

On multi-inertial extrapolations and forward-backward-forward algorithms

PAPATSARA INKRONG and PRASIT CHOLAMJIAK

ABSTRACT. In this work, we propose inclusion problems based on a novel class of forward-backward-forward algorithms. Our approach incorporates multi-inertial extrapolations and utilizes a self-adaptive technique to eliminate the need for explicitly selecting Lipschitz assumptions to enhance the speed convergence of the algorithm. We establish a weak convergence theorem under suitable assumptions. Furthermore, we conduct numerical tests on image deblurring as a practical application. The experimental results demonstrate that our algorithm surpasses some existing methods in the literature, which shows its superior performance and effectiveness.

1. INTRODUCTION

Let \mathcal{H} be a real Hilbert space. We assume $\mathcal{F} : \mathcal{H} \rightarrow \mathcal{H}$ and $\mathcal{G} : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ are monotone operators. The monotone inclusion problem is to seek $u^* \in \mathcal{H}$ such that

$$(1.1) \quad 0 \in \mathcal{F}u^* + \mathcal{G}u^*.$$

The set of all solutions to (1.1) is denoted as $(\mathcal{F} + \mathcal{G})^{-1}(0)$ and consistently assume that this set is nonempty. Problems (1.1) hold significant importance in resolving diverse challenges, including variational inequalities, minimization problems, machine learning, convex programming, and split feasibility problems. These problems have applications in various fields where monotone inclusion problems are fundamental for discovering solutions and tackling optimization challenges. (see [1, 4, 12, 13, 17, 22, 23, 28, 32, 34]).

A fascinating particular instance of problem (1.1) corresponds to a nonsmooth convex minimization problem. Let $f, g : \mathcal{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex and lower semicontinuous that f is differentiable on \mathcal{H} . The nonsmooth convex minimization can be stated as follows:

$$(1.2) \quad \min_{u \in \mathcal{H}} \{f(u) + g(u)\}.$$

By substituting $\mathcal{F} = \nabla f$ and $\mathcal{G} = \partial g$ into problem (1.1), where ∇f denotes the gradient of the function f , and ∂g denotes the subdifferential of g , then problem (1.1) reduces to problem (1.2).

The forward-backward algorithm is the most widely recognized approach for solving (1.1). The classical one was initially suggested by Passty [28] and later extended by Lions and Mercier [20]. It is described as follows:

$$(1.3) \quad u_{n+1} = (I + \alpha_n \mathcal{G})^{-1}(I - \alpha_n \mathcal{F})(u_n), \quad n \geq 1,$$

where $\alpha_n > 0$, $u_1 \in \mathcal{H}$, \mathcal{F} is L -Lipschitz and \mathcal{G} is maximally monotone. If $\alpha_n \in (0, \frac{2}{L})$, then (1.3) exhibits weak convergence to a solution of (1.1).

Received: 04.09.2023 . In revised form: 01.03.2024. Accepted: 08.03.2024

2020 *Mathematics Subject Classification.* 47H09, 47H10, 65K15, 65Y05, 68W10.

Key words and phrases. Inclusion problem, Forward-backward-forward algorithm, Multi-inertial extrapolations, Weak convergence.

Corresponding author: Prasit Cholamjiak; prasit.ch@up.ac.th

Tseng [34] introduced a modification to (1.1) by integrating a two-step method within the forward-backward algorithm. It can be represented in the following manner:

$$\begin{cases} u_1 \in \mathcal{H}, \\ w_n = (I + \alpha_n \mathcal{G})^{-1}(I - \alpha_n \mathcal{F})(u_n), \\ u_{n+1} = w_n - \alpha_n(\mathcal{F}w_n - \mathcal{F}u_n), \end{cases}$$

where $\alpha_n = \alpha l^{t_n}$ and t_n is the smallest non-negative integer satisfying

$$\alpha_n \|\mathcal{F}(u_{n+1}) - \mathcal{F}(u_n)\| \leq \mu \|u_{n+1} - u_n\|,$$

where $\mu, l \in (0, 1)$ and $\alpha > 0$. It has been proven that $\{u_n\}$ converges weakly to a point in $(\mathcal{F} + \mathcal{G})^{-1}(0)$. This is often called the forward-backward-forward algorithm.

The origin of inertial-type algorithms can be traced back to the heavy ball method, an implicit discretization of a second-order time dynamical system [2, 29]. These algorithms have recently gained increasing attention from researchers due to their ability to speed up the convergence of algorithms (see [14, 16]). A vital characteristic of these algorithms is that the next iterate is determined by using the previous two iterates. In the context of solving the structured monotone inclusion problem (1.1), several researchers have introduced inertial operator splitting methods. Notable examples include the inertial forward-backward methods [3, 23], the inertial Douglas-Rachford methods [8], the inertial proximal ADMM [10], and the inertial forward-backward-forward primal-dual algorithm [7]. These methods aim to effectively address the problem by incorporating inertia and leveraging the benefits of operator splitting techniques. In 2009, Beck and Teboulle [6] introduced a highly efficient iterative shrinkage-thresholding algorithm called FISTA (Fast Iterative Shrinkage-Thresholding Algorithm). They demonstrated the convergence rate and applied it to image processing.

As reported in [19], it was observed that the inclusion of more than two points, namely u_n and u_{n-1} , has the potential to improve acceleration. This enhancement can be demonstrated by employing a two-step inertial extrapolation as follows:

$$z_n = u_n + \theta(u_n - u_{n-1}) + \lambda(u_{n-1} - u_{n-2}),$$

with $\theta > 0$ and $\lambda > 0$ can bring about acceleration. In reference [31], the limitations of one-step inertial acceleration in ADMM were investigated, leading to the introduction of an alternative approach known as adaptive acceleration for ADMM.

Ortega and Rheinboldt [26] introduced a comprehensive iterative process as follows:

$$(1.4) \quad u_{n+1} = \theta_n(u_n, u_{n-1}, \dots, u_{n-k+1}).$$

The given iterative process (1.4) is called as the k -step method, where $n \geq 1$ and $\theta_n(\cdot)$ represents the extrapolation onto $x_n, x_{n-1}, \dots, x_{n-k+1}$. Polyak [30] also explored the potential of multi-step inertial methods in enhancing optimization speed, even though no established convergence or rate outcomes were presented in [30].

Liang [19] introduced two multi-step inertial methods. The first one is the multi-step inertial forward-backward method and another one is the multi-step inertial primal-dual method. Moreover, Liang introduced the multi-step inertial Douglas-Rachford method, although its convergence has yet to be investigated. Similar to Polyak's k -step method, the update x_{n+1} in Liang's methods using at most $k + 1$ iterations $x_n, x_{n-1}, \dots, x_{n-k+1}$. Recent studies have explored multi-step inertial methods, and specific results have been presented in [11, 15, 35, 36].

Motivated by previous research, this study introduces a novel class of forward-backward-forward algorithms designed explicitly for monotone inclusion. The proposed algorithms

achieve weak convergence through a self-adaptive technique that updates α_n . Additionally, we incorporate multiple inertial extrapolations into our algorithm to expedite convergence.

In this research, Section 2 contains the basic concepts. For the proof of our algorithm, we show it in Section 3. Section 4 demonstrates numerical tests on image deblurring to show the algorithm’s performance.

2. PRELIMINARIES

Given that $\langle \cdot, \cdot \rangle$ represents the usual inner product and $\| \cdot \|$ denotes the norm defined on \mathcal{H} , the graph of $\mathcal{G} : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ can be defined by [5]:

$$\text{Graph}(\mathcal{G}) = \{(p, q) \in \mathcal{H} \times \mathcal{H} : p \in \text{dom}(\mathcal{G}), q \in \mathcal{G}(p)\}.$$

The operator \mathcal{G} is considered to be monotone [4] if

$$\langle p - q, r - s \rangle \geq 0, \text{ for all } (p, r), (q, s) \in \text{Graph}(\mathcal{G}).$$

The operator \mathcal{G} is said to be maximally monotone if it is monotone, and its graph is not strictly contained within the graph of any other monotone operator. An operator $\mathcal{F} : \mathcal{H} \rightarrow \mathcal{H}$ is L -Lipschitz if

$$\|\mathcal{F}(u) - \mathcal{F}(z)\| \leq L\|u - z\|, \text{ for all } u, z \in \mathcal{H}.$$

For $\alpha > 0$, the *resolvent* operator of $\mathcal{G} : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ can be defined as [24]:

$$J_{\alpha}^{\mathcal{G}}(u) = (I + \alpha\mathcal{G})^{-1}(u), \text{ for all } u \in \mathcal{H}.$$

It is known that $J_{\alpha}^{\mathcal{G}}$ is firmly nonexpansive, single-valued and $\text{Dom}(J_{\alpha}^{\mathcal{G}}) = \mathcal{H}$.

For $\alpha > 0$, let $T_{\alpha} = (I + \alpha\mathcal{G})^{-1}(I - \alpha\mathcal{F})$. It is also known that [5]

$$\text{Fix}(T_{\alpha}) = (\mathcal{F} + \mathcal{G})^{-1}(0),$$

where $\text{Fix}(T) = \{u \in \mathcal{H} : u = Tu\}$.

Lemma 2.1. [33] *We know that*

(i) $\|cu + (1 - c)w\|^2 = c\|u\|^2 + (1 - c)\|w\|^2 - c(1 - c)\|u - w\|^2$, for all $c \in [0, 1]$ and for all $u, w \in \mathcal{H}$;

(ii) $\|u \pm w\|^2 = \|u\|^2 \pm 2\langle u, w \rangle + \|w\|^2$, for all $u, w \in \mathcal{H}$.

Lemma 2.2. [9] *Assuming $\mathcal{F} : \mathcal{H} \rightarrow \mathcal{H}$ is Lipschitz continuous, and $\mathcal{G} : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ is a maximally monotone operator, it follows that the operator $\mathcal{F} + \mathcal{G}$ is also maximally monotone.*

Lemma 2.3. [27] *Let $\{b_n\}$, $\{\varphi_n\}$ and $\{a_n\}$ be nonnegative sequences with the following condition*

$$b_{n+1} = (1 + a_n)b_n + \varphi_n, \quad n \geq 1.$$

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

Lemma 2.4. [25] *Let Ω be a subset of \mathcal{H} and $\{u_n\} \subseteq \mathcal{H}$ such that:*

(i) *for every $u \in \Omega$, $\lim_{n \rightarrow \infty} \|u_n - u\|$ exists;*

(ii) *each weak cluster point of $\{u_n\}$ is in Ω .*

Then $\{u_n\}$ converges weakly to a point in Ω .

3. ALGORITHM AND CONVERGENCE RESULTS

Next, we propose a new type of forward-backward-forward algorithm for solving (1.1). Given $\mathcal{F} : \mathcal{H} \rightarrow \mathcal{H}$ be monotone and L -Lipschitz, and $\mathcal{G} : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ be maximally monotone. We denote the set $(\mathcal{F} + \mathcal{G})^{-1}(0)$ as Ω which is nonempty. We begin by presenting the following crucial lemma which is motivated by [18].

Lemma 3.5. *Let $k \in \{1, 2, \dots, b\}$ for some $b \in \mathbb{N}$. Let $\Psi_{1-b}, \Psi_{2-b}, \dots, \Psi_0$ be non-negative real numbers. Let $\{\Psi_n\}_{n=1}^{\infty}$ and $\{\theta_{k,n}\}_{n=1}^{\infty}$ be non-negative real sequences. If*

$$\Psi_{n+1} \leq \Psi_n + \sum_{n=1}^b (\Psi_{n-k+1} + \Psi_{n-k})\theta_{k,n}, \quad n \in \mathbb{N},$$

then

$$\Psi_{n+1} \leq K \cdot \prod_{j=1}^n (1 + 2\theta_{1,j} + 2\theta_{2,j} + \dots + 2\theta_{b-1,j} + 2\theta_{b,j}), \quad n \in \mathbb{N},$$

where $K = \max\{\Psi_{1-b}, \Psi_{2-b}, \dots, \Psi_0, \Psi_1\}$. Moreover, if for each k , $\sum_{n=1}^{\infty} \theta_{k,n} < +\infty$ then $\{\Psi_n\}$ is bounded.

Proof. Let $K = \max\{\Psi_{1-b}, \Psi_{2-b}, \dots, \Psi_0, \Psi_1\}$. We show that it is true for $n = 1$.

$$\begin{aligned} \Psi_2 &\leq (1 + \theta_{1,1})\Psi_1 + (\theta_{1,1} + \theta_{2,1})\Psi_0 + \dots + (\theta_{b-1,1} + \theta_{b,1})\Psi_{2-b} + \theta_{b,1}\Psi_{1-b} \\ &\leq (1 + 2\theta_{1,1} + 2\theta_{2,1} + \dots + 2\theta_{b-1,1} + 2\theta_{b,1})K. \end{aligned}$$

Assume that it is true for $n = m$ where $m \in \mathbb{N}$. Then we have

$$\begin{aligned} \Psi_{m+1} &\leq K(1 + 2\theta_{1,1} + 2\theta_{2,1} + \dots + 2\theta_{b-1,1} + 2\theta_{b,1})(1 + 2\theta_{1,2} + 2\theta_{2,2} + \dots \\ &\quad + 2\theta_{b-1,2} + 2\theta_{b,2}) \cdots (1 + 2\theta_{1,m} + 2\theta_{2,m} + \dots + 2\theta_{b-1,m} + 2\theta_{b,m}) \\ &= \mathcal{M}. \end{aligned}$$

Considering $n = m + 1$, we obtain

$$\begin{aligned} \Psi_{m+2} &\leq (1 + \theta_{1,m+1})\Psi_{m+1} + (\theta_{1,m+1} + \beta_{2,m+1})\Psi_m + \dots + (\theta_{b-1,m+1} + \theta_{b,m+1})\Psi_{m+2-b} \\ &\quad + \theta_{b,m+1}\Psi_{m+1-b} \\ &\leq (1 + \theta_{1,m+1})\mathcal{M} + (\theta_{1,m+1} + \theta_{2,m+1})\Psi_m + \dots + (\theta_{b-1,m+1} + \theta_{b,m+1})\Psi_{m+2-b} \\ &\quad + \theta_{b,m+1}\Psi_{m+1-b} \\ &\leq (1 + \theta_{1,m+1})\mathcal{M} + (\theta_{1,m+1} + \theta_{2,m+1})\mathcal{M} + \dots + (\theta_{b-1,m+1} + \theta_{b,m+1})\mathcal{M} \\ &\quad + \theta_{b,m+1}\mathcal{M} \\ &= (1 + 2\theta_{1,m+1} + 2\theta_{2,m+1} + \dots + 2\theta_{b-1,m+1} + 2\theta_{b,m+1})\mathcal{M} \\ &= K(1 + 2\theta_{1,1} + 2\theta_{2,1} + \dots + 2\theta_{b-1,1} + 2\theta_{b,1})(1 + 2\theta_{1,2} + 2\theta_{2,2} + \dots \\ &\quad + 2\theta_{b-1,2} + 2\theta_{b,2}) \cdots (1 + 2\theta_{1,m} + 2\theta_{2,m} + \dots + 2\theta_{b-1,m} + 2\theta_{b,m})(1 + \\ &\quad + 2\theta_{1,m+1} + 2\theta_{2,m+1} + \dots + 2\theta_{b-1,m+1} + 2\theta_{b,m+1}). \end{aligned}$$

By mathematical induction, we can conclude that

$$\Psi_{n+1} \leq K \cdot \prod_{j=1}^n (1 + 2\theta_{1,j} + 2\theta_{2,j} + \dots + 2\theta_{b-1,j} + 2\theta_{b,j}), \quad n \in \mathbb{N}.$$

Moreover, the sequence $\{\Psi_n\}$ is bounded if $\sum_{n=1}^{\infty} \theta_{k,n} < +\infty$ for all $k = 1, 2, \dots, b$. □

Algorithm 3.1. *Multi-inertial forward-backward-forward algorithm*

Initialization : Let $\alpha_1 > 0$, $\mu \in (0, 1)$ be arbitrarily. Fix $b \in \mathbb{N}$ and for each $k \in \{1, 2, \dots, b\}$, let $\{\theta_{k,n}\}_{n=1}^\infty$ be a real non-negative sequence. Let $u_{1-b}, u_{2-b}, \dots, u_0, u_1 \in \mathcal{H}$. Let $\{\beta_n\}$ be a non-negative sequence.

Step 1. Compute

$$z_n = u_n + \sum_{k=1}^b \theta_{k,n}(u_{n-k+1} - u_{n-k}),$$

$$w_n = (I + \alpha_n \mathcal{G})^{-1}(I - \alpha_n \mathcal{F})z_n.$$

If $w_n = z_n$ then stop and $w_n \in \Omega$. Otherwise, go to **Step 2**.

Step 2. Compute

$$u_{n+1} = w_n - \alpha_n(\mathcal{F}w_n - \mathcal{F}z_n),$$

where

$$\alpha_{n+1} = \begin{cases} \min \left\{ \frac{\mu \|z_n - w_n\|}{\|\mathcal{F}z_n - \mathcal{F}w_n\|}, \alpha_n + \beta_n \right\} & \text{if } \|\mathcal{F}z_n - \mathcal{F}w_n\| \neq 0; \\ \alpha_n + \beta_n & \text{otherwise.} \end{cases}$$

Then set $n = n + 1$ and go back to **Step 1**.

Remark 3.1. We define Algorithm 3.1 based on Tseng’s algorithm and multi-inertial extrapolations. Moreover, the step size using a self-adaptive technique can eliminate the need to select L -Lipschitz constant and enhance the convergence speed of the proposed algorithm.

Lemma 3.6. [21] Let $\{\alpha_n\}$ be generated by Algorithm 3.1. If $\sum_{n=1}^\infty \beta_n < +\infty$, then

$$\lim_{n \rightarrow \infty} \alpha_n = \alpha \in \left[\min \left\{ \alpha_1, \frac{\mu}{L} \right\}, \alpha_1 + \beta \right]$$

where $\beta = \sum_{n=1}^\infty \beta_n$ and

$$(3.5) \quad \|\mathcal{F}z_n - \mathcal{F}w_n\| \leq \frac{\mu}{\alpha_{n+1}} \|z_n - w_n\|.$$

Theorem 3.1. Let $\{u_n\}$ be defined by Algorithm 3.1. For each $k \in \{1, 2, \dots, b\}$, if $\sum_{n=1}^\infty \theta_{k,n} < +\infty$

and $\sum_{n=1}^\infty \beta_n < +\infty$, then

$$(i) \text{ for each } x^* \in \Omega, \text{ we have } \|u_{n+1} - x^*\| \leq K \cdot \prod_{j=1}^n (1 + 2\theta_{1,j} + 2\theta_{2,j} + \dots + 2\theta_{b-1,j} + 2\theta_{b,j}),$$

where $K = \max\{\|u_{n_0-b} - x^*\|, \|u_{n_0-b+1} - x^*\|, \dots, \|u_{n_0-1} - x^*\|, \|u_{n_0} - x^*\|\}$ for some $n_0 \in \mathbb{N}$.

(ii) $\{u_n\}$ converges weakly to a point in Ω .

Proof. (i) Let $x^* \in \Omega$. Putting $w_n = (I + \alpha_n \mathcal{G})^{-1}(I - \alpha_n \mathcal{F})z_n$, we can deduce that $(I - \alpha_n \mathcal{F})z_n \in (I + \alpha_n \mathcal{G})w_n$. Since \mathcal{G} is a maximal monotone operator, there is $v_n \in \mathcal{G}w_n$ with

$$(I - \alpha_n \mathcal{F})z_n = w_n + \alpha_n v_n.$$

Therefore

$$(3.6) \quad v_n = \frac{1}{\alpha_n}(z_n - w_n - \alpha_n \mathcal{F}z_n).$$

In addition, we have $0 \in (\mathcal{F} + \mathcal{G})x^*$ and $\mathcal{F}w_n + v_n \in (\mathcal{F} + \mathcal{G})w_n$. Since $\mathcal{F} + \mathcal{G}$ is monotone, we obtain

$$(3.7) \quad \langle \mathcal{F}w_n + v_n, w_n - x^* \rangle \geq 0.$$

By combining equations (3.6) and (3.7), we have

$$(3.8) \quad \frac{1}{\alpha_n} \langle z_n - w_n - \alpha_n \mathcal{F}z_n + \alpha_n \mathcal{F}w_n, w_n - x^* \rangle \geq 0.$$

It can be observed that

$$\begin{aligned}
 \|u_{n+1} - x^*\|^2 &= \|w_n - \alpha_n(\mathcal{F}w_n - \mathcal{F}z_n) - x^*\|^2 \\
 &= \|w_n - x^*\|^2 + \alpha_n^2 \|\mathcal{F}w_n - \mathcal{F}z_n\|^2 - 2\alpha_n \langle w_n - x^*, \mathcal{F}w_n - \mathcal{F}z_n \rangle \\
 &= \|z_n - x^*\|^2 + \|z_n - w_n\|^2 + 2\langle w_n - z_n, z_n - x^* \rangle + \alpha_n^2 \|\mathcal{F}w_n - \mathcal{F}z_n\|^2 \\
 &\quad - 2\alpha_n \langle w_n - x^*, \mathcal{F}w_n - \mathcal{F}z_n \rangle \\
 &= \|z_n - x^*\|^2 + \|z_n - w_n\|^2 - 2\langle w_n - z_n, w_n - z_n \rangle + 2\langle w_n - z_n, w_n - x^* \rangle \\
 &\quad + \alpha_n^2 \|\mathcal{F}w_n - \mathcal{F}z_n\|^2 - 2\alpha_n \langle w_n - x^*, \mathcal{F}w_n - \mathcal{F}z_n \rangle \\
 &= \|z_n - x^*\|^2 - \|z_n - w_n\|^2 + 2\langle w_n - z_n, w_n - x^* \rangle \\
 &\quad + \alpha_n^2 \|\mathcal{F}w_n - \mathcal{F}z_n\|^2 - 2\alpha_n \langle w_n - x^*, \mathcal{F}w_n - \mathcal{F}z_n \rangle \\
 &= \|z_n - x^*\|^2 - \|z_n - w_n\|^2 - 2\langle z_n - w_n - \alpha_n(\mathcal{F}z_n - \mathcal{F}w_n), w_n - x^* \rangle \\
 (3.9) \quad &+ \alpha_n^2 \|\mathcal{F}w_n - \mathcal{F}z_n\|^2.
 \end{aligned}$$

Using (3.5), (3.8) and (3.9), we obtain

$$\begin{aligned}
 \|u_{n+1} - x^*\|^2 &\leq \|z_n - x^*\|^2 - \|z_n - w_n\|^2 + \left(\frac{\mu^2 \alpha_n^2}{\alpha_{n+1}^2}\right) \|w_n - z_n\|^2 \\
 (3.10) \quad &= \|z_n - x^*\|^2 - \left(1 - \frac{\mu^2 \alpha_n^2}{\alpha_{n+1}^2}\right) \|w_n - z_n\|^2.
 \end{aligned}$$

Moreover, we have

$$\begin{aligned}
 \|z_n - x^*\| &= \left\| u_n + \sum_{k=1}^b \theta_{k,n}(u_{n-k+1} - u_{n-k}) - x^* \right\| \\
 (3.11) \quad &\leq \|u_n - x^*\| + \sum_{k=1}^b \theta_{k,n} \|u_{n-k+1} - u_{n-k}\|.
 \end{aligned}$$

Since $\lim_{n \rightarrow \infty} \left(1 - \mu^2 \frac{\alpha_n^2}{\alpha_{n+1}^2}\right) = 1 - \mu^2 > 0$, we can choose $n_0 \in \mathbb{N}$ such that $1 - \mu^2 \frac{\alpha_n^2}{\alpha_{n+1}^2} > 0$ for all $n \geq n_0$.

By using (3.10) and (3.11), we can deduce

$$\begin{aligned}
 \|u_{n+1} - x^*\| &\leq \|z_n - x^*\| \\
 (3.12) \quad &\leq \|u_n - x^*\| + \sum_{k=1}^b \theta_{k,n} \|u_{n-k+1} - u_{n-k}\| \\
 &\leq \|u_n - x^*\| + \sum_{k=1}^b \theta_{k,n} [\|u_{n-k+1} - x^*\| + \|u_{n-k} - x^*\|].
 \end{aligned}$$

By Lemma 3.5, we obtain

$$(3.13) \quad \|u_{n+1} - x^*\| \leq K \cdot \prod_{j=1}^n (1 + 2\theta_{1,j} + 2\theta_{2,j} + \dots + 2\theta_{b-1,j} + 2\theta_{b,j}), \quad n \geq n_0.$$

where $K = \max\{\|u_{n_0-b} - x^*\|, \|u_{n_0-b+1} - x^*\|, \dots, \|u_{n_0-1} - x^*\|, \|u_{n_0} - x^*\|\}$.

(ii) We next establish that $\{u_n\}$ converges weakly to a point in Ω . Because $\sum_{n=1}^{\infty} \theta_{k,n} < +\infty$, the sequence $\{u_n\}$ is bounded according to Lemma 3.5 and (3.13). Consequently, we have $\sum_{n=1}^{\infty} \theta_{1,n} \|u_n - u_{n-1}\| < +\infty$, $\sum_{n=1}^{\infty} \theta_{2,n} \|u_{n-1} - u_{n-2}\| < +\infty$, ..., $\sum_{n=1}^{\infty} \theta_{b-1,n} \|u_{n-b+2} - u_{n-b+1}\| < +\infty$, and $\sum_{n=1}^{\infty} \theta_{b,n} \|u_{n-b+1} - u_{n-b}\| < +\infty$. Applying Lemma 2.3 in (3.12), we conclude that $\lim_{n \rightarrow \infty} \|u_n - x^*\|$ exists.

Now, we see that

$$\begin{aligned}
 \|z_n - x^*\|^2 &= \left\| u_n + \sum_{k=1}^b \theta_{k,n} (u_{n-k+1} - u_{n-k}) - x^* \right\|^2 \\
 &= \|u_n - x^*\|^2 + \left\| \sum_{k=1}^b \theta_{k,n} (u_{n-k+1} - u_{n-k}) \right\|^2 \\
 &\quad + 2 \left\langle u_n - x^*, \sum_{k=1}^b \theta_{k,n} (u_{n-k+1} - u_{n-k}) \right\rangle \\
 &\leq \|u_n - x^*\|^2 + \left(\sum_{k=1}^b \theta_{k,n} \|u_{n-k+1} - u_{n-k}\| \right)^2 \\
 (3.14) \quad &\quad + 2 \sum_{k=1}^b \theta_{k,n} \|u_n - x^*\| \|u_{n-k+1} - u_{n-k}\|.
 \end{aligned}$$

Replacing (3.14) into (3.10), we get

$$\begin{aligned}
 \|u_{n+1} - x^*\|^2 &\leq \|u_n - x^*\|^2 + \left(\sum_{k=1}^b \theta_{k,n} \|u_{n-k+1} - u_{n-k}\| \right)^2 \\
 &\quad + 2 \sum_{k=1}^b \theta_{k,n} \|u_n - x^*\| \|u_{n-k+1} - u_{n-k}\| \\
 (3.15) \quad &\quad - \left(1 - \frac{\mu^2 \alpha_n^2}{\alpha_{n+1}^2} \right) \|z_n - w_n\|^2.
 \end{aligned}$$

This gives

$$(3.16) \quad \left(1 - \frac{\mu^2 \alpha_n^2}{\alpha_{n+1}^2}\right) \|z_n - w_n\|^2 \leq [\|u_n - x^*\|^2 - \|u_{n+1} - x^*\|^2] + \left(\sum_{k=1}^b \theta_{k,n} \|u_{n-k+1} - u_{n-k}\|\right)^2 + 2 \sum_{k=1}^b \theta_{k,n} \|u_n - x^*\| \|u_{n-k+1} - u_{n-k}\|.$$

Since $\lim_{n \rightarrow \infty} \|u_n - x^*\|$ exists, and $\sum_{n=1}^{\infty} \theta_{k,n} < +\infty$, we can deduce from (3.16) that

$$(3.17) \quad \lim_{n \rightarrow \infty} \|z_n - w_n\| = 0.$$

Observe that

$$(3.18) \quad \begin{aligned} \|z_n - u_n\| &= \left\| u_n + \sum_{k=1}^b \theta_{k,n} (u_{n-k+1} - u_{n-k}) - u_n \right\| \\ &\leq \sum_{k=1}^b \theta_{k,n} \|u_{n-k+1} - u_{n-k}\| \\ &\rightarrow 0. \end{aligned}$$

From (3.17) and (3.18), we have

$$\begin{aligned} \|u_n - w_n\| &= \|u_n - z_n\| + \|z_n - w_n\| \\ &\rightarrow 0. \end{aligned}$$

If $(s, t) \in \text{Graph}(\mathcal{F} + \mathcal{G})$, then $t - \mathcal{F}s \in \mathcal{G}s$. For $n_k \subset n$, we have

$$w_{n_k} = (I + \alpha_{n_k} \mathcal{G})^{-1} (I - \alpha_{n_k} \mathcal{F}) z_{n_k},$$

which implies

$$(I - \alpha_{n_k} \mathcal{F}) z_{n_k} \in (I + \alpha_{n_k} \mathcal{G}) w_{n_k}.$$

This shows that

$$\frac{1}{\alpha_{n_k}} (z_{n_k} - w_{n_k} - \alpha_{n_k} \mathcal{F} z_{n_k}) \in \mathcal{G} w_{n_k}.$$

By the monotonicity of \mathcal{G} , we get

$$\left\langle s - w_{n_k}, t - \mathcal{F}s - \frac{1}{\alpha_{n_k}} (z_{n_k} - w_{n_k} - \alpha_{n_k} \mathcal{F} z_{n_k}) \right\rangle \geq 0.$$

It is therefore

$$(3.19) \quad \begin{aligned} \langle s - w_{n_k}, t \rangle &\geq \left\langle s - w_{n_k}, \mathcal{F}s + \frac{1}{\alpha_{n_k}} (z_{n_k} - w_{n_k} - \alpha_{n_k} \mathcal{F} z_{n_k}) \right\rangle \\ &= \langle s - w_{n_k}, \mathcal{F}s - \mathcal{F}z_{n_k} \rangle + \left\langle s - w_{n_k}, \frac{1}{\alpha_{n_k}} (z_{n_k} - w_{n_k}) \right\rangle \\ &= \langle s - w_{n_k}, \mathcal{F}s - \mathcal{F}w_{n_k} \rangle + \langle s - w_{n_k}, \mathcal{F}w_{n_k} - \mathcal{F}z_{n_k} \rangle \\ &\quad + \left\langle s - w_{n_k}, \frac{1}{\alpha_{n_k}} (z_{n_k} - w_{n_k}) \right\rangle \\ &\geq \langle s - w_{n_k}, \mathcal{F}w_{n_k} - \mathcal{F}z_{n_k} \rangle + \left\langle s - w_{n_k}, \frac{1}{\alpha_{n_k}} (z_{n_k} - w_{n_k}) \right\rangle. \end{aligned}$$

From (3.17) and the L -Lipschitz continuity of \mathcal{F} , we obtain $\lim_{k \rightarrow \infty} \|\mathcal{F}z_{n_k} - \mathcal{F}w_{n_k}\| = 0$. Since $\{u_n\}$ is bounded, there is a subsequence $\{u_{n_i}\}$ of $\{u_n\}$ and $\hat{z} \in \mathcal{H}$ such that $u_{n_i} \rightharpoonup \hat{z}$. Moreover, $y_{n_i} \rightharpoonup \hat{z}$. Therefore, from (3.19), we have

$$\langle s - \hat{z}, t \rangle \geq 0.$$

Hence, by the maximal monotonicity of $\mathcal{F} + \mathcal{G}$, we have $0 \in (\mathcal{F} + \mathcal{G})(\hat{z})$, i.e., $\hat{z} \in (\mathcal{F} + \mathcal{G})^{-1}(0) = \Omega$. By Lemma 2.4, it is concluded $\{u_n\}$ converges weakly to a point in Ω . We finish the proof. \square

Our Contributions

We apply the forward-backward-forward algorithm with multi-inertial extrapolations to solve the monotone inclusion problem (1.1). The convergence need a simple condition that $\sum_{n=1}^{\infty} \theta_{k,n} < +\infty$. This is quite different from those in [15].

4. NUMERICAL EXPERIMENTS

Next, we present numerical results to evaluate the performance of Algorithm 3.1 and to compare it with DHLCR Algorithm as suggested in ([15], Algorithm 43, P. 816), ZDC Algorithm as proposed in ([35], Algorithm 3.1 with self-adaptive), ZDCL Algorithm ([35], Algorithm 3.1 with L -Lipschitz), and ZZT Algorithm as presented in ([36], Algorithm 1). The experiments were conducted by MATLAB R2022A on an HP Laptop with an Intel(R) Core(TM) i7-1165G7 processor and 16.00 GB RAM.

Next, we address the problem of image recovery, which can be formulated as follows:

$$(4.20) \quad b = Dx + v,$$

where an original image is denoted as $x \in \mathbb{R}^{M \times 1}$, a degraded image is represented by $b \in \mathbb{R}^{M \times 1}$, a noise term is given by $v \in \mathbb{R}^{M \times 1}$, and $D \in \mathbb{R}^{M \times M}$ is a blurring matrix. It is worth noting that the problem (4.20) can be formulated equivalently as the subsequent convex minimization model:

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \|b - Dx\|_2^2 + \rho \|x\|_1$$

This example considers the case where $\rho > 0$. We define $\mathcal{F} = D^t(Dx - b)$ and $\mathcal{G} = \partial(\rho \|x\|_1)$ as the operators involved. To evaluate the quality of the restored images, we utilize two values: the peak signal-to-noise ratio (PSNR) and the structural similarity index measure (SSIM). These formulas are defined by:

$$\text{PSNR} := 20 \log_{10} \left(\frac{255^2}{\|x^n - x\|_2^2} \right),$$

and

$$\text{SSIM} := \frac{(2\theta_x \theta_{x_r} + c_1)(2\sigma_{xx_r} + c_2)}{(\theta_x^2 + \theta_{x_r}^2 + c_1)(\sigma_x^2 + \sigma_{x_r}^2 + c_2)}.$$

In these equations, x represents the original image, x_r represents the recovery image, θ_x and θ_{x_r} are the mean values of the original and restored images respectively, σ_x^2 and $\sigma_{x_r}^2$ are the variances, $\sigma_{xx_r}^2$ is the covariance of the two images, c_1 and c_2 are constants defined as $c_1 = (0.01L)^2$ and $c_2 = (0.03L)^2$, and L is the dynamic range of pixel values.

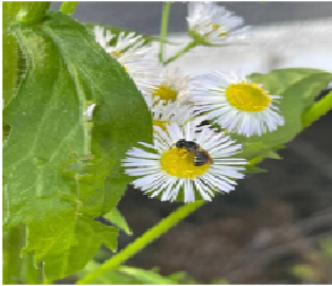
In this example, the initial points u_{-2} , u_{-1} , u_0 , and u_1 are set to be the blurred image. All parameters are selected as follows:

- Algorithm 3.1 if $1 \leq n \leq 500$, we set $s_1 = 1$, $\theta_{1,n} = \frac{s_n - 1}{s_{n+1}}$ where $s_{n+1} = \frac{1 + \sqrt{1 + 4s_n^2}}{2}$ and set $\theta_{1,n} = \frac{1}{(5n+1)^2}$ if $n > 500$, $\theta_{2,n} = \frac{1}{(10n^3 + n)^5}$, $\theta_{3,n} = \frac{1}{(5n+1)^2}$, $\beta_n = \frac{1}{(n+1)^2}$, $\lambda_1 = 2.0$ and $\mu = 0.95$;
- DHLCR: $a_{1,n} = 1.1$, $a_{2,n} = 1.8$, $a_{3,n} = 1.5$, $b_{1,n} = 0.1$, $b_{2,n} = 0.8$, $b_{3,n} = 0.5$, $\lambda_n = 0.05$ and $\gamma = 1.2$;
- ZDC: $a_{1,k} = 0.1$, $a_{2,k} = 0.6$, $a_{3,k} = 0.2$, $\eta = 0.88$, $\sigma = 0.35$, $\mu = 0.5$ and $\gamma = 1.5$;
- ZDCL: $a_{1,k} = 0.1$, $a_{2,k} = 0.6$, $a_{3,k} = 0.2$, $\lambda_k = 0.9$, $\mu = 0.5$ and $\gamma = 1.5$;
- ZZT: $a_{1,k} = 0.1$, $a_{2,k} = 0.6$, $a_{3,k} = 0.2$ and $\gamma_k = 0.1$.

Firstly, we will use three types of blur to reduce the quality of the original image (Bee, Dog and Cat, and View images) in Figure 1. The details are as follows:

- Out-of-focus blur (disk): radius 6;
- Motion blur: a motion length 10 and an angle 180;
- Gaussian blur: filter size $[5, 5]$ with standard deviation 5.

In the following sections, we will show the restored images at 1000 iterations for each algorithm, as depicted in Figure 2. Additionally, we will provide performance evaluations for each algorithm's restoration of the original image, which is expressed as PSNR and SSIM values to determine the absolute error between the pixels and to measure the differences between the properties (luminance, contrast, and structure) of the pixels of the recovered images, respectively. The reports are shown in Table 1 and Figure 3.



(A) Bee



(B) Dog and Cat



(C) View

FIGURE 1. The original images

TABLE 1. The numerical PSNR and SSIM values of each algorithm in different blurs

	Blur	Bee (640*853)		Dog and Cat (820*1093)		View (1280*960)	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Algorithm 3.1	Disk	37.2126	0.9474	35.4521	0.9270	38.7077	0.9407
	Motion	34.7877	0.9251	38.5623	0.9644	40.5999	0.9771
	Gaussian	40.2976	0.9743	37.9386	0.9559	41.7718	0.9715
DHLCR	Disk	26.8375	0.7887	24.5920	0.7322	28.2109	0.8318
	Motion	30.9980	0.8923	28.7540	0.8769	33.6119	0.9490
	Gaussian	31.0900	0.9208	28.9823	0.8880	33.3727	0.9362
ZDC	Disk	31.8220	0.9054	29.8296	0.8655	33.0769	0.9122
	Motion	34.1055	0.9221	34.7845	0.9520	38.2283	0.9729
	Gaussian	35.4449	0.9601	33.5516	0.9381	37.5319	0.9605
ZDCL	Disk	33.9348	0.9304	32.0118	0.8999	35.3048	0.9337
	Motion	34.5973	0.9251	36.3975	0.9592	39.6169	0.9771
	Gaussian	37.1349	0.9688	35.0776	0.9470	39.0032	0.9645
ZZT	Disk	28.8385	0.8486	26.7033	0.7980	30.2904	0.8741
	Motion	32.7644	0.9104	31.3271	0.9148	35.9033	0.9633
	Gaussian	32.9683	0.9424	31.0054	0.9155	35.3439	0.9480

From Table 1, we see the image recovery performance of our algorithm using different blurs. When measured with PSNR values, it was the highest among all algorithms. The View image has the highest PSNR value of 41.7718 if using Gaussian blur. If using motion and disk blur, the values are 40.5999 and 38.7077, respectively. Moreover, we also found that the SSIM values of our algorithm are higher than those of other algorithms. The recovered image with the highest SSIM value is the View image.

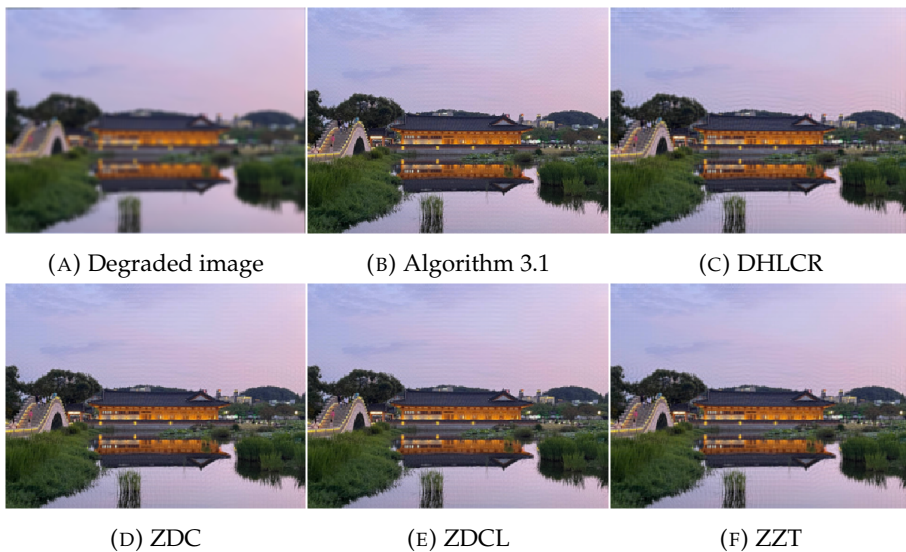


FIGURE 2. The View images blurred by gaussian and recovered by each algorithm

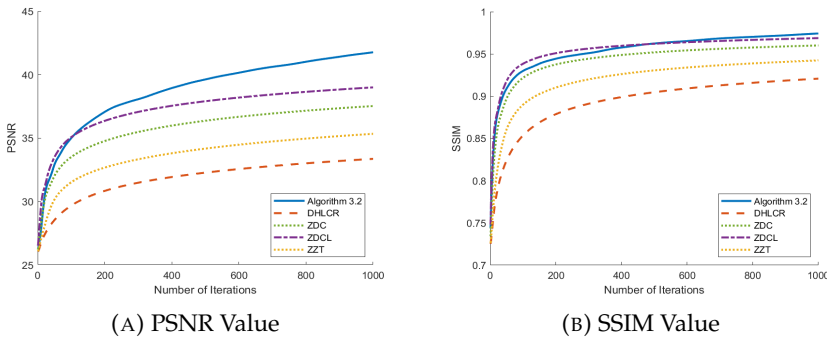


FIGURE 3. PSNR and SSIM in recovery for View image using Gaussian blur

Based on the numerical experiments, our proposed algorithm exhibits superior convergence performance compared to DHLCR Algorithm, ZDC Algorithm, ZDCL Algorithm, and ZZT Algorithm in terms of PSNR and SSIM. It reveals that multi-inertial extrapolations defined in our algorithm can speed up its convergence behavior.

5. CONCLUSIONS

In this study, we have presented a novel forward-backward-forward algorithm for inclusion problems. Our approach incorporates a self-adaptive technique to eliminate the need for explicit Lipschitz assumptions and employs multi-inertial extrapolations to accelerate the convergence of the algorithm. Furthermore, we have established a weak convergence theorem under reasonable assumptions. Additionally, we have conducted numerical tests in image deblurring, demonstrating the superiority of our algorithm over existing methods in the literature. The experimental results validate the effectiveness and improved performance of our proposed approach.

Acknowledgments. The authors sincerely thank the anonymous reviewers for their suggestions that substantially improved the manuscript. Papatsara Inkrong has received funding support from School of Science, University of Phayao, a grant no. PBTSC66032. Moreover, Prasit Cholamjiak was supported by the National Research Council of Thailand under grant no. N41A640094, the Thailand Science Research and Innovation Fund and the University of Phayao (FF67).

REFERENCES

- [1] Abaidoo, R.; Agyapong, E. K. Financial development and institutional quality among emerging economies. *J. Econ. Dev.*, **24** (2022), No. 3, 198–216.
- [2] Alvarez, F. Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space. *SIAM J. Optim.* **14** (2004), no. 3, 773–782.
- [3] Attouch, H.; Peyrouquet, J.; Redont, P. A dynamical approach to an inertial forward-backward algorithm for convex minimization. *SIAM J. Optim.* **24** (2014), no. 1, 232–256.
- [4] Bauschke, H. H.; Combettes, P. L. *Convex Analysis and Monotone Operator Theory in Hilbert spaces*, 2011. *CMS books in mathematics*. DOI, 10, 978-1.
- [5] Bauschke, H. H.; Combettes, P. L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Second edition. 2017. pp. C1–C4. New York, Springer.
- [6] Beck, A.; Teboulle, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2** (2009), no. 1, 183–202.
- [7] Boţ, R. I.; Csetnek, E. R. An inertial forward-backward-forward primal-dual splitting algorithm for solving monotone inclusion problems. *Numer. Algorithms* **71** (2016), 519–540.

- [8] Boţ, R. I.; Csetnek, E. R.; Hendrich, C. Inertial Douglas–Rachford splitting for monotone inclusion problems. *Appl. Math. Comput.* **256** (2015), 472–487.
- [9] Brezis, H. Operateurs Maximaux Monotones. Chapitre II. *North-Holland Mathematics Studies* **5** (1973), 19–51.
- [10] Chen, C.; Chan, R. H.; Ma, S.; Yang, J. Inertial proximal ADMM for linearly constrained separable convex optimization. *SIAM J. Imaging Sci.* **8** (2015), no. 4, 2239–2267.
- [11] Combettes, P. L.; Glaudin, L. E. Quasi-nonexpansive iterations on the affine hull of orbits: from Mann’s mean value algorithm to inertial methods. *SIAM J. Optim.* **27** (2017), no. 4, 2356–2380.
- [12] Combettes, P. L.; Wajs, V. R. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling and Simulation* **4** (2005), no. 4, 1168–1200.
- [13] Daubechies, I.; Defrise, M.; De Mol, C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun Pure Appl Math.: A Journal Issued by the Courant Institute of Mathematical Sciences* **57** (2004), no. 11, 1413–1457.
- [14] Dong, Q. L.; Cho, Y. J.; Zhong, L. L.; Rassias, T. M. Inertial projection and contraction algorithms for variational inequalities. *J Glob Optim.* **70** (2018), 687–704.
- [15] Dong, Q. L.; Huang, J. Z.; Li, X. H.; Cho, Y. J.; Rassias, T. M. MiKM: multi-step inertial Krasnosel’skiĭ–Mann algorithm and its applications. *J Glob Optim.* **73** (2019), 801–824.
- [16] Dong, Q. L.; Yuan, H. B.; Cho, Y. J.; Rassias, T. M. Modified inertial Mann algorithm and inertial CQ-algorithm for nonexpansive mappings. *Optim. Lett.* **12** (2018), 87–102.
- [17] Duchi, J.; Singer, Y. Efficient online and batch learning using forward backward splitting. *J. Mach. Learn. Res.* **10** (2009), 2899–2934.
- [18] Hanjing, A.; Suantai, S. A fast image restoration algorithm based on a fixed point and optimization method. *Mathematics* **8** (2020), no. 3, 378.
- [19] Liang, J. *Convergence rates of first-order operator splitting methods* (Doctoral dissertation, Normandie Université; GREYC CNRS UMR 6072), 2016.
- [20] Lions, P. L.; Mercier, B. Splitting algorithms for the sum of two nonlinear operators. *SIAM J Numer Anal.* **16** (1979), no. 6, 96–979.
- [21] Liu, H.; Yang, J. Weak convergence of iterative methods for solving quasimonotone variational inequalities. *Comput Optim Appl.* **77** (2020), no. 2, 491–508.
- [22] López, G.; Martín-Márquez, V.; Wang, F.; Xu, H. K. (2012, January). Forward-backward splitting methods for accretive operators in Banach spaces. In *Abstract and Applied Analysis* (Vol. 2012). Hindawi.
- [23] Lorenz, D. A.; Pock, T. An inertial forward-backward algorithm for monotone inclusions. *J Math Imaging Vis.* **51** (2015), 311–325.
- [24] Minty, G. J. Monotone (nonlinear) operators in Hilbert space. *Duke Math. J.* **29** (1962), 341–346.
- [25] Opial, Z. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Am. Math. Soc.* **73** (1967), no. 4, 591–597.
- [26] Ortega, J. M.; Rheinboldt, W. C. Iterative solution of nonlinear equations in several variables. Academic, New York, 1970.
- [27] Osilike, M. O.; Aniagbosor, S. C. Weak and strong convergence theorems for fixed points of asymptotically nonexpansive mappings. *Math Comput Model Dyn Syst.* **32** (2000), no. 10, 1181–1191.
- [28] Passty, G. B. Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *J. Math. Anal. Appl.* **72** (1979), no. 2, 383–390.
- [29] Polyak, B. T. Some methods of speeding up the convergence of iteration methods. *USSR Comput. Math. and Math. Phys.* **4** (1964), no. 5, 1–17.
- [30] Polyak, B. T. Introduction to optimization. optimization software. Inc., Publications Division, New York, **1** (1987), 32.
- [31] Poon, C.; Liang, J. Trajectory of alternating direction method of multipliers and adaptive acceleration. *Adv. Neural Inf. Process. Syst.*, **32** (2019).
- [32] Raguet, H.; Fadili, J.; Peyr e, G. A generalized forward-backward splitting. *SIAM J. Imaging Sci.*, **6**(2013), no. 3, 1199–1226.
- [33] Takahashi, W. *Introduction to nonlinear and convex analysis*. Yokohama Publishers, 2009.
- [34] Tseng, P. A modified forward-backward splitting method for maximal monotone mappings. *SIAM J Control Optim.*, **38** (2000), no. 2, 431–446.
- [35] Zhang, C.; Dong, Q. L.; Chen, J. Multi-step inertial proximal contraction algorithms for monotone variational inclusion problems. *Carpathian J. Math.* **36** (2020), no. 1, 159–177.
- [36] Zong, C.; Zhang, G.; Tang, Y. Multi-step inertial forward-backward-half forward algorithm for solving monotone inclusion. *Linear and Multilinear Algebra*, **71** (2023), no. 4, 631–661.

SCHOOL OF SCIENCE, UNIVERSITY OF PHAYAO, PHAYAO 56000, THAILAND

Email address: papatsara.inkrong@gmail.com

Email address: prasit.ch@up.ac.th