J. Appl. Numer. Optim. 6 (2024), No. 1, pp. 97-113
Available online at http://jano.biemdas.com
https://doi.org/10.23952/jano.6.2024.1.06

AN ACCELERATED REGULARIZATION METHOD FOR VARIATIONAL INCLUSIONS

NGUYEN HAI HA[1], DANG VAN HIEU[2,*], LE DUNG MUU[3]

[1] *Faculty of Basic Sciences, University of Transport and Communications, Hanoi, Vietnam*
[2] *Department of Mathematics, University of Air Force, Nha Trang, Vietnam*
[3] *TIMAS - Thang Long University, Hanoi, Vietnam*

Abstract. The paper proposes an iterative method for solving a variational inclusion with the sum of two operators in a Hilbert space. The method can be considered as a combination of the proximal contraction method, the regularization method, and the multi-step inertial technique. Theorem of strong convergence is established under mild conditions imposed on cost operators and control parameters.

Keywords. Multi-step inertial method; Proximal contraction method; Regularization; Variational inclusion.

2020 Mathematics Subject Classification. 65Y05, 65K15, 90C30.

1. INTRODUCTION

The main purpose of the paper is to introduce a numerical approach for finding a solution of the following variational inclusion (VI) in a real Hilbert space $\mathscr{H}$, namely

$$\text{Find } u^* \in \mathscr{H} \text{ such that } 0 \in \mathscr{A}u^* + fu^*, \tag{1.1}$$

where $\mathscr{A} : \mathscr{H} \to 2^{\mathscr{H}}$ is a maximally monotone multi-valued operator and $f : \mathscr{H} \to \mathscr{H}$ is a L-Lipschitz continuous, monotone, and single-valued operator. Throughout this paper, we denote $\Omega = (\mathscr{A} + f)^{-1}(0)$ the solution set of the VI and assume that it is nonempty.

The VI (for the case $f = 0$) was early studied by Rockafellar [1] with the celebrated proximal point algorithm. This problem plays a central role in optimization field as well as nonlinear analysis. It involves some known problems such as variational inequality problems, fixed point problems, operator equations, and equilibrium problems [1, 2], and it has received a lot of attention from mathematicians because of its broad application in applied sciences such as image recovery, deep learning, and data analysis; see, e.g., [3, 4, 5, 6, 7].

Consider the following optimization problem (OP),

$$\min_{u \in \mathscr{H}} (\Gamma(u) + \Theta(u)),$$

where $\Gamma : \mathscr{H} \to \mathbb{R}$ is a subdifferentiable convex function with the subdifferential $\partial\Gamma$, and $\Theta : \mathscr{H} \to \mathbb{R}$ is a differentiable convex function with the gradient $\nabla\Theta$. The problem OP can be reformulated under

*Corresponding author.
E-mail address: hanguyenhai@utc.edu.vn (N.H. Ha), dv.hieu83@gmail.com (D.V. Hieu), ldmuu@math.ac.vn (L.D. Muu).

Received 20 December 2022; Accepted 1 September 2023; Published online 5 March 2024.

the problem VI with $\mathscr{A} = \partial\Gamma$ and $f = \nabla\Theta$. Another example is for the constraint optimization problem (COP),

$$\min_{u\in\mathscr{C}} \Theta(u),$$

where $\mathscr{C}$ is a nonempty, closed, and convex subset of $\mathscr{H}$ and, $\Theta : \mathscr{H} \to \mathbb{R}$ is a function which is convex on $\mathscr{C}$ and differentiable on a neighborhood of $\mathscr{C}$. In this case, the COP is a special case of the VI with $\mathscr{A} = N_{\mathscr{C}}$ and $f = \nabla\Theta$ (the detail of this claim is analyzed in Section 5), where $N_{\mathscr{C}}$ is the normal cone of $\mathscr{C}$, given by

$$N_{\mathscr{C}}(u) = \{w \in \mathscr{H} : \langle w, v-u\rangle \le 0, \forall v \in \mathscr{C}\}.$$

Some notable methods were proposed for solving problem VI (1.1) such as the forward-backward splitting method [6, 8], the modified forward-backward method [7], the proximal contraction method [9], the forward-reflected-backward splitting method [10], and the others [5, 11, 12, 13]. In Hilbert spaces, without such an additional condition, these methods in general only provide the weak convergence, while the strong convergence is more desirable, especially in infinite dimensional spaces. In order to get this aim, the methods are often combined with one or more techniques as the viscosity method, the Halpern method, the hybrid (shrinking) projection method or the regularization method.

In 2018, Zhang and Wang [9] proposed the following proximal contraction methods (PCM) for solving problem VI (1.1):

$$\begin{cases} v_n = J^{\mathscr{A}}_{\lambda_n}(u_n - \lambda_n f(u_n)), \\ d(u_n, v_n) = u_n - v_n - \lambda_n(f(u_n) - f(v_n)), \\ u_{n+1} = u_n - \gamma\beta_n d(u_n, v_n), \end{cases} \tag{1.2}$$

where $\gamma \in (0,2)$,

$$\beta_n = \frac{\phi(u_n, v_n)}{||d(u_n, v_n)||^2},$$

$$\phi(u_n, v_n) = \langle u_n - v_n, d(u_n, v_n)\rangle,$$

$\{\lambda_n\}$ satisfies prediction step-size conditions and $\liminf_{n\to\infty} \lambda_n \ge \underline{\lambda} > 0$, and $J^{\mathscr{A}}_{\lambda_n} = (I + \lambda_n\mathscr{A})^{-1}$ is the resolvent of $\mathscr{A}$ associated with the parameter λ_n. Zhang and Wang [9] proved the weak convergence of their proposed method.

Let $\mathscr{F} : \mathscr{H} \to \mathscr{H}$ be a strongly monotone and Lipschitz continuous operator. In 2021, the authors in [14] introduced the regularization proximal contraction method (RPCM) for solving problem VI (1.1) in Hilbert spaces:

$$\begin{cases} v_n = J^{\mathscr{A}}_{\lambda_n}(u_n - \lambda_n(f(u_n) + \alpha_n\mathscr{F}(u_n))), \\ u_{n+1} = u_n - r\beta_n d(u_n, v_n), \end{cases} \tag{1.3}$$

where $r \in (0,2)$ and

$$\beta_n = \min\left\{\beta, \frac{\phi(u_n, v_n)}{||d(u_n, v_n)||^2}\right\},$$

where $\beta > 0$. Method RPCM (1.3) is a combination between method PCM (1.2) and the regularization technique [14, Lemma 3]. Thanks to the incorporated regularization, the method RPCM provides the strong convergence [14, Theorem 1].

One notable aspect, which draws our attention to the acceleration of the original method, is the role of inertial-type technique. Alvarez and Attouch in [15] pointed out that it is from the so-called "heavy ball with friction" second order dynamical system. The method was intensively and widely utilized for the purpose of speeding up the convergence rate of various iterative algorithms [15, 16, 17, 18].

The main purpose of this paper is to introduce a new numerical method for solving problem VI (1.1) in Hilbert spaces. The method is developed from method RPCM (1.3) incorporated with the multi-step inertial effect. More precisely, let N be a given natural number and the previous iterates $u_0, u_1, \cdots, u_n$ be known. We use an intermediate combination

$$w_n = u_n + \sum_{i=1}^{\min\{n,N\}} \theta_{i,n}(u_{n-i+1} - u_{n-i}),$$

to find the next iterate u_{n+1} based on the computations of method RPCM (1.3), where the parameter $\theta_{i,n} \geq 0$ is suitably chosen. The next iterate u_{n+1} is computed from the information of N previous iterates. In the case when $\theta_{i,n} = 0$, the new method is reduced to the original method RPCM (1.3). Some one-step inertial methods ($N = 1$) can be found in [15, 16, 17] and multi-step inertial methods ($N > 1$) are in [19, 20, 21]. Almost these methods provide the weak convergence and/or for the special cases of problem VI (1.1). In this paper, we establish the strong convergence of the new method under mild conditions imposed on cost operators and control parameters.

The paper is organized as follows: Section 2 supplies basic notions and lemmas used in latter parts. Section 3 presents some results regarding the regularization technique. Section 4 introduces our main algorithm and proves the strong convergence of the method. Section 5 discusses the application of our method to solve a couple of optimization problems.

2. Preliminaries

We take the following concepts in a real Hilbert space $\mathscr{H}$. Let $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ denote the inner product of $\mathscr{H}$ and the norm generated by $\langle \cdot, \cdot \rangle$, respectively. In any Hilbert space, we have the equality

$$2\langle u, v\rangle = ||u||^2 + ||v||^2 - ||u-v||^2, \ \forall u, v \in \mathscr{H},$$

and the inequality

$$2||u||\,||v|| \leq \nu ||u||^2 + \frac{1}{\nu}||v||^2, \ \forall \nu > 0, \ \forall u, v \in \mathscr{H}.$$

Let $u_n \to x$ stand for the fact that u_n is convergent in norm (or strongly converges) to x while $u_n \rightharpoonup x$ means that u_n converges weakly to x as $n \to \infty$.

An multi-valued operator $\mathscr{A} : \mathscr{H} \to 2^{\mathscr{H}}$ is called:

(i) *monotone* if $\langle u - v, x - y\rangle \geq 0$ for all $x, y \in \mathscr{H}$ and $u \in \mathscr{A}x$, $v \in \mathscr{A}y$;

(ii) *strongly monotone* if there exists a number $\gamma > 0$ such that $\langle u - v, x - y\rangle \geq \gamma||x-y||^2$ for all $x, y \in \mathscr{H}$ and $u \in \mathscr{A}x$, $v \in \mathscr{A}y$;

(iii) *inverse strongly monotone* if there exists $c > 0$ such that $\langle u - v, x - y\rangle \geq c||u-v||^2$ for all $x, y \in \mathscr{H}$ and $u \in \mathscr{A}x$, $v \in \mathscr{A}y$. If $c = 1$, then $\mathscr{A}$ is called *firmly-nonexpansive*.

An operator $f : \mathscr{H} \to \mathscr{H}$ is called *Lipschitz continuous* if there exists a number $L > 0$ such that $||f(u) - f(v)|| \leq L||u - v||$ for all $u, v \in \mathscr{H}$. f is called *nonexpansive* if $L = 1$, and *contractive* if $0 \leq L < 1$.

The graph of an operator $\mathscr{A}:\mathscr{H}\to 2^{\mathscr{H}}$ is defined by

$$Graph(\mathscr{A})=\{(x,u)\in\mathscr{H}\times\mathscr{H}: u\in\mathscr{A}x\}.$$

A multi-valued operator $\mathscr{A}:\mathscr{H}\to 2^{\mathscr{H}}$ is called *maximally monotone* if it is monotone and its graph is not properly contained in the graph of any other monotone operator. It follows from the maximal monotonicity of $\mathscr{A}$ that, for each pair $(x,u)\in\mathscr{H}\times\mathscr{H}$, $\langle x-y,u-v\rangle\geq 0$ for all $(y,v)\in Graph(\mathscr{A})$, one has $v\in\mathscr{A}y$. Let $\mathscr{A}:\mathscr{H}\to 2^{\mathscr{H}}$ be maximally monotone. The resolvent of $\mathscr{A}$ associated with a number $\lambda>0$ is defined by

$$J_\lambda^{\mathscr{A}}(u)=(I+\lambda\mathscr{A})^{-1}(u),\ u\in\mathscr{H}.$$

The resolvent operator $J_\lambda^{\mathscr{A}}$ is single-valued, nonexpansive, and firmly-nonexpansive.

Lemma 2.1. [22, Lemma 2.4] *Let $\mathscr{A}:\mathscr{H}\to 2^{\mathscr{H}}$ be a maximally monotone operator, and let $f:\mathscr{H}\to\mathscr{H}$ be a monotone Lipschitz continuous operator. Then, $\mathscr{B}=\mathscr{A}+f$ is maximally monotone.*

Lemma 2.2. [23, Sect. 4] *Let $\mathscr{A}:\mathscr{H}\to 2^{\mathscr{H}}$ be a maximally monotone operator, and let $f:\mathscr{H}\to\mathscr{H}$ be an operator. For each $\lambda>0$, define the mapping*

$$\mathscr{T}_\lambda(u):=J_\lambda^{\mathscr{A}}(u-\lambda f(u))$$

for all $u\in\mathscr{H}$. Then

$$u^*\in(\mathscr{A}+f)^{-1}(0)\Leftrightarrow u^*\in Fix(\mathscr{T}_\lambda),$$

where $Fix(\mathscr{T}_\lambda)$ is the fixed point set of $\mathscr{T}_\lambda$.

For solving problem VI (1.1), we consider the following assumptions:
(A1) $\mathscr{A}$ is maximally monotone;
(A2) f is monotone and Lipschitz continuous;
(A3) The solution set $\Omega:=(\mathscr{A}+f)^{-1}(0)$ of problem VI (1.1) is nonempty.

Let $\mathscr{F}:\mathscr{H}\to\mathscr{H}$ be a γ - strongly monotone and L - Lipschitz continuous operator. In order to solve problem VI (1.1), i.e., to select an element in Ω, we are interested in solving the following variational inequality (VIP):

$$\text{Find } u^\dagger\in\Omega \text{ such that } \left\langle \mathscr{F}u^\dagger,u^*-u^\dagger\right\rangle\geq 0,\ \forall u^*\in\Omega. \tag{2.1}$$

Thanks to the given properties of $\mathscr{F}$ and the fact that Ω is convex, closed, and nonempty (assumed), problem VIP (2.1) has a unique solution $u^\dagger$. The iterative sequence $\{u_n\}$ generated by our proposed method is proved to be convergent to the solution $u^\dagger$ of problem VIP (2.1). If $\mathscr{F}u=u$, then $u^\dagger$ is the smallest norm solution of problem VI (1.1), while if $\mathscr{F}u=u-u^g$, where u^g is a suggested point in $\mathscr{H}$, the solution $u^\dagger$ of (2.1) is $u^\dagger=P_\Omega(u^g)$, the point in Ω is the nearest to u^g. Considering the operator $\mathscr{F}$ as in problem (2.1) helps us to find a solution of problem VI with a desired property. Problem (2.1) can be considered as a bilevel problem that the constraint is the solution set to problem VI (1.1).

Finally, we need the following technical lemma.

Lemma 2.3. [24] *Let $\{\Psi_n\}$ be a sequence of nonnegative real numbers. Suppose that*

$$\Psi_{n+1}\leq(1-p_n)\Psi_n+q_n$$

for all $n\geq 0$, where the sequences $\{p_n\}$ in $(0,1)$ and $\{q_n\}$ in $\Re$ satisfy the conditions:

(i) $\lim_{n\to\infty} p_n = 0$,
(ii) $\sum_{n=1}^{\infty} p_n = \infty$, *and*
(iii) $\limsup_{n\to\infty} \frac{q_n}{p_n} \leq 0$.
Then $\lim_{n\to\infty} \Psi_n = 0$.

3. REGULARIZATION

In this section, we recall the Tikhonov-type regularization to the class of monotone variational inclusion problem [14]. For each $\alpha > 0$, we associate our problem VI (1.1) with the following regularized variational inclusion (RVI) problem:

$$\text{Find } u \in \mathscr{H} \text{ such that } 0 \in \mathscr{A}u + fu + \alpha\mathscr{F}u, \tag{3.1}$$

where $\mathscr{F} : \mathscr{H} \to \mathscr{H}$ is a γ-strongly monotone and L-Lipschitz continuous operator Under our assumptions, problem RVI (3.1) has a unique solution for each $\alpha > 0$, denoted by u_α. Furthermore, we remark that, for each $\lambda > 0$, the point u_α is a solution to problem RVI (3.1) if and only if it is a fixed point of the mapping $J_\lambda^{\mathscr{A}}(I - \lambda(f + \alpha\mathscr{F}))$, i.e.,

$$u_\alpha = J_\lambda^{\mathscr{A}}(u_\alpha - \lambda(fu_\alpha + \alpha\mathscr{F}u_\alpha)).$$

This can follow directly from Lemma 2.2. We have the following result.

Lemma 3.1. [14, Lemma 3] (i) *The sequence* $\{u_\alpha\}$ *is bounded.*

(ii) *There exists a number* $M > 0$ *such that, for all* $\alpha_1 > 0$ *and* $\alpha_2 > 0$,

$$||u_{\alpha_1} - u_{\alpha_2}|| \leq \frac{|\alpha_2 - \alpha_1|}{\alpha_1} M.$$

(iii) $\omega(u_\alpha) \subset \Omega$, *where* $\omega(u_\alpha)$ *is the set of weak cluster points of the sequence* $\{u_\alpha\}$.

(iv) $\lim\limits_{\alpha\to 0^+} u_\alpha = u^\dagger \in \Omega$, *the unique solution of problem* VIP (2.1).

We remark that if $\{\alpha_n\}$ is a sequence of positive real numbers such that $\lim\limits_{n\to\infty} \alpha_n = 0$, then we have by Lemma 3.1(iv) that $\lim_{n\to\infty} u_{\alpha_n} = u^\dagger$. However, in practice, finding u_{α_n} for each $n \geq 1$ can be expansive and time-consuming. In the next section, we introduce an iterative-regularization procedure which generates a sequence $\{u_n\} \subset \mathscr{H}$ satisfying $||u_n - u_{\alpha_n}|| \to 0$ as $n \to \infty$.

4. REGULARIZATION MULTI-STEP INERTIAL PROXIMAL METHOD

In this section, we introduce an iterative-regularization method with multi-step inertial effect. In order to design the method, we need an operator $\mathscr{F} : \mathscr{H} \to \mathscr{H}$ being γ-strongly monotone and k-Lipschitz continuous. In addition, we take the sequence $\{\alpha_n\} \subset (0, +\infty)$ such that

$$\text{(C1)} : \lim_{n\to\infty} \alpha_n = 0,$$

$$\text{(C2)} : \sum_{n=1}^{\infty} \alpha_n = +\infty, \textit{and}$$

$$\text{(C3)} : \lim_{n\to\infty} (\alpha_n - \alpha_{n+1})\alpha_n^{-2} = 0.$$

Conditions (C1)-(C3) hold for sequences such as $\alpha_n = \frac{1}{n^p}$ with $0 < p < 1$. For simplicity, we employ the following conventions: $\frac{0}{0} = +\infty$ and $\frac{1}{0} = +\infty$.

Algorithm 1 Regularization Multi-Step Inertial Proximal Method - RMSIPM

Initialization: Take $u_0,\ u_1 \in \mathscr{H}$ arbitrarily, $r \in (0,2)$, $\theta_0 > 0$, $\sigma > 0$, and a positive integer $N > 0$. For each $i = 1,2,\dots,N$, choose a sequence $\mu_{i,n}$ of positive numbers such that

$$(\text{C4}) : \lim_{n\to\infty} \frac{\mu_{i,n}}{\alpha_n} = 0 \ \text{ and } \ \sum_{n=1}^{+\infty} \mu_{i,n} < +\infty.$$

Iterative Steps: Compute u_{n+1} for $n \geq 1$ as follows

Step 1. Compute

$$w_n = u_n + \sum_{i=1}^{\min\{n,N\}} \theta_{i,n}(u_{n-i+1} - u_{n-i}),$$

where

$$\theta_{i,n} = \begin{cases} \dfrac{\mu_{i,n}}{||u_{n-i+1} - u_{n-i}||} & \text{if } \ u_{n-i+1} \neq u_{n-i} \\ \theta_0 & \text{otherwise.} \end{cases}$$

Step 2. Compute $v_n = J_{\lambda_n}^{\mathscr{A}}(w_n - \lambda_n(fw_n + \alpha_n \mathscr{F} w_n))$, where $\lambda_n > 0$.

Step 3. Compute $u_{n+1} = w_n + r\sigma_n q(v_n, w_n)$, where

$$\begin{cases} q(v_n, w_n) = v_n - w_n - \lambda_n(fv_n - fw_n) \\ D(v_n, w_n) = \langle v_n - w_n, q(v_n, w_n)\rangle \\ \sigma_n = \min\left\{\sigma, \dfrac{D(v_n, w_n)}{||q(v_n, w_n)||^2}\right\}. \end{cases}$$

As in [14], we say that the sequence $\{\lambda_n\}$ in Algorithm 1 satisfies *Predicted Stepsize Conditions* (PSC) if there exist four positive numbers c_1, c_2, $\underline{\lambda}$, and $\bar{\lambda}$ and an integer $n_0 > 0$ such that the following inequalities hold

$$D(v_n, w_n) \geq c_1 ||v_n - w_n||^2 \text{ and } \sigma_n \geq c_2 \tag{4.1}$$

and

$$0 < \underline{\lambda} \leq \lambda_n \leq \bar{\lambda} < +\infty \tag{4.2}$$

for all $n \geq n_0$.

The following lemma gives some cases where condition PSC holds.

Lemma 4.1. *The condition PSC is satisfied if one of the following cases holds:*

(i) $\{\lambda_n\} \subset [a,b] \subset \left(0, \frac{1}{L}\right)$, *where L is the Lipschitz constant of f.*

(ii) *Let $\sigma > 0$, $l \in (0,1)$, $\mu \in (0,1)$. For each $n \geq 0$, λ_n is the largest $\lambda \in \left\{\sigma, \sigma l, \sigma l^2, \cdots\right\}$ such that*

$$\lambda_n ||f(v_n) - f(w_n)|| \leq \mu ||v_n - w_n||.$$

(iii) *Let $\lambda_0 > 0$, $\mu \in (0,1)$, $\{\kappa_n\} \subset [0,+\infty)$ be a summable sequence. For each $n \geq 0$, we take*

$$\lambda_{n+1} = \min\left\{\lambda_n + \kappa_n, \frac{\mu ||v_n - w_n||}{||f(v_n) - f(w_n)||}\right\}. \tag{4.3}$$

Proof. The proof of items (i) and (ii) can be found, for example, in [14, Lemmas 5 and 6]. We give a proof for item (iii). First, since $0<\mu<1$, we are able to choose a number $\xi>0$ such that $(1+\xi)\mu<1$. Since $\sum_{n=1}^{\infty}\kappa_n<+\infty$, we have $\lim_{n\to\infty}\kappa_n=0$, which implies from (4.3) that

$$\begin{cases}\lambda_1 & \le \lambda_0+\kappa_0,\\ \lambda_{n+1} & \le \lambda_n+\kappa_n.\end{cases} \tag{4.4}$$

Thus it inductively follows that

$$\lambda_{n+1}\le \lambda_0+\sum_{i=0}^{n}\kappa_i, \quad \forall n. \tag{4.5}$$

It follows from (4.5) that

$$\lambda_n\le \lambda_0+\sum_{i=0}^{\infty}\kappa_i \tag{4.6}$$

for every n. Besides, the Lipschitz continuity of f results in

$$\frac{\mu||v_n-w_n||}{||f(v_n)-f(w_n)||}\ge \frac{\mu||v_n-w_n||}{L||v_n-w_n||}=\frac{\mu}{L}.$$

It is obvious that

$$\begin{aligned}\lambda_1 &= \min\left\{\lambda_0+\kappa_0, \frac{\mu||y_0-z_0||}{||f(y_0)-f(z_0)||}\right\},\\ &\ge \min\left\{\lambda_0+\kappa_0, \frac{\mu}{L}\right\}.\end{aligned}$$

Hence, we once again use induction to obtain

$$\lambda_{n+1}\ge \min\left\{\lambda_n+\kappa_n, \frac{\mu}{L}\right\}\ge \min\left\{\lambda_n, \frac{\mu}{L}\right\}\ge \min\left\{\lambda_0+\kappa_0, \frac{\mu}{L}\right\} \tag{4.7}$$

for each $n\ge 0$. Now, we let

$$\bar{\lambda}:=\lambda_0+\sum_{n=0}^{+\infty}\kappa_n \text{ and } \underline{\lambda}:=\min\left\{\lambda_0+\kappa_0, \frac{\mu}{L}\right\}.$$

Thus condition (4.2) is directly achieved from (4.6) and (4.7).

Next, it is necessary to indicate that $\{\lambda_n\}$ converges to some number λ. We take a sequence $\{s_n\}_{n=1}^{\infty}$ defined by

$$s_{n+1}:=\lambda_{n+1}-\sum_{i=0}^{n}\kappa_i$$

for each $n\ge 0$. We see that $\{s_n\}$ is non-increasing. Indeed, due to (4.4), we have

$$s_{n+1}-s_n=\lambda_{n+1}-(\lambda_n+\kappa_n)\le 0.$$

Moreover, $\{s_n\}$ is lower bounded since

$$s_n=\lambda_n-\sum_{i=0}^{n-1}\kappa_i\ge \underline{\lambda}-\sum_{i=0}^{\infty}\kappa_i$$

for each $n \geq 1$. The fact that $\{s_n\}$ is non-increasing and lower bounded implies that it converges. The convergence of two sequences $\{s_n\}$ and $\{\sum_{i=0}^{n-1} \kappa_i\}_{n=1}^{\infty}$ simultaneously infers that $\{\lambda_n\}$ also converge. Moreover, from the definition of $D(v_n, w_n)$, we have

$$\begin{aligned} D(v_n, w_n) &= \langle v_n - w_n, v_n - w_n - \lambda_n(fv_n - fw_n)\rangle \\ &= ||v_n - w_n||^2 - \lambda_n \langle v_n - w_n, fv_n - fw_n\rangle \\ &= ||v_n - w_n||^2 - \frac{\lambda_n}{\lambda_{n+1}} \lambda_{n+1} \langle v_n - w_n, fv_n - fw_n\rangle \\ &\geq ||v_n - w_n||^2 - \frac{\lambda_n}{\lambda_{n+1}} \cdot \lambda_{n+1} ||fv_n - fw_n||.||v_n - w_n||. \end{aligned} \tag{4.8}$$

It is obvious from the definition of $\{\lambda_n\}$ that

$$\lambda_{n+1} ||fv_n - fw_n|| \leq \mu ||v_n - w_n||. \tag{4.9}$$

On the other hand, since $\lim_{n\to\infty} \lambda_n = \lambda \geq \underline{\lambda}$, we easily see that

$$\lim_{n\to\infty} \frac{\lambda_n}{\lambda_{n+1}} = 1.$$

Consequently, there exists a positive integer n_0 large enough such that, for all $n \geq n_0$,

$$\frac{\lambda_n}{\lambda_{n+1}} < 1 + \xi. \tag{4.10}$$

We combine inequalities (4.8), (4.9), and (4.10) to reach

$$\begin{aligned} D(v_n, w_n) &\geq ||v_n - w_n||^2 - (1+\xi)\mu ||v_n - w_n||^2 \\ &= c_1 ||v_n - w_n||^2 \end{aligned} \tag{4.11}$$

for all $n \geq n_0$, where $c_1 := 1 - (1+\xi)\mu$, which is already shown to be positive by $(1+\xi)\mu < 1$.. Let $c_2 := \min\{\sigma, c_1\}$. Using (4.11), one sees that

$$\sigma_n = \min\left\{\sigma, \frac{D(v_n, w_n)}{||q(v_n, w_n)||^2}\right\} \geq \min\{\sigma, \frac{c_1}{1+\lambda L}\} = c_2 \tag{4.12}$$

for each $n \geq n_0$. Finally, for each $n \geq n_0$, we derive (4.1) from (4.11) and (4.12). The lemma is proved. □

Theorem 4.1. *Assume that the assumptions* (C1)-(C4) *hold and* $\{\lambda_n\}$ *satisfies condition PSC. Then the sequence* $\{u_n\}$ *generated by Algorithm 1 converges strongly to a solution* $u^\dagger \in \Omega$ *of problem* VI (1.1)*, which uniquely solves problem* VIP (2.1).

Proof. First of all, condition (4.1) ensures that $D(v_n, w_n)$ and σ_n are positive numbers for every n.

Next, because $\gamma > 0$ and $k > 0$, we are able to find the (enough small) numbers $\varepsilon_1, \varepsilon_2, \varepsilon_3 > 0$ which satisfy $2\gamma - k\varepsilon_1 > 0$ and

$$2\gamma r c_2 - k\varepsilon_1 r\sigma - \varepsilon_2 - \varepsilon_3 > 0. \tag{4.13}$$

From now on, let u_{α_n} denote the solution of problem RVI (3.1) with $\alpha = \alpha_n$. We have

$$\begin{aligned}||u_{n+1}-u_{\alpha_n}||^2 &= ||w_n-u_{\alpha_n}||^2+2r\sigma_n\langle w_n-u_{\alpha_n},q(v_n,w_n)\rangle+r^2\sigma_n^2||q(v_n,w_n)||^2\\ &= ||w_n-u_{\alpha_n}||^2+2r\sigma_n\langle v_n-u_{\alpha_n},q(v_n,w_n)\rangle+2r\sigma_n\langle w_n-v_n,q(v_n,w_n)\rangle\\ &\quad +r^2\sigma_n^2||q(v_n,w_n)||^2\\ &= ||w_n-u_{\alpha_n}||^2+2r\sigma_n\langle v_n-u_{\alpha_n},q(v_n,w_n)\rangle-2r\sigma_nD(v_n,w_n)\\ &\quad +r^2\sigma_n^2||q(v_n,w_n)||^2. \end{aligned}\tag{4.14}$$

Next, it follows from the establishment of σ_n that

$$\sigma_n \le \frac{D(v_n,w_n)}{||v_n-w_n||^2},$$

or equivalently,

$$\sigma_n||v_n-w_n||^2 \le D(v_n,w_n). \tag{4.15}$$

We combine (4.14) and (4.15) to find that

$$\begin{aligned}&||u_{n+1}-u_{\alpha_n}||^2\\ &\le ||w_n-u_{\alpha_n}||^2+2r\sigma_n\langle v_n-u_{\alpha_n},q(v_n,w_n)\rangle-2r\sigma_nD(v_n,w_n)+r^2\sigma_nD(v_n,w_n)\\ &= ||w_n-u_{\alpha_n}||^2+2r\sigma_n\langle v_n-u_{\alpha_n},q(v_n,w_n)\rangle-r\sigma_n(2-r)D(v_n,w_n).\end{aligned}$$

On the other hand, since u_{α_n} is the solution to problem RVI (3.1), then $0\in\mathscr{A}x_{\alpha_n}+fx_{\alpha_n}+\alpha_n\mathscr{F}x_{\alpha_n}$, which immediately leads to

$$-\alpha_n\mathscr{F}u_{\alpha_n}\in(\mathscr{A}+f)x_{\alpha_n}. \tag{4.16}$$

Furthermore, by $v_n=J_{\lambda_n}^{\mathscr{A}}(w_n-\lambda_n(fw_n+\alpha_n\mathscr{F}w_n))$, we obtain

$$w_n-\lambda_n(fw_n+\alpha_n\mathscr{F}w_n)-v_n\in\lambda_n\mathscr{A}v_n,$$

or equivalently,

$$w_n-v_n-\lambda_n(fw_n-fv_n)-\lambda_n\alpha_n\mathscr{F}w_n\in\lambda_n(\mathscr{A}+f)v_n.$$

Consequently, we deduce

$$-q(v_n,w_n)-\lambda_n\alpha_n\mathscr{F}w_n\in\lambda_n(\mathscr{A}+f)v_n. \tag{4.17}$$

We combine (4.16), (4.17), and the monotonicity of $\mathscr{A}+f$ to arrive at

$$\langle v_n-x_{\alpha_n},-q(v_n,w_n)-\lambda_n\alpha_n\mathscr{F}w_n+\lambda_n\alpha_n\mathscr{F}u_{\alpha_n}\rangle\ge 0,$$

which can be rewritten as

$$\langle v_n-u_{\alpha_n},q(v_n,w_n)\rangle\le-\lambda_n\alpha_n\langle v_n-u_{\alpha_n},\mathscr{F}w_n-\mathscr{F}u_{\alpha_n}\rangle. \tag{4.18}$$

Besides, we use the Lipschitz continuity and the γ-strong monotinicity of $\mathscr{F}$, as well as the inequality $ab \leq \frac{\varepsilon_1}{2}a^2 + \frac{1}{2\varepsilon_1}b^2$ to obtain

$$\begin{aligned}\langle \mathscr{F}w_n - \mathscr{F}u_{\alpha_n}, v_n - u_{\alpha_n}\rangle &= \langle \mathscr{F}w_n - \mathscr{F}u_{\alpha_n}, v_n - w_n\rangle + \langle \mathscr{F}w_n - \mathscr{F}u_{\alpha_n}, w_n - u_{\alpha_n}\rangle \\ &\geq -k||w_n - u_{\alpha_n}||||v_n - w_n|| + \gamma||w_n - u_{\alpha_n}||^2 \\ &\geq -\frac{k\varepsilon_1}{2}||w_n - u_{\alpha_n}||^2 - \frac{k}{2\varepsilon_1}||v_n - w_n||^2 + \gamma||w_n - u_{\alpha_n}||^2 \\ &= \frac{2\gamma - k\varepsilon_1}{2}||w_n - u_{\alpha_n}||^2 - \frac{k}{2\varepsilon_1}||v_n - w_n||^2. \end{aligned} \tag{4.19}$$

The inequality (4.18) together with (4.19) yields that

$$\langle v_n - u_{\alpha_n}, q(v_n, w_n)\rangle \leq -\lambda_n\alpha_n\left(\frac{2\gamma - k\varepsilon_1}{2}||w_n - u_{\alpha_n}||^2 - \frac{k}{2\varepsilon_1}||v_n - w_n||^2\right). \tag{4.20}$$

Substituting (4.20) into (4.16), we obtain that

$$\begin{aligned}||u_{n+1} - u_{\alpha_n}||^2 &\leq ||w_n - u_{\alpha_n}||^2 - 2r\sigma_n\lambda_n\alpha_n\left(\frac{2\gamma - k\varepsilon_1}{2}||w_n - u_{\alpha_n}||^2 - \frac{k}{2\varepsilon_1}||v_n - w_n||^2\right) \\ &\quad -r\sigma_n(2-r)D(v_n, w_n) \\ &= (1 - r\sigma_n\lambda_n\alpha_n(2\gamma - k\varepsilon_1))||w_n - u_{\alpha_n}||^2 + \frac{kr\sigma_n\lambda_n\alpha_n}{\varepsilon_1}||v_n - w_n||^2 \\ &\quad -r\sigma_n(2-r)D(v_n, w_n). \end{aligned} \tag{4.21}$$

We combine (4.21) and the condition (4.1) to reach

$$\begin{aligned}||u_{n+1} - u_{\alpha_n}||^2 &\leq (1 - r\sigma_n\lambda_n\alpha_n(2\gamma - k\varepsilon_1))||w_n - u_{\alpha_n}||^2 + \frac{kr\sigma_n\lambda_n\alpha_n}{\varepsilon_1}||v_n - w_n||^2 \\ &\quad -r\sigma_n(2-r)c_1||v_n - w_n||^2 \\ &= (1 - r\sigma_n\lambda_n\alpha_n(2\gamma - k\varepsilon_1))||w_n - u_{\alpha_n}||^2 \\ &\quad -\sigma_n\left(r(2-r)c_1 - \frac{kr\lambda_n\alpha_n}{\varepsilon_1}\right)||v_n - w_n||^2. \end{aligned} \tag{4.22}$$

From the definition of $\{\sigma_n\}$ and (4.1), it is easy to see that

$$c_2 \leq \sigma_n \leq \sigma \quad \forall n. \tag{4.23}$$

Next, since $\lim_{n\to\infty}\frac{\mu_{i,n}}{\alpha_n} = 0$ for $i = 1, 2, \ldots, N$, $\lim\limits_{n\to\infty}\alpha_n = 0$, $\{\lambda_n\} \subset [\underline{\lambda}, \bar{\lambda}]$ by (4.2), and $\{\sigma_n\} \subset [c_2, \sigma]$ as seen in (4.23), there exists a number K large enough such that, for every $n > K$,

$$\begin{cases} \sum_{i=1}^{N}\mu_{i,n} < \varepsilon_2\lambda_n\alpha_n, \\ 1 - \lambda_n\alpha_n(2\gamma r\sigma_n - k\varepsilon_1) > 0, \\ 1 - r\sigma_n\lambda_n\alpha_n(2\gamma - k\varepsilon_1) > 0, \\ r(2-r)c_1 - \frac{kr\lambda_n\alpha_n}{\varepsilon_1} > 0. \end{cases} \tag{4.24}$$

Therefore, by the definiton of $\{\theta_{i,n}\}$, we observe that, for all $n > \max\{K,N\}$,

$$
\begin{aligned}
&||w_n - u_{\alpha_n}||^2 \\
&= \left\| (u_n - u_{\alpha_n}) + \sum_{i=1}^{\min\{n,N\}} \theta_{i,n}(u_{n-i+1} - u_{n-i}) \right\|^2 \\
&\le \left(||u_n - u_{\alpha_n}|| + \sum_{i=1}^{N} \theta_{i,n} ||u_{n-i+1} - u_{n-i}|| \right)^2 \\
&= ||u_n - u_{\alpha_n}||^2 + 2\sum_{i=1}^{N} \theta_{i,n} ||u_n - u_{\alpha_n}|| ||u_{n-i+1} - u_{n-i}|| \\
&\quad + 2\sum_{1<i<j<N} \theta_{i,n}\theta_{j,n} ||u_{n-i+1} - u_{n-i}|| ||u_{n-j+1} - u_{n-j}|| + \sum_{i=1}^{N} \theta_{i,n}^2 ||u_{n-i} - u_{n-i+1}||^2 \\
&\le ||u_n - u_{\alpha_n}||^2 + 2\sum_{i=1}^{N} \mu_{i,n} ||u_n - u_{\alpha_n}|| + 2\sum_{1<i<j<N} \mu_{i,n}\mu_{j,n} + \sum_{i=1}^{N} \mu_{i,n}^2 \\
&\le ||u_n - u_{\alpha_n}||^2 + \sum_{i=1}^{N} (\mu_{i,n} ||u_n - u_{\alpha_n}||^2 + \mu_{i,n}) + 2\sum_{1<i<j<N} \mu_{i,n}\mu_{j,n} + \sum_{i=1}^{N} \mu_{i,n}^2 \\
&\le ||u_n - u_{\alpha_n}||^2 + \varepsilon_2\lambda_n\alpha_n ||u_n - u_{\alpha_n}||^2 + \sum_{i=1}^{N} \mu_{i,n} + 2\sum_{1<i<j<N} \mu_{i,n}\mu_{j,n} + \sum_{i=1}^{N} \mu_{i,n}^2 \\
&= (1 + \varepsilon_2\lambda_n\alpha_n) ||u_n - u_{\alpha_n}||^2 + \tilde{\mu}_n,
\end{aligned}
$$

where

$$
\tilde{\mu}_n = \sum_{i=1}^{N} \mu_{i,n} + 2\sum_{1<i<j<N} \mu_{i,n}\mu_{j,n} + \sum_{i=1}^{N} \mu_{i,n}^2.
$$

Combining the above relation and (4.22), we obtain

$$
\begin{aligned}
&||u_{n+1} - u_{\alpha_n}||^2 \\
&\le (1 - r\sigma_n\lambda_n\alpha_n(2\gamma - k\varepsilon_1))(1 + \varepsilon_2\lambda_n\alpha_n) ||u_n - u_{\alpha_n}||^2 + (1 - r\sigma_n\lambda_n\alpha_n(2\gamma - k\varepsilon_1))\tilde{\mu}_n \\
&\quad - \sigma_n \left(r(2-r)c_1 - \frac{kr\lambda_n\alpha_n}{\varepsilon_1} \right) ||v_n - w_n||^2 \\
&\le (1 - r\sigma_n\lambda_n\alpha_n(2\gamma - k\varepsilon_1))(1 + \varepsilon_2\lambda_n\alpha_n) ||u_n - u_{\alpha_n}||^2 \\
&\quad + (1 - r\sigma_n\lambda_n\alpha_n(2\gamma - k\varepsilon_1))\tilde{\mu}_n
\end{aligned} \tag{4.25}
$$

for all $n > \max\{K,N\}$.

On the other hand, using the inequality

$$
ab \le \frac{\varepsilon_3}{2} a^2 + \frac{1}{2\varepsilon_3} b^2
$$

and the result from Lemma 3.1(ii), we have, for all $n > \max\{K,N\}$,

$$\begin{aligned}
&||u_{n+1}-u_{\alpha_{n+1}}||^2\\
&= ||u_{n+1}-u_{\alpha_n}||^2 + ||u_{\alpha_{n+1}}-u_{\alpha_n}||^2 + 2\left\langle u_{n+1}-u_{\alpha_n}, u_{\alpha_n}-u_{\alpha_{n+1}}\right\rangle\\
&\le ||u_{n+1}-u_{\alpha_n}||^2 + ||u_{\alpha_{n+1}}-u_{\alpha_n}||^2 + 2||u_{n+1}-u_{\alpha_n}||\,||u_{\alpha_{n+1}}-u_{\alpha_n}||\\
&\le ||u_{n+1}-u_{\alpha_n}||^2 + ||u_{\alpha_{n+1}}-u_{\alpha_n}||^2 + \varepsilon_3\alpha_n\lambda_n||u_{n+1}-u_{\alpha_n}||^2\\
&\quad + \frac{1}{\varepsilon_3\alpha_n\lambda_n}||u_{\alpha_{n+1}}-u_{\alpha_n}||^2\\
&= (1+\varepsilon_3\alpha_n\lambda_n)||u_{n+1}-u_{\alpha_n}||^2 + \left(\frac{1}{\varepsilon_3\alpha_n\lambda_n}+1\right)||u_{\alpha_{n+1}}-u_{\alpha_n}||^2\\
&\le (1+\varepsilon_3\alpha_n\lambda_n)||u_{n+1}-u_{\alpha_n}||^2 + \frac{(1+\varepsilon_3\alpha_n\lambda_n)M^2}{\varepsilon_3\alpha_n\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^2}\\
&= (1+\varepsilon_3\alpha_n\lambda_n)||u_{n+1}-u_{\alpha_n}||^2 + \frac{(1+\varepsilon_3\alpha_n\lambda_n)M^2}{\varepsilon_3\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^3}.
\end{aligned} \tag{4.26}$$

We deduce from (4.25) and (4.26) that, for all $n > \max\{K,N\}$,

$$\begin{aligned}
&||u_{n+1}-u_{\alpha_{n+1}}||^2\\
&\le (1-r\sigma_n\lambda_n\alpha_n(2\gamma-k\varepsilon_1))(1+\varepsilon_2\lambda_n\alpha_n)(1+\varepsilon_3\lambda_n\alpha_n)||u_n-u_{\alpha_n}||^2\\
&\quad + (1-r\sigma_n\lambda_n\alpha_n(2\gamma-k\varepsilon_1))(1+\varepsilon_3\lambda_n\alpha_n)\tilde{\mu}_n\\
&\quad + \frac{M^2(1+\varepsilon_3\lambda_n\alpha_n)(\alpha_{n+1}-\alpha_n)^2}{\varepsilon_3\lambda_n\alpha_n^3}.
\end{aligned} \tag{4.27}$$

Furthermore, by a simple calculation, we easily find that, for every $n > \max\{K,N\}$,

$$\begin{aligned}
&(1-r\sigma_n\lambda_n\alpha_n(2\gamma-k\varepsilon_1))(1+\varepsilon_3\lambda_n\alpha_n)\\
&= 1-\lambda_n\alpha_n(2\gamma r\sigma_n-k\varepsilon_1 r\sigma_n-\varepsilon_3) - r\sigma_n\lambda_n^2\alpha_n^2\varepsilon_3(2\gamma-k\varepsilon_1)\\
&< 1-\lambda_n\alpha_n(2\gamma r\sigma_n-k\varepsilon_1 r\sigma_n-\varepsilon_3) < 1,
\end{aligned} \tag{4.28}$$

where $2\gamma r\sigma_n - k\varepsilon_1 r\sigma_n - \varepsilon_2 > 0$ results from (4.13) and (4.23) as

$$\begin{aligned}
2\gamma r\sigma_n - k\varepsilon_1 r\sigma_n - \varepsilon_2 &\ge 2\gamma rc_2 - k\varepsilon_1 r\sigma - \varepsilon_2\\
&\ge 2\gamma rc_2 - k\varepsilon_1 r\sigma - \varepsilon_2 - \varepsilon_3 > 0.
\end{aligned}$$

Similarly, we have, for every $n > \max\{K,N\}$,

$$\begin{aligned}
&(1-\lambda_n\alpha_n(2\gamma-k\varepsilon_1))(1+\varepsilon_2\lambda_n\alpha_n)(1+\varepsilon_3\lambda_n\alpha_n)\\
&< (1-\lambda_n\alpha_n(2\gamma r\sigma_n-k\varepsilon_1 r\sigma_n-\varepsilon_3))(1+\varepsilon_2\lambda_n\alpha_n)\\
&< 1-\lambda_n\alpha_n(2\gamma r\sigma_n-k\varepsilon_1 r\sigma_n-\varepsilon_2-\varepsilon_3).
\end{aligned} \tag{4.29}$$

It follows from (4.27), (4.28) and (4.29) that

$$\begin{aligned}
||u_{n+1}-u_{\alpha_{n+1}}||^2 &\leq (1-\lambda_n\alpha_n(2\gamma r\sigma_n - k\varepsilon_1 r\sigma_n - \varepsilon_2-\varepsilon_3))||u_n-u_{\alpha_n}||^2 \\
&\quad + (1-\lambda_n\alpha_n(2\gamma r\sigma_n - k\varepsilon_1 r\sigma_n - \varepsilon_3))\tilde{\mu} \\
&\quad + \frac{M^2(1+\varepsilon_3\lambda_n\alpha_n)}{\varepsilon_3\lambda_n}\cdot\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^3} \\
&\leq (1-\lambda_n\alpha_n(2\gamma r\sigma_n - k\varepsilon_1 r\sigma_n - \varepsilon_2-\varepsilon_3))||u_n-u_{\alpha_n}||^2+\tilde{\mu} \\
&\quad + \frac{M^2(1+\varepsilon_3\lambda_n\alpha_n)}{\varepsilon_3\lambda_n}\cdot\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^3}.
\end{aligned} \tag{4.30}$$

for all $n>\max\{M,N\}$. Set $\Psi_n=||u_n-u_{\alpha_n}||^2$, $p_n=(2\gamma r\sigma_n-k\varepsilon_1 r\sigma_n-\varepsilon_2-\varepsilon_3)\lambda_n\alpha_n$, and

$$q_n=\tilde{\mu_n}+\frac{M^2(1+\varepsilon_3\lambda_n\alpha_n)}{\varepsilon_3\lambda_n}\cdot\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^3}.$$

Then, relation (4.30) can be rewritten as follows:

$$\Psi_{n+1}\leq(1-p_n)\Psi_n+q_n,\ \forall n>\max\{M,N\}. \tag{4.31}$$

It implies from condition (C1) and relations (4.2), (4.23) that

$$p_n\to 0 \qquad \text{as} \qquad n\to\infty. \tag{4.32}$$

Indeed, for $n>\max\{M,N\}$,

$$\begin{aligned}
p_n &= (2\gamma r\sigma_n-k\varepsilon_1 r\sigma_n-\varepsilon_2-\varepsilon_3)\lambda_n\alpha_n \\
&\leq (2\gamma r\sigma - k\varepsilon_1 rc_2-\varepsilon_2-\varepsilon_3)\bar{\lambda}\,\alpha_n\to 0
\end{aligned}$$

since $\alpha_n\to 0$. Besides, from conditions (C2) and (C4) and relations (4.2) and (4.23), we see that, for some positive number $\overline{N}>\max\{M,N\}$

$$\begin{aligned}
\sum_{n=\overline{N}}^{\infty}p_n &= \sum_{n=\overline{N}}^{\infty}\lambda_n\alpha_n(2\gamma r\sigma_n-k\varepsilon_1 r\sigma_n-\varepsilon_2-\varepsilon_3) \\
&\geq (2\gamma rc_2-k\varepsilon_1 r\sigma-\varepsilon_2-\varepsilon_3)\underline{\lambda}\sum_{n=\overline{N}}^{\infty}\alpha_n=+\infty
\end{aligned}$$

and hence,

$$\sum_{n=\overline{N}}^{\infty}p_n=+\infty. \tag{4.33}$$

It can be seen that $\lim_{n\to\infty}\frac{\tilde{\mu_n}}{\alpha_n}=0$. Moreover, by conditions (C3) and (C4), one has

$$\begin{aligned}
&\limsup_{n\to\infty}\frac{q_n}{p_n} \\
&=\limsup_{n\to\infty}\frac{1}{2\gamma r\sigma_n-k\varepsilon_1 r\sigma_n-\varepsilon_2-\varepsilon_3}\left(\frac{M^2(1+\varepsilon_2\lambda_n\alpha_n)}{\varepsilon_2\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^4}+\frac{\tilde{\mu_n}}{\alpha_n}\right) \\
&\leq\frac{1}{2\gamma rc_2-k\varepsilon_1 r\sigma-\varepsilon_2-\varepsilon_3}\limsup_{n\to\infty}\left(\frac{M^2(1+\varepsilon_2\lambda_n\alpha_n)}{\varepsilon_2\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^4}+\frac{\tilde{\mu_n}}{\alpha_n}\right)=0.
\end{aligned} \tag{4.34}$$

We use the combination of inequality (4.31), relations (4.32)-(4.34), and Lemma 2.3 to conclude that $\Psi_n = ||u_n - u_{\alpha_n}||^2 \to 0$ as $n \to \infty$. By Lemma 3.1(iv) and condition (C1), we obtain $u_{\alpha_n} \to u^\dagger$. Thus $u_n \to u^\dagger$ as $n \to \infty$, which completes the proof. □

5. Applications

5.1. **Optimization problem.** Let $\Gamma : \mathscr{H} \to \mathbb{R}$ be a subdifferentiable function on $\mathscr{H}$, i.e, for each $u \in \mathscr{H}$, the *subdifferential* of Γ at u,

$$\partial\Gamma(u) = \{w \in \mathscr{H} : \langle v - u, w\rangle + \Gamma(u) \le \Gamma(v)\}$$

is nonempty. According to [25, Theorem 20.40], $\partial\Gamma$ is maximally monotone.

Let prox_Γ denote the proximal mapping of Γ, which is defined by

$$\text{prox}_\Gamma(u) = \arg\min_{v \in \mathscr{H}} \left\{\Gamma(v) + \frac{1}{2}||v - u||^2\right\} \quad \text{for each } u \in \mathscr{H}.$$

By Fermat's rule [25, Theorem 16.2], $s = \text{prox}_{\lambda\Gamma}(u)$ if and only if

$$0 \in \partial\left(\lambda\Gamma(\cdot) + \frac{1}{2}||\cdot - u||^2\right)(s) = \lambda\partial\Gamma(s) + s - u,$$

or equivalenty, $s = (I + \lambda\partial\Gamma)^{-1}(u) = J_\lambda^{\partial\Gamma}$, where $\lambda > 0$. Thus, we have that

$$\text{prox}_{\lambda\Gamma} = J_\lambda^{\partial\Gamma}.$$

Also, let $\Theta : \mathscr{H} \to \mathbb{R}$ be a differentiable function with its gradient $\nabla\Theta$ being Lipschitz continuous. Evidently, $\partial\Theta$ is single-valued and $\partial\Theta = \nabla\Theta$. Recall the convex optimization problem (OP) mentioned in Section 1

$$\text{Find } u^* \in \mathscr{H} \text{ such that } u^* \in \arg\min_{u \in \mathscr{H}} (\Gamma(u) + \Theta(u)). \tag{5.1}$$

Applying Fermat's rule, we see that the OP (5.1) can be equivalently written as

$$\text{Find } u^* \in \mathscr{H} \text{ such that } 0 \in \partial\Gamma(u^*) + \nabla\Theta(u^*).$$

Thus optimization problem (5.1) is reformulated to the problem VI (1.1) with $\mathscr{A} = \partial\Gamma$ and $f = \nabla\Theta$. The following theorem follows directly from Theorem 4.1.

Theorem 5.1. *Let Γ: $\mathscr{H} \to \mathbb{R}$ be a subdifferential function and $\Theta : \mathscr{H} \to \mathbb{R}$ be differentiable with a Lipschitz continuous gradient. Furthermore, assume that the solution set $\Omega = \arg\min_{\mathscr{H}}(\Gamma + \Theta)$ is nonempty. Let positive numbers $r \in (0,2)$, θ_0, σ, N arbitrarily and the sequences $\{\alpha_n\}$, $\{\mu_{i,n}\}$ for $i = 1,2,\ldots,N$ satisfy conditions* (C1)-(C4). *Choose a γ-strongly monotone and Lipschitz continuous operator $\mathscr{F} : \mathscr{H} \to \mathscr{H}$ and two starting points u_0, $u_1 \in \mathscr{H}$. Let $\{u_n\}$ be a sequence in $\mathscr{H}$ generated by the following procedure*

$$\begin{cases} w_n = u_n + \sum_{i=1}^{\min\{n,N\}} \theta_{i,n}(u_{n-i+1} - u_{n-i}), \\ v_n = \text{prox}_{\lambda_n\Gamma}(w_n - \lambda_n(\nabla\Theta w_n + \alpha_n\mathscr{F}w_n)), \\ u_{n+1} = w_n + r\sigma_n q(v_n, w_n), \end{cases}$$

where

$$\begin{cases} q(v_n,w_n) = v_n - w_n - \lambda_n(\nabla\Theta v_n - \nabla\Theta w_n), \\ D(v_n,w_n) = \langle v_n - w_n, q(v_n,w_n)\rangle, \\ \sigma_n = \min\left\{\sigma, \dfrac{D(v_n,w_n)}{||q(v_n,w_n)||^2}\right\}, \end{cases}$$

where $\lambda_n > 0$ *and satisfies the condition* PSC *and* $\theta_{i,n}$ *is taken exactly the same as Algorithm 1. Then, sequence* $\{u_n\}$ *converges in norm to* $u^\dagger \in \Omega$ *which satisfies the variational inequality* $\langle \mathscr{F}(u^\dagger), u^* - u^\dagger\rangle \geq 0$, $\forall u^* \in \Omega$.

5.2. **Bilevel optimization problem.** Let $\mathscr{C}$ be a nonempty, closed, and convex subset of a Hilbert space $\mathscr{H}$. Let $g : \mathscr{H} \to \mathbb{R}$ be a function, which is convex on $\mathscr{C}$ and differentiable on a neighborhood of $\mathscr{C}$ with ∇g being Lipschitz continuous. Let $h : \mathscr{H} \to \mathbb{R}$ be a strongly convex and differentiable function such that ∇h is Lipschitz continuous. We are interested in the following bilevel optimization problem (BOP):

$$\min_{u\in S} h(u), \tag{5.2}$$

where $S = \arg\min_{u\in\mathscr{C}} g(u)$ which is assumed to be nonempty. Let $\iota_{\mathscr{C}}$ be the indicator function of $\mathscr{C}$, i.e.,

$$\iota_{\mathscr{C}} : \mathscr{H} \to (-\infty, +\infty]$$
$$u \mapsto \begin{cases} 0, & \text{if } u \in \mathscr{C}, \\ +\infty & \text{otherwise.} \end{cases}$$

Therefore, S is the solution to the optimization problem

$$\min_{u\in\mathscr{H}} \left(\iota_{\mathscr{C}}(u) + g(u)\right).$$

Problem (5.2) can be reformulated as

$$\text{Find } u^\dagger \in S \text{ such that } \left\langle \nabla h(u^\dagger), u^* - u^\dagger \right\rangle \geq 0,\ \forall u^* \in S, \tag{5.3}$$

where S is the solution to the variational inclusion problem: Find $u^* \in \mathscr{H}$ such that

$$0 \in \partial\left(\iota_{\mathscr{C}}(u^*) + g(u^*)\right) = N_{\mathscr{C}}(u^*) + \nabla g(u^*). \tag{5.4}$$

We are now ready to solve problems (5.3) and (5.4) by applying Algorithm 1 with $\mathscr{A} = N_{\mathscr{C}}$, $f = \nabla g$ and $\mathscr{F} = \nabla h$. Note that, for every $\lambda > 0$, $J_\lambda^{N_{\mathscr{C}}} = \mathrm{prox}_{\lambda\iota_{\mathscr{C}}} = P_{\mathscr{C}}$, where $P_{\mathscr{C}}$ is the metric projection from $\mathscr{H}$ onto $\mathscr{C}$. We have the following result which follows from Theorem 4.1.

Theorem 5.2. *Let* g*:* $\mathscr{H} \to \mathbb{R}$ *be a function, which is convex on* $\mathscr{C}$ *and differentiable on a neighborhood of* $\mathscr{C}$ *such that* ∇g *is Lipschitz continuous. Let* $h : \mathscr{H} \to \mathbb{R}$ *be a strongly convex and differentiable function with* ∇h *being Lipschitz continuous. Furthermore, assume that the solution set* $S = \arg\min_{\mathscr{C}} g$ *is nonempty. Choose positive numbers* $r \in (0,2)$, θ_0, σ, N *arbitrarily and the sequences* $\{\alpha_n\}$, $\{\mu_{i,n}\}$ *for* $i = 1,2,\ldots,N$ *satisfying conditions* (C1)-(C4). *Choose two certain starting points* u_0, $u_1 \in \mathscr{H}$*. Let* $\{u_n\}$ *be a sequence in* $\mathscr{H}$ *generated by the following procedure*

$$\begin{cases} w_n = u_n + \sum_{i=1}^{\min\{n,N\}} \theta_{i,n}(u_{n-i+1} - u_{n-i}), \\ v_n = P_{\mathscr{C}}(w_n - \lambda_n(\nabla g w_n + \alpha_n \nabla h w_n)), \\ u_{n+1} = w_n + r\sigma_n q(v_n,w_n), \end{cases}$$

where

$$\begin{cases} q(v_n, w_n) = v_n - w_n - \lambda_n(\nabla g v_n - \nabla g w_n), \\ D(v_n, w_n) = \langle v_n - w_n, q(v_n, w_n)\rangle, \\ \sigma_n = \min\left\{\sigma, \dfrac{D(v_n, w_n)}{||q(v_n, w_n)||^2}\right\}, \end{cases}$$

where $\lambda_n > 0$ *and satisfies the condition* PSC *and* $\theta_{i,n}$ *is taken exactly the same as Algorithm 1. Then,* $\{u_n\}$ *converges in norm to* $u^\dagger \in S$*, which is the unique solution to problem (5.2).*

Acknowledgments

Nguyen Hai Ha would like to thank the Master, PhD Scholarship Programme of Vingroup Innovation Foundation (VINIF) for the financial support under code VINIF.2022.TS032. The authors would like to appreciate the editor and the reviewers for their constructive cooperation which substantially improves this paper. This research was funded by University of Transport and Communications (UTC) under grant number T2023-CB-006.

REFERENCES

[1] R.T. Rockafellar, Monotone operators and the proximal point algorithms, SIAM J. Control Optim. 14(5) (1976), 877-898.
[2] A. Antipin, Gradient approach of computing fixed points of equilibrium problems, J. Glob. Optim. 24 (2002), 285-309.
[3] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imaging Sci. 2(1) (2009), 183-202.
[4] R.I. Bot, E.R. Csetnek, C. Hendrich, Inertial Douglas-Rachford splitting for monotone inclusion problems, Appl. Math. Comput. 256 (2015), 472-487.
[5] D. Davis, W.T. Yin, A three-operator splitting scheme and its optimization applications, Set Valued Var. Anal. 25 (2017), 829-858.
[6] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numer. Anal. 16 (1979), 964-979.
[7] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, SIAM J. Control Optim. 38 (2000), 431-446.
[8] 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.
[9] C. Zhang, Y. Wang, Proximal algorithm for solving monotone variational inclusion, Optimization 67 (2018), 1197-1209.
[10] Y. Malitsky, M.K. Tam, A forward-backward splitting method for monotone inclusions without cocoercivity, SIAM J. Optim. 30 (2020), 1451-1472.
[11] D.V. Hieu, P.K. Anh, L.D. Muu, J.J. Strodiot, Iterative regularization methods with new stepsize rules for solving variational inclusions, J. Appl. Math. Comput. 68 (2021), 571-599.
[12] L.D. Muu, N.V. Quy, A new algorithm for solving strongly monotone equilibrium and multivalued variational inequality problems, Appl. Set-Valued Anal. Optim. 4 (2022), 73-82.
[13] H.K. Xu, N. Altwaijry, I. Alzughaibi, S. Chebbi, The viscosity approximation method for accretive operators in Banach spaces, J. Nonlinear Var. Anal. 6 (2022), 37-50.
[14] D.V. Hieu, P.K. Anh,N.H. Ha, Regularization proximal method for monotone variational inclusions, Networks and Spatial Economics 21 (2021), 905-932.
[15] F. Alvarez, 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.
[16] L. Liu, S.Y. Cho, J.C. Yao, Convergence analysis of an inertial Tseng's extragradient algorithm for solving pseudomonotone variational inequalities and applications, J. Nonlinear Var. Anal. 5 (2021), 627-644.

[17] R.I. Bot, E.R. Csetnek, N. Nimana, An inertial proximal-gradient penalization scheme for constrained convex optimization problems, Vietnam J. Math. 46(1) (2018), 53-71.

[18] L.V. Long, D.V. Thong, V.T. Dung, New algorithms for the split variational inclusion problems and application to split feasibility problems, Optimization 68 (2019), 2339-2367.

[19] J. Liang, Convergence rates of first-order operator splitting methods, Optim. Cont [math.OC], Normandie Université, Greyc Cnrs Umr 6072 (2016).

[20] C. Zhang, Q-L. Dong, J. Chen, Multi-step inertial proximal contraction algorithms for monotone variational inclusion problems, Carpathian J. Math. 36(1) (2020), 159-177.

[21] X. Li, Q.L. Dong, A. Gibali, PMiCA – Parallel multi-step inertial contracting algorithm for solving common variational inclusions, J. Nonlinear Funct. Anal. 2022 (2022) 7.

[22] H. Brézis, Chapitre II Operateurs maximaux monotones, North-Holland Math. Stud. 5 (1973), 19-51.

[23] W. Takahashi, Nonlinear Functional Analysis-Fixed Point Theory and Its Applications, Yokohama Publishers, Yokohama, 2000.

[24] H.K. Xu, Another control condition in an iterative method for nonexpansive mappings, Bull. Austral. Math. Soc. 65 (2002), 109-113.

[25] H.H. Bauschke, P.L. Combettes, Convex analysis and monotone operator theory in Hilbert spaces, Springer, New York, 2011.