Journal of Nonlinear and Convex Analysis
Volume 26, Number 10, 2025, 2915–2928

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

ACCELERATED COMMON FIXED POINT ALGORITHM FOR CONVEX MINIMIZATION PROBLEMS WITH APPLICATIONS IN DATA CLASSIFICATION

JIRAYUT BUTWANG* AND SUTHEP SUANTAI†

ABSTRACT. In this work, a new accelerated algorithm using double inertial technique for finding a common fixed point of a countable family of nonexpansive mappings in a Hilbert space is introduced and analyzed. A weak convergence theorem of the proposed algorithm under some suitable conditions is established. As a consequence, our main result can be applied for solving a convex minimization problem. Furthermore, we employ our proposed algorithm for solving data classification and compare its performance with the existing algorithms in the literature. Our experiments show that our proposed algorithm has a better performance than the others.

1. INTRODUCTION

Fixed point theory plays very important role in solving many real world problems. It can be used to prove the existence of a solution of the interseted problem and provides various methds for approximation such solutions. It can be applied to study various problems in science, applied science, engineering , ecomonics, data science and medical science and so on. It is well-known that fixed point theory is also crucial in artificail intelligence (AI) because it can be used to construct fast machine learnings algorithm which can be employed for building such intelligence AI for data prediction and classification. Most of AI models or machine learning algorithms are in the form of the optimization models which can be solved by fixed point theory. The mathematical models using in machine learning are in the form of convex minimization problems of the sum of two convex functions for the following form:

$$\min_{x\in\mathcal{H}} \psi_1(x) + \psi_2(x), \tag{1.1}$$

where $\psi_1 : \mathcal{H} \to \mathbb{R}$ is a proper lower semi-continuous and convex differentiable function such that $\nabla\psi_1$ is Lipschitz continuous with constant $L, \psi_2 : \mathcal{H} \to \mathbb{R}$ is a proper lower semi-continuous and convex function.

We denote $\arg\min(\psi_1 + \psi_2)$ the set of all solutions to problems (1) It is known that x is a solution of (1) if and only if x is a fixed point of the operator $T \coloneqq \text{prox}_{\lambda\psi_2}(I - \lambda\nabla\psi_1(x))$, where $\lambda > 0$ and $\text{prox}_{\lambda\psi_2}$ is the proximity operator of ψ_2

2020 *Mathematics Subject Classification.* 47H10, 49K35, 65K10.

Key words and phrases. Forward-backward algorithms, nonexpansive mapping, common fixed point, convex minimization, double inertial technique.

This work was supported by Chiang Mai University, Chiang Mai, Thailand.

*J. Butwang was supported by TA/RA Scholarship from Chiang Mai University.

†The second author would like to thank Chiang Mai University for financial support and Sripat medical center for supporting Hypertension dataset for our experiment.

defined by $\text{prox}_{\lambda\psi_2} = (\text{Id} - \lambda\partial\psi_2)^{-1}$ and $\partial\psi_2$ is the subdifferential of ψ_2. The operator $T \coloneqq \text{prox}_{\lambda\psi_2}(I - \lambda\nabla\psi_1)$ is called forward-backward operator and it is nonexpansive if $\lambda \in \left(0, \frac{2}{L}\right)$.

The calssical method for solving (1) is the forward-backward splitting algorithm (FBS) which was defined by $x_1 \in \mathbb{R}^n$ and $x_{k+1} = \text{prox}_{\lambda_k\psi_2}(\text{Id} - \lambda_k\nabla\psi_1)(x_k), k \geqslant 0$ where $\lambda_k \in \left(0, \frac{2}{L}\right)$ is a step size.

Fixed point methods are effective for solving convex minimization problems of the form of (1). In order to find a fixed point of a nonlinear operator $S : C \to C$ where C is a nonempty closed and convex subset of a real Hilbert space $\mathcal{H}$, several itertive methods were introduced and studied. One of the famous iterative methods is the Picard iteration process, defined by: $x_0 \in C$

$$x_{k+1} = Sx_k.$$

It is well known that $\{x_k\}$ converges strongly to a unique fixed point x^* of a contraction mapping S.

The Picard iteration process may not converge when S in not a contraction, such as S is nonexpansive. In the case taht S is nonexpansive, Mann [15] introduced an iteration process, called the Mann iteration process, defined as follows: $x_1 \in C$

$$x_{k+1} = (1 - \alpha_k)x_k + \alpha_k Sx_k$$

where $\{\alpha_k\} \subset (0, 1)$. He showed that $\{x_k\}$ converges weakly to a fixed poitn of S under certian control condition on $\{\alpha_k\}$.

Since Mann iteraton process provided only weak convergence, to obtain a strong convergence, Halpern [11] proposed the following fixed point algorithm: $x_0, x \in C$: $x_{k+1} = (1 - \alpha_k)x + \alpha_k Sx_k$, where $\{\alpha_k\} \subset (0, 1)$ and showed taht the sequence $\{x_k\}$ converges strongly to a fixed point of S under some control condition. It is worth to mentianing that when S is a pseudo contractive mapping, Mann iteration might not yield a convergence result. To overcome this problem, Ishikawa [13] proposed an iteration process known as Ishikawa iterative process, as follows: $x_1 \in C$,

$$\begin{cases} y_k = (1 - \alpha_k)x_k + \alpha_k Sx_k \\ x_{k+1} = (1 - \beta_k)x_k + \beta_k Sy_k \end{cases}$$

where $\{\alpha_k\}$ and $\{\beta_k\}$ are sequences in $(0, 1)$. He showed that the Ishikawa iteratoin process converges strongly to a fixed point of S.

Currently, by some modification of Ishikawas teraton, Agarwal et al. [2] introduced a new iteration process, called S-iteration process, as the following: $x_1 \in C$,

$$\begin{cases} y_k = (1 - \alpha_k)x_k + \alpha_k Sx_k \\ x_{k+1} = (1 - \beta_k)Sx_k + \beta_k Sy_k \end{cases}$$

where $\{\alpha_k\}$ and $\{\beta_k\}$ are sequences in $(0, 1)$. The convergence behavior of S-iteration was shown to be better than that of Mann and Ishikawa iterations.

For finding common fixed points of two mappings, Das and Debata [9] and Takahashi and Tamura [25] modified the Ishikawa iteration for two mapping S and T as follows: $x_1 \in C$

$$\begin{cases} x_{k+1} = (1-\beta_k)x_k + \beta_k S y_k \\ y_k = (1-\alpha_k)x_k + \alpha_k T x_k \end{cases}$$

where $\{\alpha_k\}$ and $\{\beta_k\}$ are sequences in $(0,1)$. Using a hybrid method in mathematical programming, Nakajo et al. [18] introduced an iterative methods for finding a common fixed point of a family of nonexpansive mappings. On the other hand, Aoyama et al. [4] modified the Halpern iteration by constructing the sequence $\{x_k\}$ by $x_{k+1} = (1-\alpha_k)x + \alpha_k S_k x_k$, where $\{\alpha_k\} \subset (0,1)$ and $x_1, x \in C$. Then, the strong convergence of the Halpern type algorithm was established under some suitable condition. For finding a common fixed point of a ountable family of nonexpansive mappings, Takahashi et al. [24] introduced an important condition known as the NST-condition (I). After that this condition was used continuously by many mathematicians.

In optimization theory, the inertial technique a useful tool for enhancing the convergence behavior of the algorithms. This technique was initially introduced by Polyak [21] for solving some convex minimization problems and many inertial-type algorithms have been proposed to solve various problems, see [1,7,8,30] for examples. To accelerate to convergence of FBS algorithm for solving (1), the inertial forward-backward splitting (IFBS) algorithm was proposed by Moudafi and Al-Shemas [16] as follows:

$$\begin{cases} y_k = x_k + \theta_k(x_k - x_{k-1}) \\ x_{k+1} = \operatorname{prox}_{\lambda_k \psi_2}(y_k - \lambda_k \nabla \psi_1(x_k)) \end{cases}$$

where $x_{-1}, x_0, x_1 \in \mathbb{R}^n, \theta_k \in (0,\infty)$ and $\lambda_k \in \left(0, \frac{2}{L}\right)$. They proved the convergence of IFBS under some suitable conditions on λ_k and θ_k.

The inertial parameter θ_k plays an important role for accellation the convergence behavior of the algorithms, Beck and Teboulle [7] suggested the fast iterative shrinkage-thresholding algorithm (FISTA) as defined by the following:

$$\begin{cases} y_k = x_k - \frac{1}{L}\nabla \psi_1(x_k), \\ t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}, \\ \theta_k = \frac{t_k - 1}{t_{k+1}}, \\ x_{k+1} = y_k + \theta_k(y_k - y_{k-1}), \end{cases}$$

where $k \in \mathbb{N}, x_1 = y_0 \in \mathbb{R}^n$ and $t_1 = 1$. they also examined the convergence rate of FISTA and applied it to image restoration problems. Their work highlights that the LASSO model is well-suited for addressing image restoration challenges.

Recently, Hanjing and Suantai [12] proposed an inertial Picard-Mann forward-backward splitting (iPM-FBS) method for solving (1) as follows:

$$\begin{cases} v_k = x_k + \alpha_k(x_k - x_{k-1}), \\ w_k = v_k + \beta_k(\operatorname{prox}_{\lambda_k \psi_2}(\mathrm{Id} - \lambda_k \nabla \psi_1)v_k - v_k), \\ x_{k+1} = \operatorname{prox}_{\lambda_k \psi_2}(\mathrm{Id} - \lambda_k \nabla \psi_1)w_k, \end{cases}$$

where $k \in \mathbb{N}, x_0, x_1 \in \mathbb{R}^n, \beta_k \in (0,1), \lambda_k \in \left(0, \frac{2}{L}\right)$ and $\alpha_k \in (0,\infty)$. They proved a convergence theorem of (iPM-FBS) under some suitable conditions and employed it for solving image inpainting.

Beside one step inertial technique which was used to accelerate the convergence behavior of the algorithm, for more faster convergence double inertial technique was used by many authors in the literature, see [20,22,29]. Recently, Wattanataweekul, Janngam and Suantai [28] proposed a new algorithm by using double inertial technique. Such algorithm was known as Two-Step Inertial Forward-Backward Bilevel Gradient Method (TIFB-BiGM) and defined by:

$$\begin{cases} \vartheta_k = \begin{cases} \min\left\{\mu_k, \frac{\tau_k}{\|x_k - x_{k-1}\|}\right\} & \text{if } x_k \neq x_{k-1}, \\ \mu_k & \text{otherwise,} \end{cases} \\ \delta_k = \begin{cases} \max\left\{-\rho_k, \frac{-\tau_k}{\|x_{k-1} - x_{k-2}\|}\right\} & \text{if } x_{k-1} \neq x_{k-2}, \\ -\rho_k & \text{otherwise,} \end{cases} \\ w_k = x_k + \vartheta_k(x_k - x_{k-1}) + \delta_k(x_{k-1} - x_{k-2}), \\ z_k = (1-\gamma_k)\operatorname{prox}_{c_k\psi_2}(\mathrm{Id} - c_k\nabla\psi_1)w_k + \gamma_k(\mathrm{Id} - s\nabla\omega)(w_k), \\ x_{k+1} = (1-\beta_k)\operatorname{prox}_{c_k\psi_2}(\mathrm{Id} - c_k\nabla\psi_1)w_k + \beta_k \operatorname{prox}_{c_k\psi_2}(\mathrm{Id} - c_k\nabla\psi_1)z_k, \end{cases}$$

where $k \in \mathbb{N}, x_{-1}, x_0, x_1 \in \mathbb{R}^n, \beta_k, \gamma_k \in [0,1], c_k \subset \left(0, \frac{2}{L_{\psi_1}}\right)$ and $\mu_k, \rho_k \subset (0,\infty)$. they proved the convergence of this algorithm and applied the main result to image recovery.

Motivated by the results discussed above, this paper introduces a new accelerated algorithm by using a two-step inertial technique for finding a common fixed point of two countable families of nonexpansive mappings. A convergence theorem of the proposed algorithm is established under certain conditions on the inertial parameters and control sequences. Additionally, we employ our algorithm to solve some convex minimization problems and data classification.

2. Preliminaries

In this section, we provide some basic definitions and useful results for proving the main results of our work.

Let $\mathcal{H}$ be a real Hilbert space with norm $\|\cdot\|$ and inner product $\langle\cdot,\cdot\rangle$. Let $S : \mathcal{H} \to \mathcal{H}$ be a mapping on $\mathcal{H}$. A point $x \in \mathcal{H}$ is said to be a fixed point of S if $Sx = x$ and we denote $\mathrm{Fix}(S)$, the set of all fixed pont of S. We say that S is Lipschitz continuous if there exists $L > 0$ such that

$$\|Su - Sv\| \leqslant L\|u - v\|, \forall u, v \in \mathcal{H}.$$

When $L = 1$, then S is said to be nonexpansive.

Let $\{S_k : \mathcal{H} \to \mathcal{H}$ and $S : \mathcal{H} \to \mathcal{H}\}$ be such that $\varnothing \neq \mathrm{Fix}(S) \subseteq \bigcap_{k=1}^{\infty} \mathrm{Fix}(S_k)$. Then, $\{S_k\}$ is said to satisfy NST-condition (I) with S [17] if, for all bounded sequence $\{x_k\}$ in $\mathcal{H}$,

$$\lim_{k\to\infty} \|x_k - S_k x_k\| = 0 \Longrightarrow \lim_{k\to\infty} \|x_k - Sx_k\| = 0.$$

A sequence $\{S_k\}$ is said to satisfy the condition (Z) [3, 5] if $\{x_k\}$ is a bounded sequence in $\mathcal{H}$ with $\lim_{k\to\infty}\|x_k - S_k x_k\| = 0$, we haave every weak cluster point of $\{x_k\}$ belongs to $\bigcap_{k=1}^{\infty}\mathrm{Fix}(S_k)$. The following identities are crucial for our main results.

Lemma 2.1 ([6])). *For any $y, z \in \mathcal{H}$ and $\eta \in [0,1]$, we have the following:*

(i) $\|y \pm z\|^2 = \|y\|^2 \pm 2\langle y, z\rangle + \|z\|^2$,

(ii) $\|\eta y + (1-\eta)z\|^2 = \eta\|y\|^2 + (1-\eta)\|z\|^2 - \eta(1-\eta)\|y - z\|^2$.

Lemma 2.2 ([8])). *Let $\psi_1 : \mathcal{H} \to \mathbb{R}$ be a convex and differentiable function such that $\nabla\psi_1$ is Lipschitz continuous with constant L and $\psi_2 : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ be a convex proper lower semi-continuous function. Let $S_k := \mathrm{prox}_{\lambda_k\psi_2}(Id - \lambda_k\nabla\psi_1)$ and $S := \mathrm{prox}_{\lambda\psi_2}(Id - \lambda\nabla\psi_1)$, when $0 < \lambda_k, \lambda < \frac{2}{L}$ with $\lim_{k\to\infty}\lambda_k = \lambda$. Then $\{S_k\}$ satisfies NST-condition (I) with S.*

Lemma 2.3 ([26]). *Let $\{x_k\}$ and $\{y_k\}$ be sequence of nonnegative real numbers such taht $x_{k+1} \leqslant x_k + y_k$, $\forall k \in \mathbb{N}$. If $\sum_{k=1}^{\infty} y_k < \infty$ then $\lim_{k\to\infty} x_k$ exists.*

Lemma 2.4 ([16,19]). *Let $\{x_k\}$ be a sequence in $\mathcal{H}$ such that there exists a nonempty set $\Gamma \subset \mathcal{H}$ satisfying:*

(i) *For every $x^* \in \Gamma, \lim_{k\to\infty}\|x_k - x^*\|$ exists;*

(ii) *$\omega_w(x_k) \subset \Gamma$, where $\omega_w(x_k)$ is the set of all weak-cluster points of $\{x_k\}$.*

Then, $\{x_k\}$ converges weakly to a point in Γ.

3. Main results

In this section, we first propose an new accelerated algorithm by using double inertial technique for finding a common fixed point of two countable families of nonexpansive mappings in real Hilbert and then analyze its convergence result under suitable conditions. Through out this section, we assume the following conditions:

(A1) $\{S_k : \mathcal{H} \to \mathcal{H}\}$ and $\{T_k : \mathcal{H} \to \mathcal{H}\}$ are two countable family of nonexpansive mappings which satisfy condition (Z);

(A2) $\Gamma := \bigcap_{k=1}^{\infty}\mathrm{Fix}(S_k) \cap \bigcap_{k=1}^{\infty}\mathrm{Fix}(T_k) \neq \varnothing$.

Our algorithm is called (TSFPA), for short, and defined as follows:

Algorithm 1: Two step Inertial Fixed Point Algorithm (TSFPA)

Initialization: Let $x_{-1}, x_0, x_1 \in \mathcal{H}$. Choose $\{\beta_k\}$ and $\{\gamma_k\}$, let $\{\tau_k\} \subset (0,\infty)$ and let $\{\mu_k\}, \{\rho_k\} \subset (0,\infty)$ be bounded sequences.

Step 1: Compute the inertial step:

$$\alpha_k = \begin{cases} \min\left\{\mu_k, \dfrac{\tau_k}{\|x_k - x_{k-1}\|}\right\}, & \text{if } x_k \neq x_{k-1} \\ \mu_k, & \text{Otherwise} \end{cases}$$

and

$$\theta_k = \begin{cases} \max\left\{-\rho_k, \dfrac{-\tau_k}{\|x_{k-1} - x_{k-2}\|}\right\}, & \text{if } x_{k-1} \neq x_{k-2} \\ -\rho_k, & \text{Otherwise} \end{cases}$$

$$v_k = x_k + \alpha_k(x_k - x_{k-1}) + \theta_k(x_{k-1} - x_{k-2}). \tag{3.1}$$

Step 2: Compute

$$w_k = (1 - \beta_k)v_k + \beta_k T_k v_k. \tag{3.2}$$

Step 3: Compute x_{k+1}

$$x_{k+1} = (1 - \gamma_k)w_k + \gamma_k S_k w_k. \tag{3.3}$$

Then go to Step 1.

Next, we prove a weak convergence result of Algorithm 1.

Theorem 3.1. *Let a sequence $\{x_k\}$ be generated by Algorithm 1. Suppose the sequences $\{\beta_k\}, \{\gamma_k\}$ and $\{\tau_k\}$ satisfy the following conditions:*

(C1) $\beta_k \in (a, b), \gamma_k \in (c, d)$ *for all* $k \in \mathbb{N}$ *and for some* $a, b, c, d \in \mathbb{R}$ *with* $0 < a < b < 1$ *and* $0 < c < d < 1$.

(C2) $\tau_k \geqslant 0,$ *for all* $k \in \mathbb{N}$ *and* $\sum_{k=1}^{\infty} \tau_k < \infty$.

Then $x_k \rightharpoonup x^* \in \Gamma$

Proof. Let $x^* \in \Gamma$, Using (3.1), we have

$$\|v_k - x^*\| \leqslant \|x_k - x^*\| + \alpha_k\|x_k - x_{k-1}\| + |\theta_k|\,\|x_{k-1} - x_{k-2}\| \tag{3.4}$$

Using (3.2) and nonexpansiveness of T_k, we have

$$\begin{aligned}\|w_k - x^*\| &\leqslant (1 - \beta_k)\|v_k - x^*\| + \beta_k\|T_k v_k - x^*\| \\ &\leqslant \|v_k - x^*\| \end{aligned} \tag{3.5}$$

Using (3.3) and nonexpansiveness of S_k, we get

$$\begin{aligned}\|x_{k+1} - x^*\| &\leqslant (1 - \gamma_k)\|w_k - x^*\| + \gamma_k\|S_k w_k - x^*\| \\ &\leqslant \|w_k - x^*\| \end{aligned} \tag{3.6}$$

From (3.4) - (3.6), we obtain

$$\begin{aligned}\|x_{k+1} - x^*\| &\leqslant \|w_k - x^*\| \\ &\leqslant \|v_k - x^*\| \\ &\leqslant \|x_k - x^*\| + \alpha_k\|x_k - x_{k-1}\| + |\theta_k|\,\|x_{k-1} - x_{k-2}\| \end{aligned} \tag{3.7}$$

By definition of $\{\alpha_k\}$ and $\{\theta_k\}$ with (C2), we conclude that $\{x_k\}$ is bounded and also $\{v_k\}$ and $\{w_k\}$. This implies $\sum_{k=1}^{\infty} \alpha_k\|x_k - x_{k-1}\| < \infty$ and $\sum_{k=1}^{\infty}|\theta_k|\,\|x_{k-1} - x_{k-2}\| < \infty$. Using (3.7) and Lemma 2.3, we obtain that $\lim_{k\to\infty}\|x_k - x^*\|$ exists. Using Lemma 2.1 (i) and (3.1), we obtain

$$\begin{aligned}\|v_k - x^*\|^2 &\leqslant \|x_k - x^*\|^2 + \|\alpha_k(x_k - x_{k-1}) + \theta_k(x_{k-1} - x_{k-2})\|^2 \\ &\quad + 2\|x_k - x^*\|\|\alpha_k(x_k - x_{k-1}) + \theta_k(x_{k-1} - x_{k-2})\| \\ &\leqslant \|x_k - x^*\|^2 + \|\alpha_k(x_k - x_{k-1}) + \theta_k(x_{k-1} - x_{k-2})\|^2\end{aligned}$$

$$+ 2\|x_k - x^*\|(\alpha_k\|x_k - x_{k-1}\| + |\theta_k|\,\|x_{k-1} - x_{k-2}\|) \tag{3.8}$$

Using Lemma 2.1 (ii), (3.2), (3.3), we get

$$\begin{aligned}\|w_k - x^*\|^2 &= (1-\beta_k)\|v_k - x^*\|^2 + \beta_k\|T_k v_k - x^*\|^2 - \beta_k(1-\beta_k)\|v_k - T_k v_k\|^2 \\ &\leqslant \|v_k - x^*\|^2 - \beta_k(1-\beta_k)\|v_k - T_k v_k\|^2 \end{aligned} \tag{3.9}$$

and

$$\begin{aligned}\|x_{k+1} - x^*\|^2 &= (1-\gamma_k)\|w_k - x^*\|^2 + \gamma_k\|S_k w_k - x^*\|^2 - \gamma_k(1-\gamma_k)\|w_k - S_k w_k\|^2 \\ &\leqslant \|w_k - x^*\|^2 - \gamma_k(1-\gamma_k)\|w_k - S_k w_k\|^2 \end{aligned} \tag{3.10}$$

From (3.8) - (3.10), we have

$$\begin{aligned}\|x_{k+1} - x^*\|^2 &\leqslant \|w_k - x^*\|^2 - \gamma_k(1-\gamma_k)\|w_k - S_k w_K\|^2 \\ &\leqslant \|v_k - x^*\|^2 - \beta_k(1-\beta_k)\|v_k - T_k v_k\|^2 - \gamma_k(1-\gamma_k)\|w_k - S_k w_K\|^2 \\ &\leqslant \|x_k - x^*\|^2 + \|\alpha_k(x_k - x_{k-1}) + \theta_k(x_{k-1} - x_{k-2})\|^2 \\ &\quad + 2\|x_k - x^*\|(\alpha_k\|x_k - x_{k-1}\| + |\theta_k|\,\|x_{k-1} - x_{k-2}\|) \\ &\quad - \beta_k(1-\beta_k)\|v_k - T_k v_k\|^2 - \gamma_k(1-\gamma_k)\|w_k - S_k w_K\|^2 \end{aligned} \tag{3.11}$$

From (3.11) and by condition (C1), $\sum_{k=1}^{\infty}\alpha_k\|x_k - x_{k-1}\| < \infty, \sum_{k=1}^{\infty}|\theta_k|\,\|x_{k-1} - x_{k-2}\| < \infty$, and $\lim_{k\to\infty}\|x_k - x^*\|$ exists, we have

$$\lim_{k\to\infty}\|v_k - T_k v_k\| = 0 \text{ and } \lim_{k\to\infty}\|w_k - S_k w_k\| = 0$$

From (3.2), we have

$$\begin{aligned}\|w_k - v_k\| &= \|(1-\beta_k)v_k + \beta_k T_k v_k - v_k\| \\ &= \| - \beta_k v_k + \beta_k T_k v_k\| \\ &= \beta_k\|T_k v_k - v_k\|\end{aligned}$$

It follows that

$$\lim_{k\to\infty}\|w_k - v_k\| = 0. \tag{3.12}$$

Since $\sum_{k=1}^{\infty}\alpha_k\|x_k - x_{k-1}\| < \infty$ and $\sum_{k=1}^{\infty}|\theta_k|\,\|x_{k-1} - x_{k-2}\| < \infty$, we have by (3.1) that

$$\|v_k - x_k\| \leqslant \alpha_k\|x_k - x_{k-1}\| + |\theta_k|\,\|x_{k-1} - x_{k-2}\| \to 0 \text{ as } k \to \infty.$$

Therefore

$$\lim_{k\to\infty}\|v_k - x_k\| = 0. \tag{3.13}$$

Since $\{x_k\}$ is bounded, we know that $\omega_w(x_k) \neq \varnothing$. Using (3.12) and (3.13), we get $\omega_w(x_k) \subseteq \omega_w(v_k) \subseteq \omega_w(w_k)$. Since $\{S_k\}$ and $\{T_k\}$ satisfy the condition (Z), $\lim_{k\to\infty}\|v_k - T_k v_k\| = \lim_{k\to\infty}\|w_k - S_k w_k\| = 0$, we have $\omega_w(v_k) \subseteq \bigcap_{k=1}^{\infty}\mathrm{Fix}(T_k)$ and $\omega_w(w_k) \subseteq \bigcap_{k=1}^{\infty}\mathrm{Fix}(S_k)$. That is $\omega_w(x_k) \subseteq \bigcap_{k=1}^{\infty}\mathrm{Fix}(S_k) \cap \bigcap_{k=1}^{\infty}\mathrm{Fix}(T_k)$. By Lemma 2.4 we can conclude that $\{x_k\}$ converges weakly to a point in Γ. □

As a consequence of Theorem 3.1, we are ready to introduce the double inertial forward-backward splitting method for solving the problem (1). In order to do this, the following assumptions are given:

- **(D1)** $\xi_1 : \mathcal{H} \to \mathbb{R}$ and $\psi_1 : \mathcal{H} \to \mathbb{R}$ are differentiable and convex functions;
- **(D2)** $\nabla\xi_1$ and $\nabla\psi_1$ are Lipschitz continuous with constants L_1 and L_2, respectively;
- **(D3)** $\xi_2 : \mathcal{H} \to \mathbb{R}$ and $\psi_2 : \mathcal{H} \to \mathbb{R}$ are convex lower semi-continuous functions;
- **(D4)** $\Theta := \arg\min(\xi_1 + \xi_2) \cap \arg\min(\psi_1 + \psi_2) \neq \emptyset$

Remark 1. Let $T_k := \operatorname{prox}_{\lambda_k \xi_2}(\mathrm{Id} - \lambda_k \nabla\xi_1)$ and $T := \operatorname{prox}_{\lambda\xi_2}(\mathrm{Id} - \lambda\nabla\xi_1)$. It is known that T_k and T are nonexpansive when $\lambda_k, \lambda \in \left(0, \frac{2}{L_1}\right)$ and $\mathrm{Fix}(T) = \arg\min(\psi_1 + \psi_2) = \bigcap_{k=1}^{\infty} \mathrm{Fix}(T_k)$. By Lemma (2.2), we also knew that $\{T_k\}$ satisfies NST-condition (I) with T if $\lambda_k \to \lambda$.

Algorithm 2:

Initialization: Let $x_{-1}, x_0, x_1 \in \mathcal{H}$. Choose $\{\beta_k\}, \{\gamma_k\}, \{\lambda_k\}$ and $\{\lambda_k^*\}$, let $\{\tau_k\} \subset (0, \infty)$ and let $\{\mu_k\}, \{\rho_k\} \subset (0, \infty)$ be bounded sequences.

Step 1: Compute the inertial step:

$$\alpha_k = \begin{cases} \min\left\{\mu_k, \dfrac{\tau_k}{\|x_k - x_{k-1}\|}\right\}, & \text{if } x_k \neq x_{k-1} \\ \mu_k, & \text{Otherwise} \end{cases}$$

and

$$\theta_k = \begin{cases} \max\left\{-\rho_k, \dfrac{-\tau_k}{\|x_{k-1} - x_{k-2}\|}\right\}, & \text{if } x_{k-1} \neq x_{k-2} \\ -\rho_k, & \text{Otherwise} \end{cases}$$

$$v_k = x_k + \alpha_k(x_k - x_{k-1}) + \theta_k(x_{k-1} - x_{k-2})$$

Step 2: Compute

$$w_k = (1 - \beta_k)v_k + \beta_k \operatorname{prox}_{\lambda_k \xi_2}(\mathrm{Id} - \lambda_k \nabla\xi_1)v_k$$

Step 3: Compute x_{k+1}

$$x_{k+1} = (1 - \gamma_k)w_k + \gamma_k \operatorname{prox}_{\lambda_k^* \psi_2}(\mathrm{Id} - \lambda_k^* \nabla\psi_1)w_k$$

Now, we prove the convergence result of Algorithm 2.

Theorem 3.2. *Let $\{x_k\}$ be a sequence generated by Algorithm 2. Suppose that the sequences $\{\beta_k\}, \{\gamma_k\}$ and $\{\tau_k\}$ satisfy the following conditions:*

- (C1) $\beta_k \in (a, b), \gamma_k \in (c, d)$, *for some* $a, b, c, d \in \mathbb{R}$ *with* $0 < a < b < 1$ *and* $0 < c < d < 1$.
- (C2) $\tau_k \geqslant 0,\ \forall k \in \mathbb{N}$ *and* $\sum_{k=1}^{\infty} \tau_k < \infty$.
- (C3) $0 < \lambda_k, \lambda < \frac{2}{L_1}, 0 < \lambda_k^*, \lambda^* < \frac{2}{L_1}, \forall k \in \mathbb{N}$ *such that* $\lambda_k \to \lambda$ *and* $\lambda_k^* \to \lambda^*$ *as* $k \to \infty$

Then $x_k \rightharpoonup x^* \in \Theta$

Proof. For each $k \in \mathbb{N}$, let $T_k, S_k, T, S : \mathcal{H} \to \mathcal{H}$ be defined by:

$$T_k := \text{prox}_{\lambda_k \xi_2}(\text{Id} - \lambda_k \nabla \xi_1) \text{ and } T := \text{prox}_{\lambda \xi_2}(\text{Id} - \lambda \nabla \xi_1)$$
$$S_k := \text{prox}_{\lambda_k^* \psi_2}(\text{Id} - \lambda_k^* \nabla \psi_1), \;\; S := \text{prox}_{\lambda^* \psi_2}(\text{Id} - \lambda^* \nabla \psi_1).$$

By Remark 1, we know that $\{T_k\}$ and $\{S_k\}$ are two families of nonexpansive mappings satisfying NST-condition (I) with T and S, respectively. Hence $\{T_k\}$ and $\{S_k\}$ satisfy the condition (Z). By theorem 3.1, we obtain that $x_k \rightharpoonup x$ as $k \to \infty$ for some $x \in \Theta$. □

Remark 2. If $\xi_1 = \psi_1, \xi_2 = \psi_2$ and $\lambda_k = \lambda_k^*$ then the Algorithm 2 is reduced to Algorithm 3 which can be used for solving the convex minimization problem (1).

Algorithm 3: Two step Inertial Forward-Backward Algorithm (TSFBA)

Initialization: Let $x_{-1}, x_0, x_1 \in \mathcal{H}$. Choose $\{\beta_k\}$ and $\{\gamma_k\}$, et $\{\tau_k\} \subset (0, \infty)$ and let $\{\mu_k\}, \{\rho_k\} \subset (0, \infty)$ be bounded sequences.

Step 1: Compute the inertial step:

$$\alpha_k = \begin{cases} \min\left\{\mu_k, \dfrac{\tau_k}{\|x_k - x_{k-1}\|}\right\}, & \text{if } x_k \neq x_{k-1} \\ \mu_k, & \text{Otherwise} \end{cases}$$

and

$$\theta_k = \begin{cases} \max\left\{-\rho_k, \dfrac{-\tau_k}{\|x_{k-1} - x_{k-2}\|}\right\}, & \text{if } x_{k-1} \neq x_{k-2} \\ -\rho_k, & \text{Otherwise} \end{cases}$$

$$v_k = x_k + \alpha_k(x_k - x_{k-1}) + \theta_k(x_{k-1} - x_{k-2})$$

Step 2: Compute

$$w_k = (1 - \beta_k)v_k + \beta_k \, \text{prox}_{\lambda_k \psi_2}(\text{Id} - \lambda_k \nabla \psi_1)v_k$$

Step 3: Compute x_{k+1}

$$x_{k+1} = (1 - \gamma_k)w_k + \gamma_k \, \text{prox}_{\lambda_k \psi_2}(\text{Id} - \lambda_k \nabla \psi_1)w_k$$

---

4. Applications

In this section, we employ Algorithm 3 as a machine learning algorithm for data classification of some data by using a model of Single Hidden Layer Feedforward Neural Networks (SLFNs) and Extreme Learning Machine.

We first recall some basic concept of the extreme learning machine (ELM) for data classification problems and then employ our proposed algorithm for solving data classification and compare its performance with the other methods.

Let $D = \{(x_i, t_i) : x_i \in \mathbb{R}^n, t_i \in \mathbb{R}^m, i = 1, 2, \ldots, N\}$ be a training set with N distinct samples, x_i and t_i are input data and target, respectively. In this aplication, we use the standard single hidden layer feedforward neural networks with M hidden nodes and activation function $g(x)$, as a mathematical model for data classification. The model is given by

$$\sum_{j=1}^{M} \beta_j g(\langle w_j, x_i\rangle + b_j) = o_i,\ i = 1, \dots, N,$$

where β_j is the weight vector connecting the jth hidden node and the output node, b_j is the threshold of the jth hidden node, and w_j is the weight vector connecting the jth hidden node and the input node. The target of SLFNs is to find β_j, w_j, b_j such that

$$\sum_{j=1}^{M} \beta_j g(\langle w_j, x_i\rangle + b_j) = t_i,\ i = 1, \dots, N.$$

We can transform above system of N equations into a simple matrix form as the following:

$$H\beta = T, \tag{4.1}$$

where

$$H = \begin{bmatrix} g(\langle w_1, x_1\rangle + b_1) & \cdots & g(\langle w_M, x_1\rangle + b_M) \\ \vdots & \ddots & \vdots \\ g(\langle w_1, x_N\rangle + b_1) & \cdots & g(\langle w_M, x_N\rangle + b_M) \end{bmatrix}_{N\times M}$$

$$\beta = \left[\beta_1^T, \dots, \beta_M^T\right]^T_{m\times M},\ T = \left[t_1^T, \dots, t_N^T\right]^T_{m\times N}.$$

In the standard SLFNs, we want to estimate β_j, w_j and b_j for solving (4.1) while ELM aims to find only β_j with randomly w_j and b_j.

In order to solve problem (4.1), we consider the following optimization problem

$$\min_{\beta} \|H\beta - T\|_2^2 + \lambda\|\beta\|_1, \tag{4.2}$$

where $\|x\|_1 = \sum_{i=1}^{n} |x_i|, x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $\lambda > 0$ is a ragulalized parameter. The model (4.2) is known as the least absolute shrinkage and selection operator (LASSO model) [27] which is very useful and opular model to overcome the over-fitting problem in prediction and classification.

We use our algorithm, Algorithm 3, as a machine learning algorithm for finding the optimal weight w_j of the convex minimzation problem (4.2) and compare its performance for data classification with the other methods. The data sets used for our experiment obtained from UC Irvine Machine Learning Repository (UCI) and Sriphat Medicine Center, faculty of medicine, Chiang Mai university.

In our experiment, we use the following datasets:

- Iris dataset: This dataset is a plant species which contains 3 classes of 150 instances where each class refers to a type of iris plant. Iris plant consists of three type as follows: iris setosa, iris versicolour, and iris virginica . They

can be separated from sepal and teta length. This dataset was created by Fisher [30] and it can be download via https://doi.org/10.24432/C56C76
- Diabetes: This is a medical dataset which contains 436 examples and each example has 8 features, this data set has two types. The source of this dataset is referred to [31], Diabetes dataset, Kaggle, 1990. Available from: https://www.kaggle.com/datasets/mathchi/diabetes-data-set/data.
- Hypertension: This is a medical dataset which contains 6138 examples and each example has 9 features, this data set has two classes. The data was collected by Sripat Medical Center, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand.

Table 1 provides information of the studied datasets, number of attributes and number of samples for training (around 70% of data) and testing (remainder 30% of data) sets.

Table 1. Training and testing sets of each dataset.

Dataset	**Attributes**	Sample	
		training	testing
Iris	4	105	45
Hypertension	9	4276	1832
Diabetes	8	306	130

For data classification, an accuracy of the output data is calculated by

$$\text{accuracy} = \frac{TP + TN}{TP + FN + TN + FP} \times 100$$

where TP, FP, TN and FN stand for number of samples correctly predicted as positive, number of samples wrongly predicted as positive, number of samples correctly predicted as negative and number of samples wrongly predicted as negative, respectively.

We set control parameters with $\lambda_k = 1 \times 10^{-5}$ as in the table 2.

Table 2. Details of parameters for each method.

Methods	Setting
TSFBA	$\tau_k = \frac{10^{18}}{k^2}, \mu_k = \frac{k}{k+1}, \rho_k = \frac{-1}{k}$ $\beta_k, \gamma_k = \begin{cases} \frac{0.5k}{k+1}, & \text{Iris and Hypertention data;} \\ \frac{2}{k}, & \text{Diabetes data.} \end{cases}$
iPM-FBS	$\alpha_k = \begin{cases} \frac{k}{k+1}, & \text{if } 1 \leqslant k \leqslant 4000; \\ \frac{1}{2^k}, & \text{otherwise.} \end{cases}$ $\beta_k = \frac{0.99k}{k+1}$
FISTA	$t_1 = 1, t_{n+1} = (1 + \sqrt{1 + 4t_n^2})/2,$ $\theta_n = (t_n - 1)/t_{n+1}$

Note that all of the parameters in Table 2 satisfy the convergence theorems for each method. We use sigmoid as an activated function and the number of hidden nodes $M = 100$.

Table 3 shows the performance of each algorithm in term of accuracy of training set and accuracy of testing set.

Table 3. Accuracy for each algorithm.

Dataset	Iris		Hypertension		Diabetes	
	Train	Test	Train	Test	Train	Test
TSFBA	98.1	100	89.45	90.23	90.16	98.47
iPM-FBS	97.14	97.78	89.38	89.47	96.39	96.18
FISTA	86.67	93.33	87.16	88.16	85.9	91.6

Table 3, provides comparision results of the performance of the studied algorithms (TSFBA, iPM-FBS, and FISTA) in terms of accuracy for both training and testing datasets across three datasets: Iris, Hypertension, and Diabetes. TSFBA consistently shows the highest accuracy for both training and testing sets across all datasets. For Iris dataset, TSFBA achieves perfect test accuracy, outperforming iPM-FBS and FISTA. For Hypertension dataset, TSFBA performs slightly better than the other algorithms. And for Diabetes dataset, TSFBA again shows the highest accuracy on both training and testing sets.

And, Table 4 shows us another score of each algorithm where each score calculated

Table 4. Another score of each algorithm.

Dataset	Algorithm	Precision Score		Recall Score		F1 Score	
		Training	Testing	Training	Testing	Training	Testing
Iris	TSFBA	0.98	1	0.98	1	0.98	1
	iPM-FBS	0.97	0.98	0.97	0.98	0.97	0.98
	FISTA	0.87	0.94	0.87	0.93	0.87	0.93
Hypertension	TSFBA	0.87	0.89	0.91	0.91	0.89	0.9
	iPM-FBS	0.87	0.89	0.9	0.89	0.88	0.89
	FISTA	0.8	0.83	0.95	0.94	0.87	0.88
Diabetes	TSFBA	0.92	0.99	0.89	0.99	0.91	0.99
	iPM-FBS	0.94	0.96	1	0.98	0.97	0.97
	FISTA	0.94	0.97	0.8	0.89	0.86	0.93

by

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1} = \frac{2 \times (\text{Precision}) \times (\text{Recall})}{\text{Precision} + \text{Recall}}$$

This table shows that TSFBA consistently outperforms iPM-FBS and FISTA across all datasets and achieves the highest precision, recall, and F1 scores. This table demonstrates effectiveness of each algorithm for data classification tasks in various scenarios.

5. Conclusions

In this work, a new fast algorithm for approximating a common fixed points of a countable family of nonexpansive mappings and the weak convergence of our algorithm is established. In addition, we employ our algorithm for solving data classification problems. Moreover, our numerical experiments assert that our proposed algorithm provides more accuracy, precision, recall and F1-score than iPM-FBS, and FISTA.

Acknowledgments

The authors would like to thank the referees for their helpful comments and valuable suggestions. This work was supported by Chiang Mai University, Chiang Mai, Thailand.

References

[1] H. Adisak and S. Suantai, *A fast image restoration algorithm based on a fixed point and optimization method*, Mathematics **8** (2020): 378.

[2] R. P. Agarwal, D. O'Regan and D. R. Sahu, *Iterative construction of fixed point of nearly asymptotically nonexpansive mapping*, J. Nonlinear Convex Anal. **8** (2017), 61–79.

[3] K. Aoyama and Y. Kimura, *Strong convergence theorems for strongly nonexpansive sequences*, Appl. Math. Comput. **217** (2011), 7537–7545.

[4] K. Aoyama, Y. Kimura, W. Takahashi and M.Toyoda, *Approximation of common fixed points of a countable family of nonexpansive mappings in a Banach space*, Nonlinear Anal. **67** (2007), 2350–2360.

[5] K. Aoyama, F. Kohsaka and W. Takahashi, *Strong convergence theorems by shrinking and hybrid projection methods for relatively nonexpansive mappings in Banach spaces*, in: Proceedings of the 5th International Conference on Nonlinear Analysis and Convex Analysis; Yokohama Publishers : Yokohama, 2009, pp. 7–26.

[6] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hibert Spaces*, Springer, New York, 2011.

[7] A. Beck and M.Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAMJ. Imaging Sci. **2** (2009), 183–202.

[8] L. Bussaban, S. Suantai and A. Kaewkhao, *A parallel inertial S-iteration forward-backward algorithm for regression and classification problems*, Carpathian J. Math. **36** (2020), 35–44.

[9] G. Das and J. P. Debata, *Fixed point of quasi-non-expansive mappings*, Indian J. Pure Appl. Math. **17** (1968), 1263–1269.

[10] R. A. Fisher, *Iris, UCI machine learning repository*, 1988. `https://doi.org/10.24432/C56C76`.

[11] B. Halpern, *Fixed points of nonexpansive maps*, Bull. Amer. Math. Soc. **73** (1967), 957– 961.

[12] A. Hanjing and S. Suantai, *Novel algorithms with inertial techniques for solving constrained convex minimization problems and applications to image inpainting*, Mathematics **11** (2023): 1813.

[13] S. Ishikawa, *Fixed Point by a new iteration method*, Proc. Amer. Math. Soc. **44** (1974), 147–150.

[14] Kaggle, Diabetes dataset, Kaggle, 1990. Available from:`https://www.kaggle.com/datasets/mathchi/diabetes-data-set/data`.

[15] W. R. Mann, *Mean value methods in iteration*, Proc. Amer. Math. Soc. **4** (1953), 506–510.

[16] A. Moudafi and E. Al-Shemas, *Simulataneous iterative methods for split equality problemm*, Trans. Math. Program. Appl. **1** (2013), 1–11.

[17] K. Nakajo, K. Shimoji and W. Takahashi, *On strong convergence by the hybrid method for families of mappings in Hilbert spaces*, Nonlinear Anal. Theory Methods Appl. **71** (2009), 112–119.

[18] K. Nakajo, K. Shimoji and W. Takahashi, *Strong convergence theorems by the hybrid method for families of nonexpansive mappings in Hilbert spaces*, Taiwanese J. Math. **10** (2006), 339–360.

[19] Z. Opial, *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bull. Am. Math. Soc. **73** (1967), 591–597.

[20] P. Sae-Jia and S. Suantai, *A double inertial fixed point algorithm with linesearch and its application to machine learning for data classification*, CARPATHIANJ.MATH. 42, 09-26.

[21] B. Polyak, *Some methods of speeding up the convergence of iteration methods*, USSR Comput. Math. Phys. **4** (1964), 1–17.

[22] B. T. Polyak, *Introduction to Optimization; Optimization Software*, Publication Division, New York, 1987.

[23] W. Takahashi, *Iterative methods for approximation of fixed point and their applications*, J. Oper. Res. Soc. Jpn. **43** (2000), 87–108.

[24] W. Takahashi, Y. Takeuchi and R. Kubota, *Strong convergence theorems by hybrid methods for families of nonexpansive mappings in Hilbert Spaces*, J. Math. Anal. Appl. **341** (2008), 276–286.

[25] W. Takahashi and T. Tamura, *Convergence theorems for a pair of non-expansive mappings*, J. Convex Anal. **5** (1998), 45–58.

[26] K. Tan and H. K. Xu, *Approximating fixed points of nonexpansive mappings by the ishikawa iteration process*, J. Math. Anal. Appl. **178** (1993), 301–308.

[27] R. Tibshirani, *Regression shrinkage and selection via the lasso*, Journal of the Royal Statistical Society, Series B (Methodological) **58** (1996), 267–288.

[28] R. Wattanataweekul, K. Janngam and S. Suantai, *A novel two-step inertial viscosity algorithm for bilevel optimization problems applied to image recovery*, Mathematics **11** (2023): 3518.

[29] R. Wattanataweekul, K. Janngam and S. Suantai, *A novel double inertial viscosity algorithm for convex bilevel optimization problems applied to image restoration problems*, Optimization **74** (2024), 2635–2656.

[30] M. Verma and K. K. Shikla, *A new accelerated proximal gradient technique for regularized multitask learning framework*, Pattern Recogn. Lett. **95** (2017), 98–103.

Manuscript received February 24, 2025
revised October 2, 2025

J. Butwang
Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand
E-mail address: jirayut1butwang@gmail.com

S. Suantai
Excellence Center in Optimization and Computational Intelligence for Big Data Prediction, Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand
E-mail address: suthep.s@cmu.ac.th