ARCHIVUM MATHEMATICUM (BRNO)
Tomus 55 (2019), 167–194

A VISCOSITY-PROXIMAL GRADIENT METHOD WITH INERTIAL EXTRAPOLATION FOR SOLVING CERTAIN MINIMIZATION PROBLEMS IN HILBERT SPACE

L.O. Jolaoso, H.A. Abass, and O.T. Mewomo

Abstract. In this paper, we study the strong convergence of the proximal gradient algorithm with inertial extrapolation term for solving classical minimization problem and finding the fixed points of δ-demimetric mapping in a real Hilbert space. Our algorithm is inspired by the inertial proximal point algorithm and the viscosity approximation method of Moudafi. A strong convergence result is achieved in our result without necessarily imposing the summation condition $\sum_{n=1}^{\infty}\beta_n\|x_{n-1}-x_n\|<+\infty$ on the inertial term. Finally, we provide some applications and numerical example to show the efficiency and accuracy of our algorithm. Our results improve and complement many other related results in the literature.

1. Introduction

Let H be a real Hilbert space and C be a nonempty, closed and convex subset of H. Let $T\colon H\to H$ be a nonlinear mapping, a point $x\in H$ is called a fixed point of T if $Tx=x$. We denote the set of all fixed points of T by $F(T)$. Let $D(T)\subset H$, then T is said to be

(1) a contraction if there exists $\alpha\in[0,1)$ such that

$$\|Tx-Ty\|\le\alpha\|x-y\|\,,\quad \forall x,\,y\in D(T)\,.$$

If $\alpha=1$, then T is called a nonexpansive mapping;

(2) quasi-nonexpansive if $F(T)\neq\emptyset$ and

$$\|Tx-p\|\le\|x-p\|\,,\quad x\in D(T)\quad\text{and}\quad p\in F(T)\,;$$

(3) firmly nonexpansive if for all $x,y\in D(T)$, we have

$$\|Tx-Ty\|^2\le\langle Tx-Ty,x-y\rangle\,;$$

2010 *Mathematics Subject Classification*: primary 47H10; secondary 46N10, 47J25, 65K10, 65K15.

Key words and phrases: proximal gradient algorithm, proximal operator, demimetric mappings, inertial algorithm, viscosity approximation, Meir Keeler contraction, fixed point theory.

Received September 27, 2018, revised February 2019. Editor M. Feistauer.

DOI: 10.5817/AM2019-3-167

(4) β-inverse strongly monotone (shortly β-ism) if there exists $\beta > 0$ such that

$$\langle x-y, Tx-Ty\rangle \geq \beta\|Tx-Ty\|^2\,, \quad \forall x,\, y \in D(T)\,;$$

(5) k-strictly pseudo-contraction if there exists $k \in [0,1)$ such that

$$\|Tx-Ty\|^2 \leq \|x-y\|^2 + k\|x-Tx-(y-Ty)\|^2\,, \quad \forall x,\, y \in D(T)\,;$$

(6) δ-demimetric, if there exist $\delta \in (-\infty, 1)$ such that

$$\langle x-p, x-Tx\rangle \geq \frac{1-\delta}{2}\|x-Tx\|^2\,, \quad \forall x \in D(T) \quad \text{and} \quad p \in F(T)\,. \tag{1.1}$$

Equivalently, T is δ-demimetric, if there exists $\delta \in (-\infty, 1)$ such that

$$\|Tx-p\|^2 \leq \|x-p\|^2 + \delta\|x-Tx\|^2\,, \quad \forall x \in D(T) \quad \text{and} \quad p \in F(T)\,. \tag{1.2}$$

It is easy to see that every firmly nonexpansive mapping is 1-ism. The class of δ-demimetric was recently introduced by Takahashi [46] as a generalization of k-strictly pseudo-contraction, firmly nonexpansive, quasi-nonexpansive and nonexpansive mappings in a real Hilbert space.

We give the following examples of δ-demimetric mapping in real Hilbert space.

Example 1.1. Let $H = \mathbb{R}$ (the real line with usual metric). Define $T\colon \mathbb{R} \to \mathbb{R}$ by $Tx = \frac{x}{2}$, for all $x \in \mathbb{R}$. Clearly, $F(T) = \{0\}$. Thus

$$\begin{aligned}
\langle x-p, x-Tx\rangle &= \langle x-0, x-\frac{x}{2}\rangle = \langle x, \frac{x}{2}\rangle = \frac{1}{2}\langle x, x\rangle \\
&= \frac{1}{2}|x|^2 \geq |\frac{x}{2}|^2 \\
&= \frac{1-\delta}{2}|\frac{x}{2}|\frac{1-\delta}{2}|x-Tx|^2\,,
\end{aligned}$$

where $\delta = -1$. From (1.1), we see that T is -1-demimetric.

Example 1.2. Let H be the real line and $C = [-2, 1]$. Define

$$Tx = \begin{cases} \dfrac{x+9}{10}\,, & x \in [0,1]\,, \\ \dfrac{3+x}{4}\,, & x \in [-2, 0)\,. \end{cases}$$

Obviously, $F(T) = \{1\}$. We will show that there exists $\delta \in (-\infty, 1)$ such that

$$|Tx-1|^2 \leq |x-1|^2 + \delta|x-Tx|^2\,, \qquad \forall x \in [-2, 1]\,.$$

Consider the following two cases:

Case (i): Let $x \in [0, 1]$, then

$$|x-Tx|^2 = \left|x - \frac{x+9}{10}\right|^2 = \left|\frac{9}{10}(x-1)\right|^2 = \frac{81}{100}|x-1|^2.$$

Also

$$\begin{aligned}|Tx-1|^2 &= \left|\frac{x+9}{10}-1\right|^2 = \frac{1}{100}|x-1|^2\\ &= |x-1|^2 - \frac{99}{100}|x-1|^2\\ &= |x-1|^2 - \frac{99}{81}\times\frac{81}{100}|x-1|^2\\ &\le |x-1|^2 + \delta_1 . \frac{81}{100}|x-1|^2 ,\end{aligned}$$

for any $\delta_1 \in [-\frac{99}{81}, 1)$. Hence $|Tx-1|^2 \le |x-1|^2 + \delta_1|x-Tx|^2$.

Case (ii): Let $x \in [-2, 0)$, thus

$$|x-Tx|^2 = \left|x - \frac{3+x}{4}\right|^2 = \left|\frac{3(x-1)}{4}\right|^2 = \frac{9}{16}|x-1|^2 .$$

Then

$$\begin{aligned}|Tx-1|^2 &= \left|\frac{3+x}{4}-1\right|^2 = \left|\frac{x-1}{4}\right|^2 = \frac{1}{16}|x-1|^2\\ &= |x-1|^2 - \frac{15}{16}|x-1|^2\\ &= |x-1|^2 - \frac{15}{9}\cdot\frac{9}{16}|x-1|^2\\ &\le |x-1|^2 + \delta_2\cdot\frac{9}{16}|x-1|^2 ,\end{aligned}$$

for any $\delta_2 \in [-\frac{15}{9}, 1)$. Hence $|Tx-1|^2 \le |x-1|^2 + \delta_1|x-Tx|^2$. In particular, choose $\delta = \min\{\delta_1, \delta_2\}$. Thus, T is $-\frac{15}{9}$-demimetric.

Consider the following minimization problem

$$\text{minimize } \big\{g(x) + h(x)\big\} , \tag{1.3}$$

where $h\colon H \to \mathbb{R}\cup\{+\infty\}$ is a proper, closed and convex function which is possibly nonsmooth and $g\colon H \to \mathbb{R}$ is a proper, closed, convex and continuously differentiable function and its gradient $\nabla g(\cdot)$ is Lipschitz continuous on H, i.e. there exists a constant $\alpha > 0$ such that

$$\|\nabla g(x) - \nabla g(y)\| \le \alpha\|x-y\| , \quad \forall x,\, y \in H .$$

Throughout this paper, we assume that Problem (1.3) has a solution and denote its set of solutions by Ω. The Proximal Gradient Method (PGM) which has been effective in approximating solutions of (1.3) can be formulate as follows: Given the initial point $x_1 \in H$, compute

$$x_{n+1} = \mathrm{prox}_{\gamma_n h}(x_n - \gamma_n\nabla g(x_n)) , \quad n \ge 1 , \tag{1.4}$$

where $\mathrm{prox}_{\gamma_n h}(x) := \underset{u\in H}{\mathrm{argmin}}\{h(x) + \frac{1}{2\gamma_n}||x-u||^2\}$ and $\gamma_n > 0$ is a stepsize. The $prox_{\gamma h}$ operator is firmly nonexpansive and when $g = 0$ in (1.3), the PGM reduces to the classical proximal point algorithm, see [18]. The PGM can be shown to converge with rate $O(\frac{1}{k})$ when a fixed stepsize $\gamma_n = \gamma \in (0, \frac{1}{\alpha}]$ is used (see [14, 37]).

If α is unknown, the stepsize γ_n can be found by line searching method (see [3]). More so, if the condition

$$0 < \liminf_{n\to\infty} \gamma_n \leq \limsup_{n\to\infty} \gamma_n < \frac{2}{\alpha}$$

is satisfied, then the sequence $\{x_n\}$ converges weakly to a point in Ω. The PGM can also be interpreted as a fixed point iteration. A point x^* is a solution of (1.3) if and only if it is a fixed point of the operator $\mathrm{prox}_{\gamma h}(I - \gamma\nabla g)$ (see Section 4.2.1 in [37] and Proposition 3.2 in [50]).

When $h = I_C$ (the indicator function on a nonempty closed convex subset of H), the PGM reduces to the well known gradient projection algorithm which is defined as follows. For an initial guess $x_1 \in H$,

$$x_{n+1} = P_C\big(x_n - \gamma_n \nabla g(x_n)\big)\,, \quad n \geq 1\,, \tag{1.5}$$

where P_C is the metric projection from H onto C. The convergence of algorithm (1.5) depends on the behaviour of the gradient ∇g. It is known that if ∇g is ν-strongly monotone operator, i.e. there exists $\alpha > 0$ such that

$$\langle \nabla g(x) - \nabla g(y), x - y\rangle \geq \nu\|x - y\|^2, \quad \forall\ x,\ y \in C\,,$$

then, the operator $T := P_C(I - \gamma\nabla g)$ is a contraction; hence, the sequence $\{x_n\}$ defined by (1.5) converges strongly to a solution of (1.3) for $h = I_C$. More general, if the sequence $\{\gamma_n\}$ is chosen to satisfy the property

$$0 < \liminf_{n\to\infty} \gamma_n \leq \limsup_{n\to\infty} \gamma_n < \frac{2\nu}{\alpha^2}\,,$$

then the sequence $\{x_n\}$ defined by (1.5) converges in norm to the unique solution of (1.3) for $h = I_C$. However, if the gradient ∇g fails to be strongly monotone, then the operator $T := P_C(I - \gamma\nabla g)$ would fail to be a contraction. Consequently, the sequence $\{x_n\}$ generated by (1.5) may fail to converge strongly (see Section 4 in [49]). The gradient projection algorithm (1.5) has been studied extensively by many authors, see for instance [8, 9, 19, 20, 21, 44, 49] and reference therein.

In 2000, Moudafi [29] introduced the viscosity approximation method for approximating fixed points of nonexpansive mappings. Let f be a contraction on H, starting with an arbitrary $x_0 \in H$, define a sequence $\{x_n\}$ recursively by

$$x_{n+1} = \lambda_n f(x_n) + (1 - \lambda_n)Tx_n\,, \quad n \geq 0\,, \tag{1.6}$$

where $\{\lambda_n\}$ is a sequence in $(0, 1)$. Xu [48] proved that if $\{\lambda_n\}$ satisfies some certain conditions, the sequence $\{x_n\}$ generated by (1.6) converges strongly to the unique solution $x^\dagger \in F(T)$ of the variational inequality

$$\langle (I - f)x^\dagger, x - x^\dagger\rangle \geq 0\,, \quad \forall x \in F(T)\,.$$

Also, based on the heavy ball methods of the order-two time dynamical system, Polyak [39] first proposed an inertial extrapolation as an acceleration process to solve the smooth convex minimization. The inertial algorithm is a two-step iteration where the next iterate is defined by making use of the previous two iterates. Recently, alot of researchers have constructed some fast iterative algorithm by using inertial extrapolation which includes inertial proximal method [2], inertial forward-backward method [23], inertial proximal ADMM [12] and fast iterative shrinkage thresholding algorithm FISTA [4, 11]. Using the technique of inertial extrapolation, in 2008, Mainge [24] introduced the following inertial Mann algorithm:

$$\begin{cases} y_n = x_n + \beta_n(x_n - x_{n-1})\,, \\ x_{n+1} = (1-\lambda_n)y_n + \lambda_n T y_n\,, \end{cases} \tag{1.7}$$

for each $n \geq 1$. Mainge [24] showed that the iterative sequence $\{x_n\}$ converges weakly to a fixed point of T under the following conditions:

(A1) $\beta_n \in [0, \alpha)$ for each $n \geq 1$, where $\alpha \in [0, 1)$;

(A2) $\sum_{n=1}^{\infty} \beta_n \|x_n - x_{n-1}\|^2 < +\infty$;

(A3) $0 < \inf \lambda_n \leq \sup \lambda_n < 1$.

Moudafi and Oliny [30] proposed the following inertial proximal point algorithm for finding the zero point of the sum of two monotone operators in real Hilbert space: for $x_1 \in H$,

$$\begin{cases} y_n = x_n + \beta_n(x_n - x_{n-1})\,, \\ x_{n+1} = (I + \lambda_n B)^{-1}(y_n - \lambda_n A x_n)\,, \quad n \geq 1\,, \end{cases}$$

where $A\colon H \to H$ and $B\colon H \to 2^H$ are monotone operators and $\{\beta_n\} \subset [0, 1)$. They obtained a weak convergence theorem provided $0 < \lambda_n < \frac{2}{c}$ with c being the Lipschitz constant of A and that the condition (A2) holds.

Note that for the condition (A2) to be satisfied, one needs to first calculate β_n at each step of the iterations (see [30]). Other iterative methods involving the inertial extrapolation process which have been introduced include the works of Beck and Teboulle [4], Bot et.al [5, 6] and Pesquet and Putselnik [38].

Recently, Chembolle and Dossel [11] proved the weak convergence of the following modified PGM with inertial extrapolation term in a real Hilbert space

$$\begin{cases} x_n = T(y_{n-1}), \\ y_n = \big(1 - \frac{1}{t_{n+1}}\big)x_n + \frac{1}{t_{n+1}} u_n\,, \\ u_n = x_{n-1} + t_n(x_n - x_{n-1})\,, \qquad n \geq 1\,, \end{cases} \tag{1.8}$$

equivalently, (1.8) can be written as

$$\begin{aligned} & x_n = T(y_{n-1})\,, \\ & y_n = x_n + \alpha_n(x_n - x_{n-1})\,, \qquad \alpha_n = \frac{t_{n-1}}{t_{n+1}}\,, \quad \text{for } n \geq 1\,, \end{aligned}$$

where $a > 2$ is a positive real number, $t_n = \frac{n+a-1}{a}$ for all $n \in \mathbb{N}$ and $Tx = \operatorname{prox}_{\gamma h}(x - \gamma \nabla g(x))$.

More recently, Guo and Cui [17] proposed the following PGM with perturbations for solving (1.3):

$$x_{n+1} = \alpha_n f(x_n) + (1-\alpha_n)\operatorname{prox}_{\gamma_n h}(I - \gamma_n \nabla g)x_n + e_n\,, \tag{1.9}$$

where $\{\alpha_n\} \subset [0, 1]$, $0 < a \leq \liminf_{n\to\infty} \gamma_n < \frac{2}{\alpha}$, $f\colon H \to H$ is a contraction and $e\colon H \to H$ is a perturbation operator satisfying $\sum_{n=0}^{\infty} \|e(x_n)\| < +\infty$. They obtained a strong convergence theorem for the sequence generated by (1.9) for approximating solution of (1.3) in a real Hilbert space.

Motivated by the above works, our interest in this paper is to introduce a new relaxed proximal gradient algorithm for approximating a common solution of the minimization problem (1.3) and fixed point of δ-demimetric mapping in a real Hilbert space. Our algorithm is developed by combining the proximal gradient algorithm (1.4) and the viscosity approximation method of Moudafi [29] with an inertial extrapolation term. We obtain a strong convergence result for the approximation of common solution of (1.3) and a fixed point of δ-demimetric mapping in a real Hilbert space. Finally, we give a numerical example to illustrate the effectiveness of our algorithm. Our results complement and improve some other related results in the literature.

2. Preliminaries

In this section, we give some basic definitions and results which will be used in the sequel. We denote the strong convergence of $\{x_n\}$ to z by $x_n \to z$ and the weak convergence of $\{x_n\}$ to z by $x_n \rightharpoonup z$.

Let H be a real Hilbert space and C be a nonempty, closed and convex subset of H. Recall that the metric projection of $x \in H$ onto C is the necessarily unique vector $P_C x \in C$ satisfying

$$\|P_C x - x\| \leq \|x - y\|\,,\ \forall y \in C,\ x \in H\,.$$

It is well known that P_C satisfies the following property,

$$\langle x - y, P_C x - P_C y\rangle \geq \|P_C x - P_C y\|^2\,,\ \forall x,\ y \in H\,.$$

Also, the metric projection have the following characterization.

Lemma 2.1. *Let H be a real Hilbert space and let C be a nonempty, closed and convex subset of H. Then for $x \in H$ and $w \in C$, the following conditions are equivalent:*

(i) $w = P_C(x)$;

(ii) $\langle x - w, y - w\rangle \leq 0$, *for all* $y \in C$;

(iii) $\|x - w\|^2 + \|y - w\|^2 \leq \|x - y\|^2$ *for all* $y \in C$.

Lemma 2.2. *Let H be a real Hilbert space. Then the following hold: for all x, $y \in H$,*

(i) $\|x + y\|^2 \leq \|y\|^2 + 2\langle x, x + y\rangle$;

(ii) $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x, y\rangle$;

(iii) $\|\alpha x + (1-\alpha)y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x - y\|^2$, *for* $\alpha \in (0, 1)$.

Let (X, d) be a complete metric space. A mapping $f\colon X \to X$ is called a Meir-Keeler contraction [27] if for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$d(x, y) < \epsilon + \delta \quad \text{implies} \quad d\big(f(x), f(y)\big) < \epsilon\,,$$

for all x, $y \in X$. It is well-known that the Meir-Keeler contraction is a generalization of the contraction.

Lemma 2.3 ([27])**.** *A Meir-Keeler contraction defined on a complete metric space has a unique fixed point.*

Lemma 2.4 ([45])**.** *Let f be a Meir-Keeler contraction on a convex subset C of a Banach space E. Then for every $\epsilon > 0$, there exists $r_\epsilon \in (0,1)$ such that*

$$\|x-y\| \geq \epsilon \;\Rightarrow\; \|f(x)-f(y)\| \leq r_\epsilon \|x-y\|$$

for all x, $y \in C$.

A point $x^* \in C$ is said to be an asymptotic fixed point of T if C contains a sequence $\{x_n\}$ which converges weakly to x^* and $\lim_{n\to\infty} \|x_n - Tx_n\| = 0$. The set of asymptotic fixed points of T is denoted by $\hat{F}(T)$.

A mapping $T\colon H \to H$ is said to be an α-averaged mapping if $T = (1-\alpha)I + \alpha S$, where $\alpha \in (0,1)$ and $S\colon H \to H$ is nonexpansive. Many nonlinear operators belong to the class of averaged mapping. For instance, the class of firmly nonexpansive mapping is $\frac{1}{2}$-averaged. The following lemmas will be used in the sequel.

Lemma 2.5 ([7, 13])**.** *Let S, T, $: H \to H$ be given nonlinear operators:*

(i) *If $T = (1-\alpha)S + \alpha V$, for some $\alpha \in (0,1)$ and if S is averaged and V is nonexpansive, then T is averaged.*

(ii) *The composition of finitely many averaged mapping is averaged. In particular, if T_1 is α_1-averaged and T_2 is α_2-averaged, where $\alpha_1, \alpha_2 \in (0,1)$, then, the composition T_1T_2 is α-averaged, where $\alpha = \alpha_1 + \alpha_2 - \alpha_1\alpha_2$.*

(iii) *If $\{T_i\}$ is a finite family of averaged mappings and have a common fixed point, then*

$$\bigcap_{i=1}^{N} F(T_i) = F(T_1 \dots T_N)\,.$$

Lemma 2.6 ([7, 26])**.** *Let $U\colon H \to H$ be a given operator, we have*

(i) *U is nonexpansive if and only if the complement $I - U$ is $\frac{1}{2}$-ism.*

(ii) *If U is κ-ism, then for $\gamma > 0$, κU is $\frac{\kappa}{\gamma}$-ism.*

(iii) *U is averaged if and only if the complement $I - U$ is κ-ism for some $\kappa > \frac{1}{2}$. Indeed, for $\alpha \in (0,1)$, U is averaged if and only if $I - U$ is $\frac{1}{2\alpha}$-ism.*

Lemma 2.7 ([16] (Demiclosedness Principle))**.** *Let C be a closed and convex subset of a Hilbert space H and $T\colon C \to C$ be a nonexpansive mapping with $F(T) \neq \emptyset$. If $\{x_n\}$ is a sequence in C weakly converging to p and if $\{(I-T)x_n\}$ converges strongly to q, then $(I-T)p = q$. In particular, if $q = 0$, then $p \in F(T)$.*

Lemma 2.8 ([24])**.** *Let $\{\alpha_n\}$ and $\{\gamma_n\}$ be sequences of nonnegative real numbers such that*

$$\alpha_{n+1} \leq (1-\delta_n)\alpha_n + \beta_n + \gamma_n\,, \qquad n \geq 1\,,$$

where $\{\delta_n\}$ is a sequence in $(0,1)$ and $\{\beta_n\}$ is a real sequence. Assume that $\sum_{n=0}^{\infty} \beta_n < \infty$. Then, the following results hold:

(i) *If $\beta_n \leq \delta_n M$ for some $M \geq 0$, then $\{\alpha_n\}$ is a bounded sequence.*

(ii) *If* $\sum_{n=0}^{\infty} \delta_n = \infty$ *and* $\limsup_{n\to\infty} \dfrac{\beta_n}{\delta_n} \le 0$, *then* $\lim_{n\to\infty} \alpha_n = 0$.

Lemma 2.9 ([25])**.** *Let $\{a_n\}$ be a sequence of real numbers such that there exists a subsequence $\{n_i\}$ of $\{n\}$ with $a_{n_i} < a_{n_i+1}$ for all $i \in \mathbb{N}$. Consider the integer $\{m_k\}$ defined by*

$$m_k = \max\{j \le k : a_j < a_{j+1}\}\,.$$

Then $\{m_k\}$ is a nondecreasing sequence verifying $\lim_{n\to\infty} m_n = \infty$, and for all $k \in \mathbb{N}$, the following estimate hold:

$$a_{m_k} \le a_{m_k+1} \quad \textit{and} \quad a_k \le a_{m_k+1}\,.$$

3. Main result

In this section, we modify the proximal gradient algorithm combine with an inertial extrapolation term and prove a strong convergence theorem for approximating solution of (1.3) and fixed point of δ-demimetric mapping in a real Hilbert space. First, we proof the following lemma which plays a crucial role in the proof of the main theorem.

Lemma 3.1. *Assume that the minimization problem* (1.3) *is consistent and gradient ∇g is Lipschitz continuous with Lipschitz constant $L > 0$. Let $\gamma > 0$ such that $0 < \gamma < \frac{2}{L}$, then the following inequality holds:*

$$\|\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x - x\|^2 \le 2\big\langle x-y, x-\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x\big\rangle\,. \tag{3.1}$$

Proof. Since $\operatorname{prox}_{\gamma h}$ is firmly nonexpansive, then it is $\frac{1}{2}$-averaged. Also, the Lipschitz condition on ∇g implies that ∇g is $\frac{1}{L}$-ism and by Lemma 2.6(ii), $\gamma\nabla g$ is $\frac{1}{\gamma L}$-ism. Hence, by Lemma 2.6(iii), we have that $I-\gamma\nabla g$ is $\frac{\gamma L}{2}$-averaged. It follows from Lemma 2.5(ii) that the $\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)$ is averaged with constant $\frac{2+\gamma L}{4}$. In particular, $\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)$ is nonexpansive. Then, for any $x \in C$ and $y \in \Omega$, we have

$$\begin{aligned}
\|\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x - y\|^2 &= \|\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x - \operatorname{prox}_{\gamma h}(I-\gamma\nabla g)y\|^2 \\
&\le \|x-y\|^2 \\
&= \langle x-y, x-\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x \\
&\quad + \operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x - y\rangle \\
&= \langle x-y, x-\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x\rangle \\
&\quad + \langle x-y, \operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x - y\rangle\,.
\end{aligned}$$

This implies that

$$\langle \operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x - x, \operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x - y\rangle \le \langle x-y, x-\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x\rangle\,.$$

Thus

$$\begin{aligned}
&\langle \operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x - x, \operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x - x + x - y\rangle \\
&\qquad\qquad \le \langle x-y, x-\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x\rangle\,,
\end{aligned}$$

which gives that

$$\begin{aligned}\langle \operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x - x, \operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x-x\rangle &\leq \langle x-y, x-\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x\rangle \\ &+ \langle x-y, x-\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x\rangle,\end{aligned}$$

therefore

$$\|\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x - x\|^2 \leq 2\big\langle x-y, x-\operatorname{prox}_{\gamma h}(I-\gamma\nabla g)x\big\rangle\,.$$

□

Theorem 3.2. *Let C be a nonempty, closed and convex subset of a real Hilbert space H. Let g, $h\colon H \to \mathbb{R}\cup\{+\infty\}$ be two proper convex lower semicontinuous functions such that h is nonsmooth and the gradient ∇g is $\frac{1}{L}$-ism with $L>0$. Let $f\colon C\to C$ be a Meir Keeler contraction mapping, $B\colon C\to H$ be a strongly positive bounded linear operator with coefficient $\tau>0$ such that $0<\xi<\frac{\tau}{2}$ and $T\colon C\to C$ be a δ-demimetric mapping for $\delta\in(-\infty,1)$ and $\hat{F}(T)=F(T)$. Suppose $\Gamma=\Omega\cap F(T)\neq\emptyset$, let $\alpha_n\in[0,1]$, $\beta_n\in[0,1)$, w_n, $\theta_n\in(0,1)$ and $\gamma_n>0$. Choose initial points x_0, $x_1\in H$ arbitrarily and let $\{x_n\}$, $\{y_n\}$ and $\{u_n\}$ be generated by*

$$\begin{cases} y_n = x_n + \beta_n(x_n - x_{n-1})\,, \\ u_n = (1-w_n)y_n + w_n \operatorname{prox}_{\gamma_n h}(y_n - \gamma_n\nabla g(y_n))\,, \\ x_{n+1} = P_C\big(\alpha_n\xi f(x_n) + \theta_n x_n + ((1-\theta_n)I - \alpha_n B)T_{\lambda_n}u_n\big)\,, \quad n\geq 1\,, \end{cases} \tag{3.2}$$

where $T_{\lambda_n}=(1-\lambda_n)I+\lambda_n T$ for $\lambda_n\in(0,1)$. Assume that the following conditions are satisfied:

(C1) $\lim_{n\to\infty}\alpha_n = 0$ *and* $\sum_{n=1}^{\infty}\alpha_n=\infty$,

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

(C3) $0<\liminf_{n\to\infty} w_n \leq \limsup_{n\to\infty} w_n < 1$,

(C4) $0<\liminf_{n\to\infty} \gamma_n \leq \limsup_{n\to\infty} \gamma_n < \frac{2}{L}$,

(C5) $0<\liminf_{n\to\infty} \lambda_n \leq \limsup_{n\to\infty} \lambda_n < 1-\delta$.

Then, $\{x_n\}$ converges strongly to a point $\bar{x}$, where $\bar{x}=P_\Gamma(I-B+\xi f)(\bar{x})$ is the unique solution of the variational inequality

$$\langle (B-\xi f)\bar{x}, \bar{x}-y\rangle \leq 0\,, \quad y\in\Gamma\,. \tag{3.3}$$

Proof. Firstly, we show that $\{x_n\}$ is bounded. Let $x^*\in\Gamma$ and a number $\varepsilon>0$. Suppose $\|x_n-x^*\|\leq\varepsilon$, then we can easily see that $\{x_n\}$ is bounded. On the other hand, let $\|x_n-x^*\|\geq\varepsilon$, then by Lemma 2.4, there exists $\rho_\varepsilon\in(0,1)$ such that

$$\|f(x_n)-f(x^*)\|\leq\rho_\varepsilon\|x_n-x^*\|\,. \tag{3.4}$$

From (3.2), we have

$$\begin{aligned}\|y_n-x^*\| &= \|x_n-x^*+\beta_n(x_n-x_{n-1})\| \\ &\leq \|x_n-x^*\|+\beta_n\|x_n-x_{n-1}\|\,. \end{aligned} \tag{3.5}$$

Also

$$\begin{aligned}\|u_n - x^*\|^2 &= \|(1-w_n)y_n + w_n \operatorname{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n - x^*\|^2\\ &= \|(y_n - x^*) + w_n(\operatorname{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n - y_n)\|^2\\ &= \|y_n - x^*\|^2 + 2w_n\langle y_n - x^*, \operatorname{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n - y_n\rangle\\ &\quad + w_n^2\|\operatorname{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n - y_n\|^2, \end{aligned}\tag{3.6}$$

and from Lemma 3.1, we have that

$$\begin{aligned}\|u_n - x^*\|^2 &\le \|y_n - x^*\|^2 - w_n(1-w_n)\|\operatorname{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n - y_n\|^2\\ &\le \|y_n - x^*\|^2 . \end{aligned}\tag{3.7}$$

Moreover, from the definition of δ-demimetric (1.1), we have

$$\begin{aligned}\|T_{\lambda_n}u_n - x^*\|^2 &= \|(u_n - x^*) + \lambda_n(Tu_n - u_n)\|^2\\ &= \|u_n - x^*\|^2 - 2\lambda_n\langle u_n - x^*, u_n - Tu_n\rangle + \lambda_n^2\|u_n - Tu_n\|^2\\ &\le \|u_n - x^*\|^2 - \lambda_n(1-\delta)\|u_n - Tu_n\|^2 + \lambda_n^2\|u_n - Tu_n\|^2\\ &= \|u_n - x^*\|^2 - \lambda_n(1-\delta-\lambda_n)\|u_n - Tu_n\|^2 , \end{aligned}\tag{3.8}$$

and by condition (C5), we get

$$\|T_{\lambda_n}u_n - x^*\|^2 \le \|u_n - x^*\|^2 . \tag{3.9}$$

Thus, we have from (3.2) that

$$\begin{aligned}\|x_{n+1} - x^*\| &= \|P_C(\alpha_n\xi f(x_n) + \theta_n x_n + ((1-\theta_n)I - \alpha_n B)T_{\lambda_n}u_n) - P_C x^*\|\\ &\le \|\alpha_n\xi f(x_n) + \theta_n x_n + ((1-\theta_n)I - \alpha_n B)T_{\lambda_n}u_n - x^*\|\\ &= \|\alpha_n(\xi f(x_n) - Bx^*) + \theta_n(x_n - x^*) + ((1-\theta_n)I - \alpha_n B)(T_{\lambda_n}u_n - x^*)\|\\ &\le \alpha_n(\xi\|f(x_n) - f(x^*)\| + \|\xi f(x^*) - Bx^*\|) + \theta_n\|x_n - x^*\|\\ &\quad + ((1-\theta_n)I - \alpha_n\tau)\|T_{\lambda_n}u_n - x^*\|\\ &\le \alpha_n\xi\rho_\varepsilon\|x_n - x^*\| + \alpha_n\|\xi f(x^*) - Bx^*\| + \theta_n\|x_n - x^*\|\\ &\quad + ((1-\theta_n)I - \alpha_n\tau)\|u_n - x^*\|\\ &\le \alpha_n\xi\rho_\varepsilon\|x_n - x^*\| + \alpha_n\|\xi f(x^*) - Bx^*\| + \theta_n\|x_n - x^*\|\\ &\quad + ((1-\theta_n)I - \alpha_n\tau)[\|x_n - x^*\| + \beta_n\|x_n - x_{n-1}\|]\\ &= (1-\alpha_n(\tau - \xi\rho_\varepsilon))\|x_n - x^*\| + \alpha_n\|\xi f(x^*) - Bx^*\|\\ &\quad + ((1-\theta_n)I - \alpha_n\tau)\beta_n\|x_n - x_{n-1}\|\\ &= (1-\alpha_n(\tau - \xi\rho_\varepsilon))\|x_n - x^*\| + \alpha_n(\tau - \xi\rho_\varepsilon)\\ &\quad \times\Big\{\frac{\xi f(x^*) - Bx^*\|}{\tau - \xi\rho_\varepsilon} + \frac{((1-\theta_n)I - \alpha_n\tau)\beta_n\|x_n - x_{n-1}\|}{\alpha_n(\tau - \xi\rho_\varepsilon)}\Big\}. \end{aligned}\tag{3.10}$$

Putting

$$\sigma_n = \Big(\frac{(1-\theta_n)I - \alpha_n\tau}{\tau - \xi\rho_\varepsilon}\Big)\frac{\beta_n}{\alpha_n}\|x_n - x_{n-1}\| ,$$

from condition (C2), it is easy to see that $\lim_{n\to\infty}\sigma_n = 0$, which implies that the sequence $\{\sigma_n\}$ is bounded. Let

$$M = \max\left\{\frac{\|\xi f(x^*) - Bx^*\|}{\tau - \xi\rho_\varepsilon}, \sup_{n\in\mathbb{N}}\sigma_n\right\},$$

by using Lemma 2.8(i) and (3.10), we have that the sequence $\{\|x_n - x^*\|\}$ is bounded. This shows that $\{x_n\}$ is bounded and consequently, $\{u_n\}$ and $\{y_n\}$ are bounded. Note that

$$\begin{aligned}\|y_n - x^*\|^2 &= \|x_n - x^* + \beta_n(x_n - x_{n-1})\|^2\\ &= \|x_n - x^*\| + 2\beta_n\langle x_n - x^*, x_n - x_{n-1}\rangle + \beta_n^2\|x_n - x_{n-1}\|^2\,,\end{aligned} \tag{3.11}$$

and from Lemma 2.2(ii), we have

$$2\langle x_n - x^*, x_n - x_{n-1}\rangle = \|x_n - x^*\|^2 + \|x_n - x_{n-1}\|^2 - \|x_{n-1} - x^*\|^2\,, \tag{3.12}$$

therefore, by substituting (3.12) into (3.11), we have

$$\begin{aligned}\|y_n - x^*\|^2 = \|x_n - x^*\|^2 + \beta_n\big[\|x_n - x^*\|^2 + \|x_n - x_{n-1}\|^2 - \|x_{n-1} - x^*\|^2\big]\\ + \beta_n^2\|x_n - x_{n-1}\|^2 \le \|x_n - x^*\|^2 + \beta_n\big[\|x_n - x^*\|^2 - \|x_{n-1} - x^*\|^2\big]\\ + 2\beta_n\|x_n - x_{n-1}\|^2\,.\end{aligned} \tag{3.13}$$

Now, put $m_n = \alpha_n\xi f(x_n) + \theta_n x_n + ((1-\theta_n)I - \alpha_n B)T_{\lambda_n}u_n$, using Lemma 2.2(i) and (3.2), we have

$$\begin{aligned}\|x_{n+1} - x^*\|^2 &\le \|\alpha_n\xi f(x_n) + \theta_n x_n + ((1-\theta_n)I - \alpha_n B)T_{\lambda_n}u_n - x^*\|^2\\
&= \|\alpha_n(\xi f(x_n) - Bx^*) + \theta_n(x_n - x^*)\\
&\quad + ((1-\theta_n)I - \alpha_n B)(T_{\lambda_n}u_n - x^*)\|^2\\
&\le \|((1-\theta_n)I - \alpha_n B)(T_{\lambda_n}u_n - x^*) + \theta_n(x_n - x^*)\|^2\\
&\quad + 2\alpha_n\langle\xi f(x_n) - Bx^*, m_n - x^*\rangle\\
&= \|((1-\theta_n)I - \alpha_n B)(T_{\lambda_n}u_n - x^*)\|^2 + \theta_n^2\|x_n - x^*\|^2\\
&\quad + 2\theta_n\big\langle((1-\theta_n)I - \alpha_n B)(T_{\lambda_n}u_n - x^*), x_n - x^*\big\rangle\\
&\quad + 2\alpha_n\langle\xi f(x_n) - Bx^*, m_n - x^*\rangle\\
&\le ((1-\theta_n)I - \alpha_n\tau)^2\|T_{\lambda_n}u_n - x^*\|^2 + \theta_n^2\|x_n - x^*\|^2\\
&\quad + 2\theta_n((1-\theta_n)I - \alpha_n\tau)\|T_{\lambda_n}u_n - x^*\|\,\|x_n - x^*\|\\
&\quad + 2\alpha_n\langle\xi f(x_n) - Bx^*, m_n - x^*\rangle\\
&\le ((1-\theta_n)I - \alpha_n\tau)^2\|T_{\lambda_n}u_n - x^*\|^2 + \theta_n^2\|x_n - x^*\|^2\\
&\quad + \theta_n((1-\theta_n)I - \alpha_n\tau)[\|T_{\lambda_n}u_n - x^*\|^2 + \|x_n - x^*\|^2]\\
&\quad + 2\alpha_n\langle\xi f(x_n) - Bx^*, m_n - x^*\rangle\\
&\le ((1-\theta_n)I - \alpha_n\tau)\|T_{\lambda_n}u_n - x^*\|^2 + \theta_n\|x_n - x^*\|^2\\
&\quad + 2\alpha_n\langle\xi f(x_n) - Bx^*, m_n - x^*\rangle\,.\end{aligned} \tag{3.14}$$

Thus, from (3.9) and (3.13), we have

$$\begin{aligned}
\|x_{n+1}-x^*\|^2 &\leq ((1-\theta_n)I-\alpha_n\tau)[||u_n-x^*||^2-\lambda_n(1-\lambda_n-\delta)\|u_n-Tu_n\|^2] \\
&\quad+\theta_n\|x_n-x^*\|^2+2\alpha_n\langle\xi f(x_n)-Bx^*,m_n-x^*\rangle \\
&\leq ((1-\theta_n)I-\alpha_n\tau) \\
&\quad\times\left\{\|x_n-x^*\|^2+\beta_n[\|x_n-x^*\|^2-\|x_{n-1}-x^*\|^2]\right. \\
&\quad\left.+2\beta_n\|x_n-x_{n-1}\|^2\right\}-\lambda_n(1-\lambda_n-\delta)\|u_n-Tu_n\|^2 \\
&\quad+\theta_n\|x_n-x^*\|^2+2\alpha_n\langle\xi f(x_n)-Bx^*,m_n-x^*\rangle \\
&\leq (1-\alpha_n\tau)\|x_n-x^*\|^2+\beta_n\left[\|x_n-x^*\|^2-\|x_{n-1}-x^*\|^2\right] \\
&\quad+2\beta_n\|x_n-x_{n-1}\|^2-\lambda_n(1-\lambda_n-\delta)\|u_n \\
&\quad-Tu_n\|^2+2\alpha_n\langle\xi f(x_n)-Bx^*,m_n-x^*\rangle\,. \qquad (3.15)
\end{aligned}$$

Set $D_n=\|x_n-x^*\|^2$ and consider the following two cases.

Case I: Suppose there exists a natural number N such that $D_{n+1}\leq D_n$ for all $n\geq\mathbb{N}$. In this case, $\{D_n\}$ is convergent. Since $\{x_n\}$ is bounded, it is easy to see that condition (C2) implies $\beta_n\|x_n-x_{n-1}\|\to 0$.
From (3.15), we have

$$\begin{aligned}
\lambda_n(1-\lambda_n-\delta)\|u_n-Tu_n\|^2 &\leq (1-\alpha_n\tau)\|x_n-x^*\|^2-\|x_{n+1}-x^*\|^2 \\
&\quad+\beta_n[\|x_n-x^*\|^2-\|x_{n-1}-x^*\|^2] \\
&\quad+2\beta_n\|x_n-x_{n-1}\|^2+2\alpha_n\langle\xi f(x_n)-Bx^*,m_n-x^*\rangle \\
&=(D_n-D_{n+1})+\beta_n(D_n-D_{n-1})+2\beta_n\|x_n-x_{n-1}\|^2 \\
&\quad-\alpha_n\tau D_n+2\alpha_n\langle\xi f(x_n)-Bx^*,m_n-x^*\rangle\,.
\end{aligned}$$

Since $\{D_n\}$ is convergent and $\alpha_n\to 0$, we have

$$\lim_{n\to\infty}\lambda_n(1-\lambda_n-\delta)\|u_n-Tu_n\|^2=0\,,$$

hence, by using condition (C5), we have

$$\lim_{n\to\infty}\|u_n-Tu_n\|=0\,. \qquad (3.16)$$

This implies that

$$\begin{aligned}
\lim_{n\to\infty}\|T_{\lambda_n}u_n-u_n\| &= \lim_{n\to\infty}\|(1-\lambda_n)u_n+\lambda_nTu_n-u_n\| \\
&= \lim_{n\to\infty}|\lambda_n|\,\|u_n-Tu_n\|=0\,. \qquad (3.17)
\end{aligned}$$

Also, from (3.7) and (3.15), we see that

$$\begin{aligned}
\|x_{n+1}-x^*\|^2 &\le ((1-\theta_n)I-\alpha_n\tau)[\|u_n-x^*\|^2-\lambda_n(1-\lambda_n-\delta)\|u_n-Tu_n\|^2] \\
&\quad +\theta_n\|x_n-x^*\|^2+2\alpha_n\langle \xi f(x_n)-Bx^*, m_n-x^*\rangle \\
&\le ((1-\theta_n)I-\alpha_n\tau)\|u_n-x^*\|^2+\theta_n\|x_n-x^*\|^2 \\
&\quad +2\alpha_n\langle \xi f(x_n)-Bx^*, m_n-x^*\rangle \\
&\le ((1-\theta_n)I-\alpha_n\tau)[\|y_n-x^*\|^2-w_n(1-w_n)\| \\
&\quad \times \operatorname{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n-y_n\|^2] \\
&\quad +\theta_n\|x_n-x^*\|^2+2\alpha_n\langle \xi f(x_n)-Bx^*, m_n-x^*\rangle \\
&\le ((1-\theta_n)I-\alpha_n\tau)\big\{\|x_n-x^*\|^2+\beta_n[\|x_n-x^*\|^2-\|x_{n-1}-x^*\|^2] \\
&\quad +2\beta_n\|x_n-x_{n-1}\|^2\big\} \\
&\quad -w_n(1-w_n)\|\operatorname{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n-y_n\|^2+\theta_n\|x_n-x^*\|^2 \\
&\quad +2\alpha_n\langle \xi f(x_n)-Bx^*, m_n-x^*\rangle \\
&\le (1-\alpha_n\tau)\|x_n-x^*\|^2+\beta_n[\|x_n-x^*\|^2-\|x_{n-1}-x^*\|^2] \\
&\quad +2\beta_n\|x_n-x_{n-1}\|^2-w_n(1-w_n)\|\operatorname{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n-y_n\|^2 \\
&\quad +2\alpha_n\langle \xi f(x_n)-Bx^*, m_n-x^*\rangle\,.
\end{aligned}$$

Therefore,

$$\begin{aligned}
w_n(1-w_n)\|\operatorname{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n-y_n\|^2 &\le (1-\alpha_n\tau)\|x_n-x^*\|^2 \\
&\quad -\|x_{n+1}-x^*\|^2+\beta_n[\|x_n-x^*\|^2-\|x_{n-1}-x^*\|^2] \\
&\quad +2\beta_n\|x_n-x_{n-1}\|^2+2\alpha_n\langle \xi f(x_n)-Bx^*, m_n-x^*\rangle \\
&= (D_n-D_{n+1})+\beta_n(D_n-D_{n-1})+2\beta_n\|x_n-x_{n-1}\|^2-\alpha_n\tau\|x_n-x^*\|^2 \\
&\quad +2\alpha_n\langle \xi f(x_n)-Bx^*, m_n-x^*\rangle\,.
\end{aligned}$$

Since $\{D_n\}$ is convergent and $\alpha_n\to 0$, we have that

$$\lim_{n\to\infty} w_n(1-w_n)\|\operatorname{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n-y_n||^2=0\,,$$

and by using condition (C3), we obtain

$$\lim_{n\to\infty}\|\operatorname{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n-y_n\|=0\,. \tag{3.18}$$

Furthermore, it is easy to see from (3.2) that

$$\|y_n-x_n\|\le \beta_n\|x_n-x_{n-1}\|\to 0\,,\quad \text{as}\quad n\to\infty\,, \tag{3.19}$$

and

$$\|u_n-y_n\|\le w_n\|\operatorname{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n-y_n\|\to 0,\quad \text{as}\quad n\to\infty,$$

hence

$$\lim_{n\to\infty}\|u_n-x_n\|\le \lim_{n\to\infty}(\|u_n-y_n\|+\|y_n-x_n\|)=0\,. \tag{3.20}$$

Also from (3.2), we have that

$$\lim_{n\to\infty}\|m_n-u_n\|\le\lim_{n\to\infty}\big(\alpha_n\|\xi f(x_n)-Bu_n\|+\theta_n\|x_n-u_n\| \\ +((1-\theta_n)I-\alpha_n\tau)\|T_{\lambda_n}u_n-u_n\|\big)=0\,,$$

then from (3.20), we get

$$\|m_n-x_n\|\le\|m_n-u_n\|+\|u_n-x_n\|\to 0\,,\quad \text{as}\quad n\to\infty\,. \tag{3.21}$$

More so, by the firmly nonexpansivity of the P_C and Lemma 2.1(iii), we have that

$$\begin{aligned}\|x_{n+1}-x^*\|^2&=\|P_Cm_n-P_Cx^*\|^2\\ &\le\|m_n-x^*\|^2-\|P_Cm_n-m_n\|^2\,.\end{aligned}\tag{3.22}$$

Substituting (3.15) into (3.22), we get

$$\|x_{n+1}-x^*\|^2\le(1-\alpha_n\tau)\|x_n-x^*\|^2+\beta_n\big[\|x_n-x^*\|^2-\|x_{n-1}-x^*\|^2\big] \\ +2\beta_n\|x_n-x_{n-1}\|^2+2\alpha_n\langle\xi f(x_n)-Bx^*,m_n-x^*\rangle-\|P_Cm_n-m_n\|^2\,,$$

therefore

$$\begin{aligned}\|P_Cm_n-m_n\|^2&\le(1-\alpha_n\tau)\|x_n-x^*\|^2-\|x_{n+1}-x^*\|^2\\ &\quad+\beta_n[\|x_n-x^*\|^2-\|x_{n-1}-x^*\|^2]\\ &\quad+2\beta_n\|x_n-x_{n-1}\|^2+2\alpha_n\langle\xi f(x_n)-Bx^*,m_n-x^*\rangle\\ &=(D_n-D_{n+1})+\beta_n(D_n-D_{n-1})-\alpha_n\tau\|x_n-x^*\|^2\\ &\quad+2\beta_n\|x_n-x_{n-1}\|^2+2\alpha_n\langle\xi f(x_n)-Bx^*,m_n-x^*\rangle\,,\end{aligned}$$

then

$$\lim_{n\to\infty}\|P_Cm_n-m_n\|=0\,. \tag{3.23}$$

Thus, we have from (3.21) and (3.18)

$$\lim_{n\to\infty}\|x_{n+1}-x_n\|=\lim_{n\to\infty}(\|x_{n+1}-m_n\|+\|m_n-x_n\|)=0\,. \tag{3.24}$$

Since $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_j}\}$ of $\{x_n\}$ such that $x_{n_j}\rightharpoonup\bar{x}\in C$. It follows from (3.19) and (3.20) that $y_{n_j}\rightharpoonup\bar{x}$ and $u_{n_j}\rightharpoonup\bar{x}$ respectively. Since $\mathrm{prox}_{\gamma_n h}(I-\gamma_n\nabla g)$ is nonexpansive and $\lim_{n\to\infty}\|y_n-\mathrm{prox}_{\gamma_n h}(I-\gamma_n\nabla g)y_n\|=0$, so by Lemma 2.7, we have that $\bar{x}\in F(\mathrm{prox}_{\gamma_n h}(I-\gamma_n\nabla g))$. Hence, $\bar{x}$ is a solution of the minimization problem (1.3), that is, $\bar{x}\in\Omega$. Also, since $\lim_{n\to\infty}\|u_n-Tu_n\|=0$ and $\hat{F}(T)=F(T)$, we have that $\bar{x}\in F(T)$. Therefore $\bar{x}\in\Omega\cap F(T)$.

We now show that $\limsup_{n\to\infty}\langle(B-\xi f)z,z-x_{n+1}\rangle\le 0$, where $z=P_\Gamma(I-B+\xi f)z$. Since $x_{n_j}\rightharpoonup\bar{x}$ and from Lemma 2.1(ii), we have

$$\begin{aligned}\limsup_{n\to\infty}\langle(B-\xi f)z,z-x_{n+1}\rangle&=\lim_{j\to\infty}\langle(B-\xi f)z,z-x_{n_j+1}\rangle\\ &=\langle(B-\xi f)z,z-\bar{x}\rangle\le 0.\end{aligned}\tag{3.25}$$

Next, we show that $x_n\to z$ as $n\to\infty$. Assume that $\{x_n\}$ does not converges strongly to z. Then, there exists $\epsilon>0$ and a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that

$\|x_{n_k} - z\| \geq \epsilon$ for all $k \in \mathbb{N}$ and by Lemma 2.4, there exists a number $r_\epsilon \in (0,1)$ such that

$$\|f(x_{n_k}) - f(z)\| \leq r_\epsilon \|x_{n_k} - z\| .$$

Thus, we have

$$\begin{aligned}
\|x_{n_k+1} - z\|^2 &= \langle P_C m_{n_k} - z, P_C m_{n_k} - z\rangle \\
&= \langle P_C m_{n_k} - m_{n_k} + m_{n_k} - z, P_C m_{n_k} - z\rangle \\
&= \langle P_C m_{n_k} - m_{n_k}, P_C m_{n_k} - z\rangle + \langle m_{n_k} - z, x_{n_k+1} - z\rangle \\
&\leq \langle m_{n_k} - z, x_{n_k+1} - z\rangle \\
&= \langle \alpha_{n_k} \xi f(x_{n_k}) + \theta_n x_{n_k} + ((1-\theta_{n_k})I - \alpha_{n_k} B)T\lambda_{n_k} u_{n_k} - z, x_{n_k+1} - z\rangle \\
&= \alpha_{n_k} \langle \xi f(x_{n_k}) - \xi f(z), x_{n_k+1} - z\rangle + \alpha_{n_k}\langle \xi f(z) - B(z), x_{n_k+1} - z\rangle \\
&\quad + \theta_{n_k}\langle x_{n_k} - z, x_{n_k+1} - z\rangle \\
&\quad + \langle ((1-\theta_{n_k})I - \alpha_{n_k} B)(T_{\lambda_{n_k}} u_{n_k} - z), x_{n_k+1} - z\rangle \\
&\leq \alpha_{n_k} \xi r_\epsilon \|x_{n_k} - z\| \, \|x_{n_k+1} - z\| + \theta_{n_k}\|x_{n_k} - z\| \, \|x_{n_k+1} - z\| \\
&\quad + ((1-\theta_{n_k})I - \alpha_{n_k}\tau)\|T_{\lambda_n} u_{n_k} - z\| \, \|x_{n_k+1} - z\| \\
&\quad + \alpha_{n_k}\langle \xi f(z) - Bz, x_{n_k+1} - z\rangle \\
&\leq \alpha_{n_k} \xi r_\epsilon \|x_{n_k} - z\| \, \|x_{n_k+1} - z\| + \theta_{n_k}\|x_{n_k} - z\| \, \|x_{n_k+1} - z\| \\
&\quad + ((1-\theta_{n_k})I - \alpha_{n_k}\tau)\|u_{n_k} - z\| \, \|x_{n_k+1} - z\| \\
&\quad + \alpha_{n_k}\langle \xi f(z) - Bz, x_{n_k+1} - z\rangle \\
&\leq \alpha_{n_k} \xi r_\epsilon \|x_{n_k} - z\| \, \|x_{n_k+1} - z\| + \theta_{n_k}\|x_{n_k} - z\| \, \|x_{n_k+1} - z\| \\
&\quad + ((1-\theta_{n_k})I - \alpha_{n_k}\tau)[\|x_{n_k} - z\| + \beta_n \|x_{n_k} - x_{n_k-1}\|]\|x_{n_k+1} - z\| \\
&\quad + \alpha_n\langle \xi f(z) - Bz, x_{n_k+1} - z\rangle \\
&= (1 - \alpha_{n_k}(\tau - \xi r_\epsilon))\|x_{n_k} - z\| \, \|x_{n_k+1} - z\| + ((1-\theta_{n_k})I - \alpha_{n_k}\tau) \\
&\quad \times \beta_n \|x_{n_k} - x_{n_k-1}\| \, \|x_{n_k+1} - z\| + \alpha_{n_k}\langle \xi f(z) - Bz, x_{n_k+1} - z\rangle \\
&\leq (1 - \alpha_{n_k}(\tau - \xi r_\epsilon))\frac{1}{2}(\|x_{n_k} - z\|^2 + \|x_{n_k+1} - z\|^2) \\
&\quad + ((1-\theta_{n_k})I - \alpha_{n_k}\tau)\beta_{n_k}\|x_{n_k} - x_{n_k-1}\| \, \|x_{n_k+1} - z\| \\
&\quad + \alpha_{n_k}\langle \xi f(z) - Bz, x_{n_k+1} - z\rangle .
\end{aligned}$$

This implies that

$$\begin{aligned}
\|x_{n_k+1} - z\|^2 &\leq \frac{(1 - \alpha_{n_k}(\tau - \xi r_\epsilon))}{1 + \alpha_{n_k}(\tau - \xi r_\epsilon)}\|x_{n_k} - z\|^2 + \frac{2((1-\theta_{n_k})I - \alpha_{n_k}\tau)\beta_{n_k}}{1 + \alpha_{n_k}(\tau - \xi r_\epsilon)} \\
&\quad \times \|x_{n_k} - x_{n_k-1}\| \, \|x_{n_k+1} - z\| + \frac{2\alpha_{n_k}}{1 + \alpha_{n_k}(\tau - \xi r_\epsilon)}\langle \xi f(z) - Bz, x_{n_k+1} - z\rangle \\
&\leq (1 - \alpha_{n_k}(\tau - \xi r_\epsilon))\|x_{n_k} - z\|^2 + \frac{2\beta_{n_k}}{1 + \alpha_{n_k}(\tau - \xi r_\epsilon)}\|x_{n_k} - x_{n_k+1}\| \, \|x_{n_k+1} - z\| \\
&\quad + \frac{2\alpha_{n_k}}{1 + \alpha_{n_k}(\tau - \xi r_\epsilon)}\langle \xi f(z) - Bz, x_{n_k+1} - z\rangle
\end{aligned}$$

$$
\begin{aligned}
&= (1-\alpha_{n_k}(\tau-\xi r_\epsilon))\|x_{n_k}-z\|^2 + \frac{2\alpha_n(\tau-\xi r_\epsilon)}{(1+\alpha_{n_k}(\tau-\xi r_\epsilon))(\tau-\xi r_\epsilon)} \\
&\quad \times \Big(\frac{\beta_{n_k}}{\alpha_{n_k}}\|x_{n_k}-x_{n_k-1}\|\,\|x_{n_k+1}-z\| + \langle \xi f(z)-Bz, x_{n_k+1}-z\rangle\Big) \\
&= (1-p_{n_k})\|x_{n_k}-z\|^2 + p_{n_k}q_{n_k}\,, \qquad (3.26)
\end{aligned}
$$

where $p_{n_k}=\alpha_{n_k}(\tau-\xi r_\epsilon)$ and $q_{n_k}=\Big(\frac{2\|x_{n_k+1}-z\|}{(1+\alpha_{n_k}(\tau-\xi r_\epsilon))(\tau-\xi r_\epsilon)}\Big)\frac{\beta_{n_k}}{\alpha_{n_k}}\|x_{n_k}-x_{n_k-1}\| + \frac{2}{(1+\alpha_{n_k}(\tau-\xi r_\epsilon))(\tau-\xi r_\epsilon)}\langle \xi f(z)-Bz, x_{n_k+1}-z\rangle$. Applying Lemma 2.8 and using conditions (C1), (C2) and (3.25), we conclude that the sequence $\{x_{n_k}\}$ converges strongly to z. The contradiction permits us to conclude that $x_n\to z$, where $z=P_\Gamma(I-B+\xi f)z$ which is the unique solution to the variational inequality (3.3).

Case II: Suppose there exists a subsequence $\{n_i\}$ of $\{n\}$ such that $D_{n_i}\le D_{n_i+1}$ for all $i\in\mathbb{N}$. Then, by Lemma 2.9, there exists a decreasing sequence $\{m_k\}\subset\mathbb{N}$ such that $m_k\to\infty$, $D_{m_k}<D_{m_k+1}$, for all $k\in\mathbb{N}$. Let $\epsilon>0$ and $\|x_{m_k}-x^*\|>\epsilon$, then, by Lemma 2.4, there exists $r_\epsilon\in(0,1)$ such that

$$
\|f(x_{m_k})-f(x^*)\|\le r_\epsilon\|x_{m_k}-x^*\|\,.
$$

Following similar argument as in Case I, we obtain $\|y_{m_k}-\mathrm{prox}_{\gamma_{m_k}h}(I-\gamma_{m_k}\nabla g)y_{m_k}\|\to 0$, $\|u_{m_k}-Tu_{m_k}\|\to 0$, $\|u_{m_k}-x_{m_k}\|\to 0$ and $\|x_{m_k+1}-x_{m_k}\|\to 0$. Since $\{x_{m_k}\}$ is bounded, there exists a subsequence of $\{x_{m_k}\}$ still denoted by $\{x_{m_k}\}$ which converges weakly to $\bar{x}$. Suppose $\{x_{m_k}\}$ is such that

$$
\limsup_{k\to\infty}\langle \xi f(x^*)-Bx^*, x_{m_k+1}-x^*\rangle = \lim_{k\to\infty}\langle \xi f(x^*)-Bx^*, x_{m_k+1}-x^*\rangle\,.
$$

It follows from Lemma 2.1 that

$$
\begin{aligned}
\limsup_{k\to\infty}\langle \xi f(x^*)-Bx^*, x_{m_k+1}-x^*\rangle &= \lim_{k\to\infty}\langle \xi f(x^*)-Bx^*, x_{m_k+1}-x^*\rangle \\
&= \langle \xi f(x^*)-Bx^*, \bar{x}-x^*\rangle \le 0\,.
\end{aligned}
$$

Hence

$$
\limsup_{k\to\infty}\langle \xi f(x^*)-Bx^*, x_{m_k+1}-x^*\rangle \le 0\,. \qquad (3.27)
$$

Similarly as in (3.26), we obtain

$$
\begin{aligned}
\|x_{m_k+1}-x^*\|^2 &\le (1-\alpha_{m_k}(\tau-\xi r_\epsilon))\|x_{m_k}-x^*\|^2 + \frac{2\alpha_{m_k}}{(1+\alpha_{m_k}(\tau-\xi r_\epsilon))} \\
&\quad \times \Big(\frac{\beta_{m_k}}{\alpha_{m_k}}\|x_{m_k}-x_{m_k+1}\|\,\|x_{m_k+1}-x^*\| + \langle \xi f(x^*)-Bx^*, x_{m_k+1}-x^*\rangle\Big)\,. \qquad (3.28)
\end{aligned}
$$

Since $D_{m_k} \leq D_{m_k+1}$, then from (3.28), we have

$$0 \leq \|x_{m_k+1} - x^*\|^2 - \|x_{m_k} - x^*\|^2 \leq (1 - \alpha_{m_k}(\tau - \xi r_\epsilon))\|x_{m_k} - x^*\|^2 + \frac{2\alpha_{m_k}}{(1 + \alpha_{m_k}(\tau - \xi r_\epsilon))}\Big(\frac{\beta_{m_k}}{\alpha_{m_k}}\|x_{m_k} - x_{m_k+1}\| \, \|x_{m_k+1} - x^*\| + \langle \xi f(x^*) - Bx^*, x_{m_k+1} - x^*\rangle\Big) - \|x_{m_k} - x^*\|^2.$$

This implies that

$$\alpha_{m_k}(\tau - \xi r_\epsilon)\|x_{m_k} - x^*\|^2 \leq \frac{2\alpha_{m_k}}{(1 + \alpha_{m_k}(\tau - \xi r_\epsilon))} \times \Big(\frac{\beta_{m_k}}{\alpha_{m_k}}\|x_{m_k} - x_{m_k+1}\| \, \|x_{m_k+1} - x^*\| + \langle \xi f(x^*) - Bx^*, x_{m_k+1} - x^*\rangle\Big). \tag{3.29}$$

Hence, from condition (C2) and (3.27), we obtain

$$\lim_{n\to\infty} \|x_{m_k} - x^*\| = 0\,. \tag{3.30}$$

As a consequence, we obtain

$$\|x_{m_k+1} - x^*\| \leq \|x_{m_k+1} - x_{m_k}\| + \|x_{m_k} - x^*\| \to 0\,,$$

as $n \to \infty$. By Lemma 2.9, we have $D_n \leq D_{m_k+1}$ and thus

$$D_n = \|x_n - x^*\|^2 \leq \|x_{m_k+1} - x^*\|^2 \to 0\,, \tag{3.31}$$

as $n \to \infty$. This implies that $\{x_n\}$ converges strongly to x^*. This complete the proof. □

The following consequences can easily be obtained from our main result.

1. Suppose $h = i_C$, the indicator operator on C, and $w_n = 1$, we obtain the following result which improve and complement the results of Cai and Shehu [8] and Tian and Huang [47].

Corollary 3.3. *Let C be a nonempty, closed and convex subset of a real Hilbert space H. Let $g\colon H \to \mathbb{R} \cup \{+\infty\}$ be a proper convex lower semicontinuous function such that the gradient ∇g is $\frac{1}{L}$-ism with $L > 0$. Let $f\colon C \to C$ be a Meir Keeler contraction mapping, $B\colon C \to H$ be a strongly positive bounded linear operator with coefficient $\tau > 0$ such that $0 < \xi < \frac{\tau}{\rho}$ and $T\colon C \to C$ be a δ-demimetric mapping for $\delta \in (-\infty, 1)$. Suppose $\Gamma = \Omega \cap F(T) \neq \emptyset$, let $\alpha_n \in [0,1]$, $\beta_n \in [0,1)$, $\theta_n \in (0,1)$ and $\gamma_n > 0$. Choose initial points $x_0, x_1 \in H$ arbitrarily and let $\{x_n\}, \{y_n\}$ and $\{u_n\}$ be generated by*

$$\begin{cases} y_n = x_n + \beta_n(x_n - x_{n-1})\,, \\ u_n = P_C(y_n - \gamma_n \nabla g(y_n))\,, \\ x_{n+1} = P_C[\alpha_n \xi f(x_n) + \theta_n x_n + ((1-\theta_n)I - \alpha_n B)T_{\lambda_n} u_n]\,, \quad n \geq 1\,, \end{cases} \tag{3.32}$$

where $T_{\lambda_n} = (1-\lambda_n)I + \lambda_n T$ for $\lambda_n \in (0,1)$. Assume that the following conditions are satisfy:

(C1) $\lim_{n\to\infty} \alpha_n = 0$ *and* $\sum_{n=1}^{\infty} \alpha_n = \infty$,

(C2) $\lim_{n\to\infty} \frac{\beta_n}{\alpha_n}||x_n - x_{n-1}|| = 0$, (C3) $0 < \liminf_{n\to\infty} \gamma_n \leq \limsup_{n\to\infty} \gamma_n < \frac{2}{L}$,

(C4) $0 < \liminf_{n\to\infty} \lambda_n \leq \limsup_{n\to\infty} \lambda_n < 1-\delta$.

Then, $\{x_n\}$ converges strongly to a point $\bar{x}$, where $\bar{x} = P_\Gamma(I - B + \xi f)(\bar{x})$ is the unique solution of the variational inequality

$$\langle (B - \xi f)\bar{x}, \bar{x} - y\rangle \leq 0\,, \quad y \in \Gamma\,. \tag{3.33}$$

2. If $h = i_C$, $\theta_n = 0, \xi = \lambda_n = w_n = 1$, B be identity operator on H and $T\colon C \to C$ is nonexpansive, we obtain the following result from our Theorem 3.2 which improve the corresponding results of Xu [49, Theorem 5.2] and Shehu [42, Theorem 1].

Corollary 3.4. *Let C be a nonempty, closed and convex subset of a real Hilbert space H. Let $g\colon H \to \mathbb{R} \cup \{+\infty\}$ be a proper convex lower semicontinuous function such that the gradient ∇g is $\frac{1}{L}$-ism continuous with $L > 0$. Let $f\colon C \to C$ be a Meir Keeler contraction mapping and $T\colon C \to C$ be a nonexpansive mapping. Suppose $\Gamma = \Omega \cap F(T) \neq \emptyset$, let $\alpha_n \in [0,1]$, $\beta_n \in [0,1)$ and $\gamma_n > 0$. Choose initial points x_0, $x_1 \in H$ arbitrarily and let $\{x_n\}$, $\{y_n\}$ and $\{u_n\}$ be generated by*

$$\begin{cases} y_n = x_n + \beta_n(x_n - x_{n-1})\,, \\ u_n = P_C(y_n - \gamma_n \nabla g(y_n))\,, \\ x_{n+1} = P_C[\alpha_n f(x_n) + (1-\alpha_n)Tu_n]\,, \quad n \geq 1\,. \end{cases} \tag{3.34}$$

Assume that the following conditions are satisfy:

(C1) $\lim_{n\to\infty} \alpha_n = 0$ *and* $\sum_{n=1}^{\infty} \alpha_n = \infty$,

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

(C3) $0 < \liminf_{n\to\infty} \gamma_n \leq \limsup_{n\to\infty} \gamma_n < \frac{2}{L}$,

(C4) $0 < \liminf_{n\to\infty} \lambda_n \leq \limsup_{n\to\infty} \lambda_n < 1-\delta$.

Then, $\{x_n\}$ converges strongly to a point $\bar{x}$, where $\bar{x} = P_\Gamma(I - B + \xi f)(\bar{x})$ is the unique solution of the variational inequality

$$\langle (I - f)\bar{x}, \bar{x} - y\rangle \leq 0\,, \quad y \in \Gamma\,. \tag{3.35}$$

3. Also in Theorem 3.2, we obtained a strong convergence result using a proximal gradient algorithm with an inertial extrapolation term, this improve the weak convergence result proved by Chambolle and Dossal in [11].

Remark 3.5. In [17], the authors proposed a modified proximal gradient algorithm with perturbation for approximating solutions of the minimization problem (1.3), where as in this paper, we presented a strong convergence result using a modified proximal gradient algorithm with inertial extrapolation term without imposing the summation condition (A2). This result improve other recent results on inertial algorithms in the literature.

4. Applications and numerical example

In this section, we present some applications of Theorem 3.2 and give a numerical example to show the efficiency of the iterative scheme (4.10).

4.1 Application to monotone variational inequality problem

Let H be a real Hilbert space and C be a nonempty closed convex subset of H. The Monotone Variational Inequality Problem (MVIP) can be formulated as finding a point $x^* \in C$ such that

$$\langle Mx^*, z - x^* \rangle \geq , \quad \forall\, z \in C , \tag{4.1}$$

where $M\colon C \to H$ is a monotone operator. The set of solutions of the MVIP is denoted by $VIP(M,C)$. The MVIP was first initiated independently by Fichéra [15] and Stampacchia [43] in the early 1960's to study the problems in the elasticity and potential theory respectively. However, the existence and uniqueness of solutions of MVIP was proved by Lions and Stampacchia [22] in 1967.

One method for solving the MVIP (4.1) is by using the projection gradient algorithm which generate a sequence $\{x_n\}$ in H starting with an arbitrary point $x_0 \in H$ by the formula

$$x_{n+1} = P_C(x_n - \lambda M x_n) , \tag{4.2}$$

where $\lambda > 0$ is properly chosen as a stepsize. If M is ν-ism, then the iteration (4.2) with $0 < \lambda < 2\nu$ converges weakly to a point in $VIP(M,C)$. The MVIP (4.1) is equivalent to finding a point $x^* \in C$ such that (see [40])

$$0 \in (M + N_C)x^* ,$$

where N_C is the normal cone operator of C. Note that the resolvent of the normal cone is the projection operator and that if M is ν-ism, then the set $VIP(M,C)$ is closed and convex. Also, if $M : C \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex and lower semicontinuous function, then, the sugradient ∂M which is defined by

$$\partial M := \{u \in C : M(y) \geq M(x) + \langle u, y - x \rangle , \quad \forall\, y \in C\}$$

is maximal monotone operator (see [41]). Thus, setting $M = g$ and $N_C = h$ in our Theorem 3.2, we get the following strong convergence theorem for finding a common solution of MVIP (4.1) and fixed point of δ-demimetric mappings in a real Hilbert space.

Theorem 4.1. *Let C be a nonempty, closed and convex subset of a real Hilbert space H. Let $M\colon C \to \mathbb{R} \cup \{+\infty\}$ be a proper convex lower semicontinuous function such that the gradient ∇M is $\frac{1}{L}$-ism with $L > 0$. Let $f\colon C \to C$ be a Meir Keeler contraction mapping, $B\colon C \to H$ be a strongly positive bounded linear operator with coefficient $\tau > 0$ such that $0 < \xi < \frac{\tau}{2}$ and $T\colon C \to C$ be a δ-demimetric mapping for $\delta \in (-\infty, 1)$ and $\hat{F}(T) = F(T)$. Suppose $\Gamma = VIP(M,C) \cap F(T) \neq \emptyset$, let $\alpha_n \in [0,1]$, $\beta_n \in [0,1)$, $\{w_n\}$ and $\{\theta_n\}$ are sequences in $(0,1)$ and $\gamma_n > 0$. Choose* *initial points* $x_0, x_1 \in H$ *arbitrarily and let* $\{x_n\}, \{y_n\}$ *and* $\{u_n\}$ *be generated by*

$$\begin{cases} y_n = x_n + \beta_n(x_n - x_{n-1})\,, \\ u_n = (1-w_n)y_n + w_n \operatorname{prox}_{\gamma_n h}(y_n - \gamma_n \nabla M(y_n))\,, \\ x_{n+1} = P_C[\alpha_n \xi f(x_n) + \theta_n x_n + ((1-\theta_n)I - \alpha_n B)T_{\lambda_n} u_n]\,, \quad n \geq 1\,, \end{cases} \tag{4.3}$$

where $T_{\lambda_n} = (1-\lambda_n)I + \lambda_n T$ *for* $\lambda_n \in (0,1)$. *Assume that the following conditions are satisfy:*

(C1) $\lim_{n\to\infty} \alpha_n = 0$ *and* $\sum_{n=1}^{\infty} \alpha_n = \infty$,

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

(C3) $0 < \liminf_{n\to\infty} w_n \leq \limsup_{n\to\infty} w_n < 1$,

(C4) $0 < \liminf_{n\to\infty} \gamma_n \leq \limsup_{n\to\infty} \gamma_n < \frac{2}{L}$,

(C5) $0 < \liminf_{n\to\infty} \lambda_n \leq \limsup_{n\to\infty} \lambda_n < 1 - \delta$.

Then, $\{x_n\}$ *converges strongly to a point* $\bar{x}$*, where* $\bar{x} = P_\Gamma(I - B + \xi f)(\bar{x})$ *is the unique solution of the variational inequality*

$$\langle (B - \xi f)\bar{x}, \bar{x} - y\rangle \leq 0\,, \quad y \in \Gamma\,. \tag{4.4}$$

4.2 Application to proximal split feasibility problem

Let H_1 and H_2 be real Hilbert spaces, C and Q be nonempty closed and convex subset of H_1 and H_2 respectively. Let $R\colon H_1 \to \mathbb{R}\cup\{+\infty\}$ and $S\colon H_2 \to \mathbb{R}\cup\{+\infty\}$ be proper, convex and lower semicontinuous functions and let $A\colon H_1 \to H_2$ be a bounded linear operator. The Proximal Split Feasibility Problem (PSFP) is to find a point x^* with the property

$$x^* \in \text{argmin } R \quad \text{such that} \quad Ax^* \in \text{argmin } S\,, \tag{4.5}$$

where

$$\begin{aligned} &\text{argmin } S := \{x \in H_1 : S(x) \leq S(y)\,,\ \forall\ y \in H_1\}\,, \quad \textit{and} \\ &\text{argmin } R := \{u \in H_2 : R(u) \leq R(v)\,,\ \forall\ v \in H_2\}\,. \end{aligned}$$

We denote the solution set of the PSFP (4.5) by Λ. The PSFP was first introduced by Moudafi and Thakur in [31]. By taking $S = i_C$ and $R = i_Q$, the indicator functions on C and Q respectively, the PSFP reduces to the split feasibility problem introduced by Censor and Elfving [10]. The SFP have been applied to model inverse problem arising in machine learning, signal processing, medical radiation therapy, etc [10]. To solve the PSFP, it is very important to investigate the following minimization problem: Find a solution $x^* \in H_1$ such that

$$\underset{x\in H_1}{\text{minimize}}\{R(x) + S_\mu(Ax)\}\,, \tag{4.6}$$

where $S_\mu(y) = \underset{u\in H_2}{\text{argmin}}\{S(u) + \dfrac{1}{2\mu}\|u-y\|^2\}$ stands for the Moreau-Yosida approximation of S with parameter μ [31]. By the differentiability of the Yosida approximation

S_μ (see for instance [41]), we have the additive of the subdifferentials and thus, we can write

$$\begin{aligned}\partial(R(x)+S_\mu(Ax)) &= \partial R(x) + A^*\nabla S_\mu(Ax)\\ &= \partial R(x) + A^*\Big(\frac{I-\mathrm{prox}_{\mu S}}{\mu}\Big)(Ax)\,.\end{aligned}$$

This implies that the optimality condition of (4.6) can then be written as

$$0 \in \ \mu\partial R(x) + A^*(I-\mathrm{prox}_{\mu S})Ax\,, \tag{4.7}$$

where ∂R stands for the subdifferential of R at x, i.e.

$$\partial R := \{u \in H_1 : R(y) \ge R(x) + \langle u, y-x\rangle,\ \forall y \in H_1\}\,.$$

This inclusion in (4.7) yields the following equivalent fixed point formulation (see [31])

$$\mathrm{prox}_{\gamma\mu R}(x^* - \gamma A^*(I-\mathrm{prox}_{\mu S}))Ax^* = x^*. \tag{4.8}$$

Hence, to solve (4.6), (4.8) suggest we consider the following split proximal algorithm:

$$x_{n+1} = \mathrm{prox}_{\gamma\mu R}(x_n - \gamma_n A^*(I-\mathrm{prox}_{\mu S}))Ax_n\,. \tag{4.9}$$

Several other iterative methods have been introduced for solving the PSFP in Hilbert spaces, see for instance [1, 28, 33, 34, 35, 36] and references therein.

Setting $\nabla g(x) = A^*(I-\mathrm{prox}_{\mu S})Ax$ in Theorem 3.2, then ∇g is $\frac{1}{\nu}$-ism with $\nu = \|A\|$ (see [7], Page 113). This implies that we can apply Theorem 3.2 to obtain solution of PSFP in real Hilbert space. Thus, we give the following result which complement other results in literature on finding solution of PSFP.

Theorem 4.2. *Let C and Q be nonempty, closed and convex subsets of real Hilbert spaces H_1 and H_2 respectively. Let $A\colon H_1 \to H_2$ be a bounded linear operator, $R\colon H_1 \to \mathbb{R}\cup\{+\infty\}$ and $S\colon H_2 \to \mathbb{R}\cup\{+\infty\}$ be two proper convex lower semicontinuous functions such that $\Lambda \ne \emptyset$. Let $f\colon C \to C$ be a Meir Keeler contraction mapping, $B\colon C \to H$ be a strongly positive bounded linear operator with coefficient $\tau > 0$ such that $0 < \xi < \frac{\tau}{\rho}$ and $T\colon C \to C$ be a δ-demimetric mapping for $\delta \in (-\infty, 1)$. Suppose $\Gamma = \Lambda \cap F(T) \ne \emptyset$, let $\alpha_n \in [0,1]$, $\beta_n \in [0,1)$, w_n, $\theta_n \in (0,1)$ and $\gamma_n > 0$. Choose initial points x_0, $x_1 \in H_1$ arbitrarily and let $\{x_n\}, \{y_n\}$ and $\{u_n\}$ be generated by*

$$\begin{cases} y_n = x_n + \beta_n(x_n - x_{n-1})\,,\\ u_n = (1-w_n)y_n\\ \qquad\qquad + w_n\,\mathrm{prox}_{\mu_n\gamma_n R}(y_n - \gamma_n A^*(I-\mathrm{prox}_{\mu_n S})Ay_n)\,,\\ x_{n+1} = P_C[\alpha_n\xi f(x_n) + \theta_n x_n + ((1-\theta_n)I - \alpha_n B)T_{\lambda_n}u_n]\,,\\ \qquad n \ge 1\,, \end{cases} \tag{4.10}$$

where $T_{\lambda_n} = (1-\lambda_n)I + \lambda_n T$ for $\lambda_n \in (0,1)$. Assume that the following conditions are satisfy:

(C1) $\lim_{n\to\infty}\alpha_n = 0$ *and* $\sum_{n=1}^{\infty}\alpha_n = \infty$, (C2) $\lim_{n\to\infty} \frac{\beta_n}{\alpha_n}||x_n - x_{n-1}|| = 0$,

(C3) $0 < \liminf_{n\to\infty} w_n \le \limsup_{n\to\infty} w_n < 1$,

(C4) $0 < \liminf_{n\to\infty} \gamma_n \le \limsup_{n\to\infty} \gamma_n < \frac{2}{||A||^2}$,

(C5) $0 < \liminf_{n\to\infty} \lambda_n \le \limsup_{n\to\infty} \lambda_n < 1 - \delta$.

Then, $\{x_n\}$ converges strongly to a point $\bar{x}$, where $\bar{x} = P_\Gamma(I - B + \xi f)(\bar{x})$ is the unique solution of the variational inequality

$$\langle (B - \xi f)\bar{x}, \bar{x} - y\rangle \le 0\,, \quad y \in \Gamma\,. \tag{4.11}$$

4.3 Numerical example

In this subsection, we give a numerical example to show the efficiency and implementation of our algorithm (4.10) for solving PSFP. All codes were written in Matlab 2016(a) and run on HP EliteBook 6930p laptop.

Example 4.3. Let $H_1 = \mathbb{R}^N = H_2$ and $S := \|\cdot\|_2$, the Euclidean norm on $\mathbb{R}^N$. It is obvious that we can project onto the Euclidean unit ball B_r as follows:

$$P_{B_r}(x) = \begin{cases} \frac{x}{\|x\|_2}, & \text{if } \|x\|_2 > 1\,, \\ x\,, & \text{if } \|x\|_2 \le 1\,. \end{cases} \tag{4.12}$$

In this case, the proximal operator is given by

$$\text{prox}_S(x) = \begin{cases} \big(1 - \frac{1}{\|x\|_2}\big)x\,, & if \;\; \|x\|_2 \ge 1\,, \\ 0\,, & \text{if } \;\; \|x\|_2 < 1\,. \end{cases} \tag{4.13}$$

This proximal operator is called the block soft thresholding. Also, let $x_i \in \mathbb{R}$, $i = 1, 2, \dots, N$. Define

$$i_j(x_j) = \max\Big\{|x_j| - 1, 0\Big\}, \quad j = 1, 2, \dots, N\,,$$

and

$$R(x) = \sum_{j=1}^{N} i_j(x_j)\,.$$

Then (see [14])

$$\text{prox}_{i_j}(x_j) = \begin{cases} x_j\,, & if \;|x_j| < 1, \\ \text{sign}(x_j)\,, & if \;1 \le |x_j| \le 2, \\ \text{sign}(x_j - 1)\,, & \text{otherwise}, \end{cases} \tag{4.14}$$

and

$$\text{prox}_R(x) = \big(\,\text{prox}_{i_1}(x_1), \text{prox}_{i_2}(x_2), \dots, \text{prox}_{i_N}(x_N)\big)\,.$$

Suppose $Ax = x \in \mathbb{R}^N$. We consider the following PSFP:

$$\text{find} \quad x^* \in \text{argmin}\, R \quad \text{such that} \quad Ax^* \in \text{argmin}\, S\,. \tag{4.15}$$

Chosen $\alpha_n = \dfrac{1}{n+1}$, $\beta_n = \dfrac{1}{(n+1)^3}$, $\theta_n = \dfrac{n}{2(n+3)}$, $w_n = \dfrac{1}{5(1+\frac{1}{n})}$ and $\lambda_n = \dfrac{n}{2n+3}$. Let $f(x) = \frac{x}{2}$, $B(x) = x$, $T(x) = \frac{x}{2}$, $\xi = 1$, $x_0 = 0.5 \times \text{randn}(50, N)$ and

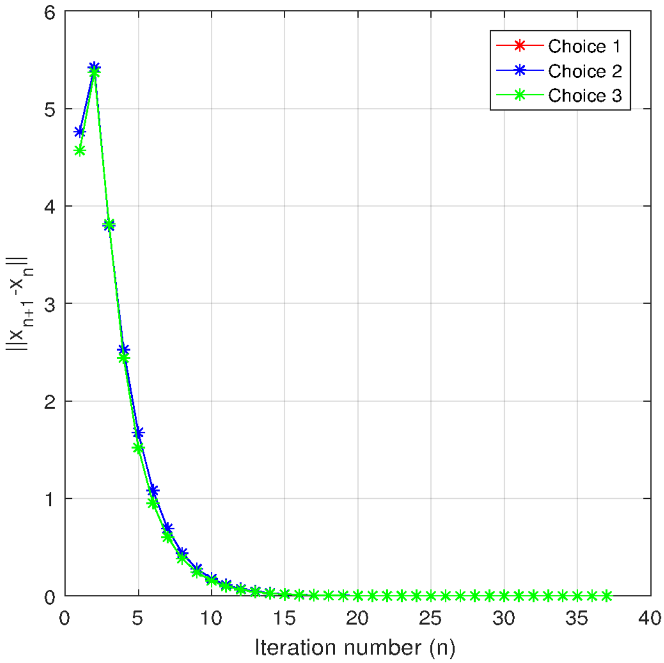

FIG. 1: Example 4.3, Case (i): $N = 100$.

$x_1 = 2 \times \mathrm{randn}(50, N)$ (randomly generated vectors in $\mathbb{R}^{\mathbb{N}}$). Using $\dfrac{\|x_{n+1} - x_n\|_2}{\|x_2 - x_1\|_2} < 10^{-6}$ as the stopping criterion, we consider various values of N and choices of γ_n as follows:

Case (i): $N = 100$, Case (ii): $N = 500$, Case (iii): $N = 1000$, Case (iv): $N = 2000$,

and

$$\text{Choice (i): } \gamma_n = \frac{n}{n+1}\,, \quad \text{Choice (ii) } \gamma_n = \frac{n}{5n+7}\,, \quad \text{Choice (iii) } \gamma_n = 0.7\,.$$

Remark 4.4. Figures 1, 2, 3, 4 and Table 1 show that there is no significant change in the CPU time taken and the number of iterations for different values of N and the stepsizes.

Acknowledgement. The second author acknowledge with thanks the bursary and financial support from Department of Science and Technology and National Research Foundation, Republic of South Africa Center of Excellence in Mathematical and Statistical Sciences (DST-NRF COE-MaSS) Doctoral Bursary. Opinions expressed and conclusions arrived are those of the authors and are not necessarily to be attributed to the CoE-MaSS.

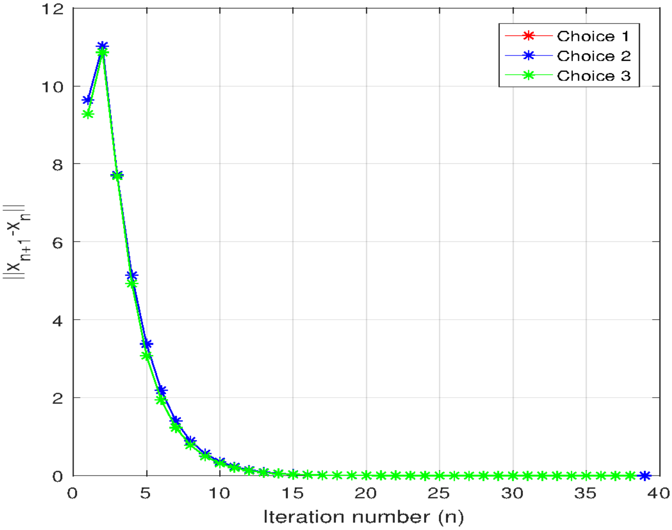

FIG. 2: Example 4.3, Case (ii): $N = 500$.

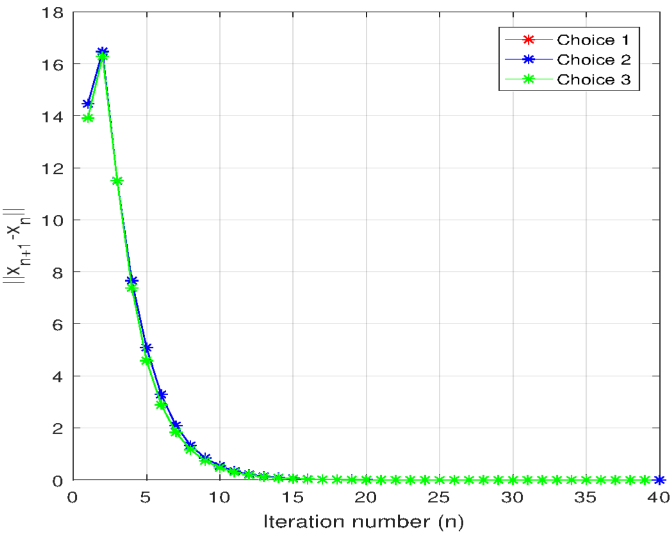

FIG. 3: Example 4.3, Case (iii), $N = 1000$.

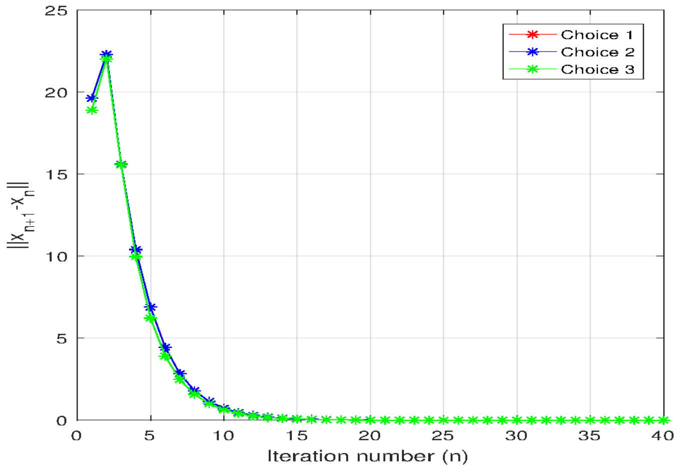

FIG. 4: Example 4.3, Case (iv), $N = 2000$.

N		Choice (i)	Choice (ii)	Choice (iii)
100	Number of Iter.	25	27	27
	CPU time (sec)	0.0356	0.0399	0.0476
500	Number of Iter.	27	29	29
	CPU time (sec)	0.2244	0.3913	0.5061
1000	Number of Iter.	29	30	30
	CPU time (sec)	0.4697	0.4095	0.5235
2000	Number of Iter.	30	30	30
	CPU time (sec)	1.0731	1.0912	0.8785

Table 1 : Showing the number of iteration and CPU time (sec) for each values of the stepsize and N in Example 4.3

References

[1] Abass, H.A., Ogbuisi, F.U., Mewomo, O.T., *Common solution of split equilibrium problem and fixed point problem with no prior knowledge of operator norm*, U.P.B. Sci. Bull., Series A **80** (1) (2018), 175–190.

[2] Alvarez, F., Attouch, H., *An inertial proximal method for monotone operators via discretization of a nonlinear oscillator with damping*, Set-Valued Anal. **9** (2001), 3–11.

[3] Beck, A., Teboull, M., *Gradient-based algorithms with applications to signal-recovery problems*, Convex optimization in signal processing and communications (Palomar, D., Elder, Y., eds.), Cambridge Univ. Press, Cambridge, 2010, pp. 42–88.

[4] Beck, A., Teboulle, M., *A fast iterative shrinkage-thresholding algorithm for linear inverse problem*, SIAM J. Imaging Sci. **2** (1) (2009), 183–202.

[5] Bot, R.I., Csetnek, E.R., *An inerial Tseng's type proximal point algorithm for nonsmooth and nonconvex optimization problem*, J. Optim. Theory Appl. **171** (2016), 600–616.

[6] Bot, R.I., Csetnek, E.R., Laszlo, S.C., *An inertial forward-backward algorithm for the minimization of the sum of two nonconvex functions*, EJCO **4** (2016), 3–25.

[7] Byrne, C., *A unified treatment of some iterative algorithms in signal processing and image reconstruction*, Inverse Problems **20** (2004), 103–120.

[8] Cai, G., Shehu, Y., *An iterative algorithm for fixed point problem and convex minimization problem with applications*, Fixed Point Theory and Appl. 2015 **123** (2015), 18 pp.

[9] Ceng, L.-C., Ansari, Q.H., Ya, J.-C., *Some iterative methods for finding fixed points and for solving constrained convex minimization problems*, Nonlinear Anal. **74** (2011), 5286–5302.

[10] Censor, Y., Elfving, T., *A multiprojection algorithm using Bregman projections in a product space*, Numer. Algorithms **8** (2–4) (1994), 221–239.

[11] Chambolle, A., Dossal, C., *On the convergence of the iterates of the fast iterative shrinkage thresholding algorithm*, J. Optim. Theory Appl. **166** (2015), 968–982.

[12] Chan, R.H., Ma, S., Jang, J.F., *Inertial proximal ADMM for linearly constrained separable convex optimization*, SIAM J. Imaging Sci. **8** (4) (2015), 2239–2267.

[13] Combettes, P.L., *Solving monotone inclusions via compositions of nonexpansive averaged operators*, Optimization **53** (2004), 475–504.

[14] Combettes, P.L., Pesquet, J.-C., *Proximal Splitting Methods in Signal Processing*, Fixed-Point Algorithms for Inverse Problems in Science and Engineering, Springer, New York, 2011, pp. 185–212.

[15] Fichera, G., *Problemi elastostatic con vincoli unilaterli: II Problema di signorini con ambigue condizioni al contorno*, Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez. I (8) **7** (1963/1964), 91–140.

[16] Geobel, K., Kirk, W.A., *Topics in Metric Fixed Point Theory*, Cambridge Studies in Advanced Mathematics, vol. 28, Cambridge University Press, Cambridge, 1990.

[17] Guo, Y., Cui, W., *Strong convergence and bounded perturbation resilence of a modified proximal gradient algorithm*, J. Ineq. Appl. **2018** (2018).

[18] Izuchukwu, C., Ugwunnadi, G.C., Mewomo, O.T., Khan, A.R., Abbas, M., *Proximal-type algorithms for split minimization problem in P-uniformly convex metric spaces*, Numer. Algorithms (2018), `https://doi.org/10.1007/s11075-018-0633-9`.

[19] Jolaoso, L.O., Ogbuisi, F.U., Mewomo, O.T., *An iterative method for solving minimization, variational inequality and fixed point problems in reflexive Banach spaces*, Adv. Pure Appl. Math. **9** (3) (2018), 167–183.

[20] Jolaoso, L.O., Oyewole, K.O., Okeke, C.C., Mewomo, O.T., *A unified algorithm for solving split generalized mixed equilibrium problem and fixed point of nonspreading mapping in Hilbert space*, Demonstratio Math. **51** (2018), 211–232.

[21] Jolaoso, L.O., Taiwo, A., Alakoya, T.O., Mewomo, O.T., *A strong convergence theorem for solving variational inequalities using an inertial viscosity subgradient extragradient algorithm with self adaptive stepsize*, Demonstratio Math. **52** (1) (2019), 183–203.

[22] Lions, J.L., Stampacchia, G., *Variational inequalities*, Comm. Pure Appl. Math. **20** (1967), 493–519.

[23] Lorenz, D., Pock, T., *An inertial forward-backward algorithm for monotone inclusions*, J. Math. Imaging Vision **51** (2) (2015), 311–325.

[24] Maingé, P.E., *Approximation methods for common fixed points of nonexpansive mappings in Hilbert spaces*, J. Math. Anal. Appl. **325** (2007), 469–479.

[25] Maingé, P.E., *Strong convergence of projected subgradient methods for nonsmooth and nonstrictly convex minimization*, Set-Valued Anal. **16** (2008), 899–912.

[26] Martinez-Yanes, C., Xu, H.K., *Strong convergence of the CQ method for fixed-point iteration processes*, Nonlinear Anal. **64** (2006), 2400–2411.

[27] Meir, A., Keeler, E., *A theorem on contraction mappings*, J. Math. Anal. Appl. **28** (1969), 326–329.

[28] Mewomo, O.T., Ogbuisi, F.U., *Convergence analysis of iterative method for multiple set split feasibility problems in certain Banach spaces*, Quaestiones Math. **41** (1) (2018), 129–148.

[29] Moudafi, A., *Viscosity approximation method for fixed-points problems*, J. Math. Anal. Appl. **241** (1) (2000), 46–55.

[30] Moudafi, A., Oliny, M., *Convergence of a splitting inertial proximal method for monotone operators*, J. Comput. Appl. Math. **155** (2003), 447–454.

[31] Moudafi, A., Thakur, B.S., *Solving proximal split feasibility problems without prior knowledge of operator norms*, Optim. Lett. **8** (7) (2014), 2099–2110.

[32] Nesterov, Y., *A method for solving the convex programming problem with convergence rate* $0(\frac{1}{k^2})$, Dokl. Akad. Nauk SSSR **269** (3) (1983), 543–547.

[33] Ogbuisi, F.U., Mewomo, O.T., *On split generalized mixed equilibrium problems and fixed point problems with no prior knowledge of operator norm*, J. Fixed Point Theory Appl. **19** (3) (2016), 2109–2128.

[34] Ogbuisi, F.U., Mewomo, O.T., *Iterative solution of split variational inclusion problem in a real Banach space*, Afrika Mat. (3) **28** (1–2) (2017), 295–309.

[35] Ogbuisi, F.U., Mewomo, O.T., *Convergence analysis of common solution of certain nonlinear problems*, Fixed Point Theory **19** (1) (2018), 335–358.

[36] Okeke, C.C., Mewomo, O.T., *On split equilibrim problem, variational inequality problem and fixed point problem for multi-valued mappings*, Ann. Acad. Rom. Sci. Ser. Math. Appl. **9** (2) (2017), 255–280.

[37] Parith, N., Boyd, S., *Proximal algorithms*, Foundations and Trends in Optimization **1** (3) (2013), 123–231.

[38] Pesquet, J.-C., Putselnik, N., *A parallel inertial proximal optimization method*, Pacific J. Optim. **8** (2) (2012), 273–306.

[39] Polyak, B.T., *Some methods of speeding up the convergence of iteration methods*, U.S.S.R. Comput. Math. Math. Phys. **4** (5) (1964), 1–17.

[40] Rockafellar, R.T., *On the maximality of sums of nonlinear monotone operators*, Trans. Amer. Math. Soc. **149** (1970), 75–88.

[41] Rockafellar, R.T., Wets, R., *Variational Analysis*, Springer, Berlin, 1988.

[42] Shehu, Y., *Approximation of solutions to constrained convex minimization problem in Hilbert spaces*, Vietnam J. Math. (2014), `DOI10.1007/s10013-014-0091-1`.

[43] Stampacchia, G., *Formes bilinearies coercitives sur les ensembles convexes*, Comput. Rend. Acad. Sci. Paris **258** (1964), 4413–4416.

[44] Su, M., Xu, H.K., *Remarks on the gradient-projection algorithm*, J. Nonlinear Anal. Optim. **1** (1) (2010), 35–43.

[45] Suzuki, T., *Moudai's viscosity approximations with Meir-Keeler contractions*, J. Math. Anal. Appl. **325** (2007), 342–352.

[46] Takahashi, W., Wen, C.-F., Yao, J.-C., *The shrinking projection method for a finite family of demimetric mappings with variational inequality problems in a Hilbert space*, Fixed Point Theory **19** (1) (2018), 407–420.

[47] Tian, M., Huang, L.H., *A general approximation method for a kind of convex optimization problems in Hilbert spaces*, J. Appl. Math. **2014** (2014), 9 pages, ArticleID156073.

[48] Xu, H.K., *Viscosity approximation method for nonexpansive mappings*, J. Math. Anal. Appl. **298** (1) (2004), 279–291.

[49] Xu, H.K., *Average mappings and the gradient projection algorithm*, J. Optim. Theory Appl. **150 rm (2)** (2011), 360–378.

[50] Xu, H.K., *Properties and iterative methods for the LASSO and its variants*, Chin. Ann. Math., Ser. B **35** (2014), 501–518.

Corresponding author: O.T. Mewomo,
School of Mathematics, Statistics and Computer Science,
University of KwaZulu-Natal,
Durban, South Africa
E-mail: mewomoo@ukzn.ac.za
216074984@stu.ukzn.ac.za lateefjolaoso89@gmail.com
216075727@stu.ukzn.ac.za hammedabass548@gmail.com