Journal of Applied Analysis and Computation
Volume 10, Number 5, October 2020, 2184–2197
Website:http://jaac-online.com
DOI:10.11948/20190363

ACCELERATED PROJECTION-BASED FORWARD-BACKWARD SPLITTING ALGORITHMS FOR MONOTONE INCLUSION PROBLEMS

Bing Tan[1], Zheng Zhou[1] and Xiaolong Qin[2,†]

Abstract In this paper, based on inertial and Tseng's ideas, we propose two projection-based algorithms to solve a monotone inclusion problem in infinite dimensional Hilbert spaces. Solution theorems of strong convergence are obtained under the certain conditions. Some numerical experiments are presented to illustrate that our algorithms are efficient than the existing results.

Keywords Monotone operator, forward-backward splitting algorithm, strong convergence, inclusion problem.

MSC(2010) 47H05, 49J40, 65K10, 47J20.

1. Introduction

In this paper, we start with the general optimization problem

$$\min\{\Phi(\mathbf{x}) = F(\mathbf{x}) + G(\mathbf{x}) : \mathbf{x} \in H\}, \tag{P}$$

where H is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$, $F : H \to (-\infty, \infty)$ is a continuously differentiable function and $G : H \to (-\infty, \infty]$ is a convex and closed function, which is assumed to be subdifferentiable on $\operatorname{dom} G$, the domain of G. If $\mathbf{x}^\star \in H$ is a local minimum of (P), then it is a stationary point of (P), i.e.,

$$\mathbf{0} \in \nabla F(\mathbf{x}^\star) + \partial G(\mathbf{x}^\star), \tag{1.1}$$

where $\partial G(\cdot)$ stands for the subdifferential of G. Note that if F is also convex, then $\mathbf{x}^\star$ is the global minimum of (P). For any $t > 0$, one sees from (1.1) that

$$\begin{aligned}\mathbf{0} \in t\nabla F(\mathbf{x}^\star) + t\partial G(\mathbf{x}^\star) &\Leftrightarrow (I + t\partial G)(\mathbf{x}^\star) \in (I - t\nabla F)(\mathbf{x}^\star)\\ &\Leftrightarrow \mathbf{x}^\star = (I + t\partial G)^{-1}(I - t\nabla F)(\mathbf{x}^\star),\end{aligned}$$

from which a fixed-point scheme naturally arises to generate the following iterative sequence $\{\mathbf{x}_k\}$:

$$\mathbf{x}_k = (I + t_k\partial G)^{-1}(I - t_k\nabla F)(\mathbf{x}_{k-1}), \quad \mathbf{x}_0 \in \mathbb{R},\ t_k > 0. \tag{1.2}$$

[†]The corresponding author. Email address: qxlxajh@163.com (X. Qin)
[1]Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China
[2]Department of Mathematics, Zhejiang Normal University, Zhejiang, China

Actually, (1.2) is a special case of the forward-backward (FB) algorithm which was originally designed to find a zero of the more general inclusion problem:

$$\mathbf{0} \in A\left(\mathbf{x}^{\star}\right) + B\left(\mathbf{x}^{\star}\right), \tag{1.3}$$

where A and B are set-valued maximal monotone maps. (1.3) is reduced to (1.1) if both F and G are convex and $A := \nabla F$ and $B := \partial G$.

A classic algorithm to solve (1.3) is the known forward-backward splitting algorithm, which was first introduced by Passty [20] and Lions and Mercier [15]. In recent years, this method has been widely investigated in various problems, such as, coupled monotone inclusions, constrained variational inequalities, signal processing, image recovery, machine learning, convex optimization problems, etc; see, e.g., [1, 6, 8, 22] and the references therein. It is known that the FB method converges provided that the inverse of forward mapping A^{-1} is strongly monotone and B is maximal monotone [11]. In 1997, Chen and Rockafellar [7] gave the convergence rates analysis of the FB method. In 2000, Tseng [23] obtained a modified FB algorithm for zeros of maximal monotone mappings. This method achieves convergence only by assuming that the forward mapping is continuous over a closed convex subset of its domain.

Tseng (2000): Modified Forward-Backward Splitting Algorithm.

$$\begin{cases} y_n = \left(I + \gamma_n G\right)^{-1}\left(x_n - \gamma_n F x_n\right), \\ x_{n+1} = y_n - \gamma_n\left(F y_n - F x_n\right). \end{cases} \tag{1.4}$$

Polyak [21] first proposed the inertial idea to improve the convergence of the algorithms. Inertial-type methods, which are considered as a method to accelerate the convergence of Tseng-type iterative methods, are based on a discrete version of a second-order dissipative dynamical system [2]. In recent years, some authors constructed various fast iterative algorithms via inertial extrapolation techniques on some classical methods, such as, inertial proximal point algorithms, inertial Mann algorithms, inertial Douglas-Rachford splitting algorithms, inertial alternating direction method of multipliers, inertial forward-backward splitting algorithms and inertial extragradient algorithms, etc. On the other hand, Nesterov [18] developed an acceleration scheme which improves the convergence speed of the forward-backward algorithm from the standard $O\left(k^{-1}\right)$ to $O\left(k^{-2}\right)$. In addition, Attouch and Peypouquet [3] proved that the Nesterov's accelerated forward-backward method is actually $o\left(k^{-2}\right)$ rather than $O\left(k^{-2}\right)$.

In 2015, Lorenz and Pock [16] proposed the following inertial forward-backward algorithm by combining the inertial idea with the forward-backward algorithm for monotone operators. It should be noted that Algorithm (1.5) is still weakly convergent.

Lorenz and Pock (2015): Inertial Forward-Backward Algorithm.

$$\begin{cases} y_n = x_n + \alpha_n\left(x_n - x_{n-1}\right), \\ x_{n+1} = \left(I + \gamma_n G\right)^{-1}\left(y_n - \gamma_n F y_n\right). \end{cases} \tag{1.5}$$

In practical applications, many problems, such as, quantum physics and image reconstruction, are in infinite dimensional spaces. To investigate these problems, norm convergence is usually preferable to the weak convergence. In 2003, Nakajo and Takahashi [19] established strong convergence of the Mann iteration with the aid of projections. Indeed, they considered the following algorithm:

Nakajo and Takahashi (2003): Hybrid Projection Method.

$$\begin{cases} y_n = \alpha_n x_n + (1-\alpha_n) T x_n \,, \\ C_n = \{u \in C : \|y_n - u\| \leqslant \|x_n - u\|\} \,, \\ Q_n = \{u \in C : \langle x_n - u, x_0 - x_n\rangle \geqslant 0\} \,, \\ x_{n+1} = \mathcal{P}_{C_n \cap Q_n} x_0 \,, \quad n \in \mathbb{N} \,, \end{cases} \tag{1.6}$$

where $\{\alpha_n\} \subset [0,1)$, T is a nonexpansive mapping on C and $\mathcal{P}_{C_n \cap Q_n}$ is the nearest point projection from C onto $C_n \cap Q_n$. This method is now referred to as the hybrid projection method. Inspired by Nakajo and Takahashi [19], Takahashi, Takeuchi and Kubota [24] also proposed a projection-based method and obtain the strong convergence of the method, which is now called the shrinking projection method. In recent years, many authors studied these projection-based methods in various spaces; see, e.g., [9, 10, 25, 26].

Inspired and motivated by the above works, we propose two new projection-based inertial solution methods with adaptive stepsizes, which are more flexible than the fixed stepsizes. Solution theorems of strong convergence are established in the framework of real Hilbert spaces. Numerical examples to illustrate the efficiency and robustness of the proposed algorithms are provided.

Our paper is organized as follows. In Section 2, we give some useful and necessary preliminaries for our convergence analysis and numerical experiments. In Section 3, we propose our new algorithms and obtain solution theorems of strong convergence under some mild conditions. In Section 4, we give some numerical results in convex minimization problems to show the efficient and robust of our algorithms. Section 5 ends this paper.

2. Preliminaries

Let C be a non-empty, convex and closed set in a real Hilbert space H. For a given sequence $\{x_n\} \subset H$, let $\omega_w(x_n) := \{x : \exists x_{n_j} \rightharpoonup x\}$ denote the weak w-limit set of $\{x_n\}$. For any $x, y \in H$, we have

(1) $\|x-y\|^2 + 2\langle x-y, y\rangle = \|x\|^2 - \|y\|^2$;

(2) $\|x+y\|^2 \le \|x\|^2 + 2\langle y, x+y\rangle$;

(3) $\|tx + (1-t)y\|^2 + t(1-t)\|x-y\|^2 = t\|x\|^2 + (1-t)\|y\|^2$, $\forall t \in \mathbb{R}$.

Let $F : H \to H$ be an operator. The fixed-pint set of F is denoted by $\mathrm{Fix}(F)$, where $\mathrm{Fix}(F) := \{x \in H \mid Fx = x\}$. F is said to be L-Lipschitz continuous with $L > 0$ if

$$\|Fx - Fy\| \le L\|x-y\| \,, \forall x, y \in H.$$

If $L = 1$, then F is said to be nonexpansive. F is said to be monotone if

$$\langle Fx - Fy, x - y\rangle \geq 0\,,\ \forall x, y \in H.$$

F is said to be strongly monotone with $\alpha > 0$ if

$$\langle Fx - Fy, x - y\rangle \geq \alpha\|x - y\|^2,\ \forall x, y \in H.$$

For any $x \in H$, there exists a unique nearest point in C, denoted by $\mathcal{P}_C x$, such that

$$\mathcal{P}_C(x) := \operatorname{argmin}_{y\in C} \|x - y\|\,,$$

where $\mathcal{P}_C$ is called the metric projection of H onto C. It has such an equivalent form $\langle \mathcal{P}_C x - x, \mathcal{P}_C x - y\rangle \leq 0, \forall y \in C$, and can also be converted to $\|y - \mathcal{P}_C x\|^2 + \|x - \mathcal{P}_C x\|^2 \leq \|x - y\|^2$. It can be calculated that the projection of x_0 on a polyhedron is described by linear inequalities $Ax \preceq b$ via the following quadratic programming (QP)

$$\text{minimize } \|x - x_0\|_2^2\,, \qquad \text{subject to } Ax \preceq b\,.$$

We next give some special cases with simple analytical solutions.

(i) The Euclidean projection of x_0 onto an affine subspace $\Omega = \{x : Ax = b\}$ with $A \in \mathbb{R}^{m\times n}$ and $\operatorname{rank}(A) = m < n$ is given by

$$\mathcal{P}_\Omega(x_0) = x_0 + A^\mathsf{T}\left(AA^\mathsf{T}\right)^{-1}(b - Ax_0)\,.$$

(ii) The Euclidean projection of x_0 onto a halfspace $\Omega = \left\{x : a^\mathsf{T} x \leq b\ (a \neq 0)\right\}$ is given by

$$\mathcal{P}_\Omega(x_0) = \begin{cases} x_0\,, & \text{if } a^\mathsf{T} x_0 \leq b\,; \\ x_0 + \dfrac{b - a^\mathsf{T} x_0}{\|a\|^2} a\,, & \text{if } a^\mathsf{T} x_0 > b. \end{cases}$$

Let G be a proper, lower semi-continuous and convex function.

$$\operatorname{prox}_{\gamma G}(y) := \operatorname{argmin}_{x\in H}\left\{\frac{1}{2}\|x - y\|^2 + \gamma G(x)\right\},\ \forall y \in H\,,$$

where γ is a positive real number, is called the proximity operator.

Note that it has the closed-form expression in some important cases. For example, if the Euclidean norm $G(x) = \|x\|_1$, then one has the shrinkage-threshold operator $\mathcal{T}_\gamma(y)$

$$\operatorname{prox}_{\gamma G}(y) = \left(\mathcal{T}_\gamma(y)\right)_i = \begin{cases} \operatorname{sign}(y_i)\cdot(|y_i| - \gamma)_+\,, & \text{if } |y_i| > \gamma\,; \\ 0\,, & \text{if } |y_i| \leq \gamma. \end{cases}$$

Let $G : H \to 2^H$ be a multivalued operator on H. G is said to be monotone iff $\langle p - q, x - y\rangle \geq 0$ for any $x, y \in H$, $p \in Gx$ and $q \in Gy$. Recall that a mutivalued operator $G : H \to 2^H$ is said to be maximal iff its Graph is not contained in the graph of any other monotone operator properly. One knows that a monotone $G : H \to 2^H$ is maximal iff for any $(x, p) \in H \times H$, $\langle p - q, x - y\rangle \geq 0$ for every $(y, q) \in \operatorname{Graph}(G)$ yields $p \in Gx$.

Lemma 2.1 ([14]). *Let $F : H \to H$ be a operator and $G : H \to 2^H$ be a maximal monotone operator. Define $T_\gamma := (I + \gamma G)^{-1}(I - \gamma F), \gamma > 0$. Then,* $\operatorname{Fix}(T_\gamma) = (F + G)^{-1}(0)$.

Lemma 2.2 ([4]). *Let $F : H \to H$ be a Lipschitz continuous and monotone mapping, and let $G : H \to 2^H$ be a maximal monotone mapping. Then $F + G$ is maximally monotone.*

Lemma 2.3 ([13]). *Let C be a convex and closed set in a real Hilbert space H. Given $x, y, z \in H$ and $a \in \mathbb{R}$, $\left\{v \in C : \|y - v\|^2 \le \|x - v\|^2 + \langle z, v\rangle + a\right\}$ is convex and closed.*

Lemma 2.4 ([17]). *Let C be a convex and closed set in a real Hilbert space H, $\{x_n\} \subset H$ and $u \in H$. Let $q = \mathcal{P}_C u$. If the weak ω-limit set $\omega_w(x_n) \subset C$ and $\|x_n - u\| \le \|u - q\|$, $\forall n \in \mathbb{N}$, then $\{x_n\}$ converges to q in norm.*

3. Main Results

In this section, we assume that the following conditions are satisfied for our convergence analysis.

(A1) The solution set of the inclusion problem (1.3) is nonempty, i.e., $\Omega := (F + G)^{-1}(0) \neq \emptyset$.

(A2) The mapping $G : H \to 2^H$ is maximal monotone, $F : H \to H$ is L-Lipschitz continuous and monotone.

3.1. The Inertial Hybrid Projection Algorithm

Algorithm 3.1: Inertial Hybrid Projection Algorithm (IHPA).
Input: $x_{-1} = x_0$, $\gamma_0 > 0$, $\mu \in (0, 1)$, $\alpha_n \in [0, 1)$.
for $n = 0 : \mathit{Maxiters}$ **do**

$$\begin{cases} w_n = x_n + \alpha_n (x_n - x_{n-1}), \\ y_n = (I + \gamma_n G)^{-1} (I - \gamma_n F) w_n, \\ z_n = y_n - \gamma_n (F y_n - F w_n), \\ C_n = \left\{u \in H : \|z_n - u\|^2 \le \|w_n - u\|^2 - \left(1 - \mu^2 \frac{\gamma_n^2}{\gamma_{n+1}^2}\right) \|w_n - y_n\|^2 \right\}, \\ Q_n = \{u \in H : \langle x_n - u, x_n - x_0\rangle \le 0\}, \\ x_{n+1} = \mathcal{P}_{C_n \cap Q_n} x_0,\ n \ge 0. \end{cases} \tag{3.1}$$

Update γ_n by (3.2),
end

where $\{\gamma_n\}$ is the step size generated by

$$\gamma_{n+1} = \begin{cases} \min\left\{\dfrac{\mu \|w_n - y_n\|}{\|F w_n - F y_n\|}, \gamma_n\right\}, & \text{if } F w_n - F y_n \neq 0; \\ \gamma_n, & \text{otherwise.} \end{cases} \tag{3.2}$$

Remark 3.1. The inertial parameter $\{\alpha_n\}$ in (3.1) can be selected as an arbitrary sequence in $[0,1)$ to produce acceleration. Notice that the parameter $\{\alpha_n\}$ in (3.1) was generated by the expression $\left(\frac{t_{n-1}-1}{t_n}\right)$ in [5] and $\frac{n-1}{n+3}$ in [3]. In this paper, $\{\alpha_n\}$ will also be adaptively updated by

$$\alpha_n = \begin{cases} \min\left\{\alpha, \dfrac{\xi_n}{\|x_n - x_{n-1}\|}\right\}, & \text{if } x_n \neq x_{n-1}\,; \\ \alpha\,, & \text{otherwise,} \end{cases} \tag{3.3}$$

where $\alpha \in [0,1)$, the sequence $\{\xi_n\}$ satisfies $\lim_{n\to\infty} \xi_n = 0$ and $\sum_{n=1}^{\infty} \xi_n = \infty$.

The following lemmas play a significant role in this paper for the convergence analysis.

Lemma 3.1. *Let $\{z_n\}$ be a sequence generated by Algorithm 3.1. If conditions (A1) and (A2) hold, then*

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - \left(1 - \mu^2 \frac{\gamma_n^2}{\gamma_{n+1}^2}\right) \|w_n - y_n\|^2\,,\ \forall p \in \Omega\,. \tag{3.4}$$

Proof. Setting $a_n = \gamma_n^2 \|Fy_n - Fw_n\|^2 - 2\gamma_n \langle y_n - p, Fy_n - Fw_n\rangle$, one has

$$\begin{aligned} \|z_n - p\|^2 &= \|y_n - p\|^2 + \gamma_n^2 \|Fy_n - Fw_n\|^2 - 2\gamma_n \langle y_n - p, Fy_n - Fw_n\rangle \\ &= \|w_n - p\|^2 + \|y_n - w_n\|^2 + 2\langle w_n - p, y_n - w_n\rangle + a_n \\ &= \|w_n - p\|^2 + \|y_n - w_n\|^2 - 2\langle y_n - w_n, y_n - w_n\rangle + 2\langle y_n - w_n, y_n - p\rangle + a_n \\ &= \|w_n - p\|^2 - \|y_n - w_n\|^2 - 2\langle y_n - p, w_n - y_n + \gamma_n (Fy_n - Fw_n)\rangle \\ &\quad + \gamma_n^2 \|Fy_n - Fw_n\|^2\,. \end{aligned} \tag{3.5}$$

Note that

$$\gamma_{n+1} = \min\left\{\frac{\mu\|w_n - y_n\|}{\|Fw_n - Fy_n\|}, \gamma_n\right\} \leq \frac{\mu\|w_n - y_n\|}{\|Fw_n - Fy_n\|}\,,$$

which means that

$$\|Fw_n - Fy_n\| \leq \frac{\mu}{\gamma_{n+1}} \|w_n - y_n\|\,. \tag{3.6}$$

If $Fw_n = Fy_n$, then inequality (3.6) holds obviously. Combining (3.5) and (3.6), one obtains

$$\begin{aligned} \|z_n - p\|^2 \leq \|w_n - p\|^2 &- \left(1 - \mu^2 \frac{\gamma_n^2}{\gamma_{n+1}^2}\right) \|w_n - y_n\|^2 \\ &- 2\langle y_n - p, w_n - y_n + \gamma_n (Fy_n - Fw_n)\rangle\,. \end{aligned} \tag{3.7}$$

Next, one proves

$$\langle y_n - p, w_n - y_n + \gamma_n (Fy_n - Fw_n)\rangle \geq 0\,. \tag{3.8}$$

From $y_n = (I + \gamma_n G)^{-1} (I - \gamma_n F) w_n$, one obtains $(I - \gamma_n F) w_n \in (I + \gamma_n G) y_n$. Since G is maximally monotone, one concludes that there exists $u_n \in G y_n$ such that $(I - \gamma_n F) w_n = y_n + \gamma_n u_n$. This means that

$$u_n = \frac{1}{\gamma_n} (w_n - \gamma_n F w_n - y_n)\,. \tag{3.9}$$

On the other hand, one has $0 \in (F+G)p$ and $Fy_n + u_n \in (F+G)y_n$. Since $F+G$ is maximally monotone, one gets

$$\langle Fy_n + u_n, y_n - p\rangle \geq 0 \,. \tag{3.10}$$

Substituting (3.9) into (3.10), one gets

$$\frac{1}{\gamma_n}\langle w_n - \gamma_n Fw_n - y_n + \gamma_n Fy_n, y_n - p\rangle \geq 0 \,,$$

which means that $\langle w_n - y_n + \gamma_n (Fy_n - Fw_n), y_n - p\rangle \geq 0$. From (3.7) and (3.8), one concludes (3.4) immediately. □

Lemma 3.2. *Let $\{x_n\}, \{w_n\}$ and $\{y_n\}$ be three sequences generated by Algorithm 3.1. Assume that conditions (A1) and (A2) hold. If $\lim_{n\to\infty} \|w_n - x_n\| = 0$, $\lim_{n\to\infty} \|w_n - y_n\| = 0$, and $\{x_{n_k}\}$, which is a subsequence of $\{x_n\}$, converges weakly to some $q \in H$, then $q \in \Omega$, where $\Omega = (F+G)^{-1}(0)$.*

Proof. Let $(h, g) \in \mathrm{Graph}(F+G)$, i.e., $g - Fh \in Gh$. From the definition of y_n, we have $y_{n_k} = (I + \gamma_{n_k} G)^{-1} (I - \gamma_{n_k} F) w_{n_k}$, one obtains $(I - \gamma_{n_k} F) w_{n_k} \in (I + \gamma_{n_k} G) y_{n_k}$, which implies

$$\frac{1}{\gamma_{n_k}} (w_{n_k} - y_{n_k} - \gamma_{n_k} Fw_{n_k}) \in Gy_{n_k} \,.$$

On the other hand, by the maximal monotonicity of G, one has

$$\langle h - y_{n_k}, g - Fh - (w_{n_k} - y_{n_k} - \gamma_{n_k} Fw_{n_k}) / \gamma_{n_k}\rangle \geq 0 \,.$$

Therefore,

$$\begin{aligned}
\langle h - y_{n_k}, g\rangle &\geq \langle h - y_{n_k}, Fh + (w_{n_k} - y_{n_k} - \gamma_{n_k} Fw_{n_k}) / \gamma_{n_k}\rangle \\
&= \langle h - y_{n_k}, Fh - Fw_{n_k}\rangle + \langle h - y_{n_k}, (w_{n_k} - y_{n_k}) / \gamma_{n_k}\rangle \\
&= \langle h - y_{n_k}, Fh - Fy_{n_k}\rangle + \langle h - y_{n_k}, Fy_{n_k} - Fw_{n_k}\rangle \\
&\quad + \langle h - y_{n_k}, (w_{n_k} - y_{n_k}) / \gamma_{n_k}\rangle \\
&\geq \langle h - y_{n_k}, Fy_{n_k} - Fw_{n_k}\rangle + \langle h - y_{n_k}, (w_{n_k} - y_{n_k}) / \gamma_{n_k}\rangle \,.
\end{aligned}$$

Since $\lim_{n\to\infty} \|w_n - x_n\| = 0$, $\lim_{n\to\infty} \|w_n - y_n\| = 0$, and F is Lipschitz continuous, one gets $\lim_{k\to\infty} \|Fy_{n_k} - Fw_{n_k}\| = 0$. By $\lim_{n\to\infty} \gamma_n = \gamma \geq \min\left\{\gamma_0, \frac{\mu}{L}\right\}$, one obtains

$$\lim_{k\to\infty} \langle h - y_{n_k}, g\rangle = \langle h - q, g\rangle \geq 0 \,.$$

With the aid of the maximal monotonicity of $F+G$, one obtains $0 \in (F+G)q$, that is, $q \in \Omega$. □

Theorem 3.1. *Assume that both F and G satisfy conditions (A1)–(A2). Then the sequence $\{x_n\}$ generated by Algorithm 3.1 converges to an element $q^* \in \Omega$ strongly, where $q^* = \mathcal{P}_\Omega x_0$.*

Proof. The proof is divided into three steps.

Step 1. It is obvious that C_n and Q_n are convex closed for all $n \geq 0$. Next one shows that $\Omega \subset C_n \cap Q_n, \forall n \geq 0$ and $\{x_n\}$ is well defined. Lemma 3.1 implies that $\Omega \subset C_n, \forall n \geq 0$. From the definition of Q_n in Algorithm 3.1, one has $Q_0 = H$. Further, $\Omega \subset C_0 \cap Q_0$ and $x_1 = \mathcal{P}_{C_0 \cap Q_0} x_0$ is well defined. Without loss of generality, one assumes that x_n is given and $\Omega \subset C_n \cap Q_n$ for some n. This shows that $x_{n+1} = \mathcal{P}_{C_n \cap Q_n} x_0$ is well defined. It follows from the property of the projection that $\langle z - x_{n+1}, x_0 - x_{n+1} \rangle \leq 0, \forall z \in C_n \cap Q_n$. Since $\Omega \subset C_n \cap Q_n$, one concludes $\langle u - x_{n+1}, x_0 - x_{n+1} \rangle \leq 0, \forall u \in \Omega$. This implies that $\Omega \subset Q_{n+1}$ and thus $\Omega \subset C_{n+1} \cap Q_{n+1}$.

Step 2. One shows that $\{x_n\}$ is bounded and $\lim_{n\to\infty} \|w_n - y_n\| = 0$. Since $\Omega \subset C_n \cap Q_n$ and $x_{n+1} = \mathcal{P}_{C_n \cap Q_n} x_0$, one gets $\|x_{n+1} - x_0\| \leq \|q^* - x_0\|, \forall n \geq 0$. This means that $\{x_n\}$ is bounded, so are $\{w_n\}$ and $\{z_n\}$. Combining the definition of Q_n and the property of the projection, one has $x_n = \mathcal{P}_{Q_n} x_0$. Since $x_{n+1} \in Q_n$, one further has

$$\|x_n - x_0\| \leq \|x_{n+1} - x_0\|, \forall n \geq 0.$$

Thus $\lim_{n\to\infty} \|x_n - x_0\|$ exists. It follows that

$$\|x_n - x_{n+1}\|^2 \leq \|x_{n+1} - x_0\|^2 - \|x_n - x_0\|^2 .$$

We see that $\lim_{n\to\infty} \|x_n - x_{n+1}\| = 0$. Since $\|x_{n+1} - z_n\| \leq \|w_n - x_{n+1}\|$ ($x_{n+1} \in C_n$) and $\|w_n - x_n\| \leq |\alpha_n| \|x_n - x_{n-1}\|$, one arrives at $\lim_{n\to\infty} \|z_n - w_n\| \leq \lim_{n\to\infty}\{\|z_n - x_n\| + \|x_n - w_n\|\} = 0$. From Lemma 3.1, we obtain

$$\left(1 - \mu^2 \frac{\gamma_n^2}{\gamma_{n+1}^2}\right) \|w_n - y_n\|^2 \leq \|w_n - p\|^2 - \|z_n - p\|^2 \leq (\|w_n - p\| + \|z_n - p\|) \|z_n - w_n\| .$$

It is clear to see that $\lim_{n\to\infty} \|w_n - y_n\| = 0$.

Step 3. One shows that $\{x_n\}$ converges to $q^* \in \Omega$ strongly, where $q^* = \mathcal{P}_\Omega x_0$. So far, we have shown that

(1) If $q^* = \mathcal{P}_\Omega x_0$, then $\|x_n - x_0\| \leq \|x_0 - q^*\|, \forall n \in \mathbb{N}$.

(2) Every sequential weak cluster point of the sequence $\{x_n\}$ is in Ω, i.e., $\omega_w(x_n) \subset \Omega$.

By Lemma 2.4, one concludes that $\{x_n\}$ converges to the point $q^* \in \Omega$ strongly, where $q^* = \mathcal{P}_\Omega x_0$. The proof is completed. □

3.2. The Inertial Shrinking Projection Algorithm

Algorithm 3.2: Inertial Shrinking Projection Algorithm (ISPA).
Input: $x_{-1} = x_0$, $\gamma_0 > 0$, $\mu \in (0,1)$, $C_0 = H$, $\alpha_n \in [0,1)$.
for $n = 0 : Maxiters$ **do**

$$\begin{cases} w_n = x_n + \alpha_n (x_n - x_{n-1}), \\ y_n = (I + \gamma_n G)^{-1} (I - \gamma_n F) w_n , \\ z_n = y_n - \gamma_n (F y_n - F w_n), \\ C_{n+1} = \left\{u \in C_n : \|z_n - u\|^2 \leq \|w_n - u\|^2 - \left(1 - \mu^2 \frac{\gamma_n^2}{\gamma_{n+1}^2}\right) \|w_n - y_n\|^2 \right\}, \\ x_{n+1} = \mathcal{P}_{C_{n+1}} x_0, \ n \geq 0 . \end{cases}$$

Update γ_n by (3.2).
end

Theorem 3.2. *Assume that both F and G satisfy conditions (A1)–(A2). Then the sequence $\{x_n\}$ generated by Algorithm 3.2 converges to an element $q^* \in \Omega$ strongly, where $q^* = \mathcal{P}_\Omega x_0$.*

Proof. From Lemma 3.1, one easily concludes that

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - \left(1 - \mu^2 \frac{\gamma_n^2}{\gamma_{n+1}^2}\right) \|w_n - y_n\|^2 , \forall p \in \Omega .$$

Since $x_n = \mathcal{P}_{C_n} x_0$ and $x_{n+1} = \mathcal{P}_{C_{n+1}} x_0 \in C_{n+1} \subset C_n$, we obtain $\|x_n - x_0\| \leq \|x_{n+1} - x_0\|$. On the other hand, from $\Omega \subset C_n$, we get $\|x_n - x_0\| \leq \|q^* - x_0\|$. It implies that $\{x_n\}$ is bounded and nondecreasing. Thus, $\lim_{n\to\infty} \|x_n - x_0\|$ exists. From Step 3 in Theorem 3.1, $\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0$ and $\lim_{n\to\infty} \|w_n - y_n\| = 0$ hold. From Lemma 3.2 and Lemma 2.4, $\{x_n\}$ converges to the point $q^* \in \Omega$ strongly, where $q^* = \mathcal{P}_\Omega x_0$. This completes the proof. □

4. Numerical Results

In this section, we give some numerical examples to illustrate the effectiveness and robustness of the proposed algorithms in Section 3. We compare the two strong convergence algorithms, proposed by Gibali and Thong [12], Mann Tseng-type algorithm and Viscosity Tseng-type algorithm. All the programs are performed in MATLAB2018a on a PC Desktop Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz 1.800 GHz, RAM 8.00 GB.

Based on Mann and Viscosity ideas, Gibali and Thong [12] presented two modifications of the forward-backward splitting method in real Hilbert spaces as follows:

Algorithm 4.1: Mann Tseng-type modification (MTTM).

$$\begin{cases} y_n = (I + \gamma_n G)^{-1} (I - \gamma_n F) x_n , \\ z_n = y_n - \gamma_n (F y_n - F x_n) , \\ x_{n+1} = (1 - \delta_n - \theta_n) x_n + \theta_n z_n . \end{cases}$$

Update γ_n by (3.2),

and

Algorithm 4.2: Viscosity Tseng-type modification (VTTM).

$$\begin{cases} y_n = (I + \gamma_n G)^{-1} (I - \gamma_n F) x_n , \\ z_n = y_n - \gamma_n (F y_n - F x_n) , \\ x_{n+1} = \delta_n f(x_n) + (1 - \delta_n) z_n . \end{cases}$$

Update γ_n by (3.2),

where $\{\delta_n\}$ and $\{\theta_n\}$ are two real sequences in $(0,1)$ such that $\{\theta_n\} \subset (a,b) \subset (0, 1 - \delta_n)$ for some $a > 0, b > 0$, $\lim_{n\to\infty} \delta_n = 0$, $\sum_{n=1}^{\infty} \delta_n = \infty$ and mapping $f : H \to H$ is contraction.

Example 4.1. Let $x = (x_1, x_2, \ldots, x_{10}) \in \mathbb{R}^{10}$ and define $F : \mathbb{R}^{10} \to \mathbb{R}^{10}$ and $G : \mathbb{R}^{10} \to \mathbb{R}^{10}$ by $Fx = 2x + (1, 1, \ldots, 1)$ and $Gx = 5x$, respectively. It is clear to see that G is maximally monotone, and F is 2-Lipschitz continuous and monotone. After simple calculations, we obtain

$$(I + \gamma_n G)^{-1} (x_n - \gamma_n F x_n) = \frac{1 - 2\gamma_n}{1 + 5\gamma_n} x_n - \frac{\gamma_n}{1 + 5\gamma_n} (1, 1, \ldots, 1).$$

Our parameters are set as follows. The stepsizes of the four algorithms are updated by (3.2) with $\gamma_0 = 0.4$ and $\mu = 0.5$. Algorithm 3.1 updates the inertial parameters by $\alpha_n = \frac{n-1}{n+3}$. Algorithm 4.1, Algorithm 4.2 and Algorithm 3.2 update the inertial parameters by (3.3) with $\alpha = 0.6$ and $\xi_n = \frac{1}{(n+1)^2}$. In Algorithm 4.1 and Algorithm 4.2, we set $\delta_n = \frac{1}{n+1}$, $\theta_n = \frac{n}{2(n+1)}$, $f(x) = 0.5x$. The maximum iteration of 100 as the stopping criterion. Fig. 1 shows the convergence behavior of $\{\|x_n - x^*\|\}$, where $x^* = -(1, 1, \ldots, 1)/7$. The numerical results illustrate that the inertial parameters play a positive role in the convergence speed and the precision of the algorithms.

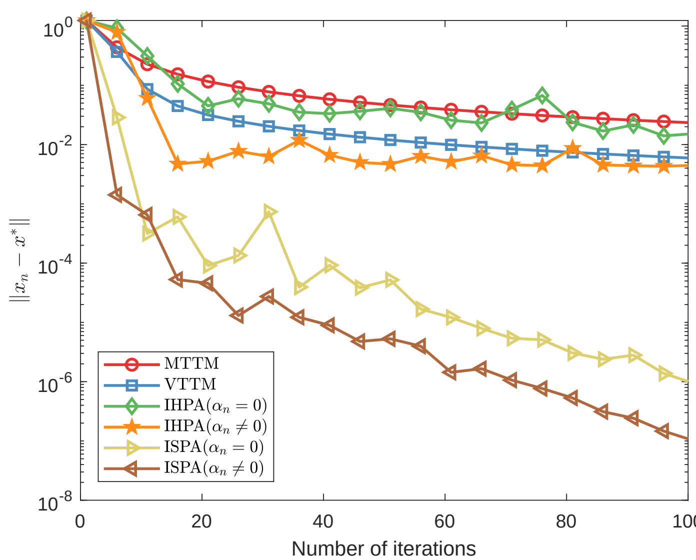

Figure 1. Convergence behavior of iterative sequence $\{\|x_n - x^*\|\}$.

Example 4.2. Find a solution of the following convex minimization problem:

$$\min_{x \in \mathbb{R}^2} \|x\|_2^2 + (3, 5)x + \|x\|_1,$$

where $x = (x_1, x_2) \in \mathbb{R}^2$. We know the exact solution x^* is $(-1, -2)$ and the minimum vaule is -5.

Next, we use our algorithms to solve the minimization problem in Example 4.2. Set $F(x) = \|x\|_2^2 + (3, 5)x$, $G(x) = \|x\|_1$ and $\Phi(x) = F(x) + G(x)$. It is clear that F is convex differentiable with $\nabla F = 2x + (3, 5)$, G is convex lower semicontinuous but not differentiable. Note that

$$(I+\gamma\partial G)^{-1}(x)=(\max\{|x_1|-\gamma,0\}\operatorname{sign}(x_1),\max\{|x_2|-\gamma,0\}\operatorname{sign}(x_2))\ .$$

Our parameter settings are the same as in Example 4.1. Fig. 2 shows the convergence behavior of the iterative sequence $\{\|x_n-x_0\|\}$. Fig. 3 shows the convergence behavior of the sequence $\{\|\Phi(x_n)-\Phi(x^*)\|\}$.

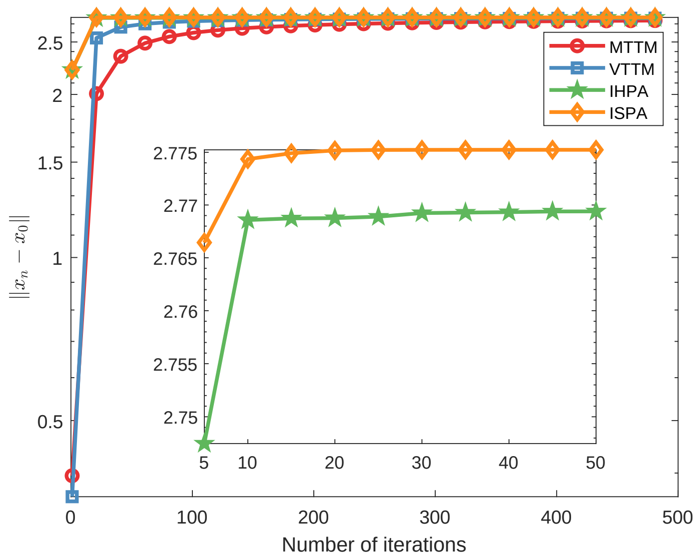

Figure 2. Convergence behavior of iterative sequence $\{\|x_n-x_0\|\}$.

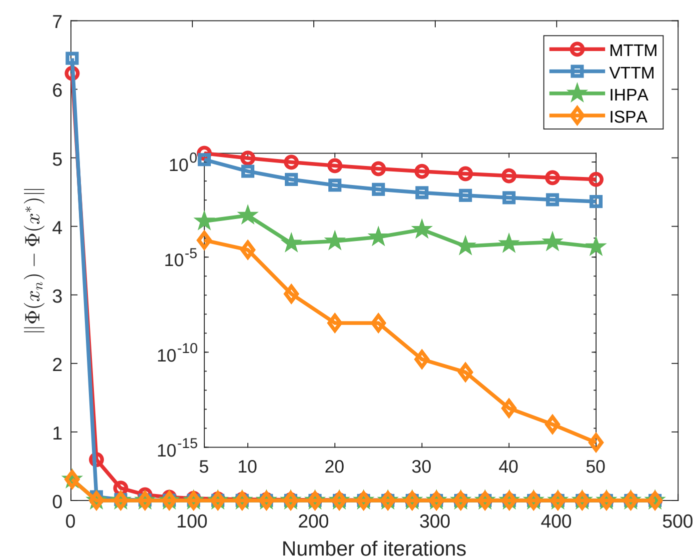

Figure 3. Convergence behavior of iterative sequence $\{\|\Phi(x_n)-\Phi(x^*)\|\}$.

As shown in Figs. 2 and 3, sequence $\{\|\Phi(x_n) - \Phi(x^*)\|\}$ converges to 0 means that the function value converges to the optimal value. In addition, it is easy to see that the iterative sequences $\{x_n\}$ and $\{\Phi(x_n)\}$ of Algorithm 3.1 and Algorithm 3.2 converge faster than Algorithm 4.1 and Algorithm 4.2.

Further, we show the numerical results in Table 1. The function value $\{\Phi(x_n)\}$ converges to the optimal value $\Phi(x^*) = -5$ as the number of iterations increases. We find that our proposed Algorithm 3.1 and Algorithm 3.2 enjoy higher precision than Algorithm 4.1 and Algorithm 4.2. It should be pointed out that our Algorithm 3.1 and Algorithm 3.2 require only a few iterations to achieve convergence (cf. Table 1).

Table 1. Comparison of four algorithms in Example 4.2.

iter n	$\|\Phi(x_n) - \Phi(x^*)\|$			
	MTTM	VTTM	IHPA	ISPA
1	$6.2330e+00$	$6.4518e+00$	$3.0808e-01$	$3.0808e-01$
10	$1.6077e+00$	$3.2709e-01$	$1.5545e-03$	$2.4074e-05$
20	$6.4395e-01$	$6.0776e-02$	$6.8905e-05$	$3.4213e-09$
100	$3.1369e-02$	$2.0442e-03$	$4.1663e-05$	$2.5848e-10$
300	$3.5320e-03$	$2.2375e-04$	$2.2536e-05$	$1.7764e-15$
500	$1.2749e-03$	$8.0326e-05$	$1.6109e-05$	$8.8818e-16$

To show that our algorithms are robust, four different initial values were tested, and the experimental results are reported in Table 2.

Table 2. Function value errors at different initial values.

Start point x_0	$\|\Phi(x_n) - \Phi(x^*)\|$			
	MTTM	VTTM	IHPA	ISPA
$[0.6787, 0.7577]$	$1.2749e-03$	$8.0326e-05$	$2.1152e-05$	$8.8818e-16$
$[-0.6739, -0.2305]$	$1.2749e-03$	$8.0326e-05$	$2.7860e-05$	$8.8818e-16$
$[0.4218, -0.9157]$	$1.2749e-03$	$8.0326e-05$	$8.4837e-06$	$1.7764e-15$
$[-0.9575, 0.9649]$	$1.2749e-03$	$8.0326e-05$	$1.4506e-05$	$1.7764e-15$

In addition, we also plot the convergence behavior of $\{\|\Phi(x_n) - \Phi(x^*)\|\}$ with different initial points in Fig. 4. Note that the projection-type algorithms converge faster than the others. These results are independent of the choice of initial values. This shows that our algorithms are effective and robust.

5. Conclusion

Forward-Backward splitting algorithms are efficient and powerful to monotone inclusion problems. In this paper, we investigated the problem of finding a zero of the sum of two monotone operators in real Hilbert spaces by proposing two projection-based algorithms with inertial effects. Our algorithms use a new step size rule which makes them more efficient and robust.

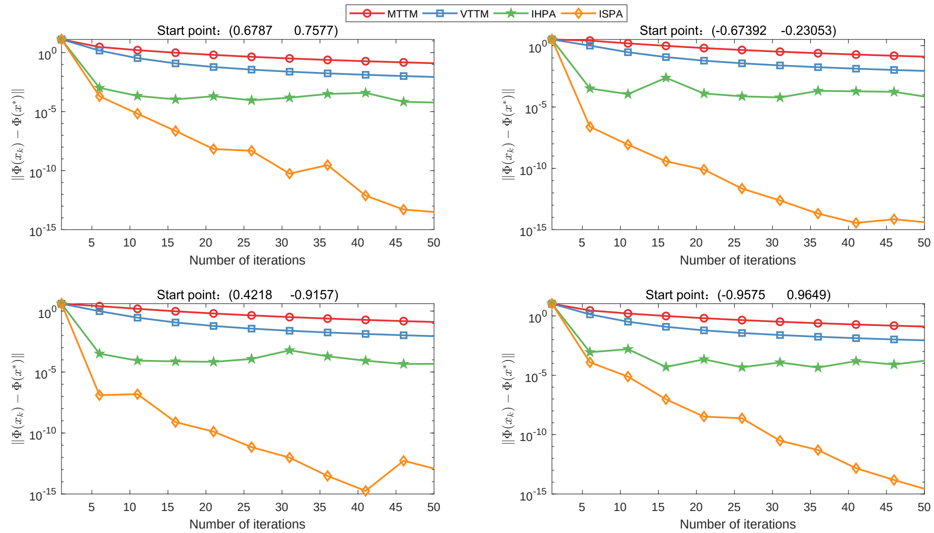

Figure 4. Component convergence behavior of $\|\Phi(x_n) - \Phi(x^*)\|$ with different initial values.

References

[1] Q.H. Ansari, M. Islam and J.C. Yao, *Nonsmooth variational inequalities on Hadamard manifolds,* Appl. Anal., 2020, 99, 340–358.

[2] H. Attouch, X. Goudou and P. Redont, *The heavy ball with friction method, I. the continuous dynamical system: global exploration of the local minima of a real-valued function by asymptotic analysis of a dissipative dynamical system*, Commun. Contemp. Math., 2000, 2(1), 1–34.

[3] H. Attouch and J. Peypouquet, *The rate of convergence of Nesterov's accelerated forward-backward method is actually faster than* $1/k^2$, SIAM J. Optim., 2016, 26(3), 1824–1834.

[4] H. Brezis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, Elsevier, North Holland, 1973.

[5] A. Beck and M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAM J. Imaging Sci., 2009, 2(1), 183–202.

[6] T.H. Cuong, J.C. Yao and N.D. Yen, *Qualitative properties of the minimum sum-of-squares clustering problem,* Optimization, (2020), 10.1080/02331934.2020.1778685.

[7] G. Chen and R. T. Rockafellar, *Convergence rates in forward–backward splitting*, SIAM J. Optim., 1997, 7(2), 421–444.

[8] S.Y. Cho and S.M. Kang, *Approximation of common solutions of variational inequalities via strict pseudocontractions,* Acta Math. Sci. 2012, 32, 1607–1618.

[9] S.Y. Cho, W. Li and S.M. Kang, *Convergence analysis of an iterative algorithm for monotone operators,* J. Inequal. Appl., 2013, 2013, Article ID 199.

[10] S. Y. Cho, *Strong convergence analysis of a hybrid algorithm for nonlinear operators in a Banach space*, J. Appl. Anal. Comput., 2018, 8(1), 19–31.

[11] D. Gabay, *Augmented Lagrangian methods: applications to the numerical solution of boundary-value problems*, Elsevier, North Holland, 2000.

[12] A. Gibali and D.V. Thong, *Tseng type methods for solving inclusion problems and its applications*, Calcolo, 2018, 55(4), 49.

[13] T. H. Kim and H. Xu, *Strong convergence of modified Mann iterations for asymptotically nonexpansive mappings and semigroups*, Nonlinear Anal., 2006, 64(5), 1140–1152.

[14] L. Liu, *A hybrid steepest descent method for solving split feasibility problems involving nonexpansive mappings*, J. Nonlinear Convex Anal., 2019, 20(3), 471–481.

[15] P. L. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Numer. Anal., 1979, 16(6), 964–979.

[16] D. A. Lorenz and T. Pock, *An inertial forward-backward algorithm for monotone inclusions*, J. Math. Imaging Vision, 2015, 51(2), 311–325.

[17] C. Martinez-Yanes and H. Xu, *Strong convergence of the CQ method for fixed point iteration processes*, Nonlinear Anal., 2006, 64(11), 2400–2411.

[18] Y. Nesterov, *A method for solving the convex programming problem with convergence rate* $O(1/k^2)$, Dokl. Akad. Aauk SSSR, 1983, 269, 543–547.

[19] K. Nakajo and W. Takahashi, *Strong convergence theorems for nonexpansive mappings and nonexpansive semigroups*, J. Math. Anal. Appl., 2003, 279(2), 372–379.

[20] G. B. Passty, *Ergodic convergence to a zero of the sum of monotone operators in Hilbert space*, J. Math. Anal. Appl., 1979, 72(2), 383–390.

[21] B. T. Polyak, *Some methods of speeding up the convergence of iteration methods*, Comput. Math. Math. Phys., 1964, 4(5), 1–17.

[22] D.R. Sahu, J.C. Yao, M. Verma and K.K. Shukla, *Convergence rate analysis of proximal gradient methods with applications to composite minimization problems*, Optimization, (2020), 10.1080/02331934.2019.1702040

[23] P. Tseng, *A modified forward-backward splitting method for maximal monotone mappings*, SIAM J. Control Optim., 2000, 38(2), 431–446.

[24] W. Takahashi, Y. Takeuchi and R. Kubota, *Strong convergence theorems by hybrid methods for families of nonexpansive mappings in Hilbert spaces*, J. Math. Anal. Appl., 2008, 341(1), 276–286.

[25] W. Takahahsi and J.C. Yao, *The split common fixed point problem for two finite families of nonlinear mappings in Hilbert spaces*, J. Nonlinear Convex Anal., 2019, 20, 173–195.

[26] X. Zhao, K.F. Ng, C. Li and J.C. Yao, *Linear regularity and linear convergence of projection-based methods for solving convex feasibility problems,* Appl. Math. Optim., 2018, 78, 613–641.