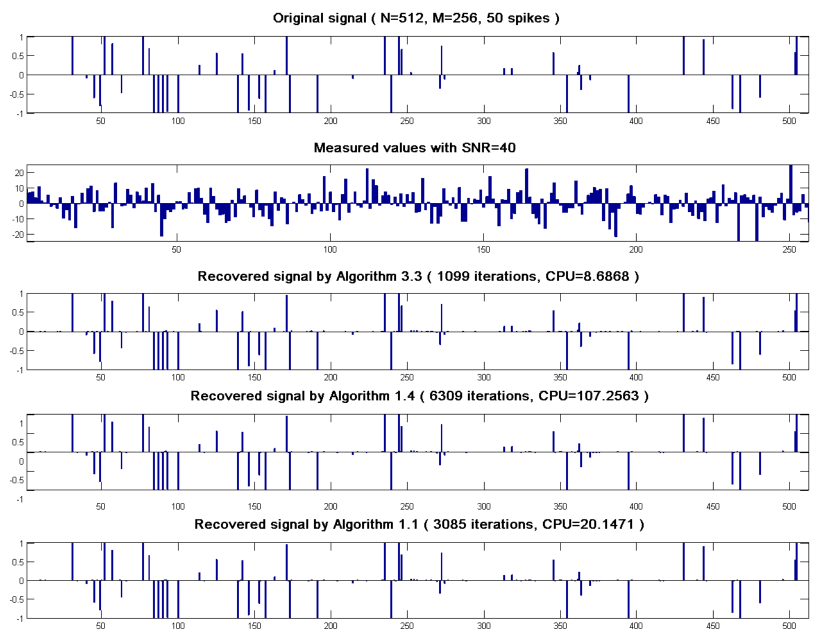

Figure 5: From top to bottom: original signal, observation data, recovered signal by Algorithm 3.2, Algorithm 1.4 and Algorithm 1.1 with $N = 512$ and $M = 256$, respectively.

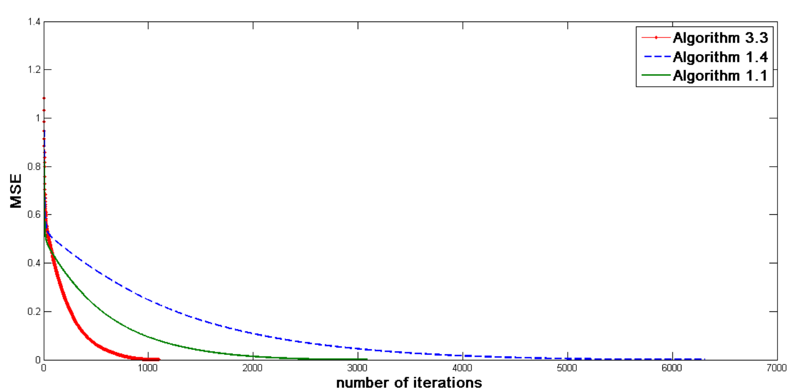

Figure 6: The objective function value versus number of iterations in case N=512, M=256.

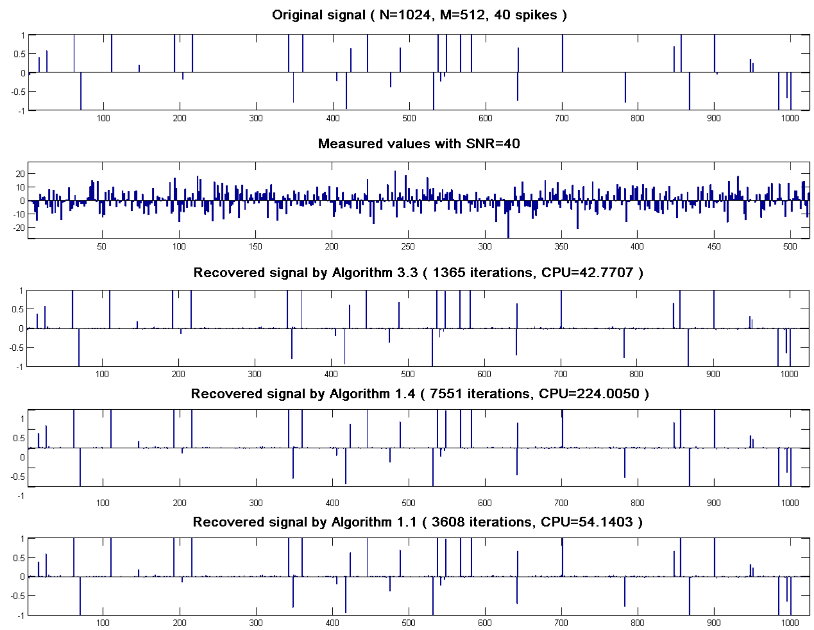

Figure 7: From top to bottom: original signal, observation data, recovered signal by Algorithm 3.2, Algorithm 1.4 and Algorithm 1.1 with $N = 512$ and $M = 256$, respectively.

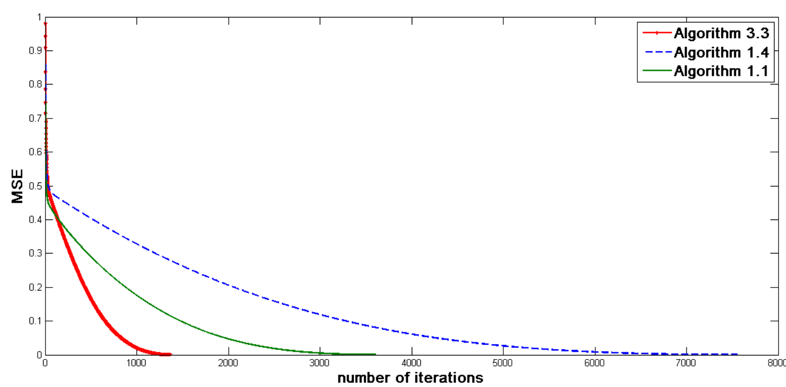

Figure 8: The objective function value versus number of iterations in case N=512, M=256.

5 Conclusions

In this work, we discuss the modified forward-backward splitting method involving linesearches for solving minimization problems of two convex functions. We prove strong convergence theorems by using projection method and shrinking projection method. All the results are compared, in compressed sensing, with different kinds of forward-backward methods. It is found that algorithm using linesearch has a better convergence behavior than other methods through experiments. Our algorithms do not require to compute the Lipschitz constant of the gradient of functions. This advantage is very useful and convenient in practice.

Acknowledgements : This research was supported by Thailand Research Fund and University of Phayao under the project RSA6180084 and UOE62001. This work was partially supported by Thailand Science Research and Innovation under the project IRN62W0007.

References

[1] P.L. Combettes, J.C. Pesquet, Proximal splitting methods in signal processing, in Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Springer, New York (2011) 185-212.

[2] P.L. Combettes, V.R. Wajs, Signal recovery by proximal forward-backward splitting, Multiscale Modeling Simulation 4 (4) (2005) 1168-1200.

[3] K. Kankam, N. Pholasa, P. Cholamjiak, On convergence and complexity of the modified forward backward method involving new linesearches for convex minimization, Mathematical Methods in the Applied Sciences 42 (5) (2019) 1352-1362.

[4] N. Parikh, S. Boyd, Proximal algorithms, Foundations and Trends in Optimization 1 (3) (2014) 127-239.

[5] O. Guler, On the convergence of the proximal point algorithm for convex minimization, SIAM Journal on Control and Optimization 29 (2) (1991) 403-419.

[6] B. Martinet, Brvèv communication, Régularisation d'inéquations variationnelles par approximations successives, Revue franaise d'informatique et de recherche opérationnelle, Série rouge 4 (1970) 154-158.

[7] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM Journal on Control and Optimization 14 (5) (1976) 877-898.

[8] J.C. Dunn, Convexity, monotonicity, and gradient processes in Hilbert space, Journal of Mathematical Analysis and Applications 53 (1) (1976) 145-158.

[9] K. Sitthithakerngkiet, J. Deepho, P. Kumam, Modified hybrid steepest method for the split feasibility problem in image recovery of inverse problems, Numerical Functional Analysis and Optimization 38 (4) (2017) 507-522.

[10] C. Wang, N. Xiu, Convergence of the gradient projection method for generalized convex minimization, Computational Optimization and Applications 16 (2) (2000) 111-120.

[11] H.K. Xu, Averaged mappings and the gradient-projection algorithm, Journal of Optimization Theory and Applications 150 (2) (2011) 360-378.

[12] P. Cholamjiak, A generalized forward-backward splitting method for solving quasi inclusion problems in Banach spaces, Numerical Algorithms 71 (4) (2016) 915-932.

[13] W. Cholamjiak, P. Cholamjiak, S. Suantai, An inertial forward-backward splitting method for solving inclusion problems in Hilbert spaces, Journal of Fixed Point Theory and Applications 20 (1) (2018) 42.

[14] Q. Dong, D. Jiang, P. Cholamjiak, Y. Shehu, A strong convergence result involving an inertial forward-backward algorithm for monotone inclusions, Journal of Fixed Point Theory and Applications 19 (4) (2017) 3097-3118.

[15] S.A. Khan, S. Suantai, W. Cholamjiak, Shrinking projection methods involving inertial forward-backward splitting methods for inclusion problems, Revista de la Real Academia de Ciencias Exactas, Fsicas y Naturales. Serie A. Matemticas (2018) 1-12.

[16] S.A. Khan, W. Cholamjiak, K.R. Kazmi, An inertial forward-backward splitting method for solving combination of equilibrium problems and inclusion problems, Computational and Applied Mathematics 37 (5) (2018) 6283-6307.

[17] N. Pholasa, P. Cholamjiak, Y.J. Cho, Modified forward-backward splitting methods for accretive operators in Banach spaces, J. Nonlinear Sci. Appl 9 (2016) 2766-2778.

[18] K. Nakajo, W. Takahashi, Strong convergence theorems for nonexpansive mappings and nonexpansive semigroups, Journal of Mathematical Analysis and Applications, 279 (2) (2003) 372-379.

[19] W. Takahashi, K. Zembayashi, Strong convergence theorem by a new hybrid method for equilibrium problems and relatively nonexpansive mappings, Fixed Point Theory and Applications 2008 (2018) DOI: 10.1155/2008/528476.

[20] L.J. Lin, W. Takahashi, A general iterative method for hierarchical variational inequality problems in Hilbert spaces and applications, Positivity 16 (3) (2012) 429-453.

[21] J.Y. Bello Cruz, T.T. Nghia, On the convergence of the forward-backward splitting method with linesearches, Optimization Methods and Software 31 (6) (2016) 1209-1238.

[22] H.H. Bauschke, P.L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer, New York, 2011.

[23] R.S. Burachik, A.N. Iusem, Set-Valued Mappings and Enlargements of Monotone Operators, in Springer Optimization and Its Applications, Springer-Verlag, US, 2007.

(Received 6 June 2019)
(Accepted 21 August 2019)