Thai Journal of Mathematics
Vol. 18, No. 1 (2020),
Pages 63 - 76

ISSN 1686-0209

A DOUBLE FORWARD-BACKWARD ALGORITHM USING LINESEARCHES FOR MINIMIZATION PROBLEMS

Prasit Cholamjiak[†], **Kunrada Kankam**[†], **Phootares Srinet**[†], **Nattawut Pholasa**[‡],
[†] *School of Science, University of Phayao, Phayao 56000, Thailand,*
e-mail : prasitch2008@yahoo.com (P. Cholamjiak), kunradazzz@gmail.com (K. Kankam),
Pnootres2559@gmail.com (P. srient).
[‡] *Demonstration school University of Phayao, Phayao 56000, Thailand,*
e-mail : nattawut_math@hotmail.com (N. Pholasa).

Abstract In this paper, we introduce a novel forward-backward algorithm involving linesearches for solving nonsmooth optimization problems in Hilbert spaces. The convergence including the complexity are proved under mild conditions. Further, some numerical experiments are tested to show the efficiency and the implementation of our algorithms. It reveals that the proposed method has a good convergence in terms of CPU time and number of iterations.

MSC: 47H09; 47H10
Keywords: minimization problem; forward-backward algorithm; Hilbert space; linesearch

Submission date: 27.10.2019 / Acceptance date: 13.12.2019

1. Introduction

Let H be a real Hilbert space with the inner product $\langle\cdot,\cdot\rangle$ and the induced norm $\|\cdot\|$. Let $f, g : H \to \mathbb{R} \cup \{+\infty\}$ be two proper, lower-semicontinuous and convex functions in which f is Fréchet differentiable on open set containing the domain of g. We are interested in solving problems for the minimization problem:

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

The solution set of this problem (1.1) will be denoted by S_*. Recently problem (1.1) has received much attention due to its applications in optimal control, signal processing, system identification, machine learning, and image analysis; see, e.g. [10, 11, 27]. It is well-known that, for any $\alpha > 0$, x is an optimal solution to problem (1.1) if and only if

$$x = \operatorname{prox}_{\alpha g}(x - \alpha \nabla f(x)). \tag{1.2}$$

Published by The Mathematical Association of Thailand.

The above equation shows an equivalency between convex minimization problems and fixed point problems.This alternative equivalent formulation has played a major role in study of minimization problem. In particular, the minimizers of (1.1) can be approximated by using the following proximal technique:

$$x^{k+1} = \underbrace{\text{prox}_{\alpha_k g}}_{\text{backward step}} \underbrace{(x^k - \alpha_k \nabla f(x^k))}_{\text{forward step}}, \tag{1.3}$$

where α_k is a suitable stepsize. This method is called the forward-backward splitting algorithm. The forward-backward method based on iteration (1.3) has been studied by many authors; see, e.g. [5, 6, 10, 12, 14, 19, 23, 24, 28, 34]. Moreover, scheme (1.3) may reduce to many popular optimization methods as particular cases including the projected gradient method for smooth constrained minimization, the proximal point method, the CQ algorithm for the split feasibility problem, the projected Landweber algorithm for constrained least squares; the iterative soft thresholding algorithm for linear inverse problems; decomposition methods for solving variational inequalities; and the simultaneous orthogonal projection algorithm for the convex feasibility problem; see, e.g. [2, 7, 8, 13, 15, 17, 29, 33, 34] and the references therein. Combettes and Wajs [11] introduced the following iterative sequence which is based on the classical forward-backward iteration (1.3).

Algorithm 1.1. [11] *Given* $\epsilon \in (0, \min\{1, \frac{1}{\beta}\})$ *and let* $x^0 \in \mathbb{R}^N$. *For* $k \geq 1$, *calculate*

$$\begin{aligned} y^k &= x^k - \alpha_k \nabla f(x^k), \\ x^{k+1} &= x^k + \lambda_k(\text{prox}_{\alpha_k g} y^k - x^k), \end{aligned} \tag{1.4}$$

where $\alpha_k \in [\epsilon, \frac{2}{\beta} - \epsilon]$ *and* $\lambda_k \in [\epsilon, 1]$. *Here* β *is the Lipschitz constant of the gradient of* f.

It should be remarked that, in general, the Lipschitz constant is unknown even if the convex objective function is given. To obtain the convergence, it is usually assumed that the gradient of the function is Lipschitz continuous and also the stepsize is bounded below and less than some constants related to the Lipschitz constant. This leads to a difficulty in computing the iterative sequence to a solution. As pointed out, it is interested to study and develop the forward-backward algorithm for solving minimization problems without information of the Lipschitz constant.

In 2016, J.Y. Bello Cruz and T.T.A Nghia [3] proposed the following forward-backward method using the linesearch technique which does not depend on the Lipschitz constant.

Algorithm 1.2. *Given* $\sigma > 0, \theta \in (0,1)$ *and* $\delta \in (0, \frac{1}{2})$. *Let* $x^0 \in \text{dom}g$. *For* $k \geq 1$, *calculate*

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

where $\alpha_k = \sigma\theta^{m_k}$ *and* m_k *is the smallest nonnegative integer such that*

$$\alpha_k \|\nabla f(x^{k+1}) - \nabla f(x^k)\| \leq \delta \|x^{k+1} - x^k\|.$$

It was shown that the sequence $(x^k)_{k\in\mathbb{N}}$ converges weakly to a solution in S_* and $\lim_{k\to\infty} (f+g)(x^k) = \min_{x\in H}(f+g)(x)$. In this work linesearches are used to eliminate the undesired Lipschitz assumption. As far as we observe, the theory of convergence and complexity for the forward-backward is almost complete under such a Lipschitz assumption. However, the Lipschitz condition fails in many natural circumstances; see, e.g. [9].

It is quite interesting to question the convergence of the method without the Lipschitz assumption aforementioned. So, we aim to modify Algorithm 1.2 for solving (1.1) with a new linesearch. The main advantage is that our scheme do not require the information of the Lipschitz constant of the gradient of functions which makes proposed algorithm more practical for computing.

This paper is organized as follows: In section 2, we present some preliminary results that will be used in the proof. In Section 3, we prove the weak convergence and the complexity of the proposed algorithm. In Section 4, we provide some numerical experiments to support our main theorem and also show its efficiency comparing with other methods. In section 5, we complete the paper with the conclusion in this work.

2. Preliminaries

In this section, we recall some basic concepts and lemmas which will be used in our proof. 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. The subdifferential of 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 2.2. [[1], 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. [4] *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 that the proximal operator $\mathrm{prox}_g : H \to \mathrm{dom} g$ is defined by $\mathrm{prox}_g(z) = (Id + \partial g)^{-1}(z), z \in H$. Here Id denotes the identity operator. It is well-known that the proximal operator is single-valued with full domain. Furthermore, note that [3]

$$\frac{z - \mathrm{prox}_{\alpha g}(z)}{\alpha} \in \partial g(\mathrm{prox}_{\alpha g}(z)) \text{ for all } z \in H,\ \alpha > 0. \tag{2.1}$$

Definition 2.4. Let S be a nonempty subset of H. A sequence $(x^k)_{k\in\mathbb{N}}$ in H is said to be quasi-Fejér convergent to S if and only if for all $x \in S$ there exists a positive sequence $(\epsilon_k)_{k\in\mathbb{N}}$ such that $\sum_{k=0}^{\infty} \epsilon_k < +\infty$ and $\|x^{k+1} - x\|^2 \leq \|x^k - x\|^2 + \epsilon_k$ for all $k \in \mathbb{N}$. When $(\epsilon_k)_{k\in\mathbb{N}}$ is a null sequence, we say that $(x^k)_{k\in\mathbb{N}}$ is Fejér convergent to S.

Fact 2.5. [[20],Theorem 4.1] *If $(x^k)_{k\in\mathbb{N}}$ is quasi-Fejér convergent to S, then one has:*
(i) *The sequence $(x^k)_{k\in\mathbb{N}}$ is bounded.*
(ii) *If all weak accumulation points of $(x^k)_{k\in\mathbb{N}}$ belong to S, then $(x^k)_{k\in\mathbb{N}}$ is weakly convergent to a point in S.*

Following [3], 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 $\mathrm{dom} g \subseteq \mathrm{dom} f$.

(A2) The function f is Fréchet differentiable on an open set containing dom$g$. The gradient ∇f is uniformly continuous on any bounded subset of dom$g$ and maps any bounded subset of dom$g$ to a bounded set in H.

Remark 2.6. The second part of (A2) still holds when ∇f is Lipschitz continuous on dom$g$.

3. Main Results

In this section, we present our algorithm and prove its convergence and complexity. Throughout this work, we denote by S_* the solution set of (1.1) and assume that S_* is nonempty.

Algorithm 3.1. *Let $\sigma > 0, \theta \in (0,1)$ and $\delta \in (0, 1/4)$, take $x^0 \in$ dom$g$ and*

$$y^k = \text{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)), \tag{3.1}$$

$$x^{k+1} = \text{prox}_{\alpha_k g}(y^k - \alpha_k \nabla f(y^k)), \tag{3.2}$$

where $\alpha_k = \sigma\theta^{m_k}$ and m_k is the smallest nonnegative integer such that

$$\alpha_k \|\nabla f(y^k) - \nabla f(x^k)\| \leq \delta \|y^k - x^k\| \tag{3.3}$$

and

$$\alpha_k \|\nabla f(x^{k+1}) - \nabla f(y^k)\| \leq \delta \|x^{k+1} - y^k\|. \tag{3.4}$$

Lemma 3.2. [[3], Lemma 3.1] *The linesearch (3.3) and (3.4) stops after finitely many steps.*

Lemma 3.3. *Let (x^k) be defined by Algorithm 3.1. For all $k \in \mathbb{N}$ and $x \in$ dom$g$, we have*
(i) $\|x^k - x\|^2 - \|x^{k+1} - x\|^2 \geq 2\alpha_k[(f+g)(x^{k+1}) - (f+g)(x) + (f+g)(y^k) - (f+g)(x)]$
$+ (1-2\delta)\|x^k - y^k\|^2 + (1-2\delta)\|x^{k+1} - y^k\|^2$;
(ii) $((f+g)(x^k))_{k\in\mathbb{N}}$ *is decreasing.*

Proof. First we will prove (i). From (2.1) and (3.1), we have

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

By the convexity of g, it follows that

$$g(x) - g(y^k) \geq \langle \frac{x^k - y^k}{\alpha_k} - \nabla f(x^k), x - y^k \rangle, \forall x \in \text{dom} g. \tag{3.5}$$

From (2.1) and (3.2), we obtain

$$\frac{y^k - x^{k+1}}{\alpha_k} - \nabla f(y^k) = \frac{y^k - \text{prox}_{\alpha_k g}(y^k - \alpha_k \nabla f(y^k))}{\alpha_k} - \nabla f(y^k) \in \partial g(x^{k+1}).$$

By the convexity of g, we also have

$$g(x) - g(x^{k+1}) \geq \langle \frac{y^k - x^{k+1}}{\alpha_k} - \nabla f(y^k), x - x^{k+1} \rangle, \forall x \in \text{dom} g. \tag{3.6}$$

By Fact 2.2, we see that

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

For any $x \in \mathrm{dom} g \subseteq \mathrm{dom} f$ and $y = x^k$ in (3.7), we see that

$$f(x) - f(x^k) \geq \langle \nabla f(x^k), x - x^k \rangle. \tag{3.8}$$

Taking $y = y^k$ in (3.7), we obtain

$$f(x) - f(y^k) \geq \langle \nabla f(y^k), x - y^k \rangle. \tag{3.9}$$

Using (3.5), (3.6), (3.8) and (3.9), we have

$$\begin{aligned}
& g(x) - g(x^{k+1}) + g(x) - g(y^k) + f(x) - f(y^k) + f(x) - f(x^k) \\
\geq\ & \langle \frac{y^k - x^{k+1}}{\alpha_k} - \nabla f(y^k), x - x^{k+1} \rangle + \langle \frac{x^k - y^k}{\alpha_k} - \nabla f(x^k), x - y^k \rangle + \langle \nabla f(y^k), x - y^k \rangle \\
& + \langle \nabla f(x^k), x - x^k \rangle \\
=\ & \frac{1}{\alpha_k}\langle y^k - x^{k+1}, x - x^{k+1} \rangle + \langle \nabla f(y^k), x^{k+1} - x \rangle + \frac{1}{\alpha_k}\langle x^k - y^k, x - y^k \rangle \\
& + \langle \nabla f(x^k), y^k - x \rangle + \langle \nabla f(y^k), x - y^k \rangle + \langle \nabla f(x^k), x - x^k \rangle \\
=\ & \frac{1}{\alpha_k}\langle y^k - x^{k+1}, x - x^{k+1} \rangle + \frac{1}{\alpha_k}\langle x^k - y^k, x - y^k \rangle + \langle \nabla f(y^k), x^{k+1} - y^k \rangle \\
& + \langle \nabla f(x^k), y^k - x^k \rangle \\
=\ & \frac{1}{\alpha_k}\langle y^k - x^{k+1}, x - x^{k+1} \rangle + \frac{1}{\alpha_k}\langle x^k - y^k, x - y^k \rangle + \langle \nabla f(y^k) - \nabla f(x^{k+1}), x^{k+1} - y^k \rangle \\
& + \langle \nabla f(x^{k+1}), x^{k+1} - y^k \rangle + \langle \nabla f(x^k) - \nabla f(y^k), y^k - x^k \rangle + \langle \nabla f(y^k), y^k - x^k \rangle \\
\geq\ & \frac{1}{\alpha_k}\langle y^k - x^{k+1}, x - x^{k+1} \rangle + \frac{1}{\alpha_k}\langle x^k - y^k, x - y^k \rangle - \|\nabla f(y^k) - \nabla f(x^{k+1})\|\|x^{k+1} - y^k\| \\
& + f(x^{k+1}) - f(y^k) - \|\nabla f(x^k) - \nabla f(y^k)\|\|y^k - x^k\| + f(y^k) - f(x^k),
\end{aligned} \tag{3.10}$$

where the last inequality follows from Fact 2.2. This shows that

$$\begin{aligned}
& \frac{1}{\alpha_k}[\langle y^k - x^{k+1}, x^{k+1} - x \rangle + \langle x^k - y^k, y^k - x \rangle] \\
\geq\ & (f+g)(x^{k+1}) - (f+g)(x) + (f+g)(y^k) - (f+g)(x) - \|\nabla f(y^k) - \nabla f(x^{k+1})\|\|x^{k+1} - y^k\| \\
& - \|\nabla f(x^k) - \nabla f(y^k)\|\|y^k - x^k\|.
\end{aligned}$$

This shows that, by (3.3) and (3.4),

$$\begin{aligned}
& \frac{1}{\alpha_k}[\langle y^k - x^{k+1}, x^{k+1} - x \rangle + \langle x^k - y^k, y^k - x \rangle] \\
\geq\ & (f+g)(x^{k+1}) - (f+g)(x) + (f+g)(y^k) - (f+g)(x) - \frac{\delta}{\alpha_k}\|y^k - x^{k+1}\|\|x^{k+1} - y^k\| \\
& - \frac{\delta}{\alpha_k}\|x^k - y^k\|\|y^k - x^k\|. \\
=\ & (f+g)(x^{k+1}) - (f+g)(x) + (f+g)(y^k) - (f+g)(x) - \frac{\delta}{\alpha_k}\|y^k - x^{k+1}\|^2 \\
& - \frac{\delta}{\alpha_k}\|x^k - y^k\|^2.
\end{aligned} \tag{3.11}$$

We know that

$$2\langle y^k - x^{k+1}, x^{k+1} - x \rangle = \|y^k - x\|^2 - \|y^k - x^{k+1}\|^2 - \|x^{k+1} - x\|^2 \tag{3.12}$$

and

$$2\langle x^k - y^k, y^k - x\rangle = \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - x\|^2. \tag{3.13}$$

Replacing (3.12) and (3.13) in (3.11), we have

$$\begin{aligned}\|x^k - x\|^2 - \|x^{k+1} - x\|^2 \quad &\geq \quad 2\alpha_k[(f+g)(x^{k+1}) - (f+g)(x) + (f+g)(y^k) - (f+g)(x)] \\ &\quad +(1-2\delta)\|x^k - y^k\|^2 + (1-2\delta)\|x^{k+1} - y^k\|^2. \qquad (3.14)\end{aligned}$$

This proves (i). Setting $x = x^k$ in (3.14), we have

$$\begin{aligned}-\|x^{k+1} - x^k\|^2 \quad &\geq \quad 2\alpha_k[(f+g)(x^{k+1}) - (f+g)(x) + (f+g)(y^k) - (f+g)(x)] \\ &\quad +(1-2\delta)\|x^k - y^k\|^2 + (1-2\delta)\|x^{k+1} - y^k\|^2. \qquad (3.15)\end{aligned}$$

Also setting $x = y^k$ in (3.14), we have

$$\begin{aligned}& \|x^k - y^k\|^2 - \|x^{k+1} - y^k\|^2 \qquad (3.16) \\ \geq \quad & 2\alpha_k[(f+g)(x^{k+1}) - (f+g)(x) + (f+g)(y^k) - (f+g)(x)] \\ & +(1-2\delta)\|x^k - y^k\|^2 + (1-2\delta)\|x^{k+1} - y^k\|^2. \qquad (3.17)\end{aligned}$$

Summing (3.15) and (3.16), it follows that

$$\begin{aligned}& -\|x^{k+1} - x^k\|^2 - \|x^k - y^k\|^2 - \|x^{k+1} - y^k\|^2 \\ \geq \quad & 4\alpha_k[(f+g)(x^{k+1}) - (f+g)(x) + (f+g)(y^k) - (f+g)(x)] \\ & +(2-4\delta)\|x^k - y^k\|^2 + (2-4\delta)\|x^{k+1} - y^k\|^2. \qquad (3.18)\end{aligned}$$

This shows that

$$(f+g)(x^{k+1}) - (f+g)(x^k) \leq -\frac{1}{4\alpha_k}[(3-4\delta)\|y^k - x^{k+1}\|^2 + (1-4\delta)\|x^k - y^k\|^2 + \|x^{k+1} - x^k\|^2].$$

Hence, $((f+g)(x^k))_{k\in\mathbb{N}}$ is decreasing. We thus complete the proof of (ii). ■

Next, we prove weak convergence theorem of Algorithm 3.1.

Theorem 3.4. *Let $(x^k)_{k\in\mathbb{N}}$ and $(\alpha_k)_{k\in\mathbb{N}}$ be the sequences generated by Algorithm 3.1. If there exists $\alpha > 0$ such that $\alpha_k \geq \alpha > 0$ for all $k \in \mathbb{N}$, then $(x^k)_{k\in\mathbb{N}}$ is weakly convergent to a point in S_*. Moreover,*

$$\lim_{k\to\infty}(f+g)(x^k) = \min_{x\in H}(f+g)(x). \tag{3.19}$$

Proof. Let $x_* \in S_*$. Using Lemma 3.3(i), we see that

$$\begin{aligned}& \|x^k - x_*\|^2 - \|x^{k+1} - x_*\|^2 \qquad (3.20) \\ \geq \quad & 2\alpha_k[(f+g)(x^{k+1}) - (f+g)(x_*) + (f+g)(y^k) - (f+g)(x_*)] \\ & +(1-2\delta)\|y^k - x^{k+1}\|^2 + (1-2\delta)\|x^k - y^k\|^2 \\ \geq \quad & (1-2\delta)\|y^k - x^{k+1}\|^2 + (1-2\delta)\|x^k - y^k\|^2 \\ \geq \quad & 0. \qquad (3.21)\end{aligned}$$

Then the sequence $(x^k)_{k\in\mathbb{N}}$ is Féjer convergent to S_*. Hence, by Fact 2.5 (i), it is bounded. From (3.20), we get

$$\begin{aligned} 0 &\le 2\alpha_k[(f+g)(x^{k+1})-(f+g)(x_*)+(f+g)(y^k)-(f+g)(x_*)] \\ &\le \|x^k-x_*\|^2-\|x^{k+1}-x_*\|^2 \\ &= (\|x^k-x_*\|-\|x^{k+1}-x_*\|)(\|x^k-x_*\|+\|x^{k+1}-x_*\|) \\ &\le 2M(\|x^k-x_*\|-\|x^{k+1}-x_*\|) \\ &\le 2M\|x^k-x^{k+1}\|, \text{ where } M:=\sup_{k\in\mathbb{N}}\{\|x^k-x_*\|\}<+\infty. \end{aligned} \tag{3.22}$$

It follows that

$$2\alpha_k[(f+g)(x^{k+1})-(f+g)(x_*)]\le 2M\|x^k-x^{k+1}\|.$$

Hence

$$(f+g)(x^{k+1})-(f+g)(x_*)\le \frac{M\|x^k-x^{k+1}\|}{\alpha_k}. \tag{3.23}$$

Since $(\|x^k-x_*\|)_{k\in\mathbb{N}}$ is convergent, by (3.20), we have $\|y^k-x^{k+1}\|\to 0$ and $\|y^k-x^k\|\to 0$ as $k\to\infty$. Thus $\|x^k-x^{k+1}\|\to 0$ as $k\to\infty$.

Since $(x^k)_{k\in\mathbb{N}}$ is bounded, the set of its weak accumulation points is nonempty. Take any weak accumulation point $\bar{x}$ of $(x^k)_{k\in\mathbb{N}}$. So there is a subsequence $(x^{n_k})_{k\in\mathbb{N}}$ of $(x^k)_{k\in\mathbb{N}}$ weakly converging to $\bar{x}$. Moreover, x^{n_k+1} also weakly converges to $\bar{x}$. Since $(x^{n_k})_{k\in\mathbb{N}}$ is bounded and $\|x^{n_k+1}-y^{n_k}\|\to 0$, we get from Assumption (A2) that

$$\lim_{k\to\infty}\|\nabla f(x^{n_k+1})-\nabla f(y^{n_k})\|=0. \tag{3.24}$$

Since $x^{n_k+1}=\operatorname{prox}_{\alpha_{n_k}g}(y^{n_k}-\alpha_{n_k}\nabla f(y^{n_k}))$, it follows from (2.1) that

$$\frac{y^{n_k}-\alpha_{n_k}\nabla f(y^{n_k})-x^{n_k+1}}{\alpha_{n_k}}\in\partial g(x^{n_k+1}) \tag{3.25}$$

which implies that

$$\frac{y^{n_k}-x^{n_k+1}}{\alpha_{n_k}}+\nabla f(x^{n_k+1})-\nabla f(y^{n_k})\in\nabla f(x^{n_k+1})+\partial g(x^{n_k+1})\in\partial(f+g)(x^{n_k+1}) \tag{3.26}$$

By passing $k\to\infty$ in (3.26), we get from (3.24) and Fact 2.5 that $0\in\partial(f+g)(\bar{x})$. Thus $\bar{x}\in S_*$. Furthermore, since the sequence $((f+g)(x^k))_{k\in\mathbb{N}}$ is decreasing due to Lemma 3.3(ii), (3.19) is a consequence of (3.23). ■

We next discuss the complexity of Algorithm 3.1.

Theorem 3.5. *Let $(x^k)_{k\in\mathbb{N}}$ and $(\alpha_k)_{k\in\mathbb{N}}$ be the sequences generated by Algorithm 3.1. If there exists $\alpha>0$ such that $\alpha_k\ge\alpha>0$ for all $k\in\mathbb{N}$, then*

$$(f+g)(x^k)-\min_{x\in H}(f+g)(x)\le\frac{1}{2\alpha}\frac{[\operatorname{dist}(x^0,S_*)]^2}{k}.$$

Proof. Let $x_* \in S_*$. Then we have, by Lemma 3.3(i),

$$\begin{aligned} 0 &\geq 2\alpha_l[(f+g)(x_*) - (f+g)(x^{l+1}) + (f+g)(x_*) - (f+g)(y^l)] \\ &\geq \|x^{l+1} - x_*\|^2 - \|x^l - x_*\|^2 + (1-2\delta)\|y^l - x^{l+1}\|^2 + (1-2\delta)\|y^l - x^l\|^2 \\ &\geq \|x^{l+1} - x_*\|^2 - \|x^l - x_*\|^2 \end{aligned} \tag{3.27}$$

for any $l \in \mathbb{N}$. Since $\alpha_l \geq \alpha$, we get from (3.27) that

$$0 \geq (f+g)(x_*) - (f+g)(x^{l+1}) \geq \frac{1}{2\alpha}(\|x^{l+1} - x_*\|^2 - \|x^l - x_*\|^2). \tag{3.28}$$

Summing the above inequality over $l = 0, 1, ..., k-1$ implies that

$$k(f+g)(x_*) - \sum_{l=0}^{k-1}(f+g)(x^{l+1}) \geq \frac{1}{2\alpha}(\|x^k - x_*\|^2 - \|x^0 - x_*\|^2). \tag{3.29}$$

Since $(f+g)(x^l)$ is decreasing by Lemma 3.3(ii), it follows that

$$k[(f+g)(x^k) - (f+g)(x_*)] \leq \frac{1}{2\alpha}(\|x_* - x^0\|^2 - \|x^k - x_*\|^2) \leq \frac{1}{2\alpha}\|x_* - x^0\|^2. \tag{3.30}$$

From (3.30), we obtain

$$(f+g)(x^k) - \min_{x\in H}(f+g)(x) \leq \frac{1}{2\alpha}\inf_{y\in S_*}\frac{\|y - x^0\|^2}{k} = \frac{1}{2\alpha}\frac{[\text{dist}(x^0, S_*)]^2}{k}. \tag{3.31}$$

This completes the proof. ∎

4. Numerical Experiments

In this section, we give some numerical examples to the signal recovery in compressed sensing. We provide a comparison among Algorithm 1.1, Algorithm 1.2 and Algorithm 3.1. 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 m 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 $g(x) = \lambda\|x\|_1$.

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

$$E_k = \|x^{k+1} - x^k\|_2 < 10^{-7}.$$

The step size α_k in Algorithm (1.1) is $\frac{0.2}{\|A\|^2}$ and $\lambda_k = 1$ and $\sigma = 0.02, \theta = 0.3$, and $\delta = \frac{1}{6}$ in both Algorithm (1.2) and Algorithm (3.1). We denote by CPU the time using in CPU and Iter the number of iterations. The numerical results are reported as follows:

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 1.1	0.1609	4260	0.5498	4447
	Algorithm 1.2	0.0704	432	0.2283	352
	Algorithm 3.1	0.0631	234	0.1789	203
m=30	Algorithm 1.1	0.1835	4569	0.6916	5450
	Algorithm 1.2	0.0729	470	0.2692	418
	Algorithm 3.1	0.0605	256	0.2386	235
m=40	Algorithm 1.1	0.2279	5883	1.0277	7888
	Algorithm 1.2	0.0841	545	0.4062	587
	Algorithm 3.1	0.0757	304	0.3474	313
m=50	Algorithm 1.1	0.3673	9899	1.1935	9480
	Algorithm 1.2	0.1169	819	0.4463	679
	Algorithm 3.1	0.0919	437	0.3806	365
m=60	Algorithm 1.1	0.4382	11079	1.4771	10232
	Algorithm 1.2	0.1527	1045	0.4697	756
	Algorithm 3.1	0.1127	547	0.3919	403

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 [22]. and Algorithm 1.2 of [3].

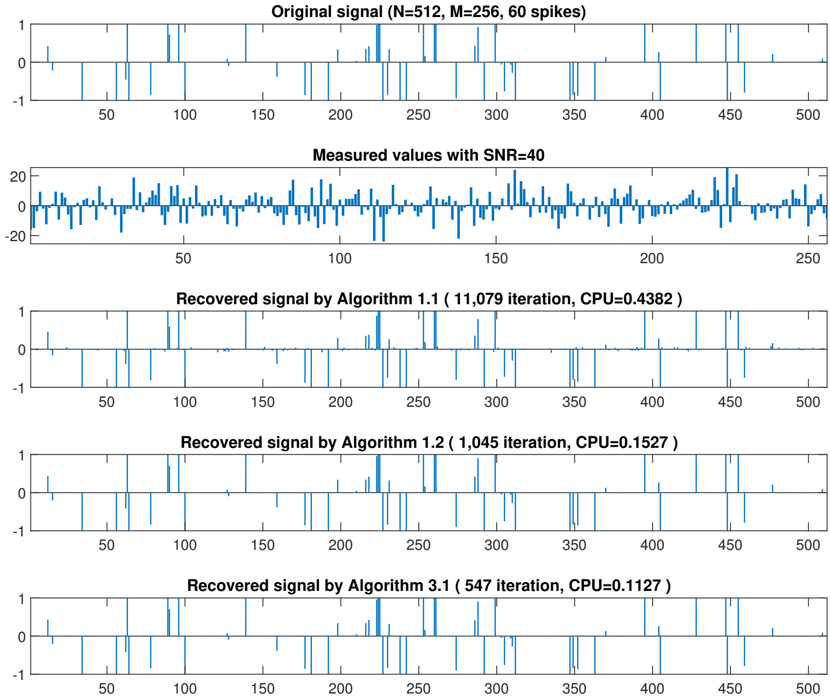

FIGURE 1. From top to bottom: original signal, observation data, recovered signal by Algorithm 1.1, Algorithm 1.2 and Algorithm 3.1 with $N = 512$ and $M = 256$, respectively.

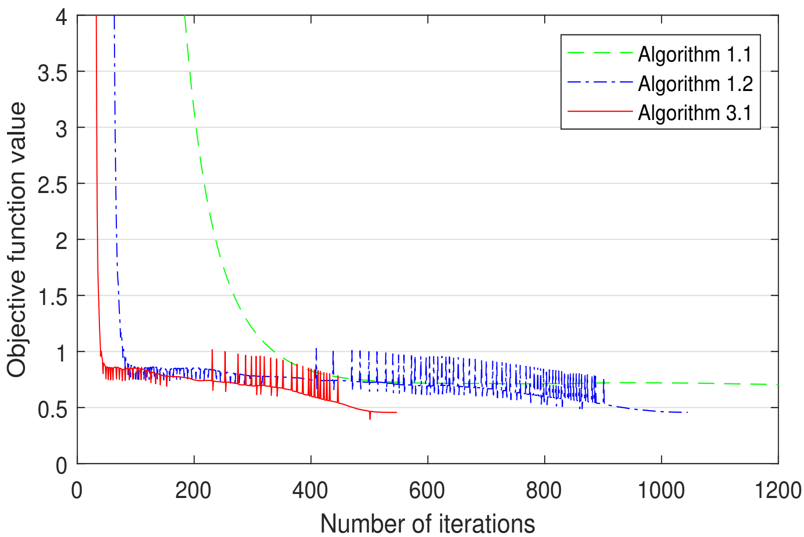

FIGURE 2. The objective function value versus number of iterations in case N=512, M=256.

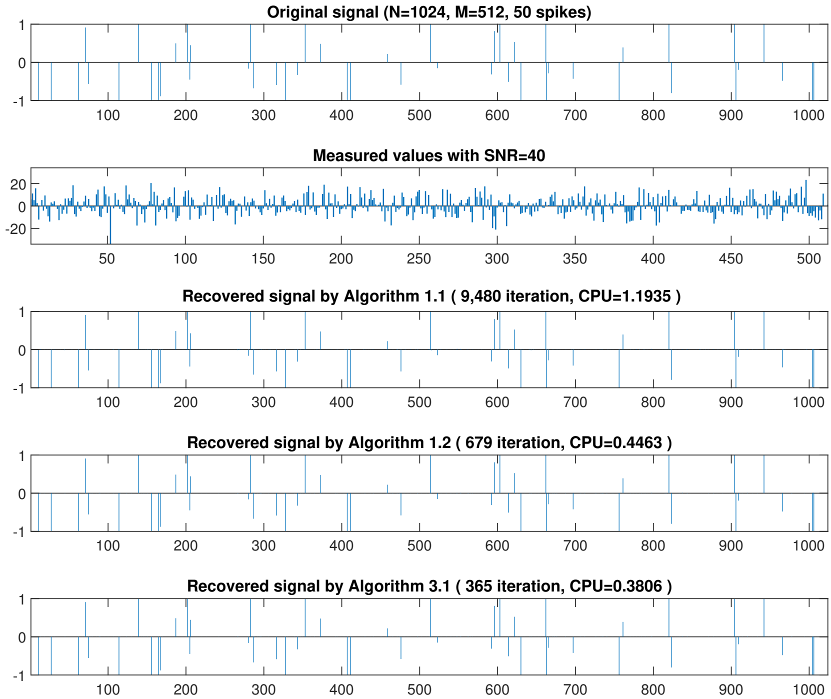

FIGURE 3. From top to bottom: original signal, observation data, recovered signal by Algorithm 1.1, Algorithm 1.2 and Algorithm 3.1 with $N = 1024$ and $M = 512$, respectively.

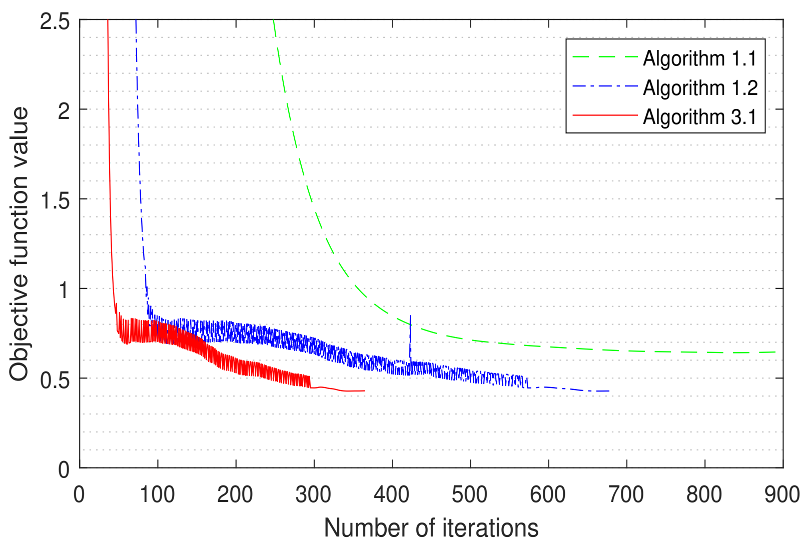

FIGURE 4. The objective function value versus number of iterations in case N=1024, M=512.

5. Conclusion

In this work, we discuss the modified forward-backward splitting method involving new linesearches for solving minimization problems of two convex functions. Our algorithm need not compute the Lipschitz constant of the gradient of functions. Also, we show in a simple and novel way how the sequence generated by the method weakly converges to a solution of the minimization problem. We also discuss the complexity of our defined algorithm. All the results are compared, in compressed sensing, with forward-backward method of [11] in Algorithm 1.1 and forward-backward method with linesearches [3] in Algorithm 1.2. One interesting and challenge topic is to design new linesearches that reduce the number of iterations and the CPU time in computing.

Acknowledgments

The authors would like to thank reviewers and editor for their comments on the manuscript which helped us very much in improving and presenting the original version of this paper. This work is supported by School of Science, University of Phayao, Phayao, Thailand (Grant No. UoE 62001).

References

[1] H.H. Bauschke, P.L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces (Vol. 408), New York, Springer, 2011.

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

[3] 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.

[4] R.S. Burachik, A.N. Iusem, Set-valued mappings and enlargements of monotone operators (Vol. 8), Springer Science Business Media, 2007.

[5] G.H. Chen, R.T. Rockafellar, Convergence rates in forward-backward splitting, SIAM Journal on Optimization 7(2)(1997) 421–444.

[6] 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.

[7] P. Cholamjiak, The modified proximal point algorithm in CAT(0) spaces, Optim. Lett. 9(2015) 1401–1410.

[8] P. Cholamjiak, Y.J. Cho, S. Suantai, Composite iterative schemes for maximal monotone operators in reflexive Banach spaces, Fixed Point Theory Appl. 2011 (2011); https://doi.org/10.1186/1687-1812-2011-7.

[9] P.L. Combettes, J.C. Pesquet, A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery, IEEE J.Sel. Top. Signal Process 1(2007) 564–574.

[10] 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, NY. 2011, 185–212.

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

[12] A. Padcharoen, P. Sukprasert, Nonlinear operators as concerns convex programming and applied to signal processing, Mathematics 7(9)(2019) 866; https://doi.org/10.3390/math7090866.

[13] P.L. Combettes, Inconsistent signal feasibility problems: Least-squares solutions in a product space, IEEE Transactions on Signal Process 42(1994) 2955–2966.

[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] I. Daubechies, M. Defrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, Commun. Pure Appl. Math. 57(2004) 1413–1457.

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

[17] E. Eicke, Iteration methods for convexly constrained ill-posed problems in Hilbert space, Numer. Funct. Anal. Optim. 13(1992) 413–429.

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

[19] 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. Matemáticas (2018) 1–12; https://doi.org/10.1007/s13398-018-0504-1.

[20] A.N. Iusem, B.F. Svaiter, M. Teboulle, Entropy-like proximal methods in convex programming, Mathematics of Operations Research 19(4)(1994) 790–814.

[21] P. Latafat, P. Patrinos, Asymmetric forward-backward-adjoint splitting for solving monotone inclusions involving three operators, Computational Optimization and Applications 68(1)(2017) 57–93.

[22] 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.

[23] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM Journal on Numerical Analysis 16(6)(1979) 964–979.

[24] D.A. Lorenz, T. Pock, An inertial forward-backward algorithm for monotone inclusions, Journal of Mathematical Imaging and Vision 51(2)(2015) 311–325.

[25] B. Martinet, Brvèv communication. Régularisation d'inéquations variationnelles par approximations successives, Revue franaise d'informatique et de recherche opérationnelle. Série rouge 4(R3)(1970) 154–158.

[26] C. Molinari, J. Peypouquet, Lagrangian Penalization Scheme with Parallel Forward-Backward Splitting, Journal of Optimization Theory and Applications 177(2)(2018) 413–447.

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

[28] G.B. Passty, Ergodic convergence to a zero of the sum of monotone operators in Hilbert space, Journal of Mathematical Analysis and Applications 72(1979) 383–290.

[29] R.T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming, Math. Oper. Res. 1(1976) 97–116.

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

[31] Y. Shehu, Iterative approximations for zeros of sum of accretive operators in Banach spaces, Journal of Function Spaces 2016(2016); http://dx.doi.org/10.1155/2016/5973468.

[32] Y. Shehu, G. Cai, Strong convergence result of forward-backward splitting methods for accretive operators in banach spaces with applications, Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas 112(1)(2018) 71–87.

[33] R. Suparatulatorn, P. Cholamjiak, S. Suantai, On solving the minimization problem and the fixed point problem for nonexpansive mappings in CAT(0) spaces, Optim. Meth. Softw. 32(2017) 182–192.

[34] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, SIAM Journal on Control and Optimization 38(2)(2000) 431–446.

[35] Y. Wang, F. Wang, Strong convergence of the forward-backward splitting method with multiple parameters in Hilbert spaces, Optimization 67(4)(2018) 493–505.

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

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