mathematics

MDPI

Article

A New Accelerated Viscosity Iterative Method for an Infinite Family of Nonexpansive Mappings with Applications to Image Restoration Problems

Jenwit Puangpee [1] and Suthep Suantai [2,*]

1 PhD Degree Program in Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand; jenwit.pp@hotmail.com

2 Data Science Research Center, Research Center in Mathematics and Applied Mathematics, Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand

* Correspondence: suthep.s@cmu.ac.th

Received: 16 March 2020; Accepted: 9 April 2020; Published: 16 April 2020

Abstract: The image restoration problem is one of the popular topics in image processing which is extensively studied by many authors because of its applications in various areas of science, engineering and medical image. The main aim of this paper is to introduce a new accelerated fixed algorithm using viscosity approximation technique with inertial effect for finding a common fixed point of an infinite family of nonexpansive mappings in a Hilbert space and prove a strong convergence result of the proposed method under some suitable control conditions. As an application, we apply our algorithm to solving image restoration problem and compare the efficiency of our algorithm with FISTA method which is a popular algorithm for image restoration. By numerical experiments, it is shown that our algorithm has more efficiency than that of FISTA.

Keywords: image restoration problem; viscosity; inertial; nonexpansive mapping

1. Introduction

Let us first consider a simple linear inverse problem as the following form:

$$Ax = b + w, \tag{1}$$

where $x \in \mathbb{R}^{n\times 1}$ is the solution of the problem to be approximated, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^{m\times 1}$ are known and $w \in \mathbb{R}^{m\times 1}$ is an additive noise vector. Such problems (1) arise in various applications such as the image and signal processing problems, astrophysical problems and data classification problems.

Further, one of their well-known applications is the problem of approximating the original image from the observed blurred and noisy image which is known as the image restoration problem. In this problem, x, A and b represent the original image, blur operator and observed image, respectively.

The purpose of the image restoration problem is to minimize the additive noise in which the classical estimator is the least squares (LS) given as follows:

$$\hat{x} := \arg\min_x \|Ax - b\|_2^2, \tag{2}$$

where $\|\cdot\|_2$ is ℓ^2-norm. However, this model still has some ill-conditions in the case that the least square solution has a huge norm which is thus meaningless. In 1977, Tikhonov and Arsenin [1] improved this ill-posed problem by introducing the regularization techniques which are known as the Tikhonov regularization (TR) model and it is of the following form:

$$\hat{x} := \arg\min_{x}\{\|Ax - b\|_2^2 + \lambda\|Lx\|_2^2\}, \tag{3}$$

for some regularization parameter $\lambda > 0$ and Tikhonov matrix L.

On the other hand, another successful regularization method for improvement of Tikhonov regularization is known as the least absolute shrinkage and selection operator (LASSO) which was introduced by Tibshirani (1996). The method is to find a solution

$$\hat{x} := \arg\min_{x}\{\|Ax - b\|_2^2 + \lambda\|x\|_1\}, \tag{4}$$

where $\|\cdot\|_1$ is ℓ^1-norm. The LASSO can be applied to regression problems and image restoration problems (see [2,3] for examples).

For solving (3) and (4), we extend them to a general naturally formulation, that is, the problem of finding the minimizer of sum of two functions:

$$\hat{x} := \arg\min_{x}\{h(x) + g(x)\}. \tag{5}$$

In order to solve (5), we assume the following:

1. $h : \mathbb{R}^n \to \mathbb{R}$ is a smooth convex loss function and differentiable with L-Lipschitz continuous gradient, where $L > 0$, i.e.,
$$\|\nabla h(x) - \nabla h(y)\| \le L\|x - y\|, \quad \forall x, y \in \mathbb{R}^n;$$
2. $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a proper convex and lower semi-continuous function.

We here denote the set of all solutions of above problem by $\arg\min(h + g)$. It is well-known that the solution of (5) can be reformulated as the problem of finding a zero-solution $\hat{x}$ such that

$$0 \in \partial g(\hat{x}) + \nabla h(\hat{x}), \tag{6}$$

where ∂g is the subdifferential of function g and ∇h is the gradient operator of function h (see [4] for more details). Moreover, the problem (6) can be solved by using the proximal gradient technique which was presented by Parikh and Boyd [5], i.e., if $\hat{x}$ is a solution of (6), then it is a fixed point of a forward-backward operator T defined by $T := \mathrm{prox}_{\lambda g}(I - \lambda\nabla h)$ for $\lambda > 0$. The operator $\mathrm{prox}_{\lambda g}$ is called the proximity operator with respect to λ and function g. We know that T is a nonexpansive mapping whenever $\lambda \in \left(0, \frac{2}{L}\right)$. It is easily seen that $\mathrm{prox}_{\lambda g}$ is an example of the resolvent of ∂g, that is, $\mathrm{prox}_{\lambda g} = J_{\lambda}^{\partial g} = (I + \lambda\partial g)^{-1}$, see Section 2 for more details.

We have seen from above fact that fixed point theory plays very important role in solving and developing of the image and signal processing problems which can be applied to solving many real-world problems in digital image processing such as medical image and astronomy as well as image processing for security sections. Fixed point theory focuses on two important problems. The first one is an existence problem of a solution of many kind of real-world problems while the other problem is a problem of how to approximate such solutions of the interested problems. For the past two decades, a lot of fixed point iteration processes were introduced and studied to solving many practical problems. It is well-known by Banach Contraction Principle that every contraction map from a complete metric space X into itself has a unique fixed point.

A mapping T from a metric space (X, d) into itself is called a contraction if there is a $k \in [0, 1)$ such that $d(Tx, Ty) \le kd(x, y)$ for all $x, y \in X$.

It is well-known that the Picard iteration process, defined by $x_1 \in X$ and

$$x_{n+1} = Tx_n, \quad \forall n \geq 1,$$

converges to a unique fixed point x^* of T.

It is observed that when $k = 1$ in above inequality, we have a new nonlinear mapping, called nonexpansive mapping. This type of mapping plays a crucial role to solving many optimization problems and economics.

From now on, we would like to provide some background concerning various iteration methods for finding a fixed point of nonexpansive and other nonlinear mappings.

Mann [6] was the first who introduced a modified iterative method known as Mann iteration process in Hilbert space H as follows: $x_1 \in H$,

$$x_{n+1} = \alpha_n x_n + (1 - \alpha_n)Tx_n, \quad \forall n \geq 1,$$

where $\{\alpha_n\}$ is a real sequence in $[0, 1]$. In 1974, Ishikawa extended Mann iteration, called the Ishikawa iteration process, by the following method: For an initial point x_1,

$$\begin{cases} y_n = (1 - \alpha_n)x_n + \alpha_n Tx_n, \\ x_{n+1} = (1 - \beta_n)x_n + \beta_n Ty_n, \quad \forall n \geq 1, \end{cases} \tag{7}$$

where $\{\alpha_n\}$, $\{\beta_n\} \subset [0, 1]$. Agarwal et al. employed the idea of the Ishikawa method to introduce S-iteration process as follows: For an initial point x_1,

$$\begin{cases} y_n = (1 - \alpha_n)x_n + \alpha_n Tx_n, \\ x_{n+1} = (1 - \beta_n)Tx_n + \beta_n Ty_n, \quad \forall n \geq 1, \end{cases} \tag{8}$$

where $\{\alpha_n\}$ and $\{\beta_n\}$ are sequences in $[0, 1]$. They showed that the convergence behavior of S-iteration is better than that of Mann and of Ishikawa iterations.

Because Mann iteration obtained only weak convergence (see [7] for more details). In 2000, Moufafi [8] introduced a well-known viscosity approximation was defined as follows: For $x_1 \in H$,

$$x_{n+1} = \alpha_n f(x_n) + (1 - \alpha_n)Tx_n, \quad \forall n \geq 1, \tag{9}$$

where $\{\alpha_n\} \subset [0, 1]$ and f is a contraction mapping. Under some suitable control conditions, he proved that $\{x_n\}$ converges strongly to a fixed point of T, when T is a nonexpansive mapping. Recently, authors in [9] proposed the viscosity-based inertial forward-backward algorithm (VIFBA), for solving (5) by finding a common fixed point of an infinite family $\{T_n\}$ of forward-backward operators. For initial points $x_0, x_1 \in H$, they define their method as follows:

$$\begin{cases} y_n = x_n + \theta_n(x_n - x_{n-1}), \\ z_n = (1 - \beta_n)y_n + \beta_n f(y_n), \\ x_{n+1} = T_n z_n, \quad \forall n \geq 1, \end{cases} \tag{10}$$

where f is a contraction mapping on H and $\{\alpha_n\}, \{\beta_n\}$ are sequences in $[0, 1]$. Here, the inertial term is represented by the term $\theta_n(x_n - x_{n-1})$ which was firstly introduced by Nesterov [10]. This algorithm is also applied to solve the regression and recognition problems.

In 2009, Beck and Teboulle introduced a fast iterative shrinkage-thresholding algorithm (FISTA) which was defined by

$$\begin{cases} y_n = Tx_n, \\ t_{n+1} = \dfrac{1+\sqrt{1+4t_n^2}}{2}, \\ \theta_n = \dfrac{t_n - 1}{t_{n+1}}, \\ x_{n+1} = y_n + \theta_n(y_n - y_{n-1}), \quad \forall n \geq 1, \end{cases} \tag{11}$$

where $T = \text{prox}_{\lambda g}(I - \lambda \nabla h)$ for $\lambda > 0$ and the initial points $x_1 = y_0 \in \mathbb{R}^n$ and $t_1 = 1$. Moreover, they applied their algorithm to the image restoration problems (see [3] for more details). It is pointed out from this work that the LASSO model is a suitable model for image restoration problems.

Motivated and inspired by all of these researches going on in this direction, in this paper, we introduce a new accelerated algorithm for finding a common fixed point of a family of nonexpansive mappings $\{T_n\}$ in Hilbert spaces based on the concept of inertial forward-backward, of Mann and of viscosity algorithms. Then a strong convergence theorem is established under some control conditions. Moreover, we apply the main results to solving image restoration problems and compare efficiency of our proposed algorithm with others. The presented results in this work also improve some well-known results in the literature.

This paper is organized as follows: In Section 2, Preliminaries, we recall some definitions and the useful facts which will be used in the later sections. We prove and analyze a strong convergence of the proposed algorithm in Section 3, Main Results. In the next section, Section 4 (Applications), we apply our main result to solving image restoration problems. Finally, the last section, Section 5 (Conclusions), is the summary of our work.

2. Preliminaries

Throughout this paper, we let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. Let $\{x_n\}$ be a sequence in H. We use $x_n \to x$ stands for $\{x_n\}$ converges strongly to x and $x_n \rightharpoonup x$ stands for $\{x_n\}$ converges weakly to x. Let $T : C \to C$ be a mapping from a nonempty closed convex subset of H into itself. A fixed point of T is a point $x \in C$ such that $x = Tx$. The set of all fixed points of T is denoted by $F(T)$, that is,

$$F(T) := \{x \in C : x = Tx\}.$$

A mapping $T : C \to H$ is said to be L-Lipschitzian, if there exists a constant $L > 0$ such that

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

If $L = 1$, then T is said to be a nonexpansive mapping. It is well-known that if T is nonexpansive, then $F(T)$ is closed and convex.

We call a mapping $f : C \to H$ a contraction, if there exists a constant $k \in [0, 1)$ such that

$$\|f(x) - f(y)\| \leq k\|x - y\|, \;\; \forall x, y \in C.$$

Here, we say that f is a k-contraction mapping.

Let $A : H \to 2^H$. The domain of A is the set $D(A) := \{x \in H : Ax \neq \emptyset\}$ and the range of A is the set $R(A) := \bigcup\{Az : z \in D(A)\}$. The inverse of A is denoted by A^{-1} is defined as follows: $x \in A^{-1}y$ if and only if $y \in Ax$. The graph of A is denoted by $G(A)$ and $G(A) := \{(x, u) : u \in Ax\}$.

An operator $A : H \to 2^H$ is said to be monotone if $\langle u - v, x - y \rangle \geq 0,$ for all $u \in Ax$ and $v \in Ay$. A monotone operator A on H is said to be maximal if the graph of A is not properly contained in any graph of other monotone operators on H. It is well-known that A is maximal if and only if for $x \in D(A)$ and $u \in Ax$, $\langle u - v, x - y \rangle \geq 0$ implies $(y, v) \in G(A)$.

Moreover, A is a maximally monotone operator if and only if $R(I + \lambda A) = H$ for every $\lambda > 0$, where I is an identity operator. We also know that the subdifferential of a proper lower semicontinuous convex function is a nice example of a maximal monotone.

For a function $g : H \to [-\infty, +\infty]$. The subdifferential $\partial g : H \to 2^H$ of g at $x \in H$, with $g(x) \in \mathbb{R}$, is the set $\partial g(x) := \{x^* \in H : g(x) + \langle y - x, x^* \rangle \leq g(y), \forall y \in H\}$. We take by convention $\partial g(x) := \emptyset$, if $g(x) \in \{\pm\infty\}$. If $g \in \Gamma_0(H)$, the set of proper lower semicontinuous convex functions from H to $(-\infty, +\infty]$, then ∂g is maximally monotone (see [11] for more details).

For a maximally monotone operator A and $\lambda > 0$, the resolvent of A for λ is defined to be a single-valued operator $J_\lambda^A : R(I + \lambda A) \to D(A)$, where $J_\lambda^A = (I + \lambda A)^{-1}$. It is well-known that J_λ^A is a nonexpansive mapping and $F(J_\lambda^A) = A^{-1}0$, where $A^{-1}0 := \{x \in H : 0 \in Ax\}$ and it is called the set of all zero (or null) points of A.

Let $A : H \to 2^H$ be a multi-valued mapping and $B : H \to H$ a single-valued nonlinear mapping. The quasi-variational inclusion problem is the problem of finding a point $x \in H$ such that

$$0 \in Ax + Bx. \tag{12}$$

The set of all solutions of the problem (12) is denoted by $(A + B)^{-1}0$.

A classical method for solving the problem (12) is the forward-backward method [12–14] which was first introduced by Combettes and Hirstoaga [15] in the following manner: $x_1 \in H$ and

$$x_{n+1} = J_\lambda^A(x_n - \lambda Bx_n), \ \ \forall n \geq 1, \tag{13}$$

where $\lambda > 0$. Moreover, we have from [16], if A is a maximally monotone operator and B is an L-Lipschitz continuous, then $F(J_\lambda^A(I - \lambda B)) = (A + B)^{-1}0$.

Definition 1. *Let $g \in \Gamma_0(H)$ and $\lambda > 0$. The proximity operator of parameter λ of g at $x \in H$ is denoted by $prox_{\lambda g}$ and it is defined by*

$$prox_{\lambda g}x := \arg\min_{y \in H} \left\{ g(y) + \frac{1}{2\lambda}\|y - x\|^2 \right\}.$$

It is well-known that if $g \in \Gamma_0(H)$, then $J_\lambda^{\partial g} = \text{prox}_{\lambda g}$, that is, the proximity operator is an example of resolvent operator. Moreover, if $g = \|\cdot\|_1$, then

$$\text{prox}_{\lambda\|\cdot\|_1}x = \text{sign}(x)\max\{\|x\|_1 - \lambda, 0\},$$

where sign is a signum function (see [4] for more details).

The following basic definitions and well-known results are also needed for proving our main results.

Lemma 1. *([17,18]) Let H be a real Hilbert space. For $x, y \in H$ and any arbitrary real number λ in $[0, 1]$, the following hold:*

1. $\|x \pm y\|^2 = \|x\|^2 \pm 2\langle x, y\rangle + \|y\|^2$;
2. $\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y\rangle$;
3. $\|\lambda x + (1 - \lambda)y\|^2 = \lambda\|x\|^2 + (1 - \lambda)\|y\|^2 - \lambda(1 - \lambda)\|x - y\|^2$.

The identity in Lemma 1(3) implies that the following equality holds:

$$\|\alpha x + \beta y + \gamma z\|^2 = \alpha\|x\|^2 + \beta\|y\|^2 + \gamma\|z\|^2 - \alpha\beta\|x - y\|^2 - \beta\gamma\|y - z\|^2 - \alpha\gamma\|x - z\|^2, \tag{14}$$

for all $x, y, z \in H$ and $\alpha, \beta, \gamma \in [0, 1]$ with $\alpha + \beta + \gamma = 1$.

Let C be a nonempty closed convex subset of a Hilbert space H. We know that for each element $x \in H$, there exists a unique point in C, say $P_C x$, such that

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

Such a mapping P_C is called the metric projection of H onto C. It is well-known that P_C is a nonexpansive mapping. Moreover, P_C can be characterized by the following inequality

$$\langle x - P_C x, y - P_C x \rangle \leq 0 \tag{15}$$

holds for all $x \in H$ and $y \in C$ (see [19] for more details).

We next recall the following properties which are useful for proving our main result, we refer to [20,21].

Let $\{T_n\}$ and $\mathcal{T}$ be families of nonexpansive mappings of H into itself such that $\emptyset \neq F(\mathcal{T}) \subset \bigcap_{n=1}^{\infty} F(T_n)$, where $F(\mathcal{T})$ is the set of all common fixed points of $\mathcal{T}$. We say that $\{T_n\}$ satisfies NST-condition(I) with $\mathcal{T}$ if for each bounded sequence $\{x_n\}$ such that $\lim_{n\to\infty} \|x_n - T_n x_n\| = 0$, it follows

$$\lim_{n\to\infty} \|x_n - T x_n\| = 0 \quad \text{for all } T \in \mathcal{T}.$$

In particular, if $\mathcal{T}$ consists of one mapping T, i.e., $\mathcal{T} = \{T\}$, then $\{T_n\}$ is said to satisfy NST-condition(I) with T.

Lemma 2. *Let $\{T_n\}$ be a family of nonexpansive mappings of H into itself and $T : H \to H$ a nonexpansive mapping with $\emptyset \neq F(T) \subset \bigcap_{n=1}^{\infty} F(T_n)$. One always has, if $\{T_n\}$ satisfies NST-condition(I) with T, then $\{T_t\}$ also satisfies NST-condition(I) with T, for any subsequences $\{t\}$ of positive integers.*

Proof. Let $\{x_t\}$ be a bounded sequence such that $\|x_t - T_t x_t\| \to 0$ as $t \to +\infty$. Take $u \in F(T)$. Define the sequence $\{x_n\}$ by

$$x_n := \begin{cases} x_t & \text{if } n = t; \\ u & \text{otherwise.} \end{cases}$$

Then $\{x_n\}$ is bounded. Moreover, we have that

$$\lim_{n\to\infty} \|x_n - T_n x_n\| = \lim_{t\to\infty} \|x_t - T_t x_t\| = 0,$$

due to u is a fixed point of T_n for all $n \in \mathbb{N}$. By the NST-condition(I) with T on $\{T_n\}$, we obtain that

$$\lim_{t\to\infty} \|x_t - T x_t\| = \lim_{n\to\infty} \|x_n - T x_n\| = 0.$$

Thus, $\{T_t\}$ satisfies NST-condition(I) with T. □

Proposition 1. *([22]) Let H be a Hilbert space. Let $A : H \to 2^H$ be a maximally monotone operator and $B : H \to H$ an L-Lipschitz operator, where $L > 0$. Let $T_n = J_{\lambda_n}^A (I - \lambda_n B)$, where $0 < \lambda_n < \frac{2}{L}$ for all $n \geq 1$ and let $T = J_\lambda^A (I - \lambda B)$, where $0 < \lambda < \frac{2}{L}$ with $\lambda_n \to \lambda$. Then $\{T_n\}$ satisfies the NST-condition(I) with T.*

The following lemmas are crucial for proving our main results.

Lemma 3. *([23]) Let H be a real Hilbert space and $T : H \to H$ a nonexpansive mapping with $F(T) \neq \emptyset$. Then the mapping $I - T$ is demiclosed at zero, i.e., for any sequences $\{x_n\}$ in H such that $x_n \rightharpoonup x \in H$ and $\|x_n - T x_n\| \to 0$ imply $x \in F(T)$.*

Lemma 4. *([24,25]) Let $\{s_n\}$, $\{\xi_n\}$ be sequences of nonnegative real numbers, $\{\delta_n\}$ a sequence in [0,1] and $\{t_n\}$ a sequence of real numbers such that*

$$s_{n+1} \le (1-\delta_n)s_n + \delta_n t_n + \xi_n,$$

for all $n \in \mathbb{N}$. If the following conditions hold:

1. $\sum_{n=1}^{\infty} \delta_n = \infty$;
2. $\sum_{n=1}^{\infty} \xi_n < \infty$;
3. $\limsup_{n\to\infty} t_n \le 0$.

   Then $\lim_{n\to\infty} s_n = 0$.

Lemma 5. *([26]) Let $\{\Theta_n\}$ be a sequence of real numbers that does not decrease at infinity in the sense that there exists a subsequence $\{\Theta_{n_i}\}$ of $\{\Theta_n\}$ which satisfies $\Theta_{n_i} < \Theta_{n_i+1}$ for all $i \in \mathbb{N}$. Define the sequence $\{\tau(n)\}_{n\ge n_0}$ of integers as follows:*

$$\tau(n) := \max\{k \le n : \Theta_k < \Theta_{k+1}\},$$

where $n_0 \in \mathbb{N}$ such that $\{k \le n_0 : \Theta_k < \Theta_{k+1}\} \neq \emptyset$. Then the following hold:

1. $\tau(n_0) \le \tau(n_0+1) \le ...$ *and* $\tau(n) \to \infty$;
2. $\Theta_{\tau(n)} \le \Theta_{\tau(n)+1}$ *and* $\Theta_n \le \Theta_{\tau(n)+1}$ *for all* $n \ge n_0$.

3. Main Results

In this section, we first give a new algorithm for finding a common fixed point of a family of nonexpansive mappings in a real Hilbert space. We then prove its strong convergence under some suitable conditions.

We now propose a new accelerated algorithm for approximating a solution of our common fixed point problem as the following.

Let H be a real Hilbert space. Let $\{T_n\}$ be a family of nonexpansive mappings on H into itself. Let f be a k-contraction mapping on H with $k \in (0,1)$ and let $\{\eta_n\} \subset (0,\infty)$ and $\{\sigma_n\}, \{\alpha_n\}, \{\beta_n\}, \{\gamma_n\} \subset (0,1)$.

We next prove the convergence of the sequence generated by Algorithm 1. To this end, we assume that the algorithm does not stop after finitely many iterations.

Algorithm 1: NAVA (New Accelerated Viscosity Algorithm).

Initialization: Take $x_0, x_1 \in H$. Choose $\theta \ge 0$.

For $n \ge 1$**:**

Set

$$\theta_n := \begin{cases} \min\left\{\theta, \dfrac{\eta_n \alpha_n}{\|x_n - x_{n-1}\|}\right\} & \text{if } x_n \neq x_{n-1}; \\ \theta & \text{otherwise.} \end{cases}$$

Compute

$$\begin{aligned} y_n &:= x_n + \theta_n(x_n - x_{n-1}), \\ z_n &:= (1-\sigma_n)y_n + \sigma_n T_n y_n, \\ x_{n+1} &:= \alpha_n f(x_n) + \beta_n T_n y_n + \gamma_n T_n z_n. \end{aligned}$$

Theorem 1. *Let $\{T_n\}$ be a family of nonexpansive mappings and $T : H \to H$ a nonexpansive mapping such that $\emptyset \neq F(T) \subset \bigcap_{n=1}^{\infty} F(T_n)$. Suppose that $\{T_n\}$ satisfies NST-condition(I) with T. Let $\{x_n\}$ be the sequence generated by Algorithm 1 such that the following additional conditions hold:*

1. $\alpha_n + \beta_n + \gamma_n = 1$;
2. $0 < a \leq \sigma_n \leq a' < 1$;
3. $0 < b \leq \beta_n \leq b' < 1$;
4. $0 < c \leq \gamma_n \leq c' < 1$;
5. $\lim_{n\to\infty} \eta_n = 0$;
6. $\lim_{n\to\infty} \alpha_n = 0$ *and* $\sum_{n=1}^{\infty} \alpha_n = \infty$,

for some positive real numbers a, b, c, a', b', c'. Then the sequence $\{x_n\}$ converges strongly to $u \in F(T)$, where $u = P_{F(T)} f(u)$.

Proof. Let $u \in F(T)$ be such that $u = P_{F(T)} f(u)$. First of all, we show that $\{x_n\}$ is bounded. By the definition of y_n and of z_n, we have

$$\begin{aligned} \|y_n - u\| &= \|x_n + \theta_n(x_n - x_{n-1}) - u\| \\ &\leq \|x_n - u\| + \theta_n \|x_n - x_{n-1}\|, \quad \forall n \geq 1, \end{aligned} \tag{16}$$

and

$$\begin{aligned} \|z_n - u\| &= \|(1-\sigma_n)y_n + \sigma_n T_n y_n - u\| \\ &\leq (1-\sigma_n)\|y_n - u\| + \sigma_n \|T_n y_n - u\| \\ &= (1-\sigma_n)\|y_n - u\| + \sigma_n \|T_n y_n - T_n u\| \\ &\leq (1-\sigma_n)\|y_n - u\| + \sigma_n \|y_n - u\| \\ &= \|y_n - u\|, \quad \forall n \geq 1. \end{aligned} \tag{17}$$

From (16) and (17), we also have that

$$\begin{aligned} \|x_{n+1} - u\| &= \|\alpha_n f(x_n) + \beta_n T_n y_n + \gamma_n T_n z_n - u\| \\ &= \|\alpha_n(f(x_n) - u) + \beta_n(T_n y_n - u) + \gamma_n(T_n z_n - u)\| \\ &\leq \alpha_n \|f(x_n) - u\| + \beta_n \|T_n y_n - u\| + \gamma_n \|T_n z_n - u\| \\ &= \alpha_n \|f(x_n) - u\| + \beta_n \|T_n y_n - T_n u\| + \gamma_n \|T_n z_n - T_n u\| \\ &\leq \alpha_n \|f(x_n) - f(u)\| + \alpha_n \|f(u) - u\| + \beta_n \|y_n - u\| + \gamma_n \|z_n - u\| \\ &\leq \alpha_n k \|x_n - u\| + \alpha_n \|f(u) - u\| + (\beta_n + \gamma_n)\|y_n - u\| \\ &\leq \alpha_n k \|x_n - u\| + \alpha_n \|f(u) - u\| + (\beta_n + \gamma_n)(\|x_n - u\| + \theta_n \|x_n - x_{n-1}\|) \\ &\leq (\alpha_n k + \beta_n + \gamma_n)\|x_n - u\| + \alpha_n \|f(u) - u\| + (\beta_n + \gamma_n)\theta_n \|x_n - x_{n-1}\| \\ &\leq (\alpha_n k + \beta_n + \gamma_n)\|x_n - u\| + \alpha_n \|f(u) - u\| + (b' + c')\theta_n \|x_n - x_{n-1}\| \\ &= (1 - \alpha_n(1-k))\|x_n - u\| + \alpha_n \|f(u) - u\| \\ &\quad + (b' + c')\alpha_n \cdot \frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\|, \quad \forall n \geq 1. \end{aligned} \tag{18}$$

According to the definition of θ_n and the assumption (5), we have

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

Then there exists a positive constant $M_1 > 0$ such that

$$\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| \leq M_1, \quad \forall n \geq 1.$$

From (18), we obtain

$$\begin{aligned}
\|x_{n+1}-u\| &\leq (1-\alpha_n(1-k))\|x_n-u\| + \alpha_n\|f(u)-u\| + \alpha_n(b'+c')M_1 \\
&= (1-\alpha_n(1-k))\|x_n-u\| + \alpha_n(\|f(u)-u\| + (b'+c')M_1) \\
&= (1-\alpha_n(1-k))\|x_n-u\| + \alpha_n(1-k)\left[\frac{\|f(u)-u\|+(b'+c')M_1}{1-k}\right] \\
&\leq \max\left\{\|x_n-u\|, \frac{\|f(u)-u\|+(b'+c')M_1}{1-k}\right\} \\
&\vdots \\
&\leq \max\left\{\|x_1-u\|, \frac{\|f(u)-u\|+(b'+c')M_1}{1-k}\right\}, \quad \forall n \geq 1.
\end{aligned}$$

This implies $\{x_n\}$ is bounded and so are $\{y_n\}, \{z_n\}, \{f(x_n)\}$ and $\{T_n y_n\}$.
Indeed, we have that for all $n \geq 1$,

$$\begin{aligned}
\|y_n-u\|^2 &= \|x_n+\theta_n(x_n-x_{n-1})-u\|^2 \\
&= \|(x_n-u)+\theta_n(x_n-x_{n-1})\|^2 \\
&\leq \|x_n-u\|^2 + 2\theta_n\|x_n-u\|\|x_n-x_{n-1}\| + \theta_n^2\|x_n-x_{n-1}\|^2.
\end{aligned} \tag{19}$$

By Lemma 1(2), (14) and (17) we have

$$\begin{aligned}
\|x_{n+1}-u\|^2 &= \|\alpha_n f(x_n) + \beta_n T_n y_n + \gamma_n T_n z_n - u\|^2 \\
&= \|\alpha_n(f(x_n)-f(u)) + \beta_n(T_n y_n - u) + \gamma_n(T_n z_n - u) + \alpha_n(f(u)-u)\|^2 \\
&\leq \|\alpha_n(f(x_n)-f(u)) + \beta_n(T_n y_n - u) + \gamma_n(T_n z_n - u)\|^2 + 2\alpha_n\langle f(u)-u, x_{n+1}-u\rangle \\
&\leq \alpha_n\|f(x_n)-f(u)\|^2 + \beta_n\|T_n y_n - u\|^2 + \gamma_n\|T_n z_n - u\|^2 + 2\alpha_n\langle f(u)-u, x_{n+1}-u\rangle \\
&= \alpha_n\|f(x_n)-f(u)\|^2 + \beta_n\|T_n y_n - T_n u\|^2 + \gamma_n\|T_n z_n - T_n u\|^2 + 2\alpha_n\langle f(u)-u, x_{n+1}-u\rangle \\
&\leq \alpha_n k^2\|x_n-u\|^2 + \beta_n\|y_n-u\|^2 + \gamma_n\|z_n-u\|^2 + 2\alpha_n\langle f(u)-u, x_{n+1}-u\rangle \\
&\leq \alpha_n k^2\|x_n-u\|^2 + (\beta_n+\gamma_n)\|y_n-u\|^2 + 2\alpha_n\langle f(u)-u, x_{n+1}-u\rangle, \quad \forall n \geq 1
\end{aligned} \tag{20}$$

It follows from (19) with $0 < k < 1$ that

$$\begin{aligned}
\|x_{n+1}-u\|^2 &\leq \alpha_n k\|x_n-u\|^2 \\
&\quad + (\beta_n+\gamma_n)\big(\|x_n-u\|^2 + 2\theta_n\|x_n-u\|\|x_n-x_{n-1}\| + \theta_n^2\|x_n-x_{n-1}\|^2\big) \\
&\quad + 2\alpha_n\langle f(u)-u, x_{n+1}-u\rangle \\
&= (1-\alpha_n(1-k))\|x_n-u\|^2 \\
&\quad + (\beta_n+\gamma_n)\theta_n\|x_n-x_{n-1}\|\big(2\|x_n-u\| + \theta_n\|x_n-x_{n-1}\|\big) \\
&\quad + 2\alpha_n\langle f(u)-u, x_{n+1}-u\rangle, \quad \forall n \geq 1.
\end{aligned} \tag{21}$$

Since

$$\theta_n\|x_n-x_{n-1}\| = \alpha_n \cdot \frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\| \to 0, \text{ as } n \to \infty,$$

there exists a positive constant $M_2 > 0$ such that

$$\theta_n\|x_n-x_{n-1}\| \leq M_2, \quad \forall n \geq 1.$$

From the inequality (21), we derive that for all $n \geq 1$,

$$\begin{aligned} \|x_{n+1}-u\|^2 &\leq (1-\alpha_n(1-k))\|x_n-u\|^2 + 3M_3(\beta_n+\gamma_n)\theta_n\|x_n-x_{n-1}\| + 2\alpha_n\langle f(u)-u, x_{n+1}-u\rangle \\ &\leq (1-\alpha_n(1-k))\|x_n-u\|^2 + 3M_3(b'+c')\theta_n\|x_n-x_{n-1}\| + 2\alpha_n\langle f(u)-u, x_{n+1}-u\rangle \\ &\leq (1-\alpha_n(1-k))\|x_n-u\|^2 \\ &\quad + \alpha_n(1-k)\left[\frac{3M_3(b'+c')}{1-k}\cdot\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\| + \frac{2}{1-k}\langle f(u)-u, x_{n+1}-u\rangle\right], \end{aligned} \tag{22}$$

where $M_3 := \sup_{n\geq 1}\{\|x_n-u\|, M_2\}$. From above inequality, we set

$$s_n := \|x_n-u\|^2, \quad \delta_n := \alpha_n(1-k)$$

and

$$t_n := \frac{3M_3(b'+c')}{1-k}\cdot\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\| + \frac{2}{1-k}\langle f(u)-u, x_{n+1}-u\rangle, \quad \forall n \geq 1.$$

So, we obtain

$$s_{n+1} \leq (1-\delta_n)s_n + \delta_n t_n, \tag{23}$$

for all $n \geq 1$.

Now, we consider two cases for the proof as follows:

Case 1. Suppose that there exists a natural number n_0 such that the sequence $\{\|x_n-u\|\}_{n\geq n_0}$ is nonincreasing. Hence, $\{\|x_n-u\|\}$ converges due to it is bounded from below by 0. Using the assumption (6), we get that $\sum_{n=1}^{\infty}\delta_n = \infty$. From Lemma 4, we next claim that

$$\limsup_{n\to\infty}\langle f(u)-u, x_{n+1}-u\rangle \leq 0.$$

Coming back to the definition of z_n, by Lemma 1(3), one has that

$$\begin{aligned} \|z_n-u\|^2 &= \|(1-\sigma_n)y_n + \sigma_n T_n y_n - u\|^2 \\ &= \|(1-\sigma_n)(y_n-u) + \sigma_n(T_n y_n - u)\|^2 \\ &= (1-\sigma_n)\|y_n-u\|^2 + \sigma_n\|T_n y_n - u\|^2 - \sigma_n(1-\sigma_n)\|y_n - T_n y_n\|^2 \\ &= (1-\sigma_n)\|y_n-u\|^2 + \sigma_n\|T_n y_n - T_n u\|^2 - \sigma_n(1-\sigma_n)\|y_n - T_n y_n\|^2 \\ &\leq \|y_n-u\|^2 - \sigma_n(1-\sigma_n)\|y_n - T_n y_n\|^2, \quad \forall n \geq 1. \end{aligned} \tag{24}$$

Using the facts that (14), (17), (19) and (24) yield

$$\begin{aligned} \|x_{n+1}-u\|^2 &= \|\alpha_n f(x_n) + \beta_n T_n y_n + \gamma_n T_n z_n - u\|^2 \\ &= \|\alpha_n(f(x_n)-u) + \beta_n(T_n y_n - u) + \gamma_n(T_n z_n - u)\|^2 \\ &\leq \alpha_n\|f(x_n)-u\|^2 + \beta_n\|T_n y_n - u\|^2 + \gamma_n\|T_n z_n - u\|^2 \\ &= \alpha_n\|f(x_n)-u\|^2 + \beta_n\|T_n y_n - T_n u\|^2 + \gamma_n\|T_n z_n - T_n u\|^2 \\ &\leq \alpha_n\|f(x_n)-u\|^2 + \beta_n\|y_n-u\|^2 + \gamma_n\|z_n-u\|^2 \\ &\leq \alpha_n\|f(x_n)-u\|^2 + \beta_n\|y_n-u\|^2 + \gamma_n\left(\|y_n-u\|^2 - \sigma_n(1-\sigma_n)\|y_n-T_n y_n\|^2\right) \\ &= \alpha_n\|f(x_n)-u\|^2 + (\beta_n+\gamma_n)\|y_n-u\|^2 - \gamma_n\sigma_n(1-\sigma_n)\|y_n-T_n y_n\|^2 \\ &\leq \alpha_n\|f(x_n)-u\|^2 + (\beta_n+\gamma_n)\|x_n-u\|^2 + 2(\beta_n+\gamma_n)\theta_n\|x_n-u\|\|x_n-x_{n-1}\| \\ &\quad + (\beta_n+\gamma_n)\theta_n^2\|x_n-x_{n-1}\|^2 - \gamma_n\sigma_n(1-\sigma_n)\|y_n-T_n y_n\|^2 \\ &= \alpha_n\|f(x_n)-u\|^2 + (1-\alpha_n)\|x_n-u\|^2 + 2(\beta_n+\gamma_n)\theta_n\|x_n-u\|\|x_n-x_{n-1}\| \\ &\quad + (\beta_n+\gamma_n)\theta_n^2\|x_n-x_{n-1}\|^2 - \gamma_n\sigma_n(1-\sigma_n)\|y_n-T_n y_n\|^2, \quad \forall n \geq 1. \end{aligned} \tag{25}$$

It implies that for all $n \geq 1$,

$$\gamma_n\sigma_n(1-\sigma_n)\|y_n - T_ny_n\|^2 \leq \alpha_n(\|f(x_n)-u\|^2 - \|x_n-u\|^2) + \|x_n-u\|^2 - \|x_{n+1}-u\|^2 \\ + (\beta_n+\gamma_n)\theta_n\|x_n-x_{n-1}\|(2\|x_n-u\| + \theta_n\|x_n-x_{n-1}\|). \tag{26}$$

It follows from the assumptions (2), (3), (4), (6) and the convergence of the sequences $\{\|x_n - u\|\}$ and of $\{\theta_n\|x_n - x_{n-1}\|\}$ that

$$\|y_n - T_ny_n\| \to 0 \text{ as } n \to \infty. \tag{27}$$

According to $\{T_n\}$ satisfies NST-condition(I) with T, we obtain that

$$\|y_n - Ty_n\| \to 0 \text{ as } n \to \infty. \tag{28}$$

From the definition of y_n and of z_n, we obtain

$$\|y_n - x_n\| = \theta_n\|x_n - x_{n-1}\| \to 0 \text{ as } n \to \infty, \tag{29}$$

and

$$\|z_n - y_n\| \leq \sigma_n\|y_n - T_ny_n\| \to 0 \text{ as } n \to \infty. \tag{30}$$

Using (27) and (30) with the assumption (6), we have

$$\begin{aligned} \|x_{n+1}-y_n\| &\leq \|x_{n+1}-T_ny_n\| + \|T_ny_n - y_n\| \\ &\leq \|\alpha_nf(x_n) + \beta_nT_ny_n + \gamma_nT_nz_n - T_ny_n\| + \|T_ny_n-y_n\| \\ &\leq \|\alpha_n(f(x_n)-T_ny_n) + \beta_n(T_ny_n - T_ny_n) + \gamma_n(T_nz_n - T_ny_n)\| + \|T_ny_n - y_n\| \\ &\leq \alpha_n\|f(x_n)-T_ny_n\| + \beta_n\|T_ny_n - T_ny_n\| + \gamma_n\|T_nz_n - T_ny_n\| + \|T_ny_n-y_n\| \\ &\leq \alpha_n\|f(x_n)-T_ny_n\| + \gamma_n\|z_n - y_n\| + \|T_ny_n - y_n\|, \quad \forall n \geq 1, \end{aligned} \tag{31}$$

which implies

$$\|x_{n+1} - y_n\| \to 0 \text{ as } n \to \infty. \tag{32}$$

Hence

$$\|x_{n+1} - x_n\| \leq \|x_{n+1} - y_n\| + \|y_n - x_n\| \to 0 \text{ as } n \to \infty. \tag{33}$$

Let

$$v = \limsup_{n\to\infty}\langle f(u) - u, x_{n+1} - u\rangle.$$

So, there exists a subsequence $\{x_t\}$ of $\{x_n\}$ such that

$$v = \lim_{t\to\infty}\langle f(u) - u, x_{t+1} - u\rangle.$$

Since $\{x_t\}$ is bounded, there exists a subsequence $\{x_{t'}\}$ of $\{x_t\}$ such that $x_{t'} \rightharpoonup w$ for some $w \in H$. Without loss of generality, we may assume that $x_t \rightharpoonup w$ and

$$v = \lim_{t\to\infty}\langle f(u) - u, x_{t+1} - u\rangle.$$

From (28) and (29), we derive

$$\begin{aligned} \|x_n - Tx_n\| &\leq \|x_n - y_n\| + \|y_n - Ty_n\| + \|Ty_n - Tx_n\| \\ &\leq 2\|x_n - y_n\| + \|y_n - Ty_n\|, \quad \forall n \geq 1, \end{aligned} \tag{34}$$

and hence,

$$\|x_n - Tx_n\| \to 0 \text{ as } n \to \infty. \tag{35}$$

It implies by Lemma 3 that $w \in F(T)$. Since $\|x_{n+1} - x_n\| \to 0$, we get $x_{t+1} \rightharpoonup w$. Moreover, using $u = P_{F(T)}f(u)$ and (15), we obtain

$$v = \lim_{t\to\infty} \langle f(u) - u, x_{t+1} - u\rangle = \langle f(u) - u, w - u\rangle \leq 0. \tag{36}$$

Then

$$\limsup_{n\to\infty} \langle f(u) - u, x_{n+1} - u\rangle \leq 0. \tag{37}$$

It implies from (37) with the fact of $\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| \to 0$ that $\limsup_{n\to\infty} t_n \leq 0$. Coming back to (23), using Lemma 4, we conclude that $x_n \to u$.

Case 2. Suppose that the sequence $\{\|x_n - u\|\}_{n\geq n_0}$ is not a monotonically decreasing sequence for some n_0 large enough. Set

$$\Theta_n := \|x_n - u\|^2.$$

So, there exists a subsequence $\{\Theta_{n_j}\}$ of $\{\Theta_n\}$ such that $\Theta_{n_j} \leq \Theta_{n_j+1}$ for all $j \in \mathbb{N}$. In this case, we define $\tau : \{n : n \geq n_0\} \to \mathbb{N}$ by

$$\tau(n) := \max\{k \in \mathbb{N} : k \leq n, \Theta_k < \Theta_{k+1}\}.$$

By Lemma 5, we have that $\Theta_{\tau(n)} \leq \Theta_{\tau(n)+1}$ for all $n \geq n_0$. That is,

$$\|x_{\tau(n)} - u\| \leq \|x_{\tau(n)+1} - u\|, \quad \forall n \geq n_0.$$

As in Case 1, we can conclude that for all $n \geq n_0$,

$$\begin{aligned}
\gamma_{\tau(n)}\sigma_{\tau(n)}(1-\sigma_{\tau(n)})\|y_{\tau(n)} - T_{\tau(n)}y_{\tau(n)}\|^2 &\leq \alpha_{\tau(n)}\big(\|f(x_{\tau(n)}) - u\|^2 - \|x_{\tau(n)} - u\|^2\big) \\
&\quad + \|x_{\tau(n)} - u\|^2 - \|x_{\tau(n)+1} - u\|^2 \\
&\quad + (\beta_{\tau(n)} + \gamma_{\tau(n)})\theta_{\tau(n)}\|x_{\tau(n)} - x_{\tau(n)-1}\| \\
&\quad \times \big(2\|x_{\tau(n)} - u\| + \theta_{\tau(n)}\|x_{\tau(n)} - x_{\tau(n)-1}\|\big) \\
&\leq \alpha_{\tau(n)}\big(\|f(x_{\tau(n)}) - u\|^2 - \|x_{\tau(n)} - u\|^2\big) \\
&\quad + (\beta_{\tau(n)} + \gamma_{\tau(n)})\theta_{\tau(n)}\|x_{\tau(n)} - x_{\tau(n)-1}\| \\
&\quad \times \big(2\|x_{\tau(n)} - u\| + \theta_{\tau(n)}\|x_{\tau(n)} - x_{\tau(n)-1}\|\big)
\end{aligned}$$

and hence,

$$\|y_{\tau(n)} - T_{\tau(n)}y_{\tau(n)}\| \to 0 \text{ as } n \to \infty. \tag{38}$$

Similarly to the proof of Case 1, we get

$$\|y_{\tau(n)} - x_{\tau(n)}\| \to 0, \tag{39}$$

$$\|z_{\tau(n)} - y_{\tau(n)}\| \to 0 \tag{40}$$

and

$$\|x_{\tau(n)+1} - y_{\tau(n)}\| \to 0, \tag{41}$$

as $n \to \infty$, and hence

$$\|x_{\tau(n)+1} - x_{\tau(n)}\| \to 0 \text{ as } n \to \infty. \tag{42}$$

We next show that $\limsup_{n\to\infty}\langle f(u)-u, x_{\tau(n)+1}-u\rangle \le 0$. Put

$$v=\limsup_{n\to\infty}\langle f(u)-u, x_{\tau(n)+1}-u\rangle.$$

Without loss of generality, there exists a subsequence $\{x_{\tau(t)}\}$ of $\{x_{\tau(n)}\}$ such that $\{x_{\tau(t)}\}$ converges weakly to some point $w\in H$ and

$$v=\lim_{t\to\infty}\langle f(u)-u, x_{\tau(t)+1}-u\rangle.$$

By Lemma 2, one has $\{T_{\tau(t)}\}$ satisfies NST-condition(I) with T. So, according to the equality (38), $\|y_{\tau(t)}-T_{\tau(t)}y_{\tau(t)}\|\to 0$, we obtain that

$$\|y_{\tau(t)}-Ty_{\tau(t)}\|\to 0 \;\text{ as }\; t\to\infty, \tag{43}$$

which implies, by (39) and Lemma 3 again, that $w\in F(T)$. Since $\|x_{\tau(t)+1}-x_{\tau(t)}\|\to 0$, we get $x_{\tau(t)+1}\rightharpoonup w$. Moreover, using $u=P_{F(T)}f(u)$ and (15), we obtain

$$v=\lim_{t\to\infty}\langle f(u)-u, x_{\tau(t)+1}-u\rangle=\langle f(u)-u, w-u\rangle\le 0. \tag{44}$$

Then

$$\limsup_{n\to\infty}\langle f(u)-u, x_{\tau(n)+1}-u\rangle\le 0. \tag{45}$$

Since $\Theta_{\tau(n)}\le\Theta_{\tau(n)+1}$ and $\alpha_{\tau(n)}(1-k)>0$, as in the proof of Case 1, we have that for all $n\ge n_0$,

$$\|x_{\tau(n)}-u_0\|^2\le\frac{3M_3(b'+c')}{1-k}\cdot\frac{\theta_{\tau(n)}}{\alpha_{\tau(n)}}\|x_{\tau(n)}-x_{\tau(n)-1}\|+\frac{2}{1-k}\langle f(u)-u, x_{\tau(n)+1}-u\rangle. \tag{46}$$

It follows by the fact that $\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\|\to 0$ and (45) that

$$\limsup_{n\to\infty}\|x_{\tau(n)}-u_0\|^2\le 0,$$

and hence $\|x_{\tau(n)}-u_0\|\to 0$ as $n\to\infty$. It implies by (42) that $\|x_{\tau(n)+1}-u_0\|\to 0$ as $n\to\infty$.
By Lemma 5, we get

$$\|x_n-u_0\|\le\|x_{\tau(n)+1}-u_0\|\to 0 \text{ as } n\to\infty.$$

Hence $x_n\to u_0$. The proof is completed. □

As a direct consequence of Theorem 1, by using Proposition 1, we obtain the following corollary.

Corollary 1. *Let H be a real Hilbert space. Let $A: H\to 2^H$ be a maximally monotone operator and $B: H\to H$ an L-Lipschitz operator, where $L>0$. Let $T_n=J^A_{\lambda_n}(I-\lambda_n B)$, where $0<\lambda_n<\frac{2}{L}$ for all $n\ge 1$ and let $T=J^A_\lambda(I-\lambda B)$, where $0<\lambda<\frac{2}{L}$ with $\lambda_n\to\lambda$. Suppose that $(A+B)^{-1}0\neq\emptyset$. Let f be a k-contraction mapping on H with $k\in(0,1)$. Let $\{x_n\}$ be a sequence in H generated by Algorithm 1 under the same conditions of parameters as in Theorem 1. Then $\{x_n\}$ converges strongly to $u\in(A+B)^{-1}0$, where $u=P_{(A+B)^{-1}0}f(u)$.*

Proof. Since $(A+B)^{-1}0=F(T)\subset\bigcap_{n=1}^\infty F(T_n)$ and T, T_n are nonexpansive for each $n\in\mathbb{N}$, we can conclude that the sequence $\{x_n\}$ converges strongly to $u\in(A+B)^{-1}0$ by using Proposition 1 and Theorem 1. □

4. Applications

In this section, we first begin with presenting the algorithm obtained from our main results. We investigate throughout this section under the following setting.

- H is a real Hilbert space;
- $h : H \to \mathbb{R}$ is a differentiable and convex function with an L-Lipschitz continuous gradient ∇h where $L > 0$;
- $g \in \Gamma_0(H)$;
- $\arg\min(h+g) \neq \emptyset$;
- f is a k-contraction mapping on H with $k \in (0,1)$;
- $\lambda \in \left(0, \frac{2}{L}\right)$ and $\{\lambda_n\} \subset \left(0, \frac{2}{L}\right)$ with $\lambda_n \to \lambda$;
- $\{\eta_n\} \subset (0,\infty)$ and $\{\sigma_n\}, \{\alpha_n\}, \{\beta_n\}, \{\gamma_n\} \subset (0,1)$.

The algorithm we propose in this context has the following formulation.

We next prove the strong convergence of the sequence generated by our proposed algorithm.

Theorem 2. *Let $\{x_n\}$ be a sequence generated by Algorithm 2 under the same conditions of parameters as in Theorem 1. Then $\{x_n\}$ converges strongly to $u \in \arg\min(h+g)$.*

Algorithm 2: AVFBA (Accelerated Viscosity Forward-Backward Algorithm).

Initialization: Take $x_0, x_1 \in H$. Choose $\theta \geq 0$.

For $n \geq 1$**:**

Set

$$\theta_n := \begin{cases} \min\left\{\theta, \dfrac{\eta_n \alpha_n}{\|x_n - x_{n-1}\|}\right\} & \text{if } x_n \neq x_{n-1}; \\ \theta & \text{otherwise.} \end{cases}$$

Compute

$$\begin{aligned} y_n &:= x_n + \theta_n(x_n - x_{n-1}), \\ z_n &:= (1-\sigma_n)y_n + \sigma_n \mathrm{prox}_{\lambda_n g}(I - \lambda_n \nabla h)y_n, \\ x_{n+1} &:= \alpha_n f(x_n) + \beta_n \mathrm{prox}_{\lambda_n g}(I - \lambda_n \nabla h)y_n + \gamma_n \mathrm{prox}_{\lambda_n g}(I - \lambda_n \nabla h)z_n. \end{aligned}$$

Proof. In Corollary 1, we set $A := \partial g$ and $B := \nabla h$. So, A is a maximal operator. Then we obtain the required result directly by Corollary 1. □

We next discuss some experiment results by using our proposal algorithm to solving the image restoration problem. The image restoration problem (2) can be related to

$$\min_x \{\|Ax - b\|_2^2 + \lambda \|x\|_1\},$$

where $x \in \mathbb{R}^n$ is the original image, b is the observed image and A represents the blurring operator. In this situation, we choose the regularization parameter $\lambda = 5e^{-5}$. For this example, we look at the 256×256 Schonbrunn palace (original image). We use a Gausssian blur of size 9×9 and standard deviation $\sigma = 4$ to create the blurred and noisy image (observed image). These two images are given as in Figure 1.

(**a**) Original image (**b**) Observed image

Figure 1. The Schonbrunn palace.

In 2009, Thung and Raveendran [27] introduced Peak Signal-to-Noise Ratio (PSNR) to measure a quality of restored images for each x_n as the following:

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

where $MSE = \frac{1}{256^2} \|x_n - x\|^2$, the Mean Square Error for the original image x. We note that a higher PSNR shows a higher quality for deblurring image.

In Theorem 2, we set $h(x) = \|Ax - b\|_2^2$ and $g(x) = \lambda \|x\|_1$ and choose the Lipschitz constant L of the gradient ∇h which is the maximum value of eigenvalues of the matrix $A^T A$.

Let us begin with the first experiment. We study convergence behavior of our method by considering the following six different cases:

Parameters	**Case (a)**	**Case (b)**	**Case (c)**	**Case (d)**	**Case (e)**	**Case (f)**
α_n	$\frac{1}{33n}$	$\frac{1}{33n}$	$\frac{1}{33n}$	$\frac{1}{33n}$	$\frac{1}{33n}$	$\frac{1}{33n}$
β_n	$\frac{n}{n+1}$	$\frac{n}{300n+1}$	$\frac{n}{300n+1}$	$\frac{n}{300n+1}$	$\frac{n}{300n+1}$	$\frac{n}{300n+1}$
γ_n	$1 - \alpha_n - \beta_n$	$1 - \alpha_n - \beta_n$	$1 - \alpha_n - \beta_n$	$1 - \alpha_n - \beta_n$	$1 - \alpha_n - \beta_n$	$1 - \alpha_n - \beta_n$
σ_n	$\frac{n}{10(n+1)}$	$\frac{n}{10(n+1)}$	$\frac{99n}{100(n+1)}$	$\frac{99n}{100(n+1)}$	$\frac{99n}{100(n+1)}$	$\frac{99n}{100(n+1)}$
η_n	$\frac{33 \cdot 10^{20}}{n}$	$\frac{33 \cdot 10^{20}}{n}$	$\frac{33 \cdot 10^{20}}{n}$	$\frac{33 \cdot 10^{20}}{n}$	$\frac{33 \cdot 10^{20}}{n}$	$\frac{33 \cdot 10^{20}}{n}$
λ_n	$\frac{n}{L(n+1)}$	$\frac{n}{L(n+1)}$	$\frac{n}{L(n+1)}$	$\frac{31n}{20L(n+1)}$	$\frac{31n}{20L(n+1)}$	$\frac{31n}{20L(n+1)}$
θ	0.5	0.5	0.5	0.5	0.09	0.99

It is clear that these control parameters satisfy all conditions of Theorem 2. In this experiment, we take $f(x) = 0.25 \cdot x$. By Theorem 2, the sequence $\{x_n\}$ converges to the original image and its convergence behavior for each case is shown by the values of PSNR as seen in Table 1.

Table 1. The values of PSNR of six cases in Theorem 2 at $x_1, x_5, x_{10}, x_{25}, x_{50}, x_{100}, x_{250}, x_{500}$.

No. Iterations	**Case (a)**	**Case (b)**	**Case (c)**	**Case (d)**	**Case (e)**	**Case (f)**
1	19.435266	19.440476	19.528889	19.686937	19.686937	19.686937
5	20.539773	20.568442	20.849492	21.151064	20.888207	21.598216
10	21.126851	21.174059	21.606305	22.068696	21.554700	23.435160
25	22.136789	22.228847	22.949968	23.587612	22.788089	25.383824
50	23.125241	23.246531	24.087158	24.748261	23.893124	27.283463
100	24.198479	24.331965	25.199378	25.852069	25.000206	29.440357
250	25.605450	25.741976	26.608033	27.265159	26.408127	31.342524
500	26.645135	26.784026	27.674898	28.365891	27.466181	32.443743

The second experiment is to consider the behavior of the sequence $\{x_n\}$ for each case of contraction mappings $f(x) = k \cdot x$. We consider the following four different cases as follows:

Case (1)	$f(x) = 0.1 \cdot x$
Case (2)	$f(x) = 0.5 \cdot x$
Case (3)	$f(x) = 0.75 \cdot x$
Case (4)	$f(x) = 0.95 \cdot x$

We choose the parameters as follows:

$$\alpha_n = \frac{1}{33n}, \ \beta_n = \frac{n}{300n+1}, \ \gamma_n = 1 - \alpha_n - \beta_n, \ \sigma_n = \frac{99n}{100(n+1)}, \ \eta_n = \frac{33 \cdot 10^{20}}{n}, \ \lambda_n = \frac{31n}{20L(n+1)}.$$

Here $\theta = 0.99$ then

$$\theta_n = \begin{cases} \min\left\{0.99, \dfrac{10^{20}}{n^2\|x_n - x_{n-1}\|}\right\} & \text{if } \ x_n \neq x_{n-1}; \\ 0.99 & \text{otherwise.} \end{cases}$$

From Table 2, we get the values of PSNR at x_{500} of each case which equal to 32.212326, 32.929758, 33.580650 and 34.170032, respectively. We also observe from Table 2 and Figure 2 that when k is close to 1, the value of PSNR is higher than those of smaller k.

Table 2. The values of PSNR of four cases in Theorem 2 at x_1, x_5, x_{10}, x_{25}, x_{50}, x_{100}, x_{250}, x_{500}.

No. Iterations	Case (1)	Case (2)	Case (3)	Case (4)
1	19.645629	19.742816	19.781947	19.800875
5	21.595741	21.600924	21.601822	21.601207
10	23.430366	23.441553	23.445829	23.447628
25	25.433348	25.289174	25.178923	25.079196
50	27.329109	27.167872	26.997762	26.819210
100	29.375156	29.468395	29.352150	29.121785
250	31.091760	31.811123	32.260577	32.431003
500	32.212326	32.929758	33.580650	34.170032

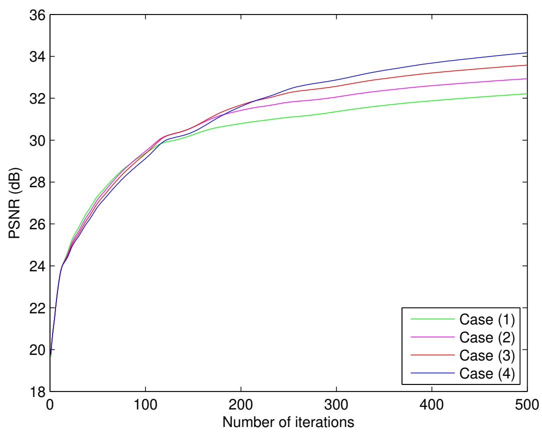


Figure 2. Comparison of four cases in Theorem 2.

On the other hand, the other experiment is to compare the quality of image restored by our algorithm and the quality of image restored by FISTA method [3]. Here, all parameters in Theorem 2 were the same as the previous experiment and we used $f(x) = 0.95 \cdot x$.

For FISTA method [3], we set

$$h(x) = \|Ax - b\|_2^2, \;\; g(x) = \lambda\|x\|_1 \text{ and } T = \text{prox}_{\lambda g}(I - \lambda\nabla h),$$

where the parameter $\lambda = \frac{1}{L}$.

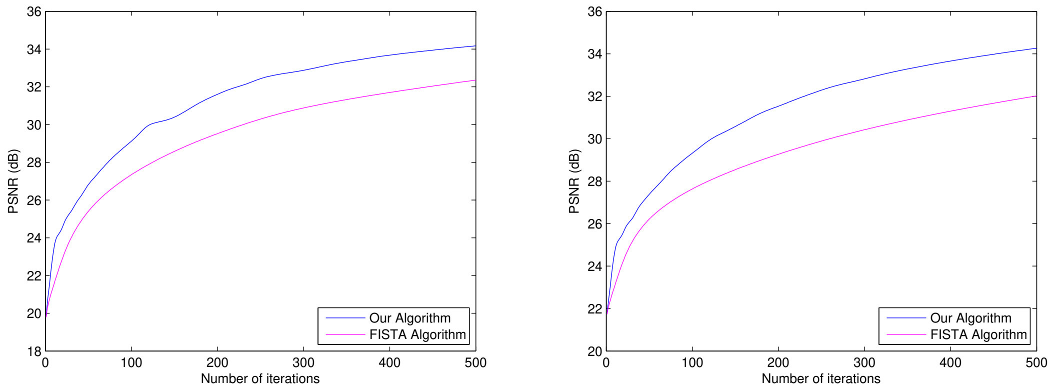
Then we obtain the PSNR values of our algorithm and of FISTA as seen in Tables 3 and 4, and Figure 3. The restoration images at 500th iteration of both algorithms are also presented in Figure 4.

Table 3. The values of PSNR at $x_1, x_5, x_{10}, x_{25}, x_{50}, x_{100}, x_{250}, x_{500}$ (Schonbrunn palace).

No. Iterations	Our Algorithm	FISTA Method
1	19.800875	19.785363
5	21.601207	20.774354
10	23.447628	21.530504
25	25.079196	23.502806
50	26.819210	25.401943
100	29.121785	27.342763
250	32.431003	30.290802
500	34.170032	32.356010

Table 4. The values of PSNR at $x_1, x_5, x_{10}, x_{25}, x_{50}, x_{100}, x_{250}, x_{500}$ (Camera man).

No. Iterations	Our Algorithm	FISTA Method
1	21.738865	21.730405
5	23.165748	22.429808
10	24.702169	23.081284
25	25.986847	24.741192
50	27.383389	26.213578
100	29.320097	27.633632
250	32.287817	29.889833
500	34.262873	32.016958



(**a**) Schonbrunn palace (**b**) Camera man

Figure 3. Plotting of the values of PSNR.

Our experiments show that our algorithm gives a better performance in restoring the blurred image than that of FISTA [3].

(**a**) Original palace

(**b**) Blurred palace

(**c**) By our algorithm

(**d**) By FISTA algorithm [3]

(**e**) Original camera man

(**f**) Blurred camera man

(**g**) By our algorithm

(**h**) By FISTA algorithm [3]

Figure 4. Original images: Schonbrunn palace (**a**), camera man (**b**), Blurred images: Schonbrunn palace (**b**), camera man (**f**), Results at x_{500} for restored images by our algorithm (**c**,**g**) and by FISTA method [3] (**d**,**h**).

5. Conclusions

In this paper, we present a new accelerated fixed point algorithm using the ideas of the viscosity and inertial technique to solving image restoration problems. A strong convergence theorem of our proposed method, Theorem 1, is established and proved under some suitable conditions. We then compare its convergence behavior with the others by considering its application to an image restoration problem. We find that our algorithm has convergence behavior better than FISTA which is a well-known and popular method using in image restoration problem.

Author Contributions: Funding acquisition and supervision, S.S.; writing—review and editing and software, J.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by Chiang Mai University, Chiang Mai, Thailand.

Acknowledgments: The authors would like to thank Chiang Mai University, Chiang Mai, Thailand for the financial support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tikhonov, A.N.; Arsenin, V.Y *Solution of Ill-Posed Problems*; V.H. Winston: Washington, DC, USA, 1977.
2. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B (Methodol.)* **1996**, *58*, 267–288, . [CrossRef]
3. Beck, A.; Teboulle, M. A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. *SIAM J. Imaging Sci.* **2009**, *2*, 183–202. [CrossRef]
4. Bauschke, H.H.; Combettes, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed.; Incorporated; Springer: New York, NY, USA, 2017.
5. Parikh, N.; Boyd, S. Proximal Algorithms. *Found. Trends® Optim.* **2014**, *1*, 127–239. [CrossRef]
6. Mann, W.R. Mean value methods in iteration. *Proc. Am. Math. Soc.* **1953**, *4*, 506–510. [CrossRef]
7. Reich, S. Weak convergence theorems for nonexpansive mappings in Banach spaces. *J. Math. Anal. Appl.* **1979**, *67*, 274–276. [CrossRef]
8. Moudafi, A. Viscosity approximation method for fixed-points problems. *J. Math. Anal. Appl.* **2000**, *241*, 46–55. [CrossRef]
9. Verma, M.; Sahu, D.R.; Shukla, K.K. VAGA: A novel viscosity-based accelerated gradient algorithm-Convergence analysis and applications. *Appl. Intell.* **2018**, *48*, 2613–2627. [CrossRef]
10. Nesterov, Y. A method for solving the convex programming problem with convergence rate (1/k2). *Dokl. Akad. Nauk SSSR* **1983**, *269*, 543–547.
11. Rockafellar, R.T. On the maximal monotonicity of subdifferential mappings. *Pac. J. Math.* **1970**, *33*, 209–216. [CrossRef]
12. Lions, P.L.; Mercier, B. Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.* **1979**, *16*, 964–979. [CrossRef]
13. Passty, G.B. Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *J. Math. Anal. Appl.* **1979**, *72*, 383–390. [CrossRef]
14. Tseng, P. A modified forward-backward splitting method for maximal monotone mappings. *SIAM J. Control Optim.* **2000**, *38*, 431–446. [CrossRef]
15. Combettes, P.L.; Hirstoaga, S.A. Equilibrium programming in Hilbert spaces. *J. Nonlinear Convex Anal.* **2005**, *6*, 117–136.
16. Lopez, G.; Martin-Marquez, V.; Wang, F.; Xu, H.K. Forward-backward splitting methods for accretive operators in Banach spaces. *Abstr. Appl. Anal.* **2012**. [CrossRef]
17. Takahashi, W. *Introduction to Nonlinear and Convex Analysis*; Yokohama Publishers: Yokohama, Japan, 2009.
18. Takahashi, W. *Nonlinear Functional Analysis*; Yokohama Publishers: Yokohama, Japan, 2000.
19. Goebel, K.; Reich, S. *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*; Marcel Dekker: New York, NY, USA, 1984.
20. Nakajo, K.; Shimoji, K.; Takahashi, W. Strong convergence to a common fixed point of families of nonexpansive mappings in banach spaces. *J. Math. Anal. Appl.* **2007**, *8*, 11–34.

21. Takahashi, W. Viscosity approximation methods for countable families of nonexpansive mappings in banach spaces. *Nonlinear Anal.* **2009**, *70*, 719–734. [CrossRef]
22. Bussaban, L.; Suantai, S.; Kaewkhao, A. A parallel inertial S-iteration forward-backward algorithm for regression and classification problems. *Carpathian. J.* **2020**, *36*, 35–44.
23. Goebel, K.; Kirk, W.A. *Topics in Metric Fixed Point Theory*; Cambridge Studies in Advanced Mathematics; Cambridge University Press: Cambridge, UK, 1990; Volume 28.
24. Aoyama, K.; Yasunori, Y.; Takahashi, W.; Toyoda, M. On a strongly nonexpansive sequence in a Hilbert space. *J. Nonlinear Convex Anal.* **2007**, *8*, 471–489.
25. Xu, H.K. Another control condition in an iterative method for nonexpansive mappings. *Bull. Aust. Math. Soc.* **2002**, *65*, 109–113. [CrossRef]
26. Mainge, P.E. Strong convergence of projected subgradient methods for nonsmooth and nonstrictly convex minimization. *Set-Valued Anal.* **2008**, *16*, 899–912. [CrossRef]
27. Thung, K.-H.; Raveendran, P. A Survey of Image Quality Measures. In Proceedings of the IEEE Technical Postgraduates (TECHPOS) International Conference, Kuala Lumpur, 14–15 December 2009; pp. 1–4.