Noname manuscript No.
(will be inserted by the editor)

New inertial factors of a splitting method for monotone inclusions

Yunda Dong · Bingbing Fang · Xiangyang Wang

Received: date / Accepted: date

Abstract In this article, we consider monotone inclusions of two operators in real Hilbert spaces, in which one is further assumed to be Lipschitz continuous, and we suggest adding an inertial term to a splitting method at each iteration. The associated weak convergence is analyzed under standard assumptions. The way of choosing steplength is self-adaptive via some Armijo-like condition, and it still works even if Lipschitz constant is unkonwn. Rudimentary experiments indicate that the inertial term can improve numerical performance for some test problems.

Keywords Monotone inclusions · Splitting method · Inertial term · Steplength · Weak convergence

Mathematics Subject Classification (2000) 58E35 · 65K15

1 Introduction

Let $\mathcal{H}$ be a real infinite-dimensional Hilbert space with usual inner product $\langle x, y\rangle$ and induced norm $\|x\| = \sqrt{\langle x, x\rangle}$ for any $x, y \in \mathcal{H}$. We consider the following monotone inclusion of finding an $x \in \mathcal{H}$ such that

$$0 \in F(x) + B(x), \tag{1}$$

where $F : \mathcal{H} \to \mathcal{H}$ is Lipschitz continuous and monotone, and $B : \mathcal{H} \rightrightarrows \mathcal{H}$ is (possibly multi-valued) maximal monotone, with the effective domain $\mathrm{dom}B := \{x \in \mathcal{H} : B(x) \neq \emptyset\}$. This problem model covers the minimization of convex

✉ Yunda Dong
School of Mathematics and Statistics, Zhengzhou University, Zhengzhou, 450001, PR China
Tel.: +86-0371-67780033
Fax: +86-0371-67780033
E-mail: ydong@zzu.edu.cn

Bingbing Fang
School of Mathematics and Statistics, Zhengzhou University, Zhengzhou, 450001, PR China

Xiangyang Wang
School of Mathematics and Statistics, Zhengzhou University, Zhengzhou, 450001, PR China

functions, computation of saddle points of convex-concave functions, solution of monotone complementarity and variational inequality problems and so on [1].

For the problem above, a very simple iterative procedure is the following forward-backward splitting method [2, 3]:

$$x^{k+1} = (I + \alpha_k B)^{-1}(x^k - \alpha_k F(x^k)),$$

where I stands for the identity mapping, and $\alpha_k > 0$ is a step length.

In the year 2000, Tseng [1] modified the forward-backward splitting method by adding an extra step at each iteration. More specifically, let $\mathcal{X}$ be some closed convex set intersecting the solution set of the problem (1), choose the starting point $x^0 \in \mathcal{X}$. At k-th iteration, for known iterate x^k, choose $\alpha_k > 0$, then the new iterate x^{k+1} is given by

$$y^k = (I + \alpha_k B)^{-1}(x^k - \alpha_k F(x^k)), \tag{2}$$

$$x^{k+1} = P_{\mathcal{X}}[y^k - \alpha_k F(y^k) + \alpha_k F(x^k)]. \tag{3}$$

This iterative scheme described by (2)-(3) is sometimes called Tseng's splitting algorithm.

Subsequently, the first author suggested a relaxed form of Tseng's splitting algorithm in his PhD dissertation [4]: Choose the step length $\alpha_k > 0$ through an Armijo-like condition used in [4, Algorithm 4.2.4]. Compute

$$y^k = (I + \alpha_k B)^{-1}(x^k - \alpha_k F(x^k)). \tag{4}$$

And compute a relaxation factor $\gamma_k > 0$. Then the new iterate is given by

$$x^{k+1} := P_{\mathcal{X}}[x^k - \gamma_k(x^k - y^k - \alpha_k F(x^k) + \alpha_k F(y^k))], \tag{5}$$

where $\mathcal{X}$ is the same set as defined in Tseng's splitting algorithm. When specialized to monotone variational inequality problems, such an iterative scheme just reduces to a projection-type method independently proposed in [5–7]. It has the same nice convergence properties as Tseng's splitting algorithm [4] and demonstrates better numerical performance [8].

Very recently, [9] further studied the projection-type method mentioned above and suggested adding an inertial term. Impressively, they designed a nice test problem and confirmed that the associated inertial term (based on a pioneering work [10]) can improve the method's numerical performance.

Note that the conditions on inertial factors there follow from [11] (see Remark 3.1 below), which seem complicated and restrictive as analyzed in [12].

To avoid this, we instead follow [12] to suggest a new way of choosing the inertial factors. Specifically speaking, we replace x^k in both (4) and (5) by

$$\hat{x}^k = x^k + t_k(x^k - x^{k-1}),$$

and require inertial factor t_k to meet some criterion (6). The other feature is that our way of choosing steplength is self-adaptive via some Armijo-like condition (see Step 2 of Algorithm 1 below), and there is no need to evaluate the corresponding Lipschitz constant in advance.

Our main goal of this article is to show that such doing has two advantages over [11, 9]: simplicity in theory and usefulness in practice.

The rest of this article is organized as follows. In Sect. 2, we specify the notation used and review some useful results. In Sect. 3, we state in details our suggested method. In Sect. 4, we analyze weak convergence of our suggested method under suitable assumptions. In Sect. 5, we did numerical experiments to confirm that our way of choosing inertial factors is useful in practice. In Sect. 6, we close this article by some concluding remarks.

2 Preliminaries

In this section, we give some useful definitions and concepts.

Recall that $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is called monotone if

$$\langle s - s', x - x' \rangle \geq 0, \qquad \text{for all } x,\ x' \in \mathrm{dom} T,\ s \in T(x),\ s' \in T(x');$$

maximal monotone if it is monotone and its graph $\{(x,\ s) : x \in \mathcal{H}, s \in T(x)\}$ can not be enlarged without loss of the monotonicity. In addition, if there exists $\mu > 0$ such that $\langle s - s', x - x' \rangle \geq \mu \|x - x'\|^2$, for all $x,\ x' \in \mathrm{dom} T,\ s \in T(x),\ s' \in T(x')$, then T is usually called μ-strongly monotone. For a mapping $F : \mathcal{H} \to \mathcal{H}$, if there exists $L > 0$ such that

$$\|F(x) - F(y)\| \leq L\|x - y\|, \qquad \text{for all } x,\ y \in \mathcal{H},$$

then F is called L-Lipschitz continuous in $\mathcal{H}$. More on the Euclidean space $\mathcal{R}^n$. Let $f : \mathcal{R}^n \to \mathcal{R} \cup \{+\infty\}$ be a closed proper convex function, then its sub-differential is defined by

$$\partial f(x) = \{s \in \mathcal{R}^n : f(y) - f(x) \geq \langle s, y - x \rangle, \text{ for all } y \in \mathcal{R}^n\},$$

for any given x in $\mathcal{R}^n$. Moreover, if f is further continuously differentiable, then $\partial f(x) = \{\nabla f(x)\}$, where $\nabla f(x)$ is the gradient of f at $x \in \mathcal{R}^n$. Let $\mathcal{C}$ be some nonempty closed convex set in $\mathcal{R}^n$, the usual projection is defined by $P_{\mathcal{C}}(u) = \mathrm{argmin}\{\|u - x\| : x \in \mathcal{C}\}$. The associated indicator function defined by

$$\delta_{\mathcal{C}}(x) = \begin{cases} 0 & \text{if } x \in \mathcal{C}, \\ \infty & \text{if } x \notin \mathcal{C}. \end{cases}$$

is a closed proper convex function. Furthermore, the normal cone operator $N_{\mathcal{C}} := \partial \delta_{\mathcal{C}}$ is maximal monotone and $P_{\mathcal{C}} = (I + \alpha N_{\mathcal{C}})^{-1}$, where $\alpha > 0$.

3 Method

In this section, we describe step by step our suggested method and make some remarks on inertial factors.

Remark 3.1 Note that our way of choosing inertial factors is given by (6). In contrast, in the setting of monotone variational inequality problem [9], the inertial factors there are required to follow [11] to satisfy the following conditions

$$t_0 = 0, \quad 0 \leq t_k \leq t_{k+1} \leq t < 1,$$

Algorithm 3.1

Step 0. Choose $x^{-1}, x^0 \in \mathcal{H}$. Choose $\alpha_{-1} = 1$, $1 \leq \theta < 2$, $\beta \in (0,1)$, $\rho \in (0,1)$ and sufficiently small $\varepsilon > 0$. Calculate

$$t := \frac{2 - \theta(1+\varepsilon)}{2+\theta}. \tag{6}$$

Set $k := 0$.
Step 1. (*Inertial step*) Choose $t_k \in [0, t]$. Compute

$$\hat{x}^k = x^k + t_k(x^k - x^{k-1}). \tag{7}$$

Step 2. (*Steplength*) Set

$$\alpha := \alpha_{k-1}\beta^j,\ j = 0, 1, 2, ..., \quad \hat{x}^k(\alpha) := (I + \alpha B)^{-1}(\hat{x}^k - \alpha F(\hat{x}^k)).$$

Find the smallest nonnegative integer j_k such that

$$\alpha\langle \hat{x}^k - \hat{x}^k(\alpha),\ F(\hat{x}^k) - F(\hat{x}^k(\alpha))\rangle \leq (1-\rho)\|\hat{x}^k - \hat{x}^k(\alpha)\|^2. \tag{8}$$

Set $\alpha_k := \alpha_{k-1}\beta^{j_k}$.
Step 3. (*Proximal step or Forward-backward step*) Compute

$$y^k := \hat{x}^k(\alpha_k) = (I + \alpha_k B)^{-1}(\hat{x}^k - \alpha_k F(\hat{x}^k)). \tag{9}$$

If $y^k = \hat{x}^k$, then stop. Otherwise go to Step 4.
Step 4. (*Descent step*) Calculate

$$d^k = \hat{x}^k - y^k - \alpha_k(F(\hat{x}^k) - F(y^k)), \tag{10}$$

and

$$\gamma_k = \theta\langle \hat{x}^k - y^k, d^k\rangle / \|d^k\|^2. \tag{11}$$

Compute

$$x^{k+1} = \hat{x}^k - \gamma_k d^k. \tag{12}$$

Set $k := k+1$, and go to Step 1.

and $\{\alpha_k\}$ satisfies

$$0 < \underline{\alpha} \leq \alpha_k \leq \frac{\delta(1-t^2) - t^2 - t^3 - t\sigma}{\delta[1 + t(1+t) + t\delta + \sigma]}, \quad \text{with } \delta > \frac{t^2 + t^3 + t\sigma}{1 - t^2}, \tag{13}$$

where $\delta > 0$ and $\sigma > 0$. As analyzed in [12], these conditions are complicated and restrictive in essence.

Remark 3.2 From (8),(10) and (11), it is not difficult to check

$$\gamma_k \geq \theta \frac{\rho}{(1 + \alpha_k L)^2} \geq \gamma_{\min} > 0$$

since the steplength sequence $\{\alpha_k\}$ is bounded above and the forward operator F has been assumed to be L-Lipschitz continuous.

Note that, for the splitting method described by (4) and (5), [13] independently suggested a conceptual one, but without the corresponding self-adaptive choice of α_k, and proved the method's convergence in the finite-dimensional space.

4 Convergence properties

In this section, we analyze convergence behaviours of our suggested Algorithm 3.1. Under standard assumptions, we prove its weak convergence.

This section begins with a well-known result [10], which is used for simplifying the proof of our main theorem in this article.

Lemma 4.1 *Let $\{\varphi_k\}$, $\{t_k\}$ and $\{\delta_k\}$ be nonnegative sequences. Assume that*

$$\varphi_{k+1} \leq \varphi_k + t_k(\varphi_k - \varphi_{k-1}) + \delta_k, \quad k = 0, 1, ...,$$

and $0 \leq t_k \leq t < 1$ and $\sum_{k=0}^{+\infty} \delta_k < +\infty$. Then $\lim_{k\to+\infty} \varphi_k$ exists.

Lemma 4.2 *Assume the sequence $\{\hat{x}^k\}$ generated by the Algorithm 3.1 and x^* is an element of the solution set (if nonempty) of the problem (1). Then*

$$\begin{aligned} &(i) \quad \alpha_k > \beta(1-\rho)/L; \\ &(ii) \quad \langle \hat{x}^k - x^*, d^k \rangle \geq \langle \hat{x}^k - y^k, d^k \rangle. \end{aligned}$$

Proof (i) Obviously, $\beta^{-1}\alpha_k = \alpha_{k-1}\beta^{j_k-1}$ fails to satisfy (8), i.e.,

$$\begin{aligned} &\beta^{-1}\alpha_k \langle \hat{x}^k - \hat{x}^k(\beta^{-1}\alpha_k), F(\hat{x}^k) - F(\hat{x}^k(\beta^{-1}\alpha_k)) \rangle \\ > &(1-\rho)\|\hat{x}^k - \hat{x}^k(\beta^{-1}\alpha_k)\|^2, \end{aligned}$$

which, together with L-Lipschitz continuity of F, implies the desired result.

(ii) It follows from (9) that

$$B(y^k) \ni \alpha_k^{-1}(\hat{x}^k - y^k) - F(\hat{x}^k), \tag{14}$$

which, together with $B(x^*) \ni -F(x^*)$ and monotonicity of B, implies

$$\begin{aligned} 0 &\leq \langle y^k - x^*, \alpha_k^{-1}(\hat{x}^k - y^k) - (F(\hat{x}^k) - F(x^*)) \rangle \\ &= \langle y^k - x^*, \alpha_k^{-1}(\hat{x}^k - y^k) - (F(\hat{x}^k) - F(y^k) + F(y^k) - F(x^*)) \rangle \\ &= \langle y^k - x^*, \alpha_k^{-1}(\hat{x}^k - y^k) - (F(\hat{x}^k) - F(y^k)) \rangle - \langle y^k - x^*, F(y^k) - F(x^*) \rangle \\ &\leq \langle y^k - x^*, \alpha_k^{-1}(\hat{x}^k - y^k) - (F(\hat{x}^k) - F(y^k)) \rangle, \end{aligned}$$

where the last inequality follows from the monotonicity of F. Thus, we get

$$\begin{aligned} &\langle \hat{x}^k - x^*, \alpha_k^{-1}(\hat{x}^k - y^k) - (F(\hat{x}^k) - F(y^k)) \rangle \\ \geq &\langle \hat{x}^k - y^k, \alpha_k^{-1}(\hat{x}^k - y^k) - (F(\hat{x}^k) - F(y^k)) \rangle. \end{aligned}$$

So, the proof is complete. □

Theorem 4.1 *Further assume that the sequence $\{t_k\}$ is nondecreasing. The sequence $\{x^k\}$ generated by the Algorithm 3.1 must converge weakly to an element of the solution set (if nonempty) of the problem (1).*

Proof Let x^* be a solution of the problem (1) above. It follows from Lemma 4.2 and (12) that

$$\begin{aligned}\|x^{k+1}-x^*\|^2 &= \|\hat{x}^k - x^* - \gamma_k d^k\|^2\\ &= \|\hat{x}^k - x^*\|^2 - 2\gamma_k\langle \hat{x}^k - x^*, d^k\rangle + \gamma_k^2\|d^k\|^2\\ &\le \|\hat{x}^k - x^*\|^2 - 2\gamma_k\langle \hat{x}^k - y^k, d^k\rangle + \gamma_k^2\|d^k\|^2.\end{aligned}$$

Combining this with (11) yields

$$\|x^{k+1}-x^*\|^2 \le \|\hat{x}^k - x^*\|^2 - (2-\theta)\gamma_k\langle \hat{x}^k - y^k, d^k\rangle.$$

Meanwhile

$$\begin{aligned}\|\hat{x}^k - x^*\|^2 &= \|x^k - x^* + t_k(x^k - x^{k-1})\|^2\\ &= \|x^k - x^*\|^2 + 2t_k\langle x^k - x^*, x^k - x^{k-1}\rangle + t_k^2\|x^k - x^{k-1}\|^2\\ &= \|x^k - x^*\|^2 + t_k(\|x^k - x^*\|^2 + \|x^k - x^{k-1}\|^2 - \|x^{k-1} - x^*\|^2) + t_k^2\|x^k - x^{k-1}\|^2\\ &= (1+t_k)\|x^k - x^*\|^2 - t_k\|x^{k-1} - x^*\|^2 + t_k(1+t_k)\|x^k - x^{k-1}\|^2.\end{aligned}$$

According to (11) and (12), we get

$$\gamma_k\langle \hat{x}^k - y^k, d^k\rangle = \gamma_k^2\|d^k\|^2/\theta = \|x^{k+1} - \hat{x}^k\|^2/\theta,$$

and

$$\begin{aligned}\|x^{k+1} - \hat{x}^k\|^2 &= \|x^{k+1} - x^k - t_k(x^k - x^{k-1})\|^2\\ &= \|x^{k+1} - x^k\|^2 - 2t_k\langle x^{k+1} - x^k, x^k - x^{k-1}\rangle + t_k^2\|x^k - x^{k-1}\|^2\\ &\ge \|x^{k+1} - x^k\|^2 - t_k\left(\|x^{k+1} - x^k\|^2 + \|x^k - x^{k-1}\|^2\right) + t_k^2\|x^k - x^{k-1}\|^2\\ &= (1-t_k)\|x^{k+1} - x^k\|^2 - (t_k - t_k^2)\|x^k - x^{k-1}\|^2.\end{aligned}$$

Thus, we have

$$\begin{aligned}\|x^{k+1}-x^*\|^2 \le{}& (1+t_k)\|x^k - x^*\|^2 - t_k\|x^{k-1} - x^*\|^2 + t_k(1+t_k)\|x^k - x^{k-1}\|^2\\ &- \frac{2-\theta}{\theta}(1-t_k)\|x^{k+1} - x^k\|^2 + \frac{2-\theta}{\theta}(t_k - t_k^2)\|x^k - x^{k-1}\|^2.\end{aligned} \tag{15}$$

Set $\varphi_k := \|x^k - x^*\|^2$. Since the inertial sequence $\{t_k\}$ is nondecreasing, we further have

$$\begin{aligned}&\varphi_{k+1} - t_k\varphi_k + \frac{2-\theta}{\theta}(1-t_k)\|x^{k+1} - x^k\|^2\\ &\le \varphi_k - t_{k-1}\varphi_{k-1} + \left(t_k(1+t_k) + \frac{2-\theta}{\theta}(t_k - t_k^2)\right)\|x^k - x^{k-1}\|^2\\ &= \varphi_k - t_{k-1}\varphi_{k-1} + \left(\frac{2}{\theta}(t_k - t_k^2) + 2t_k^2\right)\|x^k - x^{k-1}\|^2\\ &= \varphi_k - t_{k-1}\varphi_{k-1} + \lambda_k\|x^k - x^{k-1}\|^2,\end{aligned}$$

where $\lambda_k = \frac{2}{\theta}(t_k - t_k^2) + 2t_k^2$.

Set
$$\psi_k := \varphi_k - t_{k-1}\varphi_{k-1} + \lambda_k \|x^k - x^{k-1}\|^2, \tag{16}$$
then
$$\psi_{k+1} \leq \psi_k - \left(\frac{2-\theta}{\theta}(1-t_k) - \lambda_{k+1}\right)\|x^{k+1} - x^k\|^2.$$

Next, we prove
$$\frac{2-\theta}{\theta}(1-t_k) - \lambda_{k+1} \geq \varepsilon. \tag{17}$$
In fact, it follows from $0 \leq t_k \leq t_{k+1} \leq t < 1$ and $1 \leq \theta < 2$ that
$$\begin{aligned}
\frac{2-\theta}{\theta}(1-t_k) - \lambda_{k+1} &= \frac{2-\theta}{\theta}(1-t_k) - \frac{2}{\theta}t_{k+1}(1-t_{k+1}) - 2t_{k+1}^2 \\
&= (\frac{2}{\theta}-1) - (\frac{2}{\theta}-1)t_k - \frac{2}{\theta}t_{k+1} + \frac{2}{\theta}t_{k+1}^2 - 2t_{k+1}^2 \\
&\geq (\frac{2}{\theta}-1) - (\frac{2}{\theta}-1)t_{k+1} - \frac{2}{\theta}t_{k+1} + \frac{2}{\theta}t_{k+1}^2 - 2t_{k+1}^2 \\
&\geq (\frac{2}{\theta}-1) - (\frac{4}{\theta}-1)t_{k+1} - (2-\frac{2}{\theta})t_{k+1} \\
&= (\frac{2}{\theta}-1) - (\frac{2}{\theta}+1)t_{k+1},
\end{aligned}$$
which, together with (6) and $t_{k+1} \leq t$, implies (17).

Therefore, we get
$$\psi_{k+1} \leq \psi_k - \varepsilon\|x^{k+1} - x^k\|^2. \tag{18}$$
This shows that the sequence $\{\psi_k\}$ is nonincreasing. Then
$$\begin{aligned}
\psi_0 \geq \psi_k &= \varphi_k - t_{k-1}\varphi_{k-1} + \lambda_k\|x^k - x^{k-1}\|^2 \\
&\geq \varphi_k - t_{k-1}\varphi_{k-1} \geq \varphi_k - t\varphi_{k-1} \\
&\geq -t\varphi_{k-1}.
\end{aligned} \tag{19}$$
It follows from $\psi_0 \geq \varphi_k - t\varphi_{k-1}$ that
$$\varphi_k \leq \psi_0 \Sigma_{i=0}^{k-1} t^i + t^k\varphi_0 \leq \frac{\psi_0}{1-t} + t^k\varphi_0,$$
where we notice that $\psi_0 = \varphi_0$ (due to the relation $t_0 = t_{-1} = 0$). Combining this with (18) and (19) yields
$$\begin{aligned}
\varepsilon\Sigma_{i=0}^{k}\|x^{i+1} - x^i\|^2 &\leq \psi_0 - \psi_{k+1} \leq \psi_0 + t\varphi_k \\
&\leq \frac{\psi_0}{1-t} + t^{k+1}\varphi_0 \leq \frac{\psi_0}{1-t} + \varphi_0.
\end{aligned}$$

So, we further get

$$\Sigma_{i=0}^{k}\|x^{i+1}-x^{i}\|^2<\infty \quad \Rightarrow \quad \lim_{k\to\infty}\|x^{k+1}-x^{k}\|=0, \tag{20}$$

which, together with (7) and (12) and Remark 3.2, implies

$$\lim_{k\to\infty}\|d^k\|=0. \tag{21}$$

It follows from (8) and (10) that

$$\langle \hat{x}^k-y^k, d^k\rangle=\langle \hat{x}^k-y^k, \hat{x}^k-y^k-\alpha_k(F(\hat{x}^k)-F(y^k))\rangle \geq \rho\|\hat{x}^k-y^k\|^2.$$

This means

$$\|d^k\|\geq \rho\|\hat{x}^k-y^k\|.$$

Combining this with (21) yields

$$\lim_{k\to\infty}\|\hat{x}^k-y^k\|=0. \tag{22}$$

It easily follows from (15) that

$$\varphi_{k+1}\leq \varphi_k+t_k(\varphi_k-\varphi_{k-1})+\delta_k,$$

where

$$\delta_k:=\left(t(1+t)+t(2-\theta)/\theta\right)\|x^k-x^{k-1}\|^2.$$

From (20) and Lemma 4.1, we can see that the sequence $\{\|x^k-x^*\|\}$ converges. This indicates that the sequence $\{x^k\}$ is bounded in norm, thus it has at least one weak cluster point, say x^∞, i.e., there exists some subsequence $\{x^{k_j}\}$ such that it converges weakly to x^∞. Meanwhile, the iterative formula (12) tells us that

$$x^k-y^k=\hat{x}^k-y^k-\gamma_k d^k+x^k-x^{k+1},$$

which, together with (20) and (21) and (22), implies that the subsequence $\{y^{k_j}\}$ converges weakly to x^∞ as well.

Now let us prove that x^∞ is indeed a solution. In fact, in view of (14), we have

$$(F+B)(y^k)\ni \alpha_k^{-1}(\hat{x}^k-y^k)-(F(\hat{x}^k)-F(y^k)).$$

It follows from monotonicity of $F+B$ that

$$\langle y^k-x, \alpha_k^{-1}(\hat{x}^k-y^k)-(F(\hat{x}^k)-F(y^k))-w\rangle\geq 0, \quad \forall x\in\mathcal{H},\ \forall w\in(F+B)(x).$$

Taking the limit along k_j yields

$$\langle x^\infty-x, 0-w\rangle\geq 0,$$

where we have made use of (22) and the facts that F is L-Lipschitz continuous and $\{\alpha_k\}$ is bounded below stated in part (i) of Lemma 4.2. Combining this with maximality of $F+B$ proves that x^∞ is indeed a solution. As to the proof of the uniqueness of the weak cluster point, we refer to [1] for more details.

Here we would like to stress that Lipschitz continuity assumption on the forward operator F can be further weakened, and we refer to [1, 8] for pertinent discussions.

5 Rudimentary experiments

In this section, we implemented our suggested Algorithm 3.1 to solve some test problem given in [9], and we confirmed that the associated way of choosing inertial factors is practically useful and is better when compared with the one in the iPCA [9]. In our writing style, rather than striving for maximal test problems, we tried to make the basic ideas and techniques as clear as possible.

All numerical experiments were run in MATLAB R2014a (8.3.0.532) with 64-bit (win64) on a desktop computer with an Intel(R) Core(TM) i5-7400 CPU 3.00 GHz and 8.00 GB of RAM. The operating system is Windows 10.

Our test problem is to solve the following variational inequality problem [9]

$$0 \in \partial \delta_C(x) + F(x),$$

where

$$C := \{x := (x_1, x_2)^T : -10 \leq x_1 \leq 100,\ -10 \leq x_2 \leq 100\},$$

and

$$F(x) := (2x + 2y + \sin x,\ -2x + 2y + \sin y)^T.$$

As shown in [9], it is strongly monotone and Lipschitz continuous with the constant $L = \sqrt{26}$. Note that the origin $x^* = (0,0)^T$ is the unique solution.

In practical implementations, we made use of the following stopping criterion

$$\|x^k - x^*\| \leq \epsilon, \quad \text{with } \epsilon = 10^{-8}.$$

We chose $x^{-1} = x^0$. We set $\beta = 0.8$ and $\rho = 0.4$. We set $\theta = 1.5$ and $\varepsilon = 10^{-9}$ and thus $t = 0.142857$ follows from the formula (6). Thus, we set $t_k \equiv 0.14$.

We reported the corresponding numerical results in the following Table 1, where "TIME" means the elapsed time using tic and toc (in seconds). And, we wrote

iPCA: It corresponds to the algorithm suggested in [9], in which $\theta = 1.5$, $t_k \equiv 0.4$ and $\alpha_k \equiv 0.5/\sqrt{26}$ were chosen for the test problem above.

iPDA: It corresponds to our suggested Algorithm 3.1 and the associated parameters were just specified above.

iPDA0: It corresponds to the iPDA above, but without using Step 2 to determine the steplength. It instead uses $\alpha_k \equiv 0.17$ (approximately $0.87/\sqrt{26}$).

PDA: It corresponds to the iPDA above, but $t_k \equiv 0$, i.e., it is the method described by (4) and (5).

Table 1 Numerical results with different starting points

	$x^0 = (1,10)^T$			$x^0 = (-100,100)^T$		
	k	TIME	$\|x^k - x^*\|$	k	TIME	$\|x^k - x^*\|$
iPCA	39	0.00071	8.504135e-09	44	0.00082	9.090472e-09
iPDA0	14	0.00032	1.371550e-09	17	0.00043	3.822849e-09
iPDA	12	0.00124	2.072000e-09	15	0.00145	5.810037e-10
PDA	16	0.00156	3.027479e-09	19	0.00182	3.005567e-09

From Table 1, we can see that (i) our way of choosing inertial factors is simpler in theory and more efficient in practice and (ii) the inertial term can improve numerical performance in a desirable way for the test problem above. Furthermore, our chosen parameters meet all the conditions required.

From Table 1, we also can see that, if we have already known a good evaluation of Lipsctitz constant of F, then we may adopt iPDA0 since it usually can save time. Otherwise, we shall use iPDA because it can be costly, even impractical, to evaluate such constant.

By the way, we also run the iPDA for $t_k \equiv 0.01$, $t_k \equiv 0.1$ and observed similar results. This indicates that the iPDA is not sensitive to choices of the inertial factors since they may vary in a wide range of from 0.01 to 0.14.

6 Conclusions

In this article, we have considered the monotone inclusion in a real infinite-dimensional Hilbert space and a relaxed form of Tseng's splitting method. We have suggested adding an inertial term to this method at each iteration. Importantly, the conditions on the associated inertial factors are simple in theory and useful in practice when compared with the existing ones. Furthermore, we can implement it with self-adaptive steplength and thus there is no need to evaluate Lipschitz constant of the forward operator beforehand. Rudimentary experiments had supported our analysis and viewpoints.

Acknowledgments We are very grateful to Xixian Bai at Shandong University for writing Matlab codes.

References

1. Tseng P.: A modified forward-backward splitting method for maximal monotone mappings. SIAM J. Control Optim. 38(2), 431-446 (2000)
2. Lions P.L., Mercier B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. 16, 964-979 (1979)
3. Passty G.B.: Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. J. Math. Anal. Appl. 72, 383-390 (1979)
4. Dong Y.D.: Splitting Methods for Monotone Inclusions. PhD dissertation, Nanjing University (2003)
5. Sun D.F.: A class of iterative methods for solving nonlinear projection equations. J. Optim. Theory Appl. 91(1), 123-140 (1996)
6. Solodov M.V., Tseng P.: Modified projection-type methods for monotone variational inequalities. SIAM J. Control Optim. 34, 1814-1830 (1996)
7. He B.S.: A class of projection and contraction methods for monotone variational inequalities. Appl. Math. Optim. 35, 69-76 (1997)
8. Huang Y.Y., Dong Y.D.: New properties of forward-backward splitting and a practical proximal-descent algorithm. Appl. Math. Comput. 237, 60-68 (2014)
9. Dong Q. L., Cho Y. J., Zhong L. L. and Rassias Th. M.: Inertial projection and contraction algorithms for variational inequalities. J. Global Optim. 70(3), 687-704 (2018)
10. Alvarez F., Attouch H.: An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. Set-Val Anal. 9(1), 3–11 (2001)
11. Bot R.I., Csetnek E.R. and Hendrich C.: Inertial Douglas-Rachford splitting for monotone inclusion problems. Appl. Math. Comput. 256(1), 472–487 (2015)
12. Dong Y.D.: New inertial factors of the Krasnosel'skiĭ-Mann iteration. manuscript. (2017)
13. Noor M.A.: Mixed quasi-variational inequalities. Appl. Math. Comput. 146, 553-578 (2003)