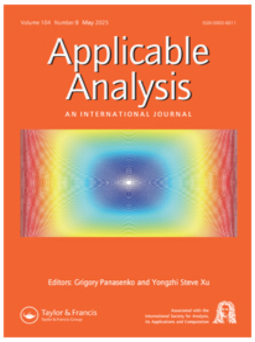

Applicable Analysis
An International Journal

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/gapa20

Convergence results for proximal point algorithm with inertial and correction terms

Chinedu Izuchukwu, Yekini Shehu & Jen-Chih Yao

APPLICABLE ANALYSIS
2025, VOL. 104, NO. 9, 1732–1752
https://doi.org/10.1080/00036811.2024.2432527

Convergence results for proximal point algorithm with inertial and correction terms

Chinedu Izuchukwu[a], Yekini Shehu[b] and Jen-Chih Yao[c,d]

[a]School of Mathematics, University of the Witwatersrand, Johannesburg, South Africa; [b]School of Mathematical Sciences, Zhejiang Normal University, Jinhua, People's Republic of China; [c]Center for General Education, China Medical University, Taichung, Taiwan; [d]UNEC Mathematical Modeling and Optimization Research Center, Azerbaijan State University of Economics (UNEC), Baku, Azerbaijan

ABSTRACT
This paper designs and studies a numerically fast proximal point algorithm to solve the monotone inclusion problem in Hilbert spaces. Our first proposed algorithm combines the proximal point algorithm, inertial term, and two correction terms. The addition of two correction terms is considered to further numerically accelerate the convergence speed of the inertial proximal point algorithm already studied in the literature. In our convergence results, we obtain both weak and linear convergence of the first proposed algorithm under some standard assumptions. Furthermore, we modify the first algorithm to obtain a strongly convergent algorithm. Applications of our proposed algorithm to mixed variational inequalities, strongly quasi-convex minimization problems, and Douglas–Rachford splitting algorithm are given. Numerical results show that our proposed algorithm outperforms other related algorithms in the literature.

ARTICLE HISTORY
Received 28 March 2024
Accepted 29 August 2024

COMMUNICATED BY
A. Zlotnik

KEYWORDS
Proximal point algorithm; correction term; inertial extrapolation; maximal monotone operators; Hilbert spaces

2020 MATHEMATICS SUBJECT CLASSIFICATIONS
90C25; 90C60; 68Q25; 49M25; 90C22

1. Introduction

Suppose H is a Hilbert space that is equipped with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. Given a maximal monotone set-valued operator $A : H \to 2^H$, consider the monotone inclusion problem

$$\text{find } x^* \in H \quad \text{such that} \quad \mathbf{0} \in A(x^*). \tag{1}$$

We shall denote the set of solutions to the monotone inclusion problem (1) by $A^{-1}(\mathbf{0})$. It is known that variational inequality problem, minimization of closed proper convex functions, and their variants and extensions can be formulated as monotone inclusion problem (1) above.

The proximal point algorithm (PPA) (see, for example, [1–5]) is a popular algorithm deployed to solve inclusion problem (1). Given $x_0 \in H$, the PPA is generated by

$$y_{k+1} = J_\lambda^A(y_k), \tag{2}$$

where the operator $J_\lambda^A := (I + \lambda A)^{-1}, \lambda > 0$, is the single-valued resolvent operator of A (see [3]). The classical augmented Lagrangian method [6,7], the Douglas-Rachford splitting method [8,9] and the alternating direction method of multipliers [10,11] are seen as special cases of PPA (2). Equivalently,

CONTACT Yekini Shehu yekini.shehu@zjnu.edu.cn, deltanougt2006@yahoo.com

PPA (2) is

$$\mathbf{0} \in \frac{y_{k+1} - y_k}{\lambda} + A(y_{k+1}), \tag{3}$$

which is an implicit discretization of the evolution differential inclusion problem

$$\mathbf{0} \in \frac{\mathrm{d}y}{\mathrm{d}t} + A(y(t)), \tag{4}$$

for which convergence of its solution trajectory to a solution of (1) has been established (see, for example, [12]). Over-relaxation version of PPA (2) has also been studied extensively in the literature (see, for example, [1,13–16]). Proximal point algorithm (2) with inertial term:

$$y_{k+1} = J_\lambda^A(y_k + \alpha_k(y_k - y_{k-1})), \tag{5}$$

with $\alpha_k \in [0, 1)$ has been studied in [17–25] and weak convergence results obtained under appropriate conditions.

Recently, Kim [26] studied accelerated proximal point algorithm, which is a combination of proximal point algorithm, inertial term and correction term. Kim [26, Theorem 4.1] proved the worst-case convergence rate, using performance estimation problem (PEP) approach of Drori–Teboulle [27], for the following algorithm:

$$\begin{cases} w_k = y_k + \dfrac{k-1}{k+1}(y_k - y_{k-1}) + \dfrac{k-1}{k+1}(w_{k-2} - y_{k-1}), \\ y_{k+1} = J_\lambda^A(w_k) \end{cases} \tag{6}$$

and obtained worst-case convergence rate $\|y_k - w_{k-1}\| = \mathcal{O}(1/k)$ in [26, Theorem 4.1]. However, no weak convergence of the sequence of iterates $\{w_k\}$ and $\{y_k\}$ generated by (6) is given. It should be noted that the results of Kim [26, Theorem 4.1] extend the results of Güler [28] from convex minimization to monotone inclusions.

Maingé in [29] further studied the following PPA with inertial term, relaxation term and correction term to solve the monotone inclusion problem (1):

$$\begin{cases} w_k = y_k + \alpha_k(y_k - y_{k-1}) + \delta_k(w_{k-1} - y_k) \\ y_{k+1} = \dfrac{1}{1+a_k} w_k + \dfrac{a_k}{1+a_k} J^A_{\lambda(1+a_k)}(w_k), \end{cases} \tag{7}$$

where α_k and δ_k are given recursively in [29, (1.23) and (1.24)]. Maingé [29, Theorem 1] obtained weak convergence results of $\{y_k\}$ generated by (7) with the fast rate $\|y_{k+1} - y_k\| = o(1/k)$ and $\frac{a_k}{1+a_k} \in (0, 1)$ but no linear rate convergence results given.

1.1. Our contributions

Motivated by the results in [26,29], we first propose a proximal point algorithm with a combination of one inertial term and two correction terms to solve the monotone inclusion problem (1). Weak and linear convergence results are obtained under standard conditions. In addition, we design a viscosity-type algorithm from the first algorithm and obtain strong convergence results. We also give applications to mixed variational inequalities, strongly quasi-convex minimization problems and Douglas–Rachford splitting algorithm. Numerical tests are carried out on our proposed algorithm, and comparisons with related algorithms show that our proposed algorithm is more efficient and performs better, in terms of number of iterations and CPU time, than other related algorithms in the literature.

1.2. Outline

Basic definitions and some needed results are given in Section 2. In Section 3, we introduce our proposed algorithm and obtain weak convergence, linear convergence rate and strong convergence. Some applications are given in Section 4 and numerical results are discussed in Section 5 with concluding remarks given in Section 6.

2. Preliminaries

In this section, we give some definitions and basic results that we will need in our analysis.

Definition 2.1: A mapping $T : H \to H$ is called

(i) nonexpansive if $\|Tx - Ty\| \le \|x - y\|$, for all $x, y \in H$;
(ii) firmly nonexpansive if $\|Tx - Ty\|^2 \le \|x - y\|^2 - \|(I - T)x - (I - T)y\|^2$ for all $x, y \in H$. Equivalently, T is firmly nonexpansive if $\|Tx - Ty\|^2 \le \langle x - y, Tx - Ty\rangle$ for all $x, y \in H$;
(iii) averaged if T can be expressed as the averaged of the identity mapping I and a nonexpansive mapping S, i.e. $T = (1 - \alpha)I + \alpha S$ with $\alpha \in (0, 1)$. Alternatively, T is α-averaged if

$$\|Tx - Ty\|^2 \le \|x - y\|^2 - \frac{1-\alpha}{\alpha}\|(I - T)x - (I - T)y\|^2, \quad \forall\, x, \quad y \in H.$$

Definition 2.2: An operator $A : H \to 2^H$ is said to be monotone if for any $x,\ y \in H$,

$$\langle x - y, f - g\rangle \ge 0,$$

where $f \in Ax$ and $g \in Ay$. The Graph of A is defined by

$$\mathrm{Gr}\,(A) := \{(x, f) \in H \times H : f \in Ax\}.$$

If $\mathrm{Gr}\,(A)$ is not properly contained in the graph of any other monotone operator, then A is maximal monotone operator. It is well-known that for each $x \in H$, and $\lambda > 0$, there is a unique $z \in H$ such that $x \in (I + \lambda A)z$.

Let $h : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a proper, lsc convex function. The proximal operator prox_h for h is defined as $\mathrm{prox}_h(z) := \mathrm{argmin}_x\{h(x) + \frac{1}{2}\|x - z\|^2\}$. Using the notions in [30,31], we say that a function $h : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ with a convex domain is strongly quasiconvex if there exists $\gamma \in (0, +\infty)$ such that for all $x, y \in \mathrm{dom}\, h$ and all $\lambda \in [0, 1]$, we have

$$h(\lambda y + (1 - \lambda)x) \le \max\{h(y), h(x)\} - \lambda(1 - \lambda)\frac{\gamma}{2}\|x - y\|^2.$$

Let K be a closed set in $\mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a proper function such that $K \cap \mathrm{dom}\, h \neq \emptyset$. We say that h is prox-convex on K [32] if there exists $\alpha > 0$ such that for every $z \in K$, $\mathrm{prox}_h(K, z) \neq \emptyset$, (where we write $\mathrm{prox}_h(K, z) := \mathrm{prox}_{(h+\delta_K)}(z)$ with δ_K being the indicator function on K) and

$$\bar{x} \in \mathrm{prox}_h(K, z) \Rightarrow h(\bar{x}) - h(x) \le \alpha\langle \bar{x} - z, x - \bar{x}\rangle, \quad \forall\, x \in K.$$

If $K = \mathbb{R}^n$, then we say that h is prox-convex. As stated in [32, Remark 3.6], the classes of strongly quasiconvex functions and prox-convex functions intersect without being included in one another.

The lemma below is quite well known, which plays an important role in proving weak convergence of sequences in a Hilbert space (see, [33]).

Lemma 2.3: *Let K be a nonempty subset of a Hilbert space H and let $\{x_n\}$ be a bounded sequence in H. Assume the following two conditions are satisfied:*

- $\lim_{n\to\infty} \|x_n - x\|$ *exists for each* $x \in K$,
- *every weak cluster point of* $\{x_n\}$ *belongs to* K.

Then $\{x_n\}$ *converges weakly to a point in* K.

Lemma 2.4 ([34]): *Suppose that* $\{t_p\}$ *is a sequence of nonnegative real numbers,* $\{\sigma_k\}$ *is a sequence of real numbers in* $(0, 1)$ *satisfying* $\sum_{k=1}^{\infty} \sigma_k = \infty$, *and* $\{h_k\}$ *is a sequence of real numbers such that*

$$t_{k+1} \le (1 - \sigma_k)t_k + \sigma_k h_k,\ k \ge 1.$$

If $\limsup_{i\to\infty} h_{k_i} \le 0$ *for each subsequence* $\{t_{k_i}\}$ *of* $\{t_k\}$ *satisfying* $\liminf_{i\to\infty}(t_{k_i+1} - t_{k_i}) \ge 0$, *then* $\lim_{k\to\infty} t_k = 0$.

Lemma 2.5 ([35, Lemma 3.1]): *Suppose that* $\{t_k\}$ *and* $\{r_k\}$ *are sequences of nonnegative real numbers such that*

$$t_{k+1} \le (1 - \sigma_k)t_k + s_k + r_k,$$

where $\{\sigma_k\}$ *is a sequence in* $(0, 1)$ *and* $\{s_k\}$ *is a real sequence. Let* $\sum_{k=1}^{\infty} r_k < \infty$ *and* $s_k \le \sigma_k M$ *for some* $M \ge 0$. *Then* $\{t_k\}$ *is bounded.*

3. Proposed methods & convergence

Assumption 3.1: In our convergence analysis, we assume that for some $\epsilon > 0$,

$$0 \le \alpha < \frac{\epsilon}{1+\epsilon}, \quad \frac{(1+\epsilon)\alpha}{1+\alpha\epsilon} < \delta < 1.$$

We now present our proposed method as follows:

Algorithm 1 (Accelerated PPA)

1: Choose α and δ with Assumption 3.1 satisfied. Pick $w_{-1} = w_{-2}, y_0, y_{-1} \in H$ arbitrarily, $\lambda > 0$ and set $k = 0$.

2: Given $w_{k-1}, w_{k-2}, y_{k-1}$ and y_k, compute

$$\begin{cases} w_k = y_k + \alpha(y_k - y_{k-1}) + \delta(1+\alpha)(w_{k-1} - y_k) \\ \qquad - \alpha\delta(w_{k-2} - y_{k-1}), \\ y_{k+1} = J_\lambda^A(w_k). \end{cases} \tag{8}$$

3: Set $k \leftarrow k + 1$, and **go to Step 2**.

Remark 3.2: When $\alpha = 0 = \delta$ in proposed Algorithm 1, we recover the proximal point algorithm. Also, if $\delta = 0$, our Algorithm 1 reduces to the inertial proximal point algorithm, for which weak convergence analysis has been established in [17–19]. We do not pursue the case of $\delta = 0$ in Algorithm 1 for our convergence analysis since the weak convergence for the case $\delta = 0$ in Algorithm 1 has been established in [17–19].

Remark 3.3: The proximal point algorithm in [17–25] contains inertial extrapolation step while our Algorithm 1 contains inertial extrapolation term (i.e. $\alpha(y_k - y_{k-1})$) and two correction terms (i.e.

$\delta(1+\alpha)(w_{k-1}-y_k)$ and $\alpha\delta(w_{k-2}-y_{k-1})$). Both the inertial extrapolation term and the two correction terms are added to the original PPA in order to improve on the numerical convergence speed of the PPA. Another interesting feature of our Algorithm 1 is that the inertial parameter α is not restricted to $[0, \frac{1}{3})$ as done in [17] and other related papers.

3.1. Weak convergence

Lemma 3.4: *Suppose $A : H \to 2^H$ is a maximal monotone operator with $A^{-1}(\mathbf{0}) \neq \emptyset$. Let $\{y_k\}$ be generated by Algorithm 1 with Assumption 3.1 fulfilled. Then $\{w_k\}$ and $\{y_k\}$ are bounded.*

Proof: Take $x^* \in A^{-1}(\mathbf{0})$. Noting that J_λ^A is firmly nonexpansive, we have

$$\|y_{k+1}-x^*\|^2 \leq \|w_k - x^*\|^2 - \|y_{k+1}-w_k\|^2. \tag{9}$$

For the rest of this proof, we define, for simplicity, $x_k = y_k + \delta(w_{k-1}-y_k)$. Then $w_k = x_k + \alpha(x_k - x_{k-1})$ in Algorithm 1. Consequently,

$$y_k = \frac{1}{1-\delta}x_k - \frac{\delta}{1-\delta}w_{k-1}$$

and thus

$$y_{k+1} = \frac{1}{1-\delta}x_{k+1} - \frac{\delta}{1-\delta}w_k.$$

Hence, we get

$$\begin{aligned}\|y_{k+1}-x^*\|^2 &= \left\|\frac{1}{1-\delta}(x_{k+1}-x^*) - \frac{\delta}{1-\delta}(w_k - x^*)\right\|^2 \\ &= \frac{1}{1-\delta}\|x_{k+1}-x^*\|^2 - \frac{\delta}{1-\delta}\|w_k - x^*\|^2 \\ &\quad + \frac{\delta}{(1-\delta)^2}\|x_{k+1}-w_k\|^2. \end{aligned} \tag{10}$$

If we combine (10) and (9), we have

$$\frac{1}{1-\delta}\|x_{k+1}-x^*\|^2 - \frac{\delta}{1-\delta}\|w_k-x^*\|^2 + \frac{\delta}{(1-\delta)^2}\|x_{k+1}-w_k\|^2 \leq \|w_k - x^*\|^2 - \|y_{k+1}-w_k\|^2.$$

Thus,

$$\begin{aligned}\frac{1}{1-\delta}\|x_{k+1}-x^*\|^2 \leq & \frac{1}{1-\delta}\|w_k - x^*\|^2 - \frac{\delta}{(1-\delta)^2}\|x_{k+1}-w_k\|^2 \\ & - \|y_{k+1}-w_k\|^2. \end{aligned} \tag{11}$$

Now,

$$\begin{aligned}\|w_k - x^*\|^2 =& \|(1+\alpha)(x_k - x^*) - \alpha(x_{k-1}-x^*)\|^2 \\ =& (1+\alpha)\|x_k - x^*\|^2 - \alpha\|x_{k-1}-x^*\|^2 + \alpha(1+\alpha)\|x_k - x_{k-1}\|^2. \end{aligned} \tag{12}$$

If we put (12) in (11), we obtain

$$\frac{1}{1-\delta}\|x_{k+1}-x^*\|^2 \leq \frac{1}{1-\delta}\Big[(1+\alpha)\|x_k-x^*\|^2-\alpha\|x_{k-1}-x^*\|^2 + \alpha(1+\alpha)\|x_k-x_{k-1}\|^2\Big]-\frac{\delta}{(1-\delta)^2}\|x_{k+1}-w_k\|^2-\|y_{k+1}-w_k\|^2.$$

Thus,

$$\begin{aligned}\frac{1}{1-\delta}\|x_{k+1}-x^*\|^2-\frac{\alpha}{1-\delta}\|x_k-x^*\|^2 \leq&\frac{1}{1-\delta}\|x_k-x^*\|^2-\frac{\alpha}{1-\delta}\|x_{k-1}-x^*\|^2\\&+\frac{\alpha(1+\alpha)}{1-\delta}\|x_k-x_{k-1}\|^2-\frac{\delta}{(1-\delta)^2}\|x_{k+1}-w_k\|^2\\&-\|y_{k+1}-w_k\|^2. \end{aligned}\tag{13}$$

Also, we have that

$$\begin{aligned}\|x_{k+1}-w_k\|^2 =&\|x_{k+1}-x_k-\alpha(x_k-x_{k-1})\|^2\\=&\|x_{k+1}-x_k\|^2-2\alpha\langle x_{k+1}-x_k,x_k-x_{k-1}\rangle+\alpha^2\|x_k-x_{k-1}\|^2\\\geq&\|x_{k+1}-x_k\|^2-\alpha\left[\|x_{k+1}-x_k\|^2+\|x_k-x_{k-1}\|^2\right]+\alpha^2\|x_k-x_{k-1}\|^2\\=&(1-\alpha)\|x_{k+1}-x_k\|^2-\alpha(1-\alpha)\|x_k-x_{k-1}\|^2. \end{aligned}\tag{14}$$

If we apply (14) in (13), we get

$$\begin{aligned}\frac{1}{1-\delta}\|x_{k+1}-x^*\|^2-\frac{\alpha}{1-\delta}\|x_k-x^*\|^2 \leq&\frac{1}{1-\delta}\|x_k-x^*\|^2-\frac{\alpha}{1-\delta}\|x_{k-1}-x^*\|^2\\&+\frac{\alpha(1+\alpha)}{1-\delta}\|x_k-x_{k-1}\|^2+\frac{\alpha\delta(1-\alpha)}{(1-\delta)^2}\|x_k-x_{k-1}\|^2\\&-\frac{\delta(1-\alpha)}{(1-\delta)^2}\|x_{k+1}-x_k\|^2-\|y_{k+1}-w_k\|^2. \end{aligned}\tag{15}$$

Similar to (14), we obtain

$$\begin{aligned}\|y_{k+1}-w_k\|^2 =&\|y_{k+1}-x_k-\alpha(x_k-x_{k-1})\|^2\\\geq&(1-\alpha)\|y_{k+1}-x_k\|^2-\alpha(1-\alpha)\|x_k-x_{k-1}\|^2.\end{aligned}$$

Hence, we have from (15) that

$$\begin{aligned}&\frac{1}{1-\delta}\|x_{k+1}-x^*\|^2-\frac{\alpha}{1-\delta}\|x_k-x^*\|^2+\frac{\delta(1-\alpha)}{(1-\delta)^2}\|x_{k+1}-x_k\|^2\\&\leq\frac{1}{1-\delta}\|x_k-x^*\|^2-\frac{\alpha}{1-\delta}\|x_{k-1}-x^*\|^2+\left[\frac{\alpha(1+\alpha)}{1-\delta}+\frac{\alpha\delta(1-\alpha)}{(1-\delta)^2}+\alpha(1-\alpha)\right]\|x_k-x_{k-1}\|^2\\&\quad-(1-\alpha)\|y_{k+1}-x_k\|^2\\&=\frac{1}{1-\delta}\|x_k-x^*\|^2-\frac{\alpha}{1-\delta}\|x_{k-1}-x^*\|^2+\frac{\delta(1-\alpha)}{(1-\delta)^2}\|x_k-x_{k-1}\|^2-(1-\alpha)\|y_{k+1}-x_k\|^2\\&\quad+\left[\frac{\alpha(1+\alpha)}{1-\delta}+\frac{\alpha\delta(1-\alpha)}{(1-\delta)^2}+\alpha(1-\alpha)\right]\|x_k-x_{k-1}\|^2-\frac{\delta(1-\alpha)}{(1-\delta)^2}\|x_k-x_{k-1}\|^2. \end{aligned}\tag{16}$$

Define for all $k \in \mathbb{N}$,

$$a_k := \frac{1}{1-\delta}\|x_k - x^*\|^2 - \frac{\alpha}{1-\delta}\|x_{k-1} - x^*\|^2 + \frac{\delta(1-\alpha)}{(1-\delta)^2}\|x_k - x_{k-1}\|^2. \tag{17}$$

Then,

$$\begin{aligned} a_{k+1} \leq & a_k - (1-\alpha)\|y_{k+1} - x_k\|^2 \\ & + \left[\frac{\alpha(1+\alpha)}{1-\delta} + \frac{\alpha\delta(1-\alpha)}{(1-\delta)^2} + \alpha(1-\alpha) - \frac{\delta(1-\alpha)}{(1-\delta)^2}\right]\|x_k - x_{k-1}\|^2. \end{aligned} \tag{18}$$

We show that $a_k \geq 0 \; \forall \, k \in \mathbb{N}$. By Peter–Paul inequality, we have

$$\begin{aligned} a_k &= \frac{1}{1-\delta}\|x_k - x^*\|^2 - \frac{\alpha}{1-\delta}\|x_{k-1} - x^*\|^2 + \frac{\delta(1-\alpha)}{(1-\delta)^2}\|x_k - x_{k-1}\|^2 \\ &\geq \frac{1}{1-\delta}\|x_k - x^*\|^2 - \frac{\alpha}{1-\delta}(1+\frac{1}{\epsilon})\|x_k - x^*\|^2 \\ &\quad - \frac{\alpha(1+\epsilon)}{1-\delta}\|x_k - x_{k-1}\|^2 + \frac{\delta(1-\alpha)}{(1-\delta)^2}\|x_k - x_{k-1}\|^2 \\ &= \frac{1}{1-\delta}\left(1 - \alpha\left(1 + \frac{1}{\epsilon}\right)\right)\|x_k - x^*\|^2 + \frac{1}{1-\delta}\left(\frac{\delta(1-\alpha)}{1-\delta} - \alpha(1+\epsilon)\right)\|x_k - x_{k-1}\|^2 \\ &\geq 0, \end{aligned}$$

by Assumption 3.1 that $\alpha < \frac{\epsilon}{1+\epsilon}$ and $\frac{\alpha(1+\epsilon)}{1+\alpha\epsilon} < \delta$.

Now, observe that

$$\frac{\alpha(1+\alpha)}{1-\delta} + \frac{\alpha\delta(1-\alpha)}{(1-\delta)^2} + \alpha(1-\alpha) - \frac{\delta(1-\alpha)}{(1-\delta)^2} < 0$$

if $\frac{2\alpha+1-\sqrt{-4\alpha^2+4\alpha+1}}{2\alpha} < \delta < \frac{2\alpha+1+\sqrt{-4\alpha^2+4\alpha+1}}{2\alpha}, 0 < \alpha < 1$ and if $\alpha = 0, 0 < \delta < 1$. If $0 < \alpha < 1$, then

$$\frac{2\alpha + 1 - \sqrt{-4\alpha^2 + 4\alpha + 1}}{2\alpha} < \frac{\alpha(1+\epsilon)}{1+\alpha\epsilon}$$

and

$$\frac{2\alpha + 1 + \sqrt{-4\alpha^2 + 4\alpha + 1}}{2\alpha} > 1.$$

Therefore, we have from (18) that $\{a_k\}$ is nonincreasing and thus $\lim_{k\to\infty} a_k$ exists. Hence, $\{a_k\}$ is bounded. Consequently, we have from (18) that

$$\lim_{k\to\infty}\|y_{k+1} - x_k\| = 0$$

and

$$\lim_{k\to\infty}\|x_k - x_{k-1}\| = 0.$$

Since $\frac{1}{1-\delta}\left(1 - \alpha\left(1 + \frac{1}{\epsilon}\right)\right)\|x_k - x^*\|^2 \leq a_k$ and $\{a_k\}$ is bounded, we obtain that $\{\|x_k - x^*\|\}$ is bounded. Thus, $\{x_k\}$ is bounded. From $w_k = x_k + \alpha(x_k - x_{k-1})$, we have that $\{w_k\}$ is bounded. Therefore, $\{y_k\}$ is also bounded. ∎

Theorem 3.5: *Suppose $A : H \to 2^H$ is a maximal monotone operator with $A^{-1}(\mathbf{0}) \neq \emptyset$. Let $\{y_k\}$ be generated by Algorithm 1 with Assumption 3.1 fulfilled. Then $\{y_k\}$ converges weakly to a point in $A^{-1}(\mathbf{0})$.*

Proof: From $w_k = x_k + \alpha(x_k - x_{k-1})$, we get

$$\lim_{k\to\infty} \|w_k - x_k\| = \alpha \lim_{k\to\infty} \|x_k - x_{k-1}\| = 0.$$

Also,

$$\|y_{k+1} - w_k\| \leq \|y_{k+1} - x_k\| + \|w_k - x_k\| \to 0.$$

Hence

$$\lim_{k\to\infty} \|J_\lambda^A(w_k) - w_k\| = \lim_{k\to\infty} \|y_{k+1} - w_k\| = 0. \tag{19}$$

Assume that $\{y_{k_j}\} \subset \{y_k\}$ such that $y_{k_j} \rightharpoonup v$, $j \to \infty$. We show that $v \in A^{-1}(\mathbf{0})$. From $x_k - y_k = \delta(w_{k-1} - y_k)$, we have that $(x_k - y_k) \to 0$. Thus,

$$\|y_k - w_k\| \leq \|x_k - y_k\| + \|w_k - x_k\| \to 0.$$

Hence, $w_{k_j} \rightharpoonup v, j \to \infty$. Thus, $J_\lambda^A(w_{k_j}) \rightharpoonup v, j \to \infty$ by (19). By $J_\lambda^A(w_k) = (I + \lambda A)^{-1} w_k$, we obtain

$$\frac{1}{\lambda}\left(w_{k_j} - J_\lambda^A(w_{k_j})\right) \in A(J_\lambda^A(w_{k_j})).$$

Monotonicity property of A implies

$$\langle x - J_\lambda^A(w_{k_j}), w - \frac{1}{\lambda}\left(w_{k_j} - J_\lambda^A(w_{k_j})\right)\rangle \geq 0, \quad \forall\, x,\ w \text{ satisfying } w \in A(x), \quad \forall\, k_j. \tag{20}$$

Using (19) in (20), we have $\langle x - v, w\rangle \geq 0$, $\forall\ x,\ w$ satisfying $w \in A(x)$. Then, A being maximal monotone gives $v \in A^{-1}(\mathbf{0})$ (see, for example, [5]).

Now, let us set

$$b_k^* = 2\langle x_{k-1} - x_k, x_k - x^*\rangle + \|x_{k-1} - x_k\|^2,$$
$$d_k^* = -\frac{\delta(1-\alpha)}{(1-\delta)^2}\|x_k - x_{k-1}\|^2.$$

Notice that

$$\|x_{k-1} - x^*\|^2 = \|x_{k-1} - x_k\|^2 + 2\langle x_{k-1} - x_k, x_k - x^*\rangle + \|x_k - x^*\|^2.$$

Consequently, we have that

$$\frac{(1-\alpha)}{1-\delta}\|x_k - x^*\|^2 = a_k + d_k^* + \frac{\alpha}{1-\delta} b_k^*. \tag{21}$$

Since $\lim_{k\to\infty} \|x_{k-1} - x_k\| = 0$, and $\{x_k\}$ is bounded, we have that $d_k^* \to 0$, and $b_k^* \to 0$. Also, since $\lim_{k\to\infty} a_k$ exists, it follows from (21) that

$$\lim_{k\to\infty} \|x_k - x^*\| \quad \text{exists for each } x^* \in A^{-1}(\mathbf{0}). \tag{22}$$

We then obtain from

$$\|y_k - x^*\|^2 = \|y_k - x_k\|^2 + 2\langle y_k - x_k, x_k - x^*\rangle + \|x_k - x^*\|^2$$

that

$$\lim_{k\to\infty} \|y_k - x^*\| \quad \text{exists for each } x^* \in A^{-1}(\mathbf{0}). \tag{23}$$

By Lemma 2.3, we conclude that $\{y_k\}$ converges weakly to a point in $A^{-1}(\mathbf{0})$. The proof is complete. ∎

3.2. Linear convergence

In this subsection, we study the linear rate of convergence of the proposed Algorithm 1. To this end, we give the following assumption (please see [36] for more details about this assumption).

Assumption 3.6 ([36, Section 3.2]): Assume that A^{-1} is Lipschitz continuous at $\mathbf{0}$ with modulus $a > 0$, i.e. (i) $A^{-1}(\mathbf{0}) = \{x^*\}$ is a singleton, where $x^* \in H$ and (ii) for some $\tau > 0$, it holds that

$$\|x - x^*\| \leq a\|w\|$$

whenever $x \in A^{-1}(w)$ and $\|w\| \leq \tau$.

Theorem 3.7: *Suppose $A : H \to 2^H$ is a maximal monotone operator with $A^{-1}(\mathbf{0}) \neq \emptyset$ and Assumption 3.6 is fulfilled. Let $\{y_k\}$ be generated by Algorithm 1. Then $\{y_k\}$ converges R-linearly to a point $x^* \in A^{-1}(\mathbf{0})$ when $\alpha = 0$ and $0 < \delta < 1$.*

Proof: By [36, Theorem 3.5], we have, under Assumption 3.6, that

$$\begin{aligned} \|y_{k+1} - x^*\|^2 \leq & \left(1 - \frac{\lambda^2}{a^2 + \lambda^2}\right) \|w_k - x^*\|^2 \\ =& \mu_1 \|w_k - x^*\|^2, \end{aligned} \tag{24}$$

where $\mu_1 := 1 - \frac{\lambda^2}{a^2+\lambda^2} \in (0, 1)$. Since $y_{k+1} = \frac{1}{1-\delta} x_{k+1} - \frac{\delta}{1-\delta} w_k$, we obtain from (10) that

$$\begin{aligned} \|y_{k+1} - x^*\|^2 =& \frac{1}{1-\delta}\|x_{k+1} - x^*\|^2 - \frac{\delta}{1-\delta}\|w_k - x^*\|^2 \\ &+ \frac{\delta}{(1-\delta)^2}\|x_{k+1} - w_k\|^2 \\ =& \frac{1}{1-\delta}\|x_{k+1} - x^*\|^2 - \frac{\delta}{1-\delta}\|w_k - x^*\|^2 + \delta\|y_{k+1} - w_k\|^2. \end{aligned} \tag{25}$$

If we combine (24) and (25), we have

$$\frac{1}{1-\delta}\|x_{k+1} - x^*\|^2 - \frac{\delta}{1-\delta}\|w_k - x^*\|^2 + \delta\|y_{k+1} - w_k\|^2 \leq \mu_1\|w_k - x^*\|^2.$$

Thus (noting that $\alpha = 0$), we obtain

$$\begin{aligned} \frac{1}{1-\delta}\|w_{k+1} - x^*\|^2 =& \frac{1}{1-\delta}\|x_{k+1} - x^*\|^2 \leq \left(\mu_1 + \frac{\delta}{1-\delta}\right)\|w_k - x^*\|^2 - \delta\|y_{k+1} - w_k\|^2 \\ \leq& (\mu_1(1-\delta) + \delta)\left(\frac{1}{1-\delta}\|w_k - x^*\|^2\right). \end{aligned} \tag{26}$$

Therefore, $\{w_k\}$ converges R-linearly to x^*. Also, it is easily seen that $\{y_k\}$ converges R-linearly to x^*. This completes the proof. ∎

Remark 3.8: It is worthy of note that the results in this section can easily be carried over to the over-relaxation case. Thus, we can obtain similar theorems to Theorems 3.5 and 3.7 when we replace $y_{k+1} = J_\lambda^A(w_k)$ in (8) of Algorithm 1 with $y_{k+1} = (1-\phi)w_k + \phi J_\lambda^A(w_k), \phi \in (0, 2)$. Consequently, we will have a replacement of (9) as

$$\|y_{k+1} - x^*\|^2 \leq \|w_k - x^*\|^2 - \frac{(2-\phi)}{\phi}\|y_{k+1} - w_k\|^2.$$

Then the rest of the proof of Lemma 3.4 follows.

3.3. Strong convergence

We now design a strongly convergent version of Algorithm 1. To this end, a mapping $g : H \to H$ is a strict contraction if there exists $\mu \in [0, 1)$ such that $\|g(x) - g(y)\| \le \mu\|x - y\|, \forall\, x, y \in H$.

We now propose our second algorithm below.

Algorithm 2 (Strongly Convergent Accelerated PPA)

1: Choose $\delta \in (0, 1), \beta_k \in (0, 1)$ and pick $w_{-1} = y_0 \in H$ arbitrarily, $\lambda > 0$ and set $k = 0$.

2: Given w_{k-1} and y_k, compute

$$\begin{cases} w_k = y_k + \delta(\beta_{k-1} g(w_{k-1}) + (1 - \beta_{k-1})w_{k-1} - y_k), \\ y_{k+1} = J_\lambda^A(\beta_k g(w_k) + (1 - \beta_k)w_k). \end{cases} \tag{27}$$

3: Set $k \leftarrow k + 1$, and **go to Step 2**.

Remark 3.9: Algorithm 2 involves a viscosity approximation and correction term. As far as we know, this is the first time a strongly convergent PPA with viscosity approximation and correction term is designed for monotone inclusion problem (1).

We show that the sequences $\{w_k\}$ and $\{y_k\}$ generated by Algorithm 2 are bounded.

Lemma 3.10: *Suppose $A : H \to 2^H$ is a maximal monotone operator with $A^{-1}(\mathbf{0}) \neq \emptyset$. Let $\{y_k\}$ be generated by Algorithm 2 with $0 < \delta < 1$. Then $\{w_k\}$ and $\{y_k\}$ are bounded if g is a strict contraction with constant $\mu \in [0, \frac{1}{2})$ and $\lim_{k\to\infty} \beta_k = 0$.*

Proof: Let $u_k := \beta_k g(w_k) + (1 - \beta_k)w_k$. Pick $x^* \in A^{-1}(\mathbf{0})$. If we use similar arguments to establish (9), we have

$$\|y_{k+1} - x^*\|^2 \le \|u_k - x^*\|^2 - \|y_{k+1} - u_k\|^2. \tag{28}$$

We have from $w_k = y_k + \delta(u_{k-1} - y_k)$ that

$$y_{k+1} = \frac{1}{1-\delta} w_{k+1} - \frac{\delta}{1-\delta} u_k. \tag{29}$$

From (29), we obtain that

$$\begin{aligned} \|y_{k+1} - x^*\|^2 &= \|\frac{1}{1-\delta} w_{k+1} - \frac{\delta}{1-\delta} u_k - x^*\|^2 \\ &= \|\frac{1}{1-\delta}(w_{k+1} - x^*) - \frac{\delta}{1-\delta}(u_k - x^*)\|^2 \\ &= \frac{1}{1-\delta}\|w_{k+1} - x^*\|^2 - \frac{\delta}{1-\delta}\|u_k - x^*\|^2 + \frac{\delta}{(1-\delta)^2}\|w_{k+1} - u_k\|^2 \\ &= \frac{1}{1-\delta}\|w_{k+1} - x^*\|^2 - \frac{\delta}{1-\delta}\|u_k - x^*\|^2 + \delta\|y_{k+1} - u_k\|^2. \end{aligned} \tag{30}$$

Plugging (30) in (28) gives

$$\frac{1}{1-\delta}\|w_{k+1} - x^*\|^2 - \frac{\delta}{1-\delta}\|u_k - x^*\|^2 + \delta\|y_{k+1} - u_k\|^2$$

$$\leq \|u_k - x^*\|^2 - \|y_{k+1} - u_k\|^2. \tag{31}$$

Also, we get

$$\begin{aligned}
\|u_k - x^*\|^2 =& \|\beta_k g(w_k) + (1-\beta_k)w_k - x^*\|^2 = \|(w_k - x^*) - \beta_k(w_k - g(w_k)\|^2 \\
=& \|w_k - x^*\|^2 + \beta_k^2\|w_k - g(w_k)\|^2 - 2\beta_k\langle w_k - x^*, w_k - g(w_k)\rangle \\
=& \|w_k - x^*\|^2 + \beta_k^2\|w_k - g(w_k)\|^2 - \beta_k\|w_k - g(w_k)\|^2 \\
& - \beta_k\|w_k - x^*\|^2 + \beta_k\|g(w_k) - x^*\|^2 \\
=& (1-\beta_k)\|w_k - x^*\|^2 + \beta_k^2\|w_k - g(w_k)\|^2 - \beta_k\|w_k - g(w_k)\|^2 \\
& + \beta_k\|g(w_k) - x^*\|^2.
\end{aligned} \tag{32}$$

By replacing x^* with y_{k+1}, we have from (32) that

$$\begin{aligned}
\|y_{k+1} - u_k\|^2 =& \|u_k - y_{k+1}\|^2 \\
=& (1-\beta_k)\|w_k - y_{k+1}\|^2 + \beta_k^2\|w_k - g(w_k)\|^2 - \beta_k\|w_k - g(w_k)\|^2 \\
& + \beta_k\|y_{k+1} - g(w_k)\|^2.
\end{aligned} \tag{33}$$

Plugging (32) and (33) into (31) gives (noting that g is a strict contraction)

$$\begin{aligned}
&\frac{1}{1-\delta}\|w_{k+1} - x^*\|^2 \leq \frac{1}{1-\delta}\|u_k - x^*\|^2 - (1+\delta)\|y_{k+1} - u_k\|^2 \\
&= \frac{1}{1-\delta}\left((1-\beta_k)\|w_k - x^*\|^2 + \beta_k^2\|w_k - g(w_k)\|^2 - \beta_k\|w_k - g(w_k)\|^2\right. \\
&\quad \left. + \beta_k\|g(w_k) - x^*\|^2\right) - (1+\delta)\left((1-\beta_k)\|w_k - y_{k+1}\|^2 + \beta_k^2\|w_k - g(w_k)\|^2\right. \\
&\quad \left. - \beta_k\|w_k - g(w_k)\|^2 + \beta_k\|y_{k+1} - g(w_k)\|^2\right) \\
&\leq \frac{1}{1-\delta}\left((1-\beta_k)\|w_k - x^*\|^2 + \beta_k^2\|w_k - g(w_k)\|^2 - \beta_k\|w_k - g(w_k)\|^2\right. \\
&\quad \left. + 2\,\beta_k\|g(w_k) - g(x^*)\|^2 + 2\beta_k\|g(x^*) - x^*\|^2\right) \\
&\quad - (1+\delta)\left((1-\beta_k)\|w_k - y_{k+1}\|^2 + \beta_k^2\|w_k - g(w_k)\|^2 - \beta_k\|w_k - g(w_k)\|^2\right. \\
&\quad \left. + \beta_k\|y_{k+1} - g(w_k)\|^2\right) \\
&\leq \frac{(1-\beta_k(1-2\mu))}{1-\delta}\|w_k - x^*\|^2 - (1+\delta)(1-\beta_k)\|w_k - y_{k+1}\|^2 \\
&\quad - (1+\delta)\beta_k\|y_{k+1} - g(w_k)\|^2 - \frac{\delta^2\beta_k(1-\beta_k)}{1-\delta}\|w_k - g(w_k)\|^2 \\
&\quad + \frac{\beta_k(1-2\mu)}{1-\delta}\frac{2\|g(x^*) - x^*\|^2}{(1-2\mu)}.
\end{aligned} \tag{34}$$

Therefore,

$$\begin{aligned}
\frac{1}{1-\delta}\|w_{k+1} - x^*\|^2 \leq& \frac{(1-\beta_k(1-2\mu)}{1-\delta}\|w_k - x^*\|^2 \\
& + \frac{\beta_k(1-2\mu)}{1-\delta}\frac{2\|g(x^*) - x^*\|^2}{(1-2\mu)}.
\end{aligned} \tag{35}$$

Take

$$t_k := \frac{1}{1-\delta}\|w_k - x^*\|^2 \geq 0.$$

Then

$$t_{k+1} \leq (1-\beta_k(1-2\mu))t_k + \frac{\beta_k(1-2\mu)}{1-\delta}\frac{2\|g(x^*)-x^*\|^2}{(1-2\mu)}.$$

By Lemma 2.5 we have that $\{t_k\}$ is bounded. It then follows that both $\{w_k\}$ and $\{s_k\}$ are bounded. ■

We now give the following strong convergence result.

Theorem 3.11: *Suppose $A : H \to 2^H$ is a maximal monotone operator with $A^{-1}(\mathbf{0}) \neq \emptyset$ with g is a strict contraction with constant $k \in [0, \frac{1}{2})$, $\lim_{k\to\infty} \beta_k = 0$, $\sum_{k=1}^{\infty} \beta_k = \infty$. Then the sequences $\{w_k\}$ and $\{s_k\}$ generated by Algorithm 2 converge strongly to $x^* = P_{A^{-1}(\mathbf{0})}g(x^*)$, where $P_{A^{-1}(\mathbf{0})}$ is the metric projection onto $A^{-1}(\mathbf{0})$.*

Proof: First observe that there exists a unique x^* such that $x^* = P_{A^{-1}(\mathbf{0})}g(x^*)$ due to contraction property of $P_{A^{-1}(\mathbf{0})}g$. Thus, we have

$$\begin{aligned}
\|u_k - x^*\|^2 =& \|\beta_k g(w_k) + (1-\beta_k)w_k - x^*\|^2 \\
=& \|\beta_k(g(w_k) - x^*) + (1-\beta_k)(w_k - x^*)\|^2 \\
=& \beta_k^2\|g(w_k) - x^*\|^2 + (1-\beta_k)^2\|w_k - x^*\|^2 \\
& + 2\beta_k(1-\beta_k)\langle g(w_k) - x^*, w_k - x^*\rangle \\
=& (1-\beta_k)^2\|w_k - x^*\|^2 + \beta_k^2\|g(w_k) - x^*\|^2 \\
& + 2\beta_k(1-\beta_k)\langle g(w_k) - g(x^*), w_k - x^*\rangle \\
& + 2\beta_k(1-\beta_k)\langle g(x^*) - x^*, w_k - x^*\rangle \\
\leq& (1-\beta_k)^2\|w_k - x^*\|^2 + \beta_k^2\|g(w_k) - x^*\|^2 \\
& + 2\beta_k(1-\beta_k)\mu\|w_k - x^*\|^2 + 2\beta_k(1-\beta_k)\langle g(x^*) - x^*, w_k - x^*\rangle \\
\leq& (1-\beta_k(1-2\mu))\|w_k - x^*\|^2 + \beta_k^2\|g(w_k) - x^*\|^2 \\
& + \beta_k^2\|w_k - x^*\|^2 + 2\beta_k(1-\beta_k)\langle g(x^*) - x^*, w_k - x^*\rangle.
\end{aligned} \tag{36}$$

Similarly,

$$\begin{aligned}
\|u_k - y_{k+1}\|^2 =& \beta_k^2\|g(w_k) - y_{k+1}\|^2 + (1-\beta_k)^2\|w_k - y_{k+1}\|^2 \\
& + 2\beta_k(1-\beta_k)\langle g(w_k) - y_{k+1}, w_k - y_{k+1}\rangle \\
\geq& \beta_k^2\|g(w_k) - y_{k+1}\|^2 + (1-\beta_k)^2\|w_k - y_{k+1}\|^2 \\
& - 2\beta_k(1-\beta_k)\|g(w_k) - y_{k+1}\|\|w_k - y_{k+1}\| \\
\geq& \beta_k^2\|g(w_k) - y_{k+1}\|^2 + (1-\beta_k)^2\|w_k - y_{k+1}\|^2 \\
& - 2\beta_k(1-\beta_k)M\|w_k - y_{k+1}\|,
\end{aligned} \tag{37}$$

where $M := \sup\|g(w_k) - y_{k+1}\| < \infty$, since $\{w_k\}$ and $\{s_k\}$ are bounded by Lemma 3.10. Plugging (36) and (37) into (31), we have

$$\frac{1}{1-\delta}\|w_{k+1} - x^*\|^2 \leq \frac{1}{1-\delta}\|u_k - x^*\|^2 - (1+\delta)\|y_{k+1} - u_k\|^2$$

$$
\begin{aligned}
&\le \frac{(1-\beta_k(1-2\mu))}{1-\delta}\|w_k - x^*\|^2 - (1+\delta)\beta_k^2\|g(w_k) - y_{k+1}\|^2 \\
&\quad + \beta_k\left(\frac{\beta_k}{1-\delta}\|w_k - x^*\|^2 + 2(1-\beta_k)\langle g(x^*) - x^*, w_k - x^*\rangle\right. \\
&\quad \left. + \frac{\beta_k}{1-\delta}\|g(w_k) - x^*\|^2 + 2(1+\delta)(1-\beta_k)M\|w_k - y_{k+1}\|\right) \\
&\quad - (1-\beta_k)^2(1+\delta)\|w_k - y_{k+1}\|^2.
\end{aligned} \tag{38}
$$

Therefore,

$$
\begin{aligned}
t_{k+1} \le& (1-\beta_k(1-2\mu))t_k + \beta_k(1-2\mu)h_k - (1+\delta)\beta_k^2\|g(w_k) - y_{k+1}\|^2 \\
&- (1-\beta_k)^2(1+\delta)\|w_k - y_{k+1}\|^2,
\end{aligned} \tag{39}
$$

where

$$
\begin{aligned}
h_k := \frac{1}{1-2\mu}&\left(\frac{\beta_k}{1-\delta}\|w_k - x^*\|^2 + 2(1-\beta_k)\langle g(x^*) - x^*, w_k - x^*\rangle\right. \\
&\left. + \frac{\beta_k}{1-\delta}\|g(w_k) - x^*\|^2 + 2(1+\delta)(1-\beta_k)M\|w_k - y_{k+1}\|\right).
\end{aligned}
$$

To conclude the proof, it suffices to show in view of Lemma 2.4 that $\limsup_{i\to\infty} h_{k_i} \le 0$ for each subsequence $\{t_{k_i}\}$ of $\{t_k\}$ such that $\liminf_{i\to\infty}(t_{k_i+1} - t_{k_i}) \ge 0$. To this end, let $\{t_{k_i}\}$ be a subsequence of $\{t_k\}$ such that $\liminf_{i\to\infty}(t_{k_i+1} - t_{k_i}) \ge 0$. Then by (39), we have

$$
\begin{aligned}
\limsup_{i\to\infty}(1-\beta_{k_i})^2(1+\delta)\|w_{k_i} - y_{k_i+1}\|^2 &\le \limsup_{i\to\infty}\left[(t_{k_i} - t_{k_i+1}) + \beta_{k_i}(h_{k_i} - t_{k_i})\right] \\
&= -\liminf_{i\to\infty}(t_{k_i+1} - t_{k_i}) \le 0.
\end{aligned}
$$

Since $\lim_{i\to\infty}(1-\beta_{k_i})^2(1+\delta) = 1+\delta > 0$, we get

$$
\lim_{i\to\infty}\|w_{k_i} - y_{k_i+1}\| = 0.
$$

We have from (27) that

$$
\delta\|u_{k_i-1} - y_{k_i}\| = \|w_{k_i} - y_{k_i}\| \to 0, \quad i \to \infty
$$

and from the definition of u_k that

$$
\|u_{k_i} - w_{k_i}\| = \beta_{k_i}\|g(w_{k_i}) - w_{k_i}\| \to 0, \quad i \to \infty.
$$

Consequently,

$$
\|w_{k_i-1} - y_{k_i}\| \le \|w_{k_i-1} - u_{k_i-1}\| + \|u_{k_i-1} - y_{k_i}\| \to 0, \quad i \to \infty. \tag{40}
$$

By Lemma 3.10 that $\{w_k\}$ and $\{y_k\}$ are bounded, we can choose a subsequence $\{w_{k_{i_j}}\}$ of $\{w_{k_i}\}$ which converges weakly to some $w^* \in H$ such that

$$
\limsup_{i\to\infty}\langle g(x^*) - x^*, w_{k_i} - x^*\rangle = \lim_{j\to\infty}\langle g(x^*) - x^*, w_{k_{i_j}} - x^*\rangle = \langle g(x^*) - x^*, w^* - x^*\rangle. \tag{41}
$$

By similar arguments in Theorem 3.5, we can show that $w^* \in A^{-1}(\mathbf{0})$.

Since $x^* = P_{\mathcal{S}}g(x^*)$, it follows from (41) and the characterization of the metric projection that

$$\limsup_{i\to\infty}\langle g(x^*) - x^*, w_{k_i} - x^*\rangle = \langle g(x^*) - x^*, w^* - x^*\rangle \le 0. \tag{42}$$

Using (40), (42) and the condition $\lim_{i\to\infty}\beta_{k_i} = 0$, we find that $\limsup_{i\to\infty} h_{p_i} \le 0$. Thus, in view of the condition $\sum_{k=1}^{\infty}\beta_k = \infty$ and Lemma 2.4, it follows from (39) that $\lim_{k\to\infty} t_k = 0$. By $t_k \ge 0$, we conclude that $\{w_k\}$ converges strongly to $x^* = P_{A^{-1}(\mathbf{0})}g(x^*)$, as asserted. ■

4. Applications

We give some applications of our results in Section 3 to mixed variational inequalities, quasi-convex minimization problems and an adaptation of our proposed Algorithm 1 to Douglas–Rachford splitting algorithm in this section.

4.1. Application to mixed variational inequalities

Let us apply our Algorithm 1 to solve the mixed variational inequality problem in $\mathbb{R}^n$. Let $C \subset \mathbb{R}^n$ be a closed and convex set, $\theta : \mathbb{R}^n \to \mathbb{R}$ be a closed convex function, and $F : \mathbb{R}^n \to \mathbb{R}^n$ be a monotone mapping. Let us consider the following mixed variational inequality problem: find $x^* \in C$ such that

$$\theta(x) - \theta(x^*) + \langle x - x^*, F(x^*)\rangle \ge 0, \quad \forall\, x \in C. \tag{43}$$

The mixed variational inequality problem (43) was also studied in [37] (see also [38] for the nonconvex case). We take G to be a symmetric positive semidefinite matrix of order n. A symmetric matrix G (i.e. $G = G^T$, where G^T means the transpose of G) is positive semidefinite if and only if its eigenvalues are nonnegative. Thus, to construct such a matrix G, it is easy to construct a matrix that is equal to its transpose and whose eigenvalues are nonnegative. For instance, for $n = 3$, $G = \begin{pmatrix} 4 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$ is a symmetric positive semidefinite matrix since $G = G^T$ and its eigenvalues 4.5616, 0.4384, 3 are nonnegative. For computational implementation, such matrix can be generated randomly as described in Appendix.

Under the fulfillment of Assumption 3.1, we propose the following version of Algorithm 1 to solve the mixed variational inequality problem (43) above.

Algorithm 3 (Accelerated PPA for mixed variational inequality problem)

1: Choose α and δ with Assumption 3.1 satisfied. Pick $w_{-1} = w_{-2}, y_0, y_{-1} \in \mathbb{R}^n$ arbitrarily, $\lambda > 0$ and set $k = 0$.

2: Given $w_{k-1}, w_{k-2}, y_{k-1}$ and y_k, find y_{k+1} such that

$$\begin{cases} w_k = y_k + \alpha(y_k - y_{k-1}) + \delta(1+\alpha)(w_{k-1} - y_k) \\ \qquad - \alpha\delta(w_{k-2} - y_{k-1}), \\ \theta(x) - \theta(y_{k+1}) + \langle x - y_{k+1}, F(y_{k+1}) + \frac{1}{\lambda}G(y_{k+1} - w_k)\rangle \ge 0, \quad \forall\, x \in C. \end{cases} \tag{44}$$

3: Set $k \leftarrow k+1$, and **go to Step 2**.

We shall denote the sets of symmetric, symmetric positive semidefinite, and symmetric positive definite matrices of order n S^n, S^n_+, and S^n_{++}. Given matrix $A \in S^n_+$ and vectors $u, v \in \mathbb{R}^n$, we let

$\langle u, v\rangle_A := u^T Av$ and $\|u\|_A := \sqrt{\langle u, u\rangle_A}$. We denote the Frobenius norm by $\|u\|_F$ and the spectral radius of a square matrix M by $\rho(M)$.

Similar to what was given in [23], we assume the following.

Assumption 4.1: (B1) The set of solutions of (43), denoted by S, is nonempty;
(B2) The mapping F is H-monotone in the sense that

$$\langle x - y, F(x) - F(y)\rangle \geq \|x - y\|_H^2, \quad \forall\, x,\, y \in \mathbb{R}^n \tag{45}$$

where $H \in S_+^n$. We observe that $H = 0$ if F is monotone, and $H \in S_{++}^n$ if F F is strongly monotone.
(B3) $M := G + H \in S_{++}^n$.

Remark 4.2: By Assumption 4.1(B2) and (B3), we have that y_{k+1} is uniquely determined in Algorithm 3. Hence, (44) is well-defined. Furthermore, Algorithm 3 reduces to Algorithm 1 if $G \in S_{++}^n$.

Theorem 4.3: *Suppose Assumption* 3.1 *and Assumption* 4.1 *hold. Then* $\{y_k\}$ *generated by Algorithm* 3 *converges to a point in S.*

Proof: Following the same line of arguments in Lemma 3.4, Theorem 3.5 and [23, Theorem 2], we can show that

$$\lim_{k\to\infty} \|y_k - w_k\|_G = 0, \quad \lim_{k\to\infty} \|y_{k+1} - w_k\|_G = 0.$$

The convergence of $\{y_k\}$ to a point in S follows from the same arguments in the proof of [23, Theorem 1]. ∎

Remark 4.4: The convergence results in Theorem 4.3 are extensions of the results in [23]. In Algorithm 3 has inertial term and correction term while [23, (2.2a)–(2.2b)] has only the inertial term.

4.2. Application to strongly quasi-convex minimization

We adapt our Algorithm 1 to solve an optimization problem involving minimization of a strongly quasiconvex function in a finite dimensional linear subspace of $\mathbb{R}^n$ and a quasiconvex objective function over closed and convex subset of $\mathbb{R}^n$. Our results are motivated by the results in [39].

Suppose C is a closed and convex subset of $\mathbb{R}^n$, and $h : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is a proper function such that $C \cap \operatorname{dom} h \neq \emptyset$. Let us consider the constrained optimization problem

$$\min_{x\in C} h(x). \tag{46}$$

Similar to [39, Section 3], we give the following assumption.

Assumption 4.5: (D1) $h : C \to \mathbb{R}$ is a continuous and strongly quasiconvex function with modulus $\gamma > 0$;
(D2) $h : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is a proper, continuous and quasiconvex function on $C \subseteq \operatorname{dom} h$;
(D3) h is 2-weakly coercive on C, i.e.

$$\liminf_{x\in C, \|x\|\to\infty} \frac{h(x)}{\|x\|^2} \geq 0.$$

We give our algorithm to solve problem (46) as follows.

Algorithm 4 (Accelerated PPA for Quasi-Convex)

1: Choose α and δ with Assumption 3.1 satisfied. Pick $w_{-1} = w_{-2}, y_0, y_{-1} \in \mathbb{R}^n$ arbitrarily, $\lambda > 0$ and set $k = 0$.

2: Given $w_{k-1}, w_{k-2}, y_{k-1}$ and y_k, find y_{k+1} such that

$$\begin{cases} w_k = y_k + \alpha(y_k - y_{k-1}) + \delta(1+\alpha)(w_{k-1} - y_k) \\ \qquad - \alpha\delta(w_{k-2} - y_{k-1}), \\ y_{k+1} \in \operatorname{prox}_{\lambda h}(C, w_k). \end{cases} \tag{47}$$

3: Set $k \leftarrow k + 1$, and **go to Step 2**.

By applying similar ideas in Lemma 3.4, Theorem 3.5 and [39, Sections 3 and 4], we can prove the following theorems.

Theorem 4.6: *Let $C \subseteq \mathbb{R}^n$ be a linear subspace, $h : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a function such that h satisfies Assumption 4.5(D1) and suppose Assumption 3.1 holds. Then $\{y_k\}$ generated by Algorithm 4 converges to $x^* = \operatorname{argmin}_C h$ and $\lim_{k\to\infty} h(y_k) = \min_C h$.*

Remark 4.7: Under assumption (D1), $\operatorname{argmin}_C h$ is singleton (see [30, Corollary 3], where the unique minimizer x^* of h is established).

Theorem 4.8: *Let $C \subseteq \mathbb{R}^n$ be a linear subspace, $h : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a function such that h satisfies Assumption 4.5(D2)–(D3), $\Omega := \{x \in C : h(x) \le h(y_k),\ \forall\ k \in \mathbb{N}\} \neq \emptyset$ and suppose Assumption 3.1 holds, h is bounded from below, and $\{y_k\}$ generated by Algorithm 4. Then the sequence $\{h(y_k)\}$ is convergent.*

Remark 4.9: The requirement for h to be bounded from below under assumptions (D2) and (D3) was also discussed in [39, Remark 3.2(ii)].

4.3. Application to Douglas–Rachford splitting algorithm

This section presents an application of Algorithm 1 to the Douglas–Rachford splitting method for finding zeros of an operator T such that T is the sum of two maximal monotone operators, i.e. $T = A + B$ with $A, B : H \to 2^H$ being maximal monotone multi-functions on a Hilbert space H. The method was originally introduced in [8] in a finite-dimensional setting, its extension to maximal monotone mappings in Hilbert spaces can be found in [9].

The Douglas–Rachford splitting method [8,9] iteratively applies the operator

$$G_{\lambda,A,B} := J_\lambda^A o(2J_\lambda^B - I) + (I - J_\lambda^B) \tag{48}$$

and has found to be effective in many applications including ADMM. In [1, Theorem 4], the Douglas–Rachford operator (48) was found to be a resolvent $J^M_{\lambda,A,B}$ of a maximal monotone operator

$$M_{\lambda,A,B} := G^{-1}_{\lambda,A,B} - I. \tag{49}$$

Therefore, the Douglas–Rachford splitting method is a special case of the proximal point method (with $\lambda = 1$), written as

$$y_{k+1} = J^M_{\lambda,A,B}(y_k) = G_{\lambda,A,B}(y_k). \tag{50}$$

Hence, we can apply our Algorithm 1 to the Douglas–Rachford splitting method as given below:

Algorithm 5 Accelerated Douglas–Rachford Splitting Algorithm

1: Choose α and δ with Assumption 3.1 satisfied. Pick $w_{-1} = w_{-2}, y_0, y_{-1} \in H$ arbitrarily, $\lambda > 0$ and set $k = 0$.

2: Given $w_{k-1}, w_{k-2}, y_{k-1}$ and y_k, compute

$$\begin{cases} w_k & = y_k + \alpha(y_k - y_{k-1}) + \delta(1+\alpha)(w_{k-1} - y_k) \\ & \quad - \alpha\delta(w_{k-2} - y_{k-1}), \\ y_{k+1} & = G_{\lambda,A,B}(w_k). \end{cases} \tag{51}$$

3: Set $k \leftarrow k + 1$, and **go to 2**.

Table 1. Comparison of algorithms with $\epsilon = 10^{-5}$.

	$n = 20$		$n = 50$		$n = 100$		$n = 500$	
Algorithms	CPU	Iter	CPU	Iter	CPU	Iter	CPU	Iter
Algorithm 1	0.0120	119	0.0412	41	0.2921	34	0.8054	37
Maingé Alg.	0.0719	1583	0.0996	2280	1.4861	2650	50.9385	4390
Kim Alg.	0.0641	136	0.1513	144	0.9183	148	1.9598	157

Table 2. Comparison of algorithms with $\epsilon = 10^{-5}$.

	$n = 20$		$n = 50$		$n = 100$		$n = 500$	
Algorithms	CPU	Iter	CPU	Iter	CPU	Iter	CPU	Iter
Algorithm 1	0.0053	48	0.1258	51	0.1239	55	0.1869	77
Maingé Alg.	0.0217	1896	1.0987	2219	1.0341	2782	1.1756	5104
Kim Alg.	0.0139	122	1.0632	117	1.0347	111	1.1356	106
GLM1 Alg 1.	0.0120	53	1.0119	55	1.0339	57	1.1267	62
GLM2 Alg 1.	0.0101	50	0.0493	51	1.0309	53	1.1329	57

5. Numerical illustrations

Let us demonstrate the numerical implementations of Algorithm 1 studied in Section 3. In particular, we give numerical comparisons with the algorithm in [29], denoted as Maingé Alg., the algorithm in [26], denoted as Kim Alg., Algorithm 1 in [40], denoted as GLM1 Algorithm 1, and Algorithm 1 in [39], denoted as GLM2 Algorithm 1. We write our codes in Matlab 2016 (b) which is running on a personal computer with an Intel(R) Core(TM) i5-2600 CPU at 2.30 GHz and 8.00 Gb-RAM.

Example 5.1: Let $B, D \in \mathbb{R}^{n\times n}$ be symmetric and positive definite matrices. Then, we consider the mixed variational inequality problem (43), where $\theta(x) = x^t Bx$ and $\langle x - x^*, F(x^*)\rangle = x^t D(x - x^*)$.

We denote the largest eigenvalue of D by $\rho(D)$ and the smallest eigenvalues of B by $\mu(B)$. By the positive definiteness assumption, we have that both $\rho(D)$ and $\mu(B)$ are positive. Using this problem, we carry out the numerical analysis of Algorithm 1 (or Algorithm 3) in comparison with Maingé Alg. [29] and Kim Alg. [26]. Numerical results for this example are displayed in Table 1 and Figure 1 below.

Example 5.2: This example is taken from [39, Example 5.1]. Suppose $h : \mathbb{R}^n \to \mathbb{R}$ is defined by $h(x) := \max\{h_1(x), h_2(x)\}$, where $h_1(x) = \sqrt{\|x\|}$ and $h_2(x) = \|x\|^2 - q, q \in \mathbb{N}$. Then h is continuous and strongly quasiconvex, but not convex. Also, we have that $\arg\min_{\mathbb{R}^n} h = \{0\}$. Using this problem, we carry out the numerical analysis of Algorithm 1 (or Algorithm 4) in comparison with Maingé Alg. [29] and Kim Alg. [26]. Kindly see our numerical results in Table 2 and Figure 2 for this example.

During the computations, we randomly generate the starting points $w_{-1}, w_{-2}, y_0, y_{-1}$ for $n = 20$, 50, 100, and 500. We choose our parameters as follows:

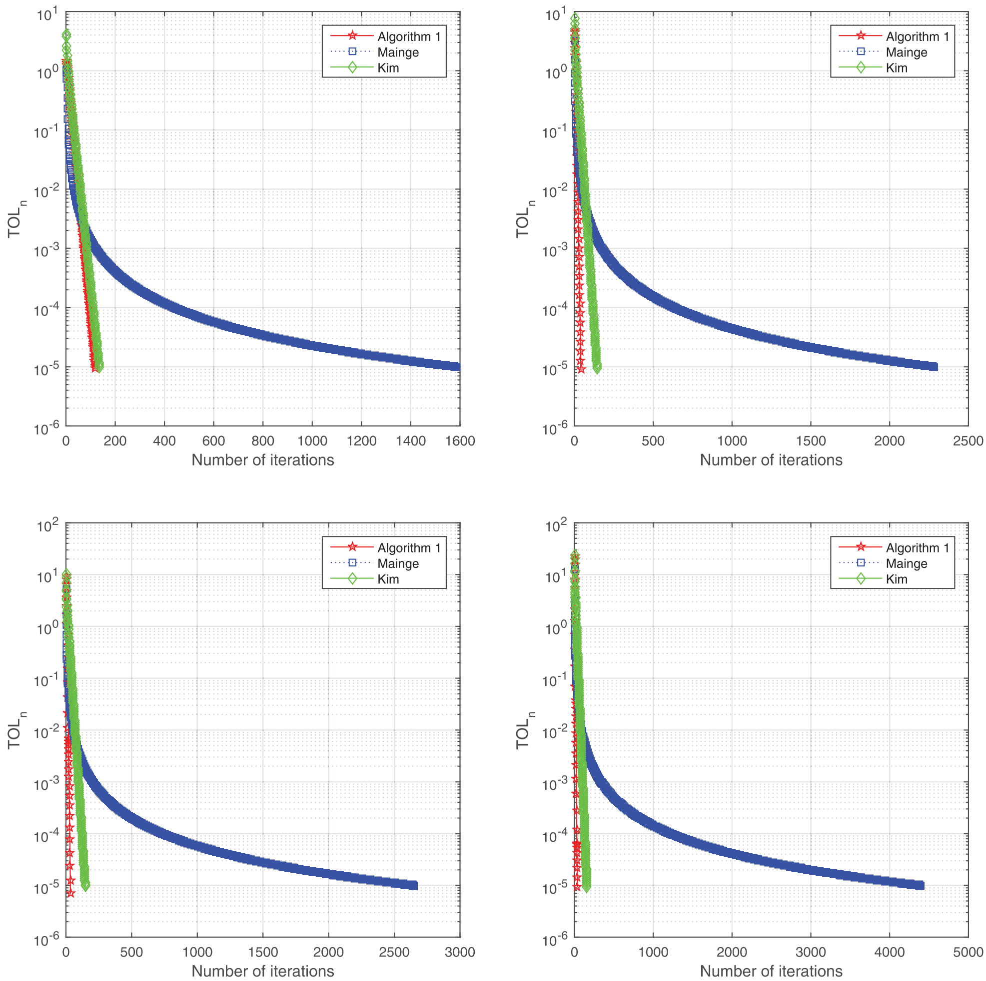

Figure 1. The behavior of TOL$_n$ with $\epsilon = 10^{-5}$ for Example 5.1: Top Left: $\boldsymbol{n} = \mathbf{20}$; Top Right: $\boldsymbol{n} = \mathbf{50}$; Bottom left: $\boldsymbol{n} = \mathbf{100}$; Bottom Right: $\boldsymbol{n} = \mathbf{500}$.

- Algorithm 1: $\alpha = 0.5, \delta = 0.9$ and $\lambda = \frac{0.99}{2\rho(B)}$.
- Maingé Alg. [29]: $a = 1$, $c = 2$, $b = 0.5$, $a_1 = 0.5$, $a_2 = 0.9$, $\bar{c} = 1$ and $\lambda = \frac{0.99}{2\rho(B)}$.
- Kim Alg. [26]: $\lambda = \frac{0.99}{2\rho(B)}$.
- GLM1 Alg 1. [40]: $\theta = 0.3$, $\theta_k = \theta - \frac{1}{5(k+1)^2}$, $\rho' = 0.6$, $\rho'' = 1.4$ and $\rho_k = \frac{1}{k}\rho' + \left(1 - \frac{1}{k}\right)\rho''$.
- GLM2 Alg 1. [39]: $c_1 = 1$, $c_{k+1} = \frac{100}{k^2} + c_k$, $\rho' = 0.9$, $\rho'' = 1.5$, $\rho_k = \frac{1}{k}\rho'' + \left(1 - \frac{1}{k}\right)\rho'$, $\alpha_{k+1} = \alpha_k + \frac{1}{900(k+1)^2}$, $\alpha_1 = 0.1$.

We then use the stopping criterion; $\text{TOL}_n := \|y_{k+1} - w_k\| < \varepsilon$, where ε is the predetermined error.

In the tables below, 'Iter' means the number of iterations. Also, in the tables and figures, Maingé Alg., Kim Alg., GLM1 Alg 1. and GLM2 Alg 1. represent the algorithm in [29], the algorithm in [26], Algorithm 1 in [40] and Algorithm 1 in [39], respectively.

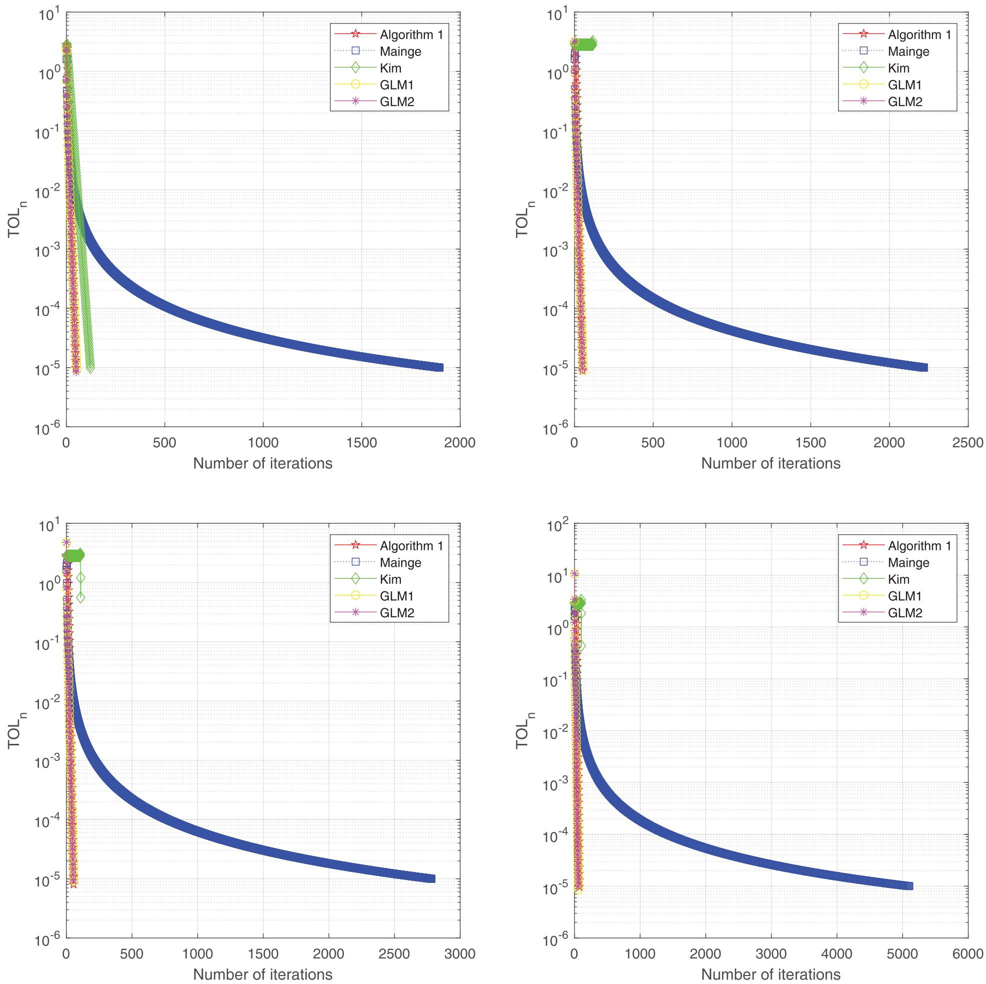

Figure 2. The behavior of TOL$_n$ with $\epsilon = 10^{-5}$ for Example 5.2: Top Left: $\boldsymbol{n} = \mathbf{20}$; Top Right: $\boldsymbol{n} = \mathbf{50}$; Bottom left: $\boldsymbol{n} = \mathbf{100}$; Bottom Right: $\boldsymbol{n} = \mathbf{500}$.

Remark 5.3: We can observe from the numerical results above that our proposed Algorithm 1 indeed accelerates the proximal point algorithm (PPA) and outperform other versions of PPA in the literature. This is due to the constellations of parameters in Algorithm 1 that make our proposed Algorithm 1 faster (in terms of the number of iterations and the CPU time) than the PPA versions in [26,29,39,40].

6. Conclusion

In this paper, we introduced a numerically fast proximal point algorithm which contains one inertial term and two correction terms. We obtained both weak and linear convergence of this proposed algorithm under some standard assumptions. Then a modification of the algorithm is designed and strong convergence results obtained. Applications of our proposed algorithm to mixed variational inequalities, strongly quasi-convex minimization problems and Douglas–Rachford splitting algorithm are given. We analyze our results using some test examples, and the numerical tests show that our proposed algorithm outperforms related state-of-the-art algorithms in the literature.

We leave developing worst-case convergence rate of our proposed algorithm using the performance estimation problem approach as future project.

Acknowledgments

The first author is supported by the DSI-NRF Centre of Excellence in Mathematical and Statistical Sciences (CoE-MaSS), South Africa (Grant Number 2024-032-OPA).

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] Eckstein J, Bertsekas DP. On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. Math Prog. 1992;55:293–318. doi: 10.1007/BF01581204
[2] Martinet B. Régularisation d'inéquations variationnelles par approximations successives[French]. Rev Fr Inf Recherche Opér. 1970;4(R-3):154–158.
[3] Moreau JJ. Proximité et dualité dans un espace Hilbertien[French]. Bull Soc Math France. 1965;93:273–299. doi: 10.24033/bsmf.1625
[4] Moudafi A, Elisabeth E. An approximate inertial proximal method using the enlargement of a maximal monotone operator. Int J Pure Appl Math. 2003;5:283–299.
[5] Rockafellar RT. Monotone operators and the proximal point algorithm. SIAM J Control Optim. 1976;14:877–898. doi: 10.1137/0314056
[6] Hestenes MR. Multiplier and gradient methods. J Optim Theory Appl. 1969;4:303–320. doi: 10.1007/BF00927673
[7] Powell MJD. A method for nonlinear constraints in minimization problems. In: Optimization (Sympos., Univ. Keele, Keele, 1968). London: Academic Press; 1969. p. 283–298.
[8] Douglas J, Rachford HH. On the numerical solution of heat conduction problems in two or three space variables. Trans Amer Math Soc. 1956;82:421–439. doi: 10.1090/tran/1956-082-02
[9] Lions PL, Mercier B. Splitting algorithms for the sum of two nonlinear operators. SIAM J Numer Anal. 1979;16:964–979. doi: 10.1137/0716071
[10] Gabay D, Mercier B. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. Comput Math Appl. 1976;2:17–40. doi: 10.1016/0898-1221(76)90003-1
[11] Glowinski R, Marrocco A. Sur l'approximation, par élḿents finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problémes de Dirichlet non linéaires [French]. Rev Fr Automat Inf Recherche Opér Sér – Rouge Anal Numér. 1975;9(R-2):41–76.
[12] Bruck RE Jr. Asymptotic convergence of nonlinear contraction semigroups in Hilbert space. J Funct Anal. 1975;18:15–26. doi: 10.1016/0022-1236(75)90027-0
[13] Corman E, Yuan X. A generalized proximal point algorithm and its convergence rate. SIAM J Optim. 2014;24:1614–1638. doi: 10.1137/130940402
[14] Gu G, Yang J. A unified and tight linear convergence analysis of the relaxed proximal point algorithm. J Ind Manage Optim. 2023;19:3742–3749. doi: 10.3934/jimo.2022107
[15] Gu G, Yang J. Tight sublinear convergence rate of the proximal point algorithm for maximal monotone inclusion problems. SIAM J Optim. 2020;30:1905–1921. doi: 10.1137/19M1299049
[16] Tao M, Yuan X. The generalized proximal point algorithm with step size 2 is not necessarily convergent. Comput Optim Appl. 2018;70:827–839. doi: 10.1007/s10589-018-9992-3
[17] Alvarez F, Attouch H. An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. Set-Valued Anal. 2001;9:3–11. doi: 10.1023/A:1011253113155
[18] Alvarez F. Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space. SIAM J Optim. 2004;14:773–782. doi: 10.1137/S1052623403427859
[19] Attouch H, Cabot A. Convergence rate of a relaxed inertial proximal algorithm for convex minimization. Optimization. 2020;69:1281–1312. doi: 10.1080/02331934.2019.1696337
[20] Attouch H, Peypouquet J, Redont P. A dynamical approach to an inertial forward–backward algorithm for convex minimization. SIAM J Optim. 2014;24:232–256. doi: 10.1137/130910294
[21] Aujol J-F, Dossal C, Rondepierre A. Optimal convergence rates for Nesterov acceleration. SIAM J Optim. 2019;29:3131–3153. doi: 10.1137/18M1186757
[22] Boț RI, Csetnek ER, Hendrich C. Inertial Douglas–Rachford splitting for monotone inclusion problems. Appl Math Comput. 2015;256:472–487.
[23] Chen C, Ma S, Yang J. A general inertial proximal point algorithm for mixed variational inequality problem. SIAM J Optim. 2015;25:2120–2142. doi: 10.1137/140980910

[24] Lorenz DA, Pock T. An inertial forward–backward algorithm for monotone inclusions. J Math Imag Vision. 2015;51:311–325. doi: 10.1007/s10851-014-0523-2
[25] Maingé P-E, Merabet N. A new inertial-type hybrid projection-proximal algorithm for monotone inclusions. Appl Math Comput. 2010;215:3149–3162.
[26] Kim D. Accelerated proximal point method for maximally monotone operators. Math Program. 2021;190:57–87. doi: 10.1007/s10107-021-01643-0
[27] Drori Y, Teboulle M. Performance of first-order methods for smooth convex minimization: a novel approach. Math Program. 2014;145:451–482. doi: 10.1007/s10107-013-0653-0
[28] Güler O. New proximal point algorithms for convex minimization. SIAM J Optim. 1992;2:649–664. doi: 10.1137/0802032
[29] Maingé P-E. Accelerated proximal algorithms with a correction term for monotone inclusions. Appl Math Optim. 2021;84:2027–2061. doi: 10.1007/s00245-021-09819-y
[30] Lara F. On strongly quasiconvex functions: existence results and proximal point algorithms. J Optim Theory Appl. 2022;192:891–911. doi: 10.1007/s10957-021-01996-8
[31] Polyak BT. Existence theorems and convergence of minimizing sequences in extremum problems with restrictions. Soviet Math. 1966;7:72–75.
[32] Grad SM, Lara F. An extension of the proximal point algorithm beyond convexity. J Global Optim. 2022;82:313–329. doi: 10.1007/s10898-021-01081-4
[33] Opial Z. Weak convergence of the sequence of successive approximations for nonexpansive mappings. Bull Am Math Soc. 1967;73:591–597. doi: 10.1090/bull/1967-73-04
[34] Saejung S, Yotkaew P. Approximation of zeros of inverse strongly monotone operators in Banach spaces. Nonlinear Anal. 2012;75:742–750. doi: 10.1016/j.na.2011.09.005
[35] Maingé PE. Approximation methods for common fixed points of nonexpansive mappings in Hilbert spaces. J Math Anal Appl. 2007;325(1):469–479. doi: 10.1016/j.jmaa.2005.12.066
[36] Tao M, Yuan X. On the optimal linear convergence rate of a generalized proximal point algorithm. J Sci Comput. 2018;74:826–850. doi: 10.1007/s10915-017-0477-9
[37] Malitsky Y. Golden ratio algorithms for variational inequalities. Math Program. 2020;184:383–410. doi: 10.1007/s10107-019-01416-w
[38] Grad SM, Lara F. Solving mixed variational inequalities beyond convexity. J Optim Theory Appl. 2021;190:565–580. doi: 10.1007/s10957-021-01860-9
[39] Grad SM, Lara F, Marcavillaca RT. Relaxed-inertial proximal point type algorithms for quasiconvex minimization. J Global Optim. 2023;85:615–635. doi: 10.1007/s10898-022-01226-z
[40] Grad SM, Lara F, Marcavillaca RT. Proximal point type algorithms with relaxed and inertial effects beyond convexity. Optimization. 2024;73(11):3393–3410. doi: 10.1080/02331934.2024.2329779

Appendix

A.1 MATLAB notation for generating a symmetric positive semidefinite matrix of order n

function A = generateSPDmatrix(n)

$A = \text{rand}\,(n, n);$
$A = 0.5 * (A + A'); \text{ OR}$
$A = A * A';$
$A = A + n * eye(n);$
end