Canad. J. Math. 2025, pp. 1–22
http://dx.doi.org/10.4153/S0008414X24000889

Two relaxed inertial forward-backward-forward algorithms for solving monotone inclusions and an application to compressed sensing

Bing Tan and Xiaolong Qin

Abstract. Two novel algorithms, which incorporate inertial terms and relaxation effects, are introduced to tackle a monotone inclusion problem. The weak and strong convergence of the algorithms are obtained under certain conditions, and the R-linear convergence for the first algorithm is demonstrated if the set-valued operator involved is strongly monotone in real Hilbert spaces. The proposed algorithms are applied to signal recovery problems and demonstrate improved performance compared to existing algorithms in the literature.

1 Introduction

In this paper, our main goal is to devise accelerated iterative algorithms for solving the classical zero point problem of the sum of two monotone operators, which is also known as the monotone inclusion problem (shortly, MIP). Recall that the MIP is formed as follows:

$$\boxed{\text{find } x^* \in \mathcal{H} \text{ such that } 0 \in (A+B)x^*,} \tag{1.1}$$

where $\mathcal{H}$ denotes a real Hilbert space with the inner product $\langle \cdot, \cdot \rangle$ and the induced norm $\|\cdot\|$, $A : \mathcal{H} \to \mathcal{H}$ is a single valued operator, and $B : \mathcal{H} \to 2^{\mathcal{H}}$ is a multi-valued operator. The monotone inclusion problem is an important problem widely studied in the fields of mathematics, engineering, physics, economics, and computer science. Its applications are extremely wide-ranging, including portfolio optimization, resource allocation, production planning, optimal control, image processing, signal recovery, and more; see, for example, [1, 2, 3, 4]. The MIP is a challenging problem and has been the subject of extensive research, leading to the development of various algorithms and methods for solving it.

Next, we give the connection between the inclusion problem and the split feasibility problem and show its application to signal processing problems.

Received by the editors January 16, 2024; revised August 30, 2024; accepted September 21, 2024.

B. Tan thanks the support of the Natural Science Foundation of Chongqing (No. CSTB2024NSCQ-MSX0354), the National Natural Science Foundation of China (No. 12471473), and the Fundamental Research Funds for the Central Universities (No. SWU-KQ24052).

AMS subject classification: **47J20**, 49J40, 65K15, 68W10, 90C33.

Keywords: Inclusion problems, monotone operator, signal recovery, forward-backward-forward method, convergence rate.

Example 1.1 The split feasibility problem (SFP) involves finding an x in the nonempty closed convex subset C of real Hilbert space $\mathcal{H}_1$ such that Tx is in another nonempty closed convex subset Q of real Hilbert space $\mathcal{H}_2$, where T is a bounded linear operator mapping from $\mathcal{H}_1$ to $\mathcal{H}_2$. This problem arises in image reconstruction and signal processing. From an optimization perspective, x^* solves SFP if it is a solution to the minimization problem with zero optimal value:

$$\min_{x \in C} h(x) = \frac{1}{2} \|Tx - P_Q Tx\|^2 ,$$

where P_Q is the metric projection of $\mathcal{H}_2$ onto Q. Note that h is a convex differentiable function and its gradient, $\nabla h(x) = T^* (I - P_Q) Tx$, is $\|T\|^2$-Lipschitz continuous and monotone. Hence, x^* solves the SFP if it finds the solutions of the following inclusion problem: find $x^* \in \mathcal{H}_1$ such that $0 \in \nabla h(x^*) + \partial \delta_C(x^*)$, where δ_C is the indicator function of C. In (1.1), setting $A := \nabla h$ and $B := \partial \delta_C$, we can obtain the SFP as a special case of the MIP.

Now we show how the signal processing problem can be modeled in the form of an SFP and thus extended to the inclusion problem. The original signal, $\mathbf{x}$, is a vector in $\mathbb{R}^N$ with only k non-zero elements where k is much smaller than N. The bounded linear operator $\mathbf{C} : \mathbb{R}^{M \times N}$ represents the transformation of the signal during transmission, and ε is the noise introduced during the process. The resulting noise signal, $\mathbf{y}$, can be modeled as $\mathbf{y} = \mathbf{C}\mathbf{x} + \varepsilon$, where $\mathbf{y}$ is a vector in $\mathbb{R}^M$. To solve this model, we can formulate an unconstrained optimization problem as follows:

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \|\mathbf{y} - \mathbf{C}\mathbf{x}\|^2 \text{ subject to } \|\mathbf{x}\|_1 \le t, \quad \text{where } t \text{ is a positive constant.}$$

By setting $C := \{\|\mathbf{x}\|_1 \le t\}$ and $Q := \{y\}$, the optimization problem defined above can be converted into an SFP model; thus, we can solve the signal processing problem by using algorithms that solve the MIP (1.1).

Splitting methods are a well known and important class of techniques for solving inclusion problems. They involve decomposing the original problem into two or simpler subproblems, which can then be solved more easily. The solutions to the subproblems are then combined to find the solution to the original problem. Splitting methods have been widely used in various fields, such as optimization, control theory, and computer science. They are known for their simplicity, efficiency, and versatility, making them a popular choice for solving inclusion problems. Among splitting methods, the forward-backward (FB) splitting algorithms [5, 6] are a prominent method for solving monotone inclusion problems. This method splits the original problem into two subproblems: a forward step, which involves finding a solution that increases the objective function, and a backward step, which involves finding a solution that decreases the objective function. The role of the forward and backward steps is to balance the trade-off between convergence and feasibility, allowing the algorithm to converge to the optimal solution while ensuring that the constraints are satisfied. By alternating between the forward and backward steps, the forward-backward splitting algorithm is able to find the optimal solution to the monotone inclusion problem in an efficient and effective manner. The forward-backward splitting algorithms have been extensively studied and are considered to be a powerful tool for solving monotone inclusion problems in various fields because they are simple to implement, computationally efficient, and can be easily extended to handle more complex problems.

The convergence of an FB algorithm can be slow or inefficient when the problem is highly nonlinear or when the objective function has multiple local minima. To address these challenges, a more advanced variant of the forward-backward splitting algorithm called the forward-backward-forthward (FBF) algorithm (sometimes known as the Tseng's algorithm) has been developed by Tseng [7]. The FBF algorithm extends the forward-backward algorithm by adding a third step, called the forward step, which involves applying an operator that maps the current solution to a new solution that is guaranteed to increase the objective function even further. By incorporating the forward step, the FBF algorithm is able to overcome the convergence limitations of the FB algorithm and find the optimal solution more quickly and efficiently, especially for problems with highly nonlinear constraints or multiple local minima. For more improved versions of the FB algorithm and the FBF algorithm for solving monotone inclusion problems and their convergence results, see, for example, [8, 9, 10, 11, 12, 13, 14, 15, 16]. However, strong convergence is considered to be a more important concept than weak convergence in infinite dimensional spaces. This is because strong convergence implies that the limit of a sequence of functions converges not only in the sense of distribution but also pointwise. This stronger convergence notion allows for a more robust and accurate analysis of the limit behavior and often provides stronger conclusions about the limit. In contrast, weak convergence only guarantees convergence in distribution, which may not capture the fine details of the limit behavior. As a result, strong convergence is often preferred in infinite dimensional spaces when a more precise and rigorous analysis is desired. In 2018, Gibali and Thong [17] proposed two modification schemes of the FBF algorithm by combining the FBF algorithm with the Mann-type method and the viscosity-type method, respectively. They established strong convergence theorems for the proposed algorithms, provided that the parameters and operators satisfy some suitable conditions. More precisely, their two iterative algorithms are shown below.

Gibali et al.'s Mann-type FBF Algorithm 1 [17]
Initialization: Given $\psi_1 > 0, \eta \in (0,1)$. Let $\{\alpha_n\}$ and $\{\delta_n\}$ be two real sequences in $(0,1)$ such that $\{\delta_n\} \subset (a,b) \subset (0, 1-\alpha_n)$ for some $a > 0$, $b > 0$ and $\lim_{n\to\infty} \alpha_n = 0, \sum_{n=1}^{\infty} \alpha_n = \infty$. Let $t_0 \in \mathcal{H}$ be arbitrary.
Iterative process:

$$\begin{cases} g_n = (I + \psi_n B)^{-1} (I - \psi_n A)\, t_n, \\ u_n = g_n - \psi_n (Ag_n - At_n), \\ t_{n+1} = (1 - \alpha_n - \delta_n) t_n + \delta_n u_n, \quad \forall n \geq 1. \end{cases}$$

Update ψ_{n+1} by

$$\psi_{n+1} = \begin{cases} \min\left\{ \frac{\eta \|t_n - g_n\|}{\|At_n - Ag_n\|}, \psi_n \right\}, & \text{if } At_n - Ag_n \neq 0, \\ \psi_n, & \text{otherwise.} \end{cases} \tag{1.2}$$

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

Gibali et al.'s viscosity-type FBF Algorithm 2 [17]
Initialization: Given $\psi_1 > 0, \eta \in (0,1)$. Let $f : \mathcal{H} \to \mathcal{H}$ be a contraction mapping with constant $\rho \in (0,1)$. Assume $\{\alpha_n\}$ is a real sequence such that $\lim_{n\to\infty} \alpha_n = 0, \sum_{n=1}^{\infty} \alpha_n = \infty$. Let $t_0 \in \mathcal{H}$ be arbitrary.
Iterative process:

$$\begin{cases} g_n = (I + \psi_n B)^{-1} (I - \psi_n A) t_n, \\ u_n = g_n - \psi_n (A g_n - A t_n), \\ t_{n+1} = \alpha_n f(t_n) + (1 - \alpha_n) u_n, \quad \forall n \geq 1. \end{cases}$$

Update ψ_{n+1} by (1.2).

Recently, in order to speed up the convergence speed of algorithms proposed by Gibali et al. [17] and also to speed up the FBF algorithm, Thong et al. [18] combined the FBF algorithm with inertial and relaxation techniques. Indeed, they proposed a modified FBF algorithm with a non-monotonic sequence of step sizes as follows.

Thong et al.'s relaxed FBF Algorithm 1 [18]
Initialization: Given $\psi_1 > 0$, $\gamma \in (0,1)$, $\eta \in (0,1)$, $\sigma \in [0, \frac{1}{2})$. Let $\{\tau_n\}$ be a nonnegative real numbers sequence such that $\sum_{n=1}^{\infty} \tau_n < +\infty$. Select starting points $t_0, t_1 \in \mathcal{H}$.
Iterative process:

$$\begin{cases} d_n = t_n + \gamma (t_n - t_{n-1}), \\ g_n = (I + \psi_n B)^{-1} (I - \psi_n A) d_n, \\ u_n = g_n - \psi_n (A g_n - A d_n), \\ t_{n+1} = (1 - \sigma) t_n + \sigma u_n, \quad \forall n \geq 1. \end{cases}$$

Update ψ_{n+1} by

$$\psi_{n+1} = \begin{cases} \min\left\{ \frac{\eta \|d_n - g_n\|}{\|A d_n - A g_n\|}, \psi_n + \tau_n \right\}, & \text{if } A d_n - A g_n \neq 0, \\ \psi_n + \tau_n, & \text{otherwise}. \end{cases} \tag{1.3}$$

Inspired and motivated by the above results, our main goal in this paper is to establish some accelerated FBF algorithms with convergence guarantee to solve the monotone inclusion problem. The structure of this paper is as follows. In Section 2, we introduce some auxiliary knowledge used throughout the paper. In Section 3, two modified FBF algorithms are presented, which incorporate the influences of inertial and relaxation and introduce a new adaptive step size scheme. Then, weak and strong convergence analyses are performed for both algorithms, respectively. Moreover, the R-linear convergence analysis of the first suggested algorithm is carried out under the condition that the involved operator is strongly monotone. In Section 4, the proposed algorithms are applied to signal recovery problems and compared with some related algorithms. Finally, we summarize the paper in Section 5, the last section.

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

2 Preliminaries

Consider a real Hilbert space $\mathcal{H}$ and a nonempty, closed, and convex subset C of $\mathcal{H}$. The symbol $t_n \rightharpoonup x$ as $n \to \infty$ represents the weak convergence of the sequence $\{t_n\}$ to x, while the expression $t_n \to x$ as $n \to \infty$ indicates the strong convergence of $\{t_n\}$ to x.

For all $x, y \in \mathcal{H}$ and for all $\gamma \in [0,1]$, we have

(i) $\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y\rangle$.
(ii) $\|\gamma x + (1-\gamma)y\|^2 = \gamma\|x\|^2 + (1-\gamma)\|y\|^2 - \gamma(1-\gamma)\|x - y\|^2$.

Definition 2.1 Let $A : \mathcal{H} \to \mathcal{H}$ denote a single-valued operator and $B : \mathcal{H} \to 2^{\mathcal{H}}$ denote a multi-valued operator.

(i) The operator A is called L-`Lipschitz continuous` with $L > 0$ if

$$\|Ax - Ay\| \leq L\|x - y\|, \quad \forall x, y \in \mathcal{H}.$$

(ii) The operator A is called `monotone` if

$$\langle Ax - Ay, x - y\rangle \geq 0, \quad \forall x, y \in \mathcal{H}.$$

(iii) The operator B is called `monotone` if

$$\langle u - v, x - y\rangle \geq 0, \quad \forall x, y \in \mathcal{H}, u \in Bx, v \in By.$$

(iv) The operator B is called μ-`strongly monotone` if there exists a number $\mu > 0$ such that

$$\langle u - v, x - y\rangle \geq \mu\|x - y\|^2, \quad \forall x, y \in \mathcal{H}, u \in Bx, v \in By.$$

(v) The operator B is called `maximal monotone` if it is monotone and if for any $(x, u) \in \mathcal{H} \times \mathcal{H}$, $\langle u - v, x - y\rangle \geq 0$ for every $(y, v) \in \mathrm{Graph}(B)$ (the graph of operator B) implies that $u \in Bx$.

Definition 2.2 For all $x \in \mathcal{H}$, there exists a unique nearest point in C, denoted by $P_C(x)$, such that

$$\|x - P_C(x)\| \leq \|x - y\|, \quad \forall y \in C,$$

where P_C is the `metric projection` of $\mathcal{H}$ onto C.

Remark 2.1 Let C be a nonempty closed convex subset of $\mathcal{H}$. The projection $P_C(x)$ of a point $x \in \mathcal{H}$ onto C is characterized by (see, for example, [19, p. 535, Eq. (29.1)])

$$\langle x - P_C(x), y - P_C(x)\rangle \leq 0, \quad \forall x \in \mathcal{H}, y \in C. \tag{2.1}$$

More information concerning the metric projection can be found in [20, Section 3].

Definition 2.3 [21] A sequence $\{t_n\}$ in $\mathcal{H}$ is said to converge R-linearly to p with rate $\rho \in [0,1)$ if there is a constant $c > 0$ such that $\|t_n - p\| \leq c\rho^n$, $\forall n \in \mathbb{N}$.

These following lemmas are crucial to the convergence analysis of main results.

Lemma 2.2 [22] *Let $A : \mathcal{H} \to \mathcal{H}$ be Lipschitz continuous and monotone and $B : \mathcal{H} \to 2^{\mathcal{H}}$ be maximal monotone. Then the operator $A + B$ is maximal monotone.*

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

Lemma 2.3 [23] *Let $\{t_n\}$, $\{\mu_n\}$, and $\{\alpha_n\}$ be nonnegative sequences such that*

$$t_{n+1} \le t_n + \alpha_n (t_n - t_{n-1}) + \mu_n, \quad \forall n \ge 1.$$

If there exists a real number α with $0 \le \alpha_n \le \alpha < 1$ for all $n \in \mathbb{N}$ and $\sum_{n=1}^{\infty} \mu_n < +\infty$, then the following hold: (i) $\sum_{n=1}^{\infty} [t_n - t_{n-1}]_+ < +\infty$, where $[t]_+ := \max\{t, 0\}$; and (ii) there exists $t^* \in [0, +\infty)$ such that $\lim_{n\to+\infty} t_n = t^*$.

Lemma 2.4 [24] *Let C be a nonempty set of $\mathcal{H}$ and $\{t_n\}$ be a sequence in $\mathcal{H}$. If $\lim_{n\to\infty} \|t_n - x\|$ exists for every $x \in C$ and every sequential weak cluster point of $\{t_n\}$ is in C, then $\{t_n\}$ converges weakly to a point in C.*

Lemma 2.5 [25] *Let $\{t_n\}$ be a nonnegative sequence, $\{\sigma_n\}$ be a sequence of real numbers, and $\{\delta_n\} \subset (0,1)$ be a sequence such that $\sum_{n=1}^{\infty} \delta_n = \infty$. Assume that*

$$t_{n+1} \le (1 - \delta_n) t_n + \delta_n \sigma_n, \quad \forall n \ge 1.$$

If $\limsup_{k\to\infty} \sigma_{n_k} \le 0$ for every subsequence $\{t_{n_k}\}$ of $\{t_n\}$ satisfying $\liminf_{k\to\infty}(t_{n_k+1} - t_{n_k}) \ge 0$, then $\lim_{n\to\infty} t_n = 0$.

3 Main results

In this section, we introduce two new iterative schemes based on the forward-backward-forward algorithm and the techniques of inertial and relaxation to solve the monotone inclusion problem. In the framework of real Hilbert spaces, the first proposed algorithm obtains weak convergence and R-linear convergence, while the second suggested algorithm can achieve strong convergence. Our algorithms directly improve on the results in [1, 2, 3, 17, 18] and many more.

3.1 Weak and linear convergence of Algorithm 3.1

The solution set of problem (1.1) is denoted as $\mathrm{MIP}(A, B)$. Before starting, we first assume that the proposed Algorithm 3.1 satisfies the following conditions.

(C1) The solution set of (1.1) is nonempty (i.e., $\mathrm{MIP}(A, B) := (A + B)^{-1}(0) \ne \emptyset$).
(C2) The operator $A : \mathcal{H} \to \mathcal{H}$ is L-Lipschitz continuous and monotone, and the operator $B : \mathcal{H} \to 2^{\mathcal{H}}$ is maximal monotone.
(C3) Choose the parameters $\eta \in (0,1)$, $\gamma \in [0,1)$, and $\sigma \in (0,1]$ satisfying the following condition:

$$\frac{\sigma(1-\eta^2)}{2 - \sigma + \sigma\eta}(1-\gamma)^2 - (1+\gamma)\gamma > 0. \tag{3.1}$$

Remark 3.1 When the parameters η and γ are fixed, by solving Equation (3.1), we can obtain a lower bound on σ as

$$\underline{\sigma} = \frac{2\omega}{1 - \eta^2 + \omega - \omega\eta}, \quad \text{where } \omega := \frac{\gamma(1+\gamma)}{(1-\gamma)^2}.$$

In other words, the parameter range of σ is $\sigma \in (\underline{\sigma}, 1]$ when η and γ are fixed. Figure 1 gives the variation of $\underline{\sigma}$ with γ when η is fixed.

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

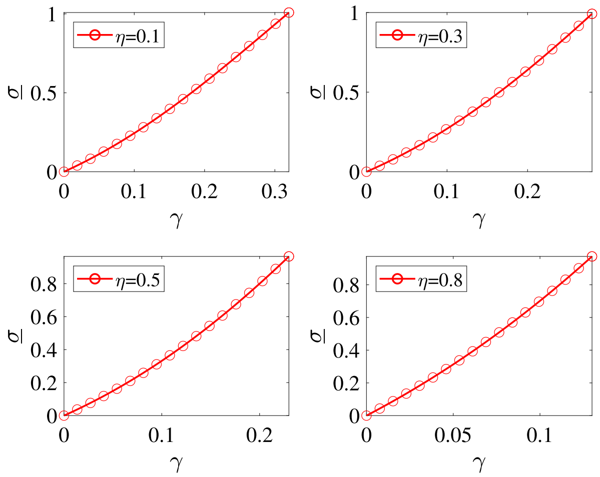


Figure 1: The relationship between the parameters γ, η, and $\underline{\sigma}$.

We now proceed to present the proposed Algorithm 3.1.

Algorithm 3.1

Initialization: Given $\psi_1 > 0$. Select $\eta \in (0,1)$, $\gamma \in [0,1)$, and $\sigma \in (0,1)$ satisfy Equation (3.1). Let $t_0, t_1 \in \mathcal{H}$ and set $n := 1$.

Iterative Steps: Given the iterates t_n, t_{n-1}, perform the following steps.

Step 1. Compute $d_n = t_n + \gamma(t_n - t_{n-1})$.

Step 2. Compute $g_n = (I + \psi_n B)^{-1}(I - \psi_n A)d_n$. If $g_n = d_n$ then stop and $g_n \in \mathrm{MIP}(A,B)$. Otherwise, go to *Step 3*.

Step 3. Compute $u_n = g_n - \psi_n(Ag_n - Ad_n)$.

Step 4. Compute $t_{n+1} = (1-\sigma)d_n + \sigma u_n$, and update ψ_{n+1} by

$$\psi_{n+1} = \begin{cases} \min\left\{\dfrac{\eta\|d_n - g_n\|}{\|Ad_n - Ag_n\|}, \xi_n\psi_n + \tau_n\right\}, & \text{if } Ad_n - Ag_n \neq 0, \\ \xi_n\psi_n + \tau_n, & \text{otherwise.} \end{cases} \tag{3.2}$$

Set $n := n+1$ and go to *Step 1*.

The following lemmas are important for the convergence analysis of Algorithm 3.1.

Lemma 3.2 [26] *Suppose that Condition (C3) holds. Then the step size sequence* $\{\psi_n\}$ *created by (3.2) is well defined, and* $\lim_{n\to\infty}\psi_n$ *exists.*

Lemma 3.3 [27] *Let the sequences* $\{d_n\}$ *and* $\{g_n\}$ *be generated by Algorithm 3.1. If* $\lim_{n\to\infty}\|d_n - g_n\| = 0$ *and* $\{d_{n_k}\}$ *converges weakly to some* $z \in \mathcal{H}$*, then* $z \in \mathrm{MIP}(A,B)$.

Theorem 3.4 (Weak convergence) *Suppose that Conditions (C1)–(C3) hold and let* $\{t_n\}$ *be any sequence formed by Algorithm 3.1. Then* $\{t_n\}$ *converges weakly to an* $p \in \mathrm{MIP}(A,B)$.

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

Proof Let $p \in \mathrm{MIP}(A, B)$. From the definition of g_n, one has $(I - \psi_n A)\, d_n \in (I + \psi_n B)\, g_n$. Since B is maximal monotone, there exists $v_n \in B g_n$ satisfying $(I - \psi_n A)\, d_n = g_n + \psi_n v_n$. This is equivalent to

$$v_n = \psi_n^{-1}\left(d_n - g_n - \psi_n A d_n\right). \tag{3.3}$$

We have $(A + B)$ is maximal monotone by means of Condition (C2) and Lemma 2.2. This combined with $Ag_n + v_n \in (A + B) g_n$ and $0 \in (A + B) p$ implies that $\langle Ag_n + v_n, g_n - p\rangle \geq 0$, which together with (3.3) further yields

$$\langle d_n - g_n - \psi_n\left(Ad_n - Ag_n\right), g_n - p\rangle \geq 0. \tag{3.4}$$

By the definition of u_n, one has

$$\begin{aligned}
\|u_n - p\|^2 &= \|g_n - \psi_n\left(Ag_n - Ad_n\right) - p\|^2 \\
&= \|g_n - p\|^2 + \psi_n^2\|Ag_n - Ad_n\|^2 - 2\psi_n\langle g_n - p, Ag_n - Ad_n\rangle \\
&= \|d_n - p\|^2 + \|d_n - g_n\|^2 + 2\langle g_n - d_n, d_n - p\rangle + \psi_n^2\|Ag_n - Ad_n\|^2 \\
&\quad - 2\psi_n\langle g_n - p, Ag_n - Ad_n\rangle \\
&= \|d_n - p\|^2 + \|d_n - g_n\|^2 - 2\langle g_n - d_n, g_n - d_n\rangle + 2\langle g_n - d_n, g_n - p\rangle \\
&\quad + \psi_n^2\|Ag_n - Ad_n\|^2 - 2\psi_n\langle g_n - p, Ag_n - Ad_n\rangle \\
&= \|d_n - p\|^2 - \|d_n - g_n\|^2 - 2\langle d_n - g_n - \psi_n\left(Ad_n - Ag_n\right), g_n - p\rangle \\
&\quad + \psi_n^2\|Ag_n - Ad_n\|^2.
\end{aligned} \tag{3.5}$$

Using (3.2), (3.4), and (3.5), we have

$$\|u_n - p\|^2 \leq \|d_n - p\|^2 - \left(1 - \eta^2\frac{\psi_n^2}{\psi_{n+1}^2}\right)\|d_n - g_n\|^2, \quad \forall p \in \mathrm{MIP}(A, B). \tag{3.6}$$

It follows from the definition of t_n that

$$\begin{aligned}
\|t_{n+1} - p\|^2 &= \|(1 - \sigma)\left(d_n - p\right) + \sigma\left(u_n - p\right)\|^2 \\
&\leq (1 - \sigma)\|d_n - p\|^2 + \sigma\|u_n - p\|^2.
\end{aligned} \tag{3.7}$$

Substituting (3.6) into (3.7) gives

$$\|t_{n+1} - p\|^2 \leq \|d_n - p\|^2 - \sigma\left(1 - \eta^2\frac{\psi_n^2}{\psi_{n+1}^2}\right)\|d_n - g_n\|^2, \quad \forall p \in \mathrm{MIP}(A, B). \tag{3.8}$$

From the definition of u_n and (3.2), we obtain

$$\|u_n - g_n\| \leq \eta\frac{\psi_n}{\psi_{n+1}}\|d_n - g_n\|.$$

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

This together with the definition u_n yields

$$\begin{aligned} \|t_{n+1}-d_n\| &\leq \|t_{n+1}-g_n\| + \|d_n-g_n\| \\ &= \|(1-\sigma)(d_n-g_n)+\sigma(u_n-g_n)\| + \|d_n-g_n\| \\ &\leq \left(2-\sigma+\sigma\eta\frac{\psi_n}{\psi_{n+1}}\right)\|d_n-g_n\|. \end{aligned} \tag{3.9}$$

Substituting (3.9) into (3.8), we obtain

$$\|t_{n+1}-p\|^2 \leq \|d_n-p\|^2 - \underbrace{\frac{\sigma\left(1-\eta^2\frac{\psi_n^2}{\psi_{n+1}^2}\right)}{2-\sigma+\sigma\eta\frac{\psi_n}{\psi_{n+1}}}}_{\alpha_n}\|t_{n+1}-d_n\|^2. \tag{3.10}$$

By the definition of d_n, one sees that

$$\begin{aligned} \|t_{n+1}-d_n\|^2 &= \|t_{n+1}-t_n-\gamma(t_n-t_{n-1})\|^2 \\ &= \|t_{n+1}-t_n\|^2+\gamma^2\|t_n-t_{n-1}\|^2-2\gamma\langle t_{n+1}-t_n, t_n-t_{n-1}\rangle \\ &\geq \|t_{n+1}-t_n\|^2+\gamma^2\|t_n-t_{n-1}\|^2-2\gamma\|t_{n+1}-t_n\|\|t_n-t_{n-1}\| \\ &\geq (1-\gamma)\|t_{n+1}-t_n\|^2+\left(\gamma^2-\gamma\right)\|t_n-t_{n-1}\|^2. \end{aligned} \tag{3.11}$$

Again using the definition of d_n, we have

$$\begin{aligned} \|d_n-p\|^2 &= \|t_n+\gamma(t_n-t_{n-1})-p\|^2 \\ &= \|(1+\gamma)(t_n-p)-\gamma(t_{n-1}-p)\|^2 \\ &= (1+\gamma)\|t_n-p\|^2-\gamma\|t_{n-1}-p\|^2+\gamma(1+\gamma)\|t_n-t_{n-1}\|^2. \end{aligned} \tag{3.12}$$

Putting (3.11) and (3.12) into (3.10), we obtain

$$\begin{aligned} \|t_{n+1}-p\|^2 \leq{}& (1+\gamma)\|t_n-p\|^2-\gamma\|t_{n-1}-p\|^2+\gamma(1+\gamma)\|t_n-t_{n-1}\|^2 \\ &-\alpha_n\left((1-\gamma)\|t_{n+1}-t_n\|^2+\left(\gamma^2-\gamma\right)\|t_n-t_{n-1}\|^2\right). \end{aligned} \tag{3.13}$$

This implies that

$$\begin{aligned} &\|t_{n+1}-p\|^2-\gamma\|t_n-p\|^2+\left[(1+\gamma)\gamma-\alpha_n\left(\gamma^2-\gamma\right)\right]\|t_{n+1}-t_n\|^2 \\ &\leq \|t_n-p\|^2-\gamma\|t_{n-1}-p\|^2+\left[(1+\gamma)\gamma-\alpha_n\left(\gamma^2-\gamma\right)\right]\|t_n-t_{n-1}\|^2 \\ &\quad-\left[\alpha_n(1-\gamma)-(1+\gamma)\gamma+\alpha_n\left(\gamma^2-\gamma\right)\right]\|t_{n+1}-t_n\|^2. \end{aligned} \tag{3.14}$$

Let

$$\Delta_n := \|t_n-p\|^2-\gamma\|t_{n-1}-p\|^2+\left[(1+\gamma)\gamma-\alpha_n\left(\gamma^2-\gamma\right)\right]\|t_n-t_{n-1}\|^2$$

and

$$\varepsilon_n := \alpha_n(1-\gamma)-(1+\gamma)\gamma+\alpha_n\left(\gamma^2-\gamma\right).$$

From (3.14), one has

$$\Delta_{n+1}-\Delta_n \leq -\varepsilon_n\|t_{n+1}-t_n\|^2.$$

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

Note that $\lim_{n\to\infty} \varepsilon_n > 0$ by Condition (3.1). This means that $\{\Delta_n\}$ is nonincreasing for all $n \geq 1$. From $\lim_{n\to\infty}\left(1-\eta^2\frac{\psi_n^2}{\psi_{n+1}^2}\right) = 1-\eta^2 > 0$, there exists $\varepsilon > 0$ and $N_1 \in \mathbb{N}$ such that $\varepsilon_n > \varepsilon$ for all $n \geq N_1$. Hence,

$$\Delta_{n+1} - \Delta_n \leq -\varepsilon \left\|t_{n+1} - t_n\right\|^2, \quad \forall n \geq N_1. \tag{3.15}$$

From $\sigma \in (0,1]$, $\eta \in (0,1)$, and $\lim_{n\to\infty}\psi_n$ exists, one obtains that α_n as defined in (3.10) is greater than 0 for all $n \geq N_1$. Since $\gamma \in [0,1)$, one can check that

$$(1+\gamma)\gamma - \alpha_n\left(\gamma^2-\gamma\right) \geq 0, \quad \forall n \geq N_1.$$

This together with the definition of Δ_n implies that

$$\Delta_n \geq \left\|t_n - p\right\|^2 - \gamma\left\|t_{n-1}-p\right\|^2.$$

By induction, we have

$$\begin{aligned}
\left\|t_n - p\right\|^2 &\leq \gamma\left\|t_{n-1}-p\right\|^2 + \Delta_n \\
&\leq \gamma\left\|t_{n-1}-p\right\|^2 + \Delta_{N_1} \\
&\leq \cdots \\
&\leq \gamma^{n-N_1}\left\|t_{N_1}-p\right\|^2 + \Delta_{N_1}\left(\gamma^{n-N_1-1}+\cdots+1\right) \\
&\leq \gamma^{n-N_1}\left\|t_{N_1}-p\right\|^2 + \frac{\Delta_{N_1}}{1-\gamma}, \quad \forall n \geq N_1.
\end{aligned} \tag{3.16}$$

It follows from the definition of Δ_{n+1} that

$$\Delta_{n+1} \geq -\gamma\left\|t_n - p\right\|^2, \quad \forall n \geq N_1. \tag{3.17}$$

By using (3.16) and (3.17), we obtain

$$-\Delta_{n+1} \leq \gamma\left\|t_n-p\right\|^2 \leq \gamma^{n-N_1+1}\left\|t_{N_1}-p\right\|^2 + \frac{\gamma\Delta_{N_1}}{1-\gamma}.$$

This together with (3.15) further yields (noting that $\gamma \in [0,1)$)

$$\begin{aligned}
\varepsilon\sum_{n=N_1}^{k}\left\|t_{n+1}-t_n\right\|^2 \leq \Delta_{N_1} - \Delta_{k+1}& \\
&\leq \gamma^{k-N_1+1}\left\|t_{N_1}-p\right\|^2 + \frac{\Delta_{N_1}}{1-\gamma} \\
&\leq \left\|t_{N_1}-p\right\|^2 + \frac{\Delta_{N_1}}{1-\gamma}, \quad \forall k > N_1.
\end{aligned}$$

This implies $\sum_{n=1}^{\infty}\left\|t_{n+1}-t_n\right\|^2 < +\infty$. Therefore, we obtain

$$\lim_{n\to\infty}\left\|t_{n+1}-t_n\right\| = 0. \tag{3.18}$$

By the definition of d_n, one sees that

$$\left\|t_{n+1}-d_n\right\|^2 = \left\|t_{n+1}-t_n\right\|^2 + \gamma^2\left\|t_n - t_{n-1}\right\|^2 - 2\gamma\left\langle t_{n+1}-t_n, t_n - t_{n-1}\right\rangle.$$

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

This combined with (3.18) gives

$$\lim_{n\to\infty} \|t_{n+1} - d_n\| = 0, \quad \lim_{n\to\infty} \|d_n - t_n\| = 0. \tag{3.19}$$

From (3.13), we obtain

$$\begin{aligned} \|t_{n+1} - p\|^2 \le \|t_n - p\|^2 + \gamma\left(\|t_n - p\|^2 - \|t_{n-1} - p\|^2\right) \\ + \left(\gamma(1+\gamma) - \alpha_n\left(\gamma^2 - \gamma\right)\right)\|t_n - t_{n-1}\|^2. \end{aligned} \tag{3.20}$$

Using $\gamma \in (0,1)$, $\sum_{n=1}^{\infty}\|t_n - t_{n-1}\|^2 < +\infty$, (3.20), and Lemma 2.3, we deduce that

$$\lim_{n\to\infty} \|t_n - p\|^2 = l. \tag{3.21}$$

By (3.18), (3.19), and (3.21), we obtain

$$\lim_{n\to\infty} \|t_{n+1} - p\|^2 = l, \quad \lim_{n\to\infty} \|d_n - p\|^2 = l. \tag{3.22}$$

From (3.8) and (3.22), we have

$$\lim_{n\to\infty} \|g_n - d_n\| = 0. \tag{3.23}$$

Since $\{t_n\}$ is a bounded sequence, there exists a subsequence $\{t_{n_k}\}$ of $\{t_n\}$ and $z \in \mathcal{H}$ such that $t_{n_k} \rightharpoonup z$. By means of (3.19), one has $d_{n_k} \rightharpoonup z$. Then it follows from (3.23) and Lemma 3.3 that $z \in \mathrm{MIP}(A,B)$. Therefore, we proved that

(i) for every $p \in \mathrm{MIP}(A,B)$, $\lim_{n\to\infty}\|t_n - p\|$ exists (see (3.21));
(ii) every sequential weak cluster point of $\{t_n\}$ is in $\mathrm{MIP}(A,B)$.

In the light of Lemma 2.4, one concludes that the sequence $\{t_n\}$ converges weakly to $p \in \mathrm{MIP}(A,B)$, as desired. ∎

Next, we show the R-linear convergence of Algorithm 3.1 in the case where the multi-valued operator involved is strongly monotone.

Theorem 3.5 (Linear convergence) *Let p be the unique solution and the sequence $\{t_n\}$ be generated by Algorithm 3.2. Assume that Conditions (C1) and (C2)′ hold.*

(C2)′ *The operator $A : \mathcal{H} \to \mathcal{H}$ is L-Lipschitz continuous and monotone, and the operator $B : \mathcal{H} \to 2^{\mathcal{H}}$ is μ-strongly monotone.*

Choose $\eta \in (0,1)$, $\sigma \in (0,1]$, $\gamma \in [0,1)$, and $\xi \in (0,1)$ such that

$$0 \le \gamma \le \frac{\omega\varepsilon}{\omega\varepsilon + 2\omega + \varepsilon}, \tag{3.24}$$

where ω and ε are defined in (3.30). Then $\{t_n\}$ converges to the solution p of the problem (1.1) with an R-linear rate.

Proof Let $p \in \mathrm{MIP}(A,B)$. It follows from $g_n = (I + \psi_n B)^{-1}(I - \psi_n A)d_n$ that $(I - \psi_n A)\,d_n \in (I + \psi_n B)\,g_n$. Hence,

$$\psi_n^{-1}(d_n - g_n - \psi_n A d_n) \in B g_n.$$

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

Since $0 \in (A+B)p$, one has $-Ap \in Bp$. Using the fact that B is μ-strongly monotone, we have

$$\langle d_n - g_n - \psi_n A d_n + \psi_n A p, g_n - p\rangle \geq \mu\psi_n \|g_n - p\|^2 .$$

This combined with the monotonicity of A implies

(3.25)
$$\begin{aligned}\langle d_n - g_n - \psi_n (Ad_n - Ag_n), g_n - p\rangle &\geq \mu\psi_n \|g_n - p\|^2 + \psi_n \langle Ag_n - Ap, g_n - p\rangle \\ &\geq \mu\psi_n \|g_n - p\|^2 .\end{aligned}$$

By using (3.5), (3.7), and (3.25), we have

$$\begin{aligned}&\|t_{n+1} - p\|^2 \\ &\leq \|d_n - p\|^2 - \sigma\left(1 - \eta^2 \frac{\psi_n^2}{\psi_{n+1}^2}\right)\|d_n - g_n\|^2 - 2\sigma\mu\psi_n \|g_n - p\|^2 \\ &= \|d_n - p\|^2 - \sigma(1-\xi)\left(1 - \eta^2 \frac{\psi_n^2}{\psi_{n+1}^2}\right)\|d_n - g_n\|^2 \\ &\quad - \sigma\xi\left(1 - \eta^2 \frac{\psi_n^2}{\psi_{n+1}^2}\right)\|d_n - g_n\|^2 - 2\sigma\mu\psi_n \|g_n - p\|^2 .\end{aligned} \tag{3.26}$$

By means of (3.9), one sees that

$$\frac{1}{2 - \sigma + \sigma\eta\frac{\psi_n}{\psi_{n+1}}}\|t_{n+1} - d_n\| \leq \|d_n - g_n\| . \tag{3.27}$$

From (3.26) and (3.27), one has

$$\begin{aligned}\|t_{n+1} - p\|^2 \leq \|d_n - p\|^2 &- \frac{\sigma(1-\xi)\left(1 - \eta^2 \frac{\psi_n^2}{\psi_{n+1}^2}\right)}{\left(2 - \sigma + \sigma\eta\frac{\psi_n}{\psi_{n+1}}\right)^2}\|t_{n+1} - d_n\|^2 \\ &- \sigma\xi\left(1 - \eta^2 \frac{\psi_n^2}{\psi_{n+1}^2}\right)\|d_n - g_n\|^2 - 2\sigma\mu\psi_n \|g_n - p\|^2 .\end{aligned} \tag{3.28}$$

Let

$$\delta := \min\left\{\frac{\sigma\xi(1-\eta^2)}{2}, \sigma\mu\psi\right\}, \qquad \text{where } \psi := \lim_{n\to\infty} \psi_n . \tag{3.29}$$

Note that $\delta \in (0,1)$ by means of (3.29), $\sigma \in (0,1)$, $\xi \in (0,1)$, and $\eta \in (0,1)$. Then we obtain

$$\begin{aligned}&\lim_{n\to\infty} \frac{\sigma(1-\xi)\left(1 - \eta^2 \frac{\psi_n^2}{\psi_{n+1}^2}\right)}{\left(2 - \sigma + \sigma\eta\frac{\psi_n}{\psi_{n+1}}\right)^2} = \frac{\sigma(1-\xi)\left(1-\eta^2\right)}{(2-\sigma+\sigma\eta)^2} \geq \frac{\sigma(1-\xi)\left(1-\eta^2\right)}{(2-\sigma+\sigma\eta)^2}\sigma := \varepsilon, \\ &\lim_{n\to\infty} \sigma\xi\left(1 - \eta^2 \frac{\psi_n^2}{\psi_{n+1}^2}\right) = \sigma\xi\left(1-\eta^2\right) \geq 2\delta, \\ &\lim_{n\to\infty} \sigma\mu\psi_n = \sigma\mu\psi \geq \delta.\end{aligned}$$

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

Thus, there exists $N \in \mathbb{N}$ such that

$$\frac{\sigma(1-\xi)\left(1-\eta^2\frac{\psi_n^2}{\psi_{n+1}^2}\right)}{\left(2-\sigma+\sigma\eta\frac{\psi_n}{\psi_{n+1}}\right)^2} \geq \varepsilon, \quad \sigma\xi\left(1-\eta^2\frac{\psi_n^2}{\psi_{n+1}^2}\right) \geq 2\delta, \quad \sigma\mu\psi_n \geq \delta, \quad \forall n \geq N.$$

Let

$$\omega := 1-\delta, \quad \varepsilon := \frac{\sigma^2(1-\xi)\left(1-\eta^2\right)}{\left(2-\sigma+\sigma\eta\right)^2}. \tag{3.30}$$

Since $\xi \in (0,1)$, $\eta \in (0,1)$, and $\sigma \in (0,1]$, we have $\omega \in (0,1)$ and $\varepsilon \in (0,1)$. Using (3.27), we obtain

$$\begin{aligned}\|t_{n+1}-p\|^2 &\leq \|d_n-p\|^2 - \varepsilon\|t_{n+1}-d_n\|^2 - 2\delta\left(\|d_n-g_n\|^2 + \|g_n-p\|^2\right)\\ &\leq \|d_n-p\|^2 - \varepsilon\|t_{n+1}-d_n\|^2 - \delta\|d_n-p\|^2.\end{aligned}$$

That is,

$$\|t_{n+1}-p\|^2 \leq \omega\|d_n-p\|^2 - \varepsilon\|t_{n+1}-d_n\|^2, \quad \forall n \geq N. \tag{3.31}$$

From (3.11), (3.12), and (3.31), we deduce that

$$\begin{aligned}\|t_{n+1}-p\|^2 \leq\ & \omega(1+\gamma)\|t_n-p\|^2 - \omega\gamma\|t_{n-1}-p\|^2 + \omega\gamma(1+\gamma)\|t_n-t_{n-1}\|^2\\ & -\varepsilon(1-\gamma)\|t_{n+1}-t_n\|^2 + \varepsilon\gamma(1-\gamma)\|t_n-t_{n-1}\|^2, \quad \forall n \geq N.\end{aligned}$$

This is equivalent to

$$\begin{aligned}&\|t_{n+1}-p\|^2 - \omega\gamma\|t_n-p\|^2 + \varepsilon(1-\gamma)\|t_{n+1}-t_n\|^2\\ &\quad\leq \omega\left[\|t_n-p\|^2 - \gamma\|t_{n-1}-p\|^2 + \varepsilon(1-\gamma)\|t_n-t_{n-1}\|^2\right]\\ &\qquad - \left(\omega\varepsilon(1-\gamma) - \omega\gamma(1+\gamma) - \varepsilon\gamma(1-\gamma)\right)\|t_n-t_{n-1}\|^2, \quad \forall n \geq N.\end{aligned}$$

We set

$$\Sigma_n := \|t_n-p\|^2 - \gamma\|t_{n-1}-p\|^2 + \varepsilon(1-\gamma)\|t_n-t_{n-1}\|^2.$$

Since $\omega \in (0,1)$, we have

$$\begin{aligned}\Sigma_{n+1} &\leq \|t_{n+1}-p\|^2 - \omega\gamma\|t_n-p\|^2 + \varepsilon(1-\gamma)\|t_{n+1}-t_n\|^2\\ &\leq \omega\Sigma_n - \left(\omega\varepsilon(1-\gamma) - \omega\gamma(1+\gamma) - \varepsilon\gamma(1-\gamma)\right)\|t_n-t_{n-1}\|^2, \quad \forall n \geq N.\end{aligned}$$

Thanks to $\gamma \in [0,1)$ and (3.24), we obtain

$$\omega\varepsilon(1-\gamma) - \omega\gamma(1+\gamma) - \varepsilon\gamma(1-\gamma) \geq \omega\varepsilon(1-\gamma) - 2\omega\gamma - \varepsilon\gamma \geq 0.$$

This implies that $\Sigma_{n+1} \leq \omega\Sigma_n$ for all $n \geq N$. Now, we show that $\Sigma_n > 0$ for all $n \geq N$. From (3.24), we have

$$\gamma \leq \frac{\omega\varepsilon}{\omega\varepsilon + 2\omega + \varepsilon} < \frac{\varepsilon}{2+\varepsilon}.$$

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

This yields

$$\frac{\varepsilon(1-\gamma)}{2} - \gamma > 0.$$

Using the definition of Σ_n, we deduce that

$$\begin{aligned}\Sigma_n &= (1-\varepsilon(1-\gamma))\|t_n - p\|^2 + \varepsilon(1-\gamma)\left(\|t_n - p\|^2 + \|t_n - t_{n-1}\|^2\right) - \gamma\|t_{n-1} - p\|^2 \\ &\geq (1-\varepsilon(1-\gamma))\|t_n - p\|^2 + \left(\frac{\varepsilon(1-\gamma)}{2} - \gamma\right)\|t_{n-1} - p\|^2 \\ &\geq (1-\varepsilon(1-\gamma))\|t_n - p\|^2 > 0.\end{aligned}$$

Hence, $\Sigma_{n+1} \leq \omega\Sigma_n \leq \cdots \leq \omega^{n-N+1}\Sigma_N$. That is,

$$\|t_n - p\|^2 \leq \frac{\Sigma_N}{(1-\varepsilon(1-\gamma))\,\omega^N}\omega^n.$$

This implies that $\{t_n\}$ converges R-linearly to p. This completes the proof. ∎

3.2 Strong convergence of Algorithm 3.2

In this subsection, we present a strongly convergent version of the FBF algorithm. More precisely, the iterative scheme is shown in Algorithm 3.2.

Algorithm 3.2

Initialization: Given $\psi_1 > 0$, $\gamma > 0$, $\eta \in (0,1)$, and $\sigma \in (0,1]$. Let $\{\varepsilon_n\}$ and $\{\delta_n\} \subset (0,1)$ satisfy (3.33). Let $t_0, t_1 \in \mathcal{H}$ and set $n := 1$.
Iterative Steps: Given the iterates t_n, t_{n-1}, perform the following steps.
Step 1. Compute $d_n = (1-\delta_n)(t_n + \gamma_n(t_n - t_{n-1}))$, where

$$\gamma_n = \begin{cases} \min\left\{\gamma, \dfrac{\varepsilon_n}{\|t_n - t_{n-1}\|}\right\}, & \text{if } t_n \neq t_{n-1}, \\ \gamma, & \text{otherwise.}\end{cases} \tag{3.32}$$

Step 2. Compute $g_n = (I + \psi_n B)^{-1}(I - \psi_n A)d_n$. If $g_n = d_n$ then stop and $g_n \in \mathrm{MIP}(A,B)$. Otherwise, go to *Step 3*.
Step 3. Compute $u_n = g_n - \psi_n(Ag_n - Ad_n)$.
Step 4. Compute $t_{n+1} = (1-\sigma)d_n + \sigma u_n$ and update ψ_{n+1} by (3.2).
Set $n := n+1$ and go to *Step 1.*

Theorem 3.6 (Strong convergence) *Assume that Conditions (C1), (C2), and (C4) hold. Then any sequence $\{t_n\}$ created by Algorithm 3.2 converges strongly to an element $p \in \mathrm{MIP}(A,B)$, where $p = P_{\mathrm{MIP}(A,B)}(0)$.*

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

(C4) *Take $\psi_1 > 0$, $\gamma > 0$, $\eta \in (0,1)$, and $\sigma \in (0,1]$. Let $\{\varepsilon_n\}$ and $\{\delta_n\} \subset (0,1)$ be two positive sequences that satisfy*

$$\lim_{n\to\infty} \delta_n = 0, \quad \sum_{n=1}^{\infty} \delta_n = \infty \quad \textit{and} \quad \lim_{n\to\infty} \frac{\varepsilon_n}{\delta_n} = 0. \tag{3.33}$$

Proof Using similar arguments as (3.3)–(3.8) in the proof of Theorem 3.4, we have

$$\|t_{n+1} - p\|^2 \le \|d_n - p\|^2 - \sigma\left(1 - \eta^2 \frac{\psi_n^2}{\psi_{n+1}^2}\right)\|d_n - g_n\|^2, \quad \forall p \in \mathrm{MIP}(A,B). \tag{3.34}$$

By the definition of d_n, one has

$$\begin{aligned} \|d_n - p\| &= \|(1-\delta_n)(t_n + \gamma_n(t_n - t_{n-1})) - p\| \\ &= \|(1-\delta_n)(t_n - p) + (1-\delta_n)\gamma_n(t_n - t_{n-1}) - \delta_n p\| \\ &\le (1-\delta_n)\|t_n - p\| + (1-\delta_n)\gamma_n\|t_n - t_{n-1}\| + \delta_n\|p\| \\ &= (1-\delta_n)\|t_n - p\| + \delta_n\left[(1-\delta_n)\frac{\gamma_n}{\delta_n}\|t_n - t_{n-1}\| + \|p\|\right]. \end{aligned} \tag{3.35}$$

From (3.32) and (3.33), we have

$$\frac{\gamma_n}{\delta_n}\|t_n - t_{n-1}\| \le \frac{\varepsilon_n}{\delta_n} \to 0 \text{ as } n \to \infty.$$

This follows that $\lim_{n\to\infty}\left[(1-\delta_n)\frac{\gamma_n}{\delta_n}\|t_n - t_{n-1}\| + \|p\|\right] = \|p\|$; thus, there exists $M > 0$ such that

$$(1-\delta_n)\frac{\gamma_n}{\delta_n}\|t_n - t_{n-1}\| + \|p\| \le M, \quad \forall n \ge 1. \tag{3.36}$$

Using (3.35) and (3.36), one gives that

$$\|d_n - p\| \le (1-\delta_n)\|t_n - p\| + \delta_n M. \tag{3.37}$$

Since $\lim_{n\to\infty}\left(1 - \eta^2\frac{\psi_n^2}{\psi_{n+1}^2}\right) = 1 - \eta^2 > 0$, there exists $n_0 \in \mathbb{N}$ such that $1 - \eta^2\frac{\psi_n^2}{\psi_{n+1}^2} > 0$, $\forall n \ge n_0$. This together with (3.34) yields that

$$\|t_{n+1} - p\| \le \|d_n - p\|, \quad \forall n \ge n_0. \tag{3.38}$$

By using (3.37) and (3.38), we obtain

$$\begin{aligned} \|t_{n+1} - p\| &\le (1-\delta_n)\|t_n - p\| + \delta_n M \\ &\le \max\{\|t_n - p\|, M\} \le \cdots \le \max\{\|t_{n_0} - p\|, M\}. \end{aligned}$$

This means that the sequence $\{t_n\}$ is bounded. Thus, the sequences $\{d_n\}$, $\{g_n\}$, and $\{u_n\}$ are also bounded by means of (3.37), (3.34), and (3.6).

From (3.37) and note that $\{\delta_n\} \subset (0,1)$, one finds

$$\begin{aligned} \|d_n - p\|^2 &\le (1-\delta_n)^2 \|t_n - p\|^2 + 2\delta_n (1-\delta_n) M \|t_n - p\| + \delta_n^2 M^2 \\ &\le \|t_n - p\|^2 + \delta_n M_1, \end{aligned} \tag{3.39}$$

where $M_1 := \max\left\{2(1-\delta_n) M \|t_n - p\| + \delta_n M^2 : n \in \mathbb{N}\right\}$. Substituting (3.39) into (3.34), we have

$$\sigma\left(1 - \eta^2 \frac{\psi_n^2}{\psi_{n+1}^2}\right) \|d_n - g_n\|^2 \le \|t_n - p\|^2 - \|t_{n+1} - p\|^2 + \delta_n M_1. \tag{3.40}$$

Using the inequality (3.38) and the definition of d_n, we obtain

$$\begin{aligned} \|t_{n+1} - p\|^2 &\le \|(1-\delta_n)(t_n - p) + (1-\delta_n)\gamma_n (t_n - t_{n-1}) - \delta_n p\|^2 \\ &\le \|(1-\delta_n)(t_n - p) + (1-\delta_n)\gamma_n (t_n - t_{n-1})\|^2 + 2\delta_n \langle -p, d_n - p\rangle \\ &\le (1-\delta_n)^2 \|t_n - p\|^2 + 2(1-\delta_n)\gamma_n \|t_n - p\| \|t_n - t_{n-1}\| \\ &\quad + \gamma_n^2 \|t_n - t_{n-1}\|^2 + 2\delta_n \langle -p, d_n - t_{n+1}\rangle + 2\delta_n \langle -p, t_{n+1} - p\rangle . \end{aligned}$$

This yields that

$$\begin{aligned} &\|t_{n+1} - p\|^2 \\ &\le (1-\delta_n)\|t_n - p\|^2 + \delta_n \Big[2(1-\delta_n)\|t_n - p\| \frac{\gamma_n}{\delta_n} \|t_n - t_{n-1}\| \\ &\quad + \gamma_n \|t_n - t_{n-1}\| \frac{\gamma_n}{\delta_n} \|t_n - t_{n-1}\| + 2\|p\| \|d_n - t_{n+1}\| + 2\langle p, p - t_{n+1}\rangle \Big]. \end{aligned} \tag{3.41}$$

Next, we show that $\{\|t_n - p\|\}$ converges to zero. Indeed, by Lemma 2.5, it suffices to prove that $\limsup_{k\to\infty} \|d_{n_k} - t_{n_k+1}\| = 0$ and $\limsup_{k\to\infty} \langle p, p - t_{n_k+1}\rangle \le 0$ for every subsequence $\{\|t_{n_k} - p\|\}$ of $\{\|t_n - p\|\}$ satisfying

$$\liminf_{k\to\infty} \left(\|t_{n_k+1} - p\| - \|t_{n_k} - p\|\right) \ge 0. \tag{3.42}$$

For this, assume that $\{\|t_{n_k} - p\|\}$ is a subsequence of $\{\|t_n - p\|\}$ such that (3.42) holds. Then

$$\begin{aligned} &\liminf_{k\to\infty} \left(\|t_{n_k+1} - p\|^2 - \|t_{n_k} - p\|^2\right) \\ &\quad = \liminf_{k\to\infty} \left[\left(\|t_{n_k+1} - p\| - \|t_{n_k} - p\|\right)\left(\|t_{n_k+1} - p\| + \|t_{n_k} - p\|\right)\right] \ge 0. \end{aligned}$$

By (3.40) and $\lim_{n\to\infty} \delta_n = 0$, we obtain

$$\begin{aligned} &\limsup_{k\to\infty} \sigma\left(1 - \eta^2 \frac{\psi_{n_k}^2}{\psi_{n_k+1}^2}\right) \|d_{n_k} - g_{n_k}\|^2 \\ &\quad \le \limsup_{k\to\infty} \left[\|t_{n_k} - p\|^2 - \|t_{n_k+1} - p\|^2 + \delta_{n_k} M_1\right] \end{aligned}$$

$$
\begin{aligned}
&\le \limsup_{k\to\infty}\left[\|t_{n_k}-p\|^2-\|t_{n_k+1}-p\|^2\right]+\limsup_{k\to\infty}\delta_{n_k}M_1\\
&=-\liminf_{k\to\infty}\left[\|t_{n_k+1}-p\|^2-\|t_{n_k}-p\|^2\right]\\
&\le 0.
\end{aligned}
$$

This implies that

$$
\lim_{k\to\infty}\|g_{n_k}-d_{n_k}\|=0. \tag{3.43}
$$

From the definition of u_n and (3.2), we obtain

$$
\|u_{n_k}-g_{n_k}\|\le \eta\frac{\psi_{n_k}}{\psi_{n_k+1}}\|d_{n_k}-g_{n_k}\|.
$$

This combined with (3.43) yields that

$$
\lim_{k\to\infty}\|u_{n_k}-g_{n_k}\|=0. \tag{3.44}
$$

By the definition of t_{n+1}, (3.43), and (3.44), we have

$$
\lim_{k\to\infty}\|t_{n_k+1}-d_{n_k}\|=\lim_{k\to\infty}\sigma\|u_{n_k}-d_{n_k}\|=0. \tag{3.45}
$$

It follows from the definition of d_n that

$$
\begin{aligned}
\|t_{n_k}-d_{n_k}\|&=\|(1-\delta_{n_k})\gamma_{n_k}(t_{n_k}-t_{n_k-1})-\delta_{n_k}t_{n_k}\|\\
&\le \delta_{n_k}\left[(1-\delta_{n_k})\frac{\gamma_{n_k}}{\delta_{n_k}}\|t_{n_k}-t_{n_k-1}\|+\|t_{n_k}\|\right].
\end{aligned}
$$

Therefore, we deduce that

$$
\lim_{k\to\infty}\|t_{n_k}-d_{n_k}\|=0. \tag{3.46}
$$

From (3.45) and (3.46), we obtain

$$
\lim_{k\to\infty}\|t_{n_k+1}-t_{n_k}\|=0. \tag{3.47}
$$

Since $\{t_{n_k}\}$ is bounded, there exists a subsequence $\{t_{n_{k_j}}\}$ of $\{t_{n_k}\}$ such that $\{t_{n_{k_j}}\}$ converges weakly to z^* as $j\to\infty$. Thanks to (3.46), one has $d_{n_k}\rightharpoonup z^*$. This combined with (3.43) and Lemma 3.3 implies that $z^*\in \mathrm{MIP}(A,B)$. By using the property of projection (cf. (2.1)) and the definition of $p=P_{\mathrm{MIP}(A,B)}(0)$, we have

$$
\limsup_{k\to\infty}\langle p,p-t_{n_k}\rangle=\lim_{j\to\infty}\langle p,p-t_{n_{k_j}}\rangle=\langle p,p-z^*\rangle\le 0. \tag{3.48}
$$

From (3.47) and (3.48), one gives

$$
\limsup_{k\to\infty}\langle p,p-t_{n_k+1}\rangle\le 0. \tag{3.49}
$$

This together with (3.41), (3.45), $\lim_{n\to\infty}\frac{\gamma_n}{\delta_n}\|t_n-t_{n-1}\|=0$, and Lemma 2.5 yields that $\lim_{n\to\infty}\|t_n-p\|=0$. That is, $t_n\to p$ as $n\to\infty$. ∎

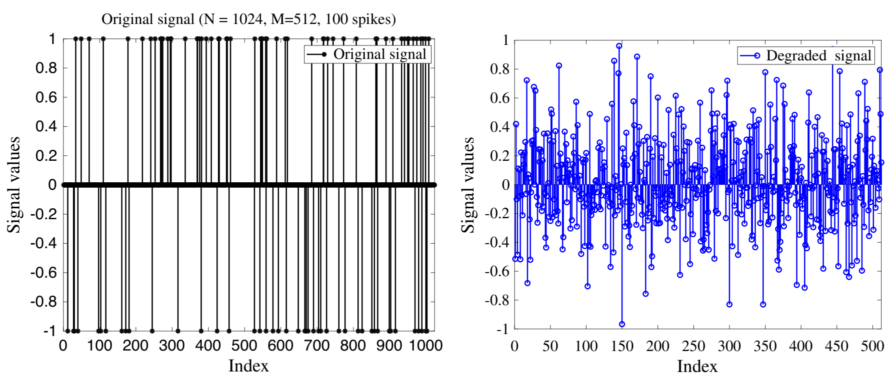
4 Numerical experiments

In this section, we apply the suggested Algorithms 3.1 and 3.2 to solve the signal recovery problem described in Example 1.1 and compare them with the ones in the literature [17, 18]. All codes were executed in MATLAB 2018a on a personal computer with RAM 8.00 GB.

Example 4.1 (Signal Recovery Problem) In our numerical experiments, the origin signal $\mathbf{x}$ is a vector of N values, where N is significantly greater than k, the number of randomly created spikes with either +1 or −1. The matrix $\mathbf{C}$ with dimensions $M \times N$ is created using a standard normal distribution with a zero mean and unit variance, and its rows are then orthonormalized. We introduce white Gaussian noise with a variance of 10^{-4}, represented by ε, to generate the observation $\mathbf{y}$ using equation $\mathbf{y} = \mathbf{Cx} + \varepsilon$. The recovery procedure starts with the initial signals $\mathbf{t_0}$ and $\mathbf{t_1}$ set to zero and terminates after 100 iterations. The accuracy of the restoration is measured utilizing the mean squared error (MSE), which is calculated as MSE $= \frac{1}{N} \|\mathbf{x}^* - \mathbf{x}\|^2$, where $\mathbf{x}^*$ is an estimated signal of the origin $\mathbf{x}$ and N denotes the dimension of $\mathbf{x}$. The parameters of our algorithms and the compared ones are set as follows.

- Set $\psi_1 = 1$, $\eta = 0.5$, $\xi_n = 1 + \frac{1}{(n+1)^2}$, and $\tau_n = \frac{1}{n+1}$ for the proposed Algorithms 3.1 and 3.2. Choose $\gamma = 0.2$ and $\sigma = 0.9$ for our Algorithm 3.1. Take $\gamma = 0.5$, $\sigma = 0.9$, $\delta_n = \frac{1}{100(n+1)}$, and $\varepsilon_n = \frac{1}{(n+1)^2}$ for our Algorithm 3.2.
- Pick $\psi_1 = 1$, $\eta = 0.5$, $\alpha_n = \frac{1}{n+1}$, and $\delta_n = 0.5(1 - \alpha_n)$ for the Algorithm 1 proposed by Gibali and Thong [17] (shortly, GT Alg. 1). Set $\psi_1 = 1$, $\eta = 0.5$, $\alpha_n = \frac{1}{n+1}$, and $f(x) = 0.1x$ for the Algorithm 2 suggested by Gibali and Thong [17] (shortly, GT Alg. 2) Take $\psi_1 = 1$, $\eta = 0.5$, $\gamma = 0.2$, $\sigma = 0.4$, and $\tau_n = \frac{1}{n+1}$ for the Algorithm 1 introduced by Thong et al. [18] (shortly, TCPDL Alg. 1).

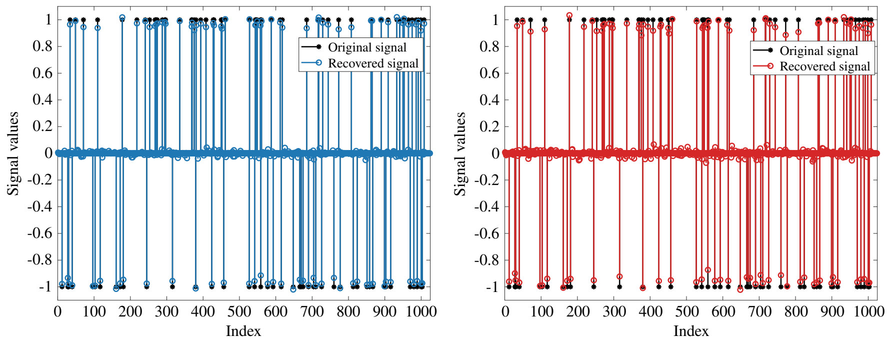
In first test, we set $M = 512$, $N = 1024$, and $k = 100$. The original and noisy signals are described in Figure 2, and the recovery results employing the proposed algorithms are shown in Figure 3. The variation of MSE with the number of iterations for all



(a) Original signal (b) Degraded signal

Figure 2: Original signal and degraded signal.

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press



(a) Recovered by our Alg. 3.1

(b) Recovered by our Alg. 3.2

Figure 3: The original signal and the signal recovered by our algorithms.

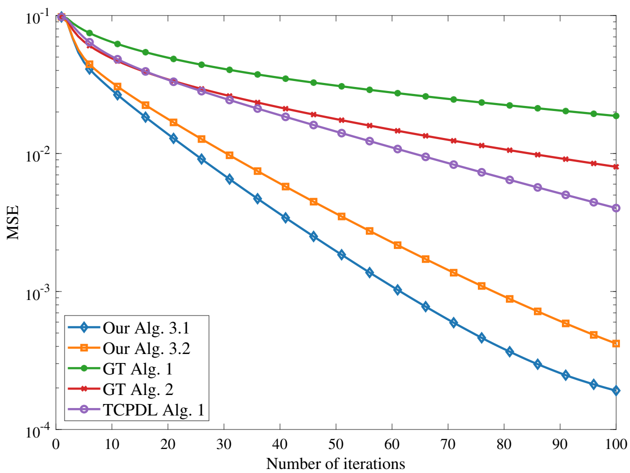


Figure 4: The variation of MSE with the number of iterations for all algorithms at $M = 512$, $N = 1024$, $k = 100$.

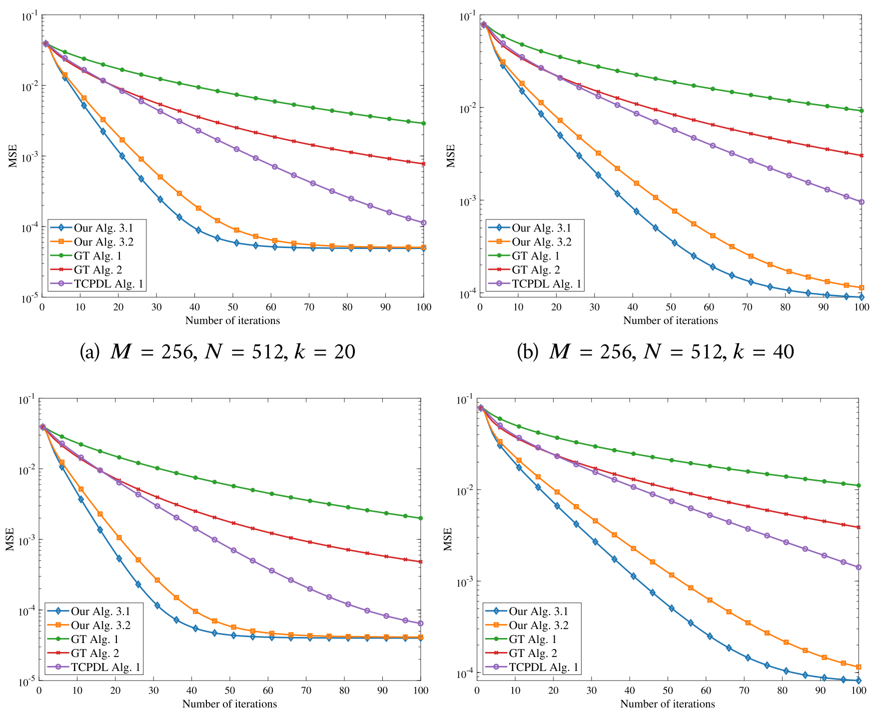
algorithms is stated in Figure 4. Finally, we applied all algorithms to solve the signal recovery problem under different dimensions and sparsity. The results are presented in Table 1 and Figure 5.

Remark 4.2 From Figures 2, 3, 4, and 5, and Table 1, it can be seen that the two algorithms proposed in this paper can effectively recover the original signal. Furthermore, the proposed algorithms perform better in terms of recovery as the original signal becomes sparser (i.e., when $k \ll N$). However, under the same dimensions and sparsity, the proposed algorithms converge faster and perform better than the algorithms in references [17, 18].

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

Algorithms	$M = 256, N = 512$ $k = 20$		$M = 256, N = 512$ $k = 40$		$M = 512, N = 1024$ $k = 40$		$M = 512, N = 1024$ $k = 80$	
	MSE ($\times 10^{-4}$)	Time (s)	MSE ($\times 10^{-3}$)	Time (s)	MSE ($\times 10^{-4}$)	Time (s)	MSE ($\times 10^{-3}$)	Time (s)
Our Alg. 3.1	0.4924	0.0138	0.0894	0.0274	0.4003	0.0377	0.0813	0.0375
Our Alg. 3.2	0.5073	0.0164	0.1123	0.0247	0.4126	0.0392	0.1127	0.0417
GT Alg. 1 [17]	28.4918	0.0310	9.0969	0.0286	19.6359	0.0401	10.9875	0.0451
GT Alg. 2 [17]	7.5977	0.0155	2.9808	0.0273	4.7326	0.0472	3.8092	0.0459
TCPDL Alg. 1 [18]	1.0968	0.0132	0.9244	0.0244	0.6320	0.0347	1.3719	0.0436

Table 1: The numerical results of all algorithms under different situations.



(a) $M = 256, N = 512, k = 20$ (b) $M = 256, N = 512, k = 40$

(c) $M = 512, N = 1024, k = 40$ (d) $M = 512, N = 1024, k = 80$

Figure 5: The variation of MSE with the number of iterations for all algorithms in four cases.

https://doi.org/10.4153/S0008414X24000889 Published online by Cambridge University Press

5 Conclusions

The paper presents two novel algorithms that solve the monotone inclusion problem by incorporating improved forward-backward-forward methods with considerations of both the inertial term and the relaxation effect. Our results demonstrate that the proposed algorithms converge weakly and strongly, respectively, in real Hilbert spaces under suitable conditions. Furthermore, the first proposed algorithm is proven to have R-linear convergence when the multi-valued operator is strongly monotone. The suggested algorithms have been applied to signal recovery problems and have shown improved performance compared to previously published methods. These findings make a significant contribution to the field and open up new opportunities for further research.

Acknowledgements The authors are very grateful to the editor and referees for their suggestions to help us improve the initial manuscript.

References

[1] D. A. T. Lorenz Pock, *An inertial forward-backward algorithm for monotone inclusions*. J. Math. Imaging Vision **51**(2015), 311–325.
[2] R. I. Boţ and E. R. Csetnek, *An inertial forward-backward-forward primal-dual splitting algorithm for solving monotone inclusion problems*. Numer. Algorithms **71**(2016), 519–540.
[3] L.O. Jolaoso, Y. Shehu, J. C. Yao and R. Xu, *Double inertial parameters forward-backward splitting method: Applications to compressed sensing, image processing, and SCAD penalty problems*. J. Nonlinear Var. Anal. 7(2023), 627–646.
[4] C. Izuchukwu, S. Reich, Y. Shehu and A. Taiwo, *Strong convergence of forward-reflected-backward splitting methods for solving monotone inclusions with applications to image restoration and optimal control*. J. Sci. Comput. **94**(2023), article no. 73.
[5] P. L. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*. SIAM J. Numer. Anal. **16(1979)**, 964–979.
[6] G. B. Passty, *Ergodic convergence to a zero of the sum of monotone operators in Hilbert space*. J. Math. Anal. Appl. **72**(1979), 383–390.
[7] P. Tseng, *A modified forward-backward splitting method for maximal monotone mappings*. SIAM J. Control Optim. **38**(2000), 431–446.
[8] L. C. Ceng and Q. Yuan, *Composite inertial subgradient extragradient methods for variational inequalities and fixed point problems*. J. Inequal. Appl. **2019**(2019), 274.
[9] W. Cholamjiak, P. Cholamjiak and S. Suantai, *An inertial forward-backward splitting method for solving inclusion problems in Hilbert spaces*. J. Fixed Point Theory Appl. **20**(2018), article no. 42.
[10] Y. Shehu, *Convergence results of forward-backward algorithms for sum of monotone operators in Banach spaces*. Results Math. **74**(2019), article no. 138.
[11] Y. Malitsky and M. K. Tam, *A forward-backward splitting method for monotone inclusions without cocoercivity*. SIAM J. Optim. **30**(2020), 1451–1472.
[12] L. C. Ceng and Q. Yuan, *Strong convergence of a new iterative algorithm for split monotone variational inclusion problems*. Mathematics 7(2019), 123.
[13] P. Cholamjiak, D. V. Hieu and Y. J. Cho, *Relaxed forward-backward splitting methods for solving variational inclusions and applications*. J. Sci. Comput. **88**(2021), article no. 85.
[14] O. S. Iyiola, C. D. Enyi and Y. Shehu, *Reflected three-operator splitting method for monotone inclusion problem*. Optim. Methods Softw. **37**(2022), 1527–1565.
[15] C. Izuchukwu, S. Reich and Y. Shehu, *Convergence of two simple methods for solving monotone inclusion problems in reflexive Banach spaces*. Results Math. 77(2022), article no. 143.
[16] A. Taiwo and S. Reich, *Bounded perturbation resilience of a regularized forward-reflected-backward splitting method for solving variational inclusion problems with applications*. Optimization 73(2024), 2089–2122.
[17] A. Gibali and D. V. Thong, *Tseng type methods for solving inclusion problems and its applications*. Calcolo **55**(2018), article no. 49.

[18] D. V. Thong, P. Cholamjiak, N. Pholasa, V. T. Dung and L. V. Long, *A new modified forward-backward-forward algorithm for solving inclusion problems*. Comput. Appl. Math. **41**(2022), article no. 405.
[19] H. H. Bauschke and P. L. Combettes, Convex analysis and monotone Operator Theory in Hilbert spaces, second edition, Springer, Berlin, 2017.
[20] K. Goebel and S. Reich, Uniform convexity, hyperbolic geometry, and nonexpansive mappings, Marcel Dekker, New York and Basel, 1984.
[21] J. M. Ortega and W. C. Rheinboldt, Iterative solution of nonlinear equations in several variables, Academic Press, New York, 1970.
[22] H. Brézis, Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert, North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973.
[23] F. Alvarez and H. Attouch, *An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping*. Set-Valued Anal. **9**(2001), 3–11.
[24] Z. Opial, *Weak convergence of the sequence of successive approximations for nonexpansive mappings*. Bull. Amer. Math. Soc. 73(1967), 591–597.
[25] S. Saejung and P. Yotkaew, *Approximation of zeros of inverse strongly monotone operators in Banach spaces*. Nonlinear Anal. **75**(2012), 742–750.
[26] B. Tan, A. Petruşel, X. Qin and J. C. Yao, *Global and linear convergence of alternated inertial single projection algorithms for pseudo-monotone variational inequalities*. Fixed Point Theory **23**(2022), 391–426.
[27] B. Tan and S. Y. Cho, *Strong convergence of inertial forward-backward methods for solving monotone inclusions*. Appl. Anal. **101**(2022), 5386–5414.

School of Mathematics and Statistics, Southwest University, Chongqing, China
e-mail: bingtan@swu.edu.cn bingtan72@gmail.com
URL: https://bingtan.me/

Department of Mathematics, Hangzhou Normal University, Hangzhou, China;
Nanjing Center for Applied Mathematics, Nanjing, China
e-mail: qxlxajh@163.com