Filomat 35:3 (2021), 771–782
https://doi.org/10.2298/FIL2103771B

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Inertial S-Iteration Forward-Backward Algorithm for a Family of Nonexpansive Operators with Applications to Image Restoration Problems

Limpapat Bussaban[a], Attapol Kaewkhao[b], Suthep Suantai[b,c]

[a]*Graduate Ph.D. Degree Program in Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand*
[b]*Research Center in Mathematics and Applied Mathematics, Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand*
[c]*Data Science Research Center, Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand*

Abstract. Image restoration is an important branch of image processing which has been studied extensively while there are several methods to solve this problem by many authors with the challenges of computational speed and accuracy of algorithms. In this paper, we present two methods, called "Inertial S-iteration forward-backward algorithm (ISFBA)" and "A fast iterative shrinkage-thresholding algorithm-Siteration (FISTA-S)", for finding an approximate solution of least absolute shrinkage and selection operator problem by using a special technique in fixed point theory and prove weak convergence of the proposed methods under some suitable conditions. Moreover, we apply our main results to solve image restoration problems. It is shown by some numerical examples that our algorithms have a good behavior compared with forward-backward algorithm (FBA), a new accelerated proximal gradient algorithm (nAGA) and a fast iterative shrinkage-thresholding algorithm (FISTA).

1. Introduction

A recently emerging technique used in signal and image processing is *compressive sensing* (CS). An important brance of image/signal processing is image restoration which is one of the most popular classical inverse problems. Such problem has been extensively studied in various applications such as image debluring, astronomical imaging, remote sensing, radar imaging, digital photography, microscopic imaging. The image restoration problem can be explained in one dimensional vector by the following model:

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

where $x \in \mathbb{R}^{N\times 1}$ is an original image, $b \in \mathbb{R}^{M\times 1}$ is the observed image, w is additive noise and $A \in \mathbb{R}^{M\times N}$ is the blurring operation. In order to solve problem (1), we aim to approximate the original image, vector x, by minimizing the additive noise, which is known as the *least squares* (LS) problem, by the following model:

$$\min_x \|Ax - b\|_2^2, \tag{2}$$

2010 *Mathematics Subject Classification*. 47H09; 90C25; 65F22
Keywords. S-iteration; forward-backward algorithm; nonexpansive operator; image restoration problem
Received: 20 February 2020; Revised: 05 November 2020; Accepted: 21 April 2021
Communicated by Vasile Berinde
Corresponding author: Suthep Suantai
Email addresses: lim.bussaban@gmail.com (Limpapat Bussaban), akaewkhao@yahoo.com (Attapol Kaewkhao), suthep.s@cmu.ac.th (Suthep Suantai)

where $\|\cdot\|_2$ is l_2-norm defined by $\|x\|_2 = \sqrt{\sum_{i=1}^{N} |x_i|^2}$. The solution of (2) can be estimated by many iterations e.g. Richardson iteration, see [26] for detail. However, the number of unknown variables is much more than the observations which causes (2) to be ill-posed problem because of a huge norm result which is thus meaningless, see [8] and [9]. Therefore, in order to improve ill-conditioned least squares problem, several *regularization methods*, were introduced such as the R_1 and R_2 regularization methods. As in [8], *the method of quasisolution*, called R_1, was introduced and studied by [11] while R_2 regularization method was suggested by [18] and [7]. One of the most popular regularization methods is *Tikhonov regularization* suggested by [24]. It is defined to solve the following minimization problem:

$$\min_x \|Ax - b\|_2^2 + \lambda \|Kx\|_2^2 \tag{3}$$

where $\lambda > 0$, is called regularization parameter, and $K \in \mathbb{R}^{P\times N}$, is called Tikhonov matrix. In the standard form, K is set to be the identity. In statistics, (3) is known as *ridge regression*. For improving the original LS (2) and classical regularization such as subset selection and ridge regression (3), a new method for estimation a solution of (1) called *least absolute shrinkage and selection operator* (LASSO), was proposed and discussed by [23] as follows:

$$\min_x \|Ax - b\|_2^2 + \lambda \|x\|_1 , \tag{4}$$

where $\|\cdot\|_1$ is l_1-norm defined by $\|x\|_1 = \sum_{i=1}^{N} |x_i|$. Moreover, the LASSO can be applied to regression problems [23], image restoration problems [4], etc.

In general, (2)-(4) can be formulated in a general form by estimating the minimizer of sum of two functions as follows:

$$\min_x F(x) := f(x) + g(x), \tag{5}$$

where g is a convex smooth (or possible non-smooth) function and f is a smooth convex loss function with gradient having Lipschitz constant L. By using Fermat's rule, Theorem 16.3 of [3], the solution of (5) can be characterized as follows: $\bar{x}$ minimizing $(f + g)$ if and only if $0 \in \partial g(\bar{x}) + \nabla f(\bar{x})$ where $\partial g(\bar{x})$ and $\nabla f(\bar{x})$ refer to the subdifferential and gradient of g and f respectively. Moreover, Parikh and Boyd [17] showed that problem (5) can also be interpreted as a fixed point problem: $\bar{x}$ minimizing $(f + g)$ if and only if

$$\bar{x} = prox_{cg}(I - c\nabla f)(\bar{x}) = J_{c\partial g}(I - c\nabla f)(\bar{x}), \tag{6}$$

where $c > 0$, I is an identity operator, $prox_{cg}$ is the proximity operator of cg and $J_{\partial g}$ is the resolvent of ∂g defined by $J_{\partial g} = (I + \partial g)^{-1}$, more description of these operators will be mentioned in Section 2. For convenience, the equation (6) can be rewritten as:

$$\bar{x} = T\bar{x}, \tag{7}$$

where $T := prox_{cg}(I - c\nabla f)$ which is called *forward-backward operator*. It is observed that a solution of (7) is a fixed point of T and T is a nonexpansive mapping when $c \in (0, \frac{2}{L})$. The existence of a fixed point of nonexpansive mappings was guaranteed by Browder's theorem, see [1] for detail. In order to find a point $\bar{x}$ satisfying (7), many researchers proposed various methods for finding the approximate solution. One of most popular iterative methods, called *Picard iteration process*, was defined by:

$$x_{n+1} = Tx_n, \tag{8}$$

where initial point x_1 is chosen randomly. In addition, other iterative methods for improving picard iteration process have been studied extensively such as follows.

Mann iteration process [14] is defined by:

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

where initial point x_1 is chosen randomly and $\{\alpha_n\}$ is a sequence in $[0,1]$. In case of $\alpha_n = 1$ for all $n \geq 1$, this iteration process reduces to the Picard iteration process.

Ishikawa iteration process [10] is defined by:

$$\begin{cases} y_n = (1-\beta_n)x_n + \beta_n T x_n, \\ x_{n+1} = (1-\alpha_n)x_n + \alpha_n T y_n, n \geq 1, \end{cases} \tag{10}$$

where initial point x_1 is chosen randomly and $\{\alpha_n\}$, $\{\beta_n\}$ are sequences in $[0,1]$. This iteration process reduces to the Mann iteration process when $\beta_n = 0$ for all $n \geq 1$.

S-iteration process [2] is defined by:

$$\begin{cases} y_n = (1-\beta_n)x_n + \beta_n T x_n, \\ x_{n+1} = (1-\alpha_n)T x_n + \alpha_n T y_n, n \geq 1, \end{cases} \tag{11}$$

where initial point x_1 is chosen randomly and $\{\alpha_n\}$, $\{\beta_n\}$ are sequences in $[0,1]$. In 2017, Agqrwal, O'Regan and Sahu proved that this iteration process is independent of Mann and Ishikawa iteration process and converges faster than both of them. However, the processes mentioned above have a badly convergence rate. Thus, to speed up, the technique for improving speed and giving a better convergence behavior was introduced firstly by Polyak [19] by adding an *inertial step*. The following classical iterative method for finding a zero of sum of two operators, i.e. find $x^* \in H$ such that $x^* \in zer(\nabla f + \partial g)$ can be viewed as Mann interation and it is known as

Forward-backward algorithm (FBA) is defined by:

$$\begin{cases} y_n = x_n - \gamma \nabla f x_n, \\ x_{n+1} = x_n + \alpha_n (J_{\gamma \partial g} y_n - x_n). \end{cases} \tag{12}$$

where $x_0 \in H$, L is a Lipschitz constant of ∇f, $\gamma \in (0, \frac{2}{L})$, $\delta = 2 - \frac{\gamma L}{2}$ and a sequence $\{\alpha_n\}$ in $[0, \delta]$ such that $\sum_{n \in \mathbb{N}} \alpha_n (\delta - \alpha_n) = +\infty$.

The following iterative methods with inertial step can be used for improving performance of Forward-backward algorithm.

A fast iterative shrinkage-thresholding algorithm (FISTA) [4], is defined by:

$$\begin{cases} y_n = T x_n, \\ t_{n+1} = \dfrac{1 + \sqrt{1 + 4t_n^2}}{2}, \theta_n \quad = \dfrac{t_n - 1}{t_{n+1}}, \\ x_{n+1} = y_n + \theta_n (y_n - y_{n-1}), \end{cases} \tag{13}$$

where $x_1 = y_0 \in \mathbb{R}^n$, $t_1 = 1$, $T := prox_{\frac{1}{L} g}(I - \frac{1}{L} \nabla f)$ and θ_n is called inertial step size. FISTA was suggested by Beck and Teboulle. They proved that rate of convergence of FISTA is better than that of ISTA and applied FISTA to image deblurring problems [4]. The inertial step size θ_n of FISTA was firstly introduced by Nesterov [16]. Generally, FISTA was modified for improving its performance by replacing t_{n+1}. For example, Chambolle and Dossal [6] turned out t_{n+1} to be $\frac{n+a}{a}$ for $a > 2$, Liang and Schönlieb [12] interpolated t_{n+1} into a general form as $t_{n+1} = \frac{p + \sqrt{q + r t_n^2}}{2}$ where $p, q > 0$ and $0 < r \leq 4$ and proved weak convergence theorem of FISTA.

A new accelerated proximal gradient algorithm (nAGA) [25], was defined by

$$\begin{cases} y_n = x_n + \theta_n (x_n - x_{n-1}), \\ x_{n+1} = T_n [(1-\alpha_n) y_n + \alpha_n T_n y_n], \end{cases} \tag{14}$$

where $\{\theta_n\}$, $\{\alpha_n\}$ are sequences in $(0,1)$ and $\dfrac{\|x_n - x_{n-1}\|_2}{\theta_n} \to 0$. They proved a convergence theorem of nAGA and applied this method for solving the non-smooth convex minimization problem with sparsity-inducing regularizers for the multitask learning framework.

Motivated by those works mentioned above, in this paper, an iterative method for sloving (5) is proposed by employing the concepts of S-iteration process together with the inertial step for a countable family of nonexpansive mappings. This paper is organized as follows: The basic concept and mathematical background will be given in Section 2. A weak convergence theorem will be proved in Section 3. Moreover, in Section 4, we apply the proposed method for solving image restoration problems.

2. Preliminaries and lemmas

Let H be a real Hilbert space with norm $\|\cdot\|$ and inner product $\langle\cdot|\cdot\rangle$. A mapping $T : H \to H$ is said to be *L-Lipschtiz operator* if there exists $L > 0$ such that $\|Tx - Ty\| \leq L\|x - y\|$ for any $x, y \in H$. An L-Lipschitz operator is called *nonexpansive operator* if $L = 1$. A mapping $A : H \to 2^H$ is called *monotone operator* if

$$\langle x - y|u - v\rangle \geq 0, \tag{15}$$

for any $(x, u), (y, v) \in graA$, where $graA = \{(x, y) \in H \times H : x \in H, y \in Ax\}$ is the graph of A. A monotone operator A is called *maximal monotone operator* if the graph $graA$ is not properly contained in the graph of any other monotone operator. It is known that A is maximal monotone operator if and only if $R(I + \lambda A) = H$ for every $\lambda > 0$.

Let $A : H \to 2^H$ be a maximal monotone operator and $c > 0$. The *resolvent* of A is defined by $J_{cA} = (I+cA)^{-1}$ where I is an identity operator. If $A = \partial f$ for some $f \in \Gamma_0(H)$, $\Gamma_0(H)$ is denoted by the set of proper lower semicontinuous convex functions from H to $(-\infty, +\infty]$, then $J_{cA} = prox_{cf}$ where $prox_f$ is *proximity operator* [3] of f given by

$$prox_f(x) = argmin_{y\in H}(f(y) + \frac{1}{2}\|x - y\|^2). \tag{16}$$

If $f = \|\cdot\|_1$, then $prox_{cf}$ can be represented by

$$prox_{c\|\cdot\|_1} = sgn(x)max\{\|x\|_1 - c, 0\}, \tag{17}$$

see chapter 24 in [3] for detial.

Let $\{T_n\}$ and $\mathcal{T}$ be families of nonexpansive operators 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 $T \in \mathcal{T}$. Then, $\{T_n\}$ is said to satisfy *NST-condition(I) with* $\mathcal{T}$ [15, 20] if for each bounded sequence $\{x_n\}$,

$$\lim_{n\to\infty} \|x_n - T_n x_n\| = 0 \text{ implies } \lim_{n\to\infty} \|x_n - Tx_n\| = 0 \text{ for all } T \in \mathcal{T}. \tag{18}$$

If $\mathcal{T}$ is singleton, i.e. $\mathcal{T} = \{T\}$, then $\{T_n\}$ is said to satisfy NST-condition(I) with T.

Theorem 2.1. [5, Theorem 3.2] *Let H be a Hilbert space. Let $A : H \to 2^H$ be a maximal monotone operator and $B : H \to H$ be an L-Lipschitz operator. Let $c \in (0, \frac{2}{L})$ and $\{c_n\} \subset (0, \frac{2}{L})$ such that $c_n \to c$. Define $T_n = J_{c_nA}(I - c_nB)$. Then, $\{T_n\}$ satisfies the NST-condition(I) with T_c where $T_c = J_{cA}(I - cB)$.*

Lemma 2.2. [21] *Let $\{a_n\}, \{b_n\}$ and $\{\delta_n\}$ be sequences of nonnegative numbers such that*

$$a_{n+1} \leq (1 + \delta_n)a_n + b_n, \forall n \in \mathbb{N}. \tag{19}$$

If $\sum_{n=1}^{\infty} \delta_n < \infty$ and $\sum_{n=1}^{\infty} b_n < \infty$, then $\lim_{n\to\infty} a_n$ exists.

Lemma 2.3 (Opial lemma). *Let H be a Hilbert space and $\{x_n\}$ be a sequence in H such that there exists a nonempty subset Ω of H satisfying the following conditions:*

- *for all $y \in \Omega$, $\lim_{n\to\infty} \|x_n - y\|$ exists,*
- *Any weak-cluster point of $\{x_n\}$ belongs to Ω.*

Then, there exists $\bar{x} \in \Omega$ such that $x_n \rightharpoonup \bar{x}$.

3. Main results

In this section, we propose an iterative method, called *Inertial S-iteration forward-backward algorithm (ISFBA)*, for finding a solution of (5) and prove a weak convergence theorem. First of all, we rewrite problem (5) into a general problem, called a *zero of sum of two operators problem*, by finding $\bar{x}$ such that

$$\bar{x} \in zer(A + B), \tag{20}$$

where $A, B : H \to 2^H$ are two set-valued operators and $zer(A + B) := \{x : 0 \in Ax + Bx\}$. In this case, we assume that $A : H \to 2^H$ is a maximal monotone operator and $B : H \to H$ is an L-Lipschitz operator. Hence, we prove, in general, a weak convergence theorem for a countable family of nonexpansive operators by assuming NST-condition(I) as follows.

Theorem 3.1. *Let H be a Hilbert space, $\{T_n\}$ be a family of nonexpansive operators and T be a nonexpansive operator such that $\{T_n\}$ satisfies NST-condition(I) with T. Suppose that $\emptyset \neq F(T) \subset \cap_{n=1}^{\infty} F(T_n)$. Let $\{x_n\}$ be a sequence in H generated by*

$$\begin{cases} x_0, x_1 \in H, \\ y_n = x_n + \theta_n(x_n - x_{n-1}), \\ z_n = (1 - \beta_n)y_n + \beta_n T_n y_n, \\ x_{n+1} = (1 - \alpha_n)T_n y_n + \alpha_n T_n z_n, \end{cases} \tag{21}$$

where $0 < q < \alpha_n \leq 1$, $0 < s < \beta_n < r < 1$, $0 \leq \theta_n \leq 1$ and $\sum_{n=1}^{\infty} \theta_n < \infty$. Then, $\{x_n\}$ converges weakly to a point in $F(T)$.

Proof. Let $x^* \in F(T)$ and let $\{x_n\}$ be a sequence in H generated by (21). Then,

$$\left\| y_n - x^* \right\| \leq \|x_n - x^*\| + \theta_n \|x_n - x_{n-1}\| \tag{22}$$

and

$$\|z_n - x^*\| \leq (1 - \beta_n) \left\| y_n - x^* \right\| + \beta_n \left\| T_n y_n - x^* \right\| \leq \left\| y_n - x^* \right\|. \tag{23}$$

Thus,

$$\begin{aligned} \|x_{n+1} - x^*\| &\leq (1 - \alpha_n) \left\| T_n y_n - x^* \right\| + \alpha_n \|T_n z_n - x^*\| \\ &\leq (1 - \alpha_n) \left\| y_n - x^* \right\| + \alpha_n \|z_n - x^*\| \\ &\leq \left\| y_n - x^* \right\| \\ &\leq \|x_n - x^*\| + \theta_n \|x_n - x_{n-1}\| \end{aligned} \tag{24}$$

Since $\|x_3 - x^*\| \leq (1 + 2\theta_n)K$ where $K = \max\{\|x_2 - x^*\|, \|x_1 - x^*\|\}$ and by induction, we can obtain that

$$\|x_{n+1} - x^*\| \leq K \prod_{j=1}^{n} (1 + 2\theta_j). \tag{25}$$

Since $\sum_{n=1}^{\infty} \theta_n < \infty$ and by limit comparison test, we get $\{x_n\}$ bounded. Hence, $\sum_{n=1}^{\infty} \theta_n \|x_n - x_{n-1}\| < \infty$. By using Lemma 2.2 in (24), we obtain that $\lim_{n\to\infty} \|x_n - x^*\|$ exists. Thus, $\{x_n\}$ is bounded which implies that $\{y_n\}$ is also bounded. From (21), we have

$$\begin{aligned} \left\| y_n - x^* \right\|^2 &= \|x_n - x^* + \theta_n(x_n - x_{n-1})\|^2 \\ &= \|x_n - x^*\|^2 + \theta_n^2 \|x_n - x_{n-1}\|^2 + 2\theta_n \langle x_n - x^* | x_n - x_{n-1} \rangle \\ &\leq \|x_n - x^*\|^2 + \theta_n^2 \|x_n - x_{n-1}\|^2 + 2\theta_n \|x_n - x^*\| \|x_n - x_{n-1}\|. \end{aligned} \tag{26}$$

Then,

$$\begin{aligned}\|z_n - x^*\|^2 &= \left\|(1-\beta_n)(y_n - x^*) + \beta_n(T_n y_n - x^*)\right\|^2 \\ &= (1-\beta_n)\left\|y_n - x^*\right\|^2 + \beta_n\left\|T_n y_n - x^*\right\|^2 - \beta_n(1-\beta_n)\left\|y_n - T_n y_n\right\|^2 \\ &\le \left\|y_n - x^*\right\|^2 - \beta_n(1-\beta_n)\left\|y_n - T_n y_n\right\|^2,\end{aligned} \tag{27}$$

and

$$\begin{aligned}\|x_{n+1} - x^*\|^2 &= (1-\alpha_n)\left\|T_n y_n - x^*\right\|^2 + \alpha_n\|T_n z_n - x^*\|^2 - \alpha_n(1-\alpha_n)\left\|T_n y_n - T_n z_n\right\|^2 \\ &\le (1-\alpha_n)\left\|T_n y_n - x^*\right\|^2 + \alpha_n\|T_n z_n - x^*\|^2 \\ &\le \left\|y_n - x^*\right\|^2 - \alpha_n\beta_n(1-\beta_n)\left\|y_n - T_n y_n\right\|^2 \\ &\le \|x_n - x^*\|^2 + \theta_n^2\|x_n - x_{n-1}\|^2 + 2\theta_n\|x_n - x^*\|\|x_n - x_{n-1}\| - \alpha_n\beta_n(1-\beta_n)\left\|y_n - T_n y_n\right\|^2.\end{aligned} \tag{28}$$

Since $\lim_{n\to\infty}\|x_n - x^*\|$ exists, it follows that $\left\|y_n - T_n y_n\right\| \to 0$. Since $\{y_n\}$ is bounded and $\{T_n\}$ satisfies NST-conditon(I) with T, we get $\left\|y_n - T y_n\right\| \to 0$. From

$$\begin{aligned}\|x_n - Tx_n\| &\le \left\|x_n - y_n\right\| + \left\|y_n - Ty_n\right\| + \left\|Ty_n - Tx_n\right\| \\ &\le 2\left\|x_n - y_n\right\| + \left\|y_n - Ty_n\right\| \\ &\le 2\theta_n\|x_n - x_{n-1}\| + \left\|y_n - Ty_n\right\|,\end{aligned} \tag{29}$$

we obtain $\|x_n - Tx_n\| \to 0$. Let w be a weak cluster point of $\{x_n\}$. Then $w \in F(T)$ by demicloseness of $I - T$ at 0. Hence, by using Opial lemma, we conclude that there exists $\bar{x} \in F(T)$ such that $x_n \rightharpoonup \bar{x}$. □

Corollary 3.2. *Let H be a Hilbert space. Let $A : H \to 2^H$ be maximal monotone operator and $B : H \to H$ be an L-Lipschitz operator. Let $c \in (0, \frac{2}{L})$ and $\{c_n\} \subset (0, \frac{2}{L})$ such that $c_n \to c$. Define $T_n = J_{c_nA}(I - c_nB)$ and $T = J_{cA}(I - cB)$. Suppose that $\emptyset \neq F(T) \subset \cap_{n=1}^{\infty}F(T_n)$. Let $\{x_n\}$ be a sequence in H generated by (21). Then, $\{x_n\}$ converges weakly to a point in $zer(A + B)$.*

Proof. Using Proposition 26.1(iv)(a) in [3], we obtain $F(T) = zer(A + B)$ and $\{T_n\}$ and T are nonexpansive operators for all n. Then, the proof is completed by Theorem 3.1 and Theorem 2.1. □

Corollary 3.3 (ISFBA). *Let H be a Hilbert space. Let $g \in \Gamma_0(H)$ and $f : H \to \mathbb{R}$ be convex and differentiable with an L-Lipschitz continuous gradient, let $c \in (0, \frac{2}{L})$ and $\{c_n\} \subset (0, \frac{2}{L})$ such that $c_n \to c$. Define $T_n = prox_{c_ng}(I - c_n\nabla f)$ and $T = prox_{cg}(I - c\nabla f)$. Suppose that $F(T) \neq \emptyset$. Let $\{x_n\}$ be a sequence in H generated by (21). Then, $\{x_n\}$ converges weakly to a point in $argmin(f + g)$.*

Proof. Setting $A := \partial g$ and $B := \nabla f$, then A is maximal monotone operator. We know that $F(T) = \cap_{n=1}^{\infty}F(T_n) = argmin(f + g) = zer(A + B)$. By Corollary 3.2, we obtain the required result. □

The following inspired from [4, 6], [12] and [13], by combining FISTA and S-iteration process.

Corollary 3.4 (FISTA-S). *Let H be a Hilbert space. Let $g \in \Gamma_0(H)$ and $f : H \to \mathbb{R}$ be convex and differentiable with an L-Lipschitz continuous gradient, let $c \in (0, \frac{2}{L})$ and $\{c_n\} \subset (0, \frac{2}{L})$ such that $c_n \to c$. Define $T_n = prox_{c_ng}(I - c_n\nabla f)$ and $T = prox_{cg}(I - c\nabla f)$. Suppose that $F(T) \neq \emptyset$. Let $t_1 = 1$ and $t_{n+1} = \frac{1+\sqrt{1+4t_n^2}}{2}$. Let N be a large positive number and $\{\gamma_n\}$ a summable positive real sequence. Define a sequence θ_n by*

$$\theta_n = \begin{cases} \frac{t_n - 1}{t_{n+1}}, & \text{if } 1 \le n \le N \\ \gamma_n, & \text{otherwise.} \end{cases} \tag{30}$$

Then a sequence $\{x_n\}$ generated by (21) converges weakly to a point in $argmin(f + g)$.

Proof. Since $\sum_{n=1}^{\infty}\gamma_n < \infty$, we get $\sum_{i=1}^{\infty}\theta_n < \infty$. Thus, the proof is completed by Theorem 3.1. □

Remark 3.5. *Corollary 3.4 can be proved under the sequence θ_n defined as in (30) and the condition of θ_n in Theorem (3.1), however, the proof with θ_n given by (13) remains open question.*

4. Applications to image restoration problems

In this section, we apply our proposed algorithms (ISFBA and FISTA-S) to solve image restoration problems. Moreover, we compare convergence behavior and efficiency of our algorithms with the classical algorithm namely, forward-backward algorithm (FBA) and the popular algorithms namely, FISTA, nAGA, introduced by Beck and Teboulle [4], Verma and Shukla [25], respectively. We set $f(x) = \frac{1}{2}\|Ax - b\|_2^2$, $g(x) = \lambda\|x\|_1$ and assume that $zer(\nabla f + \partial g) \neq \emptyset$ in Corollary 3.3. Then, we apply the proposed iterative methods, ISFBA and FISTA-S, to solve image restoration problems. All numerical experimental results are performed on Intel Core-i7 gen 8th with 8.00 GB RAM, windows 10, under MATLAB computing environment.

Let x be an original image, b be an observed image and A be a blurring operator. Then (5) can be rewritten as:

$$\min_x \|Ax - b\|_2^2 + \lambda\|x\|_1\,, \tag{31}$$

where a regularization parameter λ was choosen to be 5×10^{-5}. In this example, we consider two gray-scale images, Cameraman and Lena with size of 256×256, as the original images, and apply ISFBA (Corollary 3.3), FISTA-S (Theorem 3.4) to evaluate image bluring with Gaussian blur of size 9×9 and $\sigma = 4$. The orignal image and observed image are ginven in Figure 1.

In this example, we use the structural similarity index (SSIM) [27] and the peak signal-to-noise ratio (PSNR) [22], as a means of decision performance at x_n which are defined as follows

$$SSIM(x_n, x) = \frac{(2\mu_{x_n}\mu_x + C_1)(2\sigma_{x_n x} + C_2)}{(\mu_{x_n}^2 + \mu_x^2 + C_1)(\sigma_{x_n}^2 + \sigma_x^2 + C_1)}, \tag{32}$$

where $\{\mu_{x_n}, \sigma_{x_n}\}$ and $\{\mu_x, \sigma_x\}$ denote the mean intensity and standard deviation set of the deblerring image x_n and the original image x, respectively. $\sigma_{x_n x}$ denote their cross correlation. C_1 and C_2 are small constants value to avoid instability problem when the denominator is too close to zero.

$$PSNR(x_n) = 10\log_{10}\left(\frac{255^2}{MSE}\right), \tag{33}$$

where $MSE = \frac{1}{256^2}\|x_n - x\|^2$. A higher PSNR indicates that the deblurring image is of higher quality, that is, $PSNR(x_n)$ increases when the deblurring image x_n tend to the original image x. In order to estimate a solution of (31) by using ISFBA, FISTA-S, nAGA, FBA and FISTA, all controllers are setting in Table 1 and then, we obtain the results of

Figure 1: Original images: cameraman(left), lena(right), and their observed images with PSNR 21.3673 dB and 23.8492 dB, respectively.

iteration 200, 1000, 2000 and 3000, respectively, in Figure 4 and illustration of behavior of PSNR and of SSIM in Figure 2 and Figure 3, respectively. The quality of image produced by FBA (Mann iteration without inertial step) needs a large number of iterations to reach quality of image produced by ISFBA (S-iteration with inertial step). However, these iterations give a low quality (low PSNR value) and need a large number of interations in computing process compared with FISTA-S, nAGA and FISTA. The images produced by FISTA-S are of a better quality and lower iterations than those created by FISTA, nAGA.

Method	Setting
ISFBA (21)	$\alpha_n = \beta_n = 0.5, c_n = \frac{1.55n}{L(n+1)}, c = \frac{1}{L}$ and $\theta_n = \dfrac{1}{2^n}$
FISTA-S (Corollary 3.4)	$\alpha_n = \beta_n = 0.5, c_n = \frac{1.55n}{L(n+1)}, c = \frac{1}{L}$, $\gamma_n = \dfrac{1}{2^n}$ and θ_n defined as in (30)
nAGA (14)	$\alpha_n = 0.5, c_n = \frac{1.55n}{L(n+1)}, c = \frac{1}{L}$ and θ_n defined as in (13)
FBA (12)	$\alpha_n = 0.5, \gamma = \frac{1}{L}$

Table 1: Algorithms and their setting controls

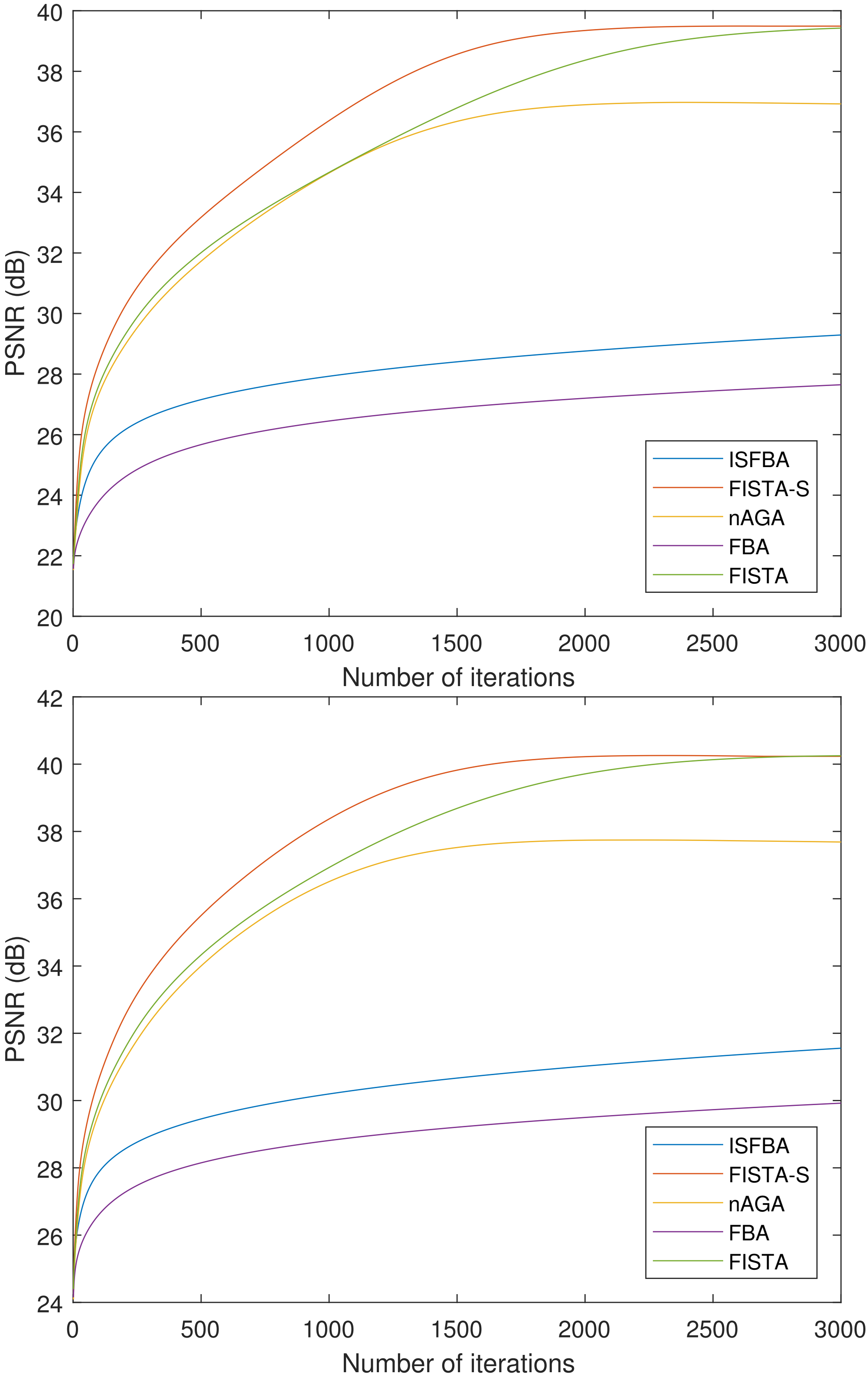

Figure 2: Comparision of efficiency of our algorithms with others using PSNR of cameraman (above) and lena (bottom)

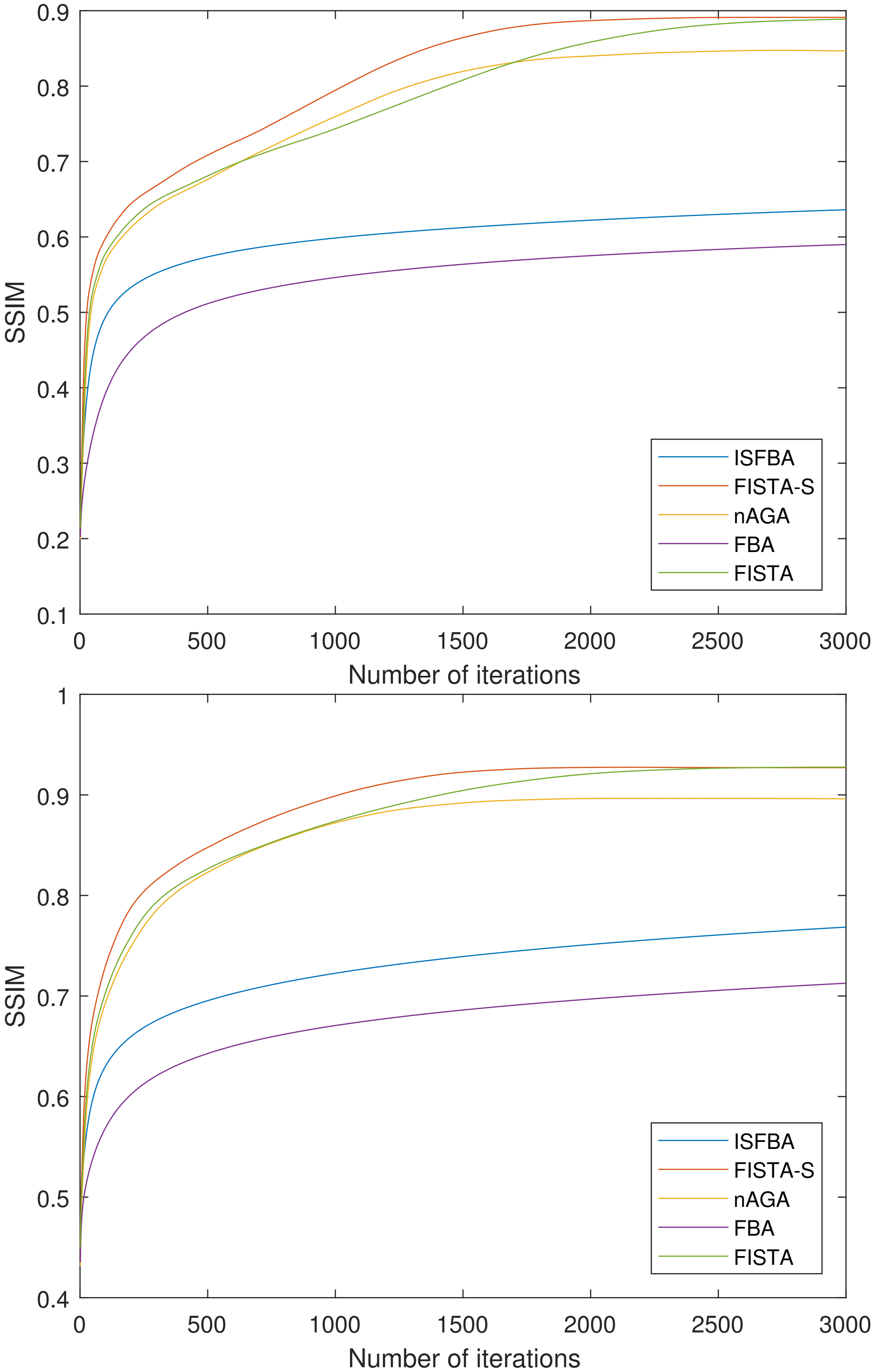

Figure 3: Comparision of efficiency of our algorithms with others using SSIM of cameraman (above) and lena (bottom)

Figure 4: $x_{200}, x_{1000}, x_{2000}, x_{3000}$ of cameraman and lena estimated by ISFBA, FISTA-S, nAGA, FBA and FISTA, respectively (from left to right)

Acknowledgments

This research was supported by Chiang Mai University and Thailand Science Research and Innovation under the project IRN62W0007. The first author would like to thanks the Thailand Research Fund through the Royal Golden Jubilee (RGJ) PhD Programme (Grant No. PHD/0184/2560).

References

[1] R. Agarwal, D. O'Regan, and D. Sahu. *Fixed Point Theory for Lipschitzian-type Mappings with Applications*. Topological Fixed Point Theory and Its Applications. Springer New York, 2009.
[2] R. P. Agarwal, D. O'Regan, and D. R. Sahu. Iterative construction of fixed point of nearly asymptotically nonexpansive mappings. *J. Nonlinear Convex Anal.*, 8(1):61–79, 2007.
[3] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer Publishing Company, Incorporated, 2nd edition, 2017.
[4] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.
[5] L. Bussaban, S. Suantai, and A. Kaewkhao. A parallel inertial s-iteration forward-backward algorithm for regression and classification problems. *Carpathian Journal of Mathematics*, 2020.
[6] A. Chambolle and C. Dossal. On the convergence of the iterates of the "fast iterative shrinkage/thresholding algorithm". *Journal of Optimization Theory and Applications*, 166(3):968–982, 2015.
[7] B. C. Cook. Least structure solution of photonuclear yield functions. *Nuclear Instruments and Methods*, 27:256–268, 1963.
[8] L. Eldén. Algorithms for the regularization of ill-conditioned least squares problems. *BIT Numerical Mathematics*, 17(2):134–145, 1977.
[9] P. C. Hansen, J. G. Nagy, and D. P. O'Leary. *Deblurring Images: Matrices, Spectra, and Filtering (Fundamentals of Algorithms 3) (Fundamentals of Algorithms)*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2006.
[10] S. Ishikawa. Fixed points by a new iteration method. *Proceedings of the American Mathematical Society*, 44:147–150, 1974.
[11] V. K. Ivanov. On linear problems which are not well-posed. *Dokl. Akad. Nauk SSSR.*, 145:270–272, 1962.
[12] J. Liang and C.-B. Schönlieb. Improving fista: Faster, smarter and greedier. http://arxiv.org/abs/1811.01430, 2018.
[13] D. A. Lorenz and T. Pock. An inertial forward-backward algorithm for monotone inclusions. *Journal of Mathematical Imaging and Vision*, 51(2):311–325, 2015.
[14] W. R. Mann. Mean value methods in iteration. *Proceedings of the American Mathematical Society*, 4:506–510, 1953.
[15] K. Nakajo, K. Shimoji, and W. Takahashi. Strong convergence to a common fixed point of families of nonexpansive mappings in banach spaces. *Journal of Nonlinear and Convex Analysis*, 8:11–34, 2007.
[16] Y. Nesterov. A method for solving the convex programming problem with convergence rate $o(1/k^2)$. *Dokl. Akad. Nauk SSSR*, 269(3):543 – 547, 1983.
[17] N. Parikh and S. Boyd. Proximal algorithms. *Found. Trends Optim.*, 1(3):127–239, 2014.
[18] D. L. Phillips. A technique for the numerical solution of certain integral equations of the first kind. *J. ACM.*, 9:84–97, 1962.
[19] B. Polyak. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1 – 17, 1964.
[20] W. Takahashi. Viscosity approximation methods for countable families of nonexpansive mappings in banach spaces. *Nonlinear Analysis: Theory, Methods & Applications*, 70(2):719 – 734, 2009.
[21] K. Tan and H. Xu. Approximating fixed points of nonexpansive mappings by the ishikawa iteration process. *Journal of Mathematical Analysis and Applications*, 178(2):301 – 308, 1993.
[22] K. Thung and P. Raveendran. A survey of image quality measures. In *2009 International Conference for Technical Postgraduates (TECHPOS)*, pages 1–4, 2009.
[23] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.
[24] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-posed problems*. W.H. Winston, 1977.
[25] M. Verma and K. Shukla. A new accelerated proximal gradient technique for regularized multitask learning framework. *Pattern Recognition Letters*, 95:98 – 103, 2017.
[26] C. R. Vogel. *Computational Methods for Inverse Problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.
[27] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 13(4):600–612, 2004.