J. Nonlinear Var. Anal. 5 (2021), No. 6, pp. 881-892
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.5.2021.6.03

WEAK AND LINEAR CONVERGENCE OF A GENERALIZED PROXIMAL POINT ALGORITHM WITH ALTERNATING INERTIAL STEPS FOR A MONOTONE INCLUSION PROBLEM

YEKINI SHEHU[1,*], JEREMIAH N. EZEORA[2]

[1]*Department of Mathematics, Zhejiang Normal University, Jinhua, China*
[2]*Department of Mathematics and Statistics, University of Port Harcourt, Rivers State, Nigeria*

Abstract. The proximal point algorithm (PPA) is a powerful tool for solving monotone inclusion problems. Recently, Tao and Yuan [On the optimal linear convergence rate of a generalized proximal point algorithm, J. Sci. Comput. 74 (2018), 826-850] proposed a generalized PPA (GPPA) for finding a zero point of a maximal monotone operator, and obtained the linear convergence rate of the generalized PPA. In this paper, we consider accelerating the GPPA with the aid of the inertial extrapolation. We propose a generalized proximal point algorithm with alternating inertial steps solving monotone inclusion problem, and obtain weak convergence results under some mild conditions. When the inverse of the involved monotone operator is Lipschitz continuous at the origin, we prove that the iterative sequence generated by our generalized proximal point algorithm is linearly convergent. The Fejér monotonicity of even subsequences of the iterative sequence is also recovered. Finally, we give some priori and posteriori error estimates of our generated sequences.
Keywords. Alternated inertial step; Maximal monotone operator; Proximal point algorithm; Weak and linear convergence.

1. Introduction

Throughout this paper, H is assumed to be a real Hilbert space with inner product $\langle , \rangle$ and induced norm $\|\cdot\|$. Given a maximal monotone set-valued operator, $T : H \to 2^H$, we consider the following inclusion problem

$$\text{find } x \in H \ \text{ such that } \ 0 \in T(x). \tag{1.1}$$

We denote by $zer(T)$ the set of solutions of inclusion problem (1.1), and assume throughout this paper that $zer(T) \neq \emptyset$. It is well known that (1.1) serves as a unifying model for many problems of fundamental importance, including fixed point problems, variational inequality problems, and their variants and extensions and has many real applications in the real world; see, e.g., [1, 2, 3, 4, 5] and the references therein.

*Corresponding author.
E-mail addresses: yekini.shehu@zjnu.edu.cn (Y. Shehu), jeremiah.ezeora@uniport.edu.ng (J.N. Ezeora).
Received April 6, 2021; Accepted May 10, 2021.

The proximal point algorithm (PPA), which was first studied by Martinet and further developed by Rockafella (see, e.g., [6, 7]), has been studied for studying problem (1.1) for many years.

Define the resolvent operator of T by $J_{cT} := (I + cT)^{-1}$, where c is a positive constant. The resolvent operator is single-valued and $Fix(J_{cT}) = zer(T)$, where $Fix(J_{cT})$ denotes the fixed point set of J_{cT}. Starting from an arbitrary point $z^0 \in H$, the exact form of the PPA iteratively generates its sequence $\{z^k\}$ as

$$z^{k+1} = J_{cT}(z^k), \tag{1.2}$$

which is equivalent to

$$0 \in cT(z^{k+1}) + z^{k+1} - z^k, \tag{1.3}$$

where c, which is called proximal parameter, is a positive real number. The inexact version of the PPA is defined as:

$$z^{k+1} \approx J_{cT}(z^k). \tag{1.4}$$

In (1.2), the tolerance of accuracy is zero, and so (1.4) includes (1.2). However, (1.2) is of interest in its own right, since it requires estimating the resolvent accurately. Under different settings, both the exact and inexact versions of the PPA have been investigated in the literature. In [8], the convergence of both the exact and inexact versions of PPA was comprehensively studied. It turns out that the PPA is a very powerful algorithmic tool and includes many known algorithms, such as, the classical augmented Lagrangian method, the Douglas-Rachford splitting method, and the alternating direction method of multipliers as special cases. For more facts about the PPA and its generalizations, we refer to [9, 10, 11].

The equivalent representation of the PPA (1.3) can be written as

$$0 \in \frac{z^{k+1} - z^k}{c} + T(z^{k+1}).$$

This can be viewed as an implicit discretization of the evolution differential inclusion problem

$$0 \in \frac{dx}{dt} + T(x(t)) \tag{1.5}$$

It has been shown that the solution trajectory of (1.5) converges to a solution of (1.1) provided that T satisfies certain conditions (see, e.g., [12]). To speed up convergence, the following second order evolution differential inclusion problem was introduced in the literature:

$$0 \in \frac{d^2x}{dt^2} + c\frac{dx}{dt} + T(x(t)), \tag{1.6}$$

where $c > 0$ is a friction parameter. If $T = \nabla f$, where $f : \mathbb{R}^2 \to \mathbb{R}$ is a differentiable convex function with attainable minimum, system (1.6) characterizes roughly the motion of a heavy ball which rolls under its own inertia over the graph of f until friction stops it at a stationary point of f. In this case, the three terms in (1.6) denote, respectively, inertial force, friction force, and gravity force. Consequently, system (1.6) is usually referred to as the heavy-ball with friction (HBF) system. In theory, the convergence of the solution trajectories of the HBF system to a solution of (1.1) can be faster than those of the first-order system (1.5), while in practice the second order inertial term $\frac{d^2x}{dt^2}$ can be exploited to design faster algorithms (see, e.g., [13, 14]).

As a result of the properties of (1.6), an implicit discretization method was proposed in [15, 16] as follows. Given z^{k-1} and z^k, the next point z^{k+1} is determined via

$$0 \in \frac{z^{k+1}-2z^k+z^{k-1}}{h^2}+\gamma\frac{z^{k+1}-z^k}{h}+T(z^{k+1}),$$

which results to an iterative algorithm of the form

$$z^{k+1}=J_{cT}(z^k+\alpha(z^k-z^{k-1})), \tag{1.7}$$

where $c=\frac{h^2}{1+ch}$ and $\alpha=\frac{1}{1+ch}$. Observe that (1.7) is the proximal point step applied to the extrapolated point $z^k+\alpha(z^k-z^{k-1})$, rather than x^k itself as in the classical PPA. Hence, iterative algorithm (1.7) is a two-step method, which is generally called the inertial PPA (iPPA). The convergence properties of (1.7) were studied in [15, 16] under some assumptions on α and c.

Since the iPPA was introduced, the inexact and other forms of iPPAs have been studied by many authors; see, e.g., [7, 17, 18] and the references therein. Recently, there are increasing interests in studying inertial-type algorithms, for example, inertial forward-backward splitting methods, inertial Douglas-Rachford splitting methods, inertial alternating method of multipliers (ADMM), and inertial forward-backward-forward methods; see, e.g., [19, 20, 21, 22, 23] and the references therein. Here, we also mention the results presented in [24, 25, 26] in which the convergence and some real applications of inertial algorithms for maximal monotone inclusion problems and variational inequalities were investigated. A major drawback of the iPPA is that Fejér monotonicity of $\|z^k-z^*\|$, where $z^*\in zer(T)$, is lost in many cases and hence, makes the sequence $\{z^k\}$ generated by the methods to swing back and forth around $zer(T)$. This situation makes these methods sometimes not converge faster than their counterpart non-inertial methods (see, e.g., [27, 28]). Furthermore, no linear convergence rate of the iPPA was obtained in the literature (see, for example, [29, 30]).

A search to overcome this drawback gave birth to the so-called *alternated iPPA* in the literature.

To overcome this drawback, the so-called alternated iPPA was considered recently. It has been shown that with the alternated iPPA, some sort of Fejér monotonicity of $\|z^k-z^*\|$ is recovered and that the method out-performs their non-inertial counterparts; see, e.g., [28, 31, 32] for more details. In [11], Tao and Yuan studied the generalized PPA for a maximal monotone set-valued operator T with T^{-1} being Lipschitz continuous at $\mathbf{0}$ in real Hilbert space, and obtained both weak and linear convergence results.

Inspired by the results mentioned above, the aim of this paper is to propose an alternated inertial generalized PPA and prove:

- the Fejér monotonicity of $\|z^k-z^*\|$, which was not obtained for iPPA in [15, 16];
- the weak convergence of the generated sequence $\{z^k\}$ to a point in $zer(T)$, which generalizes the results obtained in [32];
- the linear convergence of the generated sequence $\{z^k\}$ to a unique point z^* under the condition that T is maximal monotone and T^{-1} is Lipschitz continuous at $\mathbf{0}$, which was not obtained for iPPA in [15, 16, 32] and other related works (see, e.g., [33]);
- the priori and posteriori error estimates of the generated sequence $\{z^k\}$, which are new for alternated inertial generalized PPA.

2. Preliminaries

In this section, we present some definitions and known results needed for our convergence analysis.

Recall that a set-valued mapping $T : H \to 2^H$ is said to be monotone if, for any $x, y \in H$, $\langle x-y, f-g\rangle \geq 0$, where $f \in Tx$ and $g \in Ty$. The Graph of T is defined by

$$Gr(T) := \{(x, f) \in H \times H : f \in Tx\}.$$

If $Gr(T)$ is not properly contained in the graph of any other monotone mapping, then T is said to be maximal.

Let $T : H \to 2^H$ be a set-valued mapping. T^{-1} is said to be Lipschitz continuous at $\mathbf{0}$ with modulus $\alpha \geq 0$ if there is a unique solution z^* to $0 \in T(z)$ (i.e.$T^{-1}(0) = z^*$), and for some $\tau > 0$, $\|x - x^*\| \leq \alpha\|w\|$ whenever $x \in T^{-1}(w)$ and $\|w\| \leq \tau$.

The above definition was given in [8]. The example of a set-valued map T such that T^{-1} is Lipschitz continuous at $\mathbf{0}$ ws given in [11]. It also was shown in [11] that the Lipschitz continuity at $\mathbf{0}$ is weaker than the strong monotonicity assumed in [23, 34].

Recall that a sequence $\{z^k\}$ in H is said to converge weakly to $z^* \in H$ if, for all $q \in H$, $\lim_{k\to\infty}\langle z^k, q\rangle = \langle z^*, q\rangle$.

Let $\{z^k\}$ be a sequence in H converging to $z^* \in H$ in norm. We say that $\{z^k\}$ converges to z^* R-linearly if $\limsup_{k\to\infty}\|z^k - z^*\| < 1$. $\{z^k\}$ is said to converge to z^* Q-linearly if there exists $\sigma \in (0,1)$ such that $\|z^{k+1} - z^*\| \leq \sigma\|z^k - z^*\|$ for k sufficiently large. It is known that the Q-linear convergence implies the R-linear convergence, but the reverse implication is not true.

Lemma 2.1. *(see [11]) Let $T : H \to 2^H$ be set-valued and maximal monotone, and define $J_{cT} := (I + cT)^{-1}$ with $c > 0$. Then,*
(i) $\langle J_{cT}(z) - J_{cT}(z'), (I - J_{cT})(z) - (I - J_{cT})(z')\rangle \geq 0$, $\forall z, z' \in H$
(ii) $\|z - z'\|^2 \geq \|J_{cT}(z) - J_{cT}(z')\|^2 + \|(I - J_{cT})(z) - (I - J_{cT})(z')\|^2$, $\forall z, z' \in H$.

Lemma 2.2. *The following statements hold in H:*
(i) $\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$, $\forall\, x, y \in H$.
(ii) $\|x+y\|^2 \leq \|x\|^2 + \langle y, x+y\rangle$, $\forall x, y \in H$.
(iii) $\|\alpha x + \beta y\|^2 = \alpha(\alpha+\beta)\|x\|^2 + \beta(\alpha+\beta)\|y\|^2 - \alpha\beta\|x-y\|^2$ $\forall x, y \in H$, $\forall \alpha, \beta \in \mathbb{R}$.

Lemma 2.3. *(see [35]) Let C be a nonempty set of H, and let $\{z^k\}$ be a sequence in H such that the following two conditions hold:*
(i) for any $z^* \in C$, $\lim_{k\to\infty}\|z^k - z^*\|$ exists;
(ii) every sequential weak cluster point of $\{z^k\}$ is in C.
Then $\{z^k\}$ converges weakly to a point in C.

3. Main Results

Now, we are ready to give our main results. We first introduce the exact version of the alternated inertial PPA and then give both weak and linear convergence results. Furthermore, priori and posteriori error estimates are also obtained.

3.1. Weak convergence results.

Theorem 3.1. *Let $T : H \to 2^H$ be a maximal monotone operator and suppose the following assumptions hold:*

(i) $\gamma \in (0,2)$;
(ii) $0 \le \alpha < \frac{2-\gamma}{\gamma}$;
(iii) $T^{-1}(0) \neq \emptyset$.
For given $z^0, z^1 \in H$, *let the sequence* $\{z^k\}$ *be generated by*

$$w^k = \begin{cases} z^k, & k = \text{even}, \\ z^k + \alpha(z^k - z^{k-1}), & k = \text{odd}, \end{cases}$$

and

$$z^{k+1} \quad = \quad w^k - \gamma(w^k - J_{cT}(w^k)), \; \forall\, k \ge 1. \tag{3.1}$$

Then $\{z^k\}$ *converges weakly to* $z^* \in T^{-1}(0)$.

Proof. From (3.1), we obtain $z^{2k+2} = w^{2k+1} - \gamma(w^{2k+1} - J_{cT}(w^{2k+1}))$. It follows that

$$\begin{aligned} \|z^{2k+2} - z^*\|^2 &= \|w^{2k+1} - z^*\|^2 - 2\gamma\langle w^{2k+1} - z^*, w^{2k+1} - J_{cT}(w^{2k+1})\rangle \\ &\quad + \gamma^2 \|w^{2k+1} - J_{cT}(w^{2k+1})\|^2 \\ &\le \|w^{2k+1} - z^*\|^2 - \gamma(2-\gamma)\|w^{2k+1} - J_{cT}(w^{2k+1})\|^2. \end{aligned} \tag{3.2}$$

Using Lemma 2.2 (iii), we obtain

$$\|w^{2k+1} - z^*\|^2 = (1+\alpha)\|z^{2k+1} - z^*\|^2 - \alpha\|z^{2k} - z^*\|^2 + \alpha(1+\alpha)\|z^{2k+1} - z^{2k}\|^2, \tag{3.3}$$

and

$$\begin{aligned} \|z^{2k+1} - z^*\|^2 \quad &\le \quad \|w^{2k} - z^*\|^2 - \gamma(2-\gamma)\|w^{2k} - J_{cT}(w^{2k})\|^2 \\ &= \quad \|z^{2k} - z^*\|^2 - \gamma(2-\gamma)\|w^{2k} - J_{cT}(w^{2k})\|^2. \end{aligned} \tag{3.4}$$

Therefore,

$$\begin{aligned} \|w^{2k+1} - z^*\|^2 &\le (1+\alpha)\left[\|z^{2k} - z^*\|^2 - \gamma(2-\gamma)\|z^{2k} - J_{cT}(z^{2k})\|^2\right] \\ &\quad - \alpha\|z^{2k} - z^*\|^2 + \alpha(1+\alpha)\|z^{2k+1} - z^{2k}\|^2 \\ &= \|z^{2k} - z^*\|^2 - (1+\alpha)\gamma(2-\gamma)\|z^{2k} - J_{cT}(z^{2k})\|^2 \\ &\quad + \alpha(1+\alpha)\|z^{2k+1} - z^{2k}\|^2. \end{aligned} \tag{3.5}$$

Substituting (3.5) into (3.2), we have

$$\begin{aligned} \|z^{2k+2} - z^*\|^2 &\le \|z^{2k} - z^*\|^2 - (1+\alpha)\gamma(2-\gamma)\|z^{2k} - J_{cT}(z^{2k})\|^2 \\ &\quad + \alpha(1+\alpha)\|z^{2k+1} - z^{2k}\|^2 - \gamma(2-\gamma)\|w^{2k+1} - J_{cT}(w^{2k+1})\|^2. \end{aligned}$$

Observe that

$$\begin{aligned} \|z^{2k+1} - z^{2k}\| &= \|w^{2k} - \gamma(w^{2k} - J_{cT}(w^{2k})) - w^{2k}\| \\ &= \gamma\|w^{2k} - J_{cT}(w^{2k}))\| \\ &= \gamma\|z^{2k} - J_{cT}(z^{2k}))\|. \end{aligned} \tag{3.6}$$

Thus,

$$\begin{aligned}\|z^{2k+2}-z^*\|^2 &\leq \|z^{2k}-z^*\|^2-(1+\alpha)\gamma(2-\gamma)\|z^{2k}-J_{cT}(z^{2k})\|^2\\ &\quad +\alpha(1+\alpha)\gamma^2\|z^{2k}-J_{cT}(z^{2k})\|^2\\ &\quad -\gamma(2-\gamma)\|w^{2k+1}-J_{cT}(w^{2k+1})\|^2\\ &\leq \|z^{2k}-z^*\|^2-(1+\alpha)\gamma\big((2-\gamma)-\alpha\gamma\big)\|z^{2k}-J_{cT}(z^{2k})\|^2\\ &\quad -\gamma(2-\gamma)\|w^{2k+1}-J_{cT}(w^{2k+1})\|^2.\end{aligned} \tag{3.7}$$

Thus,

$$\|z^{2k+2}-z^*\|^2 \leq \|z^{2k}-z^*\|^2. \tag{3.8}$$

From (3.8), we conclude that $\{z^{2k}-z^*\}$ is Fejér monotone and $\lim_{k\to\infty}\|z^{2k}-z^*\|$ exists. Furthermore, $\{\|z^{2k}-z^*\|\}$ and $\{z^{2k}\}$ are bounded. From (3.7), we obtain

$$\lim_{k\to\infty}\|w^{2k+1}-J_{cT}(w^{2k+1})\|=0$$

In view of (3.6) and (3.7), we have that

$$\lim_{k\to\infty}\|z^{2k}-J_{cT}(z^{2k})\|=\lim_{k\to\infty}\|z^{2k+1}-z^{2k}\|=0, \tag{3.9}$$

which is due to $w^{2k}=z^{2k}$. From (3.1), we obtain

$$\lim_{k\to\infty}\|z^{2k+2}-w^{2k+1}\|=\gamma\lim_{k\to\infty}\|w^{2k+1}-J_{cT}(w^{2k+1})\|=0.$$

Since $\{z^{2k}\}$ is bounded, there exists a subsequence $\{z^{2k_j}\}$ of $\{z^{2k}\}$, which converges weakly to z^*, say. Let $\bar{z}^k := J_{cT}(z^k)$. Now, using the definition of $J_{cT}(z^{2k})$, we obtain $c^{-1}\big(z^{2k_j}-\bar{z}^{2k_j}\big)\in T(\bar{z}^{2k_j})$. The monotonicity of T gives that

$$\langle u-\bar{z}^{2k_j}, v-c^{-1}\big(z^{2k_j}-\bar{z}^{2k_j}\big)\rangle \geq 0,\ \forall\, u,v \text{ such that } v\in T(u). \tag{3.10}$$

From (3.9), we obtain

$$\lim_{k\to\infty}\|z^{2k}-\bar{z}^{2k}\|=0. \tag{3.11}$$

Since $\{z^{2k_j}\}$ converges weakly to z^* and (3.11) holds, then $\{\bar{z}^{2k_j}\}$ converges weakly to z^*. Hence, we obtain from (3.10) that $\langle u-z^*,v\rangle \geq 0$. Since T is maximal monotone, (see, e.g., [8]), we conclude that $z^*\in T^{-1}(0)$. Therefore, by Lemma 2.3 we have that $\{z^{2k}\}$ converges weakly to an element of $T^{-1}(0)$. Assume for contradiction that there exists $\xi\in H$ such that z^{2k_j} converges weakly to ξ. Then

$$\begin{aligned}\|z^*-\xi\|^2 &= \langle z^*-\xi,z^*-\xi\rangle=\langle z^*,z^*-\xi\rangle-\langle \xi,z^*-\xi\rangle\\ &=\lim_{k\to\infty}\langle z^{2k},z^*-\xi\rangle-\lim_{k\to\infty}\langle z^{2k},z^*-\xi\rangle\\ &=\lim_{k\to\infty}\langle z^{2k}-z^{2k},z^*-\xi\rangle=0.\end{aligned}$$

Hence, z^* is unique. The definition of the weak convergence gives $\lim_{k\to\infty}\langle z^{2k}-z^*,\mu\rangle=0$ $\forall\mu\in H$. From (3.9), we have, for all $\mu\in H$, that

$$\begin{aligned}|\langle z^{2k+1}-z^*,\mu\rangle| &= |\langle z^{2k+1}-z^*+z^{2k}-z^{2k},\mu\rangle|\\ &\leq |\langle z^{2k}-z^*,\mu\rangle|+\|z^{2k+1}-z^{2k}\|\|\mu\|\to 0,\ k\to\infty\end{aligned}$$

Hence $\{z^k\}$ converges weakly to $z^*\in T^{-1}(0)$. $\square$

Remark 3.1.
- In (3.8), the Fejér monotonicity of even terms of the iterative sequence of our proposed algorithm with respect to the solution set $zer(T)$ is obtained.
- In [16, 17], the inertial factor α is assumed to satisfy $0 \leq \alpha < \frac{1}{3}$ in the proposed vanilla inertial PPA and the choice $\alpha \geq \frac{1}{3}$ is not allowed in many PPA with vanilla inertial extrapolation step in the literature. In algorithm (3.1), the inertial factor $\alpha \geq \frac{1}{3}$ is possible. For instance, if one takes $\gamma = \frac{5}{4}$, then $\frac{(2-\gamma)}{\gamma} = \frac{3}{5} > \frac{1}{3}$. This makes our proposed method novel and interesting.
- If $\alpha = 0$, then Theorem 3.1 reduces to [11, Theorem 3.2].

3.2. **Linear convergence results.** We now give the linear convergence of our proposed alternated inertial generalized PPA (3.1) under the condition that T^{-1} is Lipschitz continuous at $\mathbf{0}$ with modulus $a > 0$. To obtain the linear convergence results, we give the following lemma, which was obtained in [11, Lemma 3.3].

Lemma 3.1. *Let $T : H \to 2^H$ be a maximal monotone, and let z^* be a solution point of (1.1). Assume that $c > 0$. If T^{-1} is Lipschitz continuous at $\mathbf{0}$ with modulus $a > 0$, then there exists a positive number τ such that*

$$\|J_{cT}(z) - z^*\| \leq \frac{a}{\sqrt{a^2+c^2}}\|z - z^*\| \text{ when } \|c^{-1}(z - J_{cT}(z))\| \leq \tau \; \forall z \in H. \tag{3.12}$$

Remark 3.2. It was shown in [11, Remark 2] that the condition that "T^{-1} is Lipschitz continuous at $\mathbf{0}$ with modulus $a > 0$" is weaker than the condition "T is $\frac{1}{a}$-strongly monotone" assumed in [34] and (3.12) is optimal in the sense that the coefficient in the right-hand side cannot be smaller.

Using Lemma 3.1 above and the similar arguments used in [11, Lemma 3.4] and [8, Theorem 2], we obtain the following result.

Lemma 3.2. *Let $\{z^k\}$ be the sequence generated by the algorithm (3.1) with $\gamma \in (0,2)$, and let z^* be a solution point of (1.1). If T^{-1} is Lipschitz continuous at $\mathbf{0}$ with modulus $a > 0$, and the proximal parameter $c > 0$, then there exists an integer $\hat{k}$ such that*

$$\|J_{cT}(w^k) - z^*\| \leq \frac{a}{\sqrt{a^2+c^2}}\|w^k - z^*\|, \; \forall\, k > \hat{k}.$$

Next, we give the following linear convergence.

Theorem 3.2. *Suppose that T^{-1} is Lipschitz continuous at $\mathbf{0}$ with modulus $a > 0$, and the proximal parameter c is positive. Then the sequence $\{z^k\}$ generated by (3.1) with $\gamma \in (0,2)$ converges strongly to z^*, the unique solution of (1.1). Moreover there exists an integer $\hat{k}$ such that*

$$\|z^{k+1} - z^*\|^2 \leq \tau\|w^k - z^*\|^2, \; \forall\, k > \hat{k} \text{ with} \tag{3.13}$$

$$\tau := 1 - \min\left\{\gamma, 2\gamma - \gamma^2\right\}\frac{c^2}{a^2+c^2} \in (0,1). \tag{3.14}$$

Furthermore,

$$\|z^k - z^*\|^2 \leq \begin{cases} \frac{\|z^2-z^*\|^2}{\tau}\tau^{\frac{k}{2}}, & k = \text{even}, \\ \\ \frac{\|z^2-z^*\|^2}{\tau}\tau^{\frac{(k-1)}{2}}, & k = \text{odd}, \end{cases}$$

and $\{z^k\}$ converges R-linearly to a solution of (1.1).

Proof. Observe that

$$
\begin{aligned}
\|z^{k+1}-z^*\|^2 &= \|(1-\gamma)(w^k-z^*)+\gamma(J_cT(w^k)-z^*)\|^2 \\
&= (1-\gamma)^2\|w^k-z^*\|^2+2\gamma(1-\gamma)\langle w^k-z^*,\bar{w}^k-z^*\rangle+\gamma^2\|\bar{w}^k-z^*\|^2 \\
&= (1-\gamma)^2\|w^k-z^*\|^2+2\gamma(1-\gamma)\|\bar{w}^k-z^*\|^2+\gamma^2\|\bar{w}^k-z^*\|^2 \\
&\quad +2\gamma(1-\gamma)\langle w^k-\bar{w}^k,\bar{w}^k-z^*\rangle \text{ with } \bar{w}^k=J_{cT}(w^k) \\
&= (1-\gamma)^2\|w^k-z^*\|^2+(2\gamma-\gamma^2)\|\bar{w}^k-z^*\|^2 \\
&\quad +2\gamma(1-\gamma)\langle w^k-\bar{w}^k,\bar{w}^k-z^*\rangle.
\end{aligned} \tag{3.15}
$$

For $\gamma=1$, assertions (3.13) and (3.14) immediately follow from Lemma 3.2. For $0<\gamma\le 1$, we have from Lemma 3.2 the following estimates

$$
\begin{aligned}
&\|z^{k+1}-z^*\|^2 \\
&= (1-\gamma)^2\|w^k-z^*\|^2+(2\gamma-\gamma^2)\|\bar{w}^k-z^*\|^2+2\gamma(1-\gamma)\langle w^k-z^*,\bar{w}^k-z^*\rangle \\
&\quad -2\gamma(1-\gamma)\langle \bar{w}^k-z^*,\bar{w}^k-z^*\rangle \\
&= (1-\gamma)^2\|w^k-z^*\|^2+(2\gamma-\gamma^2)\|\bar{w}^k-z^*\|^2+2\gamma(1-\gamma)\langle w^k-z^*,\bar{w}^k-z^*\rangle \\
&\quad -2\gamma(1-\gamma)||\bar{w}^k-z^*||^2 \\
&\le (1-\gamma)^2\|w^k-z^*\|^2+(2\gamma-\gamma^2)\|\bar{w}^k-z^*\|^2+\gamma(1-\gamma)||w^k-z^*||^2+\gamma(1-\gamma)||\bar{w}^k-z^*||^2 \\
&\quad -2\gamma(1-\gamma)||\bar{w}^k-z^*||^2 \\
&= (1-\gamma)\|w^k-z^*\|^2+\gamma\|\bar{w}^k-z^*\|^2 \\
&\le \big(1-\gamma\frac{c^2}{a^2+c^2}\big)\|w^k-z^*\|^2.
\end{aligned}
$$

That is,

$$
\|z^{k+1}-z^*\|^2 \le \Big(1-\gamma\frac{c^2}{a^2+c^2}\Big)\|w^k-z^*\|^2
$$

By using the Lemma 2.1 (ii) with $z=w^k$, $\bar{w}^k=J_{cT}(w^k)$, and $z'=z^*$, we have that

$$
\|w^k-z^*\|^2 \ge \|\bar{w}^k-z^*\|^2+\|w^k-\bar{w}^k\|^2,
$$

which implies that $\langle w^k-\bar{w}^k,\bar{w}^k-z^*\rangle\ge 0$. If $1<\gamma<2$, then

$$
2\gamma(1-\gamma)\langle w^k-\bar{w}^k,\bar{w}^k-z^*\rangle \le 0.
$$

Hence, from (3.15) and Lemma 3.2, we obtain

$$
\|z^{k+1}-z^*\|^2 \le \Big(1-(2\gamma-\gamma^2)\frac{c^2}{a^2+c^2}\Big)\|w^k-z^*\|^2,\ \forall k>\hat{k}.
$$

For $\gamma\in(0,2)$ and $c\ge a>0$, we obtain

$$
\begin{aligned}
0<1-\min\Big\{\gamma,2\gamma-\gamma^2\Big\} \le \tau &:= 1-\min\Big\{\gamma,2\gamma-\gamma^2\Big\}\frac{c^2}{a^2+c^2} \\
&< 1-\min\Big\{\gamma,2\gamma-\gamma^2\Big\}\frac{a^2}{a^2+a^2}<1.
\end{aligned}
$$

Therefore,

$$\|z^{k+1}-z^*\|^2 \leq \Big(1-\min\Big\{\gamma, 2\gamma-\gamma^2\Big\}\frac{c^2}{a^2+c^2}\Big)\|w^k-z^*\|^2.$$

This establishes (3.13).

Next, replacing k with $2k$ in (3.13), we have

$$\|z^{2k+1}-z^*\|^2 \leq \tau\|w^{2k}-z^*\|^2 = \tau\|z^{2k}-z^*\|^2 \tag{3.16}$$

Further, replacing k with $2k+1$ in (3.13), we obtain

$$\|z^{2k+2}-z^*\|^2 \leq \tau\|w^{2k+1}-z^*\|^2. \tag{3.17}$$

Using (3.3), (3.4), (3.5), and (3.6), we obtain

$$\begin{aligned}\|w^{2k+1}-z^*\|^2 &\leq \|z^{2k}-z^*\|^2-(1+\alpha)\gamma(2-\gamma)\|z^{2k}-J_{cT}(z^{2k})\|^2\\ &\qquad +\alpha(1+\alpha)\gamma^2\|z^{2k}-J_{cT}(z^{2k})\|^2\\ &= \|z^{2k}-z^*\|^2-(1+\alpha)\gamma\big((2-\gamma)-\alpha\gamma\big)\|z^{2k}-J_{cT}(z^{2k})\|^2\\ &\leq \|z^{2k}-z^*\|^2.\end{aligned} \tag{3.18}$$

Combining (3.17) and (3.18), we have

$$\begin{aligned}\|z^{2k+2}-z^*\|^2 &\leq \tau\|z^{2k}-z^*\|^2\\ &\leq \tau^2\|z^{2k-2}-z^*\|^2\\ &\vdots\\ &\leq \tau^k\|z^2-z^*\|^2.\end{aligned} \tag{3.19}$$

Thus, $\|z^{2k}-z^*\|^2 \leq \tau^{k-1}\|z^2-z^*\|^2$, $\forall k>\bar{k}$, which together with (3.16) yields that

$$\begin{aligned}\|z^{2k+1}-z^*\|^2 &\leq \tau\|z^{2k}-z^*\|^2\\ &\vdots\\ &\leq \tau^{k-1}\|z^2-z^*\|^2.\end{aligned} \tag{3.20}$$

Thus, for all $k>\bar{k}$,

$$\|z^k-z^*\|^2 \leq \begin{cases}\frac{\|z^2-z^*\|^2}{\tau}\tau^{\frac{k}{2}}, & k=\text{even},\\[2ex] \frac{\|z^2-z^*\|^2}{\tau}\tau^{\frac{(k-1)}{2}}, & k=\text{odd}.\end{cases}$$

Hence $\{z^k\}$ converges R-linearly to $z^*\in zer(T)$. This completes the proof. □

Remark 3.3. Our Theorem 3.1 and Theorem 3.2 reduce to [11, Theorem 3.2] and [11, Theorem 3.5] respectively when $\alpha=0$ in (3.1).

Remark 3.4. Up to our knowledge, there is no linear convergence result for the inertial generalized PPA. Theorem 3.2 is one of main highlights of this paper.

We next give the following priori and posteriori error estimates of the subsequences generated (3.1).

Theorem 3.3. *Let $\{z^k\}$ be a sequence generated by (3.1). Assume that T^{-1} is Lipschitz continuous at $\mathbf{0}$ with modulus $a > 0$, the proximal parameter c is positive, and z^* is the unique solution of (1.1). Then*
(i)

$$\|z^{2k+2} - z^*\| \leq \frac{\tau^{\frac{k}{2}}}{1-\sqrt{\tau}}\|z^2 - z^4\|, \ \forall k \geq 1,$$

and

$$\|z^{2k+2} - z^*\| \leq \frac{\sqrt{\tau}}{1-\sqrt{\tau}}\|z^{2k} - z^{2k+2}\|, \ \forall k \geq 1;$$

(ii)

$$\|z^{2k+1} - z^*\| \leq \frac{\tau^{\frac{k-1}{2}}}{1-\sqrt{\tau}}\|z^2 - z^4\|, \ \forall k \geq 1,$$

and

$$\|z^{2k+1} - z^*\| \leq \frac{\sqrt{\tau}}{1-\sqrt{\tau}}\|z^{2k} - z^{2k+2}\|, \ \forall k \geq 1,$$

where τ is as given in (3.14).

Proof. Observe that

$$\begin{aligned}\|z^{2k} - z^*\| &\leq \|z^{2k} - z^{2k+2}\| + \|z^{2k+2} - z^*\| \\ &\leq \|z^{2k} - z^{2k+2}\| + \sqrt{\tau}\|z^{2k} - z^*\|, \ \forall k \geq 1.\end{aligned}$$

Therefore,

$$\|z^{2k} - z^*\| \leq \frac{1}{1-\sqrt{\tau}}\|z^{2k} - z^{2k+2}\|, \ \forall k \geq 1. \tag{3.21}$$

By (3.19) and (3.21), we obtain

$$\|z^{2k+2} - z^*\| \leq \tau^{\frac{k}{2}}\|z^2 - z^*\| \leq \frac{\tau^{\frac{k}{2}}}{1-\sqrt{\tau}}\|z^2 - z^4\|, \ \forall k \geq 1.$$

Again, using (3.19) and (3.21), we have

$$\|z^{2k+2} - z^*\| \leq \sqrt{\tau}\|z^{2k} - z^*\| \leq \frac{\sqrt{\tau}}{1-\sqrt{\tau}}\|z^{2k} - z^{2k+2}\|, \ \forall k \geq 1.$$

Hence, (i) is established.

By (3.20) and (3.21), we have

$$\|z^{2k+1} - z^*\| \leq \frac{\tau^{\frac{k-1}{2}}}{1-\sqrt{\tau}}\|z^2 - z^4\|, \ \forall k \geq 1.$$

From (3.20) and (3.21), we obtain

$$\|z^{2k+1} - z^*\| \leq \sqrt{\tau}\|z^{2k} - z^*\| \leq \frac{\sqrt{\tau}}{1-\sqrt{\tau}}\|z^{2k} - z^{2k+2}\|,$$

which establishes (ii). This completes the proof. □

4. CONCLUSION

In this paper, we introduced an alternated inertial generalized PPA for solving a monotone inclusion problem in real Hilbert spaces. In addition to the weak convergence, we obtain the linear convergence of the algorithm when the inverse of the involved monotone operator is Lipschitz continuous at the origin. In the future, it is of interest to give the similar results, which were obtained in this paper, on the inexact version of the alternated inertial PPA, and estimate the worst-case convergence rate of the alternated inertial PPA in terms of the Yosida approximation operator as obtained in [34] for the generalized PPA.

Acknowledgements
The authors are grateful to the reviewers for their useful suggestions which improved the contents of this paper.

REFERENCES

[1] P. Cubiotti, J.C. Yao, On the Cauchy problem for a class of differential inclusions with applications, Appl. Anal. 99 (2020), 2543-2554.
[2] L. Liu, X. Qin, Strong convergence theorems for solving pseudo-monotone variational inequality problems and applications, Optimization, doi: 10.1080/02331934.2021.1905641.
[3] T.H. Cuong, J.C. Yao, N.D. Yen, Qualitative properties of the minimum sum-of-squares clustering problem, Optimization, 69 (2020), 2131-2154.
[4] L.V. Nguyen, Q.H. Ansari, X. Qin, Weak sharpness and finite convergence for solutions of nonsmooth variational inequalities in Hilbert spaces, Appl. Math. Optim. 84 (201), 807-828.
[5] S.Y. Cho, A convergence theorem for generalized mixed equilibrium problems and multivalued asymptotically nonexpansive mappings, J. Nonlinear Convex Anal. 21 (2020), 1017-1026.
[6] J.J. Moreau, Proximité et dualité dans un espace hilbertien, (French) Bull. Soc. Math. France 93 (1965), 273-299.
[7] A. Moudafi, E. Elisabeth, An approximate inertial proximal method using the enlargement of a maximal monotone operator, Int. J. Pure Appl. Math. 5 (2003), 283-299.
[8] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (1976), 877–898.
[9] J. Eckstein, D.P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, Math. Program. Ser. A. 55 (1992), 293-318.
[10] O. Güler, New proximal point algorithms for convex minimization, SIAM J. Optim. 2 (1992), 649-664.
[11] M. Tao, X. Yuan, On the optimal linear convergence rate of a generalized proximal point algorithm, J. Sci. Comput. 74 (2018), 826-850.
[12] R.E. Jr. Bruck, Asymptotic convergence of nonlinear contraction semigroups in Hilbert space, J. Funct. Anal. 18 (1975), 15–26.
[13] F. Aluffi-Pentini, V. Parisi, F. Zirilli, Algorithm 617. DAFNE: a differential-equations algorithm for nonlinear equations, ACM Trans. Math. Softw. 10 (1984), 317-324.
[14] A.S. Antipin, Minimization of convex functions on convex sets by means of differential equations, Differentsial'nye Uravneniya 30 (1994), 1475-1486, 1652 (translation in Differential Equations 30 (1994), 1365–1375).
[15] F. Alvarez, On the minimizing property of a second order dissipative system in Hilbert spaces, SIAM J. Control Optim. 38 (2000), 1102–1119.
[16] F. Alvarez, H. Attouch, An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, Set-Valued Anal. 9 (2001), 3–11.
[17] F. Alvarez, Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space, SIAM J. Optim. 14 (2004), 773–782.

[18] P.E. Maingé, N. Merabet, A new inertial-type hybrid projection-proximal algorithm for monotone inclusions, Appl. Math. Comput. 215 (2010), 3149-3162.
[19] L.C. Ceng, A. Petruse, X. Qin, J.C. Yao, A modified inertial subgradient extragradient method for solving pseudomonotone variational inequalities and common fixed point problems, Fixed Point Theory 21 (2020), 93-108.
[20] A. Chambolle, T. Pock, On the ergodic convergence rates of a first-order primal-dual algorithm, Math. Program. Ser. A, 159 (2016), 253–287.
[21] D. Davis, W. Yin, Faster convergence rates of relaxed Peaceman-Rachford and ADMM under regularity assumptions, Math. Oper. Res. 42 (2017), 783-805.
[22] X. Qin, L. Wang, J.C. Yao, Inertial splitting method for maximal monotone mappings, J. Nonlinear Convex Anal. 21 (2020), 2325-2333.
[23] R.I. Bot, E.R. Csetnek, C. Hendrich, Inertial Douglas-Rachford splitting for monotone inclusion problems, Appl. Math. Comput. 256 (2015), 472–487.
[24] D.A. Lorenz, T. Pock, An inertial forward-backward algorithm for monotone inclusions, J. Math. Imaging Vision 51 (2015), 311-325.
[25] Y. Shehu, Convergence rate analysis of inertial Krasnoselskii-Mann type iteration with applications, Numer. Funct. Anal. Optim. 39 (2018), 1077-1091.
[26] B. Tan, S.Y. Cho, Inertial projection and contraction methods for pseudomonotone variational inequalities with non-Lipschitz operators and applications, Appl. Anal. (2021), doi: 10.1080/00036811.2021.1979219 .
[27] P. Cholamjiak, D.V. Thong, Y.J. Cho, A novel inertial projection and contraction method for solving pseudomonotone variational inequality problems, Acta Appl. Math. 169 (2020), 217-245.
[28] Y. Shehu, O.S. Iyiola, Projection methods with alternating inertial steps for variational inequalities: Weak and linear convergence, Appl. Numer. Math. 157 (2020), 315-337.
[29] H. Attouch, A. Cabot, Convergence rate of a relaxed inertial proximal algorithm for convex minimization, Optimization 69 (2020), 1281-1312.
[30] J.-F. Aujol, C. Dossal, A. Rondepierre, Optimal convergence rates for Nesterov acceleration, SIAM J. Optim. 29 (2019), 3131-3153.
[31] C. Izuchukwu, Y. Shehu, Projection-type methods with alternating inertial steps for solving multivalued variational inequalities beyond monotonicity, J. Appl. Numer. Optim. 2 (2020), 249-277.
[32] Z. Mu, Y. Peng, A note on the inertial proximal point method. Stat. Optim. Inf. Comput. 3 (2015), 241–248.
[33] R.N. Nwokoye, C.C. Okeke, Y. Shehu, A new linear convergence method for a Lipschitz pseudomonotone variational inequality, Appl. Set-Valued Anal. Optim. 3 (2021), 215-220.
[34] E. Corman, X.M. Yuan, A generalized proximal point algorithm and its convergence rate, SIAM J. Optim. 24 (2014), 1614-1638.
[35] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, Bull. Amer. Math. Soc. 73 (1967), 591-597.