Acta Mathematica Vietnamica (2022) 47:743–753
https://doi.org/10.1007/s40306-021-00467-6

Hybrid Inertial Contraction Algorithms for Solving Variational Inequalities with Fixed Point Constraints in Hilbert Spaces

Pham Ngoc Anh[1]

Received: 5 November 2020 / Accepted: 12 October 2021 /
Pulished online: 28 January 2022

Abstract
In this paper, basing on the forward-backward method and inertial techniques, we introduce a new algorithm for solving a variational inequality problem over the fixed point set of a nonexpansive mapping. The strong convergence of the algorithm is established under strongly monotone and Lipschitz continuous assumptions imposed on the cost mapping. As an application, we also apply and analyze our algorithm to solve a convex minimization problem of the sum of two convex functions.

Keywords Variational inequality problem · Fixed point constraints · Lipschitz continuous · Strongly monotone · Forward-backward method

Mathematics Subject Classification (2010) 65K10 · 90C25 · 49J35 · 47J25 · 47J20 · 91B50

1 Introduction

Let $\mathcal{H}$ be a real Hilbert space with an inner product $\langle \cdot, \cdot \rangle$ and the induced norm $\| \cdot \|$. Denote weak and strong convergence of a sequence $\{x_n\} \subset \mathcal{H}$ to $x \in \mathcal{H}$ by $x_n \rightharpoonup x$ and $x_n \to x$, respectively. Let C be a nonempty closed convex subset in $\mathcal{H}$, and $F : \mathcal{H} \to \mathcal{H}$ be a cost mapping. The variational inequality problem $\mathrm{VI}(C, F)$ is to find a point $x^* \in C$ such that

$$\langle F(x^*), x - x^* \rangle \geq 0, \quad \forall x \in C.$$

The problem $\mathrm{VI}(C, F)$ was introduced first by Kinderleher and Stampacchia in [12]. This is an important problem that has a variety of theoretical and practical applications [16, 17]. Recently, there are very efficient algorithms for solving this problem. Some popular methods for solving the problem $\mathrm{VI}(C, F)$ are found, for instance, in [13].

✉ Pham Ngoc Anh
anhpn@ptit.edu.vn

[1] Department of Scientific Fundamentals, Posts and Telecommunications Institute of Technology, Hanoi, Vietnam

Let $S : \mathcal{H} \to \mathcal{H}$ be an operator. A *fixed point* of S is a point in $\mathcal{H}$ which is mapped to itself by S, and the set of all fixed points of S is denoted by

$$\mathrm{Fix}(S) := \{x \in \mathcal{H} : x = Sx\}.$$

In this paper, we consider the variational inequality problems with fixed point constraints VIF(F, S), which consist of the following:

$$\text{Find } x^* \in \mathrm{Fix}(S) \text{ such that } \langle F(x^*), x - x^*\rangle \geq 0, \quad x \in \mathrm{Fix}(S).$$

In the case S is the identity mapping, the problem VIF(F, S) is formulated in the form of the problem VI(C, F). In the case $F = 0$, it is written in the form of the lexicographic variational inequality problem when $Sx := Pr_C[x - \lambda G(x)]$, where $\lambda > 0$, $G : \mathcal{H} \to \mathcal{H}$, Pr_C is the metric projection on C. Many other problems can be formulated as the form of the problem VIF(F, S) [10, 11]. There are increasing interests in studying solution algorithms for a monotone class of the problem VIF(F, S) such as parallel subgradient methods [2], extragradient methods [4], subgradient extragradient methods [3], Krasnoselski-Mann iteration method [18], hybrid steepest descent methods [24, 25], and other [18, 22, 23].

Let $f : \mathcal{H} \to \mathbb{R}$ be a convex and differentiable function with a $\mathcal{L}$-coercive gradient ∇f, i.e., $\langle x - y, \nabla f(x) - \nabla f(y)\rangle \geq \mathcal{L}\|\nabla f(x) - \nabla f(y)\|^2 \ \forall x, y \in \mathcal{H}$, and let $g : \mathcal{H} \to \mathbb{R}$ be a proper lower semicontinuous and convex function. Consider the convex problem

$$\min_{x \in \mathcal{H}} \{f(x) + g(x)\}. \tag{1.1}$$

It can be shown that x^* is a solution point of the problem (1.1) if and only if it is characterized by the fixed point equation

$$x^* = \mathrm{prox}_{cg}(I - c\nabla f)(x^*),$$

where $c > 0$, the proximal operator of f is defined by $\mathrm{prox}_f(x) = \mathrm{argmin}\{f(y) + \frac{1}{2}\|y - x\|^2 : y \in \mathcal{H}\}$ and I is the identity operator. For solving the problem (1.1), the fixed point equation leads to the following iteration:

$$x^0 \in \mathcal{H}, x^{k+1} = \underbrace{\mathrm{prox}_{\varepsilon_k g}}_{\text{backward step}} \underbrace{(I - \varepsilon_k \nabla f)(x^k)}_{\text{forward step}}, \quad k \in \mathbb{N}. \tag{1.2}$$

Under the condition $\varepsilon_k \in \left(0, \frac{2}{\mathcal{L}}\right)$, Nakajo et al. [19] show that $\{x^k\}$ converges strongly to a solution of the problem (1.1). The iteration method (1.2) is known as the forward-backward splitting. By using inertial techniques and the forward-backward splitting method, Beck and Teboulle [6] proposed the fast iterative shrinkage-thresholding algorithm for solving the problem (1.1) as follows:

$$\begin{cases} y^k = \mathrm{prox}_{\frac{1}{\mathcal{L}}g}\left[x^k - \frac{1}{\mathcal{L}}\nabla f(x^k)\right], \\ x^{k+1} = y^k + \theta_k(y^k - y^{k-1}), \end{cases} \tag{1.3}$$

where $t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$, $\theta_k = \frac{t_k - 1}{t_{k+1}}$, $x^0 \in \mathcal{H}$ and $t_0 = 1$. Then, the rate of convergence is established and also applied to image restoration problems. In recent years, there have been many authors who modified some forward-backward and inertial methods for solving the other split type problems such as parallel inertial S-iteration forward-backward algorithm [8] for regression and classification problems, inertial hybrid projection-proximal point algorithms [1] for maximal monotone operators, inertial forward-backward algorithms [7] for the minimization of the sum of two nonconvex functions, inertial proximal method [9] for solving Ky Fan minimax inequalities, inertial forward-backward algorithm [15] for monotone inclusions, and other (see [20, 21] and the references therein).

It is worth noting from the above review that the convex minimization problem (1.1) is related to the fixed point problem. Also, we know that a forward-backward operator $S := \mathrm{prox}_{\varepsilon g}(I - \varepsilon\nabla f)$ is nonexpansive in the case $0 < \varepsilon < \frac{2}{L}$, i.e., $\|Sx - Sy\| \leq \|x - y\|$, $\forall x, y \in \mathcal{H}$. So the study on fixed point problems for the class of nonexpansive operators plays an important role in creating optimization methods.

The purpose of this paper is to propose a new iteration algorithm by using the forward-backward iteration scheme (1.3) and inertial techniques for solving the problem VIF(F, S), where the cost mapping F is strongly monotone and Lipschitz continuous on $\mathcal{H}$. Furthermore, we prove a strong convergence result of the proposed algorithm under the condition onto parameters. Subsequently, we apply the proposed algorithm to solving a convex unconstrained minimization problem of the sum of two convex functions by using the nonexpansiveness of the forward-backward operator S.

The paper is organized as follows. In Section 2, we present some definitions and lemmas which will be used in the paper. Section 3 deals with a new inertial forward-backward algorithm for solving the variational inequalities over the fixed point set of a nonexpansive mapping VIF(F, S) and the proof of its strong convergence in a real Hilbert space $\mathcal{H}$. As an application of our proposed algorithm, Section 4 is devoted to solve a convex unconstrained minimization problem of the sum of two convex functions in $\mathcal{H}$.

2 Preliminaries

Denote weak and strong convergence of a sequence $\{x^n\} \subset \mathcal{H}$ to $x \in \mathcal{H}$ by $x^n \rightharpoonup x$ and $x^n \to x$, respectively.

We recall that a mapping $S : \mathcal{H} \to \mathcal{H}$ is said to be

- *Strongly monotone* with constant $\beta > 0$ (shortly β-strongly monotone), if

$$\langle S(x) - S(y), x - y\rangle \geq \beta\|y - x\|^2, \quad \forall x, y \in \mathcal{H};$$

- *Lipschitz continuous* with constant $L > 0$ (shortly L-Lipschitz continuous), if

$$\|Sx - Sy\| \leq L\|x - y\|, \quad \forall x, y \in \mathcal{H};$$

- *Contraction* with constant $L > 0$, if S is L-Lipschitz continuous, where $L < 1$;

- *Nonexpansive*, if S is 1-Lipschitz continuous.

For each $x \in \mathcal{H}$, there exists a unique point in C, denoted by $Pr_C(x)$ satisfying

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

The mapping Pr_C is usually called the *metric projection* of $\mathcal{H}$ on C. An important property of Pr_C is nonexpansive on $\mathcal{H}$.

Given a function $g : \mathcal{H} \to \mathcal{R}$, the proximal mapping of g on C is the mapping given by

$$\mathrm{prox}_{(g,C)}(y) = \mathrm{argmin}\left\{g(x) + \frac{1}{2}\|y - x\|^2 \,:\, x \in C\right\}.$$

For any $x \in \mathcal{H}$, the following three claims in [5] are equivalent

(a) $u = \mathrm{prox}_{(g,C)}(x)$;

(b) $x - u \in \partial g(u) := \{w_u \in \mathcal{H} : g(x) - g(u) \geq \langle w_u, x - u\rangle, \ \forall x \in \mathcal{H}\}$;
(c) $\langle x - u, y - u\rangle \leq g(y) - g(u), \ \forall y \in C$.

Moreover, the proximal mapping of g on C is (firmly) nonexpansive and

$$\text{Fix}\left(\text{prox}_{(g,C)}\right) = \{x \in C : g(x) \leq g(y), \ \forall y \in C\}.$$

Note that, if g is the indicator function on C (defined by $\delta_C(x) = 0$ if $x \in C$; otherwise $\delta_C(x) = +\infty$), then $\text{prox}_{(g,C)} = Pr_C$.

Now we recall the following lemmas which are useful tools for proving our convergence results.

Lemma 2.1 [20, Lemma 2.6] *Let $\{s_k\}$ be a sequence of nonnegative real numbers and $\{p_k\}$ a sequence of real numbers. Let $\{\alpha_k\}$ be a sequence of real numbers in $(0, 1)$ such that $\sum_{k=1}^{\infty}\alpha_k = \infty$. Assume that*

$$s_{k+1} \leq (1 - \alpha_k)s_k + \alpha_k p_k, \quad k \in \mathbb{N}.$$

If $\limsup_{i\to\infty} p_{k_i} \leq 0$ for every subsequence $\{s_{k_i}\}$ of $\{s_k\}$ satisfying

$$\liminf_{i\to\infty}(s_{k_i+1} - s_{k_i}) \geq 0,$$

then $\lim_{k\to\infty} s_k = 0$.

Lemma 2.2 [14, Demiclosedness principle] *Assume that S is a nonexpansive self-mapping of a nonempty closed convex subset C of a real Hilbert space $\mathcal{H}$. If $\text{Fix}(S) \neq \emptyset$, then $I - S$ is demiclosed; that is, whenever $\{x^k\}$ is a sequence in C converging weakly to some $\bar{x} \in C$ and the sequence $\{(I - S)(x^k)\}$ converges strongly to some $\bar{y}$, it follows that $(I - S)(\bar{x}) = \bar{y}$. Here I is the identity operator of $\mathcal{H}$.*

3 Algorithm and Its Convergence

For solving the variational inequalities over the fixed point set VIF(F, S), we assume the mappings F and S, parameters satisfy the following conditions.

(A_1) F is β-strongly monotone, L-Lipschitz continuous such that $\beta > 0$ and $L > 0$;
(A_2) S is nonexpansive;
(A_3) The solution set of the problem VIF(F, S) is nonempty;
(A_4) For every $k \geq 0$, the positive parameters β_k, γ_k, τ_k, λ_k and $\{\mu_k\}$ satisfy the following restrictions:

$$\begin{cases} 0 < c_1 \leq \beta_k \leq c_2 < 1, \mu_k \leq \eta, \\ 0 < \gamma_k < 1, \lim\limits_{k\to\infty}\gamma_k = 0, \sum\limits_{k=1}^{\infty}\gamma_k = \infty, \\ \lim\limits_{k\to\infty}\frac{\tau_k}{\gamma_k} = 0, \lambda_k \in \left(\frac{\beta}{L^2}, \frac{2\beta}{L^2}\right), a \in (0, 1), \sqrt{1 - 2\lambda_k\beta + \lambda_k^2 L^2} < 1 - a. \end{cases} \tag{3.1}$$

Algorithm 3.1 (Hybrid inertial contraction algorithm)

Initialization: *Take* $x^0, x^1 \in \mathcal{H}$ *arbitrarily.*

Iterative steps: $k = 1, 2, \ldots$

Step 1. Compute an inertial parameter

$$\theta_k = \begin{cases} \min\left\{\mu_k, \dfrac{\tau_k}{\|x^k - x^{k-1}\|}\right\} & \textit{if } \|x^k - x^{k-1}\| \neq 0, \\ \mu_k & \textit{otherwise.} \end{cases}$$

Step 2. Compute

$$\begin{cases} w^k = x^k + \theta_k(x^k - x^{k-1}), \\ z^k = (1-\gamma_k)Sw^k + \gamma_k\left[w^k - \lambda_k F(w^k)\right], \\ x^{k+1} = (1-\beta_k)Sw^k + \beta_k S z^k. \end{cases} \tag{3.2}$$

Step 3. Set $k := k + 1$ *and return to Step 1.*

A strong convergence result is established in the following theorem.

Theorem 3.2 *Assume that the assumptions* (A_1)*–*(A_4) *are satiSfied. Then, the sequence* $\{x^k\}$ *generated by Algorithm 3.1 converges strongly to a unique solution* x^* *of the problem* $VIF(F, S)$.

Proof Since F is β-strongly monotone and L-Lipschitz continuous on $\mathcal{H}$, for each $\lambda_k > 0$, we have

$$\begin{aligned} &\|w^k - \lambda_k F(w^k) - [x^* - \lambda_k F(x^*)]\|^2 \\ &= \|w^k - x^*\|^2 - 2\lambda_k\langle F(w^k) - F(x^*), w^k - x^*\rangle + \lambda_k^2\|F(w^k) - F(x^*)\|^2 \\ &\leq \|w^k - x^*\|^2 - 2\lambda_k\beta\|w^k - x^*\|^2 + \lambda_k^2 L^2\|w^k - x^*\|^2 \\ &= (1 - 2\lambda_k\beta + \lambda_k^2 L^2)\|w^k - x^*\|^2. \end{aligned} \tag{3.3}$$

It is well-known to see that F is strongly monotone, so the problem VIF(F, S) has a unique solution, and a point $x^* \in \text{Fix}(S)$ is a solution of the problem if and only if $x^* = Pr_{\text{Fix}(S)}[x^* - \lambda_k F(x^*)]$. From the schemes (3.2) and (3.3), we get

$$\begin{aligned} \|z^k - x^*\| &= \left\|(1-\gamma_k)Sw^k + \gamma_k\left[w^k - \lambda_k F(w^k)\right] - x^*\right\| \\ &\leq \gamma_k\left\|w^k - \lambda_k F(w^k) - x^*\right\| + (1-\gamma_k)\|Sw^k - x^*\| \\ &\leq \gamma_k\|w^k - \lambda_k F(w^k) - [x^* - \lambda_k F(x^*)]\| + \gamma_k\lambda_k\|F(x^*)\| + (1-\gamma_k)\|Sw^k - Sx^*\| \\ &\leq \gamma_k\sqrt{1 - 2\lambda_k\beta + \lambda_k^2 L^2}\|w^k - x^*\| + \gamma_k\lambda_k\|F(x^*)\| + (1-\gamma_k)\|w^k - x^*\| \\ &= [1 - \gamma_k(1-\delta_k)]\|w^k - x^*\| + \gamma_k\lambda_k\|F(x^*)\|, \end{aligned} \tag{3.4}$$

where $\delta_k := \sqrt{1 - 2\lambda_k\beta + \lambda_k^2 L^2} \in (0, 1-a)$. Combining this and (3.1), we obtain

$$\begin{aligned}
&\|x^{k+1} - x^*\| \\
&= \|(1-\beta_k)Sw^k + \beta_k Sz^k - x^*\| \\
&\le (1-\beta_k)\|Sw^k - Sx^*\| + \beta_k\|Sz^k - Sx^*\| \\
&\le (1-\beta_k)\|w^k - x^*\| + \beta_k\|z^k - x^*\| \\
&\le [1 - \beta_k\gamma_k(1-\delta_k)]\|w^k - x^*\| + \beta_k\gamma_k\lambda_k\|F(x^*)\| \\
&\le [1 - \beta_k\gamma_k(1-\delta_k)]\left(\|x^k - x^*\| + \theta_k\|x^k - x^{k-1}\|\right) + \beta_k\gamma_k\frac{2\beta\|F(x^*)\|}{L^2} \\
&\le [1 - \beta_k\gamma_k(1-\delta_k)]\|x^k - x^*\| + \beta_k\gamma_k\left(\frac{\theta_k}{\beta_k\gamma_k}\|x^k - x^{k-1}\| + \frac{2\beta\|F(x^*)\|}{L^2}\right) \\
&\le [1 - \beta_k\gamma_k(1-\delta_k)]\|x^k - x^*\| + \beta_k\gamma_k(1-\delta_k)\left(\frac{\theta_k}{a\beta_k\gamma_k}\|x^k - x^{k-1}\| + \frac{2\beta\|F(x^*)\|}{aL^2}\right).
\end{aligned}$$

By using Step 1 and the conditions (3.1), we deduce

$$0 \le \frac{\theta_k}{\beta_k\gamma_k}\|x^k - x^{k-1}\| \le \frac{\tau_k}{c_1\gamma_k} \to 0 \quad \text{as } k \to \infty.$$

This implies $M = \sup_k\left\{\frac{\theta_k}{a\beta_k\gamma_k}\|x^k - x^{k-1}\| + \frac{2\beta\|F(x^*)\|}{aL^2}\right\} < +\infty$. Then,

$$\begin{aligned}
\|x^{k+1} - x^*\| &\le [1 - \beta_k\gamma_k(1-\delta_k)]\|x^k - x^*\| + \beta_k\gamma_k(1-\delta_k)M \\
&\le \max\left\{\|x^k - x^*\|, M\right\}.
\end{aligned}$$

By mathematical induction, we deduce that

$$\|x^k - x^*\| \le \max\left\{\|x^1 - x^*\|, M\right\}, \quad \forall k \ge 1.$$

So, $\{x^k\}$ is bounded. From (3.2), it follows $\|w^k - x^k\| = \theta_k\|x^k - x^{k-1}\| < +\infty$. By using (3.4), we also have that both $\{z^k\}$ and $\{w^k\}$ are bounded. By (3.3) and the relation

$$\|x + y\|^2 \le \|x\|^2 + 2\langle y, x + y\rangle, \quad \forall x, y \in \mathcal{H},$$

we get

$$\begin{aligned}
&\|z^k - x^*\|^2 \\
&= \left\|(1-\gamma_k)(Sw^k - x^*) + \gamma_k[w^k - \lambda_k F(w^k) - (x^* - \lambda_k F(x^*))] - \gamma_k\lambda_k F(x^*)\right\|^2 \\
&\le \|(1-\gamma_k)(Sw^k - x^*) + \gamma_k[w^k - \lambda_k F(w^k) - (x^* - \lambda_k F(x^*))]\|^2 \\
&\quad -2\gamma_k\lambda_k\langle F(x^*), z^k - x^*\rangle \\
&\le (1-\gamma_k)\|Sw^k - x^*\|^2 + \gamma_k\|w^k - \lambda_k F(w^k) - (x^* - \lambda_k F(x^*))\|^2 \\
&\quad -2\gamma_k\lambda_k\langle F(x^*), z^k - x^*\rangle \\
&\le (1-\gamma_k)\|w^k - x^*\|^2 + \gamma_k\delta_k^2\|w^k - x^*\|^2 - 2\gamma_k\lambda_k\langle F(x^*), z^k - x^*\rangle \\
&\le \left[1 - \gamma_k(1-\delta_k^2)\right]\|w^k - x^*\|^2 - 2\gamma_k\lambda_k\langle F(x^*), z^k - x^*\rangle. \qquad (3.5)
\end{aligned}$$

From $w^k = x^k + \theta_k(x^k - x^{k-1})$, it implies

$$\begin{aligned}
\|w^k - x^*\|^2 &= \|x^k - x^*\|^2 + \theta_k^2\|x^k - x^{k-1}\|^2 + 2\theta_k\langle x^k - x^*, x^k - x^{k-1}\rangle \\
&\le \|x^k - x^*\|^2 + \theta_k^2\|x^k - x^{k-1}\|^2 + 2\theta_k\|x^k - x^*\|\|x^k - x^{k-1}\|. \qquad (3.6)
\end{aligned}$$

Combining (3.5) and (3.6), we have

$$
\begin{aligned}
&\|x^{k+1}-x^*\|^2\\
&= \|(1-\beta_k)(Sw^k-x^*)+\beta_k(Sz^k-x^*)\|^2\\
&= (1-\beta_k)\|Sw^k-Sx^*\|^2+\beta_k\|Sz^k-Sx^*\|^2-\beta_k(1-\beta_k)\|Sw^k-Sz^k\|^2\\
&\le (1-\beta_k)\|w^k-x^*\|^2+\beta_k\|z^k-x^*\|^2-\beta_k(1-\beta_k)\|Sw^k-Sz^k\|^2\\
&\le (1-\beta_k)\|w^k-x^*\|^2+\beta_k(1-\gamma_k)\|w^k-x^*\|^2+\beta_k\gamma_k\delta_k^2\|w^k-x^*\|^2\\
&\quad -2\beta_k\gamma_k\lambda_k\langle F(x^*),z^k-x^*\rangle-\beta_k(1-\beta_k)\|Sw^k-Sz^k\|^2\\
&= [1-\beta_k\gamma_k(1-\delta_k^2)]\|w^k-x^*\|^2-2\beta_k\gamma_k\lambda_k\langle F(x^*),z^k-x^*\rangle\\
&\quad -\beta_k(1-\beta_k)\|Sw^k-Sz^k\|^2\\
&\le [1-\beta_k\gamma_k(1-\delta_k^2)]\|x^k-x^*\|^2+\theta_k^2\|x^k-x^{k-1}\|^2+2\theta_k\|x^k-x^*\|\|x^k-x^{k-1}\|\\
&\quad -2\beta_k\gamma_k\lambda_k\langle F(x^*),z^k-x^*\rangle-\beta_k(1-\beta_k)\|Sw^k-Sz^k\|^2\\
&\le [1-\beta_k\gamma_k(1-\delta_k^2)]\|x^k-x^*\|^2-\beta_k(1-\beta_k)\|Sw^k-Sz^k\|^2+\beta_k\gamma_k(1-\delta_k^2)\sigma_k,
\end{aligned}
$$

where

$$
\begin{aligned}
\sigma_k &:= \frac{1}{1-\delta_k^2}\left\{\frac{\theta_k^2}{\beta_k\gamma_k}\|x^k-x^{k-1}\|^2+\frac{2\theta_k}{\beta_k\gamma_k}\|x^k-x^*\|\|x^k-x^{k-1}\|\right.\\
&\qquad \left. -\,2\lambda_k\langle F(x^*),z^k-x^*\rangle\right\}\\
&\le \frac{1}{a(2-a)}\left\{-\,2\lambda_k\langle F(x^*),z^k-x^*\rangle+\left(\frac{\theta_k}{c_1\gamma_k}\|x^k-x^{k-1}\|\right)\theta_k\|x^k-x^{k-1}\|\right.\\
&\qquad \left. +2\|x^k-x^*\|\left(\frac{\theta_k}{c_1\gamma_k}\|x^k-x^{k-1}\|\right)\right\}.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\|x^{k+1}-x^*\|^2 \\
&\le [1-\beta_k\gamma_k(1-\delta_k^2)]\|x^k-x^*\|^2-\beta_k(1-\beta_k)\|Sw^k-Sz^k\|^2+\beta_k\gamma_k(1-\delta_k^2)\sigma,
\end{aligned}
\tag{3.7}
$$

where $\sigma := \sup_k \sigma_k \in (0,\infty)$. Now we apply Lemma 2.1 for $s_k := \|x^k-x^*\|^2, \alpha_k := \beta_k\gamma_k(1-\delta_k^2) \in (0,1)$ and $p_k := \sigma_k$. Since (3.7), we have

$$
s_{k+1} \le (1-\alpha_k)s_k+\alpha_k p_k.
$$

Assume that $\{s_{k_i}\}$ is a subsequence of $\{s_k\}$ such that

$$
\liminf_{i\to\infty}\left(s_{k_i+1}-s_{k_i}\right) \ge 0.
$$

Combining this, (3.7), and (3.1), we obtain

$$\begin{aligned}
0 &\leq c_1(1-c_2)\limsup_{i\to\infty}\left\|Sw^{k_i}-Sz^{k_i}\right\|^2\\
&\leq \limsup_{i\to\infty}\beta_{k_i}\left(1-\beta_{k_i}\right)\left\|Sw^{k_i}-Sz^{k_i}\right\|^2\\
&\leq \limsup_{i\to\infty}\left[s_{k_i}-s_{k_i+1}+\beta_{k_i}\gamma_{k_i}(1-\delta_{k_i}^2)\sigma\right]\\
&\leq \limsup_{i\to\infty}\left(s_{k_i}-s_{k_i+1}\right)\\
&= -\liminf_{i\to\infty}\left(s_{k_i+1}-s_{k_i}\right)\\
&\leq 0.
\end{aligned}$$

Consequently,

$$\lim_{i\to\infty}\left\|Sw^{k_i}-Sz^{k_i}\right\|=0. \tag{3.8}$$

From the scheme (3.2), it follows

$$\|z^k-Sw^k\|=\gamma_k\|w^k-\lambda_k F(w^k)-Sw^k\|,$$

and hence

$$\left\|z^{k_i}-Sw^{k_i}\right\|=\gamma_{k_i}\left\|w^{k_i}-\lambda_{k_i}F(w^{k_i})-Sw^{k_i}\right\|.$$

Then, using $\lim_{k\to\infty}\gamma_k=0$ and the boundedness of $\{w^k\}$, we get

$$\lim_{i\to\infty}\left\|z^{k_i}-Sw^{k_i}\right\|=0. \tag{3.9}$$

Since (3.8) and (3.9), we obtain

$$\left\|z^{k_i}-Sz^{k_i}\right\|\leq\left\|z^{k_i}-Sw^{k_i}\right\|+\left\|Sw^{k_i}-Sz^{k_i}\right\|\to 0,\quad \text{as } i\to\infty. \tag{3.10}$$

We next show that $\limsup_{i\to\infty}p_{k_i}\leq 0$. Since the condition (3.1), we have

$$\begin{aligned}
p_k &= \sigma_k\\
&\leq \frac{1}{a(2-a)}\left\{-2\lambda_k\langle F(x^*),z^k-x^*\rangle+\left(\frac{\theta_k}{c_1\gamma_k}\|x^k-x^{k-1}\|\right)\theta_k\|x^k-x^{k-1}\|\right.\\
&\quad\left.+2\|x^k-x^*\|\left(\frac{\theta_k}{c_1\gamma_k}\|x^k-x^{k-1}\|\right)\right\}\\
&\leq \frac{1}{a(2-a)}\left\{-2\lambda_k\langle F(x^*),z^k-x^*\rangle+\frac{\tau_k}{\gamma_k}\left(\frac{\mu_k\|x^k-x^{k-1}\|}{c_1}+\frac{2\|x^k-x^*\|}{c_1}\right)\right\}.
\end{aligned}$$

Since $\lambda_k\in\left(\frac{\beta}{L^2},\frac{2\beta}{L^2}\right)$, the boundedness of $\{x^k\}$ and $\{\mu_k\}$, it suffices to show that

$$\limsup_{i\to\infty}\langle F(x^*),x^*-z^{k_i}\rangle\leq 0.$$

Since $\{z^k\}$ is bounded, we can assume that there exists a subsequence $\{\bar z^{k_i}\}$ of $\{z^{k_i}\}$ such that $\bar z^{k_i}\rightharpoonup\bar x$ and

$$\limsup_{i\to\infty}\langle F(x^*),x^*-z^{k_i}\rangle=\lim_{i\to\infty}\langle F(x^*),x^*-\bar z^{k_i}\rangle.$$

Applying Lemma 2.2 for the nonexpansive mapping S with (3.10), we deduce that $\bar x\in\mathrm{Fix}(S)$. Thus

$$\limsup_{i\to\infty}\langle F(x^*),x^*-z^{k_i}\rangle=\langle F(x^*),x^*-\bar x\rangle\leq 0.$$

By Lemma 2.1, we can conclude that $x^k\to x^*$ as $k\to\infty$. The proof is complete. □

4 Application to Convex Problems

In this section, we consider the minimization problem (1.1) in the form of the sum of two convex functions in $\mathcal{H}$. Let $g : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be proper lower semi-continuous and convex. The proximal operator of g on C, in short prox_g, is formulated as follows:

$$\text{prox}_g(y) = \text{argmin}\left\{g(x) + \frac{1}{2}\|y - x\|^2 \, : \, x \in C\right\}, \quad y \in \mathcal{H}.$$

It is well-known to see that prox_g has the nonexpansiveness on $\mathcal{H}$ [5], i.e., $\|\text{prox}_g(y_1) - \text{prox}_g(y_2)\| \le \|y_1 - y_2\|$ for all $y_1, y_2 \in \mathcal{H}$.

In this situation, we put the following assumptions:

(B_1) $f : \mathcal{H} \to \mathbb{R}$ is convex and differentiable, its gradient ∇f is $\mathcal{L}$-coercive, i.e., $\langle \nabla f(x) - \nabla f(y), x - y\rangle \ge \mathcal{L}\|\nabla f(x) - \nabla f(y)\|^2$ for all $x, y \in \mathcal{H}$;
(B_2) $g : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is proper lower semicontinuous and convex;
(B_3) The solution set Ω of (1.1) is nonempty;
(B_4) $F : \mathcal{H} \to \mathcal{H}$ is β-strongly monotone and L-Lipschitz continuous.

By utilizing Algorithm 3.1, we obtain the following algorithm for solving the problem (1.1).

Algorithm 4.1
Initialization: *Take* $\varepsilon \in \left(0, \frac{2}{L}\right)$ *and two points* $x^0, x^1 \in \mathcal{H}$ *arbitrarily.*

Iterative steps: $k = 1, 2, \ldots$
Step 1. Compute an inertial parameter

$$\theta_k = \begin{cases} \min\left\{\mu_k, \dfrac{\tau_k}{\|x^k - x^{k-1}\|}\right\} & \text{if } \|x^k - x^{k-1}\| \neq 0, \\ \mu_k & \text{otherwise.} \end{cases}$$

Step 2. Compute

$$\begin{cases} w^k = x^k + \theta_k(x^k - x^{k-1}), \\ v^k = \text{prox}_{\varepsilon g}\left[w^k - \varepsilon \nabla f(w^k)\right] \\ z^k = (1 - \gamma_k)v^k + \gamma_k\left[w^k - \lambda_k F(w^k)\right], \\ x^{k+1} = (1 - \beta_k)v^k + \beta_k \text{prox}_{\varepsilon g}\left[z^k - \varepsilon \nabla f(z^k)\right]. \end{cases}$$

Step 3. Set $k := k + 1$ *and return to Step 1.*

A strong convergence result is established in the following theorem.

Theorem 4.2 *Assume that the assumptions* (B_1)–(B_4) *are satisfied. Under the conditions* (3.1) *and* $\varepsilon \in \left(0, \frac{2}{\mathcal{L}}\right)$*, the sequence* $\{x^k\}$ *generated by Algorithm 4.1 converges strongly to a solution* x^* *of the convex problem* (1.1) *Moreover,* $x^* = Pr_\Omega[x^* - \varepsilon F(x^*)]$.

Proof For each $x \in \mathcal{H}$, set $S(x) = \text{prox}_{\varepsilon g}[x - \varepsilon \nabla f(x)]$. For each $x, y \in \mathcal{H}$, since $\text{prox}_{\varepsilon g}$ is nonexpansive and the assumption (B_1), we have

$$\begin{aligned}
\|S(x) - S(y)\|^2 &= \left\|\text{prox}_{\varepsilon g}[x - \varepsilon \nabla f(x)] - \text{prox}_{\varepsilon g}[y - \varepsilon \nabla f(y)]\right\|^2 \\
&\leq \|x - \varepsilon \nabla f(x) - [y - \varepsilon \nabla f(y)]\|^2 \\
&= \|x - y\|^2 - 2\varepsilon\langle x - y, \nabla f(x) - \nabla f(y)\rangle + \varepsilon^2\|\nabla f(x) - \nabla f(y)\|^2 \\
&\leq \|x - y\|^2 - \varepsilon(2\varepsilon - \mathcal{L})\|\nabla f(x) - \nabla f(y)\|^2 \\
&\leq \|x - y\|^2,
\end{aligned}$$

where the last inequality is deduced from $\varepsilon \in \left(0, \frac{2}{\mathcal{L}}\right)$. Then, S is nonexpansive on $\mathcal{H}$. So, the convergence results are deduced from Theorem 3.2 for the nonexpansive mapping S and the cost mapping F. □

Acknowledgements We are very grateful to anonymous referees for their really helpful and constructive comments that helped us very much in improving the paper.

Funding This research is funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant Number 101.02-2019.303.

References

1. Alvarez, F.: Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space. SIAM J. Optim. **14**, 773–782 (2004)
2. Anh, P.N., Hien, N.D., Phuong, N.X., Ngoc, V.T.: Parallel subgradient methods for variational inequalities involving nonexpansive mappings. Appl. Anal. https://doi.org/10.1080/00036811.2019.1584288 (2019)
3. Anh, P.N., Le Thi, H.A.: New subgradient extragradient methods for solving monotone bilevel equilibrium problems. Optim. **68**(1), 2097–2122 (2019)
4. Anh, P.N., Kim, J.K., Muu, L.D.: An extragradient method for solving bilevel variational inequalities. J. Glob. Optim. **52**, 627–639 (2012)
5. Beck, A.: First-order Methods in Optimization. MOS-SIAM Series on Optimization, Chapter 6 (2017)
6. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imag. Sc. **2**, 183–202 (2009)
7. Bot, R.I., Csetnek, E.R., Laszlo, S.C.: An inertial forward-backward algorithm for the minimization of the sum of two nonconvex functions. EURO J. Comput. Optim. **4**, 3–25 (2016)
8. Bussaban, L., Suantai, S., Kaewkhao, A.: A parallel inertial S-iteration forward-backward algorithm for regression and classification problems. Carpathian J. Math. **36**, 35–44 (2020)
9. Chbani, Z., Riahi, H.: Weak and strong convergence of an inertial proximal method for solving Ky Fan minimax inequalities. Optim. Lett. **7**, 185–206 (2013)
10. Duc, P.M., Muu, L.D.: A splitting algorithm for a class of bilevel equilibrium problems involving nonexpansive mappings. Optim. **65**(10), 1855–1866 (2016)
11. Iiduka, H., Yamada, I.: A Subgradient-type method for the equilibrium problem over the fixed point set and its applications. Optim. **58**, 251–261 (2009)
12. Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications. Academic Press (1980)
13. Konnov, I.V.: Combined Relaxation Methods for Variational Inequalities. Springer-Verlag, Berlin (2000)
14. Liu, F., Nashed, M.Z.: Regularization of nonlinear ill-posed variational inequalities and convergence rates. Set-Valued Anal. **6**, 313–344 (1998)
15. Lorenz, D.A., Pock, T.: An inertial forward-backward algorithm for monotone inclusions. J. Math. Imaging Vis. **51**, 311–325 (2015)
16. Londono, G., Lozano, A.: A bilevel optimization program with equilibrium constraints for an urban network dependent on time. Transp. Res. Proc. **3**, 905–914 (2014)
17. Marcotte, P.: Network design problem with congestion effects: A case of bilevel programming. Math. Progr. **34**(2), 142–162 (1986)

18. Moudafi, A.: Krasnoselski–Mann iteration for hierarchical fixed-point problems. Inverse Prob. **23**, 1635–1640 (2007)
19. Nakajo, K., Shimoji, K., Takahashi, W.: On strong convergence by the hybrid method for families of mappings in Hilbert spaces. Nonl. Anal. Theory, Meth. Appl. **71**(1–2), 112–119 (2009)
20. Saejung, S., Yotkaew, P.: Approximation of zeros of inverse strongly monotone operators in Banach spaces. Nonl. Anal. **75**, 724–750 (2012)
21. Solodov, M.: An explicit descent method for bilevel convex optimization. J. Conv. Anal. **14**, 227–237 (2007)
22. Xu, M.H., Li, M., Yang, C.C.: Neural networks for a class of bilevel variational inequalities. J. Glob. Optim. **44**, 535–552 (2009)
23. Yao, Y., Marino, G., Muglia, L.: A modified Korpelevichś method convergent to the minimum-norm solution of a variational inequality. Optim. **63**, 559–569 (2014)
24. Yamada, I.: The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings. Stud. Comput. Math. **8**, 473–504 (2001)
25. Yamada, I., Ogura, N.: Hybrid steepest descent method for the variational inequality problem over the the fixed point set of certain quasi-nonexpansive mappings. Numer. Funct. Anal. Optim. **25**, 619–655 (2004)