Int. J. Anal. Appl. (2025), 23:160

International Journal of Analysis and Applications

# Accelerated Self-Adaptive Method for Solving Nonsmooth Convex Minimization Problem in Real Hilbert Spaces

**L. Mokaba[1], Hammed A. Abass[1,*], Olawale K. Oyewole[2], Koketso P. Malebana[1]**

[1]*Department of Mathematics and Applied Mathematics, Sefako Makgatho Health Science University, P.O Box 94, Pretoria 0204, South Africa*

[2]*Department of Mathematics, Tshwane University of Technology, Arcadia, PMB 00007, Pretoria, South Africa*

[*]*Corresponding author:* hammed.abass@smu.ac.za, hammedabass548@gmail.com

**Abstract.** In this manuscript, we propose a proximal gradient type algorithm together with a two step inertia method for approximating solution of convex minimization problem in real Hilbert spaces. The proposed proximal gradient type method is designed in such a way that it does not depend on the Lipschitz constant. Using a self-adaptive rule, we obtain a weak convergence result under the condition that the gradient function of one of the convex functions is uniformly continuous. Preliminary numerical results show that our proposed method has a better convergence in comparison to some other related results in the literature.

## 1. Introduction

In recent years, the importance of studying the structure of convex optimization problems have become a topic of intense research in machine learning, signal processing and intensity modulated radiotherapy. This is particularly true of techniques for non-smooth optimization, where taking advantage of the structure of non-smooth terms seems to be crucial to obtaining better performance. Let $\mathcal{H}$ be a real Hilbert space, the convex minimization problem (CMP) is to find:

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

where $h : \mathcal{H} \to (-\infty, +\infty]$ is a proper, closed and convex function which is possibly nonsmooth and $g : \mathcal{H} \to \mathbb{R}$ is a proper, closed, convex and continuously differentiable function with its gradient

Received: Apr. 29, 2025.

2020 *Mathematics Subject Classification.* 65K10.

*Key words and phrases.* convex minimization problem; inertial method; step-size; Hilbert spaces.

https://doi.org/10.28924/2291-8639-23-2025-160

ISSN: 2291-8639

$\nabla g$ being Lipschitz continuous on $\mathcal{H}$. We denote the set of minimizing of CMP (1.1) by $\Omega$. It is well-known that

$$x^* \in \Omega \Longleftrightarrow 0 \in (\nabla g + \partial h)(x^*),$$

where $\partial h$ is the subdifferential of $h$. One of the most well-studied instances of CMP (1.1) is the basic pursuit denoising (BPDN), (see [1,7,15,23]) which is of the form:

$$\min_x \Big(\frac{1}{2}\|q - Ax\|^2 + \tau\|x\|\Big),$$

where $\tau$ is a parameter that control the trade-off between sparsity and reconstruction fidelity, $x$ is an $N \times 1$ solution vector, $q$ is an $M \times 1$ vector of observation, A is an $M \times N$ transform matrix and $M < N$. The proximal gradient method (PGM) is an appealing technique for solving CMP (1.1) due to their fast theoretical convergence rates and strong practical performance. The PGM can be formulated as follows:- Given the initial point $w_1 \in \mathcal{H}$, compute

$$w_{k+1} = prox_{\tau_{kh}}(w_k - \tau_k \nabla g(w_k)), k \geq 1, \tag{1.2}$$

where $prox_{\tau_{kh}}(x) :=\arg\min_{x\in\mathcal{H}}\{h(x) + \frac{1}{2\tau_k}\|x - w\|^2\}$ and $\tau_k > 0$ is the stepsize. If $g = 0$, then the PGM reduces to the classical proximal point algorithm (PPA), (see [12]). In addition, if $\nabla g$ in (1.2) is L-Lipschitzian and $\{\tau_k\}$ is such that $0 < \liminf_{k\to\infty} \tau_k \leq \limsup_{k\to\infty} \tau_k < \frac{2}{L}$, then the sequence $\{w_k\}$ defined in (1.2) converges weakly to $x^*$, where $x^*$ is a solution of (1.1). In 2018, Guo and Cui (see [10]) proposed the following PGM with perturbation for solving CMP (1.1):

$$w_{k+1} = \delta_k f(w_k) + (1 - \alpha_k) prox_{\tau_{kh}}(I - \tau_k \nabla g) w_k + \varrho_k \tag{1.3}$$

where $\{\delta_k\} \subset [0, 1], 0 < a < \liminf_{k\to\infty} \tau_k < \frac{2}{L}$, $f : \mathcal{H} \to \mathcal{H}$ is a contraction and $\varrho : \mathcal{H} \to \mathcal{H}$ is a peturbation operator satisfying $\sum_{k=0}^{\infty} \|\varrho(w_k)\| < +\infty$. They established a strong convergence result in the setting of real Hilbert spaces.

In optimization theory, authors are not only after the convergence analysis of iterative methods but the rate at which these iterative methods converge. Fast converging iterative algorithm have gained the attention of numerous researchers in recent years, (see [4]). Many authors have employed the one step inertial method to expedite the rate of convergence of different iterative algorithms, (see [18]). Recently, Alvarez and Attouch [4] employed the heavy ball technique which was examined in (see [24]) for maximal monotone operators by applying the method of proximal points. The inertial PPA is generated as follows:

$$\begin{cases} y_k = q_k + \Theta_k (q_k - q_{k-1}) \\ q_{k+1} = (I + \tau_k B)^{-1} y_k, k \geq 1 \end{cases} \tag{1.4}$$

then (1.4) weakly converges to a zero of a maximal monotone operator $B$. In (1.4), $\Theta_k$ is called the extrapolation factor and the inertial term is represented by $\Theta_k(q_k - q_{k-1})$. It was stated in (see [14,24]) that one-step inertial does not provide acceleration. Moreover, the sequences generated by the one-step inertial iteration method were shown to converge more slowly that their non-inertial counterpart. This indicates that the inertial iteration method in one step might not be able to provide sufficient acceleration. Employing the following two-step inertial method introduced in (see [11]) could address the issue of lack of acceleration. Research on two-step inertial methods revealed that using two-step inertial method speeds up iterative method (see [11]). In 2019, Jolaoso *et al*. [12] proposed the following proximal gradient algorithm together with an inertial extrapolation term for solving CMP (1.1) and fixed point of $\delta$-demimetric mapping in the settings of real Hilbert spaces: Given $w_0, w_1 \in \mathcal{H}$ arbitrarily and let $\{w_k\}$ be generated by

$$\begin{cases} y_k = w_k + \beta_k(w_k - w_{k-1}), \\ v_k = (1-b_k)y_k + b_k prox_{\tau_{kh}}(y_k - \tau_k \nabla g(y_k)), \\ w_{k+1} = P_C(\alpha_k \epsilon_1 f(w_k) + \Theta_k w_k + ((1-\Theta_k)I - \alpha_k B)T_{\mu_k} v_k), k \geq 1 \end{cases} \tag{1.5}$$

where $T_{\mu_k}$= $(1-\mu_k)I + \mu_k T$ for $\mu_k \in (0,1), T : \mathcal{H} \to \mathcal{H}$ is a $\delta-$demimetric mapping for $\delta \in (-\infty, 1)$, $f : \mathcal{H} \to \mathcal{H}$ is a Meir keeler contraction and $B$ is a strongly positive bounded linear operator with coefficient $\tau > 0$ such that $0 < \xi < \frac{\tau}{2}$. The authors established a strong convergence results under some mild conditions. It can be easily observed that Algorithm 1.4 and 1.5 rely on a fixed constant step size, necessitating knowledge or an approximation of the Lipschitz constant of $\nabla g$. From a computational standpoint, the use of a fixed stepsize can be problematic, potentially impacting the convergence rate and the overall suitability of the method. Therefore, the reliance on a fixed step size introduces challenges that need careful consideration in practical applications. To overcome this difficulty, Chen and Duan (see [8]) proposed the following inertial self-adaptive proximal gradient methods for solving CMP (1.1):

---

**Algorithm 1.1.** *Step 0: Given $\epsilon > 0, \beta > 3$, initialize $w_0, w_1 \in \mathcal{H}$ and $\tau_1 > 0$, set $k = 1$.*
*Step 1: Given $w_{k-1}, w_k$ and compute*

$$\begin{cases} y_k = w_k + \Delta_k(w_k - w_{k-1}) \\ s_k = prox_{\tau_{kh}}(I - \tau_k \nabla g)y_k \\ w_{k+1} = \alpha_k h(y_k) + (1-\alpha_k)s_k + (1-\alpha_k)\tau_k(\nabla g(y_k) - \nabla g(s_k)), \end{cases} \tag{1.6}$$

where $\Delta_k$ satisfies $0 \leq |\Delta_k| \leq \bar{\Delta_k}$ with $\bar{\Delta_k}$ defined by

$$\overline{\bar{\Delta_k}} = \begin{cases} \min\left\{\frac{k-1}{k+\beta-1}, \quad \frac{\epsilon_k}{\|w_k - w_{k-1}\|}\right\}, & \text{if } w_k \neq w_{k-1}, \\ \frac{k-1}{k+\beta-1}, & w_k = w_{k-1}. \end{cases} \tag{1.7}$$

Step 2 : Update step-size

$$\tau_{k+1} = \min\{\tau_k, \frac{\mu_k \|y_k - s_k\|}{\|\nabla g(y_k) - \nabla g(s_k)\|}\}, \tag{1.8}$$

where $\{\mu_k\} \subset (0,1)$.

Step 3: If $\|w_{k+1} - w_k\| \leq \epsilon$, then the iterative process stops. Otherwise, set $k = k+1$ and go to step 1.

---

In Algorithm 1.1, $\nabla g$ is $L$-Lipschitz and $h$ is a $\sigma$-contraction mapping with constant $0 \leq \sigma < 1$. The authors established a strong convergence results under some suitable conditions.
In this paper, we propose a two step inertia proximal gradient method for solving CMP (1.1) where the gradient function $\nabla g$ is uniformly continuous. Our method generates variable step sizes at each iteration based on some previous iterates, without employing any line-search procedure. The convergence of our proposed method is established under suitable conditions. We present some numerical examples to illustrate the computational effectiveness of our method in comparison to some existing ones in the literature. Our results extend and complement many related results in the literature.

## 2. Preliminaries

**Definition 2.1.** *Let $\mathcal{H}$ be a real Hilbert space $\mathcal{H}$. The mapping $T : \mathcal{H} \to \mathcal{H}$ is said to be:*

(1) *L-Lipschitz continuous, where $L > 0$, if*

$$\|Tx - Ty\| \leq L\|x - y\|, \forall x, y \in H. \tag{2.1}$$

(2) *Uniformly continuous, if for every $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$, such that*

$$\|Tx - Ty\| < \epsilon \text{ whenever } \|x - y\| < \delta, \forall x, y \in H. \tag{2.2}$$

**Remark 2.1.** *It is well known that if $D$ is a convex subset of $\mathcal{H}$, then $T : D \to \mathcal{H}$ is uniformly continuous if and only if, for every $\epsilon > 0$, there exixts a constant $K < +\infty$, such that*

$$\|Tx - Ty\| \leq K\|x - y\| + \epsilon \ \forall x, y \in D. \tag{2.3}$$

**Definition 2.2.** *Let $h : \mathcal{H} \to (-\infty, +\infty]$ be proper, convex, and lower semi-continuous function and $\bar{c} > 0$. The proximity operator of $h$ of order $\bar{c}$ is defined by*

$$prox_{\bar{c}h} := \arg\min_{t \in \mathcal{H}}\{h(t) + \frac{1}{2\bar{c}}\|s - t\|^2\}, \tag{2.4}$$

*for all $s \in \mathcal{H}$.*

**Definition 2.3.** *$B$ is said to be*

(i) *monotone if for all $(s, u)$,$(t, v) \in gra(B)$ (the graph of mapping $B$),*

$$\langle u - v, s - t\rangle \tag{2.5}$$

(ii) *maximal monotone if for every $(s,u) \in \mathcal{H} \times \mathcal{H}$,*
*$(s, u) \in gra(B) \Longleftrightarrow \langle u - v, s - t\rangle \geq 0$ for all $(t, v) \in gra(B)$.*

**Definition 2.4.** *A convex function $c : \mathcal{H} \to \mathbb{R}$ is said to be subdifferentiable at a point $k \in \mathcal{H}$ if the set*

$$\partial c(x) = \{u \in \mathcal{H} | c(y) \geq c(x) + \langle u, y - x\rangle, \forall\ y \in \mathcal{H}\} \tag{2.6}$$

*is nonempty , where each element in $\partial c(x)$ is called a subgradient of $c$ at $x$ ,$\partial c(x)$ is called the subdifferential of $c$ at $x$, and the inequality in (2.6) is called the subdifferential inequality of $c$ at $x$. We say that $c$ is subdifferentiable on $\mathcal{H}$ if $c$ is subdifferentiable at each $x \in \mathcal{H}$*

**Lemma 2.1.** *[5] Let $B : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximal monotone mapping and $A : \mathcal{H} \to \mathcal{H}$ be a hemicontinuous, monotone, and bounded operator. Then, the mapping $A + B$ is a maximal monotone mapping.*

**Lemma 2.2.** *[3] The following identities holds for all $a, b, c \in \mathbb{H}$ and $\lambda\mathbb{R}$*

(i) $2\langle a, b\rangle = ||a||^2 + ||b||^2 - ||a-b||^2 = ||a+b||^2 - ||a||^2 - ||b^2$

(ii) $2\langle a-b, a-c\rangle = ||a-b||^2 + ||a-c||^2 - ||a-c||^2$

(iii) $||\lambda a + (1-\lambda)b||^2 = \lambda||a||^2 + (1-\lambda)||b||^2 - \lambda(1-\lambda)||a-b||^2.$

**Lemma 2.3.** *[22] Suppose $\lambda_k$ and $\theta_k$ are two nonegative real sequences, such that*

$$\lambda_{k+1} \leq \lambda_k + \phi_k, \forall k \geq 1. \tag{2.7}$$

*If $\sum_{k=1}^{\infty} \phi_k$, then $\lim_{k\to\infty} \lambda_k$ exists.*

**Lemma 2.4.** *[13] Suppose $\{\vartheta_k\}$,$\{\rho_k\}$ and $\{\iota_k\}$ be sequences in $[0, \infty)$ such that,*

$$\vartheta_{k+1} \leq \vartheta_k + \iota_k(\vartheta_k - \vartheta_{k-1}) + \rho_k, \forall\ k \geq 1, \sum_{k=1}^{+\infty} \rho_k < +\infty, \tag{2.8}$$

*and there exists a real number $\iota$ with $0 \leq \iota_k \leq \iota < 1$ for all $k \in \mathbb{N}$. Then one has:*

(i) $\sum_{k=1}^{+\infty} [\vartheta_k - \vartheta_{k-1}]_+ < +\infty$, *where* $[m]_+ : \max\{m, 0\}$;

(ii) $\lim_{k\to\infty} \vartheta_k = \vartheta^*$ *with* $\vartheta^* \in [0, +\infty)$.

**Lemma 2.5.** *[6] Let $\{s_n\}$ be a sequence in $\mathcal{H}$ and $\Psi$ be a nonempty subset of $\mathcal{H}$. If, for every $s^* \in \Psi$, $\{||s_n - s^*||\}$ converges and every weak sequential cluster point of $\{s_n\}$ belongs to $\Psi$, then $\{s_n\}$ converges weakly to a point in $\Psi$.*

## 3. Main Results

**Assumption 3.1.**

(A1) *$g, h : \mathbb{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ are proper, convex and lower semi-continuous functions.*

(A2) *$g$ is differentiable and its gradient $\nabla g$ is uniformly continuous.*

(A3) *Suppose that $\Gamma := arg\min(g + h)$ is nonempty and the following condition holds:*

(i) $0 \leq \theta_k \leq \theta_{k+1} \leq 1$,

(ii) $0 < \beta < \min\{\theta_1, \frac{\varepsilon - \sqrt{2\varepsilon}}{\varepsilon}\}$ *with* $\varepsilon > 2$,

(iii) $0 < \delta \leq \delta_k \leq \delta_{k+1} < \frac{1}{1+\varepsilon}$,

(iv) $\lambda_k$ *is a nonegative sequence such that* $\sum_{k=1}^{\infty} \lambda_k < +\infty$.

We present the following algorithm.

**Algorithm 3.1.** *Self adaptive step-size with two step inertial for CMP.*

*Initialization:- Given* $\tau_1 > 0$ *and* $\mu \in (0,1)$. *Let* $w_0, w_1, \in \mathbb{H}$ *be two initial points and set* $k = 1$
*Step1:- Compute*

$$\begin{cases} z_k = w_k + \beta(w_k - w_{k-1}), \\ u_k = w_k + \theta_k(w_k - w_{k-1}), \\ y_k = Prox_{\tau_k h}(I - \tau_k \nabla g)u_k. \end{cases} \tag{3.1}$$

*where*

$$\tau_{k+1} = \begin{cases} \min\left\{\frac{\mu\|u_k - u_{k+1}\|}{\|\nabla g(u_k) - \nabla g(y_k)\|}\right\}, & \text{if } \nabla g(u_k) \neq \nabla g(y_k), \\ \tau_k + \lambda_k, & \text{otherwise.} \end{cases} \tag{3.2}$$

*If* $y_k = u_k$, *then stop. Otherwise proceed to the next step.*
*Step 2:- Compute*

$$v_k = y_k + \tau_k(\nabla g(u_k) - \nabla g(y_k)) \tag{3.3}$$

*Step 3:- Compute*

$$w_{k+1} = (1 - \delta_k)z_k + \delta_k v_k. \tag{3.4}$$

*Set* $k \leftarrow k+1$, *and go to step* 1.

**Lemma 3.1.** *Suppose* $\tau_k$ *is a sequence generated by Algorithm* (3.1) *such that Assumption* (3.1) *holds. Then* $\tau_k$ *is well defined and* $\lim\limits_{k\to\infty} \tau_k = \tau \in \{min\{\frac{\mu}{K}, \tau_1\}, \tau_1 + \psi\}$, *for some* $k > 0$ *where* $\psi = \sum\limits_{k=1}^{\infty} \lambda_k$.

*Proof.* From the condition on $\nabla g$ in Assumption 3.1, we have that for any given $\varepsilon > 0$, there exists $M < +\infty$ such that

$$\|\nabla g(u_k) - \nabla g(y_k)\| \leq M\|u_k - y_k\| + \varepsilon. \tag{3.5}$$

Therefore for the case where $\nabla g(u_k) \neq \nabla g(y_k)$ and $k \geqslant 1$, we obtain

$$\frac{\mu\|u_k - y_k\|}{\|\nabla g(u_k) - \nabla g(y_k)\|} \geqslant \frac{\mu\|u_k - y_k\|}{M\|u_k - y_k\| + \varepsilon} = \frac{\mu\|u_k - y_k\|}{(M + \varepsilon_1)\|u_k - y_k\|} = \frac{\mu}{K} \tag{3.6}$$

where $\varepsilon_1 = \varepsilon\|u_k - y_k\|$ for some $\varepsilon_1 \in (0,1)$ and $K = M + \varepsilon_1$ with $k > 0$. Thus, using the definition of $\tau_{k+1}$, the sequence $\tau_k$ has a lower bound $\min\{\frac{\mu}{k}, \tau_1\}$ and upper bound $\tau_1 + \psi$. By Lemma 2.3, $\lim\limits_{k\to\infty} \tau_k$ exists and $\lim\limits_{k\to\infty} \tau_k = \tau$. Clearly, we obtain

$$\tau \in \{min\{\frac{\mu}{K}, \tau_1\}, \tau_1 + \psi\}. \tag{3.7}$$

□

**Lemma 3.2.** *Suppose that Assumption* (3.1) *hold and let* $\{w_k\}$ *be a sequence generated by Algorithm* (3.1). *Then for all* $x^* \in \Gamma$, *we have*

$$\|v_k - x^*\|^2 \leq \|u_k - x^*\|^2 - \Big(1 - \frac{\mu^2\tau_k^2}{\tau_{k+1}^2}\Big)\|u_k - y_k\|^2, \tag{3.8}$$

*and*

$$\|v_k - y_k\| \leq \frac{\mu\tau_k}{\tau_{k+1}}\|u_k - y_k\|. \tag{3.9}$$

*Proof.* Let $x^* \in \Gamma$ , then this implies that $x^* = Prox_{\tau h}(I - \tau\nabla g)x^*$, where $\tau > 0$. Also from the definition of $\tau_k$, it is obvious that for $\nabla g(u_k) = \nabla g(y_k)$,

$$\|\nabla g(u_k) - \nabla g(y_k)\| \ \leq \ \frac{\mu}{\tau_{k+1}}\|u_k - y_k\|, \forall\, k \in \mathbb{N}. \tag{3.10}$$

Now, if $\nabla g(u_k) \neq \nabla g(y_k)$, then

$$\tau_{k+1} = \min\{\frac{\mu\|u_k - y_k\|}{\|\nabla g(u_k) - \nabla g(y_k)\|}, \tau_k + \lambda_k\} \leq \frac{\mu\|u_k - y_k\|}{\|\nabla g(u_k) - \nabla g(y_k)\|}, \tag{3.11}$$

which implies that $\|\nabla g(u_k) - \nabla g(y_k)\| \leq \frac{\mu}{\tau_{k+1}}\|u_k - y_k\|$. Hence, we conclude that (3.9) holds when $\nabla g(u_k) = \nabla g(y_k)$ and $\nabla g(u_k) \neq \nabla g(y_k)$.

Now, let $x^* \in \Gamma$, then from Lemma 2.2 and (3.9), we have

$$\begin{aligned}
\|v_k - x^*\|^2 &= \|y_k + \tau_k(\nabla g(u_k)) - \nabla g(y_k) - x^*\|^2 \\
&= \|y_k - x^* + \tau_k(\nabla g(u_k) - \nabla g(y_k))\|^2 \\
&= \|(y_k - u_k) + (u_k - x^*)\|^2 + \tau_k^2\|\nabla g(u_k) - \nabla g(y_k)\|^2 + 2\tau_k\langle y_k - x^*, \nabla g(u_k) - \nabla g(y_k)\rangle \\
&= \|y_k - u_k\| + \|u_k - x^*\| + \tau_k^2\|\nabla g(u_k) - \nabla g(y_k)\|^2 + 2\langle y_k - u_k, u_k - x^*\rangle \\
&+ 2\tau_k\langle y_k - x^*, \nabla g(u_k) - \nabla g(y_k)\rangle \\
&= \|u_k - x^*\| - \|y_k - u_k\|^2 + \tau_k^2\|\nabla g(u_k) - \nabla g(y_k)\|^2 + 2\langle y_k - u_k, y_k - x^*\rangle \\
&+ 2\tau_k\langle y_k - x^*, \nabla g(u_k) - \nabla g(y_k)\rangle \\
&= \|u_k - x^*\| - \|y_k - u_k\|^2 + \tau_k^2\|\nabla g(u_k) - \nabla g(y_k)\|^2 \\
&+ 2\langle y_k - x^*, y_k - u_k, +\tau_k(\nabla g(u_k) - \nabla g(y_k))\rangle.
\end{aligned} \tag{3.12}$$

Now, we need to show that

$$2\langle y_k - x^*, y_k - u_k, +\tau_k(\nabla g(u_k) - \nabla g(y_k))\rangle \leq 0 \tag{3.13}$$

By using the definition of $y_k$, we deduce that

$$(I - \tau_k\nabla g)u_k \in (I + \tau_k\partial h)y_k.$$

Thus, we can write

$$\tau_k = \frac{u_k - y_k}{\tau_k} - \nabla g(u_k), \text{ where } \tau_k \in \partial h(y_k).$$

By applying Lemma 2.1, we that $\nabla g + \partial h$ is maximal monotone. This implies that

$$\langle y_k - x^*, \nabla g(y_K) + t_k \rangle \geqslant 0, \tag{3.14}$$

and thus

$$\langle y_k - x^*, y_k - u_k + \tau_k(\nabla g(u_k) - \nabla g(y_k)) \rangle \leq 0. \tag{3.15}$$

Hence, on substituting (3.10) and (3.15) into (3.12), we obtain

$$\|v_k - x^*\|^2 \leq \|u_k - x^*\|^2 - (1 - \frac{\mu^2 \tau_k^2}{\tau_{k+1}^2})\|u_k - y_k\|^2. \tag{3.16}$$

Using Algorithm 3.1 and (3.10), we get

$$\begin{aligned} \|v_k - y_k\| &= \|y_k + \tau_k(\nabla g(u_k)) - \nabla g(y_k) - y_k\| \\ &\leq \tau_k \|\nabla g(u_k) - \nabla g(y_k)\| \\ &\leq \tau_k \frac{\mu}{\tau_k + 1}\|u_k - y_k\|. \end{aligned} \tag{3.17}$$

Hence, the proof completes. □

**Lemma 3.3.** *Suppose that Assumption 3.1 holds and let $\{w_k\}$ be a sequence generated by Algorithm 3.1 exists. Then $\lim_{k\to\infty} \|w_k - x^*\|$, where $x^* \in \Gamma$.*

*Proof.* Using Algorithm (3.1) and Lemma (3.2) we have.

$$\begin{aligned} \|w_{k+1} - x^*\|^2 \leq \ & \|(1-\delta_k)z_k + \delta_k v_k - x^*\|^2 \\ &= (1-\delta_k)\|z_k - x^*\|^2 + \delta_k\|v_k - x^*\|^2 - \delta_k(1-\delta_k)\|z_k - v_k\|^2 \\ &\leq (1-\delta_k)\|z_k - x^*\|^2 + \delta_k\|u_k - x^*\|^2 - \delta_k(1-\delta_k)\|z_k - v_k\|^2. \end{aligned} \tag{3.18}$$

From Algorithm 3.1 and Lemma 2.2 (iii), we have

$$\begin{aligned} \|z_k - x^*\|^2 \ &= \|w_k + \beta(w_k + w_{k+1}) - x^*\|^2 \\ &= \|(1+\beta)(w_k - x^*) - \beta(w_{k-1} - x^*)\|^2 \\ &= (1+\beta)\|w_k - x^*\|^2 - \beta\|w_{k-1} - x^*\|^2 + \beta(1+\beta)\|w_k - w_{k-1}\|^2. \end{aligned} \tag{3.19}$$

Similarly,

$$\begin{aligned} \|u_k - x^*\|^2 &= \|w_k + \theta_k(w_k - w_{k+1}) - x^*\|^2 \\ &= (1+\theta_k)\|w_k - x^*\|^2 - \theta_k\|w_{k-1} - x^*\|^2 + \theta_k(1+\theta_k)\|w_k - w_{k+1}\|. \end{aligned} \tag{3.20}$$

Observe that

$$w_{k+1} = (1-\delta_k)z_k + \delta_k v_k, \ k \geq 1, \tag{3.21}$$

and this implies

$$\|v_k - z_k\|^2 = \frac{1}{\delta_k^2}\|w_{k+1} - z_k\|^2$$

$$
\begin{aligned}
&= \frac{1}{\delta_k^2}\, \|w_{k+1} - w_k - \beta(w_k - w_{k+1})\|^2 \\
&= \frac{1}{\delta_k^2}\, \{\|w_{k+1} - w_k\|^2 + \beta^2\|w_k - w_{k-1}\|^2 - 2\beta\langle w_{k+1} - w_k, w_k - w_{k-1}\rangle\} \\
&\geq \frac{1}{\delta_k^2}\, \{\|w_{k+1} - w_k\|^2 + \beta^2\|w_k - w_{k-1}\|^2 - 2\beta\|w_{k+1} - w_k\|\|w_k - w_{k-1}\} \\
&\geq \frac{1}{\delta_k^2}\, \{\|w_{k+1} - w_k\|^2 + \beta^2\|w_k - w_{k-1}\|^2 - \beta\|w_{k+1} - w_k\|^2 - \beta\|w_k - w_{k-1}\|^2\} \\
&= \frac{\|w_{k+1} - w_k\|^2}{\delta_k^2} - \frac{\beta^2}{\delta_k^2}\|w_k - w_{k-1}\|^2 - \frac{\beta}{\delta_k^2}\|w_k - w_{k-1}\|^2 \\
&= \frac{1-\beta}{\delta_k}\|w_{k+1} - w_k\|^2 + \frac{\beta^2-\beta}{\delta_k}\|w_k - w_{k-1}\|. \qquad (3.22)
\end{aligned}
$$

On substituting (3.19),(3.20) and (3.22) into (3.18), we get

$$
\begin{aligned}
\|w_{k+1} - x^*\|^2 &\leq (1-\delta_k)\|z_k - x^*\|^2 + \delta_k\|u_k - x^*\|^2 - \delta_k(1-\delta_k)\|z_k - v_k\|^2 \\
&\leq (1-\delta_k)(1+\beta)\|w_k - x^*\|^2 - \beta(1-\delta_k)\|w_{k-1} - x^*\|^2 \\
&+ (1-\delta_k)\beta(1+\beta)\|w_k - w_{k-1}\|^2 \\
&+ \delta_k\theta_k(1+\theta_k)\|w_k - w_{k-1}\|^2 - (1-\delta_k) \\
&- \frac{1-\beta}{\delta_k}\|w_{k+1} - w_k\|^2 - \frac{\beta^2-\beta}{\delta_k}(1-\delta_k)\|w_k - w_{k-1}\|^2 \\
&= \Big(1 + \delta_k\theta_k + \beta(1-\delta_k)\Big)\|w_k - x^*\|^2 - \Big(\delta_k\theta_k + \beta(1-\delta_k)\Big)\|w_{k-1} - x^*\|^2 \\
&+ \gamma_k\|w_k - w_{k-1}\|^2 - \mu_k\|w_{k+1} - w_k\|^2, \qquad (3.23)
\end{aligned}
$$

where

$$\gamma_k :=\{(1-\delta_k)\beta(1+\beta) + \delta_k\theta_k(1+\theta_k) - \frac{1-\delta_k}{\delta_k}(\beta^2 - \beta)\},$$

and

$$\mu_k = \frac{1-\delta_k}{\delta_k}(1-\beta).$$

Define

$$\Theta_k := \|w_k - x^*\|^2 - [\delta_k\theta_k + \beta(1-\delta_k)]\|w_{k-1} - x^*\|^2 + \gamma_k\|w_k - w_{k-1}\|^2 \ \forall_k \ \geq 1.$$

consider with (3.23), one obtain.

$$
\begin{aligned}
\Theta_{k+1} - \Theta_k &= \|w_{k+1} - x^*\|^2 - [\delta_{k+1}\theta_{k+1} + \beta(1-\delta_{k+1})]\|w_k - x^*\|^2 \\
&+ \gamma_{k+1}\|w_{k+1} - w_k\|^2 - \|w_k - x^*\|^2 + [\delta_k\theta_k + \beta(1-\delta_k)]\|w_{k-1} - x^*\|^2 \\
&- \gamma_k\|w_k - w_{k-1}\|^2 \\
&= \|w_{k+1} - x^*\|^2 + [\delta_k\theta_k + \beta(1-\delta_k)]\|w_{k-1} - x^*\|^2 - \gamma_k\|w_k - w_{k-1}\|^2 \\
&- [1 + \delta_{k+1}\theta_{k+1} + \beta(1-\delta_{k+1})]\|w_k - x^*\|^2 + \gamma_{k+1}\|w_{k+1} - w_k\|^2
\end{aligned}
$$

$$\leq \left([1+\delta_k\theta_k+\beta(1-\delta_k)]-[1+\delta_{k+1}\theta_{k+1}+\beta(1-\delta_{k+1})]\right)\|w_k-x^*\|^2$$
$$+\gamma_{k+1}\|w_{k+1}-w_k\|^2-\mu_k\|w_{k+1}-w_k\|^2. \tag{3.24}$$

Since $0\leq\beta\leq\theta_1\leq\theta_k$, $\forall\geq 1$, and $\theta_k\leq\theta_{k+1}$, $\delta_k=\delta_{k+1}$, we obtain $\theta_k-\beta\geq 0$ and $\theta_{k+1}-\beta\geq 0$ such that $\theta_k-\beta\leq\theta_{k+1}-\beta$ and $\delta_k(\theta_k-\beta)\leq\delta_{k+1}(\theta_{k+1}-\beta)$ for any $k\geq 1$. So

$$\delta_k(\theta_k-\beta)-\delta_{k+1}(\theta_{k+1}-\beta)\leq 0. \tag{3.25}$$

Combine (3.25) with (3.24) for any $k\geq 1$ we obtain that,

$$\begin{aligned}\Theta_{k+1}-\Theta_k &\leq [\delta_k(\theta_k-\beta)-\delta_{k+1}(\theta_{k+1}-\beta)]\|w_k-x^*\|^2\\ &\quad+\gamma_{k+1}\|w_{k+1}-w_k\|^2-\mu_k\|w_{k+1}-w_k\|^2\\ &\leq-(\mu_k-\gamma_{k+1})\|w_{k+1}-w_k\|^2. \end{aligned}\tag{3.26}$$

Based on the condition (ii) and (iii) of Assumption (3.1), one gets

$$\begin{aligned}\mu_k-\gamma_{k+1} =&\frac{1-\delta_k}{\delta_k}(1-\beta)-(1-\delta_{k+1}\beta(1+\beta)-\delta_{k+1}\theta_{k+1}(1+\theta_{k+1}))\\ &+\frac{1-\delta_{k+1}}{\delta_{k+1}}(\beta^2-\beta)\\ &\geq\varepsilon(1-\beta)-2(1-\delta_{k+1})-2\delta_{k+1}+\varepsilon(\beta^2-\beta)\\ &=\varepsilon-\varepsilon\beta-2+\varepsilon\beta^2-\varepsilon\beta\\ &=\varepsilon\beta^2-2\varepsilon\beta-2.\end{aligned}\tag{3.27}$$

We have that $\varepsilon\beta^2-2\varepsilon\beta-2>0$ since $\beta<\frac{\varepsilon-\sqrt{2\varepsilon}}{\varepsilon}$. Substituting (3.27) into (3.26), we have

$$\Theta_{k+1}-\Theta_k\leq-\eta\|w_{k+1}-w_k\|^2, \tag{3.28}$$

where $\eta:=\varepsilon\beta^2-2\varepsilon\beta-2>0$. Therefore, $\{\Theta_k\}$ is non-increasing.
By definition of $\gamma_k$, there is

$$\begin{aligned}\Theta_k &=\|w_k-x^*\|^2-[\delta_k\theta_k+\beta(1-\delta_k)]\|w_{k-1}-x^*\|^2\\ &\quad+\gamma_n\|w_k-w_{k-1}\|^2\\ &\quad\geq\|w_k-x^*\|^2-[\delta_k\theta_k+\beta(1-\delta_k)]\|w_{k-1}-x^*\|^2.\end{aligned}\tag{3.29}$$

Which can imply that

$$\begin{aligned}\|w_k-x^*\|^2 &\leq\Theta_k+[\delta_k\theta_k+\beta(1-\delta_k)]\|w_{k-1}-x^*\|^2\\ &\leq[\Theta_k+\beta(1-\delta_k)]\|w_{k-1}-x^*\|^2\\ &\leq\Theta_k+\left[\frac{1}{1+\varepsilon}+\beta(1-\delta_k)\right]\|w_{k-1}-x^*\|^2\\ &=\Theta_1+\varsigma\|w_{k-1}-x^*\|^2\end{aligned}$$

.

$$
\begin{aligned}
&.\\
&.\\
&\le \varTheta_1(1+\varsigma+....+\varsigma^{k-2}+\varsigma^{k-1})+\varsigma^k\|w_0-x^*\|^2\\
&\le \frac{\varTheta_1}{1-\varsigma}+\varsigma^k\|w_0-x^*\|^2. \qquad (3.30)
\end{aligned}
$$

Where $\varsigma := \frac{1}{1+\epsilon}+\beta(1-\delta_k)\in(0,1)$ since $\beta<\frac{\varepsilon-\sqrt{2\varepsilon}}{\varepsilon}<\frac{\varepsilon}{(1-\varepsilon)(1-\delta)},\delta<1$ from the choice of $\beta$. So, $\{\|w_k-x^*\|\}$ is bounded. In fact,

$$
-\varsigma\|w_{k-1}-x^*\|^2\le\|w_k-x^*\|^2-\varsigma\|w_{k-1}-x^*\|^2\le\varTheta_k\le...\le\varTheta_1.
$$

and,

$$
\begin{aligned}
\varTheta_{k+1}&=\|w_{k+1}-x^*\|^2-[\delta_{k+1}\theta_{k+1}+\beta(1-\delta_{k+1})]\|w_k-x^*\|^2\\
&+\gamma_{k+1}\|w_{k+1}-w_k\|^2\\
&\ge-[\delta_{k+1}\theta_{k+1}+\beta(1-\delta_{k+1})]\|w_k-x^*\|^2\\
&\ge-\varsigma\|w_k-x^*\|^2\\
&\ge-\varsigma^{k+1}\|w_0-x^*\|^2-\frac{\varTheta_1}{1-\varsigma}. \qquad (3.31)
\end{aligned}
$$

By (3.28) and (3.31), one has

$$
\begin{aligned}
\eta\sum_{k=1}^{k}\|w_{k+1}-w_k\|^2&\le\varTheta_1-\varTheta_{k+1}\\
&\le\varsigma^{k+1}\|w_0-x^*\|^2+\frac{\varTheta_1}{1-\varsigma}. \qquad (3.32)
\end{aligned}
$$

It means that

$$
\sum_{k=1}^{\infty}\|w_{k+1}-x_k\|^2\le\frac{\varTheta_1}{\eta(1-\varsigma)}<\infty. \qquad (3.33)
$$

Therefore, $\lim_{k\to\infty}\|w_{k+1}-w_k\|=0$. Similarly, according the expression of $z_k$ and (ii) of Assumption 3.1, we get

$$
\lim_{k\to\infty}\|z_k-w_k\|^2=\lim_{k\to\infty}\|w_k-w_{k-1}\|^2=\lim_{k\to\infty}\|u_k-w_k\|=0 \qquad (3.34)
$$

From (3.23), we have

$$
\begin{aligned}
\|w_{k+1}-x^*\|^2&\le[1+\delta_k\theta_k+\beta(1-\delta_k)]\|w_k-x^*\|^2\\
&-[\delta_k\theta_k+\beta(1-\delta_k)]\|w_{k-1}-x^*\|^2\\
&+\gamma_k\|w_k-w_{k-1}\|^2-\mu_k\|w_{k+1}-w_k\|^2\\
&\le[1+\delta_k\theta_k+\beta(1-\delta_k)]\|w_k-x^*\|^2\\
&-[\delta_k\theta_k+\beta(1-\delta_k)]\|w_{k-1}-x^*\|^2+\gamma_k\|w_k-w_{k-1}\|^2\\
&\le\|w_k-x^*\|^2+[\delta_k\theta_k+\beta(1-\delta_k)](\|w_k-x^*\|^2-\|w_{k-1}-x^*\|^2)
\end{aligned}
$$

$$+\left((1-\delta)\beta(1+\beta)+\frac{2}{1+\varepsilon}+\frac{1-\delta}{\delta}(\beta^2-\beta)\right)\|w_k-w_{k-1}\|^2. \tag{3.35}$$

Where $\gamma_k \leq (1-\delta)\beta(1+\beta)+\frac{2}{1+\varepsilon}+\frac{1-\delta}{\delta}(\beta^2-\beta), \forall_k \geq 1$. And also $\delta_k+\beta(1-\delta_k) < \frac{1}{1+\varepsilon}+\beta(1-\delta) < 1$ since $\beta < \frac{\varepsilon-\sqrt{2\varepsilon}}{\varepsilon} < \frac{\varepsilon}{(1-\varepsilon)(1-\delta)} for\ all\ \delta \in (0,1)$ and $\varepsilon \in (2,\infty)$.
Invoking Lemma (2.4) in (3.35) it is clear to conclude that

$$\lim_{k\to\infty}\|w_k-x^*\| = l < \infty. \tag{3.36}$$

Hence, the proof completes. □

**Lemma 3.4.** *Let $\{w_k\}$ be a sequence generated by Algorithm 3.1, then we have that,*

$$\begin{cases}\lim\limits_{k\to\infty}\|y_k-u_k\| = 0,\\ \lim\limits_{k\to\infty}\|v_k-y_k\| = 0,\end{cases} \tag{3.37}$$

*Proof.* Using (3.34) and (3.36), we have from (3.21),

$$\begin{aligned}\|v_k-z_k\| &= \frac{1}{\delta_k}\|w_{k+1}-z_k\|\\ &\leq \frac{1}{\delta_k}\|w_{k+1}-w_k\|+\frac{\beta}{\delta_k}\|w_k-w_{k+1}\| \to 0, k\to\infty.\end{aligned} \tag{3.38}$$

Also

$$\|v_k-w_k\| \leq \|v_k-z_k\|+\|z_k-w_k\| \to 0 \ as\ k\to\infty.$$
$$\|u_k-v_k\| \leq \|v_k-w_k\|+\|u_k-w_k\| \to 0 \ as\ k\to\infty.$$

Since $\{w_k\}$ is bounded, we have that both $\{u_k\}$ and $\{v_k\}$ are bounded. Hence, from (3.16), we have for some $M>0$ that,

$$\begin{aligned}\Big(1-\frac{\mu^2\tau_k^2}{\tau_{k+1}^2}\Big)\|u_k-y_k\|^2 &\leq \|u_k-x^*\|^2-\|v_k-x^*\|^2\\ &= \Big(\|u_k-x^*\|+\|v_k-x^*\|\Big)\Big(\|u_k-x^*\|-\|v_k-x^*\|\Big)\\ &\leq M\|u_k-v_k\| \to 0, k\to\infty.\end{aligned}$$

Thus,

$$\lim_{k\to\infty}\|u_k-y_k\| = 0. \tag{3.39}$$

It is obvious from (3.9) and (3.39), we have that

$$\lim_{k\to\infty}\|v_k-y_k\| = 0. \tag{3.40}$$

□

**Theorem 3.1.** *Let $\{w_k\}$ be generated by Algorithm 3.1 such that Assumption 3.1 hold. Then, every weak cluster of point $\{w_k\}$ belong to $\Gamma$.*

*Proof.* Let $\bar{x}$ be a weak sequential cluster point of $\{w_k\}$, meaning that $w_{k_l} \rightharpoonup \bar{x}$ as $l \to \infty$ for some subsequence $\{w_{k_l}\}$ of $\{w_k\}$. Let $(v, u) \in Gra(\nabla g + \partial h)$,that is , $u - \nabla g(v) \in \partial h(v)$. By the difination of $y_k$, we have that,

$$\frac{u_{k_l} - y_{k_l} - \tau_{k_l}\nabla g(u_{k_l})}{\tau_{k_l}} \in \partial h(y_{k_l})$$

Using the maximal monotonicity of $\partial h$, we have

$$\langle v - y_{k_l}, \nabla g(v) - \frac{u_{k_l} - y_{k_l} - \tau_{k_l}\nabla g(u_{k_l})}{\tau_{k_l}}\rangle \geq 0.$$

Thus, by monotonicity of $\nabla g$, we have

$$\begin{aligned}
\Big\langle v - y_{k_l}, u\rangle &\geq \langle v - y_{k_l}, \nabla g(v) + \frac{u_{k_l} - y_{k_l} - \tau_{k_l}\nabla g(u_{k_l})}{\tau_{k_l}}\Big\rangle \\
&= \langle v - y_{kl}, \nabla g(v) - \nabla g(y_{k_l})\rangle + \langle v - y_{k_l}, \nabla g(y_{k_l}) - \nabla g(u_{k_l})\rangle + \langle v - y_{k_l}, \frac{u_{k_l} - y_{k_l}}{\tau_{k_l}}\rangle. \\
&\geq \langle v - y_{k_l}, \nabla g(y_{k_l}) - \nabla g(u_{k_l})\rangle + \frac{1}{\tau_{k_l}}\langle v - y_{k_l}, u_{k_l} - y_{k_l}\rangle.
\end{aligned}$$

Using the continuity condition on $\nabla g$, $\lim\limits_{l\to\infty} \frac{1}{\tau_{k_l}} > 0$, (3.39) and (3.40) that,

$$\langle v - \bar{x}, u\rangle = \lim_{l\to\infty}\langle v - y_{k_l}, u\rangle \geq 0, \tag{3.41}$$

from which, together with the maximal monotonicity of $\nabla g + \partial h$, we obtain that $\bar{x} \in arg\min(g + h)$. Therefore, $\bar{x} \in \Gamma$, lastly by appliying Lemma 2.5, we conclude that $\{w_k\}$ converges weakly to a solution of $\Gamma$. □

## 4. Numerical Example

In this section, we showcase two numerical experiments to illustrate the effectiveness of our algorithms. We utilize the MATLAB software to carry out these experiments. We compare our proposed method with the methods given in [8] and [20]. We choose the control parameters as follows:

$\theta_k = \frac{k}{k+1}$, $\delta_k = \frac{1}{k+1}$, $\beta_k = \frac{1}{k+2}$ and $\lambda_k = \frac{1}{k^3}$. We also let $\tau_1 = 0.07$, $\mu = 0.5$. For [8, Algorithm 3], we set $\beta = 5$, $q_n = \frac{1}{k+1} + 1$, $f : C \to C$ is defined by $f(x) = \frac{1}{100}x$. We choose $\sigma = 2$, $\delta = 0.1$, $t_1 = 1$, $\gamma_k = \frac{1}{10k+1}$ and $\epsilon = \frac{1}{k^2}$ in [20, Algorithm 7].

**Example 4.1.** *Let $H = \mathbb{R}^m$. We consider the LASSO problem formulated as follows:*

$$\min_{x\in\mathbb{R}^m} \frac{1}{2}||Bx - d||^2 + \lambda||x||_1, \tag{4.1}$$

*where $\lambda \geq 0$ is a regularization parameter. To apply this formulation, we set $g(x) = \frac{1}{2}||Bx - d||^2$ and $h(x) = \lambda||x||_1$ in* (1.1), *where $B \in \mathbb{R}^{n\times m}$ is a random matrix and $d \in \mathbb{R}^n$ is a random vector. It is easy to see that $\nabla g(x) = B^T(Bx - d)$. By letting $\lambda = 1$, we also have that $prox_{\tau_k h}(x) = \arg\min_{y\in\mathbb{R}^m}\left\{\frac{1}{2\tau_k}||y - x||^2 + ||y||_1\right\}$. Furthermore,*

$$prox_{\tau_k h}(x) = [prox_{\tau_k|\cdot|}x_1, prox_{\tau_k|\cdot|}x_2, \cdots, prox_{\tau_k|\cdot|}x_m]^T$$

*where* $prox_{\tau_k|\cdot|}x_i = \max\{x_i - \tau_k, 0\}sign(x_i)$. *We select the initial points* $w_0$ *and* $w_1$ *randomly in* $\mathbb{R}^n$ *and then vary the values of* $n$ *and* $m$ *as follows:*

Case 1 $n = 400$ *and* $m = 500$;

Case 2 $n = 300$ *and* $m = 700$;

Case 3 $n = 500$ *and* $m = 800$;

Case 4 $n = 700$ *and* $m = 1000$.

*The execution of the process is stopped at* $\|w_{k+1} - w_k\| \leq 10^{-4}$. *The result of this example is given in figure 1.*

**Example 4.2.** *Suppose* $H_1 = \mathbb{R}^m$ *and* $H_2 = \mathbb{R}^n$. *Let us consider the objective function given as follows:*

$$S(x) := \partial_C x + \partial_Q(Ax),$$

*where* $A \in \mathbb{R}^{n\times m}$ *is a random matrix,* $\partial_C$ *and* $\partial_Q$ *denote the indicator function of two nonempty, closed and convex subsets* $C$ *and* $Q$ *of* $H_1$ *and* $H_2$, *respectively. Then, the problem* (1.1) *reduces to*

$$\min_{x\in C} \frac{1}{2}\|(I - P_Q)Ax\|^2.$$

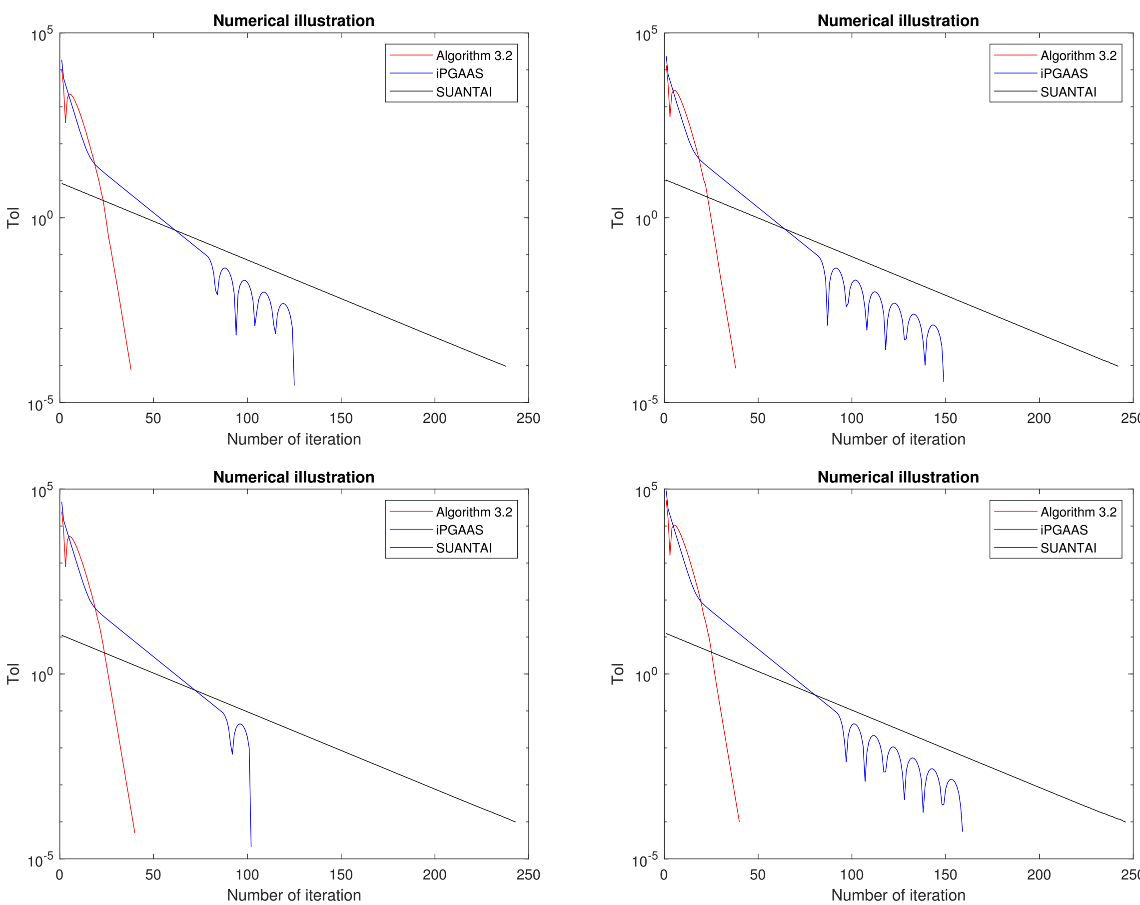


Figure 1. Numerical results for Example, Top Case I; Bottom Case II

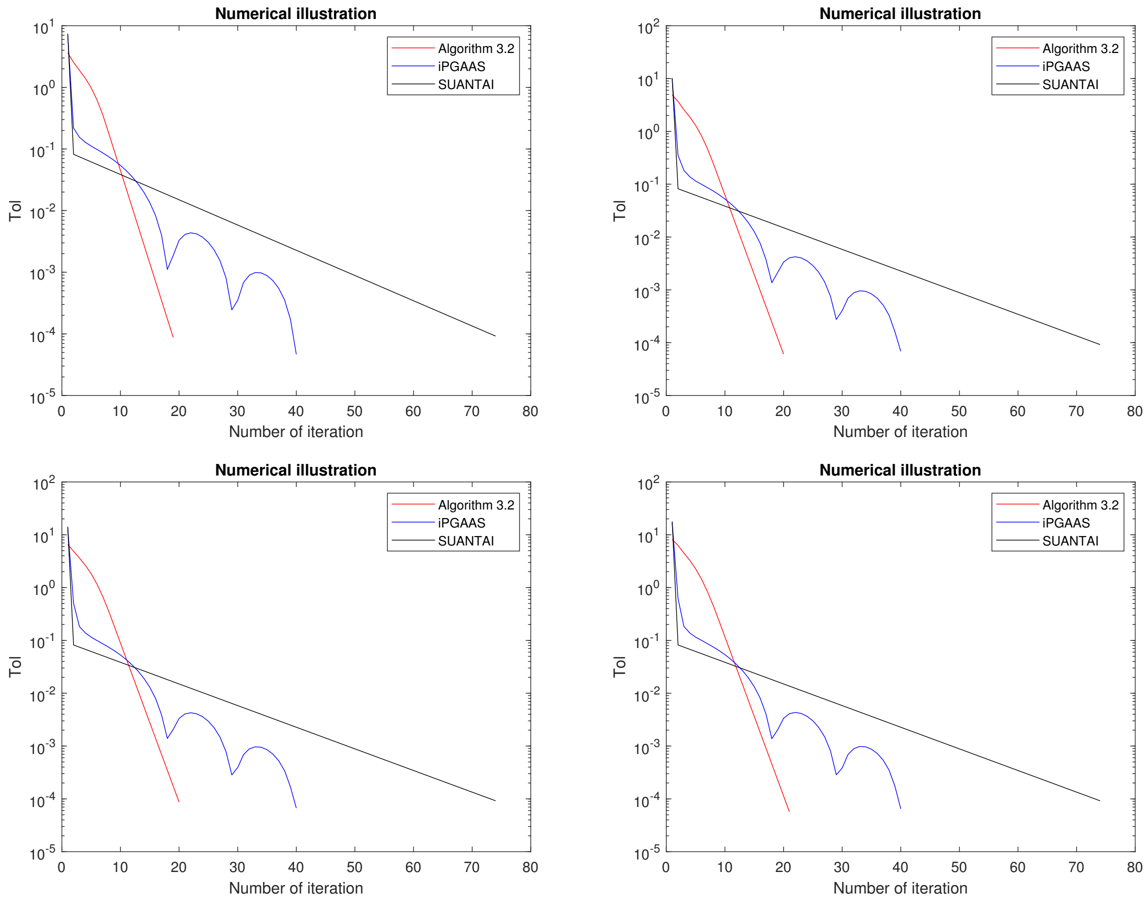


FIGURE 2. Top left: Case 1, Top right: Case 2, Bottom left: Case 3, Bottom right: Case 4.

For this formulation to hold, we set $n = m$ and $A$ is the identity matrix. Let $C = \{x \in \mathbb{R}^m : \|x\| \le r\}$ with a random number $r > 0$, $Q = \{y \in \mathbb{R}^n : \langle a, y\rangle \le b\}$ with a random vector $a \in \mathbb{R}^m$ and a random $b \in \mathbb{R}$. We choose the initial points $w_0$ and $w_1$ randomly in $\mathbb{R}^m$ and then vary the value $m$ as follows:

Case I $m = 200$;

Case II $m = 400$;

Case III $m = 700$;

Case IV $m = 1000$.

The execution of the process is stopped at $\|w_{k+1} - w_k\| \le 10^{-4}$. The result of this example is given in figure 1.

**Conflicts of Interest:** The authors declare that there are no conflicts of interest regarding the publication of this paper.

## References

[1] H.A. Abass, Linear Convergence of Alternating Inertial Tseng-type Method for Solving Inclusion Problems on Hadamard Manifolds, Proc. Edinb. Math. Soc. (2024). https://doi.org/10.1017/s0013091524000543.

[2] H.A. Abass, O.K. Oyewole, K.O. Aremu, L.O. Jolaoso, Self-adaptive Technique with Double Inertial Steps for Inclusion Problem on Hadamard Manifolds, J. Oper. Res. Soc. China (2024). https://doi.org/10.1007/s40305-024-00537-0.

[3] T.O. Alakoya, O.J. Ogunsola, O.T. Mewomo, An Inertial Viscosity Algorithm for Solving Monotone Variational Inclusion and Common Fixed Point Problems of Strict Pseudocontractions, Bol. Soc. Mat. Mex. 29 (2023), 31. https://doi.org/10.1007/s40590-023-00502-6.

[4] F. Alvarez, H. Attouch, An Inertial Proximal Method for Maximal Monotone Operators via Discretization of a Nonlinear Oscillator with Damping, Set-Valued Anal. 9 (2001), 3–11. https://doi.org/10.1023/A:1011253113155.

[5] V. Barbu, Nonlinear Semigroups and Differential Equations in Banach Spaces, Springer, 1976.

[6] H.H. Bauschke, P.L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer, 2011.

[7] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic Decomposition by Basis Pursuit, SIAM J. Sci. Comput. 20 (1998), 33–61. https://doi.org/10.1137/s1064827596304010.

[8] X. Chen, P. Duan, Inertial Self-adaptive Algorithms for Solving Non-smooth Convex Optimization Problems, Numer. Algorithms 98 (2024), 133–163. https://doi.org/10.1007/s11075-024-01788-x.

[9] P.L. Combettes, Solving Monotone Inclusions Via Compositions of Nonexpansive Averaged Operators, Optimization 53 (2004), 475–504. https://doi.org/10.1080/02331930412331327157.

[10] Y. Guo, W. Cui, Strong Convergence and Bounded Perturbation Resilience of a Modified Proximal Gradient Algorithm, J. Inequalities Appl. 2018 (2018), 103. https://doi.org/10.1186/s13660-018-1695-x.

[11] C. Izuchukwu, M. Aphane, K.O. Aremu, Two-step Inertial Forward–reflected–anchored–backward Splitting Algorithm for Solving Monotone Inclusion Problems, Comput. Appl. Math. 42 (2023), 351. https://doi.org/10.1007/s40314-023-02485-6.

[12] L. Jolaoso, H. Abass, O. Mewomo, A Viscosity-proximal Gradient Method with Inertial Extrapolation for Solving Certain Minimization Problems in Hilbert Space, Arch. Math. (2019), 167–194. https://doi.org/10.5817/am2019-3-167.

[13] P. Maingé, Convergence Theorems for Inertial Km-type Algorithms, J. Comput. Appl. Math. 219 (2008), 223–236. https://doi.org/10.1016/j.cam.2007.07.021.

[14] A.A. Mebawondu, H.A. Abass, O.K. Oyewole, An Accelerated Tseng Type Method for Solving Zero Point Problems and Certain Optimization Problems, Afr. Mat. 36 (2025), 13. https://doi.org/10.1007/s13370-024-01217-1.

[15] L. Mokaba, H.A. Abass, A. Adamu, Two Step Inertial Tseng Method for Solving Monotone Variational Inclusion Problem, Results Appl. Math. 25 (2025), 100545. https://doi.org/10.1016/j.rinam.2025.100545.

[16] A. Moudafi, M. Oliny, Convergence of a Splitting Inertial Proximal Method for Monotone Operators, J. Comput. Appl. Math. 155 (2003), 447–454. https://doi.org/10.1016/s0377-0427(02)00906-8.

[17] Y. Nesterov, Gradient Methods for Minimizing Composite Objective Function, LIDAM Discussion Papers CORE 2007076, Université Catholique de Louvain, Center for Operations Research and Econometrics, 2007.

[18] G.N. Ogwo, T.O. Alakoya, O.T. Mewomo, Inertial Iterative Method with Self-adaptive Step Size for Finite Family of Split Monotone Variational Inclusion and Fixed Point Problems in Banach Spaces, Demonstr. Math. 55 (2022), 193–216. https://doi.org/10.1515/dema-2022-0005.

[19] K. Sakurai, H. Iiduka, Acceleration of the Halpern Algorithm to Search for a Fixed Point of a Nonexpansive Mapping, Fixed Point Theory Appl. 2014 (2014), 202. https://doi.org/10.1186/1687-1812-2014-202.

[20] S. Suantai, P. Thongsri, A Fast Forward-Backward Algorithm Using Linesearch and Inertial Techniques for Convex Bi-Level Optimization Problems with Applications, Carpathian J. Math. 40 (2024), 769–788. https://doi.org/10.37193/CJM.2024.03.15.

[21] B. Tan, H.A. Abass, S. Li, O.K. Oyewole, Two Accelerated Double Inertial Algorithms for Variational Inequalities on Hadamard Manifolds, Commun. Nonlinear Sci. Numer. Simul. 145 (2025), 108734. https://doi.org/10.1016/j.cnsns.2025.108734.

[22] K. Tan, H. Xu, Approximating Fixed Points of Nonexpansive Mappings by the Ishikawa Iteration Process, J. Math. Anal. Appl. 178 (1993), 301–308. https://doi.org/10.1006/jmaa.1993.1309.

[23] R. Tibshirani, Regression Shrinkage and Selection Via the Lasso, J. R. Stat. Soc. Ser. B: Stat. Methodol. 58 (1996), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

[24] P. Tseng, A Modified Forward-Backward Splitting Method for Maximal Monotone Mappings, SIAM J. Control Optim. 38 (2000), 431–446. https://doi.org/10.1137/S0363012998338806.