Suantai et al. *Advances in Difference Equations* (2021) 2021:265
https://doi.org/10.1186/s13662-021-03422-9

Advances in Difference Equations
a SpringerOpen Journal

REVIEW Open Access

Novel forward–backward algorithms for optimization and applications to compressive sensing and image inpainting

Suthep Suantai[1,2], Muhammad Aslam Noor[3], Kunrada Kankam[4] and Prasit Cholamjiak[4*]

*Correspondence:
prasitch2008@yahoo.com
[4]School of Science, University of Phayao, Phayao 56000, Thailand
Full list of author information is available at the end of the article

Abstract

The forward–backward algorithm is a splitting method for solving convex minimization problems of the sum of two objective functions. It has a great attention in optimization due to its broad application to many disciplines, such as image and signal processing, optimal control, regression, and classification problems. In this work, we aim to introduce new forward–backward algorithms for solving both unconstrained and constrained convex minimization problems by using linesearch technique. We discuss the convergence under mild conditions that do not depend on the Lipschitz continuity assumption of the gradient. Finally, we provide some applications to solving compressive sensing and image inpainting problems. Numerical results show that the proposed algorithm is more efficient than some algorithms in the literature. We also discuss the optimal choice of parameters in algorithms via numerical experiments.

MSC: 65K05; 90C25; 90C30

Keywords: Forward–backward algorithm; Compressive sensing; Image inpainting; Minimization problem

1 Introduction

In a real Hilbert space H, the unconstrained minimization problem of the sum of two convex functions is modeled in the following form:

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

where $f,g:H\to\mathbb{R}\cup\{+\infty\}$ are proper lower semicontinuous convex functions. It is well known that (1.1) is equivalent to the problem of finding the zero of subdifferentials of $f+g$ at x. This problem is called the variational inclusion problem, see [19]. We denote by $\operatorname{argmin}(f+g)$ the solution set of (1.1). If f is differentiable on H, then (1.1) can be described by the fixed point equation

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

where $\alpha > 0$, and prox_g is the proximal operator of g defined by $\operatorname{prox}_g = (\mathrm{Id} + \partial g)^{-1}$, where Id denotes the identity operator in H, and ∂g is the subdifferential of g. In this connection, we can define a simple splitting method

$$x^{k+1} = \underbrace{\operatorname{prox}_{\alpha_k g}}_{\text{backward step}} \underbrace{(\mathrm{Id} - \alpha_k \nabla f)}_{\text{forward step}}\bigl(x^k\bigr), \quad k \geq 0, \tag{1.3}$$

where α_k is a suitable stepsize. This method is often called the forward–backward algorithm. Due to its simplicity and efficiency, there have been many modifications of (1.3) in the literature; see, for example, [1, 6–8, 12, 14]. The relaxed version of (1.3) was proposed by Combettes and Wajs [9] as follows.

Algorithm 1.1 ([9]) Given $\varepsilon \in (0, \min\{1, \frac{1}{\alpha}\})$, let $x^0 \in \mathbb{R}^N$ and, for $k \geq 1$,

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

where $\alpha_k \in [\varepsilon, \frac{2}{\alpha} - \varepsilon]$, $\lambda_k \in [\varepsilon, 1]$, and α is the Lipschitz constant of the gradient of f.

Based on the fixed point concept, there have been many optimization algorithms and fixed point algorithms for solving such problems; see [13, 15–17, 21]. In 2016, Cruz and Nghia [3] introduced a new forward–backward method using the linesearch technique. This method does not require the Lipschitz constant in computation.

Algorithm 1.2 Given $\sigma > 0, \theta \in (0,1)$, and $\delta \in (0, \frac{1}{2})$, let $x^0 \in \operatorname{dom} g$ and, for $k \geq 0$, calculate

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

where $\alpha_k = \sigma\theta^{m_k}$ with m_k the smallest nonnegative integer satisfying the following linesearch rule:

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

It was proved that the sequence (x^k) converges weakly to a minimizer of $f + g$ under suitable conditions.

In practical applications, many problems in real world such as image inpainting can be modeled as a subproblem. To investigate them, we suggest a projected forward–backward algorithm for solving the constrained convex minimization problem modeled as follows:

$$\min_{x \in \Omega} \bigl(f(x) + g(x)\bigr), \tag{1.5}$$

where Ω is a nonempty closed convex subset of H, f and g are convex functions on H, and f is differentiable on H.

In variational theory, Tseng [23] introduced the modified forward–backward splitting algorithms for finding zeros of the sum of two monotone operators. Let $X \subset \operatorname{dom} A$ be a closed convex set.

Algorithm 1.3 Given $x^0 \in \operatorname{dom} A$ and $\alpha_k \in (0, +\infty)$, calculate

$$\begin{aligned} y^k &= (\mathrm{Id} + \alpha_k B)^{-1}(\mathrm{Id} - \alpha_k A)\big(x^k\big), \\ x^{k+1} &= P_X\big[y^k - \alpha_k\big(A\big(y^k\big) - A\big(x^k\big)\big)\big], \end{aligned} \tag{1.6}$$

where A is L-Lipschitz continuous on $X \cup \operatorname{dom} B$, and $\alpha_k \in (0, 1/L)$. It was proved that (x^k) converges weakly to zeros of $A + B$ that are also contained in X.

Most of the work related to two convex minimization problems usually assume the Lipschitz condition on the gradient of f. This restriction can be relaxed by using a linesearch technique. So we suggest new forward–backward algorithms to solve the unconstrained and constrained convex minimization problems, which are based on a new linesearch technique [14]. Then we prove weak convergence theorems for the proposed algorithm. As applications, we apply our main results to solving compressed sensing and image inpainting problems. Then we compare the performance of our algorithms with Algorithms 1.1 and 1.2. Moreover, we discuss numerical results of the comparative analysis to show the optimal choice of parameters.

The content is organized as follows. In Sect. 2, we recall some the useful concepts. In Sect. 3, we establish the main theorem on our algorithms. In Sect. 4, we give numerical experiments to support the convergence of our algorithms. Finally, in Sect. 5, we end this paper by conclusions.

2 Preliminaries

In this section, we give some definitions and lemmas that play an essential role in our analysis. Let H be a real Hilbert space equipped with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. Let $h : H \to \bar{\mathbb{R}}$ be a proper lower semicontinuous convex function. We use the following notations:

- $\rightharpoonup$ denotes the weak convergence.
- $\operatorname{dom} h := \{x \in H | h(x) < +\infty\}$ denotes the domain of h.
- $\operatorname{Gph}(A) \in H \times H = \{(x, y) : y \in Ax\}$, where $A : H \to 2^H$ is a multivalued operator, denotes the graph of A.
- $\omega_w(x^k) = \{x : \exists (x^{k_n}) \subset (x^k) \text{ such that } x^{k_n} \rightharpoonup x\}$ denotes the set of all weak limit points.
- $F(T) = \{x \in C : x = Tx\}$ denotes the set of fixed points of $T : C \to C$.

We recall the following definitions:

(1) A mapping $T : H \to H$ is said to be *nonexpansive* if, for all $x, y \in H$,

$$\|Tx - Ty\| \leq \|x - y\|.$$

(2) A mapping $T : H \to H$ is said to be *firmly nonexpansive* if, for all $x, y \in H$,

$$\|Tx - Ty\|^2 \leq \langle x - y, Tx - Ty \rangle.$$

(3) A mapping $T : H \to H$ is said to be *monotone* if, for all $x, y \in H$,

$$\langle x - y, Tx - Ty \rangle \geq 0.$$

(4) An operator $A : H \to 2^H$ is said to be *maximal monotone* if there is no monotone operator $B : H \to 2^H$ such that $\mathrm{Gph}(B)$ properly contains $\mathrm{Gph}(A)$, that is, for every $(x, u) \in H \times H$,

$$(x, u) \in \mathrm{Gph}(A) \quad \Longleftrightarrow \quad \langle x - y, Ax - Ay \rangle \geq 0$$

for all $(y, v) \in \mathrm{Gph}(A)$.

(5) A function $h : H \to \mathbb{R}$ is said to be *convex* if

$$h\big(\lambda x + (1 - \lambda)y\big) \leq h(x) + (1 - \lambda)h(y)$$

for all $\lambda \in (0, 1)$ and $x, y \in H$.

(6) A differentiable function h is convex if and only if

$$h(x) + \big\langle \nabla h(x), y - x \big\rangle \leq h(y)$$

for all $y \in H$.

(7) An element $g \in H$ is said to be a *subgradient* of $h : H \to \mathbb{R}$ at x if

$$h(x) + \langle g, y - x \rangle \leq h(y)$$

for all $y \in H$.

(8) The *subdifferential* of h at x is defined by

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

(9) A function $f : H \to \mathbb{R}$ is said to be weakly lower semicontinuous at x if $x_n \rightharpoonup x$ implies

$$f(x) \leq \liminf_{n \to \infty} f(x_n).$$

(10) A *projection* of x onto a nonempty, closed and convex subset C of H is defined by

$$P_C x := \operatorname*{argmin}_{y \in C} \|x - y\|^2.$$

(11) The proximal operator $\mathrm{prox}_g : H \to H$ of g is defined by

$$\mathrm{prox}_g(z) = (\mathrm{Id} + \partial g)^{-1}(z), \quad z \in H.$$

We know that proximal operator is single-valued with full domain. Moreover, from [3] we have

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

Lemma 2.1 ([2]) *Let C be a nonempty closed convex subset of a real Hilbert space H. Then for any $x \in H$, the following statements hold*:

(i) $\langle x - P_C x, y - P_C x\rangle \le 0$ *for all* $y \in C$;
(ii) $\|P_C x - P_C y\|^2 \le \langle P_C x - P_C y, x - y\rangle$ *for all* $x, y \in H$;
(iii) $\|P_C x - y\|^2 \le \|x - y\|^2 - \|P_C x - x\|^2$ *for all* $y \in C$.

Lemma 2.2 ([4]) *The subdifferential operator* ∂h *of a convex function* 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, that is, if a sequence* $(x^k, v^k) \subset \mathrm{Gph}(\partial h)$ *is such that* (x^k) *converges weakly to* x *and* (v^k) *converges strongly to* v, *then* $(x, v) \in \mathrm{Gph}(\partial h)$.

Lemma 2.3 ([20]) *Let* H *be a real Hilbert space. Let* C *be a nonempty closed convex subset of* H, *and let* $T : C \to C$ *be a nonexpansive mapping such that* $F(T) \neq \emptyset$. *If* $(x^k) \subset C$, $x^k \rightharpoonup z$, *and* $\|Tx^k - x^k\| \to 0$, *then* $Tz = z$.

Lemma 2.4 ([2]) *Let* H *be a real Hilbert space. Let* S *be a nonempty closed convex subset of* H, *and let* (x^k) *be a sequence in* H *satisfying*:
(i) $\lim_{k\to\infty} \|x^k - x\|$ *exists for each* $x \in S$;
(ii) $\omega_w(x^k) \subset S$.
Then (x^k) *weakly converges to an element of* S.

3 Main results

In this section, we assume that the set S_* of all solutions of problem (1.1) is nonempty. We propose new algorithms by combining a new linesearch technique and prove weak convergence theorems. We assume that

(1) $f, g : H \to \mathbb{R} \cup \{+\infty\}$ are proper lower semicontinuous convex functions, f is differentiable on H and
(2) the gradient ∇f is uniformly continuous and bounded on bounded subsets of H.

Note that the latter condition holds if ∇f is Lipschitz continuous on H.

Algorithm 3.1 Given $\sigma > 0$, $\theta \in (0,1)$, $\gamma \in (0,2)$, and $\delta \in (0, \frac{1}{6})$. Let $x^0 \in H$.
Step 1. Calculate

$$y^k = \mathrm{prox}_{\alpha_k g}\bigl(x^k - \alpha_k \nabla f(x^k)\bigr)$$

and

$$z^k = \mathrm{prox}_{\alpha_k g}\bigl(y^k - \alpha_k \nabla f(y^k)\bigr),$$

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

$$\alpha_k \cdot \max\bigl\{\|\nabla f(x^k) - \nabla f(y^k)\|, \|\nabla f(z^k) - \nabla f(y^k)\|\bigr\} \le \delta\bigl(\|x^k - y^k\| + \|z^k - y^k\|\bigr). \tag{3.1}$$

Step 2. Calculate

$$x^{k+1} = x^k - \gamma\eta_k d_k,$$

where

$$d_k = x^k - z^k - \alpha_k\bigl(\nabla f(x^k) - \nabla f(z^k)\bigr) \quad \text{and} \quad \eta_k = \frac{(\frac{1}{2} - 3\delta)(\|x^k - y^k\|^2 + \|z^k - y^k\|^2)}{\|d_k\|^2}.$$

Set $k := k + 1$, and go to *Step* 1.

Remark 3.2 For variational inequality problem, this kind of method is firstly appeared in Noor [18, 19, 22]

Lemma 3.3 ([14]) *Linesearch* (3.1) *stops after finitely many steps.*

Theorem 3.4 *Let* (x^k) *and* (α_k) *be generated by Algorithm* 3.1. *Assume that there is* $\alpha > 0$ *such that* $\alpha_k \geq \alpha > 0$ *for all* $k \in \mathbb{N}$. *Then* (x^k) *weakly converges to an element of* S_*.

Proof Let x_* be a solution in S_*. Then we obtain

$$\begin{aligned}\|x^{k+1} - x_*\|^2 &= \|x^k - \gamma\eta_k d_k - x_*\|^2 \\ &= \|x^k - x_*\|^2 - 2\gamma\eta_k\langle x^k - x_*, d_k\rangle + \gamma^2\eta_k^2\|d_k\|^2.\end{aligned} \tag{3.2}$$

Since $y^k = \operatorname{prox}_{\alpha_k g}(x^k - \alpha_k\nabla f(x^k))$, we have $(\mathrm{Id} - \alpha_k\nabla f)(x^k) \in (\mathrm{Id} + \alpha_k\partial g)(y^k)$. Moreover, ∂g is maximal monotone, so there is $u^k \in \partial g(y^k)$ such that

$$(\mathrm{Id} - \alpha_k\nabla f)(x^k) = y^k + \alpha_k u^k.$$

So we have

$$u^k = \frac{1}{\alpha_k}\bigl(x^k - y^k - \alpha_k\nabla f(x^k)\bigr). \tag{3.3}$$

Note that $0 \in \nabla f(x_*) + \partial g(x_*) \subseteq \partial(f + g)(x_*)$ and $\nabla f(y^k) + u^k \in \partial(f + g)y^k$. Therefore we obtain

$$\langle\nabla f(y^k) + u^k, y^k - x_*\rangle \geq 0. \tag{3.4}$$

Using (3.3) and (3.4), we have

$$\frac{1}{\alpha_k}\langle x^k - y^k - \alpha_k\nabla f(x^k) + \alpha_k\nabla f(y^k), y^k - x_*\rangle \geq 0.$$

It follows that

$$\langle x^k - y^k - \alpha_k\bigl(\nabla f(x^k) - \nabla f(y^k)\bigr), y^k - x_*\rangle \geq 0. \tag{3.5}$$

From $z^k = \operatorname{prox}_{\alpha_k g}(y^k - \alpha_k\nabla f(y^k))$ we get $(\mathrm{Id} - \alpha_k\nabla f)(y^k) \in (\mathrm{Id} + \alpha_k\partial g)(z^k)$. Since ∂g is maximal monotone, there is $v^k \in \partial g(z^k)$ such that

$$(\mathrm{Id} - \alpha_k\nabla f)(y^k) = z^k + \alpha_k v^k.$$

This shows that

$$v^k = \frac{1}{\alpha_k}\big(y^k - z^k - \alpha_k \nabla f\big(y^k\big)\big). \tag{3.6}$$

Similarly to y^k, we can show that

$$\big\langle y^k - z^k - \alpha_k\big(\nabla f\big(y^k\big) - \nabla f\big(z^k\big)\big), z^k - x_*\big\rangle \geq 0. \tag{3.7}$$

Combining (3.5) and (3.7), we have

$$\begin{aligned}
0 &\leq \big\langle x^k - y^k - \alpha_k\big(\nabla f\big(x^k\big) - \nabla f\big(y^k\big)\big), y^k - x_*\big\rangle + \big\langle y^k - z^k - \alpha_k\big(\nabla f\big(y^k\big) - \nabla f\big(z^k\big)\big), z^k - x_*\big\rangle \\
&= \big\langle x^k - y^k - \alpha_k\big(\nabla f\big(x^k\big) - \nabla f\big(y^k\big)\big), y^k - z^k\big\rangle + \big\langle x^k - y^k - \alpha_k\big(\nabla f\big(x^k\big) - \nabla f\big(y^k\big)\big), z^k - x_*\big\rangle \\
&\quad + \big\langle y^k - z^k - \alpha_k\big(\nabla f\big(y^k\big) - \nabla f\big(z^k\big)\big), z^k - x_*\big\rangle \\
&= \big\langle x^k - y^k - \alpha_k\big(\nabla f\big(x^k\big) - \nabla f\big(y^k\big)\big), y^k - z^k\big\rangle \\
&\quad + \big\langle x^k - z^k - \alpha_k\big(\nabla f\big(x^k\big) - \nabla f\big(z^k\big)\big), z^k - x_*\big\rangle.
\end{aligned} \tag{3.8}$$

We consider

$$\begin{aligned}
&\big\langle x^k - y^k - \alpha_k\big(\nabla f\big(x^k\big) - \nabla f\big(y^k\big)\big), y^k - z^k\big\rangle \\
&\quad = \big\langle x^k - y^k, y^k - z^k\big\rangle + \alpha_k\big\langle \nabla f\big(y^k\big) - \nabla f\big(x^k\big), y^k - z^k\big\rangle \\
&\quad = \big\langle x^k - y^k, y^k - z^k\big\rangle + \alpha_k\big[\big\langle \nabla f\big(y^k\big) - \nabla f\big(z^k\big), y^k - z^k\big\rangle + \big\langle \nabla f\big(z^k\big) - \nabla f\big(x^k\big), y^k - z^k\big\rangle\big] \\
&\quad = \big\langle x^k - y^k, y^k - z^k\big\rangle + \alpha_k\big[\big\langle \nabla f\big(y^k\big) - \nabla f\big(z^k\big), y^k - z^k\big\rangle \\
&\qquad + \big\langle \nabla f\big(z^k\big), y^k - z^k\big\rangle + \big\langle \nabla f\big(x^k\big), z^k - y^k\big\rangle\big] \\
&\quad = \big\langle x^k - y^k, y^k - z^k\big\rangle + \alpha_k\big[\big\langle \nabla f\big(y^k\big) - \nabla f\big(z^k\big), y^k - z^k\big\rangle + \big\langle \nabla f\big(z^k\big), y^k - z^k\big\rangle \\
&\qquad + \big\langle \nabla f\big(x^k\big), z^k - x^k\big\rangle + \big\langle \nabla f\big(x^k\big), x^k - y^k\big\rangle\big] \\
&\quad = \big\langle x^k - y^k, y^k - z^k\big\rangle + \alpha_k\big[\big\langle \nabla f\big(y^k\big) - \nabla f\big(z^k\big), y^k - z^k\big\rangle \\
&\qquad + \big\langle \nabla f\big(z^k\big), y^k - z^k\big\rangle + \big\langle \nabla f\big(x^k\big), z^k - x^k\big\rangle \\
&\qquad + \big\langle \nabla f\big(x^k\big) - \nabla f\big(y^k\big), x^k - y^k\big\rangle + \big\langle \nabla f\big(y^k\big), x^k - y^k\big\rangle\big].
\end{aligned}$$

By the convexity of f we have

$$\begin{aligned}
&\big\langle x^k - y^k - \alpha_k\big(\nabla f\big(x^k\big) - \nabla f\big(y^k\big)\big), y^k - z^k\big\rangle \\
&\quad \leq \big\langle x^k - y^k, y^k - z^k\big\rangle + \alpha_k\big[\big\langle \nabla f\big(y^k\big) - \nabla f\big(z^k\big), y^k - z^k\big\rangle + f\big(y^k\big) - f\big(z^k\big) + f\big(z^k\big) - f\big(x^k\big) \\
&\qquad + \big\langle \nabla f\big(x^k\big) - \nabla f\big(y^k\big), x^k - y^k\big\rangle + f\big(x^k\big) - f\big(y^k\big)\big] \\
&\quad = \big\langle x^k - y^k, y^k - z^k\big\rangle + \alpha_k\big[\big\langle \nabla f\big(y^k\big) - \nabla f\big(z^k\big), y^k - z^k\big\rangle \\
&\qquad + \big\langle \nabla f\big(x^k\big) - \nabla f\big(y^k\big), x^k - y^k\big\rangle\big].
\end{aligned} \tag{3.9}$$

Using $2\langle x^k - y^k, y^k - z^k\rangle = \|x^k - z^k\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2$, (3.1), (3.8), and (3.9), we see that

$$\begin{aligned}
&-\langle x^k - z^k - \alpha_k(\nabla f(x^k) - \nabla f(z^k)), z^k - x_*\rangle \\
&\quad\le \frac{1}{2}\big[\|x^k - z^k\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2\big] + \alpha_k\big[\langle \nabla f(y^k) - \nabla f(z^k), y^k - z^k\rangle \\
&\qquad + \langle \nabla f(x^k) - \nabla f(y^k), x^k - y^k\rangle\big] \\
&\quad\le \frac{1}{2}\big[\|x^k - z^k\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2\big] + \alpha_k\big[\|\nabla f(y^k) - \nabla f(z^k)\|\|y^k - z^k\| \\
&\qquad + \|\nabla f(x^k) - \nabla f(y^k)\|\|x^k - y^k\|\big] \\
&\quad\le \frac{1}{2}\big[\|x^k - z^k\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2\big] + \delta\big[(\|x^k - y^k\| + \|z^k - y^k\|)\|y^k - z^k\| \\
&\qquad + (\|x^k - y^k\| + \|z^k - y^k\|)\|x^k - y^k\|\big] \\
&\quad\le \frac{1}{2}\big[\|x^k - z^k\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2\big] + \delta\big[\|x^k - y^k\|\|y^k - z^k\| + \|z^k - y^k\|^2 \\
&\qquad + \|x^k - y^k\|^2 + \|z^k - y^k\|\|x^k - y^k\|\big] \\
&\quad\le \frac{1}{2}\big[\|x^k - z^k\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2\big] \\
&\qquad + \delta\big[2\|x^k - y^k\|\|y^k - z^k\| + \|z^k - y^k\|^2 + \|x^k - y^k\|^2\big] \\
&\quad\le \frac{1}{2}\big[\|x^k - z^k\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2\big] + 2\delta\big[\|z^k - y^k\|^2 + \|x^k - y^k\|^2\big] \\
&\quad\le \frac{1}{2}\|x^k - z^k\|^2 - \left(\frac{1}{2} - 2\delta\right)\|x^k - y^k\|^2 - \left(\frac{1}{2} - 2\delta\right)\|y^k - z^k\|^2.
\end{aligned}$$

So we have

$$\langle d_k, z^k - x_*\rangle \ge -\frac{1}{2}\|x^k - z^k\|^2 + \left(\frac{1}{2} - 2\delta\right)\|x^k - y^k\|^2 + \left(\frac{1}{2} - 2\delta\right)\|y^k - z^k\|^2. \tag{3.10}$$

Using the definition of d_k and Linesearch (3.1), we have

$$\begin{aligned}
&\langle d_k, x^k - x_*\rangle \\
&\quad= \langle x^k - z^k - \alpha_k(\nabla f(x^k) - \nabla f(z^k)), x^k - z^k\rangle + \langle d_k, z^k - x_*\rangle \\
&\quad= \|x^k - z^k\|^2 - \alpha_k\langle x^k - z^k, \nabla f(x^k) - \nabla f(z^k)\rangle + \langle d_k, z^k - x_*\rangle \\
&\quad= \|x^k - z^k\|^2 - \alpha_k\langle x^k - z^k, \nabla f(x^k) - \nabla f(y^k)\rangle - \alpha_k\langle x^k - z^k, \nabla f(y^k) - \nabla f(z^k)\rangle \\
&\qquad + \langle d_k, z^k - x_*\rangle \\
&\quad\ge \|x^k - z^k\|^2 - \alpha_k\|x^k - z^k\|\|\nabla f(x^k) - \nabla f(y^k)\| - \alpha_k\|x^k - z^k\|\|\nabla f(y^k) - \nabla f(z^k)\| \\
&\qquad + \langle d_k, z^k - x_*\rangle \\
&\quad\ge \|x^k - z^k\|^2 - \delta\|x^k - z^k\|(\|x^k - y^k\| + \|z^k - y^k\|) \\
&\qquad - \delta\|x^k - z^k\|(\|x^k - y^k\| + \|z^k - y^k\|) \\
&\qquad + \langle d_k, z^k - x_*\rangle
\end{aligned}$$

$$
\begin{aligned}
&= \|x^k - z^k\|^2 - \delta\big(\|x^k - z^k\|\|x^k - y^k\| + \|x^k - z^k\|\|z^k - y^k\|\big) \\
&\quad - \delta\big(\|x^k - z^k\|\|x^k - y^k\| + \|x^k - z^k\|\|z^k - y^k\|\big) + \langle d_k, z^k - x_*\rangle \\
&= \|x^k - z^k\|^2 - \delta\big(2\|x^k - z^k\|\|x^k - y^k\| + 2\|x^k - z^k\|\|z^k - y^k\|\big) + \langle d_k, z^k - x_*\rangle \\
&\geq \|x^k - z^k\|^2 - \delta\big(\|x^k - z^k\|^2 + \|x^k - y^k\|^2 + \|x^k - z^k\|^2 + \|z^k - y^k\|^2\big) + \langle d_k, z^k - x_*\rangle \\
&= (1 - 2\delta)\|x^k - z^k\|^2 - \delta\big(\|x^k - y^k\|^2 + \|z^k - y^k\|^2\big) + \langle d_k, z^k - x_*\rangle. \qquad (3.11)
\end{aligned}
$$

From (3.10) and (3.11) we have

$$
\begin{aligned}
\langle d_k, x^k - x_*\rangle &\geq (1 - 2\delta)\|x^k - z^k\|^2 - \delta\big(\|x^k - y^k\|^2 + \|z^k - y^k\|^2\big) - \frac{1}{2}\|x^k - z^k\|^2 \\
&\quad + \left(\frac{1}{2} - 2\delta\right)\|x^k - y^k\|^2 + \left(\frac{1}{2} - 2\delta\right)\|y^k - z^k\|^2 \\
&= \left(\frac{1}{2} - 2\delta\right)\|x^k - z^k\|^2 + \left(\frac{1}{2} - 3\delta\right)\big(\|x^k - y^k\|^2 + \|y^k - z^k\|^2\big). \qquad (3.12)
\end{aligned}
$$

Since $\eta_k = \frac{(\frac{1}{2} - 3\delta)(\|x^k - y^k\|^2 + \|z^k - y^k\|^2)}{\|d_k\|^2}$, we have $\eta_k\|d_k\|^2 = (\frac{1}{2} - 3\delta)(\|x^k - y^k\|^2 + \|z^k - y^k\|^2)$. So

$$
\langle d_k, x^k - x_*\rangle \geq \left(\frac{1}{2} - 2\delta\right)\|x^k - z^k\|^2 + \eta_k\|d_k\|^2. \qquad (3.13)
$$

This gives

$$
-2\gamma\eta_k\langle d_k, x^k - x_*\rangle \leq -2\gamma\eta_k\left(\frac{1}{2} - 2\delta\right)\|x^k - z^k\|^2 - 2\gamma\eta_k^2\|d_k\|^2. \qquad (3.14)
$$

Therefore from (3.2) and the above we obtain

$$
\begin{aligned}
\|x^{k+1} - x_*\|^2 &\leq \|x^k - x_*\|^2 - 2\gamma\eta_k\left(\frac{1}{2} - 2\delta\right)\|x^k - z^k\|^2 - 2\gamma\eta_k^2\|d_k\|^2 + \gamma^2\eta_k^2\|d_k\|^2 \\
&= \|x^k - x_*\|^2 - 2\gamma\eta_k\left(\frac{1}{2} - 2\delta\right)\|x^k - z^k\|^2 - \frac{2 - \gamma}{\gamma}\|\gamma\eta_k d_k\|^2. \qquad (3.15)
\end{aligned}
$$

By the monotonicity of ∇f we get

$$
\begin{aligned}
\|d_k\|^2 &= \big\|x^k - z^k - \alpha_k\big(\nabla f(x^k) - \nabla f(z^k)\big)\big\|^2 \\
&= \|x^k - z^k\|^2 + \alpha_k^2\big\|\nabla f(x^k) - \nabla f(z^k)\big\|^2 - 2\alpha_k\big\langle x^k - z^k, \nabla f(x^k) - \nabla f(z^k)\big\rangle \\
&\leq \|x^k - z^k\|^2 + \alpha_k^2\big\|\nabla f(x^k) - \nabla f(y^k) + \nabla f(y^k)\nabla - f(z^k)\big\|^2 \\
&\leq \|x^k - z^k\|^2 + 2\alpha_k^2\big[\big\|\nabla f(x^k) - \nabla f(y^k)\big\|^2 + \big\|\nabla f(y^k)\nabla - f(z^k)\big\|^2\big] \\
&\leq \|x^k - z^k\|^2 + 2\alpha_k^2\big[\big(\|x^k - y^k\| + \|z^k - y^k\|\big)^2 + \big(\|x^k - y^k\| + \|z^k - y^k\|\big)^2\big] \\
&\leq \|x^k - z^k\|^2 + 4\alpha_k^2\big(\|x^k - y^k\|^2 + 2\|x^k - y^k\|\|z^k - y^k\| + \|z^k - y^k\|^2\big) \\
&\leq \|x^k - y^k\|^2 + 2\|x^k - y^k\|\|y^k - z^k\| + \|y^k - z^k\|^2 + 8\alpha_k^2\big(\|x^k - y^k\|^2 + \|z^k - y^k\|^2\big) \\
&\leq 2\big(\|x^k - y^k\|^2 + \|y^k - z^k\|^2\big) + 8\alpha_k^2\big(\|x^k - y^k\|^2 + \|z^k - y^k\|^2\big) \\
&= \big(2 + 8\delta^2\big)\big(\|x^k - y^k\|^2 + \|y^k - z^k\|^2\big) \qquad (3.16)
\end{aligned}
$$

and, equivalently,

$$\frac{1}{\|d_k\|^2} \geq \frac{1}{(2+8\delta^2)(\|x^k - y^k\|^2 + \|y^k - z^k\|^2)}. \tag{3.17}$$

Therefore we have

$$\eta_k = \frac{(\frac{1}{2} - 3\delta)(\|x^k - y^k\|^2 + \|z^k - y^k\|^2)}{\|d_k\|^2} \geq \frac{(\frac{1}{2} - 3\delta)}{(2+8\delta^2)} > 0.$$

On the other hand, we have

$$\eta_k \|d_k\|^2 = \left(\frac{1}{2} - 3\delta\right)\left(\|x^k - y^k\|^2 + \|z^k - y^k\|^2\right). \tag{3.18}$$

Thus it follows that

$$\begin{aligned} \|x^k - y^k\|^2 + \|z^k - y^k\|^2 &= \frac{1}{(\frac{1}{2} - 3\delta)} \eta_k \|d_k\|^2 \\ &= \frac{1}{(\frac{1}{2} - 3\delta)(\gamma^2 \eta_k)} \|\gamma \eta_k d_k\|^2. \end{aligned} \tag{3.19}$$

From (3.18) and (3.19) we get

$$\|x^k - y^k\|^2 + \|z^k - y^k\|^2 \leq \frac{(2+8\delta^2)}{(\frac{1}{2} - 3\delta)} \|\gamma \eta_k d_k\|^2.$$

Since $x^{k+1} = x^k - \gamma \eta_k d_k$, it follows that $\gamma \eta_k d_k = x^k - x^{k+1}$. This implies that

$$\|x^{k+1} - x_*\|^2 \leq \|x^k - x_*\|^2 - 2\gamma \eta_k \left(\frac{1}{2} - 2\delta\right)\|x^k - z^k\|^2 - \frac{2-\gamma}{\gamma}\|x^k - x^{k+1}\|^2. \tag{3.20}$$

Thus $\lim_{k\to\infty} \|x^k - x_*\|$ exists, and (x^k) is bounded. Note that by (3.20)

$$\frac{2-\gamma}{\gamma}\|x^k - x^{k+1}\|^2 + 2\gamma \eta_k \left(\frac{1}{2} - 2\delta\right)\|x^k - z^k\|^2 \leq \|x^k - x_*\|^2 - \|x^{k+1} - x_*\|^2.$$

Hence $\|x^k - z^k\| \to 0$ and $\|x^{k+1} - x^k\| \to 0$ as $k \to \infty$. It follows that $\|x^k - y^k\| \to 0$ and $\|y^k - z^k\| \to 0$ as $k \to \infty$. By the boundedness of (x^k) we know that the set of its weak limit points is nonempty. Let $x^\infty \in \omega_w(x^k)$. Then there is a subsequence (x^{k_n}) of (x^k) such that $x^{k_n} \rightharpoonup x^\infty$. Next, we show that $x^\infty \in S_*$. Let $(v,u) \in \operatorname{Gph}(\nabla f + \partial g)$, that is, $u - \nabla f(v) \in \partial g(v)$. Since $y^{k_n} = (\mathrm{Id} + \alpha_{k_n}\partial g)^{-1}(\mathrm{Id} - \alpha_{k_n}\nabla f)x^{k_n}$, we have

$$(\mathrm{Id} - \alpha_{k_n}\nabla f)x^{k_n} \in (\mathrm{Id} + \alpha_{k_n}\partial g)y^{k_n},$$

which gives

$$\frac{1}{\alpha_{k_n}}\left(x^{k_n} - y^{k_n} - \alpha_{k_n}\nabla f\left(x^{k_n}\right)\right) \in \partial g\left(y^{k_n}\right).$$

Since ∂g is maximal monotone, it follows that

$$\left\langle v - y^{k_n}, u - \nabla f(v) - \frac{1}{\alpha_{k_n}}\left(x^{k_n} - y^{k_n} - \alpha_{k_n}\nabla f\left(x^{k_n}\right)\right)\right\rangle \geq 0.$$

This shows that

$$\begin{aligned}
\langle v - y^{k_n}, u\rangle &\geq \left\langle v - y^{k_n}, \nabla f(v) + \frac{1}{\alpha_{k_n}}\left(x^{k_n} - y^{k_n} - \alpha_{k_n}\nabla f\left(x^{k_n}\right)\right)\right\rangle \\
&= \left\langle v - y^{k_n}, \nabla f(v) - \nabla f\left(x^{k_n}\right)\right\rangle + \left\langle v - y^{k_n}, \frac{1}{\alpha_{k_n}}\left(x^{k_n} - y^{k_n}\right)\right\rangle \\
&= \left\langle v - y^{k_n}, \nabla f(v) - \nabla f\left(y^{k_n}\right)\right\rangle + \left\langle v - y^{k_n}, \nabla f\left(y^{k_n}\right) - \nabla f\left(x^{k_n}\right)\right\rangle \\
&\quad + \left\langle v - y^{k_n}, \frac{1}{\alpha_{k_n}}\left(x^{k_n} - y^{k_n}\right)\right\rangle \\
&\geq \left\langle v - y^{k_n}, \nabla f\left(y^{k_n}\right) - \nabla f\left(x^{k_n}\right)\right\rangle + \left\langle v - y^{k_n}, \frac{1}{\alpha_{k_n}}\left(x^{k_n} - y^{k_n}\right)\right\rangle. \qquad (3.21)
\end{aligned}$$

Since $\lim_{k\to\infty}\|x^k - y^k\| = 0$, by the assumption we have $\lim_{k\to\infty}\|\nabla f(x^k) - \nabla f(y^k)\| = 0$. Taking the limit as $n \to \infty$ in (3.21), we have

$$\langle v - x^\infty, u\rangle \geq 0.$$

Thus $0 \in (\nabla f + \partial g)x^\infty$, and consequently $x^\infty \in S_*$. By Lemma 2.4 we conclude that (x^k) converges weakly to an element of S_*. Thus we complete the proof. □

Remark 3.5 If ∇f is L-Lipschitz continuous, then the condition on α_k in Theorem 3.4 can be removed since $\alpha_k \geq \min\{\sigma, \delta\theta/L\} > 0$; see [3].

Next, we introduce a new projected forward–backward algorithm and the convergence analysis. We denote by $\Omega \cap \operatorname{argmin}(f+g)$ the solution set of (1.5). Assume that this solution set is nonempty.

Algorithm 3.6 Given $\sigma > 0$, $\theta \in (0,1)$, $\gamma \in (0,2)$, and $\delta \in (0,\frac{1}{6})$. Let $w^0 \in H$.
Step 1. Calculate

$$x^k = \operatorname{prox}_{\alpha_k g}\left(w^k - \alpha_k \nabla f\left(w^k\right)\right)$$

and

$$y^k = \operatorname{prox}_{\alpha_k g}\left(x^k - \alpha_k \nabla f\left(x^k\right)\right),$$

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

$$\alpha_k \cdot \max\left\{\left\|\nabla f\left(w^k\right) - \nabla f\left(x^k\right)\right\|, \left\|\nabla f\left(y^k\right) - \nabla f\left(x^k\right)\right\|\right\} \leq \delta\left(\left\|w^k - x^k\right\| + \left\|y^k - x^k\right\|\right). \quad (3.22)$$

Step 2. Calculate

$$z^k = w^k - \gamma\eta_k d_k,$$

where

$$d_k = w^k - y^k - \alpha_k\big(\nabla f\big(w^k\big) - \nabla f\big(y^k\big)\big) \quad \text{and} \quad \eta_k = \frac{(\frac{1}{2} - 3\delta)(\|w^k - x^k\|^2 + \|y^k - x^k\|^2)}{\|d_k\|^2}.$$

Step 3. Calculate

$$w^{k+1} = P_\Omega\big(z^k\big).$$

Set $k := k + 1$, and go to *Step* 1.

Theorem 3.7 *Let* (x^k) *and* (α_k) *be generated by Algorithm* 3.6. *Assume that there is* $\alpha > 0$ *such that* $\alpha_k \geq \alpha > 0$ *for all* $k \in \mathbb{N}$. *Then* (x^k) *weakly converges to an element of* $\Omega \cap \operatorname{argmin}(f + g)$.

Proof Let w_* be a solution in $\Omega \cap \operatorname{argmin}(f + g)$. Then using Lemma 2.1(ii), we have

$$\begin{aligned} \big\|w^{k+1} - w_*\big\|^2 &= \big\|P_\Omega\big(z^k\big) - w_*\big\|^2 \\ &\leq \big\|z^k - w_*\big\|^2 - \big\|P_\Omega\big(z^k\big) - z^k\big\|^2. \end{aligned} \tag{3.23}$$

Since $z^k = w^k - \gamma\eta_k d_k$, we have $\gamma\eta_k d_k = w^k - z^k$. Similarly to Theorem 3.4, we can show that

$$\big\|z^k - w_*\big\|^2 \leq \big\|w^k - w_*\big\|^2 - 2\gamma\eta_k\left(\frac{1}{2} - 2\delta\right)\big\|w^k - y^k\big\|^2 - \frac{2-\gamma}{\gamma}\big\|w^k - z^k\big\|^2. \tag{3.24}$$

From (3.23) and (3.24) we obtain

$$\begin{aligned} \big\|w^{k+1} - w_*\big\|^2 \leq{}& \big\|w^k - w_*\big\|^2 - 2\gamma\eta_k\left(\frac{1}{2} - 2\delta\right)\big\|w^k - y^k\big\|^2 - \frac{2-\gamma}{\gamma}\big\|w^k - z^k\big\|^2 \\ &- \big\|P_\Omega\big(z^k\big) - z^k\big\|^2. \end{aligned} \tag{3.25}$$

Thus $\lim_{k\to\infty}\|w^k - w_*\|$ exists, and (w^k) is bounded. From (3.25) we see that

$$\begin{aligned} &2\gamma\eta_k\left(\frac{1}{2} - 2\delta\right)\big\|w^k - y^k\big\|^2 + \frac{2-\gamma}{\gamma}\big\|w^k - z^k\big\|^2 + \big\|P_\Omega\big(z^k\big) - z^k\big\|^2 \\ &\quad\leq \big\|w^k - w_*\big\|^2 - \big\|w^{k+1} - w_*\big\|^2. \end{aligned}$$

Thus $\|w^k - y^k\| \to 0$, $\|w^k - z^k\| \to 0$, and $\|P_\Omega(z^k) - z^k\| \to 0$ as $k \to \infty$. Also, we can show that $\|w^k - x^k\| \to 0$ and $\|y^k - x^k\| \to 0$ as $k \to \infty$. Let $w^\infty \in \omega_w(w_*)$. As in Theorem 3.7, we can show that $w^\infty \in \operatorname{argmin}(f + g)$. On the other hand, since $\lim_{k\to\infty}\|P_\Omega(z^k) - z^k\| = 0$ and $z^k \rightharpoonup w^\infty$, by Lemma 2.3 we have $w^\infty \in \Omega$. Therefore $w^\infty \in \Omega \cap \operatorname{argmin}(f + g)$. Using Lemma 2.4, we can conclude that Theorem 3.7 holds. □

4 Numerical experiments

In this section, we apply our result to the signal recovery in compressive sensing and image inpainting. We compare the performance of our algorithms with those of Combettes and Wajs [9] and Cruz and Nghia [3].

The numerical experiments are performed by Matlab 2020b on a 64-bit MacBook Pro Chip Apple M1 and 8 GB of RAM.

We consider the following LASSO problem:

$$\min_{x\in\mathbb{R}^N}\left(\frac{1}{2}\|Ax-y\|_2^2+\lambda\|x\|_1\right), \tag{4.1}$$

where $A:\mathbb{R}^N\to\mathbb{R}^M(M<N)$ is a bounded linear operator, $y\in\mathbb{R}^M$ is the observed data, and $\lambda>0$. Rewriting (4.1) as problem (1.1), we can set

$$f(x)=\frac{1}{2}\|y-Ax\|_2^2,\qquad g(x)=\lambda\|x\|_1.$$

In experiment, y is generated by the Gaussian noise with SNR = 40, A is generated by normal distribution with mean zero and variance one, and $x\in\mathbb{R}^N$ is generated by uniform distribution in [−2,2] that contains m nonzero components. The stopping criterion is defined by

$$\mathrm{MSE}=\frac{1}{N}\left\|x^k-x_*\right\|^2<10^{-4},$$

where x^k is an estimated signal of x_*.

The initial point x^0 is chosen to be zero. Let $\alpha=\frac{1}{\|A\|^2}$ and $\lambda_k=0.82$ in Algorithm 1.1. Let $\sigma=7$, $\delta=0.02$, $\theta=0.15$, and $\gamma=1.85$ in Algorithms 1.2 and 3.1, respectively. We now present the corresponding numerical results (the number of iterations is denoted by Iter, and CPU denotes the time of CPU) using different numbers of inequality constraints m. The numerical results are shown in Table 1.

From Table 1 we see that the experiment result of Algorithm 3.1 is better than those of Algorithms 1.1 and 1.2 in terms of CPU time and number of iterations in all cases.

Next, we provide Fig. 1 to show the convergence of each algorithm via the graph of the MSE value and number of iterations and Fig. 2 to show signal recovery in compressed sensing when $N=1024$, $M=512$, and $m=70$.

Next, we analyze the convergence and the effects of the stepsizes depending on the parameters σ, δ, θ, and γ in Algorithm 3.1.

Table 1 Computational results for compressive sensing

m-sparse signal	Methods	$N=1024$, $M=512$		$N=2048$, $M=1024$	
		CPU	Iter	CPU	Iter
$m=50$	Algorithm 1.1	17.2860	8219	60.6703	11,237
	Algorithm 1.2	15.6780	2941	54.0610	4429
	Algorithm 3.1	10.1545	1266	29.8420	1522
$m=60$	Algorithm 1.1	30.3607	11,478	82.9704	13746
	Algorithm 1.2	20.5542	3700	56.8577	4718
	Algorithm 3.1	12.4216	1622	30.7309	1742
$m=70$	Algorithm 1.1	39.9470	13,507	97.9897	15191
	Algorithm 1.2	21.8114	4079	60.7027	5035
	Algorithm 3.1	14.4815	1873	33.8620	1880
$m=90$	Algorithm 1.1	112.9716	24,608	124.0622	17,415
	Algorithm 1.2	30.4207	5683	72.1555	5793
	Algorithm 3.1	24.9734	3121	38.2926	2137

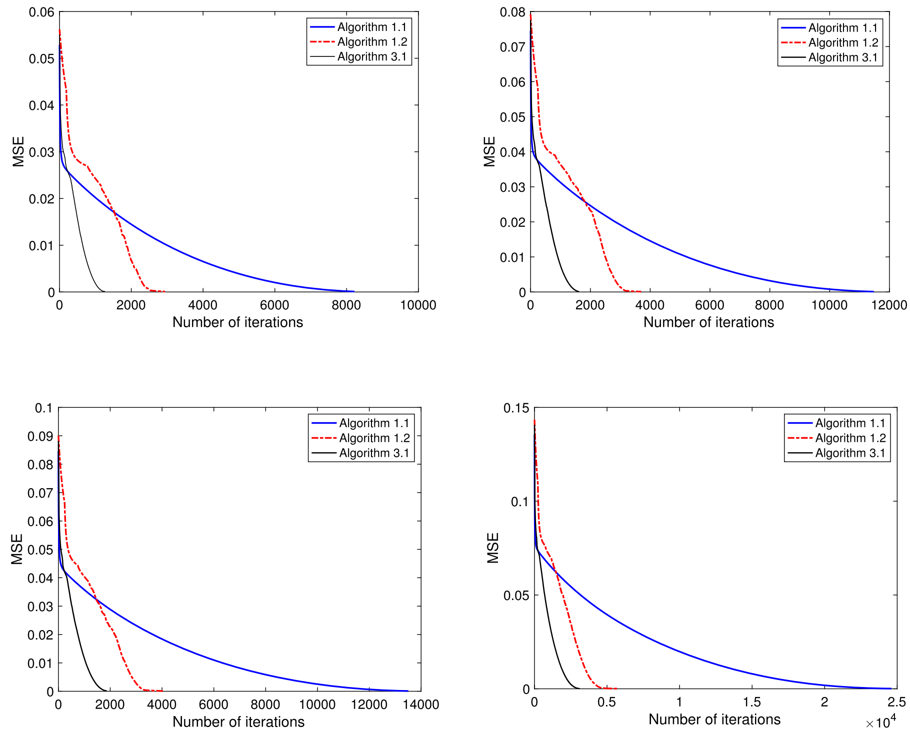

Figure 1 The MSE value and number of iterations for all cases when $N = 1024$ and $M = 512$

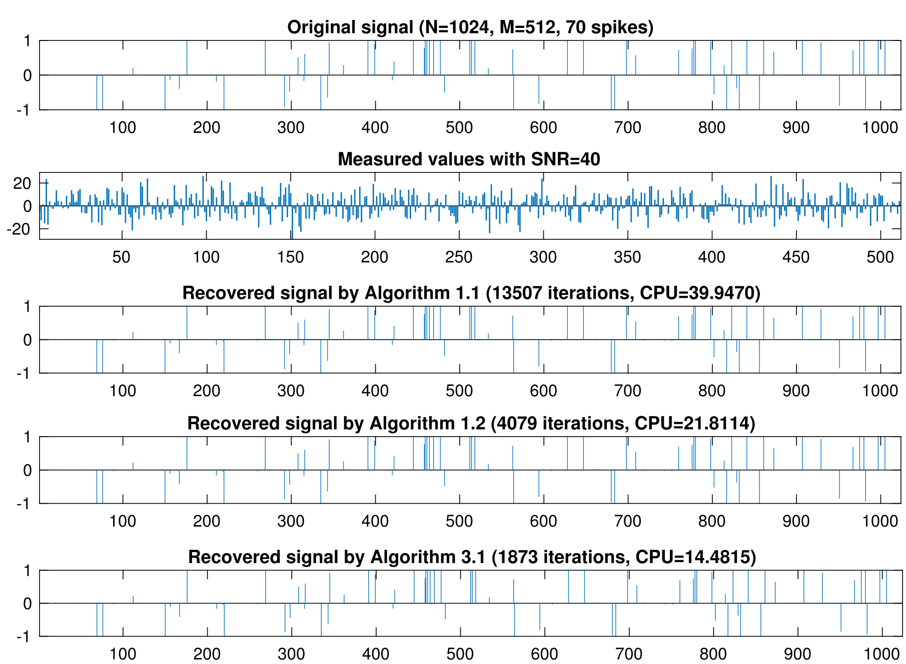

Figure 2 Original, observed, and recovered signals by Algorithms 1.1, 1.2, and 3.1, respectively

Table 2 The convergence of Algorithm 3.1 with various γ

Given: $\sigma = 1, \theta = 0.15, \delta = 0.02, m = 80$				
γ	$N = 1024, M = 512$		$N = 2048, M = 1024$	
	CPU	Iter	CPU	Iter
0.15	383.4834	33,965	860.0725	32,155
0.55	70.5710	8959	165.9668	8510
1.35	27.1354	3732	59.5584	3278
1.95	15.6307	2243	35.8540	2024

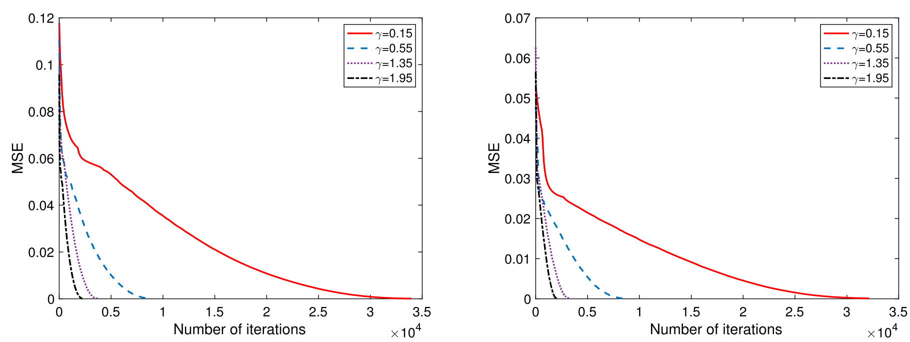

Figure 3 Graphs of MSE value and number of iterations when $N = 1024$, $M = 512$ and $N = 2048$, $M = 1024$

Table 3 The convergence of Algorithm 3.1 with each θ

Given: $\sigma = 1, \delta = 0.02, \gamma = 1.85, m = 80$				
γ	$N = 1024, M = 512$		$N = 2048, M = 1024$	
	CPU	Iter	CPU	Iter
0.15	15.6425	2203	37.4697	2045
0.25	17.8066	2195	41.1519	1950
0.75	48.5706	1818	143.6058	1896
0.95	203.4910	1428	587.1587	1530

In the first experiment, we investigate the effect of the parameter γ in the proposed algorithm. We intend to vary this parameter and study its convergence behavior. The numerical results are shown in Table 2.

From Table 2 we see that the CPU time and the number of iterations of Algorithm 3.1 decrease when the parameter γ approaches 2. We show numerical results for each case of γ in Fig. 3.

In the second experiment, we compare the performance of Algorithm 3.1 with different parameters θ in Theorem 3.4. Numerical results are shown in Table 3.

From Table 3 we observe that the CPU time of Algorithm 3.1 increases, but the number of iterations decreases when the parameter θ approaches 1. Figure 4 shows numerical results for each θ.

Next, we compare the performance of Algorithm 3.1 with different parameters σ in Theorem 3.4. Numerical results are reported in Table 4.

From Table 4 we observe that CPU increases when σ increases. However, there is no effect in terms of iterations.

Similarly, we obtain numerical results of Algorithm 3.1 with different δ in Table 5.

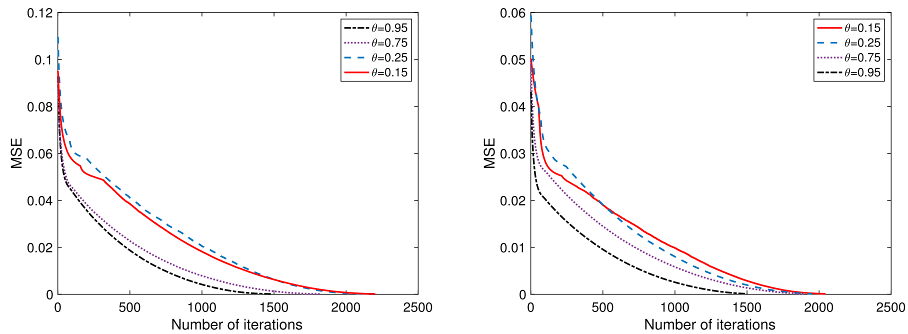

Figure 4 Graph of MSE value and number of iterations when $N = 1024$, $M = 512$ and $N = 2048$, $M = 1024$

Table 4 The convergence of Algorithm 3.1 with different σ

Given: $\delta = 0.02$, $\theta = 0.15$, $\gamma = 1.85$, $m = 80$				
γ	$N = 1024$, $M = 512$		$N = 2048$, $M = 1024$	
	CPU	Iter	CPU	Iter
0.1	13.2968	2228	26.7530	1765
1	14.3224	2069	39.0200	2177
10	19.6720	2428	54.4288	2624
100	21.7268	2261	49.6928	2165

Table 5 The convergence of Algorithm 3.1 with different δ

Given: $\sigma = 1$, $\theta = 0.02$, $\gamma = 1.85$, $m = 80$				
γ	$N = 1024$, $M = 512$		$N = 2048$, $M = 1024$	
	CPU	Iter	CPU	Iter
0.03	14.3172	2074	30.5167	1680
0.05	15.2507	2223	27.4978	1371
0.07	16.1215	2325	24.2387	1457
0.15	18.7078	2685	37.4357	2148

From Table 5 we see that the parameter δ has no effect in terms of the number of iterations and CPU time for both cases.

Next, we aim to apply our result for solving an image inpainting problem described by the following mathematical model:

$$\min_{x \in \mathbb{R}^{M \times N}} \frac{1}{2} \big\| A(x - x_0) \big\|_F^2 + \mu \|x\|_*, \tag{4.2}$$

where $x_0 \in \mathbb{R}^{M \times N} (M < N)$ is a matrix with entries that lie in the interval $[l, u]$, A is a linear map that selects a subset of the entries of an $M \times N$ matrix by setting each unknown entry in the matrix to 0, x is matrix of known entries $A(x_0)$, and $\mu > 0$ is a regularization parameter.

In particular, we consider the following image inpainting problem [10, 11]:

$$\min_x \frac{1}{2} \big\| P_\Omega(x) - P_\Omega(x_0) \big\|_F^2 + \mu \|x\|_*, \tag{4.3}$$

where $\|\cdot\|_F$ is the Frobenius norm, and $\|\cdot\|_*$ is the nuclear norm. Here we define P_Ω by

$$P_\Omega(x) = \begin{cases} x_{ij}, & (i,j) \in \Omega, \\ 0, & \text{otherwise}. \end{cases} \tag{4.4}$$

The nuclear norm has been widely used in image inpainting and matrix completion problem, which is a convex relaxation of low rank constraint. It is obvious that the optimization problem (4.3) is related to (1.5). Indeed fact, let $f(x) = \frac{1}{2}\|P_\Omega(x) - P_\Omega(x_0)\|_F^2$ and $g(x) = \mu\|x\|_*$. Then $\nabla f(x) = P_\Omega(x) - P_\Omega(x_0)$ is 1-Lipschitz continuous. The proximity operator of $g(x)$ can be computed by the singular value decomposition (SVD) [5].

To evaluate the quality of the restored images, we use the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [24] defined by

$$\text{PSNR} = 20 \log \frac{\|x\|_F}{\|x - x_r\|_F} \tag{4.5}$$

and

$$\text{SSIM} = \frac{(2u_x u_{x_r} + c_1)(2\sigma_{xx_r} + c_2)}{(u_x^2 + u_{x_r}^2 + c_1)(\sigma_x^2 + \sigma_{x_r}^2 + c_2)}, \tag{4.6}$$

where x is the original image, x_r is the restored image, u_x and u_{x_r} are the mean values of the original image x and restored image x_r, respectively, σ_x^2 and $\sigma_{x_r}^2$ are the variances, $\sigma_{xx_r}^2$ is the covariance of two images, $c_1 = (K_1L)^2$ and $c_2 = (K_2L)^2$ with $K_1 = 0.01$ and $K_2 = 0.03$, and L is the dynamic range of pixel values. SSIM ranges from 0 to 1, with 1 meaning perfect recovery. The initial point x^0 is chosen to be zero. Let $\alpha = \frac{1}{\|A\|^2}$ and $\lambda_k = 0.82$ in Algorithm 1.1. Let $\sigma = 7$, $\delta = 0.02$, $\theta = 0.15$, and $\gamma = 1.85$ in Algorithms 1.2 and 3.1, respectively. We obtain the following results.

From Table 6 we see that the experiment results of Algorithm 3.6 are better than those of Algorithms 1.1 and 1.2 in terms of PSNR and SSIM in all cases.

The original images are given in Fig. 5. The figures of inpainting images for the 250th and 350th iterations are shown in Figs. 6–7. The PSNR values and iterations are plotted in Fig. 8.

Next, we analyze the convergence and the effects of the stepsizes depending on the parameters σ, γ, θ, and δ in Algorithm 3.6.

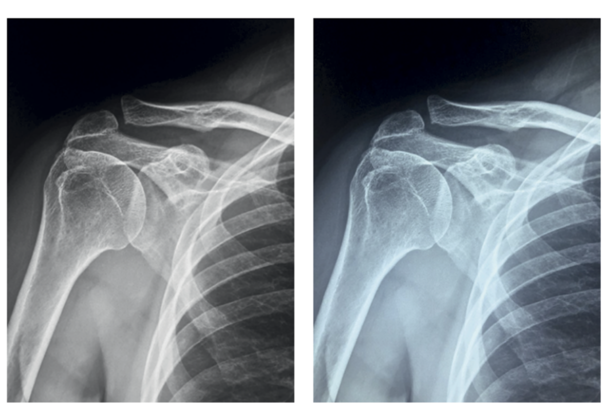

Figure 5 The original image

Table 6 Computational results for solving (4.3)

Methods	$N = 512, M = 256$, Iter = 250			$N = 1024, M = 512$, Iter = 350		
	PSNR	SSIM	CPU	PSNR	SSIM	CPU
Algorithm 1.1	31.7958	0.9777	42.9632	26.3070	0.9375	157.0621
Algorithm 1.2	31.9191	0.9780	350.0757	26.6354	0.9381	1295.4495
Algorithm 3.6	32.2537	0.9789	541.9922	27.6598	0.9400	2249.6655

Figure 6 The missing and restored images by Algorithm 1.1 (PSNR: 31.7958, SSIM: 0.9777), Algorithm 1.2 (PSNR: 31.9191, SSIM: 0.9780), and Algorithm 3.6 (PSNR: 32.2537, SSIM: 0.9789), respectively

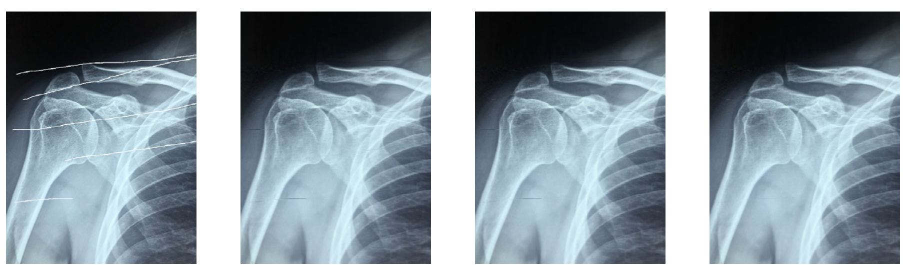

Figure 7 The missing and restored images by Algorithm 1.1 (PSNR: 26.3070, SSIM: 0.9375), Algorithm 1.2 (PSNR: 26.6354, SSIM: 0.9381), and Algorithm 3.6 (PSNR: 27.6598, SSIM: 0.9400)

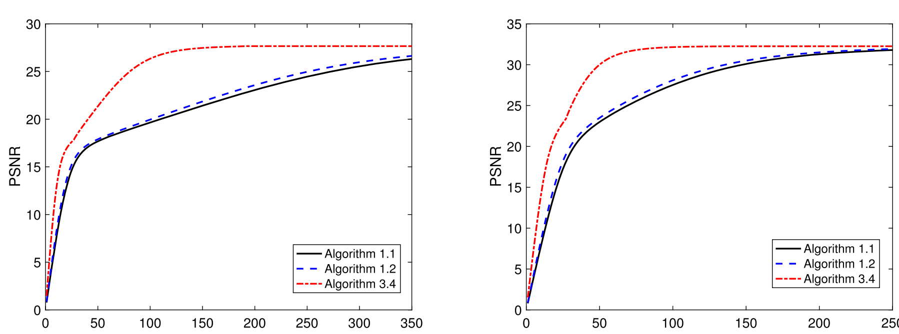

Figure 8 Graphs of PSNR value and number of iterations of Figs. 6 and 7, respectively

First, we study the effect of the parameter σ in the proposed algorithm. The numerical results are shown in Table 7.

From Table 7 we observe that the PSNR and the SSIM of Algorithm 3.6 increase when the parameter σ increases. Figure 9 shows numerical results for various σ.

Table 7 Computational results for each σ

Given: $\gamma = 1.75, \theta = 0.57, \delta = 0.17$						
σ	$N = 512, M = 256$, Iter = 80			$N = 1024, M = 512$, Iter = 120		
	PSNR	SSIM	CPU	PSNR	SSIM	CPU
1.5	31.8628	0.9777	231.7521	26.9550	0.9388	911.3357
1	31.8185	0.9778	105.0579	27.0448	0.930	390.0906
0.1	26.3606	0.9644	29.3329	20.6162	0.9292	118.4001
0.01	5.7995	0.6534	29.2479	7.9950	0.8156	119.6576

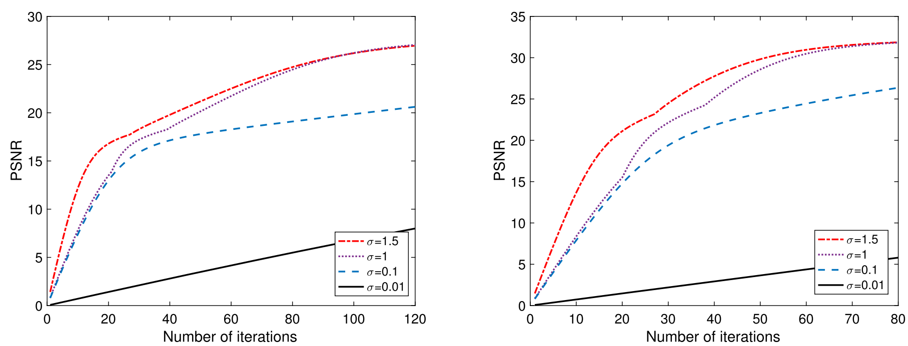

Figure 9 Graph of PSNR value and number of iterations

Table 8 Computational results for each γ

Given: $\sigma = 5, \theta = 0.57, \delta = 0.17$						
γ	$N = 512, M = 256$, Iter = 80			$N = 1024, M = 512$, Iter = 120		
	PSNR	SSIM	CPU	PSNR	SSIM	CPU
1.75	31.9129	0.9780	175.6244	27.0498	0.9390	707.5356
1.55	31.6503	0.9775	177.3801	26.6102	0.9381	722.1810
0.65	24.2457	0.9580	193.9734	20.4814	0.9292	750.3288
0.15	9.3123	0.8814	195.3802	12.0074	0.9118	789.5196

Next, we investigate the effect of the parameter γ in the proposed algorithm. We intend to vary this parameter and study its convergence behavior. The numerical results are shown in Table 8.

From Table 8 we observe that the PSNR and the SSIM of Algorithm 3.6 increase when the parameter γ approaches 2. Moreover, we see that CPU time decreases when the parameter γ approaches 2. Figure 10 shows numerical results for various γ.

Next, we study the effect of the parameter θ. The numerical results are shown in Table 9.

From Table 8 we observe that the PSNR, SSIM, and CPU time of Algorithm 3.6 increase when the parameter θ approaches 1. Figure 11 shows numerical results for various θ.

We next study the effect of the parameter δ. The results are shown in Table 10.

From Table 10 we observe that the PSNR and SSIM of Algorithm 3.6 increase when the parameter δ approaches 1/6. Moreover, we see that CPU time increases when the parameter δ approaches 0. Figure 12 shows numerical results for each δ.

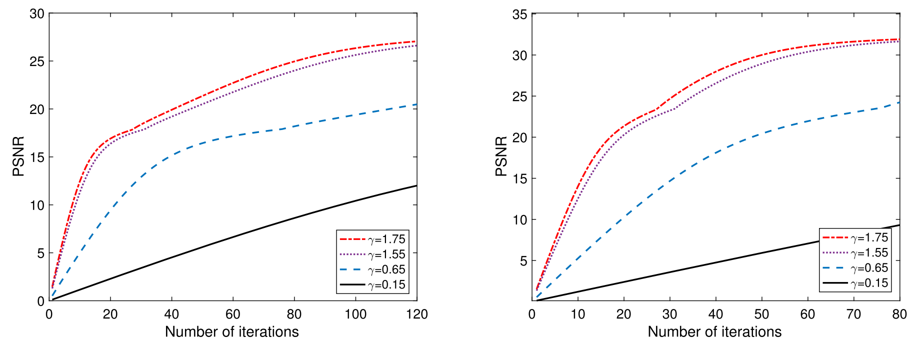

Figure 10 Graph of PSNR value and number of iterations

Table 9 Computational results for each θ

Given: $\gamma = 1.75, \sigma = 5, \delta = 0.17$						
θ	$N = 512, M = 256$, Iter = 80			$N = 1024, M = 512$, Iter = 120		
	PSNR	SSIM	CPU	PSNR	SSIM	CPU
0.15	27.2789	0.9672	85.4890	21.2435	0.9298	337.5378
0.25	27.5130	0.9681	109.2624	25.0008	0.9351	402.7532
0.75	31.8248	0.9778	302.6839	26.8473	0.9386	1261.3915
0.95	32.0885	0.9784	1569.6501	27.3524	0.9396	7198.9700

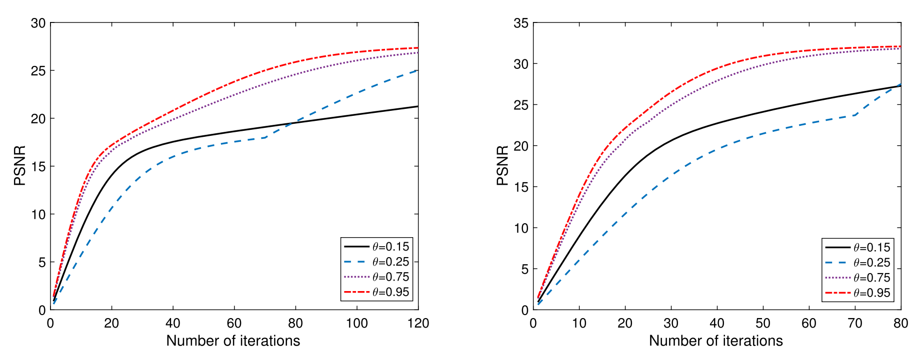

Figure 11 Graph of PSNR value and number of iterations

Table 10 Computational results for each δ

Given: $\gamma = 1.75, \sigma = 5, \theta = 0.57$						
δ	$N = 512, M = 256$, Iter = 80			$N = 1024, M = 512$, Iter = 120		
	PSNR	SSIM	CPU	PSNR	SSIM	CPU
0.03	19.1459	0.9376	296.1227	17.7976	0.9285	1155.8961
0.05	23.8391	0.9542	265.7709	20.3615	0.9289	1020.0585
0.07	27.4554	0.9676	237.0004	21.9502	0.9306	925.7105
0.15	30.1473	0.9743	202.3671	24.1919	0.9337	793.9251

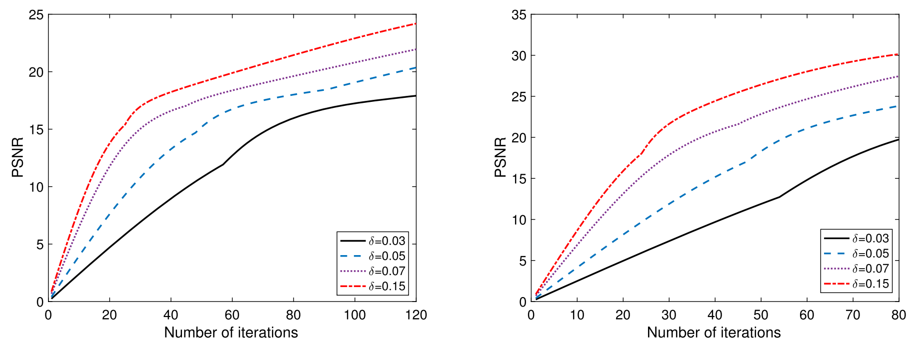

Figure 12 Graph of PSNR value and number of iterations

5 Conclusion

In this work, we proposed new forward–backward algorithms for solving convex minimization problem. We proved weak convergence theorems under some weakened assumptions on the stepsize. Our algorithms do not require the Lipschitz constant of the gradient of functions. Moreover, we proposed a new projected forward–backward splitting algorithm using new linesearch to solve constrained convex minimization problem. As a result, it can be applied effectively to solve signal recovery and image inpainting. Our algorithms have a good performance in terms of iterations and CPU times. We also have discussed the effects of all parameters in our algorithms.

Acknowledgements
The authors would like to thank editor and reviewers for value comments. The authors thank Chiang Mai University and University of Phayao for their supports.

Funding
This work was supported by Thailand Science Research and Innovation under the project IRN62W0007, the revenue budget in 2021, School of Science, University of Phayao, and Thailand Research Fund under project RSA6180084.

Availability of data and materials
Contact the author for data requests.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All the authors contributed equally to this manuscript. All authors read and approved the final manuscript.

Author details
[1]Data Science Research Center, Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand. [2]Research Center in Mathematics and Applied Mathematics, Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand. [3]Department of Mathematics, COMSATS University Islamabad, Islamabad, Pakistan. [4]School of Science, University of Phayao, Phayao 56000, Thailand.

Received: 16 February 2021 Accepted: 11 May 2021 Published online: 21 May 2021

References

1. Bauschke, H.H., Borwein, J.M.: Dykstra' s alternating projection algorithm for two sets. J. Approx. Theory **79**(3), 418–443 (1994)
2. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, vol. 408. Springer, New York (2011)
3. Bello Cruz, J.Y., Nghia, T.T.: On the convergence of the forward–backward splitting method with linesearches. Optim. Methods Softw. **31**(6), 1209–1238 (2016)

4. Burachik, R.S., Iusem, A.N.: Set-Valued Mappings and Enlargements of Monotone Operators, vol. 8. Springer, Berlin (2007)
5. Cai, J.F., Cands, E.J., Shen, Z.: A singular value thresholding algorithm for matrix completion. SIAM J. Optim. **20**(4), 1956–1982 (2010)
6. Chen, G.H., Rockafellar, R.T.: Convergence rates in forward–backward splitting. SIAM J. Optim. **7**(2), 421–444 (1997)
7. Cholamjiak, W., Cholamjiak, P., Suantai, S.: An inertial forward–backward splitting method for solving inclusion problems in Hilbert spaces. J. Fixed Point Theory Appl. **20**(1), 42 (2018)
8. Combettes, P.L., Pesquet, J.C.: Proximal splitting methods in signal processing. In: Fixed-Point Algorithms for Inverse Problems in Science and Engineering, pp. 185–212. Springer, New York (2011)
9. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward–backward splitting. Multiscale Model. Simul. **4**(4), 1168–1200 (2005)
10. Cui, F., Tang, Y., Yang, Y.: An inertial three-operator splitting algorithm with applications to image inpainting. arXiv preprint (2019). arXiv:1904.11684
11. Davis, D., Yin, W.: A three-operator splitting scheme and its optimization applications. Set-Valued Var. Anal. **25**(4), 829–858 (2017)
12. Dong, Q., Jiang, D., Cholamjiak, P., Shehu, Y.: A strong convergence result involving an inertial forward–backward algorithm for monotone inclusions. J. Fixed Point Theory Appl. **19**(4), 3097–3118 (2017)
13. Hussain, A., Ali, D., Karapinar, E.: Stability data dependency and errors estimation for a general iteration method. Alex. Eng. J. **60**(1), 703–710 (2021)
14. Kankam, K., Pholasa, N., Cholamjiak, P.: Hybrid forward–backward algorithms using linesearch rule for minimization problem. Thai J. Math. **17**(3), 607–625 (2019)
15. Li, L., Wu, D.: The convergence of Ishikawa iteration for generalized Φ-contractive mappings. Results Nonlinear Anal. **4**(1), 47–56 (2021)
16. Marino, G., Scardamaglia, B., Karapinar, E.: Strong convergence theorem for strict pseudo-contractions in Hilbert spaces. J. Inequal. Appl. **2016**(1), 134 (2016)
17. Noor, M.A.: New approximation schemes for general variational inequalities. J. Math. Anal. Appl. **251**(1), 217–229 (2000)
18. Noor, M.A.: Some developments in general variational inequalities. Appl. Math. Comput. **152**(1), 199–277 (2004)
19. Noor, M.A., Noor, K.I., Rassias, M.T.: New trends in general variational inequalities. Acta Appl. Math. **170**(1), 981–1064 (2020)
20. Opial, Z.: Weak convergence of the sequence of successive approximations for nonexpansive mappings. Bull. Am. Math. Soc. **73**(4), 591–597 (1967)
21. Rus, I.A.: Some problems in the fixed point theory. Adv. Theory Nonlinear Anal. Appl. **2**, 1–10 (2018)
22. Suantai, S., Kesornprom, S., Pholasa, N., Cho, Y.J., Cholamjiak, P.: A relaxed projection method using a new linesearch for the split feasibility problem. AIMS Math. **6**(3), 2690–2703 (2021)
23. Tseng, P.: A modified forward–backward splitting method for maximal monotone mappings. SIAM J. Control Optim. **38**(2), 431–446 (2000)
24. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)