CARPATHIAN J. MATH. Online version at https://www.carpathian.cunbm.utcluj.ro
36 (2020), No. 1, 91 - 107 Print Edition: ISSN 1584 - 2851 Online Edition: ISSN 1843 - 4401

Dedicated to Prof. Hong-Kun Xu on the occasion of his 60^{th} anniversary

An inertial extragradient method for solving bilevel equilibrium problems

JIRAPRAPA MUNKONG[1], BUI VAN DINH[2] and KASAMSUK UNGCHITTRAKOOL[1,3]

ABSTRACT. In this paper, we propose an algorithm with two inertial term extrapolation steps for solving bilevel equilibrium problem in a real Hilbert space. The inertial term extrapolation step is introduced to speed up the rate of convergence of the iteration process. Under some sufficient assumptions on the bifunctions involving pseudomonotone and Lipschitz-type conditions, we obtain the strong convergence of the iterative sequence generated by the proposed algorithm. A numerical experiment is performed to illustrate the numerical behavior of the algorithm and also comparison with some other related algorithms in the literature.

1. INTRODUCTION

Let C be a nonempty closed convex subset of a real Hilbert space H, and let f and g be bifunctions from $H \times H$ to $\mathbb{R}$ such that $f(x,x) = 0$ and $g(x,x) = 0$ for all $x \in H$. The equilibrium problem associated with g and C is denoted by EP(C, g) : Find $x^* \in C$ such that

$$g(x^*, y) \geq 0 \quad \text{for every} \quad y \in C, \tag{1.1}$$

is well known as the Ky Fan inequality early studied in [18, 39]. We denote the solution set of problem (1.1) by Ω.

The equilibrium problem can be considered as a generalization of many mathematical models such as the fixed point problem, the (generalized) Nash equilibrium problem in game theory, the saddle point problem, the variational inequality problem, the optimization problem and others (see, e.g., [4, 17, 27, 36]).

One of the most popular for solving equilibrium problems is the proximal point method. This method was first introduced by Martinet [32] for monotone variational inequality problems. After that, it was extended by many authors (see, for instant [26, 35, 42]). In 2008, Tran et al. [46] proposed the extragradient algorithm for solving the equilibrium problem by using the strongly convex minimization problem to solve at each iteration. Furthermore, Hieu [20] introduced subgradient extragradient methods for pseudomonotone equilibrium problem and the other methods (see the details in [1, 13, 21, 25, 28, 38, 48]).

In this paper, we consider the bilevel equilibrium problem, that is, the equilibrium problem whose constraints are the solution sets of equilibrium problem: find $x^* \in \Omega$ such that

$$f(x^*, y) \geq 0 \quad \text{for every} \quad y \in \Omega, \tag{1.2}$$

Received: 30.11.2019. In revised form: 13.02.2020. Accepted: 20.02.2020
2010 *Mathematics Subject Classification.* 47J25, 47H05, 47J20, 65K15, 90C25.
Key words and phrases. *Bilevel equilibrium problems, inertial method, extragradient algorithm, pseudomonotone, Lipschitz-type inequality.*
Corresponding author: Kasamsuk Ungchittrakool; kasamsuku@nu.ac.th

where Ω is the solution of the equilibrium problem associated to g and C and denoted by $\mathrm{EP}(C,g)$: find $x^* \in C$ such that

$$g(x^*, y) \geq 0 \quad \text{for every} \quad y \in C.$$

The solution set of problem (1.2) is denoted by Ω^*.

The bilevel equilibrium problems were introduced by Chadli et al. [8] in 2000. This kind of problems is a generalization class of problems including, for instance, the following particular cases: optimization problems over equilibrium constraints, variational inequality over equilibrium constraints, hierarchical minimization problems, and complementarity problems. Furthermore, the particular case of the bilevel equilibrium can be applied to a real word model such as the variational inequality over the fixed point set of a firmly nonexpansive mapping applied to the power control problem of CDMA networks which were introduced by Iiduka [23]. For more details on the relation of bilevel equilibrium with its particular cases, see [11, 24, 37].

Methods for solving bilevel equilibrium problems have been studied extensively by many authors. In 2010, Moudafi [34] introduced a simple proximal method and proved the weak convergence to a solution of problem (1.2). In 2018, Yuying et al. [49] proposed a method for finding the solution for bilevel equilibrium problems where f is strongly monotone and g is pseudomonotone and Lipschitz-type continuous. They obtained the convergent sequence by combining an extragradient subgradient method with the Halpern method. For more details and most recent works on the methods for solving bilevel equilibrium problems, we refer the reader to [2, 9, 45].

On the other hand, an inertial-type algorithm was first proposed by Polyak [40] as an acceleration process in solving a smooth convex minimization problem. An inertial-type algorithm is a two-step iterative method in which the next iterate is defined by making use of the previous two iterates. It is well known that incorporating an inertial term in an algorithm accelerates the rate of convergence of the sequence generated by the algorithm. Recently, there are growing interests in inertial-type algorithm for optimization and variational inequality problems and monotone inclusions (see e.g. [5, 6, 14–16, 29] and the references therein).

Motivated by the recent interest on inertial-type algorithms and the work of Yuying et al. [49], we propose an algorithm which is a combination of extragradient algorithm and inertial extrapolation steps for solving bilevel equilibrium problems in a real Hilbert space. Under some sufficient assumptions on the bifunctions involving pseudomonotone and Lipschitz-type conditions the strong convergence theorem of the proposed algorithm is established. A clear advantage of our results over the result of Yuying et al. [49] is that our algorithm involves two inertial extrapolation terms which are not present in [49]. The presence of these inertial extrapolation terms makes our proposed iterative algorithm faster and more efficient than Yuying et al. [49], as confirmed by the given numerical example in Sect. 4.

2. Preliminaries

Throughout this paper, H is a real Hilbert space, C is a nonempty closed convex subset of H. Denote that $x_n \rightharpoonup x$ and $x_n \to x$ are the weak convergence and the strong convergence of a sequence $\{x_n\}$ to x, respectively.

We now recall the concept of proximity operator introduced by Moreau [33]. For a proper, convex and lower semicontinuous function $g : H \to (-\infty, \infty]$ and $\gamma > 0$, the

Moreau envelope of g of parameter γ is the convex function

$$^{\gamma}g(x) = \inf_{y\in H}\left\{g(y) + \frac{1}{2\gamma}\|y-x\|^2\right\} \quad \forall x \in H.$$

For all $x \in H$, the function $y \mapsto g(y) + \frac{1}{2\gamma}\|y-x\|^2$ is proper, strongly convex and lowe semicontinuous, thus the infimum is attained, i.e. $^{\gamma}g : H \to \mathbb{R}$.

The unique minimum of $y \mapsto g(y) + \frac{1}{2\gamma}\|y-x\|^2$ is called proximal point of g at x and it is denoted by $prox_g(x)$. The operator

$$\text{prox}_g(x) : H \to H$$

$$x \mapsto \arg\min_{y\in H}\left\{g(x) + \frac{1}{2\gamma}\|y-x\|^2\right\}$$

is well-defined and is said to be the proximity operator of g. When $g = i_C$ (the indicator function of the convex set C), one has

$$\text{prox}_{i_C}(x) = P_C(x),$$

for all $x \in H$.

We also recall that the subdifferential of $g : H \to (-\infty, \infty]$ at $x \in \text{dom}g$ is defined as the set of all subgradient of g at x

$$\partial g(x) := \{w \in H : g(y) - g(x) \geq \langle w, y-x\rangle \, \forall y \in H\}.$$

The function g is called subdifferentiable at x if $\partial g(x) \neq \emptyset$, g is said to be subdifferentiable on a subset $C \subset H$ if it is subdifferentiable at each point $x \in C$, and it is said to be subdifferentiable, if it is subdifferentiable at each point $x \in H$, i.e., if $\text{Dom}(\partial g) = H$.

The normal cone of C at $x \in C$ is defined by

$$N_C(x) := \{q \in H : \langle q, y-x\rangle \leq 0, \forall y \in C\}.$$

For every $x \in H$, there exists a unique element $P_C x$ defined by

$$P_C x = \text{argmin}\{\|x-y\| : y \in C\},$$

which can be found, e.g., in [7, 12].

Lemma 2.1 ([19]). *The metric projection P_C has the following basic properties:*

(i) $\|x-y\|^2 \geq \|x-P_Cx\|^2 + \|y-P_Cx\|^2$, $\forall x \in H$ *and* $y \in C$;
(ii) $\langle x-P_Cx, P_Cx-y\rangle \geq 0$, $\forall x \in H$ *and* $y \in C$;
(iii) $\|P_C(x) - P_C(y)\| \leq \|x-y\|$, $\forall x, y \in H$.

Definition 2.1 ([43, 44]). A bifunction $\psi : H \times H \to \mathbb{R}$ is called:

(i) β-strongly monotone on C if there exists $\beta > 0$ such that

$$\psi(x,y) + \psi(y,x) \leq -\beta\|x-y\|^2, \quad \forall x, y \in C;$$

(ii) monotone on C if

$$\psi(x,y) + \psi(y,x) \leq 0, \quad \forall x, y \in C;$$

(iii) pseudomonotone on C if

$$\psi(x,y) \geq 0 \Rightarrow \psi(y,x) \leq 0, \quad \forall x, y \in C;$$

(iv) β-strongly pseudomonotone on C if there exists $\beta > 0$ such that

$$\psi(x,y) \geq 0 \Rightarrow \psi(y,x) \leq -\beta\|x-y\|^2, \quad \forall x, y \in C.$$

It is easy to see from the aforementioned definitions that the following implications hold,

$$\text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \quad \text{and} \quad \text{(i)} \Rightarrow \text{(iv)} \Rightarrow \text{(iii)}.$$

The converses in general are not true.

In this paper, we consider the bifunctions f and g under the following conditions.

Condition A

(A1) $f(x, \cdot)$ is convex, weakly lower semicontinuous and subdifferentiable on H for every fixed $x \in H$.

(A2) $f(\cdot, y)$ is weakly upper semicontinuous on H for every fixed $y \in H$.

(A3) f is δ-strongly monotone on $H \times H$.

(A4) f is Lipschitz-type continuous, i.e., there are two positive constants c_1, c_2 such that

$$f(x, y) + f(y, z) \geq f(x, z) - c_1\|x - y\|^2 - c_2\|y - z\|^2, \forall x, y, z \in H.$$

Condition B

(B1) $g(x, \cdot)$ is convex, weakly lower semicontinuous and subdifferentiable on H for every fixed $x \in H$.

(B2) $g(\cdot, y)$ is weakly upper semicontinuous on H for every fixed $y \in H$.

(B3) g is pseudomonotone on C with respect to Ω, i.e.,

$$g(x, x^*) \leq 0, \quad \forall x \in C, x^* \in \Omega.$$

(B4) g is Lipschitz-type continuous, i.e., there are two positive constants L_1, L_2 such that

$$g(x, y) + g(y, z) \geq g(x, z) - L_1\|x - y\|^2 - L_2\|y - z\|^2, \forall x, y, z \in H.$$

Example 2.1 ([49]). Let $f, g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined by $f(x, y) = 5y^2 - 7x^2 + 2xy$ and $g(x, y) = 2y^2 - 7x^2 + 5xy$. It follows that f and g satisfy Condition A and Condition B, respectively.

Lemma 2.2 ([3], Propositions 3.1, 3.2). *If the bifunction g satisfies Assumptions (B1), (B2), and (B3), then the solution set Ω is closed and convex.*

Remark 2.1. Let the bifunction f satisfy Condition A and the bifunction g satisfy Condition B. If $\Omega \neq \emptyset$, then the bilevel equilibrium problem (1.2) has a unique solution, see the details in [41].

Lemma 2.3 ([10]). *Let $\phi : C \to \mathbb{R}$ be a convex, lower semicontinuous, and subdifferentiable function on C. Then x^* is a solution to the convex optimization problem*

$$\min\{f(x) : x \in C\}$$

if and only if

$$0 \in \partial\phi(x^*) + N_C(x^*).$$

The following lemmas will be used in the proof of the convergence result.

Lemma 2.4 ([31]). *Assume $\varphi_n \in [0, +\infty)$ and $\varrho_n \in [0, +\infty)$ satisfy:*

(i) $\varphi_{n+1} - \varphi_n \leq \theta_n(\varphi_n - \varphi_{n-1}) + \varrho_n$;

(ii) $\sum_{n=1}^{+\infty} \varrho_n < +\infty$;

(iii) $\{\theta_n\} \subset [0, \theta]$, *where* $\theta \in (0, 1)$.

Then the sequence $\{\varphi_n\}$ is convergent with $\sum_{n=1}^{+\infty}[\varphi_{n+1} - \varphi_n]_+ < +\infty$, where $[t]_+ = \max\{t, 0\}$ (for any $t \in \mathbb{R}$).

Lemma 2.5 ([47]). *Let $\{a_n\}$ be a sequence of non-negative real numbers satisfying*

$$a_{n+1} \leq (1-\alpha_n)a_n + \alpha_n \xi_n + \sigma_n, \quad \forall n \geq 0,$$

where $\{\alpha_n\}$ and σ_n satisfy the conditions:

(i) $\{\alpha_n\} \subset (0,1)$ *for all* $n \in \mathbb{N}$;
(ii) $\sum_{n=0}^{+\infty} \alpha_n = +\infty$;
(iii) $\limsup_{n\to\infty} \xi_n \leq 0$;
(iv) $\sum_{n=0}^{+\infty} |\sigma_n| < +\infty$.

Then $\lim_{n\to\infty} a_n = 0$.

Lemma 2.6 ([30]). *Let $\{a_n\}$ be a sequence of real numbers that does not decrease at infinity, in the sense that there exists a subsequence $\{a_{n_j}\}$ of $\{a_n\}$ such that*

$$a_{n_j} < a_{n_j+1} \quad \textit{for all} \quad j \geq 0.$$

Also consider the sequence of integers $\{\tau(n)\}_{n\geq n_0}$ defined, for all $n \geq n_0$, by

$$\tau(n) = \max\{k \leq n \mid a_k < a_{k+1}\}.$$

Then $\{\tau(n)\}_{n\geq n_0}$ is a nondecreasing sequence verifying

$$\lim_{n\to\infty} \tau(n) = \infty,$$

and, for all $n \geq n_0$, the following two estimates hold:

$$a_{\tau(n)} \leq a_{\tau(n)+1} \quad \textit{and} \quad a_n \leq a_{\tau(n)+1}.$$

3. Main result

In this section, we propose the algorithm for finding the solution of a bilevel equilibrium problem under the strong monotonicity and Lipschitztype continuous conditions on f and the pseudomonotonicity and Lipschitztype continuous conditions on g.

Algorithm 3.1. ***Initialization***: *Choose $x_0, x_1 \in H$, $c_1 < \delta$, $\theta \in [0,1)$, $\beta \in [0,1)$, the sequences $\{\beta_n\} \subset (0,1)$, $\{\epsilon_n\} \subset [0,+\infty)$, $\{\rho_n\}$ and $\{\lambda_n\}$ are such that*

$$\begin{cases} 0 < \rho_n \leq \dfrac{1}{2c_2}, \\ \beta_n < 2\rho_n(\delta - c_1), \\ \sum_{n=0}^{+\infty}(2\rho_n(\delta - L_1)(1+\beta_n) - \beta_n) = +\infty, \\ 0 < \underline{\lambda} \leq \lambda_n \leq \bar{\lambda} < \min\left(\dfrac{1}{2L_1}, \dfrac{1}{2L_2}\right), \\ \sum_{n=0}^{+\infty} \epsilon_n < +\infty. \end{cases}$$

Iterative steps: *We have $x_{n-1}, x_n \in C$, do the following Steps.*
Step 1. *Choose θ_n such that $0 \leq \theta_n \leq \bar{\theta}_n$, where*

$$\bar{\theta}_n = \begin{cases} \min\left\{\theta, \dfrac{\epsilon_n}{\max(\|x_n - x_{n-1}\|, \|x_n - x_{n-1}\|^2)}\right\} & \textit{if } x_n \neq x_{n-1}, \\ \theta & \textit{otherwise.} \end{cases}$$

Compute $s_n = x_n + \theta_n(x_n - x_{n-1})$,

$$y_n = \arg\min_{y\in C}\left\{\lambda_n g(s_n, y) + \frac{1}{2}\|y - s_n\|^2\right\},$$

*if $y_n = s_n$ go to **Step 3**. Otherwise, go to **Step 2**.*
***Step 2.** Compute*

$$z_n = \arg\min_{y\in C}\left\{\lambda_n g(y_n, y) + \frac{1}{2}\|y - s_n\|^2\right\},$$

***Step 3.** Choose β_n such that $0 \le \beta_n \le \bar{\beta}_n$, where*

$$\beta_n = \begin{cases} \min\left\{\beta, \dfrac{\epsilon_n}{\|z_n - x_n\|^2}\right\} & \text{if } z_n \neq x_n, \\ \beta & \text{otherwise.}\end{cases}$$

Compute $u_n = z_n + \beta_n(z_n - x_n)$,

$$x_{n+1} = \arg\min_{y\in C}\left\{\rho_n f(u_n, y) + \frac{1}{2}\|y - u_n\|^2\right\}. \tag{3.3}$$

Remark 3.2. We remark here that **Step 1.** and **Step 3.** in Algorithm 3.1 are easily implemented in numerical computation since the value of $\|x_n - x_{n-1}\|$ is a priori known before choosing θ_n. Similarly, the value of $\|z_n - x_n\|$ is a priori known before choosing β_n. Furthermore, observe that by the assumption that $\sum_{n=0}^{+\infty}\epsilon_n < +\infty$, we have that $\sum_{n=0}^{+\infty}\theta_n\|x_n - x_{n-1}\| < +\infty$, $\sum_{n=0}^{+\infty}\theta_n\|x_n - x_{n-1}\|^2 < +\infty$ and $\sum_{n=0}^{+\infty}\beta_n\|z_n - x_n\|^2 < +\infty$.

Lemma 3.7. *Let bifunctions f and g satisfy Condition A and Condition B, respectively. Assume that $\Omega \neq \emptyset$. Then, the sequences $\{x_n\}$, $\{z_n\}$, $\{s_n\}$ and $\{y_n\}$ genertaed by Algorithm 3.1 satisfies the following estimate*

$$\begin{aligned}\|x_{n+1} - x^*\|^2 \le{}& \alpha_n(1+\beta_n)\|s_n - x^*\|^2 - \alpha_n\beta_n\|x_n - x^*\|^2 + 2\alpha_n\beta_n\|z_n - x_n\|^2 \\ &- \alpha_n(1+\beta_n)(1-2\lambda_n L_1)\|s_n - y_n\|^2 - \alpha_n(1+\beta_n)(1-2\lambda_n L_2)\|y_n - u_n\|^2,\end{aligned} \tag{3.4}$$

where $\alpha_n = 1 - 2\rho_n(\delta - c_1)$.

Proof. Under assumptions of two bifunctions f and g, we get the unique solution of the bilevel equilibrium problem (1.2), denoted by x^*. From the definition of y_n and Lemma 2.3 imply that

$$0 \in \partial\left\{\lambda_n g(s_n, y) + \frac{1}{2}\|y - s_n\|^2\right\}(y_n) + N_C(y_n).$$

There are $w \in \partial g(s_n, \cdot)(y_n)$ and $\bar{w} \in N_C(y_n)$ such that

$$\lambda_n w + y_n - s_n + \bar{w} = 0. \tag{3.5}$$

Since $\bar{w} \in N_C(y_n)$, we have

$$\langle \bar{w}, y - y_n\rangle \le 0 \quad \text{for all} \quad y \in C. \tag{3.6}$$

By using (3.5) and (3.6), we obtain $\lambda_n\langle w, y - y_n\rangle \ge \langle s_n - y_n, y - y_n\rangle$ for all $y \in C$. Since $z_n \in C$, we have

$$\lambda_n\langle w, z_n - y_n\rangle \ge \langle s_n - y_n, z_n - y_n\rangle. \tag{3.7}$$

It follows from $w \in \partial g(s_n, \cdot)(y_n)$ that

$$g(s_n, y) - g(s_n, y_n) \ge \langle w, y - y_n\rangle \quad \text{for all} \quad y \in H. \tag{3.8}$$

By using (3.7) and (3.8), we get

$$\lambda_n\{g(s_n, z_n) - g(s_n, y_n)\} \ge \langle s_n - y_n, z_n - y_n\rangle. \tag{3.9}$$

Similarly, the definition of z_n implies that

$$0 \in \partial\left\{\lambda_n g(y_n, y) + \frac{1}{2}\|y - s_n\|^2\right\}(z_n) + N_C(z_n).$$

There are $v \in \partial g(y_n, \cdot)(z_n)$ and $\bar{v} \in N_C(x)$ such that

$$\lambda_n v + z_n - s_n + \bar{v} = 0. \tag{3.10}$$

Since $\bar{v} \in N_C(z_n)$, we have

$$\langle \bar{v}, y - z_n \rangle \leq 0 \quad \text{for all} \quad y \in C. \tag{3.11}$$

By using (3.10) and (3.11), we obtain $\lambda_n \langle v, y - z_n \rangle \geq \langle s_n - z_n, y - z_n \rangle$ for all $y \in C$. Since $x^* \in C$, we have

$$\lambda_n \langle v, x^* - z_n \rangle \geq \langle s_n - z_n, x^* - z_n \rangle. \tag{3.12}$$

It follows from $v \in \partial g(y_n, \cdot)(z_n)$ that

$$g(y_n, y) - g(y_n, z_n) \geq \langle v, y - z_n \rangle \quad \text{for all} \quad y \in H. \tag{3.13}$$

By using (3.12) and (3.13), we get

$$\lambda_n \{ g(y_n, x^*) - g(y_n, z_n) \} \geq \langle s_n - z_n, x^* - z_n \rangle.$$

Since $x^* \in \Omega$, we have $g(x^*, y_n) \geq 0$. If follows from the pseudomonotonicity of g on C with respect to Ω that $g(y_n, x^*) \leq 0$. This implies that

$$\langle s_n - z_n, z_n - x^* \rangle \geq \lambda_n g(y_n, z_n). \tag{3.14}$$

Since g is Lipschitz-type continuous, there exist two positive constants L_1, L_2 such that

$$g(y_n, z_n) \geq g(s_n, z_n) - g(s_n, y_n) - L_1 \|s_n - y_n\|^2 - L_2 \|y_n - z_n\|^2. \tag{3.15}$$

By using (3.14) and (3.15), we get

$$\langle s_n - z_n, z_n - x^* \rangle \geq \lambda_n \{ g(s_n, z_n) - g(s_n, y_n) \} - \lambda_n L_1 \|s_n - y_n\|^2 - \lambda_n L_2 \|y_n - z_n\|^2.$$

From (3.9) and the above inequality, we obtain

$$2\langle s_n - z_n, z_n - x^* \rangle \geq 2\langle s_n - y_n, z_n - y_n \rangle - 2\lambda_n L_1 \|s_n - y_n\|^2 - 2\lambda_n L_2 \|y_n - z_n\|^2. \tag{3.16}$$

We know that

$$\begin{aligned} 2\langle s_n - z_n, z_n - x^* \rangle &= \|s_n - x^*\|^2 - \|z_n - s_n\|^2 - \|z_n - x^*\|^2 \\ 2\langle s_n - y_n, z_n - y_n \rangle &= \|s_n - y_n\|^2 + \|z_n - y_n\|^2 - \|s_n - z_n\|^2. \end{aligned}$$

From (3.16), we can conclude that

$$\|z_n - x^*\|^2 \leq \|s_n - x^*\|^2 - (1 - 2\lambda_n L_1)\|s_n - y_n\|^2 - (1 - 2\lambda_n L_2)\|y_n - z_n\|^2. \tag{3.17}$$

On the other hand, from the definitions of the proximal mapping and x_{n+1}, we can write

$$x_{n+1} = \arg\min_{y \in C} \left\{ \rho_n f(u_n, y) + \frac{1}{2}\|y - u_n\|^2 \right\} = \arg\min_{y \in C} \{ f_n(y) \}, \tag{3.18}$$

where $f_n(y) = \rho_n f(u_n, y) + \frac{1}{2}\|y - u_n\|^2$. From relation (3.18) and using Lemma 2.3, we obtain $0 \in f_n(x_{n+1}) + N_C(x_{n+1})$. Thus, there exists $w_n^* \in f_n(x_{n+1})$ such that $-w_n^* \in N_C(x_{n+1})$, i.e.,

$$\langle w_n^*, y - x_{n+1} \rangle \geq 0, \quad \text{for all} \quad y \in C. \tag{3.19}$$

Then, by the convexity of $f(u_n, \cdot)$, the function f_n is strongly convex on C with modulus 1, which implies

$$f_n(x_{n+1}) + \langle w_n, y - x_{n+1} \rangle + \frac{1}{2}\|y - x_{n+1}\|^2 \leq f_n(y), \quad \text{for all} \quad y \in C, \tag{3.20}$$

where $w_n \in \partial f_n(x_{n+1})$. Substituting $w_n = w_n^*$ and $y = x^*$ into relation (3.20) and using (3.19), we get

$$f_n(x_{n+1}) + \frac{1}{2}\|x^* - x_{n+1}\|^2 \leq f_n(x^*),$$

which together with the definition of f_n, we have

$$\|x_{n+1} - x^*\|^2 \le 2\rho_n[f(u_n, x^*) - f(u_n, x_{n+1})] + \|u_n - x^*\|^2 - \|x_{n+1} - u_n\|^2. \tag{3.21}$$

Since f is strongly monotone on C with modulus δ,

$$f(u_n, x^*) \le -f(x^*, u_n) - \delta\|u_n - x^*\|^2.$$

Substituting this inequality into (3.21), we have

$$\|x_{n+1}-x^*\|^2 \le (1-2\rho_n\delta)\|u_n-x^*\|^2+2\rho_n[-f(x^*, u_n)-f(x_{n+1}, u_n)]-\|x_{n+1}-u_n\|^2. \tag{3.22}$$

Now, applying the Lipschitz-type condition of f, we obtain.

$$\begin{aligned} -f(u_n, x_{n+1}) - f(x^*, u_n) &\le -f(x^*, x_{n+1}) + c_1\|x^* - u_n\|^2 + c_2\|u_n - x_{n+1}\|^2 \\ &= c_1\|x^* - u_n\|^2 + c_2\|u_n - x_{n+1}\|^2. \end{aligned} \tag{3.23}$$

The later inequality in (3.23) follows from $f(x^*, x_{n+1}) \ge 0$, since x^* is the solution of the bilevel equilibrium problem (1.2). Substituting into (3.22), we obtain

$$\|x_{n+1} - x^*\|^2 \le [1 - 2\rho_n(\delta - c_1)]\|u_n - x^*\|^2 - (1 - 2\rho_n c_2)\|x_{n+1} - u_n\|^2. \tag{3.24}$$

By the assumption $0 < \rho_n \le \frac{1}{2c_2}$, it follows from (3.24)

$$\|x_{n+1} - x^*\|^2 \le [1 - 2\rho_n(\delta - c_1)]\|u_n - x^*\|^2. \tag{3.25}$$

By the definition of u_n, we have

$$\begin{aligned} \|u_n - x^*\|^2 &= \|z_n + \beta_n(z_n - x_n) - x^*\|^2 \\ &= \|z_n - x^*\|^2 + 2\beta_n\langle z_n - x^*, z_n - x_n\rangle + \beta_n^2\|z_n - x_n\|^2. \end{aligned} \tag{3.26}$$

Observe that

$$2\beta_n\langle z_n - x^*, z_n - x_n\rangle = \|z_n - x^*\|^2 - \|x_n - x^*\|^2 + \|z_n - x_n\|^2. \tag{3.27}$$

Thus, from (3.26) and (3.27) and noting that $\beta_n^2 \le \beta_n$

$$\begin{aligned} \|u_n - x^*\|^2 &= \|z_n - x^*\|^2 + \beta_n(\|z_n - x^*\|^2 - \|x_n - x^*\|^2 + \|z_n - x_n\|^2) + \beta_n^2\|z_n - x_n\|^2 \\ &= \|z_n - x^*\|^2 + \beta_n(\|z_n - x^*\|^2 - \|x_n - x^*\|^2) + (\beta_n + \beta_n^2)\|z_n - x_n\|^2 \\ &\le \|z_n - x^*\|^2 + \beta_n(\|z_n - x^*\|^2 - \|x_n - x^*\|^2) + 2\beta_n\|z_n - x_n\|^2 \\ &= (1 + \beta_n)\|z_n - x^*\|^2 - \beta_n\|x_n - x^*\|^2 + 2\beta_n\|z_n - x_n\|^2. \end{aligned} \tag{3.28}$$

Hence, it follows from (3.25) and (3.28) that

$$\|x_{n+1} - x^*\|^2 \le \alpha_n(1 + \beta_n)\|z_n - x^*\|^2 - \alpha_n\beta_n\|x_n - x^*\|^2 + 2\alpha_n\beta_n\|z_n - x_n\|^2, \tag{3.29}$$

where $\alpha_n = 1 - 2\rho_n(\delta - c_1)$. Combining (3.29) with (3.17), we obtain

$$\begin{aligned} \|x_{n+1} - x^*\|^2 \le{}& \alpha_n(1 + \beta_n)\|s_n - x^*\|^2 - \alpha_n\beta_n\|x_n - x^*\|^2 + 2\alpha_n\beta_n\|z_n - x_n\|^2 \\ &- \alpha_n(1 + \beta_n)(1 - 2\lambda_n L_1)\|s_n - y_n\|^2 - \alpha_n(1 + \beta_n)(1 - 2\lambda_n L_2)\|y_n - u_n\|^2, \end{aligned}$$

where $\alpha_n = 1 - 2\rho_n(\delta - c_1)$. This yields the desired conclusion. □

Theorem 3.2. *Let bifunctions f and g satisfy Condition A and Condition B, respectively. Assume that $\Omega \neq \emptyset$. Then the sequence $\{x_n\}$ generated by Algorithm 3.1 converges strongly to the unique solution of the bilevel equilibrium problem* (1.2).

Proof. We first show that $\{x_n\}$, $\{s_n\}$, $\{y_n\}$ and $\{z_n\}$ are bounded. Since $0 < \lambda_n < a$, where $a = \min\left\{\frac{1}{2L_1}, \frac{1}{2L_2}\right\}$, we have

$$(1 - 2\lambda_n L_1) > 0 \quad \text{and} \quad (1 - 2\lambda_n L_2) > 0.$$

It follows from Lemma 3.7 and the above inequalities that

$$\|x_{n+1} - x^*\|^2 \leq \alpha_n(1 + \beta_n)\|s_n - x^*\|^2 + 2\alpha_n\beta_n\|z_n - x_n\|^2 \quad \text{for all} \quad n \in \mathbb{N}.$$

Since $\alpha_n(1 + \beta_n) \leq 1$ and $\alpha_n < 1$, we obtain that

$$\|x_{n+1} - x^*\|^2 \leq \|s_n - x^*\|^2 + 2\beta_n\|z_n - x_n\|^2 \quad \text{for all} \quad n \in \mathbb{N}. \tag{3.30}$$

By the definition of s_n, we have

$$\begin{aligned}\|s_n - x^*\|^2 &= \|x_n + \theta_n(x_n - x_{n-1}) - x^*\|^2 \\ &= \|x_n - x^*\|^2 + 2\theta_n\langle x_n - x^*, x_n - x_{n-1}\rangle + \theta_n^2\|x_n - x_{n-1}\|^2. \qquad (3.31)\end{aligned}$$

Observe that

$$2\theta_n\langle x_n - x^*, x_n - x_{n-1}\rangle = \|x_n - x^*\|^2 - \|x_{n-1} - x^*\|^2 + \|x_n - x_{n-1}\|^2. \tag{3.32}$$

Thus, from (3.31) and (3.32) and noting that $\theta_n^2 \leq \theta_n$

$$\begin{aligned}\|s_n - x^*\|^2 &= \|x_n - x^*\|^2 + \theta_n(\|x_n - x^*\|^2 - \|x_{n-1} - x^*\|^2 + \|x_n - x_{n-1}\|^2) + \theta_n^2\|x_n - x_{n-1}\|^2 \\ &= \|x_n - x^*\|^2 + \theta_n(\|x_n - x^*\|^2 - \|x_{n-1} - x^*\|^2) + (\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 \\ &\leq \|x_n - x^*\|^2 + \theta_n(\|x_n - x^*\|^2 - \|x_{n-1} - x^*\|^2) + 2\theta_n\|x_n - x_{n-1}\|^2. \qquad (3.33)\end{aligned}$$

Hence, it follows from (3.30) and (3.32) that

$$\begin{aligned}\|x_{n+1} - x^*\|^2 \leq\ &\|x_n - x^*\|^2 + \theta_n(\|x_n - x^*\|^2 - \|x_{n-1} - x^*\|^2) + 2\theta_n\|x_n - x_{n-1}\|^2 \\ &+ 2\beta_n\|z_n - x_n\|^2.\end{aligned}$$

Now, since $\sum_{n=0}^{+\infty} \theta_n\|x_n - x_{n-1}\|^2 < +\infty$ and $\sum_{n=0}^{+\infty} \beta_n\|z_n - x_n\|^2 < +\infty$, we deduce from Lemma 2.4 that the sequence $\{\|x_n - x^*\|\}$ is convergent. Thus, $\{x_n\}$ is bounded and $\sum_{n=0}^{+\infty}[\|x_{n+1} - x^*\|^2 - \|x_n - x^*\|^2]_+ < +\infty$. This implies that $\{s_n\}$ is also bounded. By using (3.17), we have $\{z_n\}$ is bounded.

Next, we will show that the sequence $\{x_n\}$ converges strongly to x^*. Let us consider two cases. Case 1: Assume that there exists $n_0 \geq 0$ such that for each $n \geq n_0$, $\{\|x_n - x^*\|\} \geq \{\|x_{n+1} - x^*\|\}$. In this case $\lim_{n\to\infty} \|x_n - x^*\|$ exists and

$$\lim_{n\to\infty}(\|x_{n+1} - x^*\| - \|x_n - x^*\|) = 0. \tag{3.34}$$

It follows from (3.4)

$$\begin{aligned}0 &\leq \alpha_n(1 + \beta_n)(1 - 2\lambda_n L_1)\|s_n - y_n\|^2 + \alpha_n(1 + \beta_n)(1 - 2\lambda_n L_2)\|y_n - z_n\|^2 \\ &\leq \alpha_n(1 + \beta_n)\|s_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 - \alpha_n\beta_n\|x_n - x^*\|^2 + 2\alpha_n\beta_n\|z_n - x_n\|^2.\end{aligned}$$

Since $\alpha_n(1 + \beta_n) \in (0, 1)$, $\alpha_n\beta_n > 0$ and $\alpha_n < 1$, from the above inequality, we get

$$\begin{aligned}0 &\leq \alpha_n(1 + \beta_n)(1 - 2\lambda_n L_1)\|s_n - y_n\|^2 + \alpha_n(1 + \beta_n)(1 - 2\lambda_n L_2)\|y_n - z_n\|^2 \\ &\leq \|s_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 + 2\beta_n\|z_n - x_n\|^2. \qquad (3.35)\end{aligned}$$

Besides, we obtain

$$\begin{aligned}\|s_n - x^*\|^2 &= \|x_n + \theta_n(x_n - x_{n-1}) - x^*\|^2 \\ &\leq (\|x_n - x^*\| + \theta_n\|x_n - x_{n-1}\|)^2 \\ &= \|x_n - x^*\|^2 + 2\theta_n\|x_n - x^*\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2\end{aligned}$$

$$\leq \|x_n - x^*\|^2 + 2\theta_n\|x_n - x^*\|\|x_n - x_{n-1}\| + \theta_n\|x_n - x_{n-1}\|^2$$
$$\leq \|x_n - x^*\|^2 + 3M_1\theta_n\|x_n - x_{n-1}\|, \tag{3.36}$$

where $M_1 = \sup_{n\in\mathbb{N}}\{\|x_n - x^*\|, \|x_n - x_{n-1}\|\}$. Thus, from (3.35) and (3.36), we have

$$\begin{aligned} 0 &\leq \|s_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 + 2\beta_n\|z_n - x_n\|^2 \\ &\leq \|x_n - x^*\|^2 + 3M_1\theta_n\|x_n - x_{n-1}\| - \|x_{n+1} - x^*\| + 2\beta_n\|z_n - x_n\|^2. \end{aligned}$$

Since $\sum_{n=0}^{+\infty}\theta_n\|x_n - x_{n-1}\| < +\infty$ and $\sum_{n=0}^{+\infty}\beta_n\|z_n - x_n\|^2 < +\infty$ together with (3.34), we get that $\lim_{n\to\infty}(\|s_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 + 2\beta_n\|z_n - x_n\|^2) = 0$. It follows from (3.35) that $\lim_{n\to\infty}\|s_n - y_n\| = 0$ and $\lim_{n\to\infty}\|y_n - z_n\| = 0$. From the definition of s_n, we have

$$\begin{aligned} \lim_{n\to\infty}\|x_n - s_n\| &= \lim_{n\to\infty}\|x_n - x_n - \theta_n(x_n - x_{n-1})\| \\ &= \lim_{n\to\infty}\theta_n\|x_n - x_{n-1}\| = 0. \end{aligned} \tag{3.37}$$

We can write inequality (3.4) in the following form

$$\begin{aligned} \|x_{n+1} - x^*\| &\leq \alpha_n(1+\beta_n)\|s_n - x^*\|^2 - \alpha_n(1+\beta_n)(1-2\lambda_nL_1)\|s_n - y_n\|^2 \\ &\quad - \alpha_n(1+\beta_n)(1-2\lambda_nL_2)\|y_n - z_n\|^2 + 2\beta_n\|z_n - x_n\|^2 \\ &\leq \alpha_n(1+\beta_n)\left[\|x_n - x^*\|^2 + 3M_1\theta_n\|x_n - x_{n-1}\|\right] \\ &\quad - \alpha_n(1+\beta_n)(1-2\lambda_nL_1)\|s_n - y_n\|^2 - \alpha_n(1+\beta_n)(1-2\lambda_nL_2)\|y_n - z_n\|^2 \\ &\quad + 2\beta_n\|z_n - x_n\|^2 \\ &\leq \alpha_n(1+\beta_n)\|x_n - x^*\|^2 + 3M_1\theta_n\|x_n - x_{n-1}\| - \alpha_n(1+\beta_n)(1-2\lambda_nL_1)\times \\ &\qquad \|s_n - y_n\|^2 - \alpha_n(1+\beta_n)(1-2\lambda_nL_2)\|y_n - z_n\|^2 + 2\beta_n\|z_n - x_n\|^2 \\ &\leq (1-\gamma_n)(1+\beta_n)\|x_n - x^*\|^2 - (1-\gamma_n)(1+\beta_n)[(1-2\lambda_nL_1)\|s_n - y_n\|^2 \\ &\quad + (1-2\lambda_nL_2)\|y_n - z_n\|^2] + 3M_1\theta_n\|x_n - x_{n-1}\| + 2\beta_n\|z_n - x_n\|^2 \\ &\leq (1-(\gamma_n + \gamma_n\beta_n - \beta_n))\|x_n - x^*\|^2 + (\gamma_n + \gamma_n\beta_n - \beta_n)[(1-2\lambda_nL_1)\times \\ &\quad \|s_n - y_n\|^2] + (1-2\lambda_nL_2)\|y_n - z_n\|^2] + 3M_1\theta_n\|x_n - x_{n-1}\| \\ &\quad + 2\beta_n\|z_n - x_n\|^2, \end{aligned} \tag{3.38}$$

where $\gamma_n := 2\rho_n(\delta - c_1)$. Since $\beta_n < \gamma_n$ and $0 < \gamma_n < 1$, so we have that $\beta_n < \gamma_n(\beta_n + 1) < \beta_n + 1$ this means that $\gamma_n + \gamma_n\beta_n - \beta_n \in (0,1)$ and $\sum_{n=0}^{+\infty}(\gamma_n + \gamma_n\beta_n - \beta_n) = +\infty$, we can apply Lemma 2.5, we can conclude that $\lim_{n\to\infty}\|x_n - x^*\| = 0$.

Case 2. Assume that there exists a subsequence $\{x_{n_j}\}$ of $\{x_n\}$ such that $\|x_{n_j} - x^*\| \leq \|x_{n_j+1} - x^*\|$ for all $j \in \mathbb{N}$. By Lemma 2.6, there exists a nondecreasing sequence $\{\tau(n)\}$ of $\mathbb{N}$ such that $\lim_{n\to\infty}\tau(n) = \infty$ and for each sufficiently large $n \in \mathbb{N}$, we have

$$\|x_{\tau(n)} - x^*\| \leq \|x_{\tau(n)+1} - x^*\| \quad \text{and} \quad \|x_n - x^*\| \leq \|x_{\tau(n)+1} - x^*\|.$$

Similarly from the Case 1, we obtain that

$$\begin{aligned} 0 &\leq \|s_{\tau(n)} - x^*\|^2 - \|x_{\tau(n)+1} - x^*\|^2 + 2\beta_{\tau(n)}\|z_{\tau(n)} - x_{\tau(n)}\|^2 \\ &\leq \|x_{\tau(n)} - x^*\|^2 + 3M_1\theta_n\|x_{\tau(n)} - x_{\tau(n)-1}\| - \|x_{\tau(n)+1} - x^*\| \\ &\quad + 2\beta_{\tau(n)}\|z_{\tau(n)} - x_{\tau(n)}\|^2. \end{aligned} \tag{3.39}$$

Since $\|x_{\tau(n)} - x^*\| \leq \|x_{\tau(n)+1} - x^*\|$, we get

$$\begin{aligned} 0 &= \|s_{\tau(n)} - x^*\|^2 - \|x_{\tau(n)+1} - x^*\|^2 + 2\beta_{\tau(n)}\|z_{\tau(n)} - x_{\tau(n)}\|^2 \\ &\leq \|x_{\tau(n)+1} - x^*\|^2 + 3M_1\theta_n\|x_{\tau(n)} - x_{\tau(n)-1}\| - \|x_{\tau(n)+1} - x^*\| \end{aligned}$$

$$+ 2\beta_{\tau(n)}\|z_{\tau(n)} - x_{\tau(n)}\|^2. \tag{3.40}$$

Since $\sum_{n=0}^{+\infty} \theta_{\tau(n)}\|x_{\tau(n)} - x_{\tau(n)-1}\| < +\infty$ and $\sum_{n=0}^{+\infty} \beta_{\tau(n)}\|z_{\tau(n)} - x_{\tau(n)}\|^2 < +\infty$, we get that $\lim_{n\to\infty}(\|s_{\tau(n)} - x^*\|^2 - \|x_{\tau(n)+1} - x^*\|^2 + 2\beta_{\tau(n)}\|z_{\tau(n)} - x_{\tau(n)}\|^2) = 0$. Thus, we obtain that $\lim_{n\to\infty}\|s_{\tau(n)} - y_{\tau(n)}\| = 0$, $\lim_{n\to\infty}\|y_{\tau(n)} - z_{\tau(n)}\| = 0$ and $\lim_{n\to\infty}\|x_{\tau(n)} - s_{\tau(n)}\| = 0$. We can write inequality (3.4) in the following form

$$\begin{aligned}
\|x_{\tau(n)+1} - x^*\| \leq\ & \alpha_{\tau(n)}(1+\beta_{\tau(n)})\|s_{\tau(n)} - x^*\|^2 - \alpha_{\tau(n)}(1+\beta_{\tau(n)})(1-2\lambda_{\tau(n)}L_1)\times \\
& \quad \|s_{\tau(n)} - y_{\tau(n)}\|^2 - \alpha_{\tau(n)}(1+\beta_{\tau(n)})(1-2\lambda_{\tau(n)}L_2)\|y_{\tau(n)} - z_{\tau(n)}\|^2 \\
& + 2\beta_{\tau(n)}\|z_{\tau(n)} - x_{\tau(n)}\|^2 \\
\leq\ & \alpha_{\tau(n)}(1+\beta_{\tau(n)})\left[\|x_{\tau(n)} - x^*\|^2 + 3M_1\theta_{\tau(n)}\|x_{\tau(n)} - x_{\tau(n)-1}\|\right] \\
& - \alpha_{\tau(n)}(1+\beta_{\tau(n)})(1-2\lambda_{\tau(n)}L_1)\|s_{\tau(n)} - y_{\tau(n)}\|^2 \\
& - \alpha_{\tau(n)}(1+\beta_{\tau(n)})(1-2\lambda_{\tau(n)}L_2)\|y_{\tau(n)} - z_{\tau(n)}\|^2 + 2\beta_{\tau(n)}\|z_{\tau(n)} - x_{\tau(n)}\|^2 \\
\leq\ & \alpha_{\tau(n)}(1+\beta_{\tau(n)})\|x_{\tau(n)} - x^*\|^2 + 3M_1\theta_{\tau(n)}\|x_{\tau(n)} - x_{\tau(n)-1}\| \\
& - \alpha_{\tau(n)}(1+\beta_{\tau(n)})(1-2\lambda_{\tau(n)}L_1)\|s_{\tau(n)} - y_{\tau(n)}\|^2 \\
& - \alpha_{\tau(n)}(1+\beta_{\tau(n)})(1-2\lambda_{\tau(n)}L_2)\|y_{\tau(n)} - z_{\tau(n)}\|^2 + 2\beta_{\tau(n)}\|z_{\tau(n)} - x_{\tau(n)}\|^2 \\
\leq\ & (1-\gamma_{\tau(n)})(1+\beta_{\tau(n)})\|x_{\tau(n)} - x^*\|^2 - (1-\gamma_{\tau(n)})(1+\beta_{\tau(n)})\times \\
& [(1-2\lambda_{\tau(n)}L_1)\|s_{\tau(n)} - y_{\tau(n)}\|^2 + (1-2\lambda_{\tau(n)}L_2)\|y_{\tau(n)} - z_{\tau(n)}\|^2] \\
& + 3M_1\theta_{\tau(n)}\|x_{\tau(n)} - x_{\tau(n)-1}\| + 2\beta_{\tau(n)}\|z_{\tau(n)} - x_{\tau(n)}\|^2 \\
\leq\ & (1-(\gamma_{\tau(n)} + \gamma_{\tau(n)}\beta_{\tau(n)} - \beta_{\tau(n)}))\|x_{\tau(n)} - x^*\|^2 \\
& + (\gamma_{\tau(n)} + \gamma_{\tau(n)}\beta_{\tau(n)} - \beta_{\tau(n)})[(1-2\lambda_{\tau(n)}L_1)\|s_{\tau(n)} - y_{\tau(n)}\|^2 \\
& + (1-2\lambda_{\tau(n)}L_2)\|y_{\tau(n)} - z_{\tau(n)}\|^2] + 3M_1\theta_{\tau(n)}\|x_{\tau(n)} - x_{\tau(n)-1}\| \\
& + 2\beta_{\tau(n)}\|z_{\tau(n)} - x_{\tau(n)}\|^2.
\end{aligned} \tag{3.41}$$

Where $\gamma_{\tau(n)} := 2\rho_{\tau(n)}(\delta - c_1)$. Since $\beta_{\tau(n)} < \gamma_{\tau(n)}$ and $0 < \gamma_{\tau(n)} < 1$, so we have that $\beta_{\tau(n)} < \gamma_{\tau(n)}(\beta_{\tau(n)} + 1) < \beta_{\tau(n)} + 1$ this means that $\gamma_{\tau(n)} + \gamma_{\tau(n)}\beta_{\tau(n)} - \beta_{\tau(n)} \in (0,1)$ and $\sum_{\tau(n)=0}^{+\infty}(\gamma_{\tau(n)} + \gamma_{\tau(n)}\beta_{\tau(n)} - \beta_{\tau(n)}) = +\infty$. By using Lemma 2.5, we can conclude that $\lim_{n\to\infty}\|x_{\tau(n)} - x^*\| = 0$. By above inequality, we have $\lim_{n\to\infty}\|x_{\tau(n)+1} - x^*\| = 0$. By using $\|x_n - x^*\| \leq \|x_{\tau(n)+1} - x^*\|$, we get that $\lim_{n\to\infty}\|x_n - x^*\| = 0$. Hence $x_n \to x^*$ as $n \to \infty$. This completes the proof. □

4. Numerical experiment

In this section, we provide a numerical example to test our algorithm. All Matlab codes were performed on a computer with CPU Intel Core i7-7500U, up to 3.5GHz, 4GB of RAM under version MATLAB R2015b. In the following example, we use the standard Euclidean norm and inner product.

Example 4.2. We compare our algorithm with Algorithm 1 proposed in Yuying et al. [49]. Let us consider a problem when $H = \mathbb{R}^n$ and $C = \{x \in \mathbb{R}^n : -5 \leq x_i \leq 5, \forall i \in \{1, 2, ..., n\}\}$. Let the bifunction $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be defined by

$$g(x,y) = \langle Px + Qy, y - x\rangle \quad \text{for all} \quad x, y \in \mathbb{R}^n,$$

where matrices P and Q are generated randomly such that Q is symmetric positive semidefinite and $Q - P$ is negative semidefinite. Then g is pseudomonotone on $\mathbb{R}^n$. In fact, let $g(x,y) \geq 0$ for every $x, y \in \mathbb{R}^n$, we have

$$g(y,x) \leq g(x,y) + g(y,x) = \langle Px + Qy, y - x\rangle + \langle Py + Qx, x - y\rangle$$

$$= \langle (Q-P)(x-y), x-y\rangle \leq 0.$$

Next, we obtain that g is Lipschitz-type continuous with $L_1 = L_2 = \frac{1}{2}\|P-Q\|$. Indeed, for each $x, y, z \in \mathbb{R}^n$,

$$\begin{aligned}
&g(x,y) + g(y,z) - g(x,z)\\
&= \langle Px + Qy, y-x\rangle + \langle Py + Qz, z-y\rangle + \langle Px + Qz, z-x\rangle\\
&= \langle Px, y-x-(z-x)\rangle + \langle Qz, z-y-(z-x)\rangle + \langle Qy, y-x\rangle + \langle Py, z-y\rangle\\
&= \langle Px, y-z)\rangle + \langle Qz, x-y\rangle + \langle Qy, y-x\rangle + \langle Py, z-y\rangle\\
&= \langle P(y-x), z-y)\rangle + \langle Q(z-y), x-y\rangle\\
&= \langle P(y-x), z-y)\rangle + \langle Q^T(x-y), z-y\rangle\\
&= \langle P(y-x), z-y)\rangle + \langle Q(x-y), z-y\rangle \quad \text{since} \quad Q = Q^T\\
&= \langle (P-Q)(y-x), z-y)\rangle\\
&\geq -2\frac{\|P-Q\|}{2}\|x-y\|\|y-z\|\\
&\geq -\frac{\|P-Q\|}{2}\|y-x\|^2 - \frac{\|P-Q\|}{2}\|z-y\|^2,
\end{aligned}$$

where $\|P-Q\|$ is the spectral norm of the matrix $\|P-Q\|$, that is, the square root of the largest eigenvalue of the positive semidefinite matrix $(P-Q)^T(P-Q)$. Furthermore, we define the bifunction $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ as

$$f(x,y) = \langle Ax + By, y-x\rangle \quad \text{for all} \quad x, y \in \mathbb{R}^n,$$

with A and B being positive definite matices defined by

$$B = N^TN + nI_n \quad \text{and} \quad A = B + M^TM + nI_n, \tag{4.42}$$

where M, N are randomly $n \times n$ matrices and I_n is the identity matrix. Then we have f is n-strongly monotone on $\mathbb{R}^n$. Indeed, let $x, y \in \mathbb{R}^n$, we get

$$\begin{aligned}
f(x,y) + f(y,x) &= \langle Ax + By, y-x\rangle + \langle Ay + Bx, x-y\rangle\\
&= -\langle (A-B)(x-y), x-y\rangle\\
&= -\langle M^TM + nI_n(x+y), x-y\rangle\\
&= -\langle M^TM(x+y), x-y\rangle - \langle nI_n(x+y), x-y\rangle\\
&= -\|M(x-y)\|^2 - n\|x-y\|^2\\
&\leq -n\|x-y\|^2.
\end{aligned}$$

Thus, we take $\delta = n$. Next, we obtain that f is Lipschitz-type continuous with $c_1 = n-1$ and $c_2 = \frac{\|A-B\|^2}{4(n-1)}$. Indeed, let $x, y, z \in \mathbb{R}^n$, we have

$$\begin{aligned}
&f(x,y) + f(y,z) - f(x,z)\\
&= \langle Ax + By, y-x\rangle + \langle Ay + Bz, z-y\rangle + \langle Ax + Bz, z-x\rangle\\
&= \langle Ax, y-x-(z-x)\rangle + \langle Bz, z-y-(z-x)\rangle + \langle By, y-x\rangle + \langle Ay, z-y\rangle\\
&= \langle Ax, y-z)\rangle + \langle Bz, x-y\rangle + \langle By, y-x\rangle + \langle Ay, z-y\rangle\\
&= \langle A(y-x), z-y)\rangle + \langle B(z-y), x-y\rangle\\
&= \langle A(y-x), z-y)\rangle + \langle B^T(x-y), z-y\rangle\\
&= \langle A(y-x), z-y)\rangle + \langle B(x-y), z-y\rangle \quad \text{since} \quad B = B^T\\
&= \langle (A-B)(y-x), z-y)\rangle,
\end{aligned}$$

from which it follows that

$$\begin{aligned} f(x,y)+f(y,z)-f(x,z) &= \langle (A-B)(y-x), z-y) \rangle \\ &\geq -2\left(\sqrt{n-1}\|y-x\|\right)\left(\frac{\|A-B\|}{2\sqrt{n-1}}\|z-y\|\right) \\ &\geq -(n-1)\|y-x\|^2 - \frac{\|A-B\|^2}{4(n-1)}\|z-y\|^2. \end{aligned}$$

Moreover, $\partial f(x,\cdot)(x) = \{(A+B)x\}$ and $\|(A+B)x-(A+B)y\| \leq \|A+B\|\|x-y\|$ for all $x, y \in \mathbb{R}^n$. Thus the mapping $x \to \partial f(x,\cdot)(x)$ is bounded and $\|A+B\|$-Lipschitz continuous on every bounded subset of H.

It is easy to see that all the conditions of Theorem 3.2 and of Theorem 3.1 in [49] are satisfied. Now, we compare the performance of our algorithm and algorithm of Yuying et al. [49], we take $\lambda_k = \frac{1}{k+5}$, $\alpha_k = \frac{1}{k+4}$, $\eta_k = \frac{k+1}{3(k+4)}$, $\mu = \frac{2}{\|A+B\|^2}$ for the algorithm of Yuying et al. [49]. We take $\rho_k = \frac{1.6(n-1)}{\|A-B\|^2}$ for our algorithm, take the starting points $x_0 = x_1 \in \{x \in \mathbb{R}^n : x_i = 3, \forall i \in \{1,2,...,n\}\}$ for both algorithms. For Algorithm 3.1, we choose $\epsilon_k = \frac{1}{k^{1.1}}$, $\theta \in [0,1)$ and θ_k such that $0 \leq \theta_k \leq \bar{\theta_k}$, $\beta \in [0,1)$ and β_k such that $0 \leq \beta_k \leq \bar{\beta_k}$, where

$$\theta_k = \begin{cases} \min\left\{\theta, \dfrac{1}{k^{1.1}\max(\|x_k-x_{k-1}\|, \|x_k-x_{k-1}\|^2)}\right\} & \text{if } x_k \neq x_{k-1}, \\ \theta & \text{otherwise.} \end{cases}$$

and

$$\beta_k = \begin{cases} \min\left\{\beta, \dfrac{1}{k^{1.1}\|z_k-x_k\|^2}\right\} & \text{if } z_k \neq x_k, \\ \beta & \text{otherwise.} \end{cases}$$

To terminate the algorithm, we used the stopping criteria $\|x_{k+1}-x_k\| < \varepsilon$ with $\varepsilon = 10^{-3}$ is a tolerance. From the result reported in the Table 1, we can see that the number of iterations (No. of Iter.) by Algorithm 3.1 with parameters $\theta = 0.7$ and $\beta = 0.1$ is less than that of Yuying et al. Algorithm [49]. Moreover, Figure 1, Figure 2 and Figure 3 illustrate the numerical behavior of both algorithms. In these figures, the values of the errors $\|x_{k+1}-x_k\|$ are represented by the y-axis, number of iterations are represented by the x-axis.

TABLE 1. Comparison: proposed Algorithm 3.1 and Yuying et al. [49] with $x_0 = x_1 \in \{x \in \mathbb{R}^n : x_i = 3, \forall i = 1,2,...,n\}$, $\theta = 0.7, \beta = 0.1$.

n	Algorithm 3.1		Yuying et al. Algorithm	
	No. of Iter.	CPU (Time)	No. of Iter.	CPU (Time)
5	20	0.1845	29	0.2069
10	33	0.3353	161	1.2337
50	42	0.6066	309	3.1672

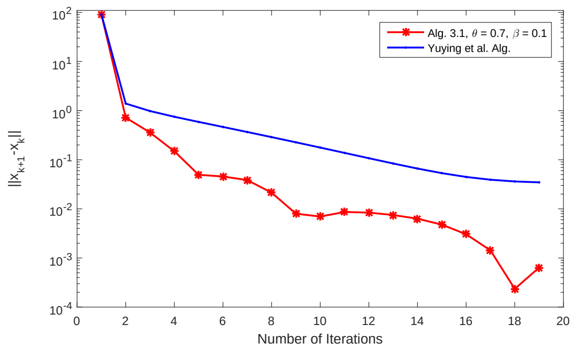

FIGURE 1. Comparison of proposed Algorithm 3.1 and Yuying et al. [49] with $x_0 = (3, 3, 3, 3, 3)^T$

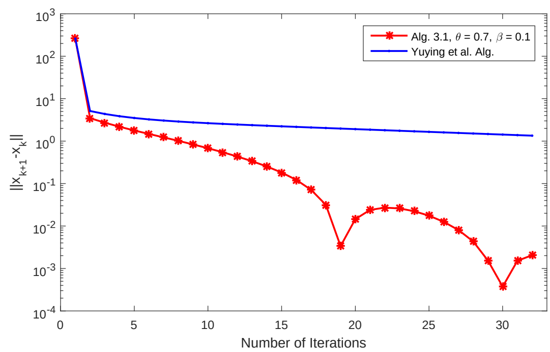

FIGURE 2. Comparison of proposed Algorithm 3.1 and Yuying et al. [49] with $x_0 = (3, 3, 3, 3, 3, 3, 3, 3, 3, 3)^T$

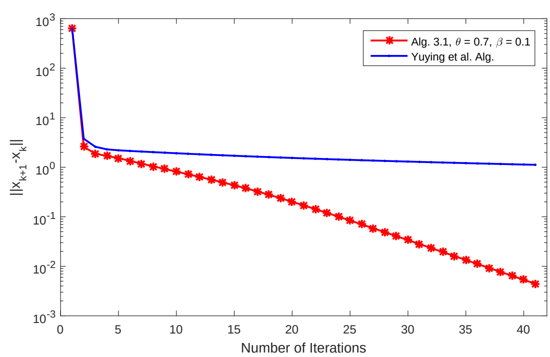

FIGURE 3. Comparison of proposed Algorithm 3.1 and Yuying et al. [49] with $x_0 = (3, 3, ..., 3)^T$ where $n = 50$

5. CONCLUSIONS

The paper has proposed an algorithms with two inertial term extrapolation steps for solving bilevel equilibrium problem in a real Hilbert space. The algorithm is a combination of the extragradient technique and inertial effects. Under some sufficient assumptions on the bifunctions involving pseudomonotone and Lipschitz-type conditions, we obtain the strong convergence of the iterative sequence generated by the proposed algorithm. A numerical results has been reported to illustrate the computational performance of the algorithm in comparison with Yuying et al. Algorithm [49] This numerical results has also confirmed that the algorithm with the inertial effects seems to work better than without inertial effects.

Acknowledgements. The authors would like to thank Naresuan University and The Thailand Research Fund for financial support. Moreover, J. Munkong is also supported by Naresuan University and The Royal Golden Jubilee Program under Grant Ph.D/0219/2556, Thailand.

Disclosure statement The authors declare that there is no conflict of interests regarding the publication of this paper.

Funding J. Munkong is supported by the Thailand Research Fund through the Royal Golden Jubilee PhD Program under Grant PHD/0219/2556, Thailand.

References

[1] Anh, P. N., Anh, T. T. H. and Hien, N. D., *Modified basic projection methods for a class of equilibrium problems*, Numer. Algor., **79** (2018), 139–152

[2] Bento, G. C., Cruz Neto, J. X., Lopes, J. O., Jr. Soares, P. A., and Soubeyran, A., *Generalized proximal distances for bilevel equilibrium problems*, SIAM J. Optim., **26** (2016), 810–830

[3] Bianchi, M. and Schaible, S., *Generalized monotone bifunctions and equilibrium problems*, J. Optim. Theory Appl., **90** (1996), 31–43

[4] Blum, E. and Oettli, W., *From optimization and variational inequalities to equilibrium problems*, Math. Program., **63** (1994), 123–145

[5] Bot, R. I., Csetnek, E. R. and Hendrich, C., *Inertial DouglasRachford splitting for monotone inclusion problems*, Appl. Math. Comput., **256** (2015), 472–487

[6] Bot, R. I., Csetnek, E. R. and Nimana, N., *Gradient-type penalty method with inertial effects for solving constrained convex optimization problems with smooth data*, Optim. Lett., **12** (2018), 17–33

[7] Cea, J., *Optimisation: thèorie et algorithmes (Dunod, Paris, 1971)*, Polish translation, Optymalizacja, Teoria i algorytmy (PWN, Warszawa, 1976)

[8] Chadli, O., Chbani, Z. and Riahi, H., *Equilibrium problems with generalized monotone bifunctions and applications to variational inequalities*, J. Optim. Theory Appl., **105** (2000), 299–323

[9] Chbani, Z. and Riahi, H., *Weak and strong convergence of proximal penalization and proximal splitting algorithms for two-level hierarchical Ky Fan minimax inequalities*, Optimization, **64** (2015), 1285–1303

[10] Daniele, P., Giannessi, F. and Maugeri, A., *Equilibrium Problems and Variational Models*, Kluwer Academic, Norwell, 2003

[11] Dempe, S., *Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints*, Optim., **52** (2003), 333–359

[12] Deutsch, F., *Best Approximation in Inner Product Spaces*, Springer, New York, 2001

[13] Dinh, B. V. and Kim, D. S., *Extragradient algorithms for equilibrium problems and symmetric generalized hybrid mappings*, Optim. Lett., **11** (2017), 537–553

[14] Dong, Q.-L., Cho, Y. J., Zhong, L. L. and Rassias, Th. M., *Inertial projection and contraction algorithms for variational inequalities*, J. Glob. Optim., **70** (2018), 687–704

[15] Dong, Q.-L., Lu, Y.-Y. and Yang, J.-F., *The extragradient algorithm with inertial effects for solving the variational inequality*, Optimization, **65** (2016), 2217–2226

[16] Dong, Q.-L., Yuan, H. B., Cho, Y. J. and Rassias, Th. M., *Modified inertial Mann algorithm and inertial CQ-algorithm for nonexpansive mappings*, Optim. Lett., **12** (2018), 87–102

[17] Facchinei, F. and Pang, J. S., *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, Berlin, 2002

[18] Fan, K., *A minimax inequality and applications*, In Shisha, O. (ed.) Inequality, III, Academic Press, New York, 1972, 103–113

[19] Goebel, K. and Reich, S., *Uniform convexity, hyperbolic geometry, and nonexpansive mappings*, Marcel Dekker, New York, 1984

[20] Hieu, D. V., *Weak and strong convergence of subgradient extragradient methods for pseudomonotone equilibrium*, Commun. Korean Math. Soc., **31** (2016), 879–893

[21] Hieu, D. V. and Moudafi, A., *A barycentric projected-subgradient algorithm for equilibrium problems*, J. Nonlinear Var. Anal., **1** (2017), 43–59

[22] Hiriart-Urruty, J.-B. and Lemaréchal, C., *Convex Analysis and Minimization Algorithms*, Vol I, Vol II, Springer, Berlin, 1993

[23] Iiduka, H., *Fixed point optimization algorithm and its application to power control in CDMA data networks*, Math. Program., **133** (2012), 227–242

[24] Iiduka, H. and Yamada, I., *A use of conjugate gradient direction for the convex optimization problem over the fixed point set of a nonexpansive mapping*, SIAM J. Optim., **19** (2009), 1881–1893

[25] Kim, D. S. and Dinh, B. V., *Parallel extragradient algorithms for multiple set split equilibrium problems in Hilbert spaces*, Numer. Algorithms, **77** (2018), 741–761

[26] Konnov, I. V., *Application of the proximal point method to nonmonotone equilibrium problems*, J. Optim. Theory Appl., **119** (2003), 317–333

[27] Konnov, I. V., *Equilibrium models and variational inequalities*, Elsevier, Amsterdam, 2007

[28] Liu, Y., *A modified hybrid method for solving variational inequality problems in Banach spaces*, J. Nonlinear Funct. Anal., (2017), Article ID 31

[29] Lorenz, D. A. and Pock, T., *An inertial forwardbackward algorithm for monotone inclusions*, J. Math. Imaging Vis., **51** (2015), 311–325

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

[31] Mainge, P. E., *Convergence theorems for inertial KM-type algorithms*, J. Comput. Appl. Math., **219** (2008), No. 1, 223–236

[32] Martinet, B., *Régularisation d'inéquations variationelles par approximations successives*, Rev. Fr. Autom. Inform. Rech. Opér., Anal. Numér., **4** (1970), 154–159

[33] Moreau, J. J., *Fonctions convexes duales et points proximaux dans un espace hilbertien*, C. R. Acad. Sci. Paris, **255** (1920), 2897–2899

[34] Moudafi, A., *Proximal methods for a class of bilevel monotone equilibrium problems*, J. Glob. Optim., **47** (2010), 287–292

[35] Moudafi, A., *Proximal point algorithm extended to equilibrium problem*, J. Nat. Geometry, **15** (1999), 91–100

[36] Muu, L. D. and Oettli, W., *Convergence of an adaptive penalty scheme for finding constrained equilibria*, Nonlinear Anal., **18** (1992), No. 12, 1159–1166

[37] Muu, L. D. and Oettli, W., *Optimization over equilibrium sets*, Optimization, **49** (2000), 179–189

[38] Muu, L. D. and Quoc, T. D., *Regularization algorithms for solving monotone Ky Fan inequalities with application to a Nash–Cournot equilibrium model*, J. Optim. Theory Appl., **142** (2009), 185–204

[39] Nikaido, H. and Isoda, K., *Note on noncooperative convex games*, Pacific J. Math., (1955), 807–815

[40] Polyak, B. T., *Some methods of speeding up the convergence of iteration methods*, USSR Comput. Math. Math. Phys., **4** (1964), No. 5, 1–17

[41] Quy, N. V., *An algorithm for a bilevel problem with equilibrium and fixed point constraints*, Optimization, **64** (2014), 1–17

[42] Rockafellar, R. T., *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim., **14** (1976), 877–898

[43] Santos, P. S. M. and Scheimberg, S., *An inexact subgradient algorithmfor equilibrium problems*, Comput Appl Math., **30** (2011), No. 1, 91–107

[44] Tan, K. K. and Xu, H. K., *Approximating fixed points of nonexpansive mappings by the Ishikawa iteration process*, J. Math. Anal. Appl., **178** (1993), No. 2, 301–308

[45] Thuy, L. Q. and Hai, T. N., *A projected subgradient algorithm for bilevel equilibrium problems and applications*, J. Optim. Theory Appl., **175** (2017), 411–431

[46] Tran, D. Q., Muu, L. D. and Nguyen, V. H., *Extragradient algorithms extended to equilibrium problems*, Optimization, **57** (2008), No. 6, 749–776

[47] Xu, H. K., *Iterative algorithms for nonlinear operators*, J. Lond. Math. Soc., **66** (2002), 240–256

[48] Yao, Y., Petrusel, A. and Qin, X., *An improved algorithm based on Korpelevich's method for variational inequalities in Banach spaces*, J. Nonlinear Convex Anal., **19** (2018), 397–406

[49] Yuying, T., Dinh, B. V., Kim, D. S. and Plubtieng, S., *Extragradient subgradient methods for solving bilevel equilibrium problems*, J. Inequal. Appl., **327** (2018)

[1]DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCE NARESUAN UNIVERSITY
PHITSANULOK 65000, THAILAND
E-mail address: Jiraprapa56@hotmail.com
E-mail address: kasamsuku@nu.ac.th

[2]DEPARTMENT OF MATHEMATICS
LE QUY DON TECHNICAL UNIVERSITY
HANOI, VIETNAM
E-mail address: vandinhb@gmail.com

[3]RESEARCH CENTER FOR ACADEMIC EXCELLENCE IN NONLINEAR ANALYSIS AND OPTIMIZATION
FACULTY OF SCIENCE
PHITSANULOK 65000, THAILAND
E-mail address: kasamsuku@nu.ac.th