C-FISTA Type Projection Algorithm for Quasi-Variational Inequalities

Yonghong Yao [*], Lateef O. Jolaoso[†], Yekini Shehu[‡]

May 17, 2024

Abstract

In this paper, we first propose a version of FISTA, called C-FISTA type gradient projection algorithm, for quasi-variational inequalities in Hilbert spaces and obtain linear convergence rate. Our results extend the results of Nesterov for C-FISTA algorithm for strongly convex optimization problem and other recent results in the literature where linear convergence results of C-FISTA are obtained for strongly convex composite optimization problems. For a comprehensive study, we also introduce a new version of gradient projection algorithm with momentum terms and give linear rate of convergence. We show the adaptability and effectiveness of our proposed algorithms through numerical comparisons with other related gradient projection algorithms that are in the literature for quasi-variational inequalities.

Keywords: Quasi-variational inequalities; C-FISTA; Strongly Monotone; Hilbert spaces

2010 MSC classification: 47H05, 47J20, 47J25, 65K15, 90C25.

1 Introduction

Let us take H as a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$ and $K \subset H$ a nonempty, closed and convex subset. Assume that $\mathcal{A} : H \to H$ is a nonlinear operator and $K : H \rightrightarrows H$ is a set-valued mapping which associates for any element $u \in H$ a closed and convex set $K(u) \subset H$. The *Quasi-Variational Inequality* (QVI), is defined by: find $x_* \in H$ such that $x_* \in K(x_*)$ and

$$\langle \mathcal{A}(x_*), x - x_* \rangle \geq 0 \text{ for all } x \in K(x_*). \quad (1)$$

[*]School of Mathematical Sciences, Tiangong University, Tianjin 300387, China; Department of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan; and Center for Advanced Information Technology, Kyung Hee University, Seoul 02447, South Korea; e-mail: yyhtgu@hotmail.com

[†]School of Mathematical Sciences, University of Southampton, SO17 1BJ, United Kingdom; Department of Mathematics and Applied Mathematics, Sefako Makgatho Health Sciences University, P.O. Box 94 Medunsa 0204, Pretoria, South Africa; e-mail: l.o.jolaoso@soton.ac.uk.

[‡](Corresponding Author) School of Mathematical Sciences, Zhejiang Normal University, Jinhua 321004, People's Republic of China; e-mail: yekini.shehu@zjnu.edu.cn

In the case of $K(x) \equiv K$ for all $x \in H$, then the QVI (1) becomes the classical variational inequality considered in ([20, 21, 26, 42]), which is to find $x_* \in K$ such that

$$\langle \mathcal{A}(x_*), x - x_* \rangle \geq 0 \text{ for all } x \in K. \tag{2}$$

Antipin et al. [3] introduced the following gradient projection algorithm to solve QVI (1)

$$x_{k+1} = P_{K(x_k)}(x_k - \gamma \mathcal{A}(x_k)) \tag{3}$$

and the following extragradient method

$$\begin{cases} y_k = P_{K(x_k)}(x_k - \gamma \mathcal{A}(x_k)), \\ x_{k+1} = P_{K(x_k)}(x_k - \gamma \mathcal{A}(y_k)) \end{cases} \tag{4}$$

Consequently, Antipin et al. [3] proved strong convergence results for the above proposed algorithms (3) and (4) to solve QVI (1) under the conditions that $K(x) := c(x) + K$, c Lipschitz continuous and $\mathcal{A}$ is strongly monotone and Lipschitz continuous. Related results to [3] can also be found in [31, 33–35].

It should be noted that the extragradient method (4) requires two projection computations and two evaluations of $\mathcal{A}$ at each iteration, which could potentially increase the computation complexities of extragradient algorithm (4). Motivated by the results in [3], Mijajlović et al. [30] designed the algorithms

$$x_{k+1} = (1 - \alpha_k)x_k + \alpha_k P_{K(x_k)}(x_k - \gamma \mathcal{A}(x_k)) \tag{5}$$

and

$$\begin{cases} y_k = (1 - \beta_k)x_k + \beta_k P_{K(x_k)}(x_k - \gamma \mathcal{A}(x_k)), \\ x_{k+1} = (1 - \alpha_k)x_k + \alpha_k P_{K(y_k)}(y_k - \gamma \mathcal{A}(y_k)) \end{cases} \tag{6}$$

with $\alpha_k \in (0, 1], \beta_k \in [0, 1]$ and strong convergence results obtained for QVI (1) where $\mathcal{A}$ is a Lipschitz continuous strongly monotone operator and condition (15) is assumed (for which the case $K(x) := c(x) + K, x \in H$ is fulfilled). It is noted that algorithm (3) is a special case of both algorithms (5) and (6). Some other related results to Mijajlović et al. [30] can also be found in [8, 17, 18, 27].

Motivated by the recent interests in iterative algorithms with inertial extrapolation step studied in [1, 2, 5–7, 9–12, 15, 28, 29, 37, 38, 40] and other related papers, Shehu et al. in [41] proposed the following gradient projection method with inertial step:

$$\begin{cases} y_k = x_k + \theta_k(x_k - x_{k-1}), \\ x_{k+1} = (1 - \alpha_k)y_k + \alpha_k P_{K(y_k)}(y_k - \gamma \mathcal{A}(y_k)), \end{cases} \tag{7}$$

where $\alpha_k \in (0, 1)$ and $0 \leq \theta_k \leq \theta < 1$. Shehu et al. in [41] obtained strong convergence results (with no linear convergence) for QVI (1) with $\mathcal{A}$ strongly monotone and Lipschitz continuous. Similarly, Çopur et al. [16] studied the gradient projection algorithm (3) with two inertial steps, which is an extension of inertial gradient projection algorithm (7).

Recently, several accelerated versions of FISTA [9] have been introduced for linear convergence of strongly convex composite optimization problems (see, for example, [13, 14, 22, 23, 43]). Quite recently, the following new variation of FISTA, called C-FISTA was presented in [24, Algorithm 1]

$$\begin{cases} w_k = \frac{1}{1+\theta} x_k + \frac{\theta}{1+\theta} z_k, \\ x_{k+1} = \text{Prox}_{\frac{R}{rL}}(w_k - \frac{1}{rL} H(Bw_k)), \\ z_{k+1} = (1-\theta) z_k + \theta w_k + \alpha(x_{k+1} - w_k) \end{cases} \tag{8}$$

(please, see [24, Algorithm 1] for the choices of r, θ, α) for a class of composite optimization model:

$$\min_{x \in X \subseteq \mathbb{R}^n} H(B(x)) + R(x), \tag{9}$$

where X is closed and convex, $H : \mathbb{R}^m \to \mathbb{R}$ is smooth and convex, $R : \mathbb{R}^n \to \mathbb{R}$ convex but potentially non-smooth, $B : \mathbb{R}^n \to \mathbb{R}^m$ a smooth mapping, and HoB is a convex function over X, and linear convergence results obtained [24, Theorem 1].

Our Contributions. Motivated by algorithm (8), our aim in this paper is to explore the linear convergence of C-FISTA type projection algorithm for QVI (1) with the possibility of improving on convergence speed of gradient projection (3) and design an algorithmic version of (8) for QVI (1). Summarily, we

- introduce a C-FISTA, which is a gradient projection algorithm with a golden ratio constant momentum and correction term for QVI (1), which can be considered as an algorithmic extension of convex composite optimization model (8) (with $B = I$) to QVI (1);
- introduce another new fast gradient projection algorithm with momentum terms;
- obtain linear convergence results for the two proposed algorithms under some standard conditions;
- provide numerical tests to confirm the superiority of our proposed algorithm over related gradient projection algorithms for QVI (1) in the literature.

Outline. We outline the paper as, viz: Section 2 entails some basic facts, concepts, and lemmas, which are needed in the linear convergence analysis. In Section 3, we introduce a C-FISTA algorithm and another fast gradient projection algorithm with momentum with their corresponding linear convergence results given. Section 4 discusses the numerical implementations of the proposed algorithms with versatility and efficiency against other related algorithms while in Section 5, we give a brief summary of our results.

2 Preliminaries

Definition 2.1. *Given an operator $\mathcal{A} : H \to H$,*

- $\mathcal{A}$ *is called* L-Lipschitz continuous *($L > 0$), if*

$$\|\mathcal{A}(x) - \mathcal{A}(y)\| \leq L\|x - y\| \text{ for all } x, y \in H. \tag{10}$$

- $\mathcal{A}$ *is called* μ-strongly monotone *($\mu > 0$), if*

$$\langle \mathcal{A}(x) - \mathcal{A}(y), x - y \rangle \geq \mu \|x - y\|^2 \text{ for all } x, y \in H. \tag{11}$$

For each $x \in H$, there exists a unique nearest point in K, denoted by $P_K(x)$, such that

$$\|x - P_K(x)\| \leq \|x - y\| \text{ for all } y \in K. \tag{12}$$

This operator $P_K : H \to K$ is called the *metric projection* of H onto K, characterized [25, Section 3] by

$$P_K(x) \in K \tag{13}$$

and

$$\langle x - P_K(x), P_K(x) - y \rangle \geq 0 \text{ for all } x \in H, \; y \in K. \tag{14}$$

We state the following sufficient conditions for the existence of solutions of QVIs (1) given in [36].

Lemma 2.2. *Let $\mathcal{A} : H \to H$ be L-Lipschitz continuous and μ-strongly monotone on H and $K(\cdot)$ be a set-valued mapping with nonempty, closed and convex values such that there exists $\lambda \geq 0$ such that*

$$\|P_{K(x)}(z) - P_{K(y)}(z)\| \leq \lambda \|x - y\|, \quad x, y, z \in H, \quad \lambda + \sqrt{1 - \frac{\mu^2}{L^2}} < 1. \tag{15}$$

Then the QVI (1) has a unique solution.

The fixed point formulation of the QVI (1) is given by

Lemma 2.3. *Let $K(\cdot)$ be a set-valued mapping with nonempty, closed and convex values in H. Then $x_* \in K(x_*)$ is a solution of the QVI (1) if and only if for any $\gamma > 0$ it holds that*

$$x_* = P_{K(x_*)}(x_* - \gamma \mathcal{A}(x_*)).$$

The following lemma is needed in our convergence analysis.

Lemma 2.4. *If $x, y \in H$, we have*

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

(ii) Assume that $x, y, z \in H$ and $\alpha, \beta, \gamma \in \mathbb{R}$ such that $\alpha + \beta + \gamma = 1$. Then

$$\begin{aligned} \|\alpha x + \beta y + \gamma z\|^2 &= \alpha\|x\|^2 + \beta\|y\|^2 + \gamma\|z\|^2 - \alpha\beta\|x - y\|^2 \\ &\quad -\alpha\gamma\|x - z\|^2 - \beta\gamma\|y - z\|^2. \end{aligned}$$

3 Main Results

In this section, we introduce our C-FISTA-type gradient projection algorithm and another fast gradient projection algorithm with momentum alongside their linear convergence results. Throughout this paper, we assume that $\gamma \geq 0$ satisfies the following condition:

Assumption 3.1.

$$\left|\gamma - \frac{\mu}{L^2}\right| < \frac{\sqrt{\mu^2 - L^2\lambda(2-\lambda)}}{L^2}. \tag{16}$$

Next is the proposed C-FISTA gradient projection algorithm below.

Algorithm 1 C-FISTA Gradient Projection Algorithm

1: Choose $\theta \geq 0$ and pick $x_0 = z_0 \in H$. Set $k := 0$.
2: Given the current iterates x_k and z_k, compute

$$\begin{cases} w_k = (1-\theta)x_k + \theta z_k, \\ x_{k+1} = P_{K(w_k)}(w_k - \gamma \mathcal{A}(w_k)), \\ z_{k+1} = \frac{\theta}{1+\theta} z_k + \frac{1}{1+\theta} w_k + \theta(x_{k+1} - w_k) \end{cases} \tag{17}$$

3: Set $k \leftarrow k+1$, and **return to 2**.

Remark 3.2.
(a) The proposed C-FISTA Gradient Projection Algorithm 1 features two momentum terms w_k and z_k; correction term $\theta(x_{k+1} - w_k)$; and the basic gradient projection step $x_{k+1} = P_{K(w_k)}(w_k - \gamma \mathcal{A}(w_k))$. The basic gradient projection step $x_{k+1} = P_{K(w_k)}(w_k - \gamma \mathcal{A}(w_k))$ for QVI (1) has been studied in [3, 4, 32, 33, 39].

(b) C-FISTA Gradient Projection Algorithm 1 can be considered as an extension of (8) for solving QVI (1).

(c) Our C-FISTA Gradient Projection Algorithm 1 reduces to the basic gradient projection algorithm for QVI (1) studied in [3,4,32,33,39] when $\theta = 0$. Consequently, in the convergence analysis of Algorithm 1, we do not need to consider the case $\theta = 0$.

Theorem 3.3. *Consider the QVI (1) with $\mathcal{A}$ being μ-strongly monotone and L-Lipschitz continuous and assume there exists $\lambda \geq 0$ such that (15) holds. Let $\{x_k\}$ be any sequence generated by Algorithm 1 with $\gamma \geq 0$ satisfying (16), and $0 < \theta \leq \frac{1}{\phi}$, $\phi := \frac{1+\sqrt{5}}{2}$. Then $\{x_k\}$ and $\{z_k\}$ converge linearly to the unique solution $x_* \in K(x_*)$ of the QVI (1).*

Proof. For the unique solution x_* of (1), we have

$$\begin{aligned} \|x_{k+1} - x_*\| &= \|P_{K(w_k)}(w_k - \gamma \mathcal{A}(w_k)) - P_{K(x_*)}(x_* - \gamma \mathcal{A}(x_*))\| \\ &\leq \|P_{K(w_k)}(w_k - \gamma \mathcal{A}(w_k)) - P_{K(x_*)}(w_k - \gamma \mathcal{A}(w_k))\| \\ &\quad + \|P_{K(x_*)}(w_k - \gamma \mathcal{A}(w_k)) - P_{K(x_*)}(x_* - \gamma \mathcal{A}(x_*))\| \\ &\leq \lambda \|w_k - x_*\| + \|w_k - x_* + \gamma(\mathcal{A}(x_*) - \mathcal{A}(w_k))\|. \end{aligned} \tag{18}$$

Since $\mathcal{A}$ is μ-strongly monotone and $L-$Lipschitz continuous, we get

$$\begin{aligned}\|w_k - x_* - \gamma(\mathcal{A}(x_*) - \mathcal{A}(w_k))\|^2 &= \|w_k - x_*\|^2 - 2\gamma\langle \mathcal{A}(w_k) - \mathcal{A}(x_*), w_k - x_*\rangle \\ &\quad +\gamma^2\|\mathcal{A}(w_k) - \mathcal{A}(x_*)\|^2 \\ &\leq (1 - 2\mu\gamma + \gamma^2 L^2)\|w_k - x_*\|^2. \end{aligned} \tag{19}$$

Combining (18) and (19), we get

$$\begin{aligned}\|x_{k+1} - x_*\| &\leq \lambda\|w_k - x_*\| + \sqrt{1 - 2\mu\gamma + \gamma^2 L^2}\|w_k - x_*\| \\ &= \beta\|w_k - x_*\|, \end{aligned} \tag{20}$$

where

$$\beta := \sqrt{1 - 2\mu\gamma + \gamma^2 L^2} + \lambda. \tag{21}$$

We next show that $\beta \in (0, 1)$. Observe that $0 < \sqrt{1 - 2\mu\gamma + \gamma^2 L^2} \Leftrightarrow 0 < 1 - 2\mu\gamma + \gamma^2 L^2 \Leftrightarrow \frac{\mu^2 - L^2}{L^4} < \left(\gamma - \frac{\mu}{L^2}\right)^2$. Since $\mathcal{A}$ is μ-strongly monotone and L-Lipschitz continuous, we have $\mu \leq L$. Therefore, $\frac{\mu^2 - L^2}{L^4} \leq 0$. Note also that $\left(\gamma - \frac{\mu}{L^2}\right)^2 > 0$ by (16). Hence, $\frac{\mu^2 - L^2}{L^4} < \left(\gamma - \frac{\mu}{L^2}\right)^2$. Noting that $\lambda \geq 0$, we then obtain $0 < \sqrt{1 - 2\mu\gamma + \gamma^2 L^2} + \lambda$.

Furthermore, by (16), we have that $\gamma - \frac{\mu}{L^2} < \frac{\sqrt{\mu^2 - L^2\lambda(2-\lambda)}}{L^2}$ which means that $\left(\gamma - \frac{\mu}{L^2}\right)^2 < \frac{L^2\lambda(\lambda-2)+\mu^2}{L^4}$ and so we have $\left(\gamma - \frac{\mu}{L^2}\right)^2 < \frac{\lambda(\lambda-2)}{L^2} + \frac{\mu^2}{L^4}$. Expanding, we have $\gamma^2 - \frac{2\mu\gamma}{L^2} < \frac{\lambda(\lambda-2)}{L^2}$. This implies $\gamma^2 L^2 - 2\mu\gamma < \lambda(\lambda-2)$. Thus, $1 - 2\mu\gamma + \gamma^2 L^2 < (1-\lambda)^2$. Hence, $\sqrt{1 - 2\mu\gamma + \gamma^2 L^2} + \lambda < 1$. Therefore, we have $\beta \in (0, 1)$.

We also have from Algorithm 1 and Lemma 2.4 (ii) that

$$\begin{aligned}\|w_k - x_*\|^2 &= \|(1-\theta)(x_k - x_*) + \theta(z_k - x_*)\|^2 \\ &= (1-\theta)\|x_k - x_*\|^2 + \theta\|z_k - x_*\|^2 \\ &\quad -\theta(1-\theta)\|x_k - z_k\|^2 \end{aligned} \tag{22}$$

and

$$\begin{aligned}\|z_{k+1} - x_*\|^2 &= \|\frac{\theta}{1+\theta}(z_k - x_*) + \Big(\frac{1}{1+\theta} - \theta\Big)(w_k - x_*) + \theta(x_{k+1} - x_*)\|^2 \\ &= \frac{\theta}{1+\theta}\|z_k - x_*\|^2 + \Big(\frac{1}{1+\theta} - \theta\Big)\|w_k - x_*\|^2 \\ &\quad +\theta\|x_{k+1} - x_*\|^2 - \frac{\theta}{1+\theta}\Big(\frac{1}{1+\theta} - \theta\Big)\|w_k - z_k\|^2 \\ &\quad -\theta\Big(\frac{1}{1+\theta} - \theta\Big)\|x_{k+1} - w_k\|^2 - \frac{\theta^2}{1+\theta}\|x_{k+1} - z_k\|^2. \end{aligned} \tag{23}$$

Observe that

$$\frac{\theta(1+\theta)\beta^2}{1 - \theta + \theta^2 + \theta^3} < \frac{(1+\theta)(1 - (1-\theta)\beta^2)}{1 - \theta + \theta^2 + \theta^3}$$

since $0 < 1 - \beta^2$. Now, choose $c > 0$ such that

$$\frac{\theta(1+\theta)\beta^2}{1-\theta+\theta^2+\theta^3} < c < \frac{(1+\theta)(1-(1-\theta)\beta^2)}{1-\theta+\theta^2+\theta^3}.$$

Note also that $1-\theta+\theta^2+\theta^3 > 0$ for $0 < \theta \leq \frac{1}{\phi}$. If we plug (22) into (20) and add the result with product of c and (23), we have

$$\begin{aligned}
\|x_{k+1}-x_*\|^2 \quad + \quad & c\|z_{k+1}-x_*\|^2 \leq (1-\theta)\beta^2\|x_k-x_*\|^2 + \theta\beta^2\|z_k-x_*\|^2 \\
& -\theta(1-\theta)\beta^2\|x_k-z_k\|^2 + \frac{c\theta}{1+\theta}\|z_k-x_*\|^2 \\
& +c\Big(\frac{1}{1+\theta}-\theta\Big)\|w_k-x_*\|^2 + c\theta\|x_{k+1}-x_*\|^2 \\
& -\frac{c\theta}{1+\theta}\Big(\frac{1}{1+\theta}-\theta\Big)\|w_k-z_k\|^2 - c\theta\Big(\frac{1}{1+\theta}-\theta\Big)\|x_{k+1}-w_k\|^2 \\
& -\frac{c\theta^2}{1+\theta}\|x_{k+1}-z_k\|^2 \\
= \quad & (1-\theta)\beta^2\|x_k-x_*\|^2 + \Big(\theta\beta^2 + \frac{c\theta}{1+\theta}\Big)\|z_k-x_*\|^2 \\
& +c\Big(\frac{1}{1+\theta}-\theta\Big)\|w_k-x_*\|^2 - \theta(1-\theta)\beta^2\|x_k-z_k\|^2 \\
& +c\theta\|x_{k+1}-x_*\|^2 - \frac{c\theta}{1+\theta}\Big(\frac{1}{1+\theta}-\theta\Big)\|w_k-z_k\|^2 \\
& -c\theta\Big(\frac{1}{1+\theta}-\theta\Big)\|x_{k+1}-w_k\|^2 - \frac{c\theta^2}{1+\theta}\|x_{k+1}-z_k\|^2.
\end{aligned}$$

Hence,

$$\begin{aligned}
& (1-c\theta)\|x_{k+1}-x_*\|^2 + c\|z_{k+1}-x_*\|^2 \leq (1-\theta)\beta^2\|x_k-x_*\|^2 \\
& +\Big(\theta\beta^2 + \frac{c\theta}{1+\theta}\Big)\|z_k-x_*\|^2 + c\Big(\frac{1}{1+\theta}-\theta\Big)(1-\theta)\|x_k-x_*\|^2 \\
& +c\Big(\frac{1}{1+\theta}-\theta\Big)\theta\|z_k-x_*\|^2 \\
= \quad & (1-\theta)\Big(\beta^2 + c\Big(\frac{1}{1+\theta}-\theta\Big)\Big)\|x_k-x_*\|^2 \\
& +\Big(\theta\beta^2 + \frac{c\theta}{1+\theta} + c\Big(\frac{1}{1+\theta}-\theta\Big)\theta\Big)\|z_k-x_*\|^2 \\
\leq \quad & \max\Bigg\{\frac{(1-\theta)\Big(\beta^2+c\Big(\frac{1}{1+\theta}-\theta\Big)\Big)}{1-c\theta}, \frac{\Big(\theta\beta^2+\frac{c\theta}{1+\theta}+c\Big(\frac{1}{1+\theta}-\theta\Big)\theta\Big)}{c}\Bigg\} \times \\
& \Big((1-c\theta)\|x_k-x_*\|^2 + c\|z_k-x_*\|^2\Big) \\
= \quad & \tau\Big((1-c\theta)\|x_k-x_*\|^2 + c\|z_k-x_*\|^2\Big), \qquad (24)
\end{aligned}$$

where

$$\tau := \max\Bigg\{\frac{(1-\theta)\Big(\beta^2+c\Big(\frac{1}{1+\theta}-\theta\Big)\Big)}{1-c\theta}, \frac{\Big(\theta\beta^2+\frac{c\theta}{1+\theta}+c\Big(\frac{1}{1+\theta}-\theta\Big)\theta\Big)}{c}\Bigg\}.$$

Observe that

$$\frac{(1-\theta)\Big(\beta^2+c\Big(\frac{1}{1+\theta}-\theta\Big)\Big)}{1-c\theta} < 1,$$

since $c < \frac{(1+\theta)(1-(1-\theta)\beta^2)}{1-\theta+\theta^2+\theta^3}$. Furthermore,

$$\frac{\left(\theta\beta^2 + \frac{c\theta}{1+\theta} + c\left(\frac{1}{1+\theta} - \theta\right)\theta\right)}{c} < 1,$$

since $\frac{\theta(1+\theta)\beta^2}{1-\theta+\theta^2+\theta^3} < c$. We also have $\frac{(1+\theta)(1-(1-\theta)\beta^2)}{1-\theta+\theta^2+\theta^3} \leq \frac{1}{\theta}$ for $0 < \theta \leq \frac{1}{\phi} = \frac{\sqrt{5}-1}{2}$ and this implies that $c < \frac{1}{\theta}$. Therefore, $\tau \in (0,1)$. Now, define

$$b_k := (1 - c\theta)\|x_k - x_*\|^2 + c\|z_k - x_*\|^2 \geq 0,$$

since $c < \frac{1}{\theta}$. Then we have from (24) that

$$\begin{aligned} b_{k+1} &\leq \tau b_k \\ &\vdots \\ &\leq \tau^{k+1} b_0. \end{aligned} \tag{25}$$

Consequently, we have that $\{b_k\}$ converges linearly to zero. Hence, we have that both $\{x_k\}$ and $\{z_k\}$ converge linearly to to the unique solution $x_* \in K(x_*)$ of the QVI (1). This completes the proof. □

In a special case when $K(x), x \in H$ is a "moving set". That is, the case when $K(x) := c(x) + K, x \in H$ where $c : H \to H$ is a λ-Lipschitz continuous mapping and $K \subset H$ is a nonempty, closed and convex subset. Then the Assumption (16) is automatically satisfied with the same value of λ (see [32]). The following result hold in this case.

Corollary 3.4. *Consider the QVI (1) with $\mathcal{A}$ being μ-strongly monotone and L-Lipschitz continuous and suppose that $K(x) := c(x) + K, x \in H$ where $c : H \to H$ is a λ-Lipschitz continuous mapping and K is a nonempty, closed and convex subset of H. Let $\{x_k\}$ be any sequence generated by Algorithm 1 with $\gamma \geq 0$ satisfying (16), and $0 < \theta \leq \frac{1}{\phi}$, $\phi := \frac{1+\sqrt{5}}{2}$. Then $\{x_k\}$ and $\{z_k\}$ converge linearly to the unique solution $x_* \in K(x_*)$ of the QVI (1).*

Next, we propose the second accelerated gradient projection algorithm with momentum terms for solving QVI (1).

Algorithm 2 Accelerated Gradient Projection Algorithm

1: Choose $x_0 = z_0 \in H$ and $\theta_k \in [0,1]$. Set $k := 0$.
2: Given the current iterates x_k and z_k, compute

$$\begin{cases} w_k = (1-\theta_k)x_k + \theta_k z_k, \\ x_{k+1} = P_{K(w_k)}(w_k - \gamma\mathcal{A}(w_k)), \\ z_{k+1} = (1-\theta_{k+1})z_k + \theta_{k+1}x_{k+1} \end{cases} \tag{26}$$

3: Set $k \leftarrow k+1$, and **return to 2**.

Theorem 3.5. *Consider the QVI (1) with $\mathcal{A}$ being μ-strongly monotone and L-Lipschitz continuous and assume there exists $\lambda \geq 0$ such that (15) holds. Let $\{x_k\}$ be any sequence generated by Algorithm 2 with $\gamma \geq 0$ satisfying (16), and $\theta_k \in \{1, 1-\beta^{2k}\}, \forall k \in \mathbb{N}$, where β is as defined in (21). Then $\{x_k\}$ and $\{z_k\}$ converge linearly to the unique solution $x_* \in K(x_*)$ of the QVI (1).*

Proof. For the unique solution x_* of (1), we have (noting (20) and (22))

$$\begin{aligned}\|x_{k+1}-x_*\|^2+\|z_{k+1}-x_*\|^2 \;\leq\; & (1-\theta_k)\beta^2\|x_k-x_*\|^2+\theta_k\beta^2\|z_k-x_*\|^2\\ & -\theta_k(1-\theta_k)\beta^2\|x_k-z_k\|^2+(1-\theta_{k+1})\|z_k-x_*\|^2\\ & +\theta_{k+1}\|x_{k+1}-x_*\|^2-\theta_{k+1}(1-\theta_{k+1})\|x_{k+1}-z_k\|^2.\end{aligned}$$

Hence,

$$\begin{aligned}(1-\theta_{k+1})\|x_{n+1}-x_*\|^2 \;+\; & \|z_{k+1}-x_*\|^2\\ & +\theta_{k+1}(1-\theta_{k+1})\|x_{k+1}-z_k\|^2\\ \leq\; & (1-\theta_k)\beta^2\|x_k-x_*\|^2+(\beta^2\theta_k+1-\theta_{k+1})\|z_k-x_*\|^2\\ & -\theta_k(1-\theta_k)\beta^2\|x_k-z_k\|^2+\theta_k(1-\theta_k)\beta^2\|x_k-z_{k-1}\|^2\\ & -\theta_k(1-\theta_k)\beta^2\|x_k-z_{k-1}\|^2\\ =\; & (1-\theta_k)\beta^2\|x_k-x_*\|^2+\beta^2\|z_k-x_*\|^2\\ & +\theta_k(1-\theta_k)\beta^2\|x_k-z_{k-1}\|^2-\theta_k(1-\theta_k)\beta^2\|x_k-z_k\|^2\\ & -\theta_k(1-\theta_k)\beta^2\|x_k-z_{k-1}\|^2, \qquad (27)\end{aligned}$$

where the last equality holds since $1-\beta^2+\beta^2\theta_k=\theta_{k+1}$. We then obtain from (27) that

$$b_{k+1} \leq \beta^2 b_k,$$

where $\{b_k\}$ is defined as

$$b_k := (1-\theta_k)\|x_k-x_*\|^2+\|z_k-x_*\|^2+\theta_k(1-\theta_k)\|x_k-z_{k-1}\|^2. \qquad (28)$$

Consequently, we have that $\{b_k\}$ converges linearly to zero. Consequently, we have that both $\{x_k\}$ and $\{z_k\}$ converge linearly to the unique solution x_* of (1). □

Corollary 3.6. *Consider the QVI (1) with $\mathcal{A}$ being μ-strongly monotone and L-Lipschitz continuous and suppose that $K(x) := c(x)+K, x \in H$ where $c : H \to H$ is a λ-Lipschitz continuous mapping and K is a nonempty, closed and convex subset of H. Let $\{x_k\}$ be any sequence generated by Algorithm 2 with $\gamma \geq 0$ satisfying (16), and and $\theta_k \in \{1, 1-\beta^{2k}\}, \forall k \in \mathbb{N}$, where β is as defined in (21). Then $\{x_k\}$ and $\{z_k\}$ converge linearly to the unique solution $x_* \in K(x_*)$ of the QVI (1).*

Remark 3.7.
(a) In the spirit of the algorithmic developments in [30, 41], the step

$$x_{k+1} = P_{K(w_k)}(w_k-\gamma\mathcal{A}(w_k))$$

in both Algorithm 1 and Algorithm 2 can be replaced with

$$x_{k+1} = (1-\alpha_k)w_k+\alpha_k P_{K(w_k)}(w_k-\gamma\mathcal{A}(w_k)), \alpha_k \in (0,1],$$

and the conclusion of Theorem 3.3 and Theorem 3.5 still obtained.

(b) Shehu et al. [41] and Çopur et al. [16] obtained strong convergence results for QVI (1) under the condition that $\mathcal{A}$ is μ-strongly monotone and L-Lipschitz continuous without linear rate of convergence. In this paper, we give linear rate of our proposed algorithms for the QVI (1) when $\mathcal{A}$ is μ-strongly monotone and L-Lipschitz continuous.

4 Numerical Examples

We give some numerical implementations of our proposed Algorithm 1 and give comparisons with some existing methods in the literature. All codes were written in MATLAB R2023a and performed on a PC Desktop Intel Core i5-8265U CPU 1.60GHz 1.80 GHz, RAM 16.00GB. We compare Algorithm 1 and Algorithm 2 with [3, 16, 41].

We choose to use the test problem library QVILIB taken from [19]; the feasible map K is assumed to be given by $K(x) := \{z \in \mathbb{R}^n : g(z,x) \leq 0\}$. We implemented Algorithm 1 and Algorithm 2 in Matlab. We implemented the projection over a convex set as the solution of a convex program. We considered the following performance measures for optimality and feasibility

$$\text{opt}(x) := -\min_z\{\mathcal{F}(x)^T(z-x) \,:\, z \in K(x)\}, \quad \text{feas}(x) := \|\max\{0, g(x,x)\}\|_\infty.$$

A point x^* is considered as a solution of the QVI if opt$(x^*) \leq$1e-3 and feas$(x^*) \leq$1e-3. As nonlinear programming solver we used the built-in function `fmincon` with the option of 'sqp' as its internal algorithm and maximum iteration `maxiter` = 1000. The QVILIB [19] comprises a diverse collection of test problems designed for evaluating algorithms used in solving QVIs (Quasi-Variational Inequalities). These problems encompass academic models, real-world applications, and discretized versions of infinite-dimensional QVIs, which model various engineering and physical phenomena. Additionally, the library provides an M-file named startinPoints, allowing users to obtain starting points for each test problem.
In our experiments, we specifically utilized problems tailored for academic purposes. These include `OutZ40, OutZ41, OutZ42, OutZ43, OutZ44, Set1A, Set2A, Box1A`, and `BiLin1A`. The feasible set $K(x)$ is defined as the intersection of a fixed set $\bar{K}$ and a moving set $\tilde{K}(x)$ that depends on the point x given by:

$$\begin{aligned} \bar{K} &:= \{y \in \mathbb{R}^k \mid g^I(y) \leq 0,\ M^I y + v^I = 0\}, \\ \tilde{K}(x) &= \{y \in \mathbb{R}^k \mid g^P(y,x) \leq 0,\ M^P(x)y + v^P(x) = 0\}. \end{aligned}$$

The comprehensive definitions of each problem can be found in [19]; however, we provide a brief description of each problem in Table 1.

In Table 1, the first column contains the name of the problem, the second column (n) contains the number of variables in the problem, column m_I contains the number of inequality constraints defining $\bar{K}$, column p_I contains the number of linear equalities in $\bar{K}$, the column m_P contains the number of inequality constraints

Table 1: Description of test problems used in the experiments

Problem name	n	m_I	p_I	m_P	p_P	n(start)
OutZ40	2	4	0	2	0	3
OutZ41	2	4	0	2	0	3
OutZ42	4	4	0	4	0	4
OutZ43	4	0	0	4	0	3
OutZ44	4	0	0	4	0	3
Set1A	5	0	0	1	0	2
Set2A	5	0	0	1	0	2
Box1A	5	0	0	10	0	2
BiLin1A	5	10	0	3	0	2

defining $\tilde{K}(x)$, the column the p_P is the number of equalities in the definition of $\tilde{K}(x)$ and the last column n(start) is the number of starting points for the problem.

Table 2 presents the results of Algorithm 1 for various combinations of γ and θ, including the number of iterations required for the algorithm to satisfy the stopping criterion. Instances of failure, where the algorithm does not converge within 1000 iterations, are also documented. The optimal performance of Algorithm 1 is observed when $\theta = 0.5$ and $\gamma = 0.05$. Generally, the algorithm exhibits improved performance when the value of θ is less than the value of γ. The table also reports the average number of iterations and the average time taken by the algorithm for all problems, further supporting the observation that the algorithm performs better when θ is less than γ.

In Table 3, we compare Algorithm 1 and Algorithm 2 performances with [3][Algorithm 1] (namely, Grad-type Alg.), [16][Algorithm 2.1] (Inertial Alg.) and [41][Algorithm 3.1] (Inertial Proj. Alg). Recall that the Gradient-type Alg. requires an additional projection onto a closed and convex set per each iteration and the Inertial Alg. computes two inertial steps at each iteration. Specifically, for Algorithm 1 we choose $\gamma = 0.05$ and $\theta = 0.5$, for Algorithm 2, we take $\gamma = 0.5$ and $\theta_k = \frac{0.9k}{k+1}$, for Grad-type Alg., we take $\alpha_k = \frac{k+1}{2(k+5)}$, for Inertial Alg., we chose $\alpha_k = \frac{k+7}{2(k+5)}, \beta_k = \frac{k-1}{4(k+6)}, \pi_k = \frac{k}{6(k+2)}$ and $\theta_k = \frac{k}{5(k+1)}$. We recorded the number of iterations and time taken by each method in Table 3. The results demonstrate that the proposed algorithm outperforms the comparative algorithms from the literature. Specifically, out of the 24 runs conducted in the experiment, the proposed algorithm achieved 15 successful runs. In comparison, the Grad-type algorithm achieved 11 successful runs, the Inertial algorithm achieved 13 successful runs, and the Inert-Proj algorithm achieved 14 successful runs. Furthermore, the average number of iterations for Algorithm 1 and 2 are 409.08 and 413.0, respectively, while the Grad-type algorithm requires 548.25 iterations on average, the Inertial algorithm takes 466.08 iterations, and the Inert-Proj algorithm takes 417.75 iterations. This highlights that the proposed algorithms converge in fewer iterations to meet the stopping criterion compared to the comparative algorithms. Similarly, the average computation time for Algorithm 1 is 5.34 seconds, Algorithm 2 takes 5.51 seconds, the Grad-type algorithm requires 7.50 seconds, the Inertial algorithm takes 6.18 seconds, and the Inert-Proj algorithm takes 5.97 seconds. In conclusion, it was observed that none of the algorithms converged even at the final iteration for problem `OutZ42`, regardless of the four different starting points used in the experiments.

Table 2: Numerical results of Algorithm 1 with different values of γ and θ: number of iterations needed for satisfying the stopping criterion.

γ	0.5	0.5	0.5	0.5	0.05	0.05	0.05	0.05
θ	0.01	0.05	0.1	0.5	0.01	0.05	0.1	0.5
OutZ40-1	1	1	1	1	1	1	1	1
OutZ40-2	18	12	10	6	25	36	26	28
OutZ40-3	19	13	11	7	14	19	14	15
OutZ41-1	1	1	1	1	1	1	1	1
OutZ41-2	34	18	12	6	338	1000	883	960
OutZ41-3	37	19	13	7	11	15	11	12
OutZ42-1	1000	1000	1000	1000	1000	1000	1000	1000
OutZ42-2	1000	1000	1000	1000	1000	1000	1000	1000
OutZ42-3	1000	1000	1000	1000	1000	1000	1000	1000
OutZ42-4	1000	1000	1000	1000	1000	1000	1000	1000
OutZ43-1	1000	1	1	1	1	3	2	1
OutZ43-2	1	1	1000	1000	2	5	1	2
OutZ43-3	1000	1000	1000	1000	1000	1000	1000	1000
OutZ44-1	1000	3	1	1000	1	1	2	5
OutZ44-2	1000	1	1000	1000	1	3	1	7
OutZ44-3	1000	1000	1000	1000	1000	1000	1000	1000
Set1A-1	1	1	1	1	1	1	1	1
Set1A-2	1	1	1	1	1	1	1	1
Set2A-1	1	1	1	1	1	1	1	1
Set2A-2	1	1	1	1	1	1	1	1
Box1A-1	3	3	3	4	4	1000	4	4
Box1A-2	2	2	2	2	6	7	1000	1000
BiLin1A-1	3	3	3	4	1000	1000	1000	1000
BiLin1A-2	2	2	2	2	1000	1000	1000	1000
# success	15	18	16	15	16	14	15	15
Av. iter	380.21	253.5	336	376.88	350.38	420.63	414.58	418.33
Av. time	4.60	3.05	4.07	4.57	4.64	5.59	5.47	5.52

5 Conclusion

We first introduced a C-FISTA type gradient projection algorithm to solve quasi-variational inequalities in Hilbert spaces and consequently obtain its linear convergence rate under strong monotonicity of the operator. This proposed algorithm is an adaptation of the Nesterov C-FISTA algorithm studied for strongly convex optimization problem to quasi-variational inequalities. Furthermore, another version of gradient projection algorithm with momentum terms is also designed and linear rate of convergence obtained. The numerical performance of the proposed algorithms

Table 3: Comparison of the performance of proposed Algorithm 1 and 2 with other methods.

Algorithms	Alg. 1		Alg. 2		Grad-type Alg.		Inertial Alg.		Iner-Proj. Alg.	
	Iter	Time	Iter	Time	Iter	Time	Iter	Time	Iter	Time
OutZ40-1	1	1.92	1	1.95	1	1.32	1	2.42	1	0.62
OutZ40-2	36	0.67	45	0.95	1000	13.56	4	0.25	3	0.16
OutZ40-3	19	0.26	23	0.39	1000	12.31	3	0.04	2	0.02
OutZ41-1	1	0.01	1	0.01	1	0.01	1	0.02	1	0.01
OutZ41-2	728	9.41	813	10.04	1000	12.22	163	2.96	1000	14.6
OutZ41-3	15	0.19	19	0.24	1000	12.01	3	0.03	2	0.02
OutZ42-1	1000	11.74	1000	12.65	1000	12.50	1000	14.08	1000	13.26
OutZ42-2	1000	11.55	1000	12.11	1000	12.41	1000	13.10	1000	12.74
OutZ42-3	1000	12.05	1000	11.89	1000	12.59	1000	13.10	1000	11.36
OutZ42-4	1000	11.24	1000	12.08	1000	14.04	1000	13.06	1000	11.46
OutZ43-1	1	0.02	2	0.05	2	0.04	2	0.03	1	0.06
OutZ43-2	1	0.02	2	0.06	1	2.91	3	0.04	1000	15.71
OutZ43-3	1000	15.15	1000	20.11	140	2.91	1000	16.23	1000	15.19
OutZ44-1	3	0.06	3	0.08	3	0.06	1000	13.89	1000	15.85
OutZ44-2	2	0.05	9	0.16	6	0.19	1000	14.12	1000	16.32
OutZ44-3	1000	15.32	1000	19.08	1000	20.95	1000	14.06	1000	15.62
Set1A-1	1	0.03	1	0.06	1	0.05	1	0.06	1	0.03
Set1A-2	1	0.02	1	0.04	1	0.03	1	0.03	1	0.03
Set2A-1	1	0.04	1	0.02	1	0.02	1	0.03	1	0.03
Set2A-2	1	0.02	1	0.02	1	0.07	1	0.02	1	0.02
Box1A-1	1000	12.62	1000	13.15	1000	13.68	1000	10.58	4	0.03
Box1A-2	7	0.09	11	0.08	1000	13.08	2	0.03	2	0.02
BiLin1A-1	1000	12.96	1000	13.88	1000	13.38	1000	10.07	4	0.03
BiLin1A-2	1000	12.78	1000	13.49	1000	13.03	1000	10.02	2	0.02
# success	15		15		11		13		14	
Av. iter	409.08		413		548.25		466.08		417.75	
Av. time	5.34		5.51		7.50		6.18		5.97	

showed that the new algorithms are efficient and outperform some popular related gradient projection algorithms in the literature for quasi-variational inequalities.

Disclosure statement

Ethical Approval and Consent to participate

All the authors gave ethical approval and consent to participate in this article.

Consent for publication

All the authors gave consent for the publication of identifiable details to be published in the journal and article.

Code availability

The Matlab codes employed to run the numerical experiments are available on request.

Availability of supporting data

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Competing interests

The authors declare no competing interests.

Funding

Not Applicable.

Authors' contributions

Y.Y. and Y.S. wrote the manuscript and L.O.J. prepared the all the figures and tables.

Acknowledgments

The authors are grateful to the anonymous referee for the excellent comments and suggestions.

References

[1] Alvarez, F.: Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space. SIAM J. Optim. **14**, 773–782 (2003)

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

[3] Antipin, A. S., Jaćimović, M., Mijajlovi, N.: Extragradient method for solving quasivariational inequalities. Optimization **67**, 103–112 (2018)

[4] Antipin, A. S., Jaćimović, M., Mijajlovi, N.: A second-order iterative method for solving quasi-variational inequalities. Comp. Math. Math. Phys. **53**, 258–264 (2013)

[5] Attouch, H., Goudon, X., Redont, P.: The heavy ball with friction. I. The continuous dynamical system. Commun. Contemp. Math. **2**, 1–34 (2000)

[6] Attouch, H., Czarnecki, M. O.: Asymptotic control and stabilization of non-linear oscillators with non-isolated equilibria. J. Diff. Equations **179**, 278–310 (2002)

[7] Attouch, H., Peypouquet, J., Redont, P.: A dynamical approach to an inertial forward-backward algorithm for convex minimization. SIAM J. Optim. **24**, 232–256 (2014)

[8] Aussel, D., Sagratella, S.: Sufficient conditions to compute any solution of a quasivariational inequality via a variational inequality. Math. Methods Oper. Res. **85**, 3–18 (2017)

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

[10] Boţ, R. I., Csetnek, E. R., Hendrich, C.: Inertial Douglas-Rachford splitting for monotone inclusion. Appl. Math. Comput. **256**, 472–487 (2015)

[11] Boţ, R. I., Csetnek, E. R.: An inertial alternating direction method of multipliers. Minimax Theory Appl. **1**, 29–49 (2016)

[12] Boţ, R. I., Csetnek, E. R.: An inertial forward-backward-forward primal-dual splitting algorithm for solving monotone inclusion problems. Numer. Algorithms **71**, 519–540 (2016)

[13] Calatroni, L.; Chambolle, A. Backtracking strategies for accelerated descent methods with smooth composite objectives. SIAM J. Optim. **29**, 1772–1798 (2019).

[14] Chambolle, A.; Pock, T. An introduction to continuous optimization for imaging. Acta Numer. **25**, 161–319 (2016).

[15] Chen, C., Chan, R. H., Ma, S., Yang, J.: Inertial proximal ADMM for linearly constrained separable convex optimization. SIAM J. Imaging Sci. **8**, 2239–2267 (2015)

[16] Ç̣opur, A. K.; Hacıoğlu, E.; Gürsoy, F.; Ertürk, M. an efficient inertial type iterative algorithm to approximate the solutions of quasi variational inequalities in real Hilbert spaces. J. Sci. Comput. 89, 50 (2021).

[17] Facchinei, F., Kanzow, C., Karl, S., Sagratella, S.: The semismooth Newton method for the solution of quasi-variational inequalities. Comput. Optim. Appl. **62**, 85–109 (2015)

[18] Facchinei, F., Kanzow, C., Sagratella, S.: Solving quasi-variational inequalities via their KKT conditions. Math. Program. **144**, 369–412 (2014)

[19] Facchinei, F., Kanzow, C., Sagratella, S.: QVILIB: a library of quasi-variational inequality test problems. Pacific J. Optim. **9**, 225-250 (2013).

[20] Fichera, G.: Sul problema elastostatico di Signorini con ambigue condizioni al contorno (English translation: "On Signorini's elastostatic problem with ambiguous boundary conditions"). Atti Accad. Naz. Lincei, VIII. Ser., Rend., Cl. Sci. Fis. Mat. Nat. **34**, 138–142 (1963)

[21] Fichera, G.: Problemi elastostatici con vincoli unilaterali: il problema di Signorini con ambigue condizioni al contorno (English translation: "Elastostatic problems with unilateral constraints: the Signorini's problem with ambiguous boundary conditions"). Atti Accad. Naz. Lincei, Mem., Cl. Sci. Fis. Mat. Nat., Sez. I, VIII. Ser. **7**, 91–140 (1964)

[22] Florea, M.I.; Vorobyov, S.A. An accelerated composite gradient method for large-scale composite objective problems. IEEE Trans. Signal Process. **67**, 444–459 (2019).

[23] Florea, M.I.; Vorobyov, S.A. A generalized accelerated composite gradient method: uniting Nesterov's fast gradient method and Fista. IEEE Trans. Signal Process. **68**, 3033–3048 (2020).

[24] Garner, C.; Zhang, S. Linearly-Convergent FISTA Variant for Composite Optimization with Duality. J. Sci. Comput 94, 65 (2022).

[25] Goebel, K., Reich, S.: Uniform convexity, hyperbolic geometry and nonexpansive mappings. Marcel Dekker Inc, U.S.A., 1984

[26] Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications. Academic Press, New York-London, 1980

[27] Latorre, V., Sagratella, S.: A canonical duality approach for the solution of affine quasi-variational inequalities. J. Global Optim. **64**, 433–449 (2016)

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

[29] Maingé, P. E.: Regularized and inertial algorithms for common fixed points of nonlinear operators. J. Math. Anal. Appl. **344**, 876–887 (2008)

[30] Mijajlović, N., Jaćimović, M., Noor, M. A.: Gradient-type projection methods for quasi-variational inequalities. Optim. Lett. **13**, 1885–1896 (2019)

[31] Mosco, U.: Implicit variational problems and quasi variational inequalities. Lecture Notes in Mathathematics, Springer, Berlin **543** (1976)

[32] Nesterov, Y., Scrimali, L.: Solving strongly monotone variational and quasi-variational inequalities. Discrete Contin. Dyn. Syst. **31**, 1383–1396 (2011)

[33] Noor, M. A.: An iterative scheme for a class of quasi variational inequalities. J. Math. Anal. Appl. **110**, 463–468 (1985)

[34] Noor, M. A.: Quasi Variational Inequalities. Appl. Math. Lett. **1**, 367-370 (1988)

[35] Noor, M. A., Noor, K. I., Khan, A. G.: Some iterative schemes for solving extended general quasi variational inequalities. Appl. Math. Inform. Sci. **7**, 917–925 (2013)

[36] Noor, M. A., Oettli, W.: On general nonlinear complementarity problems and quasi equilibria. Matematiche (Catania) **49**, 313–331 (1994)

[37] Ochs, P., Brox, T., Pock, T.: iPiasco: Inertial Proximal Algorithm for strongly convex Optimization. J. Math. Imaging Vision. **53**, 171–181 (2015).

[38] Polyak, B. T.: Some methods of speeding up the convergence of iterative methods. Zh. Vychisl. Mat. Mat. Fiz. **4**, 791–803 (1964)

[39] Ryazantseva, I. P.: First-order methods for certain quasi-variational inequalities in a Hilbert space. Comput. Math. Math. Phys. **47**, 183–190 (2007)

[40] Shehu, Y.: Convergence rate analysis of inertial Krasnoselskii-Mann-type iteration with applications. Numerical Funct. Anal. Optim. **39**, 1077–1091 (2018)

[41] Shehu, Y., Gibali, A., Sagratella, S.: Inertial projection-type methods for solving quasi-variational inequalities in real Hilbert spaces. J. Optim. Theory Appl. **184**, 877–894 (2020).

[42] Stampacchia, G.: Formes bilineaires coercitives sur les ensembles convexes. Académie des Sciences de Paris **258**, 4413–4416 (1964)

[43] Zhang, Y.; Zhang, N.; Sun, D.; Toh, K.C. An efficient hessian based algorithm for solving large-scale sparse group lasso problems. Math. Program. **179**, 223–263 (2020).