

AN INERTIAL SPLITTING METHOD FOR MONOTONE INCLUSIONS OF THREE OPERATORS

Yunda Dong* and Xiao Zhu

School of Mathematics and Statistics, Zhengzhou University, 450001, Zhengzhou, PR China,
*corresponding author. ydong@zzu.edu.cn
Zhu is the second author with email: zhuXCiao5739@163.com

ARTICLE INFO

Received: 24 February 2022
Revised: 18 April 2022
Accepted: 30 April 2022
Online: 23 June 2022

To cite this paper:
Yunda Dong & Xiao Zhu
(2022). An Inertial Splitting
Method for Monotone
Inclusions of three
Operators. *International
Journal of Mathematics,
Statistics and Operations
Research*. 2(1): pp. 43-60.
[https://DOI: 10.47509 /
IJMSOR.2022.v02i01.04](https://doi.org/10.47509/IJMSOR.2022.v02i01.04)

ABSTRACT

In this article, we consider monotone inclusions of three operators in real Hilbert spaces and suggest an inertial version of a generalized Douglas-Rachford splitting. Under standard assumptions, we prove its weak and strong convergence properties. The newly-developed proof techniques are based on the characteristic operator and thus are more self-contained and less convoluted. Rudimentary experiments demonstrated that our suggested inertial splitting method can efficiently solve some large-scale test problems.

Keywords: Monotone inclusions Characteristic operator Inverse strongly monotone Splitting method Weak convergence

Mathematics Subject Classification (2000) 58E35 . 65K15

1 Introduction

Very recently, there has been a renewed interest in adding inertial terms to various splitting methods for monotone inclusions of two operators in real Hilbert spaces Dong et al. (2018), Thong et al. (2018), Alves et al. (2020), Dong (2021a), Dong et al. (2021).

The idea of these inertial splitting methods can date back to a pioneering work Alvarez et al. (2001), where the proximal point algorithm Martinet (1970), Rockafellar (1976) was well considered by adding inertial terms.

In this article, we aim at adding inertial terms to splitting method used for solving the following monotone inclusions of three operators in real Hilbert space \mathcal{H} :

$$0 \in F(x) + B(x), \quad \text{with } F := C + A, \quad (1)$$

where the inverse of $C : \mathcal{H} \rightarrow \mathcal{H}$ is strongly monotone, $A, B : \mathcal{H} \rightrightarrows \mathcal{H}$ are maximal monotone.

Specifically speaking, to solve (1), for $k = 0, 1, \dots$, and the starting points $z^{-1}, z^0 \in \mathcal{H}$, we suggest the following inertial splitting method

$$\begin{aligned} \hat{z}^k &= z^k + t_k(z^k - z^{k-1}), \\ (J + A)(x^k) &\ni J\hat{z}^k, \\ (J + B)(y^k) &\ni 2Jx^k - J\hat{z}^k - C(x^k), \\ z^{k+1} &= \hat{z}^k - \gamma_k M^{-1}J(x^k - y^k), \end{aligned}$$

where $t_k > 0$ is an inertial factor and $J, M : \mathcal{H} \rightarrow \mathcal{H}$ are linear, bounded and strongly monotone and $\gamma_k > 0$. Henceforth, we call it iDR3 for short. For pertinent discussions, we refer to Hieu et al. (2018), Iyiola et al. (2021), Dixit et al. (2021) and the references cited therein.

If C vanishes and

$$t_k \equiv 0, \quad J := \alpha^{-1}I, \quad M := \alpha^{-1}I,$$

where $\alpha > 0$, then it just reduces to the Douglas-Rachford splitting method of Lions and Mercier Lions et al. (1979). See Eckstein et al. (1992), Dong et al. (2010) for related discussions

To guarantee weak convergence of iDR3, we assume that the inertial sequence $\{t_k\}$ satisfies

$$0 = t_0 \leq t_k \leq t_{k+1} \leq \begin{cases} t(\theta_k, \theta_{k+1}, \varepsilon), & \text{if } \theta_k \in (1, 2), \\ (1 - \varepsilon)/3, & \text{if } \theta_k \equiv 2, \end{cases} \quad (2)$$

where ε is any given sufficiently small positive number and $t(\theta_k, \theta_{k+1}, \varepsilon)$ is defined in (5) below. Notice that the first condition in (2) follows the style of Dong (2021a) and the second in (2) mimic those originally given in Alvarez et al. (2001).

The rest of this article is organized as follows. In Sect. 2, we give some useful concepts and preliminary results. In Sect. 3, we formally describe iDR3. In Sect. 4, under standard assumptions, we analyze weak and strong convergence of iDR3. For strong convergence, our proof technique follows from Dong et al. (2010), and for weak convergence, our proof appears to be more self-contained and less convoluted via introducing the characteristic operator. In Sect. 5, we compare iDR3 with existing results. In Sect. 6, we discuss how to apply iDR3 to solving more general monotone inclusion (40), in which the operator B in (1) has been linearly composed. See Algorithm 6.1 below for more details. In Sect. 7, we did rudimentary numerical experiments to confirm practical usefulness of iDR3. In Sect. 8, we close this article by some concluding remarks.

2 Preliminary Results

In this section, we first give some basic definitions and then provide some auxiliary results for later use.

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 $x, y \in \mathcal{H}$. Let $\mathcal{BL}(\mathcal{H})$ be the set of all nonzero, bounded, linear operators in \mathcal{H} . If $S \in \mathcal{BL}(\mathcal{H})$ is further self-adjoint and strongly monotone, then we use $\|x\|_S$ to stand for

$\sqrt{\langle x, Sx \rangle}$ for all $x \in \mathcal{H}$. I stands for the identity operator, i.e., $Ix = x$ for all $x \in \mathcal{H}$. $\text{dom}T$ stands for the effective domain of T , i.e., $\text{dom}T := \{x \in \mathcal{H} : Tx \neq \emptyset\}$. Throughout this article, for any given $J \in \mathcal{BL}(\mathcal{H})$, one may split it into

$$J = J^+ + J^-, \quad \text{with } J^+ := 0.5(J + J^*), \quad J^- := 0.5(J - J^*), \quad (3)$$

where J^* stands for the adjoint operator of J . notice that such adjoint operator must exist uniquely. For any given $J, J' \in \mathcal{BL}(\mathcal{H})$, we use the notation $J \succeq J'$ ($J \succ J'$) to stand for that $J - J'$ is monotone (strongly monotone). This is the corresponding Löwner partial ordering between two bounded and linear operators.

Definition 1 Let $T : \mathcal{H} \rightarrow \mathcal{H}$ be a single-valued operator. If there exists some constant number $\kappa > 0$ such that

$$\|T(x) - T(y)\| \leq \kappa \|x - y\|, \quad \forall x, y \in \mathcal{H},$$

then T is called Lipschitz continuous.

To concisely give the following definition, we agree on that the notation $(x, w) \in T$ and $x \in \mathcal{H}$, $w \in T(x)$ have the same meaning. Moreover, $w \in Tx$ if and only if $x \in T^{-1}w$, where T^{-1} stands for the inverse of T .

Definition 2 Let $\mathcal{C} \subseteq \mathcal{H}$ be a nonempty subset. An operator $T : \mathcal{C} \rightarrow \mathcal{C}$ is called non-expansive if and only if

$$\|T(x) - T(y)\| \leq \|x - y\|, \quad \forall x, y \in \mathcal{C};$$

firmly non-expansive if and only if

$$\|T(x) - T(y)\|^2 \leq \langle x - y, T(x) - T(y) \rangle, \quad \forall x, y \in \mathcal{C}.$$

Definition 3 Let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be an operator. It is called monotone if and only if

$$\langle x - x', w - w' \rangle \geq 0, \quad \forall (x, w) \in A, \quad \forall (x', w') \in A;$$

maximal monotone if and only if it is monotone and for given $\hat{x} \in \mathcal{H}$ and $\hat{w} \in \mathcal{H}$ the following implication relation holds

$$\langle x - \hat{x}, w - \hat{w} \rangle \geq 0, \quad \forall (x, w) \in A \Rightarrow \forall (\hat{x}, \hat{w}) \in A.$$

Definition 4 Let $T : \mathcal{H} \rightrightarrows \mathcal{H}$ be an operator, T is called uniformly monotone if there exists an increasing function $\phi_T : [0, +\infty) \rightarrow [0, +\infty)$ that $\phi_T(t) = 0$ if and only if $t = 0$, and

$$\langle x - x', w - w' \rangle \geq \phi_T(\|x - x'\|), \quad \forall (x, w) \in T, \quad \forall (x', w') \in T.$$

In the case of $\phi_T(t) = \mu_T t^2$ with $\mu_T > 0$, T is called μ_T -strongly monotone.

Definition 5 Let $C : \mathcal{H} \rightarrow \mathcal{H}$ be an operator. C^{-1} is called c -strongly monotone if there exists some $c > 0$ such that

$$\langle x - y, C(x) - C(y) \rangle \geq c \|C(x) - C(y)\|^2, \forall x, y \in \mathcal{H}.$$

In particular, if $C(x) = Mx + q$, where M is an $n \times n$ positive semi-definite matrix and q is an n -dimensional vector, then

$$\langle x, Mx \rangle \geq \lambda_{\max}^{-1} \|Mx\|^2, \quad \forall x \in \mathbb{R}^n,$$

where λ_{\max} is the largest eigenvalue of M .

Definition 6 Let $f : \mathcal{H} \rightarrow (-\infty, +\infty]$ be a closed, proper and convex function. Then for any given $x \in \mathcal{H}$ the sub-differential of f at x is defined by

$$\partial f(x) := \{s \in \mathcal{H} : f(y) - f(x) \geq \langle s, y - x \rangle, \forall y \in \mathcal{H}\}.$$

Each s is called a sub-gradient of f at x . 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 .

It is well known that the sub-differential of any closed proper convex function in an infinite-dimensional Hilbert space is maximal monotone as well.

For any given maximal monotone operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$, it is Minty (1962) who proved that there must exist a unique $y \in \mathcal{H}$ such that $(I + \lambda A)y \ni x$ for all $x \in \mathcal{H}$ and $\lambda > 0$, where I stands for the identity operator, i.e., $Ix = x$ for all $x \in \mathcal{H}$. This implies that the corresponding operator $J_{\lambda A} := (I + \lambda A)^{-1}$, also called the resolvent of A , is single-valued. Consider the following indicator function

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

where \mathcal{C} is some nonempty closed convex set in \mathbb{R}^n . Then, its sub-differential must be closed, proper convex. Furthermore, for any given positive number $\lambda > 0$, we have $P_{\mathcal{C}} = (I + \lambda \delta_{\mathcal{C}})^{-1}$, where $P_{\mathcal{C}}$ is usual projection onto \mathcal{C} .

Lemma 2.1 Let A, B, C be operators defined in the problem (1). Then the resulting characteristic operator via Attouch-Thera duality principle

$$T(x, u, v) := \begin{pmatrix} A \\ B^{-1} \\ C^{-1} \end{pmatrix} \begin{pmatrix} x \\ u \\ v \end{pmatrix} + \begin{pmatrix} 0 & I & I \\ -I & 0 & 0 \\ -I & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ u \\ v \end{pmatrix} \quad (4)$$

must be maximal monotone.

Proof Note that C , A and B are maximal monotone. Thus, the first operator on the right-hand side must be maximal monotone and the second must be maximal monotone as well. Maximality of T follows from Rockafellar (1970).

Lemma 2.2 Consider any maximal monotone operator $T : \mathcal{H} \rightarrow \mathcal{H}$. Assume that the sequence $\{w^k\}$ in \mathcal{H} converges weakly to w , and the sequence $\{s^k\}$ on $\text{dom}T$ converges strongly to s . If $T(w^k) \ni s^k$ for all k , then the relation $T(w) \ni s$ must hold.

3 Method

In this section, we formally state iDR3.

Now we set

$$t(\theta_k, \theta_{k+1}, \varepsilon) := \sqrt{p_k^2 + q_k} - p_k, \tag{5}$$

where

$$p_k := \frac{1}{2} \frac{\theta_k + \theta_{k+1} - 1}{2 - \theta_{k+1}}, \quad q_k := \frac{\theta_k - 1 - \varepsilon}{2 - \theta_{k+1}},$$

and ε is any given sufficiently small positive number. For the inertial sequence $\{t_k\}$, we choose it via $t_0 = 0$ and

$$t_k \leq t_{k+1} \leq \begin{cases} t(\theta_k, \theta_{k+1}, \varepsilon), & \text{if } \theta_k \in (1, 2), \\ (1 - \varepsilon)/3, & \text{if } \theta_k \equiv 2. \end{cases} \tag{6}$$

Next, we formally state

Algorithm 3.1

Step 0. Choose $J, M \in \mathcal{BC}(\mathcal{H})$. Choose $z^0, z^{-1} \in \mathcal{H}$. Choose $\theta_0 = \theta_{-1} = 2/1.9$, $t_0 = t_{-1} = 0$. Compute c and

$$D_0 := J - \frac{1}{4c}I. \tag{7}$$

Set $k := 0$.

Step 1. Compute

$$\hat{z}^k = z^k + t_k(z^k - z^{k-1}), \tag{8}$$

$$(J + A)(x^k) \ni J\hat{z}^k, \tag{9}$$

$$(J + B)(y^k) \ni 2Jx^k - J\hat{z}^k - C(x^k). \tag{10}$$

If some stopping criterion is met, then stop. Otherwise, go to Step 2.

Step 2 Compute

$$\gamma_k = 2\theta_k^{-1} \|x^k - y^k\|_{D_0}^2 / \|J(x^k - y^k)\|_{M^{-1}}^2, \tag{11}$$

$$z^{k+1} = \hat{z}^k - \gamma_k M^{-1} J(x^k - y^k). \tag{12}$$

Choose $\theta_{k+1} \in (1, 2]$ and t_{k+1} by (6). Set $k := k + 1$, and go to Step 1.

Notice that, in practical implementations, it seems convenient to directly use the following Table 1 for upper bounds of the inertial sequence.

Table 1: Numerical demonstration of the conditions (6) with $\varepsilon = 0.0001$

$2/\theta_k$	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
t	0.333	0.303	0.274	0.245	0.216	0.186	0.154	0.121	0.085	0.045

From this table, we can see that, in the case of θ_k being some constant, $t(\theta_k, \theta_{k+1}, \varepsilon) < 1/3$ holds for several special values of θ_k . In fact, it is direct to show that

$$p_k^2 + q_k < (p_k + 1/3)^2 \quad \Rightarrow \quad \sqrt{p_k^2 + q_k} < p_k + 1/3.$$

Thus, we have

$$t(\theta_k, \theta_{k+1}, \varepsilon) = \sqrt{p_k^2 + q_k} - p_k < 1/3.$$

In Algorithm 3.1, if we set

$$J := \alpha^{-1}I, \quad M := \alpha^{-1}I, \quad (13)$$

then we shall choose $\alpha \in (0, 4c)$ so that D_0 in (7) is positive definite. Meanwhile, the formula of calculating γ_k reduces to

$$\gamma_k = 2(1 - \alpha/4c)\theta_k^{-1}, \quad 1 < \theta_k \leq 2. \quad (14)$$

In this case, the corresponding algorithm coincides with a modified version Davis (2015) of the Douglas-Rachford splitting method of Lions and Mercier.

In Algorithm 3.1, if $J := \alpha^{-1}I$ and B is taken to be the differential of some nonempty closed convex subset, then (10) reduces to

$$(I + \alpha B)(y^k) \ni 2x^k - z^k - \alpha C(x^k).$$

Thus, y^k is usual projection of $2x^k - z^k - \alpha C(x^k)$ onto this subset.

4 Convergence properties

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

This section begins with the celebrated lemma due to Alvarez et al. (2001), 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*

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

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

Also, we need to make use of the following well-known results.

Lemma 4.2 *If $M \in \mathcal{BL}(\mathcal{H})$ is self-adjoint and strongly monotone. Then the following*

$$\|(1+t)u - tv\|_M^2 = (1+t)\|u\|_M^2 - t\|v\|_M^2 + t(1+t)\|u-v\|_M^2, \quad (15)$$

$$2\langle u, v \rangle \leq \|u\|_M^2 + \|v\|_M^2, \quad (16)$$

hold for all $u, v \in \mathcal{H}$ and $t \in \mathbb{R}$.

To analyze convergence behaviours of Algorithm 3.1, we make

Assumption 4.1. Assume that

- (i) The operator C is c -inverse strongly monotone.
- (ii) The backward operator $B : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone.
- (iii) A, B is uniformly monotone.

(iv) $J \in \mathcal{BL}(\mathcal{H})$ is self-adjoint and strongly monotone.

(v) $M \in \mathcal{BL}(\mathcal{H})$ is self-adjoint and strongly monotone.

For Algorithm 3.1, the corresponding (9) reads

$$A(x^k) \ni J(\hat{z}^k - x^k) \Leftrightarrow a^k = J(\hat{z}^k - x^k), \quad \text{where } a^k \in A(x^k). \quad (17)$$

Clearly, if x^* is an element of the solution set (if nonempty) of the problem (1), then there must be \hat{z}^* such that

$$A(x^*) \ni J(\hat{z}^* - x^*) \Leftrightarrow a^* = J(\hat{z}^* - x^*), \quad \text{where } a^* \in A(x^*). \quad (18)$$

Based on these observations, we introduce the following lemma.

Lemma 4.3 *If Assumption 4.1 holds, then*

$$\langle J\hat{z}^k - J\hat{z}^*, x^k - y^k \rangle \geq \|x^k - y^k\|_{D_0}^2 + \phi_A(\|x^k - x^*\|) + \phi_B(\|y^k - x^*\|).$$

Proof It follows from (9) and (10) that

$$B(y^k) \ni J(x^k - y^k) - C(x^k) - a^k, \quad \text{where } a^k \in A(x^k), \quad (19)$$

which, together with $B(x^*) \ni -C(x^*) - a^*$, $a^* \in A(x^*)$ and uniform monotonicity of B , implies

$$\begin{aligned} \phi_B(\|y^k - x^*\|) &\leq \langle y^k - x^*, J(x^k - y^k) - (C(x^k) - C(x^*)) - (a^k - a^*) \rangle \\ &= \langle y^k - x^*, J(x^k - y^k) - (a^k - a^*) \rangle - \langle y^k - x^*, C(x^k) - C(x^*) \rangle. \end{aligned}$$

Combining this with that C' 's c -inverse strong monotonicity yields

$$\begin{aligned} &\langle y^k - x^*, J(x^k - y^k) - (a^k - a^*) \rangle \\ &\geq \langle y^k - x^*, C(x^k) - C(x^*) \rangle + \phi_B(\|y^k - x^*\|) \\ &= \langle x^k - x^*, C(x^k) - C(x^*) \rangle - \langle x^k - y^k, C(x^k) - C(x^*) \rangle + \phi_B(\|y^k - x^*\|) \\ &\geq c\|C(x^k) - C(x^*)\|^2 - \frac{1}{2} \left(\frac{1}{2c}\|x^k - y^k\|^2 + 2c\|C(x^k) - C(x^*)\|^2 \right) + \phi_B(\|y^k - x^*\|) \\ &\geq -\frac{1}{4c}\|x^k - y^k\|^2 + \phi_B(\|y^k - x^*\|). \end{aligned}$$

Since J is self-adjoint, we further get

$$\begin{aligned} &-\frac{1}{4c}\|x^k - y^k\|^2 + \phi_B(\|y^k - x^*\|) \\ &\leq \langle x^k - x^* - (x^k - y^k), J(x^k - y^k) - (a^k - a^*) \rangle \\ &= \langle x^k - x^*, J(x^k - y^k) \rangle + \langle x^k - y^k, a^k - a^* \rangle - \langle x^k - y^k, J(x^k - y^k) \rangle - \langle x^k - x^*, a^k - a^* \rangle \\ &= \langle Jx^k + a^k - (Jx^* + a^*), x^k - y^k \rangle - \langle x^k - y^k, J(x^k - y^k) \rangle - \langle x^k - x^*, a^k - a^* \rangle, \end{aligned}$$

which, together with uniform monotonicity of A , indicates

$$\begin{aligned}
& \langle Jx^k + a^k - (Jx^* + a^*), x^k - y^k \rangle \\
& \geq \langle x^k - y^k, J(x^k - y^k) \rangle + \langle x^k - x^*, a^k - a^* \rangle - \frac{1}{4c} \|x^k - y^k\|^2 + \phi_B(\|y^k - x^*\|) \\
& \geq \langle x^k - y^k, J(x^k - y^k) \rangle - \frac{1}{4c} \|x^k - y^k\|^2 + \phi_A(\|x^k - x^*\|) + \phi_B(\|y^k - x^*\|) \\
& = \langle x^k - y^k, (J - \frac{1}{4c}I)(x^k - y^k) \rangle + \phi_A(\|x^k - x^*\|) + \phi_B(\|y^k - x^*\|) \\
& = \|x^k - y^k\|_{D_0}^2 + \phi_A(\|x^k - x^*\|) + \phi_B(\|y^k - x^*\|).
\end{aligned}$$

Thus, by (17) and (18), we have

$$\begin{aligned}
\langle Jz^k - Jz^*, x^k - y^k \rangle &= \langle Jx^k + a^k - (Jx^* + a^*), x^k - y^k \rangle \\
&\geq \|x^k - y^k\|_{D_0}^2 + \phi_A(\|x^k - x^*\|) + \phi_B(\|y^k - x^*\|).
\end{aligned}$$

Then, the proof is complete.

Theorem 4.1 *If Assumption 4.1 holds and there exists some $\rho > 0$ such that*

$$D_0 := J - \frac{1}{4c}I \succeq \rho I. \quad (20)$$

Then, (i) the sequence $\{x^k\}$ generated by Algorithm 3.1 must converge weakly to an element of the solution set (if nonempty) of the monotone inclusion (1); (ii) if either A or B is uniformly monotone, then this sequence is strongly convergent.

Proof (i) Let x^* be a solution of the monotone inclusion (1) above. It follows from Lemma 4.3, (11) and (12) that

$$\begin{aligned}
& \|z^{k+1} - z^*\|_M^2 \\
&= \|\hat{z}^k - z^* - \gamma_k M^{-1} J(x^k - y^k)\|_M^2 \\
&= \|\hat{z}^k - z^*\|_M^2 - 2\gamma_k \langle J\hat{z}^k - Jz^*, x^k - y^k \rangle + \gamma_k^2 \|J(x^k - y^k)\|_{M^{-1}}^2 \\
&\leq \|\hat{z}^k - z^*\|_M^2 - 2\gamma_k \|x^k - y^k\|_{D_0}^2 + \gamma_k^2 \|J(x^k - y^k)\|_{M^{-1}}^2 \\
&\quad - 2\gamma_k (\phi_A(\|x^k - x^*\|) + \phi_B(\|y^k - x^*\|)) \\
&\leq \|\hat{z}^k - z^*\|_M^2 - 2\gamma_k \|x^k - y^k\|_{D_0}^2 + \gamma_k^2 \|J(x^k - y^k)\|_{M^{-1}}^2 \\
&= \|\hat{z}^k - z^*\|_M^2 - 2(1 - \theta_k^{-1})\gamma_k \|x^k - y^k\|_{D_0}^2,
\end{aligned} \quad (21)$$

which, together with (15), implies

$$\begin{aligned}
\|\hat{z}^k - z^*\|_M^2 &= \|(1 + t_k)(z^k - \hat{z}^*) - t_k(z^{k-1} - \hat{z}^*)\|_M^2 \\
&= (1 + t_k)\|z^k - \hat{z}^*\|_M^2 - t_k\|z^{k-1} - \hat{z}^*\|_M^2 + t_k(1 + t_k)\|z^k - z^{k-1}\|_M^2.
\end{aligned} \quad (22)$$

By (11) and (12), we get

$$\gamma_k \|x^k - y^k\|_{D_0}^2 = \frac{1}{2}\theta_k \gamma_k^2 \|J(x^k - y^k)\|_{M^{-1}}^2 = \frac{1}{2}\theta_k \|\gamma_k M^{-1} J(x^k - y^k)\|_M^2. \quad (23)$$

Meanwhile, it follows from (12) that

$$\begin{aligned}
 & \|\gamma_k M^{-1} J(x^k - y^k)\|_M^2 \\
 &= \|z^{k+1} - \hat{z}^k\|_M^2 \\
 &= \|z^{k+1} - z^k - t_k(z^k - z^{k-1})\|_M^2 \\
 &= \|z^{k+1} - z^k\|_M^2 - 2t_k \langle z^{k+1} - z^k, M(z^k - z^{k-1}) \rangle + t_k^2 \|z^k - z^{k-1}\|_M^2 \\
 &\geq \|z^{k+1} - z^k\|_M^2 - t_k (\|z^{k+1} - z^k\|_M^2 + \|M(z^k - z^{k-1})\|_{M^{-1}}^2) + t_k^2 \|z^k - z^{k-1}\|_M^2 \\
 &= (1 - t_k) \|z^{k+1} - z^k\|_M^2 - (t_k - t_k^2) \|z^k - z^{k-1}\|_M^2,
 \end{aligned} \tag{24}$$

where the last inequality comes from (16).

Thus, in terms of (21), (22), (23) and (24), we have

$$\begin{aligned}
 & \|z^{k+1} - \hat{z}^*\|_M^2 \\
 &\leq (1 + t_k) \|z^k - \hat{z}^*\|_M^2 - t_k \|z^{k-1} - \hat{z}^*\|_M^2 - (\theta_k - 1)(1 - t_k) \|z^{k+1} - z^k\|_M^2 \\
 &\quad + (\theta_k t_k + (2 - \theta_k) t_k^2) \|z^k - z^{k-1}\|_M^2.
 \end{aligned} \tag{25}$$

Set

$$\begin{aligned}
 \varphi_k &:= \|z^k - \hat{z}^*\|_M^2, \quad \lambda_k := \theta_k t_k + (2 - \theta_k) t_k^2, \\
 \psi_k &:= \varphi_k - t_{k-1} \varphi_{k-1} + \lambda_k \|z^k - z^{k-1}\|_M^2.
 \end{aligned}$$

It follows from (25) and $0 \leq t_k \leq t_{k+1} \leq t < 1$, we have

$$\varphi_{k+1} \leq (1 + t_k) \varphi_k - t_{k-1} \varphi_{k-1} - (\theta_k - 1)(1 - t_k) \|z^{k+1} - z^k\|_M^2 + \lambda_k \|z^k - z^{k-1}\|_M^2.$$

Thus, we further have

$$\psi_{k+1} \leq \psi_k - ((\theta_k - 1)(1 - t_k) - \lambda_{k+1}) \|z^{k+1} - z^k\|_M^2.$$

Next, we consider

$$\begin{aligned}
 & (\theta_k - 1)(1 - t_k) - \lambda_{k+1} \\
 &= \theta_k - 1 - (\theta_k - 1)t_k - \theta_{k+1}t_{k+1} - (2 - \theta_{k+1})t_{k+1}^2 \\
 &\geq \theta_k - 1 - (\theta_k - 1)t_{k+1} - \theta_{k+1}t_{k+1} - (2 - \theta_{k+1})t_{k+1}^2 \\
 &= \theta_k - 1 - (\theta_k + \theta_{k+1} - 1)t_{k+1} - (2 - \theta_{k+1})t_{k+1}^2.
 \end{aligned}$$

To guarantee

$$\theta_k - 1 - (\theta_k + \theta_{k+1} - 1)t_{k+1} - (2 - \theta_{k+1})t_{k+1}^2 \geq \varepsilon,$$

we consider two cases. If $\theta_k \equiv 2$, then

$$t_{k+1} \leq \frac{\theta_k - 1 - \varepsilon}{\theta_k + \theta_{k+1} - 1} \quad \Rightarrow \quad t_k \leq \frac{1 - \varepsilon}{3}. \tag{26}$$

If $\theta_k \in (1, 2)$, then

$$t_{k+1} \leq t(\theta_k, \theta_{k+1}, \varepsilon), \tag{27}$$

where $t(\theta_k, \theta_{k+1}, \varepsilon)$ is given by (5). Thus, in each case, we always have $t_k \leq t < 1/3$. Therefore, we get

$$\psi_{k+1} \leq \psi_k - \varepsilon \|z^{k+1} - z^k\|_M^2, \quad (28)$$

which, together with nonincreasing property of $\{\psi_k\}$, implies

$$\begin{aligned} \psi_0 &\geq \psi_k = \varphi_k - t_{k-1}\varphi_{k-1} + \lambda_k \|z^k - z^{k-1}\|_M^2 \geq \varphi_k - t_{k-1}\varphi_{k-1} \\ &\geq \varphi_k - t\varphi_{k-1} \geq -t\varphi_{k-1}. \end{aligned} \quad (29)$$

It follows from $\psi_0 \geq \varphi_k - t\varphi_{k-1}$ that

$$\varphi_k \leq \psi_0 \sum_{i=0}^{k-1} t^i + t^k \varphi_0 \leq \frac{\psi_0}{1-t} + t^k \varphi_0,$$

where we notice that $\psi_0 = \varphi_0$ (due to the relation $t_0 = t_{-1} = 0$).

Combining (28) with (29) yields

$$\varepsilon \sum_{i=0}^k \|z^{i+1} - z^i\|_M^2 \leq \psi_0 - \psi_{k+1} \leq \psi_0 + t\varphi_k \leq \frac{\psi_0}{1-t} + t^{k+1}\varphi_0 \leq \frac{\psi_0}{1-t} + \varphi_0.$$

Thus, we know

$$\sum_{k=0}^{+\infty} \|z^{i+1} - z^i\|_M^2 < \infty \quad \Rightarrow \quad \lim_{k \rightarrow \infty} \|z^{k+1} - z^k\|_M^2 = 0. \quad (30)$$

From Assumption 4.1 (v), we further have

$$\lim_{k \rightarrow \infty} \|z^{k+1} - z^k\| = 0. \quad (31)$$

It can be easily seen from (25) that

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

where

$$\delta_k := (\theta_k t_k + (2 - \theta_k) t_k^2) \|z^k - z^{k-1}\|_M^2 \leq 2 \|z^k - z^{k-1}\|_M^2.$$

From (30) and $1 < \theta_k \leq 2$, we know $\sum_{k=0}^{+\infty} \delta_k < +\infty$. This fact, together with Lemma 4.1, indicates that the sequence $\{\|z^k - \hat{z}^*\|_M^2\}$ converges. Thus, the sequence $\{z^k\}$ is bound in norm, so does $\{z^k\}$ due to (8) and (31).

From (11) and (20), it is not difficult to check that

$$\gamma_k \geq \frac{2\rho}{\theta_k \|J\|^2 \cdot \|M^{-1}\|} > 0 \quad \Rightarrow \quad \gamma_{\min} = \liminf_{k \rightarrow \infty} \gamma_k > 0.$$

Thus, it follows from (23) that

$$\begin{aligned} 2\theta_k^{-1} \gamma_{\min} \|x^k - y^k\|_{D_0}^2 &\leq 2\theta_k^{-1} \gamma_k \|x^k - y^k\|_{D_0}^2 \\ &= \|\gamma_k M^{-1} J(x^k - y^k)\|_M^2 \\ &= \|z^{k+1} - z^k\|_M^2 \\ &= \|z^{k+1} - z^k - t_k(z^k - z^{k-1})\|_M^2, \end{aligned}$$

which, together with (31), implies

$$\lim_{k \rightarrow \infty} \|x^k - y^k\| = 0. \tag{32}$$

On the other hand, it follows from (4) that

$$T(x^k, J(x^k - y^k) - v^k - a^k, v^k) = \begin{pmatrix} A(x^k) + J(x^k - y^k) - v^k - a^k + v^k \\ -x^k + B^{-1}(J(x^k - y^k) - v^k - a^k) \\ -x^k + C^{-1}(v^k) \end{pmatrix},$$

which, together with $a^k \in A(x^k)$, $v^k = C(x^k)$ and $B^{-1}(J(x^k - y^k) - v^k - a^k) \ni y^k$, implies

$$T(x^k, J(x^k - y^k) - v^k - a^k, v^k) \ni \begin{pmatrix} Jx^k - Jy^k \\ -x^k + y^k \\ 0 \end{pmatrix}. \tag{33}$$

Next, let us prove that $\{x^k\}$, $\{a^k\}$ and $\{v^k\}$ are bounded in norm. In fact, in terms of (17), (18), self-adjoint property and strong monotonicity of J , we have

$$\begin{aligned} \|J\hat{z}^k - J\hat{z}^*\|_{J^{-1}}^2 &= \|J(x^k - x^*) + a^k - a^*\|_{J^{-1}}^2 \\ &= \|x^k - x^*\|_J^2 + 2\langle x^k - x^*, a^k - a^* \rangle + \|a^k - a^*\|_{J^{-1}}^2 \\ &\geq \|x^k - x^*\|_J^2. \end{aligned}$$

A similar discussion yields

$$\|J\hat{z}^k - J\hat{z}^*\|_{J^{-1}}^2 \geq \|a^k - a^*\|_{J^{-1}}^2.$$

These two inequalities tell us that, since $\{\hat{z}^k\}$ is bounded in norm, so do $\{x^k\}$ and $\{a^k\}$. As to $\{v^k\}$, it is certainly bounded in norm due to $v^k = C(x^k)$ and C' Lipschitz continuity. So, $\{(x^k, a^k, v^k)\}$ is bounded in norm and thus has at least one weak cluster point, say $\{(x^\infty, a^\infty, v^\infty)\}$. Consequently, there exists some subsequence of $\{(x^k, a^k, v^k)\}$ such that

$$(x^{k_j}, a^{k_j}, v^{k_j}) \rightharpoonup (x^\infty, a^\infty, v^\infty), \quad \text{as } k_j \rightarrow +\infty,$$

where the notation " \rightharpoonup " stands for weak convergence.

Now let us have a look at (33) once again. For the terms on the left-hand side, it is easy to check that

$$x^k \rightharpoonup x^\infty, \quad J(x^k - y^k) - v^k - a^k \rightharpoonup -v^\infty - a^\infty, \quad v^k \rightharpoonup v^\infty.$$

For the terms on the right-hand side, it follows from (32) and boundedness and linearity of J that $Jx^k - Jy^k$ tends to zero in norm. So, it follows from Lemma 2.2 that $(x^\infty, -v^\infty - a^\infty, v^\infty)$ must be zero of T . Thus, in view of (4), we can get

$$\begin{aligned} 0 \in v^\infty + A(x^\infty) - v^\infty - a^\infty &\Rightarrow a^\infty \in A(x^\infty), \\ 0 \in -x^\infty + B^{-1}(-v^\infty - a^\infty) &\Rightarrow -v^\infty - a^\infty \in B(x^\infty), \\ 0 = -x^\infty + C^{-1}(v^\infty) &\Rightarrow v^\infty = C(x^\infty). \end{aligned}$$

These relations on the right-hand side indicate that

$$0 \in C(x^\infty) + A(x^\infty) + B(x^\infty).$$

As to a proof of the uniqueness of weak cluster point, it is standard Dong et al. (2010) and thus is omitted.

(ii) By Lemma 4.3 and the well-known Cauchy-Schwarz inequality, we have

$$\phi_A(\|x^k - x^*\|) + \phi_B(\|y^k - x^*\|) \leq \langle Jz^k - Jz^*, x^k - y^k \rangle \leq \|Jz^k - Jz^*\| \|x^k - y^k\|.$$

So, the desired results follow from (32), boundedness of $\{z^k\}$ and boundedness and linearity of J .

Notice that our proof techniques of weak convergence are more self-contained and less convoluted via introducing the characteristic operator, and the idea can be found in Eckstein (2017) and in an early version, accepted by Optimization Online in December 2018, of a very recent article Dong (2021).

At the end of this section, we would like to point out that we are able to follow [1] to replace (6) with

$$0 \leq t_k \leq t < 1, \quad \sum_{k=0}^{+\infty} t_k \|z^k - z^{k-1}\|^2 < +\infty. \quad (34)$$

If $\alpha \in (0, 4c)$ and γ_k is given by (14), then the resulting algorithm remains weakly convergent Zhu (2020). Yet, one may argue that the assumptions shall be made on problem itself instead of iterates. Therefore, we will not discuss them any more.

5 Comparisons with existing results

In this section, we mainly compare Algorithm 3.1 with existing results Yu (2019).

The involved method Yu (2019) there is an inertial version of Algorithm 3.1 of Dong et al. (2019) there for solving the monotone inclusion of three operators (1), in which $A := L$ is further bounded and linear, and it can be stated as follows.

By comparing the description of Algorithm 5.1 with that of Algorithm 3.1, we find out that they allow for the same assumptions on inertial factors (6).

6 An application

In this section, we are mainly concerned with the following problem of finding an x in \mathcal{H} such that

$$0 \in C(x) + A(x) + Q^*B(Qx - q), \quad (40)$$

where the operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$ and the operator $B : \mathcal{G} \rightrightarrows \mathcal{G}$ are maximal monotone, and $Q : \mathcal{H} \rightarrow \mathcal{G}$ is nonzero bounded linear with its adjoint Q^* , and $q \in \mathcal{G}$.

Algorithm 5.1

Step 0. Choose $J, S \in \mathcal{BL}(\mathcal{H})$. Choose $z^0, z^{-1} \in \mathcal{H}$. Choose $\theta_0 = \theta_{-1} = 2/1.9$, $t_0 = t_{-1} = 0$. Compute c and

$$D := J^+ - L^+ - \frac{1}{4c}I. \quad (35)$$

Set $k := 0$.

Step 1. Compute

$$\hat{x}^k = x^k + t_k(x^k - x^{k-1}), \quad (36)$$

$$(J + B)(y^k) \ni (J - F)(x^k). \quad (37)$$

If some stopping criterion is met, then stop. Otherwise, go to Step 2.

Step 2 Compute

$$\gamma_k = 2\theta_k^{-1} \|\hat{x}^k - y^k\|_D^2 / \|(J - L)(\hat{x}^k - y^k)\|_{S^{-1}}^2, \quad (38)$$

$$x^{k+1} = \hat{x}^k - \gamma_k S^{-1}(J - L)(\hat{x}^k - y^k). \quad (39)$$

Choose $\theta_{k+1} \in (1, 2]$ and t_{k+1} by (6). Set $k := k + 1$, and go to Step 1.

Let

$$T(x, u) := \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} + \begin{pmatrix} A & 0 \\ 0 & B^{-1} \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} + \begin{pmatrix} 0 & Q^* \\ -Q & 0 \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} + \begin{pmatrix} 0 \\ q \end{pmatrix}. \quad (41)$$

Then (40) can be reformulated into

$$0 \in T(x, u). \quad (42)$$

Below we demonstrate how to apply Algorithm 3.1 to solving the monotone inclusion (41)-(42) above.

If the three operators in (41) correspond to C, A, B in (1), respectively, then Algorithm 3.1 becomes

Notice that, in the light of the following Moreau identity

$$(\beta I + B^{-1})^{-1} \equiv \frac{1}{\beta}I - \frac{1}{\beta}(I + \beta B)^{-1},$$

the process of (43) can be divided into

$$\hat{u}^k := \beta \hat{z}_2^k, \quad \tilde{u}^k := (I + \beta B)^{-1}(\hat{u}^k), \quad u^k := (\hat{u}^k - \tilde{u}^k)/\beta. \quad (48)$$

As to (44), we may solve this system of linear equations via

$$\begin{aligned} & (\beta\alpha^{-1}I + Q^*Q)(y^k) \\ & = 2\beta\alpha^{-1}x^k - \beta C(x^k) - 2\beta Q^*u^k - \beta\alpha^{-1}\hat{z}_1^k + \beta Q^*\hat{z}_2^k + Q^*q, \end{aligned} \quad (49)$$

$$v^k = 2u^k - \hat{z}_2^k + \beta^{-1}Qy^k - \beta^{-1}q. \quad (50)$$

Here we would like to make some remarks. When C vanishes, so do the C, c in Algorithm 6.1. The resulting algorithm seems new.

Algorithm 6.1

Step 0. Choose $M \in \mathcal{BL}(\mathcal{H})$. Choose $z_1^0, z_1^{-1} \in \mathcal{H}$, $z_2^0, z_2^{-1} \in \mathcal{G}$. Choose $\theta_0 = \theta_{-1}$, $t_0 = t_{-1} = 0$. Compute c . Choose $\alpha \in (0, 4c)$. Choose $\beta > 1/4c$. Set $k := 0$.

Step 1. Let $z^k = (z_1^k, z_2^k)$.

$$\begin{aligned} \hat{z}^k &= z^k + t_k(z^k - z^{k-1}) \\ (I + \alpha A)(x^k) &\ni z_1^k \\ (\beta I + B^{-1})(u^k) &\ni \beta z_2^k. \end{aligned} \quad (43)$$

Find y^k, v^k satisfying

$$\begin{cases} \alpha^{-1}y^k + Q^*v^k = 2\alpha^{-1}x^k - \alpha^{-1}\hat{z}_1^k - C(x^k), \\ \beta v^k - Qy^k + q = 2\beta u^k - \beta \hat{z}_2^k. \end{cases} \quad (44)$$

If some stopping criterion is met, then stop. Otherwise, go to Step 2.

Step 2 Compute

$$w_1^k := \alpha^{-1}(x^k - y^k), \quad w_2^k := \beta(u^k - v^k), \quad w^k := \begin{pmatrix} w_1^k \\ w_2^k \end{pmatrix} \quad (45)$$

$$\gamma_k = 2\theta_k^{-1} \frac{(\alpha^{-1} - \frac{1}{4c})\|x^k - y^k\|^2 + (\beta - \frac{1}{4c})\|u^k - v^k\|^2}{\|w^k\|_{M^{-1}}^2} \quad (46)$$

$$z^{k+1} = \hat{z}^k - \gamma_k M^{-1}w^k. \quad (47)$$

Choose $\theta_{k+1} \in (1, 2]$ and t_{k+1} by (6). Set $k := k + 1$, and go to Step 1.

7 Rudimentary experiments

In this section, using three different classes of test problems, we compared the empirical performance of Algorithm 3.1 and its variations, all implemented in MATLAB.

For Algorithm 3.1, we chose the associated parameters as follows.

DR3: $\theta_k \equiv 2/1.9$, $\alpha = 1.5/6$ ($t_k \equiv 0$).

iDR3(1): $\theta_k \equiv 2/1$, $\alpha = 1.5/6$, $t_k \equiv 0.333$.

iDR3(1.9): $\theta_k \equiv 2/1.9$, $\alpha = 1.5/6$, $t_k \equiv 0.045$.

iDR3plus: $\theta_k \equiv 2/1.9$, $\alpha = 1.5/6$. For t_k , we chose $t_0 = 0.333$ and

$$t_{k+1} = \begin{cases} \max\{t_k, 0.045\}, & \text{if } \|z^{k+1} - z^k\|/\|z^k - z^{k-1}\| \leq 0.9, \\ \max\{t_k/(1 + k^\tau), 0.045\}, & \text{otherwise,} \end{cases} \quad (51)$$

where $\tau \in \{0.5, 1, 1.5\}$.

Our first test problem is to solve the following linear monotone complementarity problem Dong et al. (2019)

$$0 \in F(x) + \partial\delta_\Omega(x), \quad \text{with } F := C + A,$$

where δ_Ω is the indicator function of the first orthant $\Omega = \{x : x_i \geq 0, i = 1, \dots, n\}$ and $C(x) := (1-s)Ux + q$, where $s \in [0, 1)$, and the associated matrices are $n \times n$ block tridiagonal.

$$U = \text{tridiag}(-I, \bar{Q}, -I), \quad A = sU + \frac{h\bar{c}}{2} \text{tridiag}(-I, O, I)$$

Table 2: Numerical results on the first problem: k (left) and the elapsed time

n, ϵ	DR3	iDR3(1)	iDR3(1.9)	iDR3plus
$50^2, 10^{-9}$	147 : 1.206	181 : 1.472	139 : 1.145	105 : 0.951
$100^2, 10^{-8}$	534 : 18.60	675 : 23.06	509 : 17.51	342 : 11.94
$150^2, 10^{-7}$	1120 : 96.03	1418 : 120.2	1069 : 90.68	735 : 67.04
$200^2, 10^{-6}$	1857 : 289.5	2352 : 339.1	1773 : 257.0	1228 : 178.2

and $n = m^2$, where $h = 1/(m + 1), \bar{c} = 100, I$ is an $m \times m$ identity matrix and \bar{Q} is an $m \times m$ matrix of the form $4I + Q + Q^T$ with $Q_{ij} = -1$ whenever $j = i + 1, i = 1, \dots, m - 1$, otherwise $Q_{ij} = 0$. To know the solution of this problem in advance. We set $q = -((1 - s)U + A)e_1$, where e_1 is the first column of the corresponding identity matrix. Thus, $x^* = e_1$ is the unique solution of the complementarity problem. Obviously, the problem corresponds to the monotone inclusion (1) with

$$F(x) = C(x) + A(x), \quad B(x) = \partial\delta_\Omega(x).$$

Notice that the largest eigenvalue of U is less than 6. Thus, in practical implementations, we adopted

$$c = 1/(6(1 - s)), \quad J := \alpha^{-1}I, \quad \theta_k \in (1, 2], \quad M := \alpha^{-1}I.$$

Set $s = 0.5$. The starting points were chosen as

$$x^0 = ones(n, 1), \quad z^{-1} = ones(n, 1), \quad z^0 = ones(n, 1).$$

We made use of the following stopping criterion

$$\|x^k - x^*\| \leq \epsilon \|x^0 - x^*\|.$$

We chose the parameter $\tau = 0.5$ in (51).

The corresponding numerical results were reported in Table 2.

Our second test problem is taken from Dong (2021b) and it is to find an $x \in \mathcal{R}^m$ such that

$$0 \in Dx - d + Q^*\partial\delta_C(Qx - q),$$

where

$$D = \text{tridiag}(a, b, -1), \quad a := -1 - h, \quad b := 4 + 2h, \quad h := 1/(m + 1),$$

and $q = (0, \dots, 0, -1/m)^T \in \mathcal{R}^{m+1}$ and

$$Q = \begin{pmatrix} I \\ r^T \end{pmatrix}, \quad r^T := \left(-\frac{1}{m}, \dots, -\frac{1}{m}\right),$$

and $C \subseteq \mathcal{R}^{m+1}$ is the first orthant. To ensure that $e_1 = (1, 0, \dots, 0)^T$ solves it, we set $d = De_1$ in our practical implementations. Furthermore, we chose

$$Cx = 0.5(D + D^T)x - d, \quad Ax = 0.5(D - D^T)x$$

Table 3: Numerical results on the third problem: k (left) and the elapsed time

m	Algo6.1(non)	Algo6.1(1)	Algo6.1(1.9)	Algo6.1plus
1000	147 : 1.206	181 : 1.472	139 : 1.145	105 : 0.951
5000	1857 : 289.5	2352 : 339.1	1773 : 257.0	1228 : 178.2

to match the problem (41)-(42).

For Algorithm 6.1, we set

$$z^0 = z^{-1} = 0, \quad \theta = 2/1.9, \quad c = 1/6, \quad \alpha = 0.25, \quad \beta = 4.$$

We adopted the stopping criterion $\|x_k - x^*\| \leq \epsilon \|x_0 - x^*\|$, where $\epsilon = 10^{-6}$.

Notice that, for our third test problem, Q satisfies $Q^T Q = I + rr^T$. Thus, the coefficient matrix in (49) becomes

$$\beta\alpha^{-1}I + Q^T Q = (\beta\alpha^{-1} + 1)I + rr^T.$$

Therefore, it follows from Sherman-Morrison formula that

$$((\beta\alpha^{-1} + 1)I + rr^T)^{-1} = \frac{1}{\beta\alpha^{-1} + 1} \left(I - \frac{rr^T}{\beta\alpha^{-1} + 1 + r^T r} \right).$$

In addition, (48) becomes

$$u^k = \hat{z}_2^k - \max\{0, \hat{z}_2^k\}.$$

The corresponding numerical results were reported in Table 3.

From Tables 2-3, we can see that each plus-version of both Algorithm 3.1 and Algorithm 6.1 is always the best, with less elapsed time and fewer number of iteration for the corresponding test problem.

8 Conclusions

In this article, we have considered monotone inclusions of three operators in real Hilbert spaces and have suggested an inertial splitting method: iDR3. Under standard assumptions, we have analyzed its weak and strong convergence properties. The newly-developed proof techniques are based on the characteristic operator and thus are more self-contained and less convoluted.

By writing this article, we try to break the ice from aspect of computational efficiency since we have designed a practically useful iDR3plus to significantly increase efficiency of iDR3. This is in a sharp contrast to *at most somewhat better but possibly worse* impression on inertial effects. We cherish the hope that our new findings can arouse the attention from optimization community so as to emerge more and deeper research achievements.

References

1. Alvarez, F. and Attouch, H. (2001), "An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping", *Set-Valued Analysis* **9**, 3-11.
2. Alves, M. and Marcavillaca, R. (2020), "On inexact relative-error hybrid proximal extragradient, forward-backward and Tseng's modified forward-backward methods with inertial effects", *Set-Valued and Variational Analysis* **28**, 301-325.
3. Davis, D. (2015), On the Design and Analysis of Operator-Splitting Schemes, Ph.D. Dissertation, Mathematics 0540UCLA.
4. Dixit, A., Sahu, D.R., Gautam, P. and Som, T. (2021), "Convergence analysis of two-step inertial Douglas-Rachford algorithm and application", *Journal of Applied Mathematics and Computing* DOI: 10.1007/s12190-021-01554-5
5. Dong, Q.L., He, S. and Liu, L. (2021), "A general inertial projected gradient method for variational inequality problems", *Computational and Applied Mathematics* DOI: 10.1007/s40314-021-01540-4
6. Dong, Y.D. (2021a), "New inertial factors of the Krasnosel'skiĭ-Mann iteration", *Set-Valued and Variational Analysis* **29**, 145-161.
7. Dong, Y.D. (2021b), "Weak convergence of an extended splitting method for monotone inclusions", *Journal of Global Optimization* **79**, 257-277.
8. Dong, Y.D. and Fischer, A. (2010), "A family of operator splitting methods revisited", *Nonlinear Analysis* **72**, 4307-4315.
9. Dong, Y.D., Fang, B.B. and Wang, X.Y. (2018), "New inertial factors of a splitting method for monotone inclusions", preprint. Available at <http://www.optimization-online.org/DBFILE/2018/04/6576.pdf>
10. Dong, Y.D. and Yu, X.H. (2019), "A new splitting method for monotone inclusions of three operators", *Calcolo* **56**, 3.
11. Eckstein, J. (2017), "A simplified form of block-iterative operator splitting and an asynchronous algorithm resembling the multi-block alternating direction method of multipliers", *Journal of Optimization Theory and Applications* **173**(1), 155-182.
12. Eckstein, J. and Bertsekas, D.P. (1992), "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators", *Mathematical Programming* **55**(3), 293-318.
13. Iyiola, O.S., Enyi, C.D. and Shehu, Y. (2021), "Reflected three-operator splitting method for monotone inclusion problem", *Optimization Methods and Software*, DOI: 10.1080/10556788.2021.1924715
14. Hieu, D.V., Vy, L.V. and Quy, P.K. (2020), "Three-operator splitting algorithm for a class of variational inclusion problems", *Bulletin of the Iran Mathematical Society* **46**(4): 1055-1071.

15. Latafat, P. and Patrinos, P. (2017), "Asymmetric forward-backward-adjoint splitting for solving monotone inclusions involving three operators", *Computational Optimization with Applications* **68**, 57-93.
16. Lions, P.L. and Mercier, B. (1979), "Splitting algorithms for the sum of two nonlinear operators", *SIAM Journal on Numerical Analysis* **16**, 964-979.
17. Martinet, B. (1970), "Regularisation d'inéquations variationnelles par approximations successives", *Rev Fr d'Informatique Recherche Opér* **4**, 154-158.
18. Minty, G.J. (1962), "Monotone (nonlinear) operators in Hilbert space", *Duke Mathematics Journal* **29**, 341-346.
19. Rockafellar, R.T. (1970), "On the maximality of sums of nonlinear monotone operators", *Transactions of the American Mathematical Society*, **149**, 75-88.
20. Rockafellar, R.T. (1976), "Monotone operators and the proximal point algorithm", *SIAM Journal on Control and Optimization* **14**, 877-898.
21. Thong, D.V. and Hieu, D.V. (2018), "Inertial extragradient algorithms for strongly pseudomonotone variational inequalities", *Journal of Computational and Applied Mathematics* **341**, 80-98.
22. Yu, X.H. (2019), Research on Splitting Methods with Inertial Term for Monotone Inclusion Problems (in Chinese). Thesis, Zhengzhou University.
23. Zhu, X. (2020), Inertial Splitting Methods for Monotone Inclusions of Three Operators (in Chinese). Thesis, Zhengzhou University.