

A Relaxed Proximal Point Algorithm with Double-Inertial Effects for Nonconvex Equilibrium Problems

Van Nam Tran^{1,*}

1. Faculty of Applied Sciences, Ho Chi Minh City University of Technology and Engineering, Ho Chi Minh City, Vietnam

Abstract: We present a relaxed proximal point method with double inertial effects to approximate a solution to a non-convex equilibrium problem. We establish the global convergence of the iterative sequence generated by our algorithm. Some known results are recovered as special cases of our findings. An application to mixed variational inequality problems is given. A numerical test is provided to support the theoretical results, and we show that our proposed algorithm outperforms some existing methods in certain cases.

Keywords: Non-convexity, equilibrium problem, proximal point algorithms, two-step inertial extrapolation.

2010 Mathematics Subject Classification: 47H09, 65K15, 90C30.

1. Introduction

1.1. Statement and setting of the problem

In this work, we are interested in the *equilibrium problem* (EP for short):

(EP) Find $\hat{x} \in C : f(\hat{x}, y) \geq 0, \forall y \in C,$

where C is a nonempty set in $\mathbb{R}^n$ and $f : C \times C \rightarrow \mathbb{R}$ be a bifunction satisfying $f(x, x) = 0$ for all $x \in C$. We denote by $S(C, f)$ the set of solutions of (EP).

Equilibrium problems were first studied by Ky Fan in [16] and have been further developed over the last few decades (see, for instance, [2, 5, 6]). It has been shown that many problems, such as the optimization problem, the variational inequality, the saddle point problem, the Nash equilibrium problem in non-cooperative games, and the fixed point problem, can be formulated as equilibrium problems; see, for instance, [3, 4, 5, 7, 29] and references therein. In addition, numerous applications have been effectively described using the notion of equilibrium solutions, prompting extensive research efforts in this area. We refer the readers to the monograph [6] for an excellent survey on the existence of equilibrium points and solution methods for finding them.

When developing algorithms for approximating solutions to equilibrium problems, convexity assumptions play a fundamental and often indispensable role. In this paper, we propose a relaxed proximal point method with dual inertial effects to address nonconvex equilibrium problems. Our approach is based on the notion of strong quasiconvexity, a generalized convexity concept that encompasses a broader class of functions than strong convexity.

Strongly quasiconvex functions were first introduced by Polyak in his seminal work [32]. Since then, this notion has been extensively studied in various settings; see, for instance, [18, 22, 24, 38]. More recently, (strongly) quasiconvexity has gained increasing attention in optimization and equilibrium-related problems. In this direction, [13] investigates the asymptotic convergence of saddle-point dynamics associated with bifunctions that are strongly quasiconvex in one or both variables, while [36] examines the asymptotic behavior of quasi-autonomous gradient expansive systems driven by strongly quasiconvex functions. The latter also provides a systematic study of minimizing strongly quasiconvex functions using modern

*Corresponding author.

E-mail addresses: V. N. Tran : namtv@hcmute.edu.vn

<https://doi.org/10.68381/jca34040>

Received date: March 4, 2025; Accepted date: January 26, 2026

© 2027 Published by Banach Press LLC. (Dedicated to Alexander Plakhov on the occasion of his 65th birthday.)

iterative techniques, establishing the convergence of the classical proximal point algorithm under standard assumptions. A related convergence result for a relaxed inertial proximal point method is reported in [19].

Within the equilibrium problem framework, several iterative schemes based on (strong) quasiconvexity have been proposed in recent years. Notably, [23] introduces one of the first proximal point methods under such assumptions, [39] develops a subgradient-type method, [40] proposes a proximal subgradient algorithm, and the recent work [2] extends the approach of [23] by incorporating Bregman distances.

Inertial extrapolation was originally proposed by Polyak [31] as an acceleration mechanism inspired by the heavy-ball method for second-order dynamical systems with friction. Initially designed for smooth convex minimization problems, this approach incorporates momentum terms from previous iterates with the aim of improving convergence behavior.

Subsequent research has focused on the theoretical foundations and performance improvements of inertial methods, particularly with respect to convergence properties and computational efficiency; see, for instance, [8, 12, 14, 19, 30, 35]. Significant advances were reported in [26, 34], where it was shown that two-step inertial extrapolation schemes, which utilize information from more than two previous iterates, can achieve superior acceleration compared to classical one-step inertial methods. These findings provide a rigorous justification for the advantages of multi-step inertial strategies.

More recently, multi-step inertial methods have attracted increasing attention and have been successfully applied to a wide range of problems, including optimization, equilibrium problems, and splitting algorithms; see, for example, [15, 27, 37]. This renewed interest is partly motivated by the limitations observed in one-step inertial acceleration schemes within certain frameworks, such as the alternating direction method of multipliers [34, Section 3], and by the potential of multi-step inertial techniques to further enhance convergence performance, as emphasized by Polyak [33].

1.2. Contributions

- In this paper, we apply the two inertial step proximal point method with relaxation to solve non-convex equilibrium problems. Our results extend the usage of the proximal point method from convex equilibrium problems already studied in [2, 5] to non-convex ones.
- Instead of one-step inertial extrapolation already considered in several papers in [1, 8, 9, 10, 14, 19, 20, 30] for convex equilibrium problems, we further accelerate our proximal point method by adding a two-step inertial extrapolation.
- By treating a mixed variational inequality problem (MVIP) as a special case of an equilibrium problem we obtain a relaxed two-inertial step proximal point for the MVIP.
- We show that, in certain situations, the two-step inertial extrapolation achieves a faster convergence rate than the one-step version, and this behavior is confirmed by a numerical experiment.

1.3. Organization

We organize the rest of our paper as follows: Section 2 contains basic definitions and results needed in subsequent sections. In Section 3, we present and discuss our proposed method along with its convergence results. An application to mixed variational inequality problems is given in Section 4. We give a numerical illustration in Section 5, and concluding remarks are given in Section 6.

2. Preliminaries

In this section we recall some basic notions which are used frequently in the sequel.

We denote by $\langle \cdot, \cdot \rangle$ the *inner product* in the Hilbert space $\mathbb{R}^n$, and by $\|\cdot\|$ its induced norm. Recall that the effective domain of a function $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ is defined by $\text{dom } g := \{x \in \mathbb{R}^n : g(x) < +\infty\}$. The function g is said to be *proper* if $g(x) > -\infty$ for every $x \in \mathbb{R}^n$ and $\text{dom } g$ is nonempty (clearly, $g(x) = +\infty$ for every $x \notin \text{dom } g$).

When the domain of $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ is convex, we say that the function g is

- (a) *convex* if, for any $x, y \in \text{dom } g$, then

$$(tx + (1 - t)y) \leq tg(x) + (1 - t)g(y), \quad \forall t \in [0, 1],$$

(b) *quasiconvex* if, for any $x, y \in \text{dom } g$, then

$$g(tx + (1-t)y) \leq \max\{g(x), g(y)\}, \quad \forall t \in [0, 1],$$

(c) *strongly convex* (on $\text{dom } g$) with modulus $\gamma > 0$ if there is a $\gamma \in (0, \infty)$ satisfying

$$g(tx + (1-t)y) \leq tg(x) + (1-t)g(y) - t(1-t)\frac{\gamma}{2}\|x - y\|^2, \\ \forall x, y \in \text{dom } g, \quad \forall t \in [0, 1],$$

(d) *strongly quasiconvex* (on $\text{dom } g$) with modulus $\gamma > 0$ if there exists a $\gamma \in (0, \infty)$ for which

$$g(ty + (1-t)x) \leq \max\{g(y), g(x)\} - t(1-t)\frac{\gamma}{2}\|x - y\|^2, \\ \forall x, y \in \text{dom } g, \quad \forall t \in [0, 1].$$

It is well known that a strongly convex function is also a strongly quasiconvex function; however, the inverse does not hold in general. For instance, the Euclidean norm is strongly quasiconvex on any bounded convex set but not strongly convex [24, Theorem 2]. Similarly, a convex function is also quasiconvex, while the inverse may not hold. For example, the function $g : \mathbb{R} \rightarrow \mathbb{R}$ with $g(x) = x^3$, is quasiconvex but not convex. For more relationships between different notions of convexity, we refer readers to [20]. For further studies on generalized convexity, see [11, 21], among other works.

The proximity operator becomes a very useful tool in studying numerical methods for convex optimization problems. We now recall this notion below.

The proximity operator of parameter $\lambda > 0$ of a function $g : C \rightarrow \overline{\mathbb{R}}$ at $x \in \mathbb{R}^n$ is defined by $\text{Prox}_{\lambda g}(C, \cdot) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with

$$\text{Prox}_{\lambda g}(C, x) = \arg \min_{y \in C} \left\{ g(y) + \frac{1}{2\lambda} \|y - x\|^2 \right\}, \quad \forall x \in \mathbb{R}^n,$$

where C is a nonempty subset of $\mathbb{R}^n$. When $C = \mathbb{R}^n$ we simply write $\text{Prox}_{\lambda g}(\cdot)$.

It is worth noting that the sum of a strongly quasiconvex function and half of the squared norm may not be strongly quasiconvex, as shown in [25, Remark 6] and [20]. As a result, the proximal operator of a strongly quasiconvex function is not a singleton, but a set-valued mapping. It is a singleton when the sum of a strongly quasiconvex function and half of a squared norm is strongly quasiconvex.

We now recall some notions related to bifunctions. Let C be a closed and convex set in $\mathbb{R}^n$, and $f : C \times C \rightarrow \mathbb{R}$ be a real-valued bifunction satisfying $f(x, x) = 0, \forall x \in C$. We say that f is

(a) *monotone* on C , if for every $x, y \in C$ we have

$$f(x, y) + f(y, x) \leq 0,$$

(b) *pseudomonotone* on C if

$$f(x, y) \geq 0 \implies f(y, x) \leq 0, \quad \forall x, y \in C.$$

We list here some useful lemmas that are used for analyzing the convergence of our proposed algorithm later. The first lemma presents a property of the proximal operator of a strongly quasiconvex function.

Lemma 2.1. ([24, Proposition 6]) *Let C be a closed and convex set in $\mathbb{R}^n$, $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a proper, lower semicontinuous, strongly quasiconvex function with modulus $\gamma \geq 0$ and such that $C \subseteq \text{dom } g$, $\lambda > 0$ and $x \in C$. If $\hat{x} \in \text{Prox}_{\lambda g}(C, x)$, then for all $y \in C$ and all $t \in [0; 1]$ one has*

$$g(\hat{x}) - \max\{g(y), g(\hat{x})\} \leq \frac{t}{\lambda} \langle \hat{x} - x, y - \hat{x} \rangle + \frac{t}{2} \left(\frac{t}{\lambda} - \gamma + t\gamma \right) \|y - \hat{x}\|^2.$$

The following lemma presents several useful identities that will be frequently used in the subsequent sections.

Lemma 2.2. [20, 37] Let $x, y, z \in \mathbb{R}^n$ and $a, b, \beta \in \mathbb{R}$. Then we have

- (a) $\|(1+a)x - (a-b)y - bz\|^2$
 $= (1+a)\|x\|^2 - (a-b)\|y\|^2 - b\|z\|^2 + (1+a)(a-b)\|x-y\|^2$
 $+ b(1+a)\|x-z\|^2 - b(a-b)\|y-z\|^2.$
- (b) $\langle x-z, y-x \rangle = \frac{1}{2}\|z-y\|^2 - \frac{1}{2}\|x-z\|^2 - \frac{1}{2}\|y-x\|^2.$
- (c) $\|\beta x + (1-\beta)y\|^2 = \beta\|x\|^2 + (1-\beta)\|y\|^2 - \beta(1-\beta)\|x-y\|^2.$

3. The algorithm and the convergence analysis

In this section, we introduce and discuss our proposed method to solve non-convex equivalent problems. Furthermore, a convergence analysis of the proposed method is given.

We first give some assumptions that are needed for investigating the convergence of the proposed algorithm for (EP).

- (A1) $f(\cdot, y)$ is upper semicontinuous for all $y \in C$;
- (A2) f is pseudomonotone on C ;
- (A3) f is lower semicontinuous (jointly in both arguments) on $C \times C$;
- (A4) $f(x, \cdot)$ is strongly quasiconvex on C with modulus $\gamma > 0$ for each $x \in C$;
- (A5) f satisfies the following Lipschitz-continuity type condition: there exists $\eta > 0$ such that

$$f(x, z) - f(x, y) - f(y, z) \leq \eta (\|x - y\|^2 + \|y - z\|^2), \text{ for all } x, y, z \in C.$$

As mentioned in [20], the assumptions (A1), (A2), (A3), (A5) are standard in the related literature, but (A4) is weaker than the usual one, which requires that f is convex regarding the second argument. Also, authors in [20] provided strongly quasiconvex functions which satisfy the assumptions (Ai), $i = 1, 2, 3, 4, 5$.

It is shown in [20] that the following usual assumption in related literature:

(A3'): $f(x, \cdot)$ is lower semicontinuous for any $x \in C$ is fulfilled when (A3) is satisfied.

From now on we always assume that the solution set $S(C, f)$ of (EP) is nonempty.

3.1. Proposed method

We now present our method.

Algorithm 1: Relaxed Double Inertial Proximal Point Algorithm for Non-Convex Equilibrium Problems (RDIPPA)

1. Choose $\beta \in (-\infty, 0]$, $\rho \in [0, 1)$, $\theta \in [0, \frac{1}{2})$ and $\lambda_k > 0$ for all $k = 1, 2, 3, \dots$. Pick $x_{-1}, x_0, x_1 \in C$ and set $k = 1$.
2. Given x_{k-2}, x_{k-1} and x_k , compute

$$\begin{aligned} y_k &= x_k + \theta(x_k - x_{k-1}) + \beta(x_{k-1} - x_{k-2}), \\ z_k &\in \text{Prox}_{\lambda_k f(y_k, \cdot)}(C, y_k) = \arg \min_{x \in C} (f(y_k, x) + \frac{1}{2\lambda_k} \|y_k - x\|^2). \end{aligned} \quad (1)$$

3. If $y_k = z_k$ then STOP. Otherwise, go to Step 4.
4. Choose some relaxation parameter $\rho_k \in [1 - \rho, 1 + \rho]$ and update

$$x_{k+1} = (1 - \rho_k)y_k + \rho_k z_k.$$

5. Let $k \leftarrow k + 1$ and go to Step 2.

Remark 3.1. When $\beta = 0$, our proposed algorithm reduces to Algorithm 1 of [20].

3.2. Convergence analysis

In this subsection, we present the convergence analysis of the sequence of iterates generated by our proposed Algorithm 1.

The first result gives a characterization of the solution set of (EP).

Proposition 3.2. ([20, Proposition 3.1]) *Let C be a closed and convex set in $\mathbb{R}^n$ and suppose that the bifunction f satisfies (A1), (A2), (A'3), and (A4). Then $S(f, C)$ is a singleton.*

We now show that the stopping criterion of Algorithm 1 is valid.

Proposition 3.3. *Let C be an affine subspace in $\mathbb{R}^n$, and f be a bifunction defined on $C \times C$ such that Assumptions (A1), (A3') and (A4) are valid. Let $\{\lambda_k\}_k$ and $\{\rho_k\}_k$ be sequences of positive numbers and $\{x_k\}_k, \{y_k\}_k$ and $\{z_k\}_k$ be the sequences generated by Algorithm 1. If $y_{k_0} = z_{k_0}$ for some $k_0 \in \mathbb{N}$, then y_{k_0} is a solution of (EP).*

Proof. We first observe that, by Proposition 3.2, (EP) admits a unique solution, denoted by $\hat{x}$.

If $y_{k_0} = z_{k_0}$, from (1) we derive that

$$f(y_{k_0}, z_{k_0}) + \frac{1}{2\lambda_{k_0}} \|y_{k_0} - z_{k_0}\|^2 \leq f(y_{k_0}, x) + \frac{1}{2\lambda_{k_0}} \|y_{k_0} - x\|^2, \quad \forall x \in C. \quad (2)$$

Because of the facts that $f(x, x) = 0$ for all $x \in C$ and $y_{k_0} = z_{k_0}$, it follows from (2) that

$$0 \leq f(y_{k_0}, x) + \frac{1}{2\lambda_{k_0}} \|y_{k_0} - x\|^2, \quad \forall x \in C.$$

Take $x = ty + (1-t)y_{k_0}$ with $y \in C \setminus \{\hat{x}\}$ and $t \in [0, 1]$. Since f is strongly quasiconvex in its second argument with modulus $\gamma > 0$, it holds that

$$\begin{aligned} 0 &\leq f(y_{k_0}, ty + (1-t)y_{k_0}) + \frac{1}{2\lambda_{k_0}} \|y_{k_0} - y\|^2 \\ &\leq \max\{f(y_{k_0}, y), 0\} - \frac{t(1-t)\gamma}{2} \|y_{k_0} - y\|^2 + \frac{t^2}{2\lambda_{k_0}} \|y_{k_0} - y\|^2 \\ &\leq \max\{f(y_{k_0}, y), 0\} + \frac{t}{2} \left(\frac{t}{\lambda_{k_0}} - \gamma + t\gamma \right) \|y_{k_0} - y\|^2, \quad \forall y \in C \setminus \{\hat{x}\}, \forall t \in [0, 1]. \end{aligned} \quad (3)$$

Since λ_{k_0} is positive, we take t such that $0 < t < \frac{\lambda_{k_0}\gamma}{1 + \lambda_{k_0}\gamma}$. Then $\frac{t}{\lambda_{k_0}} - \gamma + t\gamma < 0$.

Thus, it follows from (3) that

$$0 < \max\{f(y_{k_0}, y), 0\} = f(y_{k_0}, y), \quad \forall y \in C \setminus \{\hat{x}, y_{k_0}\},$$

meaning that, $y_{k_0} = \hat{x} \in S(C, f)$. □

The next result is the first step for the analysis of convergence.

Proposition 3.4. *Let C be an affine subspace in $\mathbb{R}^n$, and f be a bifunction defined on $C \times C$ such that Assumptions (Ai) with $i = 1, 2, 3', 4, 5$ hold. Let $\{\lambda_k\}_k$ and $\{\rho_k\}_k$ be sequences of positive numbers, $\{x_k\}_k, \{y_k\}_k$ and $\{z_k\}_k$ be the sequences generated by Algorithm 1. Let $\hat{x} \in S(C, f)$. Then for every $k \geq 0$, at least one of the following inequalities holds:*

$$\|x_{k+1} - \hat{x}\|^2 \leq \|y_k - \hat{x}\|^2 - \frac{2 - \rho_k}{\rho_k} \|x_{k+1} - y_k\|^2 - \frac{\rho_k(\gamma\lambda_k - 1)}{2} \|z_k - \hat{x}\|^2, \quad (4)$$

$$\|x_{k+1} - \hat{x}\|^2 \leq \|y_k - \hat{x}\|^2 - \frac{2 - 4\eta\lambda_k - \rho_k}{\rho_k} \|x_{k+1} - y_k\|^2 - \frac{\rho_k(\gamma\lambda_k - 1 - 8\eta\lambda_k)}{2} \|z_k - \hat{x}\|^2. \quad (5)$$

Proof. From the relaxation step (Step 4) of Algorithm 1 one has

$$x_{k+1} - \hat{x} = (1 - \rho_k)(y_k - \hat{x}) + \rho_k(z_k - \hat{x}).$$

Using Part (c) of Lemma 2.2, we have

$$\|x_{k+1} - \hat{x}\|^2 = (1 - \rho_k) \|y_k - \hat{x}\|^2 + \rho_k \|z_k - \hat{x}\|^2 - \rho_k(1 - \rho_k) \|y_k - z_k\|^2,$$

which is equivalent to the following

$$\|x_{k+1} - \hat{x}\|^2 - \|y_k - \hat{x}\|^2 = \rho_k \left(\|z_k - \hat{x}\|^2 - \|y_k - \hat{x}\|^2 \right) - \rho_k (1 - \rho_k) \|y_k - z_k\|^2.$$

In addition, from Part (b) of Lemma 2.2 one has

$$\|z_k - \hat{x}\|^2 - \|y_k - \hat{x}\|^2 = \|z_k - y_k\|^2 + 2 \langle z_k - y_k, y_k - \hat{x} \rangle.$$

Combining the last two identities we obtain

$$\|x_{k+1} - \hat{x}\|^2 - \|y_k - \hat{x}\|^2 = 2\rho_k \langle z_k - y_k, y_k - \hat{x} \rangle + \rho_k^2 \|y_k - z_k\|^2. \quad (6)$$

Moreover, for every $k \geq 1$, by the definition of z_k and Lemma 2.1 one has

$$\begin{aligned} & f(y_k, z_k) - \max \{f(y_k, z_k), f(y_k, \hat{x})\} \\ & \leq \frac{t}{\lambda_k} \langle z_k - y_k, \hat{x} - z_k \rangle + \frac{t}{2} \left(\frac{t}{\lambda_k} - \gamma + t\gamma \right) \|z_k - \hat{x}\|^2, \forall t \in [0, 1]. \end{aligned} \quad (7)$$

We consider two cases.

(i) If $f(y_k, z_k) \geq f(y_k, \hat{x})$, then the inequality (7) implies that

$$\frac{1}{\lambda_k} \langle z_k - y_k, z_k - \hat{x} \rangle \leq \frac{1}{2} \left(\frac{t}{\lambda_k} - \gamma + t\gamma \right) \|z_k - \hat{x}\|^2, \forall t \in (0, 1]. \quad (8)$$

Choosing $t = \frac{1}{2}$ we obtain from (8) that

$$\langle z_k - y_k, z_k - \hat{x} \rangle \leq \frac{1 - \gamma\lambda_k}{4} \|z_k - \hat{x}\|^2.$$

Consequently,

$$\langle z_k - y_k, y_k - \hat{x} \rangle \leq -\|z_k - y_k\|^2 + \frac{1 - \gamma\lambda_k}{4} \|z_k - \hat{x}\|^2. \quad (9)$$

Combining (6) and (9) with noting that $z_k - y_k = (1/\rho_k)(x_{k+1} - y_k)$, we obtain the inequality (4).

(ii) If $f(y_k, z_k) < f(y_k, \hat{x})$, then it follows from (7) that

$$\begin{aligned} 0 & \leq \frac{t}{\lambda_k} \langle z_k - y_k, \hat{x} - z_k \rangle + \frac{t}{2} \left(\frac{t}{\lambda_k} - \gamma + t\gamma \right) \|z_k - \hat{x}\|^2 \\ & \quad + f(y_k, \hat{x}) - f(y_k, z_k) \\ & \leq \frac{t}{\lambda_k} \langle z_k - y_k, \hat{x} - z_k \rangle + \frac{t}{2} \left(\frac{t}{\lambda_k} - \gamma + t\gamma \right) \|z_k - \hat{x}\|^2 \\ & \quad + f(z_k, \hat{x}) + \eta \left(\|z_k - y_k\|^2 + \|z_k - \hat{x}\|^2 \right), \forall t \in [0, 1], \end{aligned}$$

where the last inequality follows from Assumption (A5). Moreover, because we have $\{\hat{x}\} = S(C, f)$ and f is pseudomonotone, one obtains $f(z_k, \hat{x}) \leq 0$. Hence,

$$\begin{aligned} 0 & \leq \frac{t}{\lambda_k} \langle z_k - y_k, \hat{x} - z_k \rangle + \frac{t}{2} \left(\frac{t}{\lambda_k} - \gamma + t\gamma \right) \|z_k - \hat{x}\|^2 \\ & \quad + \eta \left(\|z_k - y_k\|^2 + \|z_k - \hat{x}\|^2 \right), \forall t \in [0, 1]. \end{aligned}$$

This implies that

$$\begin{aligned} \frac{t}{\lambda_k} \langle z_k - y_k, z_k - \hat{x} \rangle & \leq \frac{t}{2} \left(\frac{t}{\lambda_k} - \gamma + t\gamma \right) \|z_k - \hat{x}\|^2 + \eta \left(\|z_k - y_k\|^2 + \|z_k - \hat{x}\|^2 \right) \\ & = \eta \|z_k - y_k\|^2 + \frac{t}{2} \left(\frac{t}{\lambda_k} - \gamma + t\gamma + \frac{2\eta}{t} \right) \|z_k - \hat{x}\|^2. \end{aligned}$$

Choosing $t = \frac{1}{2}$ one obtains

$$\langle z_k - y_k, y_k - \hat{x} \rangle \leq (2\eta\lambda_k - 1) \|z_k - y_k\|^2 + \left(2\eta\lambda_k + \frac{1 - \gamma\lambda_k}{4} \right) \|z_k - \hat{x}\|^2. \quad (10)$$

Using (10) and (6) with noting that $z_k - y_k = (1/\rho_k)(x_{k+1} - y_k)$, we conclude that the inequality (5) holds.

Thus, for every $k \geq 0$, either (4) or (5) holds. $\square$

Remark 3.5. (a) When $\gamma - 8\eta > 0$, the inequality $\frac{\rho_k(\gamma\lambda_k - 1 - 8\eta\lambda_k)}{2} > 0$ holds if and only if $\lambda_k > \frac{1}{\gamma - 8\eta}$. Also, if this condition holds, then, since $\rho_k > 0$ for all $k \geq 1$, it follows that

$$\rho_k(\gamma\lambda_k - 1) > \rho_k(1 + 8\eta\lambda_k - 1) = 8\eta\rho_k > 0.$$

Thus, if $\lambda_k > \frac{1}{\gamma - 8\eta}$ holds, then we have that

$$\frac{\rho_k(\gamma\lambda_k - 1)}{2} \|z_k - \hat{x}\|^2 \geq 0 \quad \text{and} \quad \frac{\rho_k(\gamma\lambda_k - 1 - 8\eta\lambda_k)}{2} \|z_k - \hat{x}\|^2 \geq 0.$$

(b) Because $\rho_k \in [1 - \rho, 1 + \rho]$ for all $k \geq 1$, the inequality

$$\frac{2 - 4\eta\lambda_k - \rho_k}{\rho_k} \geq \frac{2 - 4\eta\lambda_k - (1 + \rho)}{1 + \rho} = \frac{1 - 4\eta\lambda_k - \rho}{1 + \rho} > 0$$

holds if and only if there exists $\epsilon > 0$ such that $\lambda_k < \epsilon \leq \frac{1}{4\eta}$ and $0 \leq \rho \leq 1 - 4\eta\epsilon$. $\square$

By the above analysis, and in order to simplify the convergence analysis of our proposed algorithm, we impose the following assumptions on the parameter sequences $\{\lambda_k\}$ and $\{\rho_k\}$, as well as on the initial parameters $\theta \in [0, 1/2)$ and $\beta \in (-\infty, 0]$.

Assumption 3.6. Let $p_k^{(1)} = \frac{2 - \rho_k}{\rho_k}$ and $p_k^{(2)} = \frac{2 - 4\eta\lambda_k - \rho_k}{\rho_k} = p_k^{(1)} - \frac{4\eta\lambda_k}{\rho_k}$, where $k = 1, 2, \dots$. We assume that there exists $\epsilon > 0$ such that the following conditions hold:

$$(C1) \quad 0 < \frac{1}{\gamma - 8\eta} < \lambda_k < \epsilon \leq \frac{1}{4\eta}, \forall k \geq 1;$$

$$(C2) \quad 0 < 1 - \rho \leq \rho_k \leq 1 + \rho, \text{ where } 0 \leq \rho \leq 1 - 4\eta\epsilon, \forall k \geq 1.$$

To ensure the convergence of the sequence generated by Algorithm 1, we further assume that:

$$(C3) \quad \max \left\{ \frac{2\theta}{p_k^{(2)}} - (1 - \theta), \frac{\theta}{1 + p_k^{(2)}} - p_k^{(2)} \frac{(\theta - 1)^2}{(1 + \theta)(1 + p_k^{(2)})} \right\} < \beta \leq 0, \forall k \geq 1;$$

$$(C4) \quad \alpha_1 := \theta^2 \left(1 - p_k^{(2)} \right) + \theta \left(1 - 2\beta + 2\beta p_k^{(2)} - 2p_k^{(2)} \right) - \beta \left(1 + 2p_k^{(1)} \right) + \beta^2 \left(1 - p_k^{(2)} \right) - p_k^{(2)} < 0, \forall k \geq 1.$$

Remark 3.7. By Remark 3.5, if Condition (C2) holds, then $0 < p_k^{(2)} \leq p_k^{(1)}$. As a result, if Condition (C3) is fulfilled, one has

$$\frac{2\theta}{p_k^{(1)}} - (1 - \theta) \leq \frac{2\theta}{p_k^{(2)}} - (1 - \theta) < \beta \leq 0, \quad \text{and} \quad (11)$$

$$\frac{\theta}{1 + p_k^{(1)}} - p_k^{(1)} \frac{(\theta - 1)^2}{(1 + \theta)(1 + p_k^{(1)})} \leq \frac{\theta}{1 + p_k^{(2)}} - p_k^{(2)} \frac{(\theta - 1)^2}{(1 + \theta)(1 + p_k^{(2)})} < \beta \leq 0, \quad \forall \theta > 0. \quad (12)$$

Indeed, (11) follows from (C3) and the fact that $0 < p_k^{(2)} \leq p_k^{(1)}$. As for (12), observe that for each $\theta \geq 0$, the function

$$\phi(t) = \frac{\theta}{1 + t} - t \frac{(\theta - 1)^2}{(1 + \theta)(1 + t)}$$

is decreasing for all $t \geq 0$. In addition, since $0 < p_k^{(2)} \leq p_k^{(1)}$ for all $k \geq 1$, it follows that $\phi(p_k^{(1)}) \leq \phi(p_k^{(2)})$ for all $k \geq 1$, which implies that (12) holds. $\square$

Because of Part (a) of Remark 3.5 and Proposition 3.4, we obtain the following corollary.

Corollary 3.8. Let C be an affine subspace in $\mathbb{R}^n$, f be a bifunction defined on $C \times C$ such that Assumptions (Ai) with $i = 1, 2, 3', 4, 5$ hold, and $\gamma - 8\eta > 0$. Let $\{\lambda_k\}_k$ and $\{\rho_k\}_k$ be sequences of positive numbers such that Assumption (C1) holds, $\{x_k\}_k$, $\{y_k\}_k$ and $\{z_k\}_k$ be the sequences generated by Algorithm 1. Taking $\{\hat{x}\} = S(C, f)$, then for every $k \geq 0$, at least one of the following inequalities holds

$$\|x_{k+1} - \hat{x}\|^2 \leq \|y_k - \hat{x}\|^2 - p_k^{(1)} \|x_{k+1} - y_k\|^2, \quad (13)$$

$$\|x_{k+1} - \hat{x}\|^2 \leq \|y_k - \hat{x}\|^2 - p_k^{(2)} \|x_{k+1} - y_k\|^2. \quad (14)$$

Proof. This follows directly from Proposition 3.3, Part (a) of Remark 3.5 and the definitions of $p_k^{(i)}, i = 1, 2$. $\square$

Notation. Let $\hat{x} \in S(C, f)$. We now define

$$\begin{aligned}\Gamma_k^{(i)} &:= \|x_k - \hat{x}\|^2 - \theta \|x_{k-1} - \hat{x}\|^2 - \beta \|x_{k-2} - \hat{x}\|^2 \\ &\quad + p_k^{(i)}(1 + \beta - \theta) \|x_k - x_{k-1}\|^2, \quad i = 1, 2; \quad \forall k \geq 1,\end{aligned}$$

and

$$\bar{\Gamma}_k^{(i)} := \Gamma_k^{(i)} + c_1^{(i)} \|x_{k-1} - x_{k-2}\|^2, \quad i = 1, 2; \quad \forall k \geq 1$$

with

$$c_1^{(i)} := -\left((\theta - \beta)(1 + \theta) - p_k^{(i)}(\theta^2 - 2\theta + \beta\theta + \beta + 1)\right), \quad i = 1, 2.$$

We also define

$$\begin{aligned}c_2^{(i)} &:= -((\theta - \beta)(1 + \theta) - p_k^{(i)}(\theta^2 - 2\theta + \beta\theta + \beta + 1) \\ &\quad - \beta(\theta - \beta) - p_k^{(i)}(\beta^2 + \beta + \beta\theta)), \\ &= c_1^{(i)} + \beta(\theta - \beta) + p_k^{(i)}(\beta^2 + \beta + \beta\theta), \quad i = 1, 2.\end{aligned}$$

Remark 3.9. (a) If Assumptions (C1)–(C3) are satisfied, one has $c_1^{(i)} > 0, i = 1, 2$.

Indeed, if Assumptions (C1)–(C3) hold then $p_k^{(i)} > 0, i = 1, 2$. Consequently, for all $0 \leq \theta < \frac{1}{2}$, for all $\beta \leq 0$ one has

$$\begin{aligned}\frac{\theta}{1 + p_k^{(2)}} - p_k^{(2)} \frac{(\theta - 1)^2}{(1 + \theta)(1 + p_k^{(2)})} &< \beta \\ \iff \theta(1 + \theta) - p_k^{(2)}(\theta - 1)^2 &< \beta(1 + \theta)(1 + p_k^{(2)}).\end{aligned}$$

By rearranging we get that $c_1^{(2)} > 0$. Using (12) we derive that $c_1^{(1)} > 0$.

(b) If the Assumptions (Ci), $i = 1, 2, 4$ are satisfied, it holds that $c_2^{(i)} \geq -\alpha_1 > 0$ for $i = 1, 2$. Indeed, by rearranging, we get

$$\begin{aligned}-c_2^{(1)} &= ((\theta - \beta)(1 + \theta) - p_k^{(1)}(\theta^2 - 2\theta + \beta\theta + \beta + 1)) \\ &= \theta^2(1 - p_k^{(1)}) + \theta(1 - 2\beta + 2\beta p_k^{(1)} - 2p_k^{(1)}) - \beta(1 + 2p_k^{(1)}) + \beta^2(1 - p_k^{(1)}) - p_k^{(1)} \\ &\leq \theta^2(1 - p_k^{(2)}) + \theta(1 - 2\beta + 2\beta p_k^{(2)} - 2p_k^{(2)}) - \beta(1 + 2p_k^{(1)}) + \beta^2(1 - p_k^{(2)}) - p_k^{(2)} < 0.\end{aligned}$$

The last inequality holds since $0 < p_k^{(2)} \leq p_k^{(1)}, \theta \geq 0, \beta \leq 0$. Thus, $c_2^{(1)} > 0$.

Similarly,

$$\begin{aligned}-c_2^{(2)} &= ((\theta - \beta)(1 + \theta) - p_k^{(2)}(\theta^2 - 2\theta + \beta\theta + \beta + 1)) \\ &= \theta^2(1 - p_k^{(2)}) + \theta(1 - 2\beta + 2\beta p_k^{(2)} - 2p_k^{(2)}) - \beta(1 + 2p_k^{(2)}) + \beta^2(1 - p_k^{(2)}) - p_k^{(2)} \\ &\leq \theta^2(1 - p_k^{(2)}) + \theta(1 - 2\beta + 2\beta p_k^{(2)} - 2p_k^{(2)}) - \beta(1 + 2p_k^{(1)}) + \beta^2(1 - p_k^{(2)}) - p_k^{(2)} < 0.\end{aligned}$$

Again, the last inequality holds since $0 < p_k^{(2)} \leq p_k^{(1)}, 0 \leq \theta < \frac{1}{2}, \beta \leq 0$.

Thus, $c_2^{(2)} > 0$. $\square$

The following technical lemma is required for proving the subsequent results. It summarizes several properties of the sequences $\{\Gamma_k^{(i)}\}_k$ and $\{\bar{\Gamma}_k^{(i)}\}_k$.

Lemma 3.10. Let C be an affine subspace in $\mathbb{R}^n$, f be a bifunction defined on $C \times C$ such that Assumptions(Ai), $i = 1, 2, 3, 4, 5$ hold, and $\gamma - 8\eta > 0$. Let $\{\lambda_k\}_k$ and $\{\rho_k\}_k$ be sequences of positive numbers such that Assumptions (Ci), $i = 1, 2$, hold.

Assume that the inertial parameters $\theta \in [0, 1/2)$ and $\beta \leq 0$ satisfy (Ci), $i = 3, 4$, and let the sequences $\{x_k\}_k$, $\{y_k\}_k$ and $\{z_k\}_k$ be generated by Algorithm 1. Taking $\{\hat{x}\} = S(C, f)$. Then, the following statements hold for either $i = 1$ or $i = 2$.

- (a) $\Gamma_k^{(i)} \geq (1 - 2\theta)\|x_k - \hat{x}\|^2 + (p_k^{(i)}(1 + \beta - \theta) - 2\theta)\|x_k - x_{k-1}\|^2 - \beta\|x_{k-2} - \hat{x}\|^2, \forall k \geq 1.$
- (b) $\Gamma_k^{(i)} \geq 0, \bar{\Gamma}_k^{(i)} \geq 0, \forall k \geq 1.$
- (c) $\bar{\Gamma}_{k+1}^{(i)} \leq \bar{\Gamma}_k^{(i)} - c_2^{(i)}\|x_{k-1} - x_{k-2}\|^2, \forall k \geq 1.$
- (d) The sequence $\{\bar{\Gamma}_k^{(i)}\}_k$ is non-increasing.
- (e) The sequence $\{\bar{\Gamma}_k^{(i)}\}_k$ is convergent.

Proof. By Corollary 3.8, at least one of the inequalities (13) or (14) holds. Suppose first that (14) holds. We will prove that the sequences $\{\Gamma_k^{(i)}\}_k$ and $\{\bar{\Gamma}_k^{(i)}\}_k$ with $i = 2$ satisfy properties (a)–(e).

(a) By the Cauchy inequality we have

$$\begin{aligned}
 \Gamma_k^{(2)} &= \|x_k - \hat{x}\|^2 - \theta\|x_{k-1} - \hat{x}\|^2 - \beta\|x_{k-2} - \hat{x}\|^2 \\
 &\quad + p_k^{(2)}(1 + \beta - \theta)\|x_k - x_{k-1}\|^2 \\
 &= \|x_k - \hat{x}\|^2 - \theta\|x_{k-1} - x_k + x_k - \hat{x}\|^2 - \beta\|x_{k-2} - \hat{x}\|^2 \\
 &\quad + p_k^{(2)}(1 + \beta - \theta)\|x_k - x_{k-1}\|^2 \\
 &\geq \|x_k - \hat{x}\|^2 - 2\theta\|x_k - x_{k-1}\|^2 - 2\theta\|x_k - \hat{x}\|^2 \\
 &\quad - \beta\|x_{k-2} - \hat{x}\|^2 + p_k^{(2)}(1 + \beta - \theta)\|x_k - x_{k-1}\|^2 \\
 &= (1 - 2\theta)\|x_k - \hat{x}\|^2 + \left[p_k^{(2)}(1 + \beta - \theta) - 2\theta\right]\|x_k - x_{k-1}\|^2 \\
 &\quad - \beta\|x_{k-2} - \hat{x}\|^2, \forall k \geq 1.
 \end{aligned}$$

and thus, Part (a) is proved for $i = 2$.

(b) Note that under Conditions (C2) and (C3) of Assumption 3.6, it holds that $p_k^{(2)}(1 + \beta - \theta) - 2\theta > 0$ for all $k \geq 1$. Also, since $0 \leq \theta < \frac{1}{2}$, $\beta \leq 0$, it follows from Part (a) that $\Gamma_k^{(2)} \geq 0, \forall k \geq 1$.

In addition, from the definition of $\bar{\Gamma}_k^{(2)}$ and Remark 3.9, one has $\bar{\Gamma}_k^{(2)} \geq 0$ for all $k \geq 1$.

(c) One has, for all $k \geq 1$

$$\begin{aligned}
 y_k - \hat{x} &= x_k + \theta(x_k - x_{k-1}) + \beta(x_{k-1} - x_{k-2}) - \hat{x} \\
 &= (1 + \theta)(x_k - \hat{x}) - (\theta - \beta)(x_{k-1} - \hat{x}) - \beta(x_{k-2} - \hat{x}).
 \end{aligned}$$

Hence, by Part (a) of Lemma 2.2, we have

$$\begin{aligned}
 \|y_k - \hat{x}\|^2 &= \|(1 + \theta)(x_k - \hat{x}) - (\theta - \beta)(x_{k-1} - \hat{x}) - \beta(x_{k-2} - \hat{x})\|^2 \\
 &= (1 + \theta)\|x_k - \hat{x}\|^2 - (\theta - \beta)\|x_{k-1} - \hat{x}\|^2 - \beta\|x_{k-2} - \hat{x}\|^2 \\
 &\quad + (1 + \theta)(\theta - \beta)\|x_k - x_{k-1}\|^2 + \beta(1 + \theta)\|x_k - x_{k-2}\|^2 \\
 &\quad - \beta(\theta - \beta)\|x_{k-1} - x_{k-2}\|^2.
 \end{aligned} \tag{15}$$

Observe that

$$\begin{aligned}
 2\theta\langle x_{k+1} - x_k, x_k - x_{k-1} \rangle &= 2\langle \theta(x_{k+1} - x_k), x_k - x_{k-1} \rangle \\
 &\leq 2|\theta|\|x_{k+1} - x_k\|\|x_k - x_{k-1}\| \\
 &= 2\theta\|x_{k+1} - x_k\|\|x_k - x_{k-1}\|
 \end{aligned}$$

and so

$$-2\theta\langle x_{k+1} - x_k, x_k - x_{k-1} \rangle \geq -2\theta\|x_{k+1} - x_k\|\|x_k - x_{k-1}\|. \quad (16)$$

In addition,

$$\begin{aligned} 2\beta\langle x_{k+1} - x_k, x_{k-1} - x_{k-2} \rangle &= 2\langle \beta(x_{k+1} - x_k), x_{k-1} - x_{k-2} \rangle \\ &\leq 2|\beta|\|x_{k+1} - x_k\|\|x_{k-1} - x_{k-2}\|, \end{aligned}$$

which implies that

$$-2\beta\langle x_{k+1} - x_k, x_{k-1} - x_{k-2} \rangle \geq -2|\beta|\|x_{k+1} - x_k\|\|x_{k-1} - x_{k-2}\|. \quad (17)$$

Similarly, we note that

$$\begin{aligned} 2\beta\theta\langle x_{k-1} - x_k, x_{k-1} - x_{k-2} \rangle &= 2\langle \beta\theta(x_{k-1} - x_k), x_{k-1} - x_{k-2} \rangle \\ &\leq 2|\beta|\theta\|x_{k-1} - x_k\|\|x_{k-1} - x_{k-2}\| \\ &= 2|\beta|\theta\|x_k - x_{k-1}\|\|x_{k-1} - x_{k-2}\| \end{aligned}$$

and thus,

$$\begin{aligned} 2\beta\theta\langle x_k - x_{k-1}, x_{k-1} - x_{k-2} \rangle &= -2\beta\theta\langle x_{k-1} - x_k, x_{k-1} - x_{k-2} \rangle \\ &\geq -2|\beta|\theta\|x_k - x_{k-1}\|\|x_{k-1} - x_{k-2}\|. \end{aligned} \quad (18)$$

By (16), (17) and (18), and the Cauchy-Schwarz inequality one has

$$\begin{aligned} \|x_{k+1} - y_k\|^2 &= \|x_{k+1} - (x_k + \theta(x_k - x_{k-1}) + \beta(x_{k-1} - x_{k-2}))\|^2 \\ &= \|x_{k+1} - x_k - \theta(x_k - x_{k-1}) - \beta(x_{k-1} - x_{k-2})\|^2 \\ &= \|x_{k+1} - x_k\|^2 - 2\theta\langle x_{k+1} - x_k, x_k - x_{k-1} \rangle \\ &\quad - 2\beta\langle x_{k+1} - x_k, x_{k-1} - x_{k-2} \rangle + \theta^2\|x_k - x_{k-1}\|^2 \\ &\quad + 2\beta\theta\langle x_k - x_{k-1}, x_{k-1} - x_{k-2} \rangle + \beta^2\|x_{k-1} - x_{k-2}\|^2 \\ &\geq \|x_{k+1} - x_k\|^2 - 2\theta\|x_{k+1} - x_k\|\|x_k - x_{k-1}\| \\ &\quad - 2|\beta|\|x_{k+1} - x_k\|\|x_{k-1} - x_{k-2}\| + \theta^2\|x_k - x_{k-1}\|^2 \\ &\quad - 2|\beta|\theta\|x_k - x_{k-1}\|\|x_{k-1} - x_{k-2}\| + \beta^2\|x_{k-1} - x_{k-2}\|^2 \\ &\geq \|x_{k+1} - x_k\|^2 - \theta\|x_{k+1} - x_k\|^2 - \theta\|x_k - x_{k-1}\|^2 \\ &\quad - |\beta|\|x_{k+1} - x_k\|^2 - |\beta|\|x_{k-1} - x_{k-2}\|^2 + \theta^2\|x_k - x_{k-1}\|^2 \\ &\quad - |\beta|\theta\|x_k - x_{k-1}\|^2 - |\beta|\theta\|x_{k-1} - x_{k-2}\|^2 + \beta^2\|x_{k-1} - x_{k-2}\|^2 \\ &= (1 - |\beta| - \theta)\|x_{k+1} - x_k\|^2 + (\theta^2 - \theta - |\beta|\theta)\|x_k - x_{k-1}\|^2 \\ &\quad + (\beta^2 - |\beta| - |\beta|\theta)\|x_{k-1} - x_{k-2}\|^2. \end{aligned} \quad (19)$$

Substituting (15) and (19) into (14) with noting that $\beta \leq 0$, we obtain

$$\begin{aligned}
\|x_{k+1} - \hat{x}\|^2 &\leq (1 + \theta)\|x_k - \hat{x}\|^2 - (\theta - \beta)\|x_{k-1} - \hat{x}\|^2 - \beta\|x_{k-2} - \hat{x}\|^2 \\
&\quad + (1 + \theta)(\theta - \beta)\|x_k - x_{k-1}\|^2 + \beta(1 + \theta)\|x_k - x_{k-2}\|^2 \\
&\quad - \beta(\theta - \beta)\|x_{k-1} - x_{k-2}\|^2 - p_k^{(2)}(1 - |\beta| - \theta)\|x_{k+1} - x_k\|^2 \\
&\quad - p_k^{(2)}(\theta^2 - \theta - |\beta|\theta)\|x_k - x_{k-1}\|^2 \\
&\quad - p_k^{(2)}(\beta^2 - |\beta| - |\beta|\theta)\|x_{k-1} - x_{k-2}\|^2 \\
&= (1 + \theta)\|x_k - \hat{x}\|^2 - (\theta - \beta)\|x_{k-1} - \hat{x}\|^2 - \beta\|x_{k-2} - \hat{x}\|^2 \\
&\quad + ((1 + \theta)(\theta - \beta) - p_k^{(2)}(\theta^2 - \theta - |\beta|\theta))\|x_k - x_{k-1}\|^2 \\
&\quad + \beta(1 + \theta)\|x_k - x_{k-2}\|^2 - p_k^{(2)}(1 - |\beta| - \theta)\|x_{k+1} - x_k\|^2 \\
&\quad - (\beta(\theta - \beta) + p_k^{(2)}(\beta^2 - |\beta| - |\beta|\theta))\|x_{k-1} - x_{k-2}\|^2 \\
&\leq (1 + \theta)\|x_k - \hat{x}\|^2 - (\theta - \beta)\|x_{k-1} - \hat{x}\|^2 - \beta\|x_{k-2} - \hat{x}\|^2 \\
&\quad + ((1 + \theta)(\theta - \beta) - p_k^{(2)}(\theta^2 - \theta + \beta\theta))\|x_k - x_{k-1}\|^2 \\
&\quad - p_k^{(2)}(1 + \beta - \theta)\|x_{k+1} - x_k\|^2 \\
&\quad - (\beta(\theta - \beta) + p_k^{(2)}(\beta^2 + \beta + \beta\theta))\|x_{k-1} - x_{k-2}\|^2.
\end{aligned}$$

Rearranging, we get

$$\begin{aligned}
&\|x_{k+1} - \hat{x}\|^2 - \theta\|x_k - \hat{x}\|^2 - \beta\|x_{k-1} - \hat{x}\|^2 + p_k^{(2)}(1 + \beta - \theta)\|x_{k+1} - x_k\|^2 \\
&\leq \|x_k - \hat{x}\|^2 - \theta\|x_{k-1} - \hat{x}\|^2 - \beta\|x_{k-2} - \hat{x}\|^2 \\
&\quad + ((1 + \theta)(\theta - \beta) - p_k^{(2)}(\theta^2 - 2\theta + \beta\theta + \beta + 1))\|x_k - x_{k-1}\|^2 \\
&\quad - (\beta(\theta - \beta) + p_k^{(2)}(\beta^2 + \beta + \beta\theta))\|x_{k-1} - x_{k-2}\|^2 \\
&\quad + p_k^{(2)}(1 + \beta - \theta)\|x_k - x_{k-1}\|^2.
\end{aligned}$$

Using the definitions of $\Gamma_k^{(2)}$, $c_1^{(2)}$, $c_2^{(2)}$, this implies that

$$\begin{aligned}
\Gamma_{k+1}^{(2)} - \Gamma_k^{(2)} &\leq ((1 + \theta)(\theta - \beta) - p_k^{(2)}(\theta^2 - 2\theta + \beta\theta + \beta + 1))\|x_k - x_{k-1}\|^2 \\
&\quad - (\beta(\theta - \beta) + p_k^{(2)}(\beta^2 + \beta + \beta\theta))\|x_{k-1} - x_{k-2}\|^2 \\
&= -((1 + \theta)(\theta - \beta) - p_k^{(2)}(\theta^2 - 2\theta + \beta\theta + \beta + 1))\|x_{k-1} - x_{k-2}\|^2 \\
&\quad - \|x_k - x_{k-1}\|^2 + ((1 + \theta)(\theta - \beta) - p_k^{(2)}(\theta^2 - 2\theta + \beta\theta + \beta + 1)) \\
&\quad - \beta(\theta - \beta) - p_k^{(2)}(\beta^2 + \beta + \beta\theta))\|x_{k-1} - x_{k-2}\|^2 \\
&= c_1^{(2)}(\|x_{k-1} - x_{k-2}\|^2 - \|x_k - x_{k-1}\|^2) - c_2^{(2)}\|x_{k-1} - x_{k-2}\|^2. \tag{20}
\end{aligned}$$

By Remark 3.9, if Assumptions (C1)–(C4) are satisfied, it holds that $c_1^{(2)} > 0$, $c_2^{(2)} > 0$. Consequently, (20) yields

$$\Gamma_{k+1}^{(2)} + c_1^{(2)}\|x_k - x_{k-1}\|^2 \leq \Gamma_k^{(2)} + c_1^{(2)}\|x_{k-1} - x_{k-2}\|^2 - c_2^{(2)}\|x_{k-1} - x_{k-2}\|^2,$$

which implies Part (c) for $i = 2$.

(d) Because $c_2^{(2)} > 0$, it follows from Part (c) that

$$\bar{\Gamma}_{k+1}^{(2)} \leq \bar{\Gamma}_k^{(2)} \quad \text{for all } k \geq 1.$$

and thus, the sequence $\{\bar{\Gamma}_k^{(2)}\}_k$ is non-increasing.

(e) By Part (b) one has $\bar{\Gamma}_k^{(2)} \geq 0$ for all $k \geq 1$, and by Part (c), the sequence $\{\bar{\Gamma}_k^{(2)}\}_k$ is non-increasing. Hence $\lim_{k \rightarrow \infty} \bar{\Gamma}_k^{(2)}$ exists.

If (13) holds, then by repeating the same arguments with $p_k^{(2)}$ replaced by $p_k^{(1)}$ and using Remark 3.7, we conclude that properties (a) – (e) also hold for $\{\Gamma_k^{(1)}\}_k$ and $\{\bar{\Gamma}_k^{(1)}\}_k$. This completes the proof of the lemma. $\square$

The lemma below shows the boundedness of the sequence of iterates $\{x_k\}_k$ generated by Algorithm 1.

Lemma 3.11. *Let C be an affine subspace in $\mathbb{R}^n$, f be a bifunction defined on $C \times C$ such that Assumptions (Ai), $i = 1, 2, 3, 4, 5$ hold, and $\gamma - 8\eta > 0$. Let $\{\lambda_k\}_k$ and $\{\rho_k\}_k$ be sequences of positive numbers such that Assumptions (Ci), $i = 1, 2$, hold. Assume that the inertial parameters $\theta \in [0, 1/2)$ and $\beta \leq 0$ satisfy (Ci), $i = 3, 4$, and let $\{x_k\}_k$, $\{y_k\}_k$ and $\{z_k\}_k$ be the sequences generated by Algorithm 1. Then the following statements hold.*

- (a) $\lim_{k \rightarrow \infty} \|x_{k-1} - x_{k-2}\| = \lim_{k \rightarrow \infty} \|x_{k+1} - y_k\| = \lim_{k \rightarrow \infty} \|x_k - y_k\| = 0$.
- (b) $\lim_{k \rightarrow \infty} \|y_k - z_k\| = 0$.
- (c) The limit $\lim_{k \rightarrow \infty} [\|x_k - \hat{x}\|^2 - \theta\|x_{k-1} - \hat{x}\|^2 - \beta\|x_{k-2} - \hat{x}\|^2]$ exists.
- (d) The sequences $\{x_k\}_k$, $\{y_k\}_k$, $\{z_k\}_k$ are all bounded.

Proof. By Proposition 3.2, the solution set $S(C, f)$ is a singleton. We denote by $\hat{x}$ the unique solution of (EP). By Corollary 3.8, either (13) or (14) holds.

We first assume that (14) holds. Then we get from Parts (c) and (e) of Lemma 3.10 and the squeeze theorem that

$$\lim_{k \rightarrow \infty} c_2 \|x_{k-1} - x_{k-2}\|^2 = 0. \quad (21)$$

- (a) Because $c_2^{(2)} > 0$ (Part (ii) of Remark 3.9), the identity (21) yields

$$\lim_{k \rightarrow \infty} \|x_{k-1} - x_{k-2}\| = 0.$$

Consequently,

$$\begin{aligned} \|x_{k+1} - y_k\| &= \|x_{k+1} - x_k - \theta(x_k - x_{k-1}) - \beta(x_{k-1} - x_{k-2})\| \\ &\leq \|x_{k+1} - x_k\| + \theta\|x_k - x_{k-1}\| + |\beta|\|x_{k-1} - x_{k-2}\| \rightarrow 0 \text{ as } k \rightarrow \infty. \end{aligned}$$

By the squeeze theorem, we derive that

$$\lim_{k \rightarrow \infty} \|x_{k+1} - y_k\| = 0. \quad (22)$$

This also implies that,

$$\|x_k - y_k\| \leq \|x_k - x_{k+1}\| + \|x_{k+1} - y_k\| \rightarrow 0 \text{ as } k \rightarrow \infty.$$

By the squeeze theorem, we derive that

$$\lim_{k \rightarrow \infty} \|x_k - y_k\| = 0.$$

- (b) From the relaxation step (Step 4) of Algorithm 1 we have $x_{k+1} - y_k = \rho_k(y_k - z_k)$. It follows from (C2) and (22) that $\lim_{k \rightarrow \infty} \|y_k - z_k\| = 0$.

(c) By the definition of $\bar{\Gamma}_k^{(2)}$, the fact that $\lim_{k \rightarrow \infty} \|x_{k-1} - x_{k-2}\| = 0$ in Part (a), and the existence of $\lim_{k \rightarrow \infty} \bar{\Gamma}_k^{(2)}$ (Part (e) of Lemma 3.10), we have that $\lim_{k \rightarrow \infty} \Gamma_k^{(2)}$ exists, and hence, the sequence $\{\Gamma_k^{(2)}\}_k$ is bounded. Now, since $\lim_{k \rightarrow \infty} \|x_k - x_{k-1}\| = 0$, we have from the definition of $\Gamma_k^{(2)}$ that

$$\lim_{k \rightarrow \infty} [\|x_k - \hat{x}\|^2 - \theta\|x_{k-1} - \hat{x}\|^2 - \beta\|x_{k-2} - \hat{x}\|^2] \text{ exists.}$$

(d) By Assumptions (C2) and (C3) one has $p_k^{(2)}(1 + \beta - \theta) - 2\theta > 0$, and since $\beta \leq 0$, and $\{\Gamma_k^{(2)}\}_k$ is bounded, it follows from Part (a) of Lemma 3.10 that

$$(1 - 2\theta)\|x_k - \hat{x}\|^2 \leq \Gamma_k^{(2)} \leq M, \quad \forall k \geq 1,$$

for some constant M . Since $0 \leq \theta < \frac{1}{2}$ and $\hat{x}$ is fixed, this implies that $\{x_k\}_k$ is bounded. The boundedness of $\{y_k\}_k$ is followed from the fact that $\lim_{k \rightarrow \infty} \|x_k - y_k\| = 0$ in Part (a) and the boundedness of $\{x_k\}_k$. Finally, it follows from Part (b) and the boundedness of $\{y_k\}_k$ that $\{z_k\}_k$ is also bounded.

If inequality (13) holds, then by repeating the same argument as above with $p_k^{(2)}$ replaced by $p_k^{(1)}$ and using Remark 3.7, we obtain the same conclusion.

Therefore, the proof is completed. $\square$

Our convergence result for Algorithm 1 is given next.

Theorem 3.12. *Let C be an affine subspace in $\mathbb{R}^n$, f be a bifunction defined on $C \times C$ such that the Assumptions (Ai), $i = 1, 2, 3, 4, 5$ hold, and $\gamma - 8\eta > 0$. Let $\{\lambda_k\}_k$ and $\{\rho_k\}_k$ be sequences of positive numbers such that the Assumptions (Ci), $i = 1, 2$ hold. Assume that the inertial parameters $\theta \in [0, 1/2)$ and $\beta \leq 0$ satisfy (Ci), $i = 3, 4$. Then the sequence $\{x_k\}_k$ generated by Algorithm 1 converges to a solution of (EP).*

Proof. By Lemma 3.11, we have that $\{x_k\}_k$ is bounded. Suppose that x^* is an accumulating point of $\{x_k\}_k$. Parts (a) and (b) of Lemma 3.11 yield that x^* is also an accumulating point of $\{y_k\}_k$ and of $\{z_k\}_k$, i.e.,

$$\lim_{m \rightarrow \infty} y_{k_m} = \lim_{m \rightarrow \infty} z_{k_m} = x^*. \quad (23)$$

From the definition of z_k , we have that for any $k \geq 1$,

$$f(y_k, z_k) + \frac{1}{2\lambda_k}\|y_k - z_k\|^2 \leq f(y_k, x) + \frac{1}{2\lambda_k}\|y_k - x\|^2, \quad \forall x \in C.$$

Because $\lambda_k > 0$, it follows that

$$f(y_k, z_k) \leq f(y_k, x) + \frac{1}{2\lambda_k}\|y_k - x\|^2, \quad \forall x \in C. \quad (24)$$

Taking $x = ty + (1 - t)x^*$ with $y \in C$ and $t \in [0, 1]$, we obtain from (24) that

$$f(y_k, z_k) \leq f(y_k, ty + (1 - t)x^*) + \frac{1}{2\lambda_k}\|t(y_k - y) + (1 - t)(y_k - x^*)\|^2. \quad (25)$$

By Assumption (A4), for each $y_k \in C$, the function $g(\cdot) := f(y_k, \cdot)$ is strongly quasiconvex with modulus $\gamma > 0$. Therefore, (25) yields

$$\begin{aligned} f(y_k, z_k) &\leq \max\{f(y_k, y), f(y_k, x^*)\} - \frac{t(1-t)\gamma}{2}\|y - x^*\|^2 \\ &\quad + \frac{t^2}{2\lambda_k}\|y_k - y\|^2 + \frac{(1-t)^2}{2\lambda_k}\|y_k - x^*\|^2 + \frac{t(1-t)}{\lambda_k}\langle y_k - y, y_k - x^* \rangle, \end{aligned} \quad (26)$$

for all $y \in C$ and for all $t \in [0, 1]$. In addition, we have

$$\|y_k - y\|^2 = \|y_k - x^* + x^* - y\|^2 = \|y_k - x^*\|^2 + \|x^* - y\|^2 + 2\langle y_k - x^*, x^* - y \rangle \quad (27)$$

and

$$\langle y_k - y, y_k - x^* \rangle = \langle y_k - x^* + x^* - y, y_k - x^* \rangle = \|y_k - x^*\|^2 + \langle x^* - y, y_k - x^* \rangle. \quad (28)$$

Combining (26), (27) and (28) we obtain

$$\begin{aligned} f(y_k, z_k) &\leq \max\{f(y_k, y), f(y_k, x^*)\} + \frac{t^2}{2\lambda_k}(\|y_k - x^*\|^2 + \|x^* - y\|^2) \\ &\quad + \frac{t^2}{\lambda_k}\langle y_k - x^*, x^* - y \rangle + \frac{(1-t)^2}{2\lambda_k}\|y_k - x^*\|^2 \\ &\quad + \frac{t(1-t)}{\lambda_k}(\|y_k - x^*\|^2 + \langle x^* - y, y_k - x^* \rangle) - \frac{t(1-t)\gamma}{2}\|y - x^*\|^2, \quad \forall k \geq 1. \end{aligned}$$

Simplifying the right-hand side, one has that

$$\begin{aligned} f(y_k, z_k) &\leq \max \{f(y_k, y), f(y_k, x^*)\} + \frac{1}{2\lambda_k} \|y_k - x^*\|^2 \\ &\quad + \frac{t}{2} \left(\frac{t}{\lambda_k} - (1-t)\gamma \right) \|x^* - y\|^2 + \frac{t}{\lambda_k} \langle x^* - y, y_k - x^* \rangle, \quad \forall k \geq 1. \end{aligned} \quad (29)$$

Take $\lambda^* := 1/(\gamma - 8\eta)$. Then, by Assumptions (C2), it holds that $0 < \lambda^* \leq \lambda_k$ for all $k \geq 1$. Substituting k by k_m in (29), then taking $\limsup$, using (23) and Assumption (A3) one has

$$\begin{aligned} 0 &= f(x^*, x^*) \leq \liminf_{m \rightarrow \infty} f(y_{k_m}, z_{k_m}) \leq \limsup_{m \rightarrow \infty} f(y_{k_m}, z_{k_m}) \\ &\leq \max \{f(x^*, y), 0\} + \frac{t}{2} \left(\frac{t}{\lambda^*} - \gamma + \gamma t \right) \|y - x^*\|^2, \quad \forall y \in C, \quad \forall t \in [0, 1]. \end{aligned}$$

Since $\gamma > 0$ we can take $t < (\lambda^* \gamma)/(1 + \lambda^* \gamma)$ such that $t/\lambda^* - \gamma + t\gamma < 0$. Then the inequality above implies that

$$0 \leq \max \{f(x^*, y), 0\} = f(x^*, y), \quad \forall y \in C \setminus \{x^*\}.$$

Thus, $x^* \in S(C, f)$.

We now suppose that there exist $\{x_{k_j}\} \subset \{x_k\}$ and $\{x_{k_m}\} \subset \{x_k\}$ such that $x_{k_j} \rightarrow \bar{x}$ as $j \rightarrow \infty$ and $x_{k_m} \rightarrow x^*$ as $m \rightarrow \infty$. We will show that $\bar{x} = x^*$.

Observe that

$$2\langle x_k, x^* - \bar{x} \rangle = \|x_k - \bar{x}\|^2 - \|x_k - x^*\|^2 - \|\bar{x}\|^2 + \|x^*\|^2, \quad (30)$$

$$2\langle x_{k-1}, x^* - \bar{x} \rangle = \|x_{k-1} - \bar{x}\|^2 - \|x_{k-1} - x^*\|^2 - \|\bar{x}\|^2 + \|x^*\|^2,$$

and

$$2\langle x_{k-2}, x^* - \bar{x} \rangle = \|x_{k-2} - \bar{x}\|^2 - \|x_{k-2} - x^*\|^2 - \|\bar{x}\|^2 + \|x^*\|^2.$$

Therefore,

$$2\langle -\theta x_{k-1}, x^* - \bar{x} \rangle = -\theta \|x_{k-1} - \bar{x}\|^2 + \theta \|x_{k-1} - x^*\|^2 + \theta \|\bar{x}\|^2 - \theta \|x^*\|^2 \quad (31)$$

and

$$2\langle -\beta x_{k-2}, x^* - \bar{x} \rangle = -\beta \|x_{k-2} - \bar{x}\|^2 + \beta \|x_{k-2} - x^*\|^2 + \beta \|\bar{x}\|^2 - \beta \|x^*\|^2. \quad (32)$$

Adding (30), (31) and (32) gives

$$\begin{aligned} 2\langle x_k - \theta x_{k-1} - \beta x_{k-2}, x^* - \bar{x} \rangle &= (\|x_k - \bar{x}\|^2 - \theta \|x_{k-1} - \bar{x}\|^2 - \beta \|x_{k-2} - \bar{x}\|^2) \\ &\quad - (\|x_k - x^*\|^2 - \theta \|x_{k-1} - x^*\|^2 - \beta \|x_{k-2} - x^*\|^2) + (1 - \theta - \beta)(\|x^*\|^2 - \|\bar{x}\|^2). \end{aligned}$$

According to Part (c) of Lemma 3.11, the limit

$$\lim_{k \rightarrow \infty} \left[\|x_k - x^*\|^2 - \theta \|x_{k-1} - x^*\|^2 - \beta \|x_{k-2} - x^*\|^2 \right]$$

exists and the limit

$$\lim_{k \rightarrow \infty} \left[\|x_k - \bar{x}\|^2 - \theta \|x_{k-1} - \bar{x}\|^2 - \beta \|x_{k-2} - \bar{x}\|^2 \right]$$

also exists.

These last three facts imply that

$$\lim_{k \rightarrow \infty} \langle x_k - \theta x_{k-1} - \beta x_{k-2}, x^* - \bar{x} \rangle$$

exists. Now,

$$\begin{aligned} \langle \bar{x} - \theta \bar{x} - \beta \bar{x}, x^* - \bar{x} \rangle &= \lim_{j \rightarrow \infty} \langle x_{k_j} - \theta x_{k_j-1} - \beta x_{k_j-2}, x^* - \bar{x} \rangle \\ &= \lim_{k \rightarrow \infty} \langle x_k - \theta x_{k-1} - \beta x_{k-2}, x^* - \bar{x} \rangle \\ &= \lim_{m \rightarrow \infty} \langle x_{k_m} - \theta x_{k_m-1} - \beta x_{k_m-2}, x^* - \bar{x} \rangle \\ &= \langle x^* - \theta x^* - \beta x^*, x^* - \bar{x} \rangle. \end{aligned}$$

Hence,

$$(1 - \theta - \beta) \|x^* - \bar{x}\|^2 = 0.$$

Since $\beta \leq 0 < 1 - \theta$, we obtain that $x^* = \bar{x}$. Therefore, every accumulation point of $\{x_k\}_k$ is a solution to (EP). This completes the proof. $\square$

4. Application to mixed variational inequality problems

In this section, we apply Algorithm 1 to a mixed variational inequality problem (MVIP): Find $x^* \in C$ such that

$$\langle T(x^*), y - x^* \rangle + h(y) - h(x^*) \geq 0, \forall y \in C, \quad (33)$$

where $T : C \longrightarrow \mathbb{R}^n$ and $h : C \longrightarrow \mathbb{R}$.

We now define
$$f(x, y) = \langle T(x), y - x \rangle + h(y) - h(x). \quad (34)$$

Then the equilibrium problem (EP) becomes the MVIP (33) and that

$$\arg \min_{x \in C} \left(f(y_k, x) + \frac{1}{2\lambda_k} \|y_k - x\|^2 \right) = \text{Prox}_{\lambda_k h} (C, y_k - \lambda_k T(y_k)).$$

As a result, Algorithm 1 is transformed into the following for MVIPs.

Algorithm 2: Relaxed Double Inertial Proximal Point Algorithm for Mixed Variational Inequality Problems (RDIPPA-MVIP)

- 1: Choose $\beta \in (-\infty, 0]$, $\rho \in [0, 1)$, $\theta \in [0, \frac{1}{2})$ and $\lambda_k > 0$ for all $k = 1, 2, 3, \dots$. Pick $x_{-1}, x_0, x_1 \in C$ and set $k = 1$.
- 2: Given x_{k-2}, x_{k-1} and x_k , compute

$$\begin{aligned} y_k &= x_k + \theta(x_k - x_{k-1}) + \beta(x_{k-1} - x_{k-2}), \\ z_k &= \text{Prox}_{\lambda_k h} (C, y_k - \lambda_k T(y_k)). \end{aligned}$$

- 3: If $y_k = z_k$ then STOP. Otherwise, go to Step 4.
- 4: Choose some relaxation parameter $\rho_k \in [1 - \rho, 1 + \rho]$ and update

$$x_{k+1} = (1 - \rho_k)y_k + \rho_k z_k.$$

- 5: Let $k \leftarrow k + 1$ and go to Step 2.

To analyze the convergence for Algorithm 2 we will make the following assumptions on the operator T and the function h .

(B1) T is h -quasimonotone on C , i.e., for all $x, y \in C$

$$\langle T(x), y - x \rangle + h(y) - h(x) \geq 0 \implies \langle T(y), x - y \rangle + h(x) - h(y) \leq 0;$$

(B2) T is L -Lipschitz continuous on C with Lipschitz constant $L > 0$;

(B3) h is strongly convex on C with modulus $\gamma > 0$;

(B4) h is continuous on C .

Theorem 4.1. Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $h : C \longrightarrow \mathbb{R}$ be such that the Assumptions (Bi), $i = 1, 2, 3, 4$ hold. Then the bifunction f defined in (34) satisfies all Assumptions (Ai), for all $i = 1, 2, 3, 4, 5$.

Proof. Because h is continuous on C , and T is L -Lipschitz continuous, it holds that f is continuous on $C \times C$. Hence, f is lower continuous in both arguments on $C \times C$ and for each $y \in C$, the function $f(\cdot, y)$ is upper continuous on C . Thus, Assumptions (A1) and (A3) are satisfied by f . We now show that f satisfies Assumption (A2).

Suppose that $f(x, y) \geq 0$ for all $x, y \in C$.

By the definition of f , we have

$$\langle T(x), y - x \rangle + h(y) - h(x) \geq 0, \forall x, y \in C.$$

Because T is h -pseudomonotone, it follows that

$$f(y, x) = \langle T(y), x - y \rangle + h(x) - h(y) \leq 0, \forall x, y \in C.$$

Hence, (A2) is satisfied by f .

As for Assumption (A4), for each $x \in C$, let

$$g(\cdot) := f(x, \cdot) = \langle T(x), \cdot - x \rangle + h(\cdot) - h(x).$$

Since h is strongly convex on C , we obtain

$$\begin{aligned} g(ty_1 + (1-t)y_2) &= \langle T(x), (ty_1 + (1-t)y_2) - x \rangle + h(ty_1 + (1-t)y_2) - h(x) \\ &\leq t\langle T(x), y_1 - x \rangle + (1-t)\langle T(x), y_2 - x \rangle + th(y_1) + (1-t)h(y_2) \\ &\quad - th(x) - (1-t)h(x) - t(1-t)\frac{\gamma}{2}\|y_1 - y_2\|^2 \\ &= tg(y_1) + (1-t)g(y_2) - t(1-t)\frac{\gamma}{2}\|y_1 - y_2\|^2, \end{aligned}$$

for all $y_1, y_2 \in C$ and for all $t \in [0, 1]$.

Hence, g is strongly convex on C , meaning that $f(x, \cdot)$ is strongly convex on C . This implies that $f(x, \cdot)$ is strongly quasiconvex on C . Thus, (A4) is satisfied by f .

We now verify that f satisfies (A5).

By the definition of f , for all $x, y, z \in C$,

$$\begin{aligned} f(x, z) - f(x, y) - f(y, z) &= \langle T(x), z - x \rangle + h(z) - h(x) - \langle T(x), y - x \rangle \\ &\quad - h(y) + h(x) - \langle T(y), z - y \rangle - h(z) + h(y) \\ &= \langle T(x) - T(y), z - y \rangle. \end{aligned} \tag{35}$$

Using the Cauchy-Schwarz inequality, the L -Lipschitz continuity of T , and the Cauchy inequality it holds that

$$\begin{aligned} \langle T(x) - T(y), z - y \rangle &\leq \|T(x) - T(y)\| \|y - z\| \leq L\|x - y\| \|y - z\| \\ &\leq \frac{L}{2} (\|x - y\|^2 + \|y - z\|^2). \end{aligned} \tag{36}$$

Combining (35) and (36), we obtain

$$f(x, z) - f(x, y) - f(y, z) \leq \frac{L}{2} (\|x - y\|^2 + \|y - z\|^2),$$

which implies that Assumption (A5) is satisfied by f with $\eta = \frac{L}{2}$.

Thus, the proof of the theorem is complete. $\square$

By Theorem 4.1, Theorem 3.12, and the transformation from Algorithm 1 to Algorithm 2, applied to the bifunction f defined in (34), we obtain the following corollary.

Corollary 4.2. *Let C be a affine, closed subset of $\mathbb{R}^n$. Let $T : C \rightarrow \mathbb{R}^n$ and $h : C \rightarrow \mathbb{R}$ be such that all Assumptions (Bi), $i = 1, 2, 3, 4$ hold. Assume further that all Conditions (Ci), $i = 1, 2, 3, 4$ in Assumption 3.6 hold. Then the sequence $\{x_k\}_k$ generated by the Algorithm 2 converges to the solution x^* of the MVIP (33).*

5. Numerical experiments

In this section, we present numerical experiments to compare our proposed Algorithm to the one of Grad, Lara, and Marcavillaca in [19]. The codes were written in Python running on <https://colab.research.google.com> on a HP ProBook 430G6 Laptop with Windows 11 Home Single Language and an Intel Core i5 8265U CPU with 1.60 GHz and 4.0GB RAM.

The numerical example below illustrates a situation in which the method proposed in this paper exhibits superior performance—in terms of both the number of iterations required and the CPU time needed to reach a solution of (EP)—compared with the Relaxed-Inertial Proximal Point Algorithm (RIPPA) proposed by Grad, Lara, and Marcavillaca in [19].

As the stopping criterion for the proposed algorithms, we consider the case where the norm (in particular, the absolute value) of the difference between y_k and z_k is not larger than a prescribed tolerance $\varepsilon > 0$; that is, the algorithm terminates when $|z_k - y_k| < \varepsilon$.

Example 5.1. [19] Let $p, q \in \mathbb{N}, p > 1, q > 1$. Let $f : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ be defined by

$$f(x, y) = p(\max\{\sqrt{|y|}, (y - q)^2 - q\}) - p(\max\{\sqrt{|x|}, (x - q)^2 - q\}) + x(y - x).$$

As shown in [19] the bifunction f satisfies all the assumptions (Ai), $i = 1, 2, 3, 4, 5$ with $\eta = 1/2$. $\square$

We implement a numerical experiment with the parameters chosen as follows:

$p = 2, q = 99, \epsilon = \frac{1}{4}, \rho = \frac{2}{5}, \varepsilon = 10^{-12}, \beta = -0.001, \theta = 0.12, \rho_k = 1.4, \lambda_k = 0.1$ for all $k = 1, 2, 3, \dots$, and with initial points $x^{-1} = 3210$ and $x_0 = x_1 = 8297$.

A straightforward verification shows that all assumptions (Ci), $i = 1, 2, 3, 4$, are satisfied for the chosen parameter values.

Table 1 reports a comparison of the performance of RDIPPA and RIPPA under these settings.

Table 1. Comparison of the performance of RDIPPA and RIPPA with $\beta = -0.001, \theta = 0.12, \rho_k = 1.4$, and $\lambda_k = 0.1$ for all $k = 1, 2, 3, \dots$

Algorithm	RDIPPA	RIPPA
	$\varepsilon = 10^{-12}$	$\varepsilon = 10^{-12}$
Approximate solution	79.19999999989813	79.19999999996573
Time	0.005859375	0.010273456573486328
Iterations	40	54

Next, the algorithm is executed with the parameters $p = 2, q = 99, \epsilon = \frac{1}{4}, \rho = \frac{2}{5}, \varepsilon = 10^{-13}, \beta = -0.0001, \theta = 0.25, \rho_k = 1.4, \lambda_k = 0.2$ for all $k = 1, 2, 3, \dots$, and with the same initial points $x^{-1} = 3210$ and $x_0 = x_1 = 8297$. The numerical results are presented in Table 2.

Table 2. Comparison of the performance of TIPPM and RIPPA with $\beta = -0.001, \theta = 0.12, \rho_k = 1.4$, and $\lambda_k = 0.1$ for all $k = 1, 2, 3, \dots$

Algorithm	TIPPM	RIPPA
	$\varepsilon = 10^{-13}$	$\varepsilon = 10^{-13}$
Approximate solution	79.1999999999766	79.1999999999986
Time	0.029584646224975586	0.04530167579650879
Iterations	152	313

From Tables 1 and 2 we observe that the proposed algorithm attains approximate solutions in less computational time and with fewer iterations than the algorithm presented in [19].

6. Conclusion

In this paper, we showed that the relaxation proximal point method with two-step inertial extrapolation can be adapted to solve non-convex equilibrium problems. We established the global convergence of the iterative sequence generated by the proposed method, and provided numerical illustrations. Our numerical experiments highlight the advantages of using two-step inertial extrapolation over the one-step approach commonly employed in related studies on (convex) equilibrium problems. In future work, we plan to extend the proposed method by incorporating a combination of inertial and correction terms, or we may develop the algorithm in more general spaces.

Acknowledgments: The author sincerely thanks the anonymous referees for their valuable comments and suggestions. The author thanks Ho Chi Minh City University of Technology and Engineering for its financial support under Project Number T2025-119.

References

- [1] J. Abubakar, P. Kumam, A. H. Ibrahim, A. Padcharoen: *Relaxed inertial Tseng's type method for solving the inclusion problem with application to image restoration*, Mathematics 8/5 (2020), art.no. 818.

- [2] Q. H. Ansari, F. Babu, M. S. Raju: *Proximal point method with Bregman distance for quasiconvex pseudomonotone equilibrium problems*, Optimization 74/9 (2025) 2035–2055.
- [3] A. S. Antipin: *Equilibrium programming: proximal methods*, Comput. Mat. Math. Phys. 37 (1997) 1285–1296 (1997); translation from Zh. Vychisl. Mat. Mat. Fiz. 37/11 (1997) 1327–1339.
- [4] A. S. Antipin: *The convergence of proximal methods to fixed points of extremal mappings and estimates of their rate of convergence*, Comput. Math. Math. Phys. 35 (1995) 539–551.
- [5] A. S. Antipin: *Equilibrium programming: proximal methods*, Comput. Mat. Math. Phys. 37 (1997) 1285–1296.
- [6] G. Bigi, M. Castellani, M. Pappalardo, M. Passacantando: *Nonlinear Programming Techniques for Equilibria*, Springer, Berlin (2019).
- [7] E. Blum, W. Oettli: *From optimization and variational inequalities to equilibrium problems*, Math. Stud. 63 (1994) 123–145.
- [8] R. I. Boţ, E. R. Csetnek: *An inertial forward-backward-forward primal-dual splitting algorithm for solving monotone inclusion problems*, Numer. Algorithms 71 (2016) 519–540.
- [9] R. I. Boţ, E. R. Csetnek: *An inertial Tseng's type proximal algorithm for nonsmooth and nonconvex optimization problems*, J. Optim. Theory Appl. 171/2 (2016) 600–616.
- [10] R. I. Boţ, M. Sedlmayer, P. T. Vuong: *A relaxed inertial forward-backward-forward algorithm for solving monotone inclusions with application to GANs*, J. Mach. Learn. Res. 24 (2023), art.no. 8, 37p.
- [11] A. Cambini, L. Martein: *Generalized Convexity and Optimization*, Springer, Berlin (2009).
- [12] C. Chen, S. Ma, J. Yang: *A general inertial proximal point algorithm for mixed variational inequality problem*, SIAM J. Optim. 25 (2015) 2120–2142.
- [13] A. Cherukuri, B. Gharesifard, J. Cortes: *Saddle-point dynamics: conditions for asymptotic stability of saddle points*, SIAM J. Control Optim. 55 (2017) 486–511.
- [14] W. Chulamjiak, P. Chulamjiak, S. Suantai: *An inertial forward-backward splitting method for solving inclusion problems in Hilbert spaces*, J. Fixed Point Theory Appl. 20/1 (2018), art.no. 42, 17p.
- [15] P. L. Combettes, L. E. Glaudin: *Quasi-nonexpansive iterations on the affine hull of orbits: from Mann's mean value algorithm to inertial methods*, SIAM J. Optim. 27 (2017) 2356–2380.
- [16] K. Fan: *A minimax inequality and applications*, in: *Inequality III*, Proc. 3rd Symp., Los Angeles (1969), O. Shisha (ed.), Academic Press, New York (1972) 103–113.
- [17] S.-M. Grad, F. Lara: *Solving mixed variational inequalities beyond convexity*, J. Optim. Theory Appl. 190 (2021) 565–580.
- [18] S.-M. Grad, F. Lara: *An extension of the proximal point algorithm beyond convexity*, J. Global Optim. 82/2 (2022) 313–329.
- [19] S.-M. Grad, F. Lara, R. T. Marcavillaca: *Relaxed-inertial proximal point type algorithms for quasiconvex minimization*, J. Global Optim. 85/3 (2023) 615–635.
- [20] S.-M. Grad, F. Lara, R. Marcavillaca: *Relaxed-inertial proximal point algorithms for nonconvex equilibrium problems with applications*, J. Optim. Theory Appl. 203/3 (2024) 2233–2262.
- [21] N. Hadjisavvas, S. Komlosi, S. Schaible: *Handbook of Generalized Convexity and Generalized Monotonicity*, Springer, Boston (2005).
- [22] A. Iusem, F. Lara: *Optimality conditions for vector equilibrium problems with applications*, J. Optim. Theory Appl. 180 (2019) 187–206.
- [23] A. Iusem, F. Lara: *Proximal point algorithms for quasiconvex pseudo-monotone equilibrium problems*, J. Optim. Theory Appl. 193 (2022) 443–461.
- [24] M. Jovanović: *A note on strongly convex and quasiconvex functions*, Math. Notes 60 (1996) 584–585.
- [25] F. Lara: *On strongly quasiconvex functions: existence results and proximal point algorithms*, J. Optim. Theory Appl. 192 (2022) 891–911.
- [26] J. Liang: *Convergence Rates of First-Order Operator Splitting Methods*, Ph. D. Thesis, Normandie University, France, GREYC CNRS UMR 6072 (2016).

- [27] O. T. Mewomo, R. N. Nwokoye, C. C. Okeke: *Two-step inertial Tseng's extragradient method for solving quasimonotone variational inequalities*, Quaestiones Mathematicae 47/7 (2024) 1505–1543.
- [28] L. D. Muu, V. H. Nguyen, N. V. Quy: *On Nash–Cournot oligopolistic market equilibrium models with concave cost functions*, J. Global Optim. 41 (2008) 351–364.
- [29] L. D. Muu, W. Oettli: *Convergence of an adaptive penalty scheme for finding constraint equilibria*, Nonlinear Analysis 18 (1992) 1159–1166.
- [30] A. Padcharoen, D. Kitkuan, W. Kumam, P. Kumam: *Tseng methods with inertial for solving inclusion problems and application to image deblurring and image recovery problems*, Comp. Math. Methods 3/3 (2020), art.no. e1088.
- [31] B. T. Polyak: *Some methods of speeding up the convergence of iteration methods*, U.S.S.R. Comput. Math. Math. Phys. 4/5 (1964) 1–17.
- [32] B. T. Polyak: *Existence theorems and convergence of minimizing sequences in extremum problems with restrictions*, Sov. Math. Dokl. 7 (1966) 72–75; translation from Dokl. Akad. Nauk SSSR 166 (1966) 287–290.
- [33] B. T. Polyak: *Introduction to Optimization*, Optimization Software Publications Division, New York (1987).
- [34] C. Poon, L. Jingwei: *Geometry of first-order methods and adaptive acceleration*, arXiv: 2003.03910 (2020).
- [35] T. D. Quoc, L. D. Muu, N. V. Hien: *Extragradient algorithms extended to equilibrium problems*, Optimization 57 (2008) 749–766.
- [36] B. D. Rouhani, M. R. Piranfar: *Asymptotic behavior for a quasi-autonomous gradient system of an expansive type governed by a quasiconvex function*, Electron. J. Diff. Equations 2021 (2021), art.no. 15, 13p.
- [37] V. N. Tran, S. Yekini, X. Renqui, T. V. Phan: *An inertial forward-backward-forward algorithm for non-convex mixed variational inequality*, J. Appl. Numer. Optim. 5/3 (2023) 335–348.
- [38] J. P. Vial: *Strong convexity of sets and functions*, J. Math. Economics 9 (1982) 187–205.
- [39] L. H. Yen, L. D. Muu: *A parallel subgradient projection algorithm for quasiconvex equilibrium problems under the intersection of convex sets*, Optimization 71 (2022) 4447–4462.
- [40] L. H. Yen, L. D. Muu: *An extragradient algorithm for quasiconvex equilibrium problems without monotonicity*, J. Global Optim. (2023).