Received 23 April 2020; | Revised: 00 Month 0000 | Accepted: 00 Month 0000

DOI: xxx/xxxx

RESEARCH ARTICLE

Adaptive hybrid steepest descent algorithms involving an inertial extrapolation term for split monotone variational inclusion problems

Zheng Zhou | Bing Tan | Songxiao Li*

Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China

Correspondence
Songxiao Li, Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China.
Email: jyulsx@163.com

MSC Classification:47H10; 47J25; 65K10; 65Y10; 90C25

In this paper, we discuss the split monotone variational inclusion problem and propose two new inertial algorithms in infinite-dimensional Hilbert spaces. As well as, the iterative sequence by the proposed algorithms converges strongly to the solution of a certain variational inequality with the help of the hybrid steepest descent method. Furthermore, an adaptive step size criterion is considered in suggested algorithms to avoid the difficulty of calculating the operator norm. Finally, some numerical experiments show that our algorithms are realistic and summarize the known results.

KEYWORDS:
adaptive step size, hybrid steepest descent method, inertial technique, signal recovery, strong convergence

1 | INTRODUCTION

As one of the important generalized forms of convex feasibility problems, the split feasibility problem (shortly, SFP) was introduced by Censor and Elfving [1] in 1994 and used to model inverse problems in phase retrievals and medical image reconstruction. In fact, the SFP is also used in signal recovery, computer tomography, radiation therapy treatment planning, etc.; for further detail, see [2,3,4] and the references therein. At the same time, many good algorithms and excellent convergence results have been produced in the study of the approximate solution of the SFP, among which the CQ method given by Byrne [5] is the most familiar iterative method. Further research, the feasible sets in the SFP are often considered as other forms, such as the fixed point set of nonlinear mappings, the solution set of variational inequality problems, the solution set of equilibrium problems, the solution set of inclusion problem and so on. Consequently, Moudafi [6] introduced the split monotone variational inclusion problem (shortly, SMVIP) that is formulated as follows:

$$\text{Find } x \in \mathcal{H}_1, \text{ such that } 0 \in f_1(x) + B_1(x) \text{ and } 0 \in f_2(Ax) + B_2(Ax),$$

where $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces, $f_1 : \mathcal{H}_1 \to \mathcal{H}_1$ and $f_2 : \mathcal{H}_2 \to \mathcal{H}_2$ are single-valued mappings, $B_1 : \mathcal{H}_1 \to 2^{\mathcal{H}_1}$ and $B_2 : \mathcal{H}_2 \to 2^{\mathcal{H}_2}$ are set-valued maximal monotone mappings, $A : \mathcal{H}_1 \to \mathcal{H}_2$ is a bounded linear operator. By means of Byrne's CQ method, Moudafi suggested the following iterative algorithm: for arbitrary $x_1 \in \mathcal{H}_1$, $\gamma > 0$, $x_{n+1} = J_\gamma^{B_1}(I - \gamma f_1)(x_n - \lambda A^*(I - J_\gamma^{B_2}(I - \gamma f_2))Ax_n)$, $n \geq 1$, where $J_\gamma^{B_i}$ is resolvent operator of B_i and is defined by $J_\gamma^{B_i} = (I + \gamma B_i)^{-1}$ for $\gamma > 0$ and $i = 1, 2$, A^* is the adjoint operator of A and I is an identity mapping. Meanwhile, the generated sequence $\{x_n\}$ converges weakly to a solution of the SMVIP under mild assumptions. It should be emphasized that the SMVIP also covers many split problem, such as the split variational inclusion problem, the split variational inequality problem, the split minimization problem and the split feasibility problem; for more detail, see Sect. 5.

Specifically, in this case that $f_1 \equiv 0 \equiv f_2$, $B_1 = N_{C_1}$ and $B_2 = N_{Q_1}$ (N_{C_1} and N_{Q_1} are normal cone of C_1 and Q_1, respectively), the SMVIP is viewed as the SFP, that is,

$$\text{find } x \in C_1 \text{ such that } Ax \in Q_1, \tag{1}$$

where C_1 and Q_1 are nonempty closed convex subsets of $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. From the perspective of optimization problems, the SFP can also be described as a constrained optimization problem as follows:

$$\min_{x \in C_1} \frac{1}{2} \|Ax - P_{Q_1}(Ax)\|^2.$$

For convenience, take the objective function $F(x) = \frac{1}{2}\|Ax - P_{Q_1}(Ax)\|^2$. Obviously, F is continuously differentiable, its gradient is given by $\nabla F(x) = A^* \left(I - P_{Q_1}\right) A(x)$ and ∇F is $\|A\|^2$-Lipschitz continuous. The gradient projection method is used to deal with this problem and the following iterative scheme is generated:

$$x_{n+1} = P_{C_1}\left(I - \lambda A^* \left(I - P_{Q_1}\right) A\right) x_n, \ n \geq 1, \tag{2}$$

where λ is a constant in $(0, \frac{2}{\|A\|^2})$. Generally, suppose that $\mathcal{H}_1 := R^N, \mathcal{H}_2 := R^M$, A is a real M by N matrix and $A^* = A^T$ (A^T is the transposition of A). So ∇F is $\rho(A^T A)$-Lipschitz continuous ($\rho(A^T A)$ is the spectral radius of the matrix $A^T A$) and Algorithm 2 is exactly the CQ algorithm proposed by Byrne.[5] Thanks to the iterative form generated by the gradient projection method, López et al.[7] constructed an adaptive step size sequence $\{\lambda_n\}$ to replace λ in Algorithm 2, that is, $\lambda_n := \frac{\sigma_n F(x_n)}{\|\nabla F(x_n)\|^2}$ with $\sigma_n \in (0, 4)$. This variable step size increases the practicability of the algorithm in applications, especially when it is not easy or possible to calculate the norm of A and the spectral radius of $A^T A$. Recently, Yao et al.[8] introduced a weakly convergent algorithm for solving the SMVIP by using the idea of this step size. But, a common flaw in[5,6,8] is that they can only guarantee the weak convergence of the algorithm. Naturally, an interesting question is how to construct a strongly convergent algorithm with an adaptive step size criterion that approximates the solution of the SMVIP.

In fact, the strong convergence of iterative sequences is better than the weak convergence for some applications in an infinite-dimensional Hilbert space, for example, CT reconstruction and machine learning. One of the most familiar strong convergence algorithms is the viscosity algorithm introduced by Moudafi[9] in 2000 which is implemented by using a contraction mapping embedded in the Krasnosel'skii-Mann iterative algorithm. Further, Marino and Xu[10] proposed a general viscosity algorithm by combining a contraction mapping and a strongly positive bounded linear operator. Meanwhile, the generated sequence converges strongly to the unique solution of a variational inequality, which is also the solution of a convex minimization problem. In addition, Yamada[11] introduced the hybrid steepest descent method for solving a variational inequality problem over the fixed point set of a nonexpansive mapping. The resulting sequence also converges in norm to the unique solution of a variational inequality. Inspiration from the above work, Tian[12] suggested a strongly convergent algorithm by combining a contraction mapping and a Lipschitz continuous and strongly monotone mapping.

On the other hand, based on the idea of the implicit discretization of a differential system of the second-order in time, Alvarez and Attouch in[13] gave an implicit weakly convergent algorithm to approximate a solution of the variational inclusion problem:

$$0 \in x_{n+1} - x_n - \alpha_n(x_n - x_{n-1}) + \gamma_n B(x_{n+1}),$$

which can also be expressed as the following explicit iterative form:

$$x_{n+1} = J^B_{\gamma_n}\left(x_n + \alpha_n(x_n - x_{n-1})\right), \tag{IPPA}$$

where $B : \mathcal{H} \to 2^{\mathcal{H}}$ is a maximal monotone mapping, $J^B_{\gamma_n}$ is the resolvent operator of B, for $\gamma_n \geq \gamma > 0$. Such an algorithm is called the inertial proximal point algorithm (for short, IPPA), and $\alpha_n(x_n - x_{n-1})$ is referred to as the inertial extrapolation term of IPPA. By means of this design, the iterative sequence $\{x_n\}$ can quickly converge to a solution of the variational inclusion problem. At the same time, the inertial technique plays an important role in accelerating the convergence speed of the algorithm for solving other mathematical problems, such as the split monotone variational inclusion problem,[8] the variational inequality problem,[14,15] the split common fixed point problem[16,17] and references therein.

Based on the above results, two novel strongly convergent inertial algorithms are proposed for solving the split monotone variational inclusion problem in infinite-dimensional Hilbert spaces. To be more precise, our contribution in this paper is twofold. The first one is that these two algorithms combine the hybrid steepest descent method and the inertial method. Thus, the strong convergence and fast convergence behavior of the suggested algorithms are guaranteed and implemented. The second one is that a variable step size is chosen in our algorithms, which is generated adaptively by calculating each iteration. Furthermore, this step size criterion effectively overcomes the case that the operator norm is not easy to calculate in the iteration process.

The rest of the article is organized as follows. Sect. 2 introduces some basic definitions and useful lemmas to explain the subsequent proofs. In Sect. 3, we give two new algorithms, namely Algorithms 1 and 2. Then in Sect. 4, the main results are presented and the corresponding proofs are given. In addition, some theoretical applications to other split problems are presented in Sect. 5. Finally, in Sect. 6, some numerical experiments including signal recovery problems are given to characterize the validity of Algorithms 1 and 2 and to compare known algorithms.

2 | PRELIMINARIES

This section will provide some relevant definitions and useful lemmas for the proofs in Sects. 3 and 4. First of all, assume that $\mathcal{H}$ is a Hilbert space embedded with the inner product $\langle\cdot,\cdot\rangle$ and the induced norm $\|\cdot\|$ and C is a nonempty closed convex subset of $\mathcal{H}$. The notations $x_n \to x$ and $x_n \rightharpoonup x$ represent strong convergence and weak convergence of the sequence $\{x_n\}$ to x, respectively. The symbol $Fix(T)$ denotes all fixed points of a mapping T. Let $B : \mathcal{H} \to 2^{\mathcal{H}}$ be a set-valued mapping with domain $D(B) = \{x \in \mathcal{H} : B(x) \neq \emptyset\}$ and graph $G(B) = \{(x, w) \in \mathcal{H}\times\mathcal{H} : x \in D(B), w \in B(x)\}$. Recall that a mapping $B : \mathcal{H} \to 2^{\mathcal{H}}$ is monotone if and only if $\langle x - y, w - v\rangle \geq 0$ for any $w \in B(x)$ and $v \in B(y)$. Further, a monotone mapping $B : \mathcal{H} \to 2^{\mathcal{H}}$ is maximal, that is, $G(B)$ is not properly contained in the graph of any other monotone mapping. In this case, B is a maximal monotone mapping if and only if for any $(x, w) \in G(B)$ and $(y, v) \in \mathcal{H}\times\mathcal{H}$, $\langle x-y, w-v\rangle \geq 0$ implies $v \in B(y)$.

Definition 1. The metric projection of $\mathcal{H}$ onto C is denoted by P_C, that is, $P_C x = \operatorname{argmin}_{y\in C} \|x - y\|$, $\forall x \in \mathcal{H}$. Meanwhile, the following conclusions are also true.

$$\langle P_C x - x, P_C x - y\rangle \leq 0, \forall y \in C \Leftrightarrow \|y - P_C x\|^2 + \|x - P_C x\|^2 \leq \|x - y\|^2. \tag{3}$$

For more detail, see Goebel and Reich[18] and Kopecká and Reich[19]. In addition, for any $x, y \in \mathcal{H}$ and $\lambda \in \mathbb{R}$, the following properties are readily available.

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \leq \|x\|^2 + 2\langle y, x + y\rangle, \tag{4}$$

$$\|\lambda x + (1 - \lambda)y\|^2 = \lambda\|x\|^2 + (1 - \lambda)\|y\|^2 - \lambda(1 - \lambda)\|x - y\|^2. \tag{5}$$

Definition 2. For any $x, y \in \mathcal{H}$, a mapping $T : \mathcal{H} \to \mathcal{H}$ is said to be

(1) contraction, if there exists $\xi \in [0, 1)$ such that
$$\|Tx - Ty\| \leq \xi\|x - y\|.$$

(2) nonexpansive, if
$$\|Tx - Ty\| \leq \|x - y\|.$$

(3) L-Lipschitz continuous with $L > 0$, if
$$\|Tx - Ty\| \leq L\|x - y\|.$$

(4) firmly nonexpansive, if
$$\|Tx - Ty\|^2 \leq \langle Tx - Ty, x - y\rangle.$$

(5) κ-averaged with $\kappa \in (0, 1)$, if there exists an identity mapping $I : \mathcal{H} \to \mathcal{H}$ and a nonexpansive mapping $S : \mathcal{H} \to \mathcal{H}$ such that
$$T = (1 - \kappa)I + \kappa S.$$

(6) η-strongly monotone, if there exists $\eta > 0$ such that
$$\eta\|x - y\|^2 \leq \langle Tx - Ty, x - y\rangle.$$

(7) ϑ-inverse strongly monotone, if there exists $\vartheta > 0$ such that
$$\vartheta\|Tx - Ty\|^2 \leq \langle Tx - Ty, x - y\rangle.$$

Remark 1. (i) The firmly nonexpansive mapping is 1-inverse strongly monotone mapping. (ii) The firmly nonexpansive mapping is 1-Lipschitz continuous mapping, i.e., nonexpansive mapping. (iii) The mapping T is firmly nonexpansive if and only if T is

$\frac{1}{2}$-averaged, i.e., $T = \frac{1}{2}(I + S)$. (iv) The ϑ-inverse strongly monotone mapping is $\frac{1}{\vartheta}$-Lipschitz continuous mapping. (v) If T is ϑ-inverse strongly monotone, then αT is $\frac{\vartheta}{\alpha}$-inverse strongly monotone for $\alpha > 0$.

Lemma 1 (Crombez[20,21])**.** Let $T : \mathcal{H} \to \mathcal{H}$ be a nonexpansive mapping. For arbitrary $x, y \in \mathcal{H}$,

$$\langle (x - Tx) - (y - Ty), Ty - Tx\rangle \leq \frac{1}{2}\|(x - Tx) - (y - Ty)\|^2$$

and consequently if $p \in Fix(T)$ then

$$\langle x - Tx, p - Tx\rangle \leq \frac{1}{2}\|x - Tx\|^2.$$

Lemma 2 (Zhou and Qin[22])**.** Let C be a nonempty closed convex subset of $\mathcal{H}$ and $T : C \to C$ be a nonexpansive mapping with $Fix(T) \neq \emptyset$. $I - T$ is demiclosed at zero, that is, for any sequence $\{x_n\} \subset C$, satisfying $x_n \rightharpoonup x^*$ and $x_n - T(x_n) \to 0$, then $x^* \in Fix(T)$.

Lemma 3 (Moudafi[6] and Byrne[23])**.** *(I)* The composite of finitely many averaged mappings is averaged;
(II) If the mappings $\{T_i\}_{i=1}^N$ are averaged and have a nonempty common fixed point, then

$$\bigcap_{i=1}^{N} Fix(T_i) = Fix\left(T_1 \cdots T_N\right);$$

(III) T is averaged if and only if its complement $I - T$ is ϑ-inverse strongly monotone for some $\vartheta > \frac{1}{2}$;
(IV) If T is averaged and N is a nonexpansive, then $(1 - a)T + aN$ is averaged for some $a \in (0, 1)$.

Lemma 4 (Moudafi[6])**.** Let $f : \mathcal{H} \to \mathcal{H}$ be a mapping and $B : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximal monotone mapping. The following properties hold.

(1) $0 \in f(x^*) + B(x^*)$ if and only if $x^* = J_\gamma^B(I - \gamma f)x^*$, i.e., $x^* \in Fix(J_\gamma^B(I - \gamma f))$, for $\gamma > 0$;

(2) If $f : \mathcal{H} \to \mathcal{H}$ is ϑ-inverse strongly monotone, then $J_\gamma^B(I - \gamma f)$ is average for $\gamma \in (0, 2\vartheta)$.

Remark 2. From Lemma 4 (1), the solution set of the split monotone variational inclusion problem is characterized as that of the fixed point problem, i.e.,

$$\Omega = \{x^* \in \mathcal{H}_1 : x^* \in Fix(J_\gamma^{B_1}(I - \gamma f_1)) \text{ and } Ax^* \in Fix(J_\gamma^{B_2}(I - \gamma f_2))\} \text{ for } \gamma > 0.$$

This implies that the solution set Ω is closed and convex.

The following lemma is an improvement of Lemma 3.1 in Yamada[11] and Lemma 2.5 in Thong et al.,[24] and also plays an important role in the convergence analysis of our algorithms.

Lemma 5. Let $D : \mathcal{H} \to \mathcal{H}$ be a L-Lipschitz continuous and η-strongly monotone mapping with $L,\ \eta > 0$. For any $\rho \in (0, 1)$, define a mapping $K_\rho(x) = (I - \rho\mu D)x, \forall x \in \mathcal{H}$. If $0 < \mu < \min\{\frac{1}{2\eta}, \frac{2\eta}{L^2}\}$, the following inequality holds:

$$\|K_\rho(x) - K_\rho(y)\| \leq (1 - \rho(1 - \sqrt{1 - \mu(2\eta - \mu L^2)}))\|x - y\|,\ \forall x, y \in \mathcal{H},$$

then K_ρ is a contraction mapping.

Proof. For any $x, y \in \mathcal{H}$, set $D_\mu = \mu D - I$. We have

$$\begin{aligned}\|D_\mu x - D_\mu y\|^2 &= \mu^2\|Dx - Dy\|^2 - 2\mu\langle Dx - Dy, x - y\rangle + \|x - y\|^2 \\ &\leq \mu^2 L^2\|x - y\|^2 - 2\mu\eta\|x - y\|^2 + \|x - y\|^2 \\ &= (1 - 2\mu\eta + \mu^2 L^2)\|x - y\|^2.\end{aligned}$$

Since $0 < \mu < \min\{\frac{1}{2\eta}, \frac{2\eta}{L^2}\}$, then $0 < 1 - 2\mu\eta + \mu^2 L^2 = 1 - \mu(2\eta - \mu L^2) < 1$. Further,

$$\begin{aligned}\|K_\rho x - K_\rho y\| &\leq \|(1 - \rho)(x - y)\| + \rho\|D_\mu x - D_\mu y\| \\ &\leq (1 - \rho)\|x - y\| + \rho\sqrt{1 - \mu(2\eta - \mu L^2)}\|x - y\| \\ &= (1 - \rho(1 - \sqrt{1 - \mu(2\eta - \mu L^2)}))\|x - y\|.\end{aligned}$$

By $\rho \in (0, 1)$ and $0 < 1 - \mu(2\eta - \mu L^2) < 1$, we get $0 < 1 - \rho(1 - \sqrt{1 - \mu(2\eta - \mu L^2)}) < 1$, this implies that the mapping K_ρ is contraction. $\square$

In particular, if $D : \mathcal{H} \to \mathcal{H}$ is a L-Lipschitz continuous and η-strongly monotone mapping with $L \geq \eta > 0$ in Lemma 5, then the following lemma can be obtained by the above proof process.

Lemma 6. Let $D : \mathcal{H} \to \mathcal{H}$ be a L-Lipschitz continuous and η-strongly monotone mapping with $L \geq \eta > 0$. For any $\rho \in (0,1)$, define a mapping $K_\rho(x) = (I - \rho\mu D)x, \forall x \in \mathcal{H}$. If $0 < \mu < \frac{2\eta}{L^2}$, the following inequality holds:

$$\|K_\rho(x) - K_\rho(y)\| \leq (1 - \rho(1 - \sqrt{1 - \mu(2\eta - \mu L^2)}))\|x - y\|, \ \forall x, y \in \mathcal{H},$$

then K_ρ is a contraction mapping.

Lemma 7 (He and Yang[25])**.** Suppose that $\{S_n\}$ and $\{c_n\}$ are sequences of nonnegative real numbers such that

$$S_{n+1} \leq (1 - a_n)S_n + a_n b_n \text{ and } S_{n+1} \leq S_n - c_n + d_n, n \geq 1,$$

where $\{a_n\}$, $\{b_n\}$ and $\{d_n\}$ are real sequences with $0 < a_n < 1$. If $\sum_{n=1}^{\infty} a_n = \infty$, $\lim\limits_{n\to\infty} d_n = 0$, and $\lim\limits_{k\to\infty} c_{n_k} = 0$ implies $\limsup\limits_{k\to\infty} b_{n_k} \leq 0$ ($\{n_k\}$ is any subsequence of $\{n\}$). The sequence $\{S_n\}$ converges to 0 as $n \to \infty$.

3 | PROPOSED ALGORITHMS

In this section, we state two adaptive algorithms with an inertial extrapolation term for finding approximate solutions of the split monotone variational inclusion problem. Suppose that the solution set Ω of the SMVIP is nonempty. To begin with, the relevant assumptions are set as follows:

(A1) $\mathcal{H}_1, \mathcal{H}_2$ are two Hilbert spaces and $A : \mathcal{H}_1 \to \mathcal{H}_2$ is a bounded linear operator with adjoint operator A^*;

(A2) $f_1 : \mathcal{H}_1 \to \mathcal{H}_1$ is a ϑ_1-inverse strongly monotone mapping and $f_2 : \mathcal{H}_2 \to \mathcal{H}_2$ is a ϑ_2-inverse strongly monotone mapping;

(A3) $B_1 : \mathcal{H}_1 \to 2^{\mathcal{H}_1}$ and $B_2 : \mathcal{H}_2 \to 2^{\mathcal{H}_2}$ are two set-valued maximal monotone mappings;

(A4) $h : \mathcal{H}_1 \to \mathcal{H}_1$ is a L_1-Lipschitz continuous mapping with $L_1 > 0$;

(A5) $D : \mathcal{H}_1 \to \mathcal{H}_1$ is a L_2-Lipschitz continuous and η-strongly monotone mapping with $L_2, \eta > 0$.

Meanwhile, setting $W_1 = J_\gamma^{B_1}(I - \gamma f_1)$ and $W_2 = J_\gamma^{B_2}(I - \gamma f_2)$. In order to ensure the convergence of the proposed algorithms, the following control conditions need to be satisfied:

(C1) $\{\beta_n\} \subset (0,1)$ such that $\lim_{n\to\infty} \beta_n = 0$ and $\sum_{n=1}^{\infty} \beta_n = \infty$;

(C2) $\{\alpha_n\} \subset [0,1)$, $\{\sigma_n\} \subset (0,1)$ and $\lim_{n\to\infty} \frac{\alpha_n}{\beta_n}\|x_n - x_{n-1}\| = 0$;

(C3) $0 < \gamma < 2\vartheta$ with $\vartheta = \min\{\vartheta_1, \vartheta_2\}$;

(C4) $0 \leq \tau L_1 < \theta = 1 - \sqrt{1 - \mu(2\eta - \mu L_2^2)}$ and $0 < \mu < \min\{\frac{1}{2\eta}, \frac{2\eta}{L_2^2}\}$.

Algorithm 1

Require: Take arbitrary starting points x_0, x_1 in $\mathcal{H}_1$. Choose sequences $\{\alpha_n\} \subset [0,1)$, $\{\sigma_n\}$ and $\{\beta_n\}$ in $(0,1)$ and $\gamma, \ \tau, \ \mu > 0$.

1: Set $n = 1$ and compute $u_n = x_n + \alpha_n(x_n - x_{n-1})$ and adaptive step size

$$\lambda_n = \begin{cases} \dfrac{\sigma_n\|(I - W_2)Au_n\|^2}{\|A^*(I - W_2)Au_n\|^2}, & Au_n \notin Fix(W_2), \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

2: Compute $y_n = W_1\left(u_n - \lambda_n A^*(I - W_2)Au_n\right)$.

3: If $y_n = u_n$, then stop. Otherwise, compute $x_{n+1} = \beta_n \tau h(y_n) + (I - \beta_n \mu D)y_n$.

4: Set $n := n + 1$ and return 1.

Algorithm 2

Require: Two arbitrary starting points x_0, x_1 in $\mathcal{H}_1$. Choose sequences $\{\alpha_n\} \subset [0,1)$, $\{\sigma_n\}$ and $\{\beta_n\}$ in $(0,1)$ and $\gamma,\ \tau,\ \mu > 0$.

1: Set $n = 1$ and compute $u_n = x_n + \alpha_n(x_n - x_{n-1})$, $z_n = W_1(u_n)$, and adaptive step size

$$\lambda_n = \begin{cases} \dfrac{\sigma_n \|(I - W_2)Az_n\|^2}{\|A^*(I - W_2)Az_n\|^2}, & Az_n \notin Fix(W_2), \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

2: Compute $y_n = z_n - \lambda_n A^*(I - W_2)Az_n$.

3: If $y_n = z_n = u_n$, then stop. Otherwise, compute $x_{n+1} = \beta_n \tau h(y_n) + (I - \beta_n \mu D)y_n$.

4: Set $n := n + 1$ and return 1.

Lemma 8. The adaptive step size sequence $\{\lambda_n\}$ defined by (6) and (7) is well-defined.

Proof. From Lemma 3 (III) and Lemma 4 (2), we have that W_2 is average and $I - W_2$ is $\bar{\vartheta}$-inverse strongly monotone for $\bar{\vartheta} > \frac{1}{2}$. Taking $x^* \in \Omega$, i.e., $x^* \in Fix(W_1)$, $Ax^* \in Fix(W_2)$. According to the definition of (6), we have

$$\begin{aligned} \|A^*(I - W_2)Au_n\|\|u_n - x^*\| &\geq \langle A^*(I - W_2)Au_n, u_n - x^*\rangle \\ &= \langle (I - W_2)Au_n, Au_n - Ax^*\rangle \\ &\geq \bar{\vartheta}\|(I - W_2)Au_n\|^2. \end{aligned}$$

So, when $Au_n \notin Fix(W_2)$, we get $\|A^*(I - W_2)Au_n\| > 0$. This means that the sequence $\{\lambda_n\}$ in (6) is well-defined. Similarly, $\{\lambda_n\}$ in (7) is also well-defined. □

Lemma 9. If $y_n = u_n$ in Algorithm 1, then u_n is a solution of SMVIP, i.e., $u_n \in \Omega$.

Proof. Since W_1 and W_2 are average, it is easy to get that W_1 and W_2 are nonexpansive. For any $x^* \in \Omega$, i.e., $x^* \in Fix(W_1)$ and $Ax^* \in Fix(W_2)$, it follows from Lemma 1 that

$$\begin{aligned} 2\lambda_n\langle (I - W_2)Au_n, Au_n - Ax^*\rangle &= 2\lambda_n\langle (I - W_2)Au_n, W_2Au_n - Ax^*\rangle + 2\lambda_n\|(I - W_2)Au_n\|^2 \\ &\geq -\lambda_n\|(I - W_2)Au_n\|^2 + 2\lambda_n\|(I - W_2)Au_n\|^2 \\ &= \lambda_n\|(I - W_2)Au_n\|^2. \end{aligned}$$

Further, we have

$$\begin{aligned} \|y_n - x^*\|^2 &\leq \|u_n - \lambda_n A^*(I - W_2)Au_n - x^*\|^2 \\ &= \|u_n - x^*\|^2 - 2\lambda_n\langle A^*(I - W_2)Au_n, u_n - x^*\rangle + \lambda_n^2\|A^*(I - W_2)Au_n\|^2 \\ &= \|u_n - x^*\|^2 - 2\lambda_n\langle (I - W_2)Au_n, Au_n - Ax^*\rangle + \lambda_n^2\|A^*(I - W_2)Au_n\|^2 \\ &\leq \|u_n - x^*\|^2 - \lambda_n(1 - \sigma_n)\|(I - W_2)Au_n\|^2. \end{aligned} \tag{8}$$

Hence,

$$\|y_n - W_1y_n\| \leq \|u_n - \lambda_n A^*(I - W_2)Au_n - y_n\| = \lambda_n\|A^*(I - W_2)Au_n\|. \tag{9}$$

By virtue of (8), (9) and $y_n = u_n$, $\lim_{n\to\infty}\|(I - W_1)y_n\| = \lim_{n\to\infty}\|(I - W_2)Au_n\| = 0$, which implies that u_n belongs to Ω. In particular, if $\lambda_n = 0$, the above results also hold. □

Lemma 10. If $y_n = z_n = u_n$ in Algorithm 2, then u_n is a solution of SMVIP, i.e., $u_n \in \Omega$.

Proof. Obviously, when $\lambda_n = 0$, $u_n \in \Omega$. On the other hand, using the proof of Lemma 8, we know that $I - W_2$ is $\bar{\vartheta}$-inverse strongly monotone for $\bar{\vartheta} > \frac{1}{2}$. In the same way, $I - W_1$ is $\bar{\upsilon}$-inverse strongly monotone for $\bar{\upsilon} > \frac{1}{2}$. For any $x^* \in \Omega$, using Lemma 1 and $y_n = z_n = u_n$ we get

$$\begin{aligned} 0 &= \langle u_n - z_n, u_n - x^*\rangle + \langle z_n - y_n, z_n - x^*\rangle \\ &= \langle u_n - W_1u_n, u_n - x^*\rangle + \lambda_n\langle A^*(I - W_2)Az_n, z_n - x^*\rangle \\ &= \langle u_n - W_1u_n, u_n - x^*\rangle + \lambda_n\langle (I - W_2)Az_n, Az_n - Ax^*\rangle \\ &\geq \bar{\upsilon}\|(I - W_1)u_n\|^2 + \lambda_n\bar{\vartheta}\|(I - W_2)Az_n\|^2. \end{aligned}$$

Then, $\lim_{n\to\infty}\|(I - W_1)u_n\| = \lim_{n\to\infty}\|(I - W_2)Az_n\| = 0$, which implies that $u_n \in \Omega$. □

4 | CONVERGENCE ANALYSIS

In what follows, the convergence analysis of Algorithms 1 and 2 are proved. Moreover, some nontrivial corollaries have also been proposed for solving the split monotone variational inclusion problem and extend the existing results.

Lemma 11. The iterative sequence $\{x_n\}$ generated by Algorithms 1 and 2 is bounded.

Proof. For any $x^* \in \Omega$, from (8) and Algorithm 1, we get

$$\|y_n - x^*\|^2 \leq \|u_n - x^*\|^2 - \lambda_n(1-\sigma_n)\|(I - W_2)Au_n\|^2.$$

This implies that $\|y_n - x^*\| \leq \|u_n - x^*\|$. Using Lemma 5 and Condition (C4) to get

$$\begin{aligned}\|x_{n+1} - x^*\| &\leq \beta_n\|\tau h(y_n) - \tau h(x^*) + \tau h(x^*) - \mu Dx^*\| + \|(I - \beta_n\mu D)y_n - (I - \beta_n\mu D)x^*)\| \\ &\leq \beta_n\tau L_1\|y_n - x^*\| + \beta_n\|\tau h(x^*) - \mu Dx^*\| + (1-\beta_n\theta)\|y_n - x^*\| \\ &\leq [1 - \beta_n(\theta - \tau L_1)]\|x_n - x^*\| + \beta_n\|\tau h(x^*) - \mu Dx^*\| + \alpha_n\|x_n - x_{n-1}\| \\ &\leq [1 - \beta_n(\theta - \tau L_1)]\|x_n - x^*\| + \beta_n(\theta - \tau L_1)\frac{\beta_n\|\tau h(x^*) - \mu Dx^*\| + \alpha_n\|x_n - x_{n-1}\|}{\beta_n(\theta - \tau L_1)}.\end{aligned}$$

In view of $0 \leq \tau L_1 < \theta$ and $\lim_{n\to\infty}\frac{\alpha_n}{\beta_n}\|x_n - x_{n-1}\| = 0$, there exists a non-negative constant $M > 0$ such that

$$M/2 = \max\left\{\frac{\|\tau h(x^*) - \mu Dx^*\|}{\theta - \tau L_1}, \frac{\alpha_n\|x_n - x_{n-1}\|}{\beta_n(\theta - \tau L_1)}\right\}.$$

Therefore, the above inequality can be characterized as follows:

$$\begin{aligned}\|x_{n+1} - x^*\| &\leq [1 - \beta_n(\theta - \tau L_1)]\|x_n - x^*\| + \beta_n(\theta - \tau L_1)M \\ &\leq \max\left\{\|x_n - x^*\|, M\right\} \leq \cdots \leq \max\left\{\|x_0 - x^*\|, M\right\}.\end{aligned} \tag{10}$$

In addition, by applying the same method as in (8) to Algorithm 2, we also have

$$\|y_n - x^*\|^2 \leq \|z_n - x^*\|^2 - \lambda_n(1-\sigma_n)\|(I - W_2)Az_n\|^2. \tag{11}$$

Since W_1 is nonexpansive, then $\|z_n - x^*\| \leq \|u_n - x^*\|$. Similarly, we can also obtain the same relation (10) above. To sum up, the sequence $\{x_n\}$ generated by Algorithms 1 and 2 is bounded. Furthermore, if $\lambda_n = 0$, the above conclusion still holds. □

Theorem 1. The iterative sequence $\{x_n\}$ generated by Algorithm 1 converges in norm to a point $x^* = P_\Omega \circ (I - \mu D + \tau h)(x^*)$, which is also a unique solution of the following variational inequality

$$\langle(\mu D - \tau h)x^*, x - x^*\rangle \geq 0, \ \forall x \in \Omega. \tag{12}$$

Proof. It follows from Remark 2 that P_Ω is well-defined. Using Lemma 5, we have that $I - \mu D$ is contraction mapping with coefficient $1 - \theta$. Further, for any $x, y \in \mathcal{H}$,

$$\|(I - \mu D + \tau h)x - (I - \mu D + \tau h)y\| \leq \tau L_1\|x - y\| + (1-\theta)\|x - y\| = (1 - (\theta - \tau L_1))\|x - y\|.$$

By the control condition (C4), we know $0 \leq \tau L_1 < \theta$. Thus, $P_\Omega \circ (I - \mu D + \tau h)$ is a contraction mapping with coefficient $1 - (\theta - \tau L_1)$. By virtue of Banach contraction principle, there exists a unique fixed point x^*, i.e., $x^* = P_\Omega \circ (I - \mu D + \tau h)(x^*)$. Furthermore, such a solution x^* is also equivalent to $\langle(\mu D - \tau h)x^*, x - x^*\rangle \geq 0, \ \forall x \in \Omega$ by (3). Besides, from u_n and (4), we obtain

$$\begin{aligned}\|u_n - x^*\|^2 &= \|x_n + \alpha_n(x_n - x_{n-1}) - x^*\|^2 \\ &\leq \|x_n - x^*\|^2 + 2\alpha_n\langle u_n - x^*, x_n - x_{n-1}\rangle \\ &\leq \|x_n - x^*\|^2 + 2\alpha_n\|u_n - x^*\|\|x_n - x_{n-1}\|.\end{aligned} \tag{13}$$

Because of W_1 is averaged, there exists a constant $\kappa \in (0, 1)$ and a nonexpansive mapping S such that $W_1 = (1-\kappa)I + \kappa S$. By means of (5), (8) and $\Delta_n = u_n - \lambda_n A^*(I - W_2)Au_n$, we have

$$\begin{aligned}\|y_n - x^*\|^2 &= \|(1-\kappa)\Delta_n + \kappa S\Delta_n - x^*\|^2 \\ &\leq (1-\kappa)\|\Delta_n - x^*\|^2 + \kappa\|S\Delta_n - x^*\|^2 - 2\kappa(1-\kappa)\|(I - S)\Delta_n\|^2 \\ &\leq \|\Delta_n - x^*\|^2 - 2\kappa(1-\kappa)\|(I - S)\Delta_n\|^2 \\ &\leq \|u_n - x^*\|^2 - \lambda_n(1-\sigma_n)\|(I - W_2)Au_n\|^2 - 2\kappa(1-\kappa)\|(I - S)\Delta_n\|^2.\end{aligned} \tag{14}$$

Set $H_n = \lambda_n(1-\sigma_n)\|(I-W_2)Au_n\|^2 + 2\kappa(1-\kappa)\|(I-S)\Delta_n\|^2$. From (4) and (13), we have

$$\begin{aligned}
\|x_{n+1}-x^*\|^2 &= \|\beta_n\tau(h(y_n)-h(x^*)) + \beta_n(\tau h(x^*)-\mu Dx^*) + (I-\beta_n\mu D)y_n - (I-\beta_n\mu D)x^*\|^2 \\
&\le \left[\beta_n\tau L_1\|y_n-x^*\| + (1-\beta_n\theta)\|y_n-x^*\|\right]^2 + 2\beta_n\langle \tau h(x^*)-\mu Dx^*, x_{n+1}-x^*\rangle \\
&\le (1-\beta_n(\theta-\tau L_1))\|y_n-x^*\|^2 + 2\beta_n\langle \tau h(x^*)-\mu Dx^*, x_{n+1}-x^*\rangle \\
&\le (1-\beta_n(\theta-\tau L_1))\left(\|u_n-x^*\|^2 - H_n\right) + 2\beta_n\langle \tau h(x^*)-\mu Dx^*, x_{n+1}-x^*\rangle \\
&\le (1-\beta_n(\theta-\tau L_1))\|x_n-x^*\|^2 + 2\beta_n\langle \tau h(x^*)-\mu Dx^*, x_{n+1}-x^*\rangle \\
&\quad + (1-\beta_n(\theta-\tau L_1))\left(2\alpha_n\|u_n-x^*\|\|x_n-x_{n-1}\| - H_n\right).
\end{aligned} \tag{15}$$

Hence, the above inequality leads to the following relations:

$$S_{n+1} \le (1-a_n)S_n + a_n b_n \text{ and } S_{n+1} \le S_n - c_n + d_n,\ n \ge 1,$$

where

$$\begin{aligned}
S_n &= \|x_n-x^*\|^2,\ a_n = \beta_n(\theta-\tau L_1),\ c_n = (1-\beta_n(\theta-\tau L_1))H_n; \\
b_n &= \frac{2\alpha_n(1-a_n)\|u_n-x^*\|\|x_n-x_{n-1}\| + 2\beta_n\langle \tau h(x^*)-\mu Dx^*, x_{n+1}-x^*\rangle}{\beta_n(\theta-\tau L_1)}; \\
d_n &= 2\alpha_n(1-a_n)\|u_n-x^*\|\|x_n-x_{n-1}\| + 2\beta_n\langle \tau h(x^*)-\mu Dx^*, x_{n+1}-x^*\rangle.
\end{aligned}$$

Since $\{x_n\}$ is bounded in Lemma 11, so $\{u_n\}$ and $\{y_n\}$ are also bounded. And then we have $\lim_{n\to\infty} a_n = 0$, $\sum_{n=1}^{\infty} a_n = \infty$, $\lim_{n\to\infty} d_n = 0$ by Conditions (C1) and (C2). After that, we need to show that $\lim_{k\to\infty} b_{n_k} \le 0$ when $\lim_{k\to\infty} c_{n_k} = 0$ for any real number sequence $\{n_k\}$ of $\{n\}$. For this purpose, suppose that $\lim_{k\to\infty} c_{n_k} = 0$, it follows from the definition of H_n that $\lim_{k\to\infty} \|(I-W_2)Au_{n_k}\| = 0$ and $\lim_{k\to\infty} \|(I-S)\Delta_{n_k}\| = 0$. Furthermore,

$$\begin{aligned}
\|y_{n_k}-u_{n_k}\| &\le \|y_{n_k}-\Delta_{n_k}\| + \|\Delta_{n_k}-u_{n_k}\| \\
&= \kappa\|(I-S)\Delta_{n_k}\| + \lambda_{n_k}\|A^*(I-W_2)Au_{n_k}\| \\
&\le \kappa\|(I-S)\Delta_{n_k}\| + \lambda_{n_k}\|A\|\|(I-W_2)Au_{n_k}\| \to 0,\ \text{as } k\to\infty,
\end{aligned}$$

and

$$\begin{aligned}
\|u_{n_k}-W_1u_{n_k}\| &\le \|u_{n_k}-y_{n_k}\| + \|y_{n_k}-W_1u_{n_k}\| \\
&\le \|u_{n_k}-y_{n_k}\| + \lambda_{n_k}\|A\|\|(I-W_2)Au_{n_k}\| \to 0,\ \text{as } k\to\infty.
\end{aligned}$$

On the other hand, from the boundedness of $\{x_{n_k}\}$, there exists a subsequence $\{x_{n_{k_i}}\}$ of $\{x_{n_k}\}$ such that $x_{n_{k_i}}$ weakly converges $\hat{x}$ and $\limsup_{k\to\infty}\langle(\tau h-\mu D)x^*, x_{n_k}-x^*\rangle = \lim_{i\to\infty}\langle(\tau h-\mu D)x^*, x_{n_{k_i}}-x^*\rangle$. By virtue of $\lim_{k\to\infty}\|u_{n_k}-x_{n_k}\| = \lim_{k\to\infty}\alpha_{n_k}\|x_{n_k}-x_{n_k-1}\| = 0$ and the bounded linear operator A, we obtain $u_{n_{k_i}} \rightharpoonup \hat{x}$ and $Au_{n_{k_i}} \rightharpoonup A\hat{x}$. It follows from Lemma 2 that $\hat{x} \in Fix(W_1)$ and $A\hat{x} \in Fix(W_2)$, i.e., $\hat{x} \in \Omega$. Subsequently, we have that $\lim_{i\to\infty}\langle(\tau h-\mu D)x^*, x_{n_{k_i}}-x^*\rangle = \langle(\tau h-\mu D)x^*, \hat{x}-x^*\rangle \le 0$ by (3). In addition,

$$\begin{aligned}
\|x_{n_k+1}-x_{n_k}\| &\le \|x_{n_k+1}-y_{n_k}\| + \|y_{n_k}-u_{n_k}\| + \|u_{n_k}-x_{n_k}\| \\
&\le \beta_{n_k}\|\tau h(y_{n_k})-\mu Dy_{n_k}\| + \|y_{n_k}-u_{n_k}\| + \|u_{n_k}-x_{n_k}\| \to 0,\ \text{as } k\to\infty.
\end{aligned}$$

As a consequence, $\limsup_{k\to\infty}\langle(\tau h-\mu D)x^*, x_{n_k+1}-x^*\rangle \le 0$ and

$$\lim_{n\to\infty}\frac{\alpha_n(1-a_n)\|u_n-x^*\|\|x_n-x_{n-1}\|}{\beta_n(\theta-\tau L_1)} \le \lim_{n\to\infty}\frac{\alpha_n\|u_n-x^*\|\|x_n-x_{n-1}\|}{\beta_n(\theta-\tau L_1)} = 0.$$

This means that $\limsup_{k\to\infty} b_{n_k} \le 0$. It follows from Lemma 7 that $\lim_{n\to\infty}\|x_n-x^*\| = 0$, i.e., the iterative sequence $\{x_n\}$ converges in norm to x^* and $x^* = P_\Omega\circ(I-\mu D+\tau h)(x^*)$. Besides, when $\lambda_n = 0$, the above strong convergence of $\{x_n\}$ is still valid. The proof is completed. □

Theorem 2. The sequence $\{x_n\}$ generated by Algorithm 2 converges in norm to a point $x^* = P_\Omega\circ(I-\mu D+\tau h)(x^*)$, which is also a unique solution of the variational inequality (12).

Proof. For any $x^* \in \Omega$, using the same approach as (14), we have

$$\begin{aligned}
\|z_n-x^*\|^2 &= \|(1-\kappa)u_n + \kappa Su_n - x^*\|^2 \\
&\le (1-\kappa)\|u_n-x^*\|^2 + \kappa\|Su_n-x^*\|^2 - 2\kappa(1-\kappa)\|(I-S)u_n\|^2 \\
&\le \|u_n-x^*\|^2 - 2\kappa(1-\kappa)\|(I-S)u_n\|^2.
\end{aligned} \tag{16}$$

Further, combining (11), (13), (15) and (16), we get

$$\begin{aligned}\|x_{n+1}-x^*\|^2 &\le (1-\beta_n(\theta-\tau L_1))\|y_n-x^*\|^2+2\beta_n\langle \tau h(x^*)-\mu Dx^*, x_{n+1}-x^*\rangle\\ &\le (1-\beta_n(\theta-\tau L_1))\|z_n-x^*\|^2+2\beta_n\langle \tau h(x^*)-\mu Dx^*, x_{n+1}-x^*\rangle\\ &\quad -(1-\beta_n(\theta-\tau L_1))\lambda_n(1-\sigma_n)\|(I-W_2)Az_n\|^2\\ &\le (1-\beta_n(\theta-\tau L_1))\left(\|u_n-x^*\|^2-E_n\right)+2\beta_n\langle \tau h(x^*)-\mu Dx^*, x_{n+1}-x^*\rangle\\ &\le (1-\beta_n(\theta-\tau L_1))\|x_n-x^*\|^2+2\beta_n\langle \tau h(x^*)-\mu Dx^*, x_{n+1}-x^*\rangle\\ &\quad +(1-\beta_n(\theta-\tau L_1))\left(2\alpha_n\|u_n-x^*\|\|x_n-x_{n-1}\|-E_n\right),\end{aligned}$$

where

$$E_n=\lambda_n(1-\sigma_n)\|(I-W_2)Az_n\|^2+2\kappa(1-\kappa)\|(I-S)u_n\|^2.$$

If $\lim_{n\to\infty} E_n=0$, we have $\lim_{n\to\infty}\|(I-W_2)Az_n\|=\lim_{n\to\infty}\|(I-S)u_n\|=0$. Further,

$$\|(I-W_1)u_n\|=\|u_n-z_n\|=\|(I-S)u_n\|\to 0,\ \text{as } n\to\infty.$$

Following the same proof of Theorem 1, we can prove $\lim_{n\to\infty}\|x_n-x^*\|=0$, which implies that the iterative sequence $\{x_n\}$ converges in norm to x^* and $x^*=P_\Omega\circ(I-\mu D+\tau h)(x^*)$. The proof is completed. □

Remark 3. (I) Algorithms 1 and 2 includes Algorithm 2.1 in Yao et al.[8] and Algorithm (8) in Moudafi.[6] These conclusions have been promoted from weak convergence to strong convergence under the condition of an adaptive step size sequence $\{\lambda_n\}$.
(II) The hybrid steepest descent method involving Lipschitz continuous mappings and strongly monotone mappings is set in our algorithms and is a broader method including the viscosity method, the Halpern method, and the Mann-type method.
(III) From Condition (C2), the coefficient α_n of the inertial extrapolation term is easy to find realistically. For example, setting the sequence $\{\alpha_n\}$ is constructed as follows:

$$\alpha_n:=\begin{cases}\min\left\{\alpha,\dfrac{\rho_n}{\|x_n-x_{n-1}\|}\right\}, & x_n\ne x_{n-1},\\ \alpha, & \text{otherwise.}\end{cases}$$

Since the value of $\|x_n-x_{n-1}\|$ is known in each iteration of Algorithms 1 and 2, $\{\rho_n\}$ can be chosen by $\rho_n=o(\beta_n)$ and $\alpha\in[0,1)$. On the other hand, from Condition (C1), we consider the sequence $\{\beta_n\}$ generated by $\beta_n=n^{-p}$ with $(0<p\le 1)$, then $\{\alpha_n\}$ is obtained by $\alpha_n=n^{-q}$ with $(q>p)$. For more detail, see[16].
(IV) When $\alpha_n=0$, Algorithms 1 and 2 are reduced to the case without inertial extrapolation terms, as well as Theorems 1 and 2 are also guaranteed under the same conditions.

In what follows, from the definition of the Lipschitz continuous mapping, we know that the Lipschitz continuous mapping includes the contraction mapping. Hence, the following corollary is obtained.

Corollary 1. Let $h:\mathcal{H}_1\to\mathcal{H}_1$ be a contraction mapping with coefficient $\xi\in[0,1)$. If Condition (C4) is replaced with $0\le\tau\xi<\theta=1-\sqrt{1-\mu(2\eta-\mu L_2^2)}$ and $0<\mu<\min\{\frac{1}{2\eta},\frac{2\eta}{L_2^2}\}$, the strong convergence of $\{x_n\}$ is still guaranteed in Algorithms 1 and 2.

In addition, when $\tau=0$, Step 3 in Algorithm 1 is reduced to the same form as in Yamada[11]. Then, the following corollary holds.

Corollary 2. Assume that (A1)-(A3), (A5) and (C1)-(C4) hold. Take any initial points $x_0,x_1\in\mathcal{H}_1$, the sequence $\{x_n\}$ is generated by the following way: u_n, y_n are constructed as in Algorithm 1. If $u_n=y_n$, then stop. Otherwise, calculate

$$x_{n+1}=y_n-\beta_n\mu D(y_n),\ n\ge 1. \tag{17}$$

The iterative sequence $\{x_n\}$ generated by the above algorithm converges in norm to a point $x^*=P_\Omega\circ(I-\mu D)(x^*)$, which is a unique solution of the following variational inequality

$$\langle Dx^*, x-x^*\rangle\ge 0,\ \forall x\in\Omega.$$

In this case that $D:\mathcal{H}_1\to\mathcal{H}_1$ is an identity mapping, $h:\mathcal{H}_1\to\mathcal{H}_1$ is a contraction mapping and $\mu=\tau=1$, Step 3 in Algorithm 1 is taken as the viscosity algorithm in Moudafi[9]. Thus, the following corollary is produced by Lemma 6 and Theorem 1.

Corollary 3. Assume that (A1)-(A3) and (C1)-(C3) hold. Let $h : \mathcal{H}_1 \to \mathcal{H}_1$ be a contraction mapping with coefficient $\xi \in [0, 1)$. Take any initial points $x_0, x_1 \in \mathcal{H}_1$, the sequence $\{x_n\}$ is generated by the following way: u_n, y_n are constructed as in Algorithm 1. If $u_n = y_n$, then stop. Otherwise, calculate

$$x_{n+1} = \beta_n h(y_n) + (1 - \beta_n)y_n, \ n \geq 1. \tag{18}$$

The iterative sequence $\{x_n\}$ converges in norm to a point $x^* = P_\Omega \circ h(x^*)$, which is a unique solution of the following variational inequality

$$\langle x^* - h(x^*), x - x^* \rangle \geq 0, \ \forall x \in \Omega.$$

Proposition 1. In Corollary 3, if contraction mapping h is a constant mapping, the formula (18) is replaced with

$$x_{n+1} = \beta_n u + (1 - \beta_n)y_n, \ n \geq 1.$$

Then the sequence $\{x_n\}$ converges strongly to a point $x^* = P_\Omega(u)$.

In addition, when $D : \mathcal{H}_1 \to \mathcal{H}_1$ is an identity mapping, $\mu = 1$ and $\tau = 0$, Algorithm 1 degenerates to a Mann-type algorithm and its strong convergence is obtained by Lemma 6 and Theorem 1.

Corollary 4. Assume that (A1)-(A3) and (C1)-(C3) hold. Take any initial points $x_0, x_1 \in \mathcal{H}_1$, $\{x_n\}$ is generated by the following scheme: u_n, y_n are constructed as in Algorithm 1. If $u_n = y_n$, then stop. Otherwise, calculate

$$x_{n+1} = (1 - \beta_n)y_n, \ n \geq 1.$$

Then $\{x_n\}$ converges in norm to a point $x^* = P_\Omega(0)$, which is the minimum-norm element of Ω.

Remark 4. The special settings in the above Corollaries can also be implemented in Algorithm 2, and the corresponding strong convergence is still satisfied.

5 | THEORETICAL APPLICATIONS

In this section, our results in Sect. 3 will be applied to other split problems, and also extend and generalize the known results. These conclusions are also helpful for their further research in the future. Moreover, some examples in practical applications are considered and solved by our algorithms. For the sake of simplicity, let $h : \mathcal{H}_1 \to \mathcal{H}_1$ be a L_1-Lipschitz continuous mapping with $L_1 > 0$ and $D : \mathcal{H}_1 \to \mathcal{H}_1$ be a L_2-Lipschitz continuous and η-strongly monotone mapping with $L_2, \eta > 0$. The related lemmas and theorems are given below.

5.1 | Split variational inclusion problems

As one of the important special cases of the SMVIP, the split variational inclusion problem has a wide range of application background, such as split minimization problems, split feasibility problems, split equilibrium problems and so on. In other words, when $f_1 \equiv 0$ and $f_2 \equiv 0$, the SMVIP is reduced to the split variational inclusion problem. For the sake of convenience, we denote by Γ the solution set of the split variational inclusion problem, i.e., $\Gamma = \{x^* \in \mathcal{H}_1 : 0 \in B_1^{-1}(x^*) \text{ and } 0 \in B_2^{-1}(Ax^*)\}$. Therefore, the results in Theorems 1 and 2 are applied to the split variational inclusion problem. Before this, the following important properties need to be reviewed. For any $\gamma > 0$, $J_\gamma^{B_i}$ represents the resolvent mapping of B_i and defined as $J_\gamma^{B_i} = (I + \gamma B_i)^{-1}$, $(i = 1, 2)$. Then $J_\gamma^{B_i}$ is a single-valued and firmly nonexpansive mapping and $Fix(J_\gamma^{B_i}) \Leftrightarrow B_i^{-1}(0) = \{x^* \in \mathcal{D}(B_i) : 0 \in B_i(x^*)\}$. For further detail, see.[22,26]

Theorem 3. Let $\mathcal{H}_1$, $\mathcal{H}_2$ be two Hilbert spaces and $A : \mathcal{H}_1 \to \mathcal{H}_2$ be a bounded linear operator with adjoint operator A^*. Let $B_1 : \mathcal{H}_1 \to 2^{\mathcal{H}_1}$ and $B_2 : \mathcal{H}_2 \to 2^{\mathcal{H}_2}$ be two set-valued maximal monotone mappings. Choose arbitrary initial points $x_0, x_1 \in \mathcal{H}_1$, $\{x_n\}$ is constructed by the following process:

$$\begin{cases} u_n = x_n + \alpha_n(x_n - x_{n-1}), \\ y_n = J_\gamma^{B_1}\left(u_n - \lambda_n A^*(I - J_\gamma^{B_2})Au_n\right), \\ x_{n+1} = \beta_n \tau h(y_n) + (I - \beta_n \mu D)y_n, \ n \geq 1, \end{cases}$$

where

$$\lambda_n = \begin{cases} \dfrac{\sigma_n \|(I - J_\gamma^{B_2})Au_n\|^2}{\|A^*(I - J_\gamma^{B_2})Au_n\|^2}, & Au_n \notin Fix(J_\gamma^{B_2}), \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that Conditions (C1)-(C2) and (C4) hold. If $y_n = u_n$, then stop and $u_n \in \Gamma$. Otherwise, the sequence $\{x_n\}$ converges in norm to a point $x^* = P_\Gamma \circ (I - \mu D + \tau h)(x^*)$, which is a unique solution of the following variational inequality

$$\langle (\mu D - \tau h)x^*, x - x^* \rangle \geq 0,\ \forall x \in \Gamma. \tag{19}$$

Theorem 4. Let $\mathcal{H}_1$, $\mathcal{H}_2$, A, A^*, B_1, B_2 be the same as those in Theorem 3. Put any initial points $x_0, x_1 \in \mathcal{H}_1$, the sequence $\{x_n\}$ is generated by the following way:

$$\begin{cases} u_n = x_n + \alpha_n(x_n - x_{n-1}), \\ z_n = J_\gamma^{B_1}(u_n), \\ y_n = z_n - \lambda_n A^*(I - J_\gamma^{B_2})Az_n, \\ x_{n+1} = \beta_n \tau h(y_n) + (I - \beta_n \mu D)y_n,\ n \geq 1, \end{cases}$$

where

$$\lambda_n = \begin{cases} \dfrac{\sigma_n \|(I - J_\gamma^{B_2})Az_n\|^2}{\|A^*(I - J_\gamma^{B_2})Az_n\|^2}, & Az_n \notin Fix(J_\gamma^{B_2}), \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that Conditions (C1)-(C2) and (C4) hold. If $y_n = z_n = u_n$, then stop and $u_n \in \Gamma$. Otherwise, the sequence $\{x_n\}$ converges in norm to a point $x^* = P_\Gamma \circ (I - \mu D + \tau h)(x^*)$, which is a unique solution of the variational inequality (19).

5.2 | Split variational inequality problems

Let C be a nonempty closed convex subset of a Hilbert space $\mathcal{H}_1$. Define the normal cone $N_C(x)$ of C at a point $x \in C$ by

$$N_C(x) = \{z \in \mathcal{H}_1 : \langle z, y - x \rangle \leq 0,\ \forall y \in C\}.$$

Obviously, $u = (I + \gamma N_C)^{-1}x \Leftrightarrow x - u \in N_C(u) \Leftrightarrow \langle x - u, y - u \rangle \leq 0,\ \forall y \in C \Leftrightarrow u = P_C x$, which implies that $(I + \gamma N_C)^{-1} = P_C$. Let C_1 and Q_1 be nonempty closed convex subsets of Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. Therefore, when $B_1 = N_{C_1}$ and $B_2 = N_{Q_1}$ in SMVIP, the following split variational inequality problem is obtained:

$$\text{find } x^* \in C_1,\ \langle f_1(x^*), x - x^* \rangle \geq 0,\ \forall x \in C_1 \text{ and } \langle f_2(Ax^*), y - Ax^* \rangle \geq 0,\ \forall y \in Q_1.$$

This is equivalent to the following form:

$$\text{find } x^* \in C_1,\ x^* \in Fix(P_{C_1}(I - \gamma f_1)) \text{ and } Ax^* \in Fix(P_{Q_1}(I - \gamma f_2)),\ \text{for any } \gamma > 0.$$

Meanwhile, we denote by Ψ the solution set of the above problem. In particular, if the mapping f_1 is ϑ_1-inverse strongly monotone and $\gamma \in (0, 2\vartheta_1)$, then $P_{C_1}(I - \gamma f_1)$ is average. Indeed, from Remark 1 (v), Lemma 3 (III) and $\gamma \in (0, 2\vartheta_1)$, the mapping $I - \gamma f_1$ is average. Furthermore, P_{C_1} is firmly nonexpansive, which means that P_{C_1} is average. So, $P_{C_1}(I - \gamma f_1)$ is average. Then, the following results can be obtained from our Theorems 1 and 2.

Theorem 5. Let $\mathcal{H}_1, \mathcal{H}_2, C_1, Q_1$ be the same as above. Let $A : \mathcal{H}_1 \to \mathcal{H}_2$ be a bounded linear operator with adjoint operator A^*. Let $f_1 : \mathcal{H}_1 \to \mathcal{H}_1$ be ϑ_1-inverse strongly monotone mapping and $f_2 : \mathcal{H}_2 \to \mathcal{H}_2$ be ϑ_2-inverse strongly monotone mapping. Select arbitrary initial points $x_0, x_1 \in \mathcal{H}_1$, $\{x_n\}$ is generated by the following scheme:

$$\begin{cases} u_n = x_n + \alpha_n(x_n - x_{n-1}), \\ y_n = P_{C_1}(I - \gamma f_1)\left(u_n - \lambda_n A^*(I - P_{Q_1}(I - \gamma f_2))Au_n\right), \\ x_{n+1} = \beta_n \tau h(y_n) + (I - \beta_n \mu D)y_n,\ n \geq 1, \end{cases} \tag{20}$$

where

$$\lambda_n = \begin{cases} \dfrac{\sigma_n \|(I - P_{Q_1}(I - \gamma f_2))Au_n\|^2}{\|A^*(I - P_{Q_1}(I - \gamma f_2))Au_n\|^2}, & Au_n \notin Fix(P_{Q_1}(I - \gamma f_2)), \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that Conditions (C1)-(C4) are satisfied. If $y_n = u_n$, then stop and $u_n \in \Psi$. Otherwise, the sequence $\{x_n\}$ converges in norm to a point $x^* = P_\Psi \circ (I - \mu D + \tau h)(x^*)$, which is a unique solution of the following variational inequality

$$\langle (\mu D - \tau h)x^*, x - x^* \rangle \geq 0, \ \forall x \in \Psi. \tag{21}$$

Theorem 6. Let $\mathcal{H}_1, \mathcal{H}_2, C_1, Q_1, A, A^*, f_1$ and f_2 be the same as those in Theorem 5 and λ_n be the same as in Algorithm 20. Take any initial points $x_0, x_1 \in \mathcal{H}_1$, the sequence $\{x_n\}$ is generated by the following process:

$$\begin{cases} u_n = x_n + \alpha_n(x_n - x_{n-1}), \\ z_n = P_{C_1}(I - \gamma f_1)(u_n), \\ y_n = z_n - \lambda_n A^*(I - P_{Q_1}(I - \gamma f_2))Az_n, \\ x_{n+1} = \beta_n \tau h(y_n) + (I - \beta_n \mu D)y_n, \ n \geq 1, \end{cases}$$

where

$$\lambda_n = \begin{cases} \dfrac{\sigma_n \|(I - P_{Q_1}(I - \gamma f_2))Az_n\|^2}{\|A^*(I - P_{Q_1}(I - \gamma f_2))Az_n\|^2}, & Az_n \notin Fix(P_{Q_1}(I - \gamma f_2)), \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that Conditions (C1)-(C4) are satisfied. If $y_n = z_n = u_n$, then stop and $u_n \in \Psi$. Otherwise, the sequence $\{x_n\}$ converges in norm to a point $x^* = P_\Psi \circ (I - \mu D + \tau h)(x^*)$, which is a unique solution of the variational inequality (21).

5.3 | Split feasibility problems

From (1), we know that the SFP is a special case of the SMVIP. According to Subsection 5.2, we have $(I + \gamma N_{C_1})^{-1} = P_{C_1}$ and $(I + \gamma N_{Q_1})^{-1} = P_{Q_1}$. Meanwhile, the solution set for SFP is called Φ. Thus, the following algorithms and theorems can be derived for finding the solution of the split feasibility problem.

Theorem 7. Let $\mathcal{H}_1, \mathcal{H}_2, C_1, Q_1, A$ and A^* be the same as above. Select arbitrary initial points $x_0, x_1 \in \mathcal{H}_1$, $\{x_n\}$ is generated by the following scheme:

$$\begin{cases} u_n = x_n + \alpha_n(x_n - x_{n-1}), \\ y_n = P_{C_1}\left(u_n - \lambda_n A^*(I - P_{Q_1})Au_n\right), \\ x_{n+1} = \beta_n \tau h(y_n) + (I - \beta_n \mu D)y_n, \ n \geq 1, \end{cases}$$

where

$$\lambda_n = \begin{cases} \dfrac{\sigma_n \|(I - P_{Q_1})Au_n\|^2}{\|A^*(I - P_{Q_1})Au_n\|^2}, & Au_n \notin Q_1, \\ 0, & \text{otherwise.} \end{cases}$$

Assume that Conditions (C1)-(C2) and (C4) hold. If $y_n = u_n$, then stop and $u_n \in \Phi$. Otherwise, the sequence $\{x_n\}$ converges in norm to a point $x^* = P_\Phi \circ (I - \mu D + \tau h)(x^*)$, which is a unique solution of the following variational inequality

$$\langle (\mu D - \tau h)x^*, x - x^* \rangle \geq 0, \ \forall x \in \Phi. \tag{22}$$

Theorem 8. Let $\mathcal{H}_1, \mathcal{H}_2, C_1, Q_1, A$ and A^* be the same as those in Theorem 7. Take any initial points $x_0, x_1 \in \mathcal{H}_1$, the sequence $\{x_n\}$ is generated by the following process:

$$\begin{cases} u_n = x_n + \alpha_n(x_n - x_{n-1}), \\ z_n = P_{C_1}(u_n), \\ y_n = z_n - \lambda_n A^*(I - P_{Q_1})Az_n, \\ x_{n+1} = \beta_n \tau h(y_n) + (I - \beta_n \mu D)y_n, \ n \geq 1, \end{cases}$$

where

$$\lambda_n = \begin{cases} \dfrac{\sigma_n \|(I - P_{Q_1})Az_n\|^2}{\|A^*(I - P_{Q_1})Az_n\|^2}, & Az_n \notin Q_1, \\ 0, & \text{otherwise.} \end{cases}$$

Assume that Conditions (C1)-(C2) and (C4) hold. If $y_n = z_n = u_n$, then stop and $u_n \in \Phi$. Otherwise, the sequence $\{x_n\}$ converges in norm to a point $x^* = P_\Phi \circ (I - \mu D + \tau h)(x^*)$, which is a unique solution of the variational inequality (22).

As an important part of the split monotone variational inclusion problem, the split feasibility problem is widely used to solve practical problems in various situations and many excellent results have been obtained. In what follows, two examples in L^2 spaces and in signal recovery problem are introduced.

Example 5.1 (The split feasibility problem in infinite-dimensional Hilbert spaces)**.** Assume that $\mathcal{H}_1 = \mathcal{H}_2 = L^2([0,1])$ with the inner product $\langle x, y\rangle := \int_0^1 x(t)y(t)\,\mathrm{d}t$ and the induced norm $\|x\| := \left(\int_0^1 |x(t)|^2\,\mathrm{d}t\right)^{1/2}$, for any $x, y \in L^2([0,1])$. Consider the following nonempty closed and convex subsets C_1 and Q_1 in $L^2([0,1])$:

$$C_1 := \left\{ x \in L_2([0,1]) \mid \int_0^1 x(t)\,\mathrm{d}t \le 1 \right\} \text{ and } Q_1 := \left\{ y \in L_2([0,1]) \mid \int_0^1 |y(t) - \sin(t)|^2\,\mathrm{d}t \le 16 \right\}.$$

Suppose that $A : L^2([0,1]) \to L^2([0,1])$ is the Volterra integration operator that is defined by $(Ax)(t) = \int_0^t x(s)\,\mathrm{d}s$, $\forall t \in [0,1]$, $x \in \mathcal{H}_1$. Hence, A is a bounded linear operator and the norm $\|A\| = \frac{2}{\pi}$. Moreover, the adjoint operator A^* of A is defined by $(A^*x)(t) = \int_t^1 x(s)\,\mathrm{d}s$. In addition, its projections on sets C_1 and Q_1 have explicit forms, i.e.,

$$P_{C_1}(x) = \begin{cases} 1 - a + x\,, & a > 1\,; \\ x\,, & a \le 1\,, \end{cases} \text{ and } P_{Q_1}(y) = \begin{cases} \sin(\cdot) + \frac{4(y - \sin(\cdot))}{\sqrt{b}}\,, & b > 16\,; \\ y\,, & b \le 16\,, \end{cases}$$

where $a := \int_0^1 x(t)\,\mathrm{d}t$ and $b := \int_0^1 |y(t) - \sin(t)|^2\,\mathrm{d}t$. Naturally, $x(t) = 0$ is a solution, i.e., the solution set is nonempty.

Example 5.2 (The split feasibility problem in signal recovery problems)**.** It is well known that compressed sensing is one of the effective methods to recover clean signals from polluted signals. In this context, the following underdetermined system problem need to be considered and resolved:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \varepsilon\,,$$

where $\mathbf{y} \in \mathbb{R}^M$ is the observed noise data, $\mathbf{A} \in \mathbb{R}^{M\times N}$ is a bounded linear observation operator, $\mathbf{x} \in \mathbb{R}^N$ with k ($k \ll N$) non-zero elements is the original and clean data that needs to be restored, and ε is the noise observation encountered during data transmission. An important consideration of this problem is that the signal $\mathbf{x}$ is sparse, that is, the number of non-zero elements in the signal $\mathbf{x}$ is much smaller than the dimension of the signal $\mathbf{x}$. To solve this situation, a classical model, convex constraint minimization problem, is used to describe the above problem, i.e.,

$$\min_{\mathbf{x}\in\mathbb{R}^N} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 \quad \text{subject to} \quad \|\mathbf{x}\|_1 \le t\,, \tag{23}$$

where t is a positive constant and $\|\cdot\|_1$ is ℓ_1 norm. It is worth noting that this problem is related to the least absolute shrinkage and selection operator problem. More precisely, the problem (23) is equivalent to the split feasibility problem when $C_1 = \left\{\mathbf{x} \in \mathbb{R}^N : \|\mathbf{z}\|_1 \le t\right\}$ and $Q_1 = \{\mathbf{y}\}$.

Remark 5. (I) All of the above theorems can be derived from the proof of Theorems 1 and 2.
(II) The above theorems generalizes many important results that are available, such as the split feasibility problem,[1,5,7] the split variational inclusion problem,[27,28,29] the split variational inequality problem[30] and so on.

6 | NUMERICAL EXPERIMENTS

In this section, we provide some numerical examples to demonstrate the effectiveness and realization of convergence behavior of Theorems 1 and 2. All the programs were implemented in Matlab 2018a on a Intel(R) Core(TM) i5-8250U CPU @1.60 GHz computer with RAM 8.00 GB.

Theorem 9 (Moudafi[6])**.** Assume that (A1)-(A3) and (C3) hold. Let l be the spectral radius of A^*A and $0 < \lambda < 1/l$. For any $x_1 \in \mathcal{H}_1$, the iterative sequence $\{x_n\}$ is generated by the following iterative scheme:

$$x_{n+1} = J_\gamma^{B_1}(I - \gamma f_1)\left(x_n - \lambda A^*(I - J_\gamma^{B_2}(I - \gamma f_2))Ax_n\right),\ n \ge 1. \tag{24}$$

Then $\{x_n\}$ converges weakly to a point $x^* \in \Omega$.

Theorem 10 (Yao et al.[8]). Assume that (A1)-(A3) and (C3) hold. For any initial points $x_0, x_1 \in \mathcal{H}_1$, the iterative sequence $\{x_n\}$ is generated by the following iterative scheme:

$$\begin{cases} u_n = x_n + \alpha_n(x_n - x_{n-1}), \\ x_{n+1} = J_\gamma^{B_1}(I - \gamma f_1)\left(u_n - \lambda_n A^*(I - J_\gamma^{B_2}(I - \gamma f_2))Au_n\right), n \geq 1. \end{cases} \tag{25}$$

where

$$\bar{\alpha}_n := \begin{cases} \min\left\{\alpha, \dfrac{\epsilon_n}{\|x_n - x_{n-1}\|^2}\right\}, & \text{if } x_n \neq x_{n-1}, \\ \alpha, & \text{if } x_n = x_{n-1}, \end{cases}$$

and

$$\lambda_n := \begin{cases} \dfrac{\sigma_n\|(I - J_\gamma^{B_2}(I - \gamma f_2))Au_n\|^2}{\|A^*(I - J_\gamma^{B_2}(I - \gamma f_2))Au_n\|^2}, & \text{if } (I - J_\gamma^{B_2}(I - \gamma f_2))Au_n \neq 0, \\ \bar{\lambda}, & \text{otherwise}, \end{cases}$$

If $0 \leq \alpha_n \leq \bar{\alpha}_n$, $\alpha \in [0, 1)$ and $\{\epsilon_n\} \subset l_1$, i.e., $\sum_{n=1}^{\infty} \alpha_n \|x_n - x_{n-1}\|^2 < \infty$, $\bar{\lambda} > 0$, $\sigma_n \in (0, 1)$ and $0 < \liminf_{n\to\infty} \sigma_n \leq \limsup_{n\to\infty} \sigma_n < 1$. The sequence $\{x_n\}$ converges weakly to a point $x^* \in \Omega$.

Example 6.1. Assume that $A, A_1, A_2 : \mathbb{R}^m \to \mathbb{R}^m$ are created from a normal distribution with mean zero and unit variance. Let $B_1 : \mathbb{R}^m \to \mathbb{R}^m$ and $B_2 : \mathbb{R}^m \to \mathbb{R}^m$ be defined by $B_1(x) = A_1^* A_1 x$ and $B_2(y) = A_2^* A_2 y$, respectively. Consider the problem of finding a point $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_m)^\top \in \mathbb{R}^m$ such that $B_1(\bar{x}) = (0, \ldots, 0)^\top$ and $B_2(A\bar{x}) = (0, \ldots, 0)^\top$. It is easy to check that the solution of the problem mentioned above is $x^* = (0, \ldots, 0)^\top$. The parameters of all algorithms are set as follows. Set $f_1 = f_2 = 0$ for all algorithms. Take $\gamma = 1$, $\beta_n = 1/(n+1)$, $\sigma_n = 0.5$, $\alpha = 0.5$ and $\epsilon_n = 1/(n+1)^2$ for the proposed Algorithms 1 and 2 and Algorithm (25). Select $h(x) = 0.5x$, $D(x) = 0.5x$, $\tau = 1$ and $\mu = 2$ for the proposed Algorithms 1 and 2. Choose $\lambda = 0.5/\|A^*A\|$ for Algorithm (24). The start points with the initial values $x_0 = x_1 = 20rand(n, 1)$. $D_n = \|x_n - x^*\|$ is used to measure the iteration error of all the algorithms. The stopping condition is $D_n < 10^{-5}$. Table 1 and Fig. 1 describe the numerical behavior of all algorithms with different dimensions.

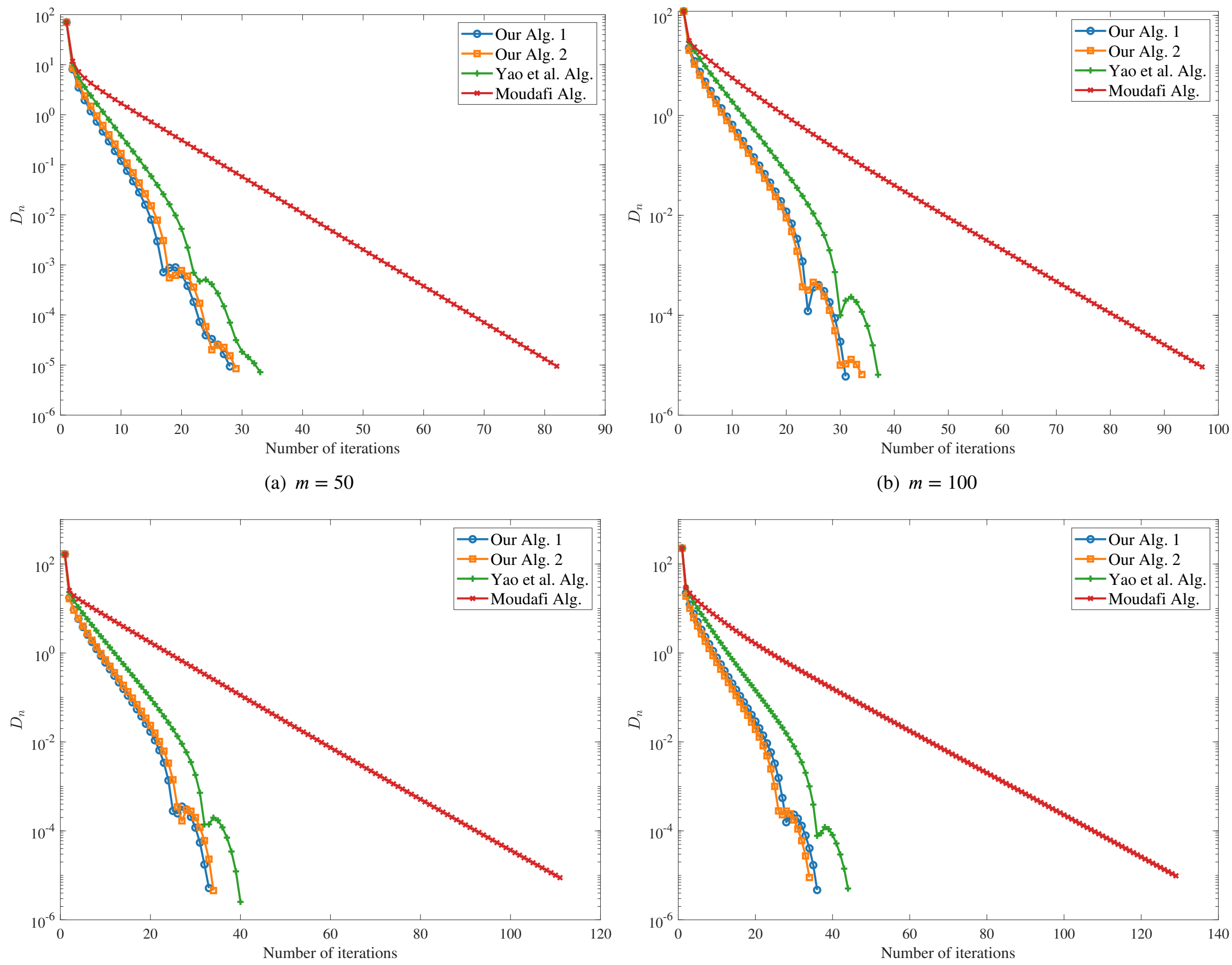
TABLE 1 Numerical results of Example 6.1

Algotithms	$m = 50$		$m = 100$		$m = 200$		$m = 400$	
	Iter.	Time (s)	Iter.	Time (s)	Iter.	Time (s)	Iter.	Time (s)
Our Alg. 1	27	0.0113	30	0.0313	32	0.0744	35	0.3091
Our Alg. 2	28	0.0116	33	0.0341	33	0.0754	33	0.2872
Yao et al. Alg.	32	0.0107	36	0.0342	39	0.0917	43	0.3693
Moudafi Alg.	81	0.0518	96	0.1883	110	0.5751	128	2.7578

Example 6.2. We apply the same algorithms and parameters as in Example 6.1 to solve Example 5.1. The stopping condition is either $D_n = \left\|(I - P_{C_1})x_n\right\|^2 + \left\|A^*(I - P_{Q_1})Ax_n\right\|^2 < 10^{-5}$ or maximum number of iterations which is set to 49. Table 2 and Fig. 2 show the numerical behavior of all algorithms with four different initial values $x_0 = x_1$.

Example 6.3. We now consider using the proposed iterative schemes to solve Example 5.2. In our numerical experiments, the matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$ is created from a standard normal distribution with zero mean and unit variance and then orthonormalizing the rows. The clean signal $\mathbf{x} \in \mathbb{R}^N$ contains k ($k \ll N$) randomly generated ± 1 spikes. The observation $\mathbf{y}$ is formed by $\mathbf{y} = \mathbf{Ax} + \epsilon$ with white Gaussian noise ϵ of variance 10^{-4}. The recovery process starts with the initial signals $\mathbf{x_0} = \mathbf{x_1} = \mathbf{0}$ and ends after 1000 iterations. We use the mean squared error $\text{MSE} = (1/N)\|\mathbf{x}^* - \mathbf{x}\|^2$ ($\mathbf{x}^*$ is an estimated signal of $\mathbf{x}$) to measure the restoration accuracy of all algorithms. In our test, we set $M = 512$, $N = 1024$ and $k = 50$. The parameters of all algorithms are the same as those set in Example 6.1. The recovery results of the suggested algorithms are shown in Fig. 3 .

Remark 6. Based on the results presented in Examples 6.1–6.3, it is easy to get the following observations.



(a) $m = 50$ (b) $m = 100$

(c) $m = 200$ (d) $m = 400$

FIGURE 1 Numerical behavior of all algorithms with different dimensions in Example 6.1

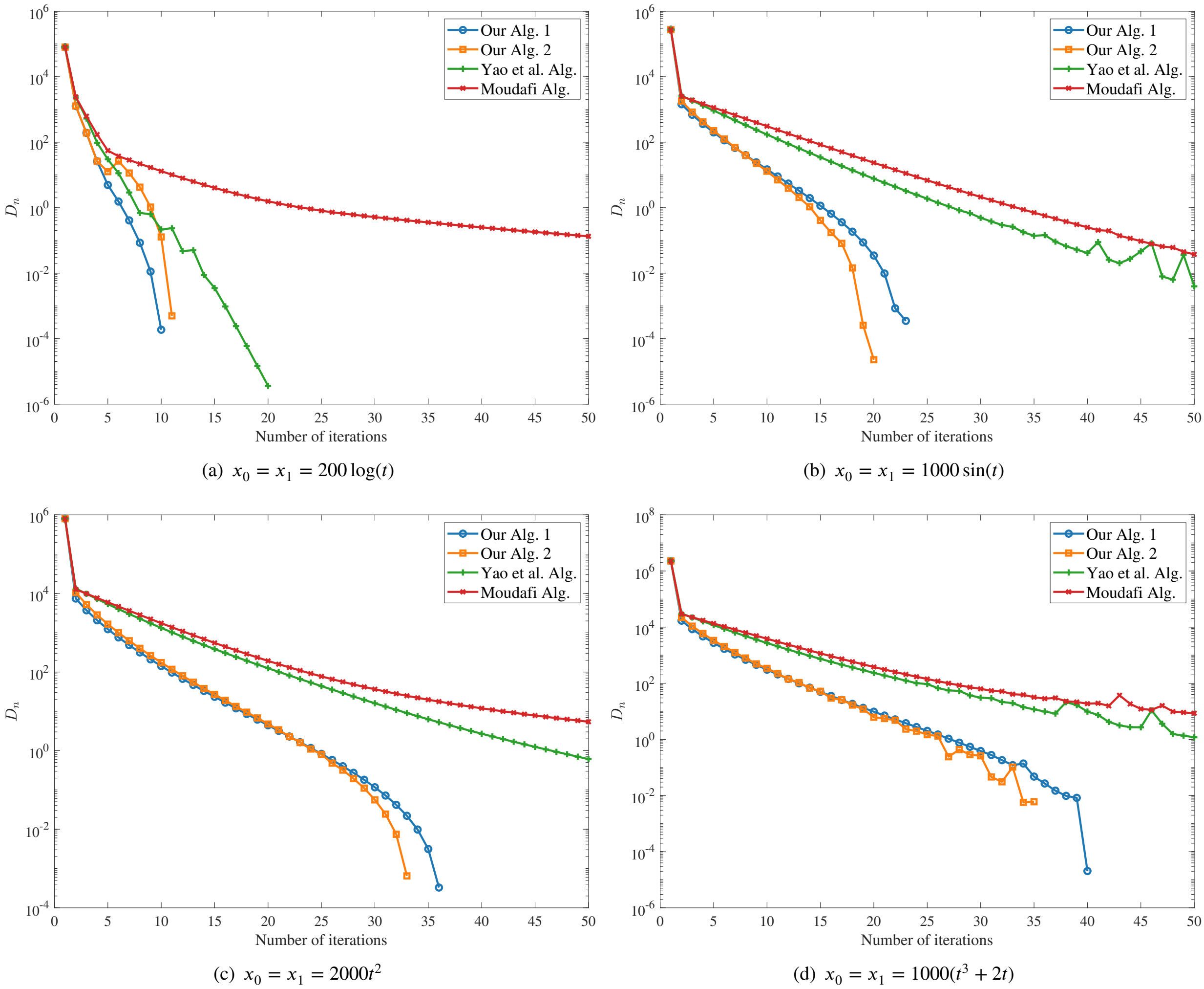
TABLE 2 Numerical results of Example 6.2

Algotithms	$x_1 = 200\log(t)$		$x_1 = 1000\sin(t)$		$x_1 = 2000t^2$		$x_1 = 1000(t^3 + 2t)$	
	Iter.	Time (s)	Iter.	Time (s)	Iter.	Time (s)	Iter.	Time (s)
Our Alg. 1	10	8.0815	23	26.2392	36	21.0198	40	32.0593
Our Alg. 2	11	8.444	20	24.3155	33	18.6057	35	30.3805
Yao et al. Alg.	19	27.1521	49	68.845	49	32.4184	49	44.4225
Moudafi Alg.	49	9.964	49	10.0659	49	8.4375	49	9.672

(1) For different initial values, our algorithms are effective under the excitation of the inertial extrapolation term and the hybrid steepest descent method.

(2) It can be seen from the figures and tables that the convergence behavior of our algorithms is better than that of the existing algorithms in Moudafi[6] and Yao et al.,[8] and these results have nothing to do with the choice of initial values and the size of the dimension.

(3) The adaptive step size is added to our algorithms and the convergence behavior is also maintained.



(a) $x_0 = x_1 = 200\log(t)$

(b) $x_0 = x_1 = 1000\sin(t)$

(c) $x_0 = x_1 = 2000t^2$

(d) $x_0 = x_1 = 1000(t^3 + 2t)$

FIGURE 2 Numerical behavior of all algorithms with different initial values in Example 6.2

7 | CONCLUSION

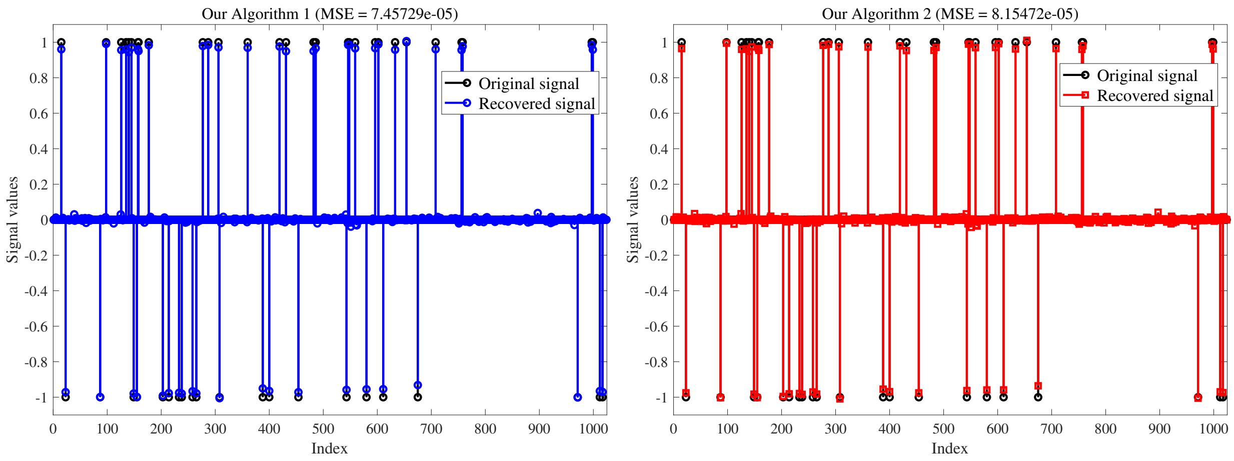
In this article, the main contribution is to introduce two novel inertial iterative algorithms for solving the split monotone variational inclusion problem. Furthermore, the suggested algorithms employs the hybrid steepest descent method, which involves a L-Lipschitz continuous mapping and a strongly monotone mapping. The strong convergence of the proposed algorithms is given through the adaptive step size criterion, which overcomes the fact that the norm of the operator is not easy to calculate in practical applications. The numerical experiment also shows the convergence behavior of our algorithms and its superiority over existing algorithms.

ACKNOWLEDGEMENT

This work has no conflicts of interest and no funds are available to report this submission. The authors also thank the reviewers for their helpful suggestions on our original manuscript.

ORCID

Zheng Zhou https://orcid.org/0000-0003-3847-436X
Bing Tan https://orcid.org/0000-0003-1509-1809



(a) Recovered by our Algorithm 1

(b) Recovered by our Algorithm 2

FIGURE 3 The original signal and the signal recovered by our algorithms

References

1. Censor Y, Elfving T. A multiprojection algorithm using bregman projections in a product space. *Numer Algorithms*. 1994;8:221–239.
2. Censor Y, Bortfeld T, Martin B, Trofimov A. A unified approach for inversion problems in intensity-modulated radiation therapy. *Phys Med Biol*. 2006;51:2353–2365.
3. Shehu Y, Gibali A. New inertial relaxed method for solving split feasibilities. *Optim Lett*. (2020). https://doi.org/10.1007/s11590-020-01603-1
4. Gibali A, Mai DT, Nguyen TV. A new relaxed CQ algorithm for solving split feasibility problems in Hilbert spaces and its applications. *J Ind Manag Optim*. 2019;15:963–984.
5. Byrne C. Iterative oblique projection onto convex sets and the split feasibility problem. *Inverse Problems*. 2002;18:441–453.
6. Moudafi A. Split monotone variational inclusions. *J Optim Theory Appl*. 2011;150:275–283.
7. López G, Martín-Márquez V, Wang F, Xu HK. Solving the split feasibility problem without prior knowledge of matrix norms. *Inverse Problems*. 2012;28:085004.
8. Yao Y, Shehu Y, Li XH, Dong QL. A method with inertial extrapolation step for split monotone inclusion problems. *Optimization*. 2021;70:741–761.
9. Moudafi A. Viscosity approximation methods for fixed-points problems. *J Math Anal Appl*. 2000;241:46–55.
10. Marino G, Xu HK. A general iterative method for nonexpansive mappings in Hilbert spaces. *J Math Anal Appl*. 2006;318:43–52.
11. Yamada I. *The hybrid steepest descent for the variational inequality problems over the intersection of fixed points sets of nonexpansive mapping*, in: D. Butnariu, Y. Censor, S. Reich (Eds.), *Inherently Parallel Algorithms in Feasibility and Optimization and Their Application*, pp. 473–504. Elservier, New York; 2001.
12. Tian M. A general iterative algorithm for nonexpansive mappings in Hilbert spaces. *Nonlinear Anal*. 2010;73:689–694.
13. Alvarez F, Attouch H. An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. *Set-Valued Analysis*. 2001;9:3–11.

14. Gibali A, Hieu DV. A new inertial double-projection method for solving variational inequalities. *J Fixed Point Theory Appl*. 2019;21:97.

15. Shehu Y, Iyiola OS. Projection methods with alternating inertial steps for variational inequalities: weak and linear convergence. *Appl Numer Math*. 2020;157:315–337.

16. Zhou Z, Tan B, Li S. A new accelerated self-adaptive stepsize algorithm with excellent stability for split common fixed point problems. *Comp Appl Math*. 2020;39:220.

17. Zhou Z, Tan B, Li S. An accelerated hybrid projection method with a self-adaptive step-size sequence for solving split common fixed point problems. *Math Methods Appl Sci*. 2021;44:7294–7303.

18. Goebel K, Reich S. *Uniform convexity, hyperbolic geometry, and nonexpansive mappings*. Marcel Dekker, New York and Basel; 1984.

19. Kopecká E, Reich S. A note on alternating projections in Hilbert space. *J Fixed Point Theory Appl*. 2012;12:41–47.

20. Crombez G. A geometrical look at iterative methods for operators with fixed points. *Numer Funct Anal Optim*. 2005;26:157–175.

21. Crombez G. A hierarchical presentation of operators with fixed points on Hilbert spaces. *Numer Funct Anal Optim*. 2006;27:259–277.

22. Zhou H, Qin X. *Fixed Points of Nonlinear Operators: Iterative Methods*. De Gruyter, Berlin; 2020.

23. Byrne C. A unified treatment of some iterative algorithms in signal processing and image reconstruction. *Inverse Problems*. 2004;20:103–120.

24. Thong DV, Triet NA, Li XH, Dong QL. Strong convergence of extragradient methods for solving bilevel pseudo-monotone variational inequality problems. *Numer Algorithms*. 2020;83:1123–1143.

25. He SN, Yang CP. Solving the variational inequality problem defined on intersection of finite level sets. *Abstr Appl Anal*. 2013;2013:942315.

26. Marino G, Xu HK. Convergence of generalized proximal point algorithm. *Commun Pure Appl Anal*. 2004;3:791–808.

27. Long LV, Thong DV, Dung VT. New algorithms for the split variational inclusion problems and application to split feasibility problems. *Optimization*. 2019;68:2339–2367.

28. Byrne C, Censor Y, Gibali A, Reich S. Weak and strong convergence of algorithms for the split common null point problem. *J Nonlinear Convex Anal*. 2011;13:759–775.

29. Anh PK, Thong DV, Dung VT. A strongly convergent Mann-type inertial algorithm for solving split variational inclusion problems. *Optim Eng*. 2021;22:159–185.

30. Censor Y, Gibali A, Reich S. The split variational inequality problem. *The Technion-Israel Institue of Technology*, Haifa September 20; 2010.