A new preconditioning algorithm for finding a zero of the sum of two monotone operators and its application to image restoration problems

Ebru Altiparmak and Ibrahim Karahan

Department of Mathematics, Faculty of Science, Erzurum Technical University, Erzurum, Turkey

ABSTRACT
Finding a zero of the sum of two monotone operators is one of the most important problems in monotone operator theory, and the forward-backward algorithm is the most prominent approach for solving this type of problem. The aim of this paper is to present a new preconditioning forward-backward algorithm to obtain the zero of the sum of two operators in which one is maximal monotone and the other one is $M$-cocoercive, where $M$ is a linear bounded operator. Furthermore, the strong convergence of the proposed algorithm, which is a broader variant of previously known algorithms, has been proven in Hilbert spaces. We also use our algorithm to tackle the convex minimization problem and show that it outperforms existing algorithms. Finally, we discuss several image restoration applications.

1. Introduction
Let $H$ be a real Hilbert space with inner product $\langle ., . \rangle$ and the induced norm $\| . \|$. One of the most important problems in monotone operator theory is the problem of finding a zero of the sum of two monotone operators so-called the monotone inclusion problem which is defined by finding $x \in H$ such that

$$0 \in (A + B)(x)$$

(1)

where $A : H \to 2^H$ is a set-valued operator and $B : H \to H$ is an operator. This problem includes many mathematical problems such as variational inequality problems, convex minimization problems, equilibrium problems and convex-concave saddle point problems see, e.g. [4,7,10,11,17,19]. More precisely, it has applications in many scientific fields such as image processing, signal processing, machine learning and statistical regression see, e.g. [3,15,16,18]. The most popular technique to solve the monotone inclusion problem is the following forward-backward splitting algorithm which is defined by Lions and Mercier [9]:

$$x_{n+1} = (I + \lambda_n A)^{-1} (I - \lambda_n B) x_n \quad \text{for all } n \in \mathbb{N}$$

(2)

where $\lambda_n$ is a step size term and $A$ and $B$ are monotone operators. If $B : H \to H$ is $1/L$-cocoercive operator and $\lambda_n \in (0, 2/L)$, the forward-backward splitting algorithm converges weakly to a solution.
of the monotone inclusion problem. It is well known that the forward-backward splitting algorithm is a generalization of the classical proximal point and proximal gradient algorithm. Let \( f : H \to \mathbb{R} \) be a differentiable convex function and let \( g : H \to \mathbb{R} \) be a proper lower semi-continuous convex function. The forward-backward splitting algorithm (2) is reduced to the proximal gradient algorithm in this scenario, which is given as follows [1]:

\[
x_{n+1} = \text{prox}_{\lambda_n g} \left( I - \lambda_n \nabla f \right) x_n \quad \text{for all } n \in \mathbb{N}
\]  

(3)

where \( \lambda_n > 0 \) is a step size. In a subsequent work, Moudafi and Oliny [13] introduced the following algorithm to solve the problem (1):

\[
\begin{align*}
  y_n &= x_n + \theta_n (x_n - x_{n-1}) \\
  x_{n+1} &= \left( I + \lambda_n A \right)^{-1} \left( y_n - \lambda_n B (x_n) \right)
\end{align*}
\]

(4)

where \( \theta_n \) is an inertial term on \([0, 1]\). They studied the weakly convergence of the algorithm, which satisfies the conditions \( \sum_{n=1}^{\infty} \theta_n \|x_n - x_{n-1}\|^2 < \infty \) and \( \lambda_n < 2/L \) where \( L \) is the Lipschitz constant of \( B \). The presence of the inertial term increases the algorithm’s performance significantly.

In optimization problems, preconditioners are often used to speed up first-order iterative optimization algorithms. For example, in gradient descent method, one takes steps in the opposite direction of the gradient of the function at the current point to find a local minimum of the real-valued function. This algorithm is given by the following way:

\[
x_{n+1} = I - \lambda_n \nabla f (x_n) \quad \text{for all } n \in \mathbb{N}.
\]

The preconditioner \( M \), which is a linear bounded operator, is applied to the algorithm as follows:

\[
x_{n+1} = I - \lambda_n M^{-1} \nabla f (x_n) \quad \text{for all } n \in \mathbb{N}.
\]

The aim of the preconditioning is to change the geometry of the space to make the level sets look like circles [6]. In this situation, the preconditioned gradient aims to get closer to the extreme point and so this accelerates the convergence. The classical splitting algorithms (2) and (4) may not generally be practical, and computing the proximal mapping \( (I + \lambda_n A)^{-1} \) could be highly costly. When we consider the preconditioned splitting algorithms with an adequate mapping \( M \), however, the algorithm becomes applicable.

In recent years, Lorenz and Pock [11] introduced the following preconditioning algorithm to solve the monotone inclusion problem:

\[
\begin{align*}
  y_n &= x_n + \theta_n (x_n - x_{n-1}) \\
  x_{n+1} &= \left( I + \lambda_n M^{-1} A \right)^{-1} \left( I - \lambda_n M^{-1} B \right) y_n
\end{align*}
\]

(5)

where \( \theta_n \) is an accelerated term on \([0, 1]\) and \( \lambda_n \) is a step size term. They proved the weak convergence of the algorithm. It is clear that Algorithm (5) is reduced to the classical forward-backward splitting algorithm (2) for \( \theta_n = 0 \) and \( M = I \).

Subsequently, in 2021, Dixit et al. [4] defined the following algorithm which is called the accelerated preconditioning forward-backward normal S-iteration (APFBNNSM):

\[
\begin{align*}
  y_n &= x_n + \theta_n (x_n - x_{n-1}) \\
  x_{n+1} &= \left( I + \lambda M^{-1} A \right)^{-1} \left( I - \lambda M^{-1} B \right) \left( (1 - \alpha_n) y_n + \alpha_n \left( I + \lambda M^{-1} A \right)^{-1} (I - \lambda M^{-1} B) y_n \right)
\end{align*}
\]

(6)

where \( \alpha_n \in (0, 1), \lambda \in [0, 1) \) and \( \theta_n \in [0, 1) \). They also proved weak convergence of the proposed algorithm under some assumptions in a real Hilbert space \( H \). For \( \theta_n = 0 \) and \( M = I \), the accelerated preconditioning forward-backward normal S-iteration (APFBNNSM) is reduced to the normal
S-iteration forward-backward splitting algorithm \[14\] (\(nS - FBSA\)):

\[
x_{n+1} = (I + \lambda A)^{-1} (I - \lambda B) \left( (1 - \alpha_n) y_n + \alpha_n (I + \lambda A)^{-1} (I - \lambda B) (y_n) \right) \quad \text{for all } n \in \mathbb{N}.
\]

In this paper, we present a new preconditioning forward-backward splitting algorithm which generalizes many existed algorithms including algorithms (2), (5) and (6), and which is more effective in image restoration. Also, we prove that the sequence generated by the proposed algorithm converges strongly to a solution of the monotone inclusion problem while the other algorithm’s sequences converge weakly to the solution of the same problem. The organization of this paper is listed as follows. In the next section, we will give some definitions and lemmas to study the convergence behaviour of the proposed algorithm. In Section 3, we will present a new preconditioning forward-backward splitting algorithm and study its convergence behaviour under mild restriction. In the last section, we will give the application of the proposed algorithm to the image restoration problem.

2. Preliminaries

In this part, we will give some definitions and lemmas which play a significant role in proving our main theorem. Let \(C\) be a nonempty subset of real Hilbert space \(H\) and \(T : C \to H\) be a mapping. A point \(x \in H\) is said to be a fixed point of \(T\) if \(Tx = x\) and the set of all fixed point of \(T\) is denoted by \(F(T)\). We use \(x_n \rightharpoonup x^*\) (resp. \(x_n \to x^*\)) to denote that \(\{x_n\}\) converges weakly (resp. strongly) to \(x^*\).

**Definition 2.1 ([2]):** Let \(C\) be a nonempty subset of a real Hilbert space \(H\) and \(x \in H\). For any \(z \in H\), if there exists a unique point \(y \in C\) such that

\[
\|y - x\| \leq \|z - x\|
\]

then \(y\) is called the metric projection of \(x\) onto \(C\) and is denoted by \(y = PCx\). If \(PCx\) exists and is uniquely determined for all \(x \in H\), then the operator \(PC : H \to C\) is called the metric projection.

It is clear that the operator \(PC\) is nonexpansive and it can be characterized by

\[
\{x - PCx, y - PCx\} \leq 0 \quad \text{for all } y \in C.
\]

Let \(A : H \to 2^H\) be a set-valued operator. If \(\langle u - v, x - y \rangle \geq 0\) for all \(u \in Ax\) and \(v \in Ay\), then \(A\) is said to be a monotone operator. If the graph of a monotone operator is not properly contained in the graph of any other monotone operators, then \(A\) is said to be a maximal monotone operator.

Let \(f : H \to (-\infty, +\infty]\) be a function. Then, \(f\) is said to be proper if there exists at least one \(x \in H\) such that \(f(x) < +\infty\). Also, the subdifferential of a proper function \(f\) is defined by

\[
\partial f (x) = \{u \in H : \langle y - x, u \rangle \leq f(y) - f(x) \quad \text{for all } y \in H\}
\]

and \(f\) is subdifferentiable at \(x \in H\), if \(\partial f(x) \neq \emptyset\). The elements of \(\partial f(x)\) are called the subgradients of \(f\) at \(x\).

**Definition 2.2 ([1]):** Let \(\Gamma_0(H)\) denote the class of all proper lower semi-continuous convex functions defined from \(H\) to \((-\infty, +\infty]\). Let \(g \in \Gamma_0(H)\) and \(\phi > 0\). The proximal operator of parameter \(\phi\) of \(g\) at \(x\) is defined by

\[
\text{prox}_{\phi g}(x) = \arg \min_{y \in H} \left\{ g(y) + \frac{1}{2\phi} \|y - x\|^2 \right\}.
\]
Example 2.3 ([1]): Let \( \phi \in (0, +\infty) \), and let \( x \in \mathbb{R}^n \). Then, the proximal operator for \( l_1 \)-norm is defined by

\[
\text{prox}_{\phi \cdot \| \cdot \|_1} (x) = (x - \phi)_+ - (-x - \phi)_+
\]

\[
= \begin{cases} 
  x_i - \phi & \text{if } x_i > \phi, \\
  0 & \text{if } -\phi \leq x_i \leq \phi, \\
  x_i + \phi & \text{if } x_i < -\phi,
\end{cases}
\]

Let \( M : H \to H \) be a bounded linear operator. \( M \) is said to be self-adjoint if \( M^* = M \) where \( M^* \) is the adjoint of operator \( M \). A self-adjoint operator is said to be positive definite if \( \langle M(x), x \rangle > 0 \) for every \( 0 \neq x \in H \) [8]. By using the self-adjoint, positive and bounded linear operator \( M \), the \( M \)-inner product is defined by

\[
\langle x, y \rangle_M = \langle x, M(y) \rangle, \quad \forall x, y \in H.
\]

In addition, the corresponding \( M \)-norm induced from the \( M \)-inner product is defined by

\[
\|x\|_M^2 = \langle x, M(x) \rangle \quad \text{for all } x \in H.
\]

Definition 2.4 ([4]): Let \( C \) be a nonempty subset of \( H, T : C \to H \) be an operator and \( M : H \to H \) be a positive-definite operator. Then \( T \) is said to be:

(i) nonexpansive operator with respect to \( M \)-norm if

\[
\|Tx - Ty\|_M \leq \|x - y\|_M, \quad \forall x, y \in H,
\]

(ii) \( M \)-cocoercive operator if \( \|Tx - Ty\|_M^2 \leq \langle x - y, Tx - Ty \rangle, \forall x, y \in H \).

Similarly, \( T \) said to be \( k \)-contraction mapping with respect to \( M \)-norm if there exists \( k \in [0, 1) \) such that

\[
\|Tx - Ty\|_M \leq k \|x - y\|_M, \quad \forall x, y \in H.
\]

Proposition 2.5 ([4]): Let \( A : H \to 2^H \) be a maximal monotone operator, \( B : H \to H \) be a \( M \)-cocoercive operator, \( M : H \to H \) be a bounded linear self-adjoint and positive-definite operator and \( \lambda \in (0, 1] \). Then we have the following properties:

(i) \( I - \lambda M^{-1}B \) is nonexpansive with respect to \( M \)-norm,
(ii) \( (I + \lambda M^{-1}A)^{-1} \) is nonexpansive with respect to \( M \)-norm,
(iii) \( f_{\lambda, M}^{A,B} = (I + \lambda M^{-1}A)^{-1}(I - \lambda M^{-1}B) \) is nonexpansive with respect to \( M \)-norm.

Proposition 2.6 ([4]): Let \( A : H \to 2^H \) be a maximal monotone operator, \( B : H \to H \) be a \( M \)-cocoercive operator, \( M : H \to H \) be a linear bounded self-adjoint and positive-definite operator and \( \lambda \in (0, \infty) \). Then \( x \in H \) is a solution of monoton inclusion problem (1) if and only if \( x \) is a fixed point of \( f_{\lambda, M}^{A,B} \).

Lemma 2.7 ([5]): Let \( C \) be a nonempty closed and convex subset of a real Hilbert space \( H \) and let \( T : C \to H \) be a nonexpansive operator with \( F(T) \neq \emptyset \). Then the mapping \( I - T \) is demiclosed at zero, that is, for any sequences \( \{x_n\} \in H \) such that \( x_n \rightharpoonup x \in H \) and \( \|x_n - Tx_n\| \to 0 \) as \( n \to \infty \), then it implies \( x \in F(T) \).

Lemma 2.8 ([1]): Let \( H \) be a real Hilbert space. Then for all \( x, y \in H \) and \( \lambda \in [0, 1] \), the following properties hold:
(i) \(|x + y| \|^2 = |x|^2 + 2(x,y) + |y|^2\),
(ii) \(|x + y| \|^2 \leq |x|^2 + 2(y,x + y)\),
(iii) \(|\lambda x + (1 - \lambda)y|^2 = \lambda|x|^2 + (1 - \lambda)|y|^2 - \lambda(1 - \lambda)|x - y|^2\).

**Lemma 2.9 ([20]):** Let \(\{s_n\}\) and \(\{\varepsilon_n\}\) be sequences of nonnegative real numbers such that

\[
s_{n+1} \leq (1 - \delta_n) s_n + \delta_n t_n + \varepsilon_n,
\]

where \(\delta_n\) is a sequence in \((0,1)\) and \(\{t_n\}\) is a real sequence. If the following conditions are hold, then

\[
\lim_{n \to \infty} s_n = 0
\]

(i) \(\sum_{n=1}^{\infty} \delta_n = \infty\),
(ii) \(\sum_{n=1}^{\infty} \varepsilon_n < \infty\),
(iii) \(\limsup_{n \to \infty} t_n \leq 0\)

**Lemma 2.10 ([12]):** Let \(\{\Phi_n\}\) be a sequence of real numbers that does not decrease at infinity such that there exists a subsequence \(\{\Phi_{n_i}\}\) of \(\{\Phi_n\}\) which satisfies \(\Phi_{n_i} < \Phi_{n_{i+1}}\) for all \(i \in \mathbb{N}\). Let \(\{\tau(n)\}_{n \geq n_0}\) be a sequence of integer which defined by:

\[
\tau(n) := \max \{l \leq n : \Phi_l < \Phi_{l+1}\}.
\]

Then the following are satisfied:

(i) \(\tau(n_0) \leq \tau(n_0 + 1) \leq \cdots \) and \(\tau(n) \to \infty\),
(ii) \(\Phi_{\tau(n)} \leq \Phi_{\tau(n)+1}\) and \(\Phi_n \leq \Phi_{\tau(n)+1}\), for all \(n \geq n_0\).

### 3. Main results

In this section, we define a new preconditioning forward-backward splitting algorithm and prove its strong convergence in real Hilbert space.

**Theorem 3.1:** Let \(M : H \to H\) be a bounded linear self-adjoint and positive-definite operator, \(A : H \to 2^H\) be a maximal monotone operator and \(B : H \to H\) be a \(M\)-cocoercive operator such that \(\Omega = (A + B)^{-1}(0)\) is nonempty. Let \(f\) be a \(k\)-contraction mapping on \(H\) with respect to \(M\)-norm and let \(\lambda \in (0,1)\). Let \(\{x_n\}\) be a sequence generated by

\[
\begin{cases}
  x_0, x_1 \in H, \\
  y_n = x_n + \theta_n (x_n - x_{n-1}), \\
  z_n = (I + \lambda M^{-1}A)^{-1} (I - \lambda M^{-1}B) \left((1 - \alpha_n) y_n + \alpha_n (I + \lambda M^{-1}A)^{-1} (I - \lambda M^{-1}B) (y_n) \right), \\
  x_{n+1} = \beta_n f(z_n) + (1 - \beta_n) (I + \lambda M^{-1}A)^{-1} (I - \lambda M^{-1}B) (z_n),
\end{cases}
\]

where \(\{\theta_n\} \subset [0,1)\) is a sequence with \(\theta \in [0,1)\) and \(\{\alpha_n\}, \{\beta_n\} \in (0,1)\) such that the following conditions hold:

(i) \(0 < a \leq \alpha_n \leq b < 1\) for some \(a, b \in \mathbb{R}\),
(ii) \(0 < c \leq \beta_n \leq d < 1\) for some \(c, d \in \mathbb{R}\),
(iii) \(\sum_{n=1}^{\infty} \theta_n \|x_n - x_{n-1}\|_M < \infty\),
(iv) \(\lim_{n \to \infty} \beta_n = 0, \sum_{n=1}^{\infty} \beta_n = \infty\).

Then the sequence \(\{x_n\}\) converges strongly to a point \(p\) in \(\Omega\) where \(p = P_\Omega f(p)\).
Proof: We will obtain the proof by dividing it into the following steps.

Step 1: In this step, we show that the sequence \( \{x_n\} \) is bounded. Let \( p \in \Omega \) such that \( p = P_{\Omega} f(p) \). From Proposition (26), we know that \( J_{\lambda,M}^{A,B}(p) = p \). Since \( J_{\lambda,M}^{A,B} \) is nonexpansive with respect to \( M \)-norm, we obtain the following from algorithm (7):

\[
\|y_n - p\|_M = \|x_n + \theta_n (x_n - x_{n-1}) - p\|_M \\
\leq \|x_n - p\|_M + \theta_n \|x_n - x_{n-1}\|_M
\]

and

\[
\|z_n - p\|_M = \|J_{\lambda,M}^{A,B} (1 - \alpha_n) y_n + \alpha_n J_{\lambda,M}^{A,B} (y_n) - J_{\lambda,M}^{A,B} (p)\|_M \\
\leq \|(1 - \alpha_n) y_n + \alpha_n J_{\lambda,M}^{A,B} (y_n) - J_{\lambda,M}^{A,B} (p)\|_M \\
= \|(1 - \alpha_n) (y_n - p) + \alpha_n (J_{\lambda,M}^{A,B} (y_n) - J_{\lambda,M}^{A,B} (p))\|_M \\
\leq (1 - \alpha_n) \|y_n - p\|_M + \alpha_n \|J_{\lambda,M}^{A,B} (y_n) - J_{\lambda,M}^{A,B} (p)\|_M \\
\leq (1 - \alpha_n) \|y_n - p\|_M + \alpha_n \|y_n - p\|_M = \|y_n - p\|_M .
\]

Since \( f \) is \( k \)-contractive mapping with respect to \( M \)-norm, we also obtain the following by combining (8) and (9):

\[
\|x_{n+1} - p\|_M = \|\beta_n f(z_n) + (1 - \beta_n) J_{\lambda,M}^{A,B} (z_n) - p\|_M \\
\leq \|\beta_n (f(z_n) - f(p) + f(p)) + (1 - \beta_n) (J_{\lambda,M}^{A,B} (z_n) - p)\|_M \\
\leq \beta_n \|f(z_n) - f(p)\|_M + \beta_n \|f(p) - p\|_M + (1 - \beta_n) \|J_{\lambda,M}^{A,B} (z_n) - p\|_M \\
\leq \beta_n k \|z_n - p\|_M + \beta_n \|f(p) - p\|_M + (1 - \beta_n) \|z_n - p\|_M \\
= (1 - \beta_n (1 - k)) \|z_n - p\|_M + \beta_n \|f(p) - p\|_M \\
\leq (1 - \beta_n (1 - k)) \|x_n - p\|_M + \beta_n \frac{\theta_n}{\beta_n} \|x_n - x_{n-1}\|_M + \beta_n \|f(p) - p\|_M .
\]

From conditions (ii) and (iii), we have \( \lim_{n \to \infty} \frac{\theta_n}{\beta_n} \|x_n - x_{n-1}\|_M = 0 \). So, there exists a positive constant \( K_1 > 0 \) such that \( \frac{\theta_n}{\beta_n} \|x_n - x_{n-1}\|_M \leq K_1 \). It follows from (10) that

\[
\|x_{n+1} - p\|_M \leq (1 - \beta_n (1 - k)) \|x_n - p\|_M + \beta_n (K_1 + \|f(p) - p\|_M) \\
= (1 - \beta_n (1 - k)) \|x_n - p\|_M + \beta_n (1 - k) \left( \frac{K_1 + \|f(p) - p\|_M}{1 - k} \right) \\
\leq \max \left\{ \|x_n - p\|_M , \frac{K_1 + \|f(p) - p\|_M}{1 - k} \right\} \\
\leq \max \left\{ \|x_1 - p\|_M , \frac{K_1 + \|f(p) - p\|_M}{1 - k} \right\}
\]

for all \( n \geq 1 \). This means that \( \{x_n\} \) is bounded so \( \{y_n\}, \{z_n\} \) are also bounded.
Step 2: Next, we have to show that $x_n \to p = P_M f(p)$. Indeed, using Lemma 2.8 we find the following for all $n \geq 1$:

$$
\|y_n - p\|_M^2 = \|x_n + \theta_n (x_n - x_{n-1}) - p\|_M^2
\leq \|x_n - p\|_M^2 + 2\theta_n \|x_n - p\|_M \|x_n - x_{n-1}\|_M + \theta_n^2 \|x_n - x_{n-1}\|_M^2
$$

and

$$
\|z_n - p\|_M^2 = \|J_{\lambda,M}^{A,B} ((1 - \alpha_n) y_n + \alpha_n J_{\lambda,M}^{A,B} (y_n)) - p\|_M^2
\leq \|(1 - \alpha_n) y_n + \alpha_n J_{\lambda,M}^{A,B} (y_n) - p\|_M^2
\leq \|y_n - p\|_M^2 - \alpha_n (1 - \alpha_n) \|J_{\lambda,M}^{A,B} (y_n) - y_n\|_M^2
\leq \|y_n - p\|_M^2 - \alpha_n (1 - \alpha_n) \|J_{\lambda,M}^{A,B} (y_n) - y_n\|_M^2
\leq \|y_n - p\|_M^2.
$$

It follows from (11), (12), and Lemma 2.8 that

$$
\|x_{n+1} - p\|_M^2 = \|\beta_n f(z_n) + (1 - \beta_n) J_{\lambda,M}^{A,B} (z_n) - p\|_M^2
\leq \|\beta_n (f(z_n) - f(p)) + (1 - \beta_n) (J_{\lambda,M}^{A,B} (z_n) - p) + \beta_n (f(p) - p)\|_M^2
\leq \|\beta_n (f(z_n) - f(p)) + (1 - \beta_n) (J_{\lambda,M}^{A,B} (z_n) - p)\|_M^2 + 2\beta_n \|f(p) - p, x_{n+1} - p\|_M
\leq \|\beta_n (f(z_n) - f(p))\|_M^2 + (1 - \beta_n) \|J_{\lambda,M}^{A,B} (z_n) - p\|_M^2 + 2\beta_n \|f(p) - p, x_{n+1} - p\|_M
\leq \|f(z_n) - f(p)\|_M^2 + (1 - \beta_n) \|z_n - p\|_M^2 + 2\beta_n \|f(p) - p, x_{n+1} - p\|_M
\leq (1 - \beta_n (1 - k^2)) \|z_n - p\|_M^2 + 2\beta_n \|f(p) - p, x_{n+1} - p\|_M
\leq (1 - \beta_n (1 - k^2)) \|x_n - p\|_M^2 + 2\theta_n \|x_n - x_{n-1}\|_M \|x_n - x_{n-1}\|_M
+ \theta_n^2 \|x_n - x_{n-1}\|_M^2 + 2\beta_n \|f(p) - p, x_{n+1} - p\|_M
\leq (1 - \beta_n (1 - k^2)) \|x_n - p\|_M^2 + \theta_n \|x_n - x_{n-1}\|_M \|x_n - p\|_M
+ \theta_n \|x_n - x_{n-1}\|_M + 2\beta_n \|f(p) - p, x_{n+1} - p\|_M
$$

for all $n \geq 1$. Since $\lim_{n \to \infty} \theta_n \|x_n - x_{n-1}\|_M = 0$, there exists a positive constant $K_2 > 0$ such that $\theta_n \|x_n - x_{n-1}\|_M \leq K_2$. From inequality (13) we observe that, for all $n \geq 1$,

$$
\|x_{n+1} - p\|_M^2 \leq (1 - \beta_n (1 - k^2)) \|x_n - p\|_M^2 + 3K_3 \|x_n - x_{n-1}\|_M
+ 2\beta_n \|f(p) - p, x_{n+1} - p\|_M
= (1 - \beta_n (1 - k^2)) \|x_n - p\|_M^2 + 3K_3 \|x_n - x_{n-1}\|_M
+ \beta_n (1 - k^2) \frac{2}{1 - k^2} \|f(p) - p, x_{n+1} - p\|_M.
$$
where \( K_3 = \sup_{n \geq 1} \{\|x_n - p\|_M, K_2\} \). In the above inequality, if we take \( \delta_n = \beta_n (1 - k^2) \), \( s_n = \|x_n - p\|_M^2 \), \( t_n = \frac{2}{(1 - k^2)}(f(p) - p, x_{n+1} - p)_M \) and \( \varepsilon_n = 3K_3\theta_n \|x_n - x_{n-1}\|_M \) then we have \( s_{n+1} \leq (1 - \delta_n)s_n + \delta_n t_n + \varepsilon_n \) for all \( n \geq 1 \).

Now, we want to show that \( \limsup_{n \to \infty} (f(p) - p, x_{n+1} - p)_M \leq 0 \). So, we take into account two cases to complete the proof.

First, we suppose that there exists \( n_0 \in \mathbb{N} \) such that \( \|x_n - p\|_M \) is a nonincreasing sequence. So, the sequence \( \{\|x_n - p\|_M\} \) is convergent since it is bounded from below by 0. By using the condition (iv), we have \( \sum_{n=1}^{\infty} \delta_n = \infty \). We claim that \( \limsup_{n \to \infty} (f(p) - p, x_{n+1} - p)_M \leq 0 \). By combining (11) and (12) with Lemma 2.8, we get

\[
\|x_{n+1} - p\|_M^2 = \|f_n f(z_n) + (1 - \beta_n) f_{\lambda, M}^{A,M}(z_n) - p\|_M^2 \\
= \beta_n \|f(z_n) - p\|_M^2 + (1 - \beta_n) \|f_{\lambda, M}^{A,M}(z_n) - p\|_M^2 - \beta_n (1 - \beta_n) \|f(z_n) - f_{\lambda, M}^{A,M}(z_n)\|_M^2 \\
\leq \beta_n \|f(z_n) - p\|_M^2 + (1 - \beta_n) \|z_n - p\|_M^2 \\
\leq \beta_n \|f(z_n) - p\|_M^2 + (1 - \beta_n) \|x_n - p\|_M^2 - \alpha_n (1 - \alpha_n) \|f_{\lambda, M}^{A,M}(y_n) - y_n\|_M^2 \\
\leq \beta_n \|f(z_n) - p\|_M^2 + (1 - \beta_n) \|x_n - p\|_M^2 + 2\theta_n \|x_n - p\|_M \|x_n - x_{n-1}\|_M \\
+ \theta_n^2 \|x_n - x_{n-1}\|_M^2 - \alpha_n (1 - \alpha_n) \|f_{\lambda, M}^{A,M}(y_n) - y_n\|_M^2 \\
= \beta_n \|f(z_n) - p\|_M^2 + (1 - \beta_n) \|x_n - p\|_M^2 + 2 (1 - \beta_n) \theta_n \|x_n - p\|_M \|x_n - x_{n-1}\|_M \\
(1 - \beta_n) \theta_n \|x_n - x_{n-1}\|_M^2 - \alpha_n (1 - \alpha_n) (1 - \beta_n) \|f_{\lambda, M}^{A,M}(y_n) - y_n\|_M^2 \\
\end{align*}

for all \( n \geq 1 \). This implies that

\[
\alpha_n (1 - \alpha_n) (1 - \beta_n) \|f_{\lambda, M}^{A,M}(y_n) - y_n\|_M^2 \\
\leq \beta_n \left( \|f(z_n) - p\|_M^2 - \|x_n - p\|_M^2 \right) - \|x_{n+1} - p\|_M^2 + \|x_n - p\|_M^2 \\
+ (1 - \beta_n) \theta_n \|x_n - x_{n-1}\|_M \left( 2 \|x_n - p\|_M^2 + \theta_n \|x_n - x_{n-1}\|_M \right).
\]

Due to conditions (iii), (iv) and the convergence of the sequence \( \{\|x_n - p\|_M\} \), we conclude that

\[
\lim_{n \to \infty} \left\| f_{\lambda, M}^{A,M}(y_n) - y_n \right\|_M = 0.
\]

(15)

On the other hand, the following are obtained:

\[
\lim_{n \to \infty} \|y_n - x_n\|_M = \lim_{n \to \infty} \theta_n \|x_n - x_{n-1}\|_M = 0
\]

(16)

and

\[
\|z_n - y_n\|_M = \|z_n - f_{\lambda, M}^{A,B}(y_n) + f_{\lambda, M}^{A,B}(y_n) - y_n\|_M \\
\leq \|z_n - f_{\lambda, M}^{A,B}(y_n)\|_M + \left\| f_{\lambda, M}^{A,B}(y_n) - y_n \right\|_M
\]

\[
\|z_n - y_n\|_M \\ 
\]
It follows from (19) that
\[ \limsup_{n \to \infty} \| \alpha \beta n J_{\lambda,\mathcal{M}}(y_n) - y_n \|_M + \| J_{\lambda,\mathcal{M}}(y_n) - y_n \|_M \]

\[ = \| \alpha_n (J_{\lambda,\mathcal{M}}(y_n) - y_n) \|_M + \| J_{\lambda,\mathcal{M}}(y_n) - y_n \|_M \]

\[ = (1 + \alpha_n) \| J_{\lambda,\mathcal{M}}(y_n) - y_n \|_M, \]

which implies
\[ \lim_{n \to \infty} \| z_n - y_n \|_M = \lim_{n \to \infty} \| J_{\lambda,\mathcal{M}}(y_n) - y_n \|_M = 0. \]  

By using (15), (16), (17) and condition (iv) we can see
\[ \| x_{n+1} - y_n \|_M = \| x_{n+1} - J_{\lambda,\mathcal{M}}(y_n) + J_{\lambda,\mathcal{M}}(y_n) - y_n \|_M \]

\[ \leq \| x_{n+1} - J_{\lambda,\mathcal{M}}(y_n) \|_M + \| J_{\lambda,\mathcal{M}}(y_n) - y_n \|_M \]

\[ = \| \beta_n f(z_n) + (1 - \beta_n) J_{\lambda,\mathcal{M}}(z_n) - J_{\lambda,\mathcal{M}}(y_n) \|_M + \| J_{\lambda,\mathcal{M}}(y_n) - y_n \|_M \]

\[ \leq \beta_n \| f(z_n) - J_{\lambda,\mathcal{M}}(z_n) \|_M + \| J_{\lambda,\mathcal{M}}(z_n) - J_{\lambda,\mathcal{M}}(y_n) \|_M + \| J_{\lambda,\mathcal{M}}(y_n) - y_n \|_M \]

which implies
\[ \lim_{n \to \infty} \| x_{n+1} - y_n \|_M = 0. \]  

(18)

So, from inequalities (16) and (18), we have
\[ \| x_{n+1} - x_n \|_M \leq \| x_{n+1} - y_n \|_M + \| y_n - x_n \|_M \lim_{n \to \infty} \| x_{n+1} - x_n \|_M = 0. \]

Now, we get
\[ \limsup_{n \to \infty} \{ f(p) - p, x_{n+1} - p \}_M = t. \]

Since the sequence \( \{ x_n \} \) is bounded, there exists a subsequence \( \{ x_{n_i} \} \) of \( \{ x_n \} \) such that \( x_{n_i} \to \nu \) and \( \lim_{i \to \infty} \{ f(p) - p, x_{n_i+1} - p \}_M = t. \)

By using (15) and (16) we can write
\[ \| J_{\lambda,\mathcal{M}}(x_n) - x_n \|_M = \| J_{\lambda,\mathcal{M}}(x_n) - x_n + y_n - y_n + J_{\lambda,\mathcal{M}}(y_n) - J_{\lambda,\mathcal{M}}(y_n) \|_M \]

\[ \leq 2 \| y_n - x_n \|_M + \| J_{\lambda,\mathcal{M}}(y_n) - y_n \|_M . \]

This implies that
\[ \lim_{n \to \infty} \| J_{\lambda,\mathcal{M}}(x_n) - x_n \|_M = 0. \]

In this case, it is clear from Lemma 2.7 that \( \nu \in F(J_{\lambda,\mathcal{M}}) \). On the other hand, since \( \| x_{n+1} - x_n \|_M \to 0 \) as \( n \to \infty \) and \( x_{n_i} \to \nu \), we have \( x_{n_i+1} \to \nu \). Moreover, by combining \( p = P_{\Omega} f(p) \) and property of the metric projection operators we can get
\[ \lim_{i \to \infty} \{ f(p) - p, x_{n_i+1} - p \}_M = \{ f(p) - p, \nu - p \}_M \leq 0. \]

Then this implies that
\[ \limsup_{n \to \infty} \{ f(p) - p, x_{n+1} - p \}_M \leq 0. \]  

(19)

It follows from (19) that \( \limsup_{n \to \infty} t_n \leq 0 \). As a result, we obtain that \( x_n \to p \).
Secondly, we assume that there exists \( n_0 \in \mathbb{N} \) such that \( \{\|x_n - p\|_M\}_{n \geq n_0} \) is a monotone decreasing sequence. Let us denote \( \Phi_n = \|x_n - p\|_M^2 \) for all \( n \geq 1 \). For this reason, there exists a subsequence \( \{\Phi_j\} \) of \( \{\Phi_n\} \) such that \( \Phi_{n_j} < \Phi_{n_{j+1}} \) for all \( n \geq n_0 \). Define \( \tau : \{n : n \geq n_0\} \rightarrow \mathbb{N} \) by

\[
\tau (n) = \max \{l \in \mathbb{N} : l \leq n, \Phi_l \leq \Phi_{l+1}\}.
\]

It is clear that the sequence \( \tau \) is nondecreasing. By Lemma 2.10 we say that \( \Phi_{\tau(n)} \leq \Phi_{\tau(n)+1} \) for all \( n \geq n_0 \). So, we have

\[
\limsup_{n \to \infty} \|\Phi_{\tau(n)} - p\|_M^2 \leq 0.
\]

Also, we have

\[
\|\Phi_{\tau(n)} - p\|_M^2 \to 0 \quad \text{and} \quad \|\Phi_{\tau(n)+1} - p\|_M \to 0 \quad \text{as} \quad n \to \infty.
\]

So, by using (20) and Lemma 2.10, we conclude that

\[
\|\Phi_n - p\|_M \leq \|\Phi_{\tau(n)+1} - p\|_M \to 0 \quad \text{as} \quad n \to \infty.
\]

Hence, we obtain that \( x_n \to p \), and the proof is completed.

\[\square\]

4. Application to convex minimization problem

Now, we consider the following convex minimization problem given as a sum of two convex functions:

\[
h(x^*) + g(x^*) = \min_{x \in H} \{h(x) + g(x)\}.
\]

Let \( h : H \to \mathbb{R} \) be differentiable with \( L_h \)-Lipschitz gradient which is Lipschitz constant of \( \nabla h \). If \( \nabla h \) is \( L_h \)-Lipschitz continuous, then the Baillon-Haddad Theorem states that \( \nabla h \) is cocoercive with respect to \( L_h^{-1} \). Furthermore, if \( g : H \to \mathbb{R} \) is a proper convex and lower semi-continuous function then \( \partial g \) is maximal monotone see, for detail [1]. A point \( x^* \) is a solution of minimization problem (21) if and only if \( 0 \in \nabla h(x^*) + \partial g(x^*) \). Then for any \( \lambda > 0 \) we have

\[
0 \in \lambda \nabla h(x^*) + \lambda \partial g(x^*)
\]

\[
\iff 0 \in \lambda L_h^{-1} \nabla h(x^*) + \lambda L_h^{-1} \partial g(x^*)
\]

\[
x^* - \lambda L_h^{-1} \nabla h(x^*) \in x^* + \lambda L_h^{-1} \partial g(x^*)
\]

\[
x^* = (I + \lambda L_h^{-1} \partial g)^{-1} (I - \lambda L_h^{-1} \nabla h)(x^*)
\]

In Theorem 3.1, set \( A = \partial g, B = \nabla h \) and \( M(x) = L_h x \). As a result, we can deduce the following corollary.

**Corollary 4.1:** Let \( h : H \to \mathbb{R} \) be a differentiable and convex function with \( L_h \)-Lipschitz gradient and \( g : H \to \mathbb{R} \) be a proper convex and lower semi-continuous function. Assume that the solution set of
convex minimization problem (21) is nonempty. The parameters \( \{\theta_n\} \subset [0, \theta] \) and \( \{\alpha_n\}, \{\beta_n\} \in (0, 1) \) satisfy the same condition as in Theorem 3.1. Let \( \{x_n\} \) be a sequence generated by

\[
\begin{aligned}
    x_0, x_1 & \in H, \\
    y_n &= x_n + \theta_n (x_n - x_{n-1}), \\
    z_n &= (I + \lambda L_h^{-1} \partial g)^{-1} \left( I - \lambda L_h^{-1} \nabla h \right) \left( (1 - \alpha_n) y_n + \alpha_n (I + \lambda L_h^{-1} \partial g)^{-1} (I - \lambda L_h^{-1} \nabla h) y_n \right), \\
    x_{n+1} &= \beta_n f(z_n) + (1 - \beta_n) (I + \lambda L_h^{-1} \partial g)^{-1} (I - \lambda L_h^{-1} \nabla h) z_n.
\end{aligned}
\]  

Then \( \{x_n\} \) converges strongly to a \( x^* \) solution of the convex minimization problem.

Figure 1. (a) Cameraman image, (b) degraded image, (c) Algorithm (5), (d) Algorithm (6) and (e) Algorithm (22).
5. Applications to image restoration problem

This section aims to show the application of the new preconditioning forward-backward algorithm to the image restoration problem. In addition, we conduct a comparison of Algorithm (22) with Algorithm (6) and Algorithm (5).

The inverse problem of the following form can be used to define a general image restoration problem:

\[ b = Ax + v, \]  

(23)

where \( x \in \mathbb{R}^d \) is original image, \( A : \mathbb{R}^d \rightarrow \mathbb{R}^m \) is a linear operator, \( b \in \mathbb{R}^m \) is observed image and \( v \) is the additive noise. It is well known that problem (23) is roughly comparable to a number of different optimization problems. Also, the \( l_1 \)-norm is commonly used as a regularization tool to solve
Table 1. SNR values for the Cameraman image.

| No. iterations | Algorithm (22) | Algorithm (6) | Algorithm (5) |
|----------------|----------------|---------------|---------------|
| 1              | 35.358278      | 34.805570     | 34.447978     |
| 5              | 39.041491      | 37.647298     | 36.739191     |
| 10             | 41.596885      | 39.737838     | 38.459428     |
| 25             | 45.483306      | 43.327179     | 41.672360     |
| 50             | 48.676063      | 46.328902     | 44.557987     |
| 100            | 52.156726      | 49.590691     | 47.648439     |
| 250            | 56.836376      | 54.268904     | 52.177390     |
| 500            | 60.117373      | 57.738798     | 55.736154     |
| 1000           | 63.000553      | 60.935109     | 59.103851     |

Table 2. PSNR values for the Cameraman image.

| No. iterations | Algorithm (22) | Algorithm (6) | Algorithm (5) |
|----------------|----------------|---------------|---------------|
| 1              | 79.479563      | 79.203224     | 79.024436     |
| 5              | 81.227461      | 80.550059     | 80.109624     |
| 10             | 82.446397      | 81.538934     | 80.921214     |
| 25             | 84.343144      | 83.273703     | 82.457524     |
| 50             | 85.906469      | 84.748400     | 83.871979     |
| 100            | 87.620117      | 86.348858     | 85.391488     |
| 250            | 89.962710      | 88.668544     | 87.623570     |
| 500            | 91.621486      | 90.416429     | 89.404532     |
| 1000           | 93.096925      | 92.035617     | 91.106244     |

Table 3. SNR values for the Mountain image.

| No. iterations | Algorithm (22) | Algorithm (6) | Algorithm (5) |
|----------------|----------------|---------------|---------------|
| 1              | 33.150494      | 33.079983     | 32.970274     |
| 5              | 33.975156      | 33.862282     | 33.758195     |
| 10             | 34.235762      | 34.094985     | 33.975192     |
| 25             | 34.593426      | 34.411929     | 34.271498     |
| 50             | 34.919419      | 34.686252     | 34.524013     |
| 100            | 35.350730      | 35.031797     | 34.822088     |
| 250            | 36.222684      | 35.692589     | 35.362060     |
| 500            | 37.213319      | 36.460288     | 35.977868     |
| 1000           | 38.502561      | 37.530003     | 36.867338     |

Table 4. PSNR values for the Cameraman image.

| No. iterations | Algorithm (22) | Algorithm (6) | Algorithm (5) |
|----------------|----------------|---------------|---------------|
| 1              | 68.908473      | 68.873217     | 68.818362     |
| 5              | 69.320821      | 69.264381     | 69.212336     |
| 10             | 69.451135      | 69.380740     | 69.320840     |
| 25             | 69.629990      | 69.539228     | 69.469005     |
| 50             | 69.793020      | 69.677597     | 69.595278     |
| 100            | 70.008724      | 69.849221     | 69.744343     |
| 250            | 70.44861       | 70.181534     | 70.014391     |
| 500            | 70.940457      | 70.563717     | 70.322399     |
| 1000           | 71.585632      | 71.098913     | 70.767362     |

these types of problems. As a result, the image restoration problem (23) may be reduced to a $l_1$-regularization problem, which can be expressed as

$$\min_{x \in \mathbb{R}^d} \left\{ \frac{1}{2} \|Ax - b\|^2 + \rho \|x\|_1 \right\}, \tag{24}$$

where $\rho > 0$ is a regularization parameter. On the other hand, for $h(x) = \frac{1}{2} \|Ax - b\|^2$ and $g(x) = \rho \|x\|_1$, the convex minimization problem can be reduced to the $l_1$-regularization problem. According
to this selection, the Lipschitz gradient of $h$ is the following form $\nabla h(x) = A^T(Ax - b)$, where $A^T$ is the transpose of $A$.

Now, we show that Algorithm (22) is used to solve the image restoration problem (23) and also that this algorithm is compared to Algorithm (6) and Algorithm (5). In all comparison, we consider the motion and gaussian blur functions and add random noise to the test images Cameraman and Mountain. We use the Matlab functions fspecial('motion',15,60) and fspecial('gaussian',9,4) to add blur and the Matlab function ‘randn’ to add random noise. We use signal-to-noise ratio (SNR) and peak signal-to-noise ratio (PSNR) values to measure the quality of the restored images where SNR is

![Figure 4. (a) Mountain image, (b) degraded image, (c) Algorithm (5), (d) Algorithm (6) and (e) Algorithm (22).](image)
Figure 5. Graphic of SNR values for the Mountain image.

![Figure 5](image5)

Figure 6. Graphic of PSNR values for the Mountain image.

![Figure 6](image6)

defined by

$$\text{SNR} = 20 \log \frac{\|x\|_2}{\|x - x_n\|_2}$$

and where PSNR is defined by

$$\text{PSNR} (x_n) = 10 \log \left( \frac{255^2}{\text{MSE}} \right),$$

where $x$ and $x_n$ are the original image and the estimated image at iteration $n$, respectively, $\text{MSE} = \frac{1}{M} \|x - x_n\|_2$ and $M$ are the number of image samples. All algorithms are implemented in MATLAB R2020a running on a Dell with Intel (R) Core (TM) i5 CPU and 8 GB of RAM.
First of all, by using Cameraman image and motion blur function, we compare Algorithm (22) with Algorithm (6) and Algorithm (5). We set $\alpha_n = \frac{1}{2}$, $\theta_n = \frac{1}{10}$, $\beta_n = \frac{1}{10^n}$, $\lambda = 0.99$, $f(x) = 0.99x$ and the regularization parameter $\rho = 0.0001$. Figures 1–3 and Tables 1 and 2 provide the visual and numerical results corresponding to these selections.

Now, using Mountain image and gaussian blur function, we compare Algorithm (22) with Algorithm (6) and Algorithm (5). We take $\alpha_n = \frac{1}{2}$, $\theta_n = \frac{1}{2}$, $\beta_n = \frac{1}{2^n}$, $\lambda = 0.99$, and $f(x) = 0.9999x$. The numerical and visual results corresponding to these selections are shown in Figures 4–6 and Tables 3 and 4.

### 6. Conclusion

In this study, we suggested a preconditioning forward-backward algorithm which generalizes some existing algorithms to handle the image restoration problem effectively. In addition, while the weak convergence theorems were proved for the other algorithms we generalized, we demonstrated the strong convergence theorems for our algorithm. Experimental results demonstrate that Algorithm (22) restores images with a greater SNR and PSNR than Algorithm (5) and Algorithm (6), indicating that its image restoration performance is superior to Algorithm (5) and Algorithm (6).

### Disclosure statement

No potential conflict of interest was reported by the author(s).

### ORCID

Ebru Altiparmak [http://orcid.org/0000-0001-6722-0807](http://orcid.org/0000-0001-6722-0807)
Ibrahim Karahan [http://orcid.org/0000-0001-6191-7515](http://orcid.org/0000-0001-6191-7515)

### References

[1] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Space*, CMS Books in Mathematics Springer, New York, 2011.

[2] A. Cegielski, *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, Lecture Notes in Mathematics, Springer, Heidelberg, 2012.

[3] P.L. Combettes and R. Wajs, *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul. 4 (2015), pp. 1168–1200.

[4] A. Dixit, D.R. Sahu, P. Gautam, T. Som, and J.C. Yao, *An accelerated forward backward splitting algorithm for solving inclusion problems with applications to regression and link prediction problems*, J. Nonlinear Var. Anal. 5(1) (2021), pp. 79–101.

[5] K. Goebel and W.A. Kirk, *Topics in Metric Fixed Point Theory*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, UK, 1990.

[6] D.M. Himmelblau, *Applied Nonlinear Programming*, McGraw-Hill, New York, 1972. pp. 78–83.

[7] D. Kitkuan, P. Kumam, and J. Martinez-Moreno, *Generalized Halpern-type forward-backward splitting methods for convex minimization problems with application to image restoration problems*, Optimization 69 (2019), pp. 1557–1581.

[8] B.V. Limaye, *Functional Analysis*, New Age, New Delhi, 1996.

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

[10] G. Lopez, V. Martin-Márquez, F. Wang, and H.K. Xu, *Forward-backward splitting methods for accretive operators in Banach spaces*, Abstr. Appl. Anal. 2012 (2012), pp. 109236.

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

[12] P.E. Mainge, *Strong convergence of projected subgradient methods for nonsmooth and nonstrictly convex minimization*, Set Valued Anal. 16 (2008), pp. 899–912.

[13] A. Moudafi and M. Oliny, *Convergence of a splitting inertial proximal method for monotone operators*, J. Comput. Appl. Math. 155 (2003), pp. 447–454.

[14] D.R. Sahu, J.C. Yao, M. Verma, and K.K. Shukla, *Convergence rate analysis of proximal gradient methods with applications to composite minimization problems*, Optimization 70 (2021), pp. 75–100.
[15] K. Sitthithakerngkiet, P. Deepho, and P. Kumam, A hybrid viscosity algorithm via modify the hybrid steepest descent method for solving the split variational inclusion image reconstruction and fixed point problems, Appl. Math. Comput. 250 (2015), pp. 986–1001.

[16] S. Sra, S. Nowozin, and S.J. Wright, Optimization for Machine Learning, MIT Press, London, 2012.

[17] P. Sunthrayuth and P. Cholamjiak, Iterative methods for solving quasi-variational inclusion and fixed point problem in q-uniformly smooth Banach spaces, Numer. Algorithms 78 (2018), pp. 1019–1044.

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

[19] P. Tseng, A modified forward–backward splitting method for maximal monotone mappings, SIAM J. Control Optim. 38 (2000), pp. 431–446.

[20] H.K. Xu, Another control condition in an iterative method for nonexpansive mappings, Bull. Aust. Math. Soc. 65 (2002), pp. 109–113.