



HAL
open science

Convergence of inertial prox-penalization and inertial forward-backward algorithms for solving bilevel monotone equilibrium problems

A Balhag, Z Mazgouri, M Théra

► **To cite this version:**

A Balhag, Z Mazgouri, M Théra. Convergence of inertial prox-penalization and inertial forward-backward algorithms for solving bilevel monotone equilibrium problems. 2023. hal-04137497

HAL Id: hal-04137497

<https://hal.science/hal-04137497v1>

Preprint submitted on 22 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Convergence of inertial prox-penalization and inertial forward-backward algorithms for solving bilevel monotone equilibrium problems

A. Balhag^a, Z. Mazgouri^b and M. Théra^c

^aInstitut de Mathématiques de Bourgogne, UMR 5584 CNRS, Université Bourgogne Franche-Comté, F-2100 Dijon, France

^bApplied Mathematics Engineering Department, National School of Applied Sciences, Sidi Mohamed Ben Abdellah University, Laboratory LAMA Mathematics, 30000 Fez, Morocco
ORCID 0000-0003-0131-4741

^cXLIM UMR-CNRS 7252, Université de Limoges, Limoges, France
ORCID 0000-0001-9022-6406.

Dedicated to Professor Stephan Dempe.

ARTICLE HISTORY

Compiled June 21, 2023

ABSTRACT

The main focus of this paper is on bilevel optimization on Hilbert spaces involving two monotone equilibrium bifunctions. We present a new achievement consisting on the introduction of inertial methods for solving this type of problems. Indeed, two several inertial type methods are suggested: a proximal algorithm and a forward-backward one. Under suitable conditions and without any restrictive assumption on the trajectories, the weak and strong convergence of the sequence generated by the both iterative methods are established. Two particular cases illustrating the proposed methods are thereafter discussed with respect to hierarchical minimization problems and equilibrium problems under a saddle point constraint. Furthermore, a numerical example is given to demonstrate the implementability of our algorithm. The algorithm and its convergence results improve and develop previous results in the field.

KEYWORDS

Bilevel Equilibrium problems; Monotone bifunctions; Proximal algorithm; Weak and strong convergence; Equilibrium Fitzpatrick transform.

AMS CLASSIFICATION

90C33; 49J40; 46N10; 65K15; 65K10

Let \mathbf{K} be a nonempty closed and convex subset of a real Hilbert space \mathbb{H} whose norm and scalar product are respectively denoted by $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ and let $f : \mathbf{K} \times \mathbf{K} \rightarrow \mathbb{R}$ be a real-valued bifunction. The equilibrium problem [1] associated with the bifunction f on \mathbf{K} consists in finding $\bar{x} \in \mathbf{K}$ such that

$$f(\bar{x}, y) \geq 0, \quad \forall y \in \mathbf{K}. \quad (\text{EP})$$

Since this abstract variational formulation constitutes a convenient unified mathematical model for many problems in the life sciences, it plays a central role in applied mathematics, for instance in optimization theory, variational and hemivariational inequalities, fixed-point and saddle point problems, network equilibrium problems, Nash equilibrium and others areas, (see for instance [1,2,3,4] and the bibliography therein). One of the most popular numerical approach for solving **(EP)** is the proximal point method (**PPM**) well known in convex optimization since Martinet and Rockafellar. Let us remind that given a bifunction f , the *resolvent* of f (see, [1]) is defined for each $\lambda > 0$ by

$$J_{\lambda}^f(x) := \{z \in \mathbf{K} : f(z, y) + \frac{1}{\lambda} \langle z - x, y - z \rangle \geq 0, \forall y \in \mathbf{K}\}.$$

Using this concept of resolvent for monotone bifunctions, Moudafi [5] proposed and analyzed a proximal method for solving general equilibrium problems. This proximal method generates the next iterate x_{n+1} , for each $n \geq 0$, by solving the subproblem $x_{n+1} = J_{r_n}^f(x_n)$, i.e.,

$$f(x_{n+1}, y) + \frac{1}{r_n} \langle x_{n+1} - x_n, y - x_{n+1} \rangle \geq 0, \quad \forall y \in \mathbf{K}, \quad (1)$$

where $\{r_n\}$ is a sequence of nonnegative numbers. Supposing the monotonicity of f , Moudafi [6] proved the weak convergence of the sequence $\{x_n\}$ generated by the algorithm (1) to a solution of **(EP)**. Thereby, a great interest has been brought to the study of **(EP)** by means of splitting proximal point (or backward) methods. One can consult [7,8,9] and the references therein.

Given its growing interest in applications to different applied fields, the problem **(EP)** is currently considered as one of the important research directions in which the optimization community is interested. Indeed, the study of the existence of a solution to this problem still falls within the scope of very recent studies concerning new methods of resolution. Let us cite in this sense the article [10], in which the authors use the Ekeland variational principle, to prove, in complete metric spaces and without any assumption of convexity, the existence of solutions to equilibrium and quasi-equilibrium problems. The bibliography of this article refers to new equilibrium concepts in which quasi-monotonicity and quasi-convexity are relaxed.

Given two real-valued bifunctions f and g defined on $\mathbf{K} \times \mathbf{K}$, in this paper we will focus on the more general problem **(BEP)**: Find $\bar{x} \in \mathbf{S}_f$ such that

$$g(\bar{x}, y) \geq 0, \quad \forall y \in \mathbf{S}_f, \quad (\mathbf{BEP})$$

where \mathbf{S}_f stands for the set of constraints and is defined by

$$\mathbf{S}_f := \{u \in \mathbf{K} : f(u, y) \geq 0, \quad \forall y \in \mathbf{K}\}. \quad (2)$$

We will assume throughout that the solution set to **(BEP)**, denoted by \mathbf{S} , is nonempty.

Problem **(BEP)** was implicitly introduced by Chadli, Chbani and Riahi [3] in the setting of the so-called *viscosity principle* for equilibrium problems. This principle aims

at a good selection of the upper equilibrium among solutions to the lower level equilibrium problem. This class of hierarchical problems covers in both levels, all the cases cited previously for an equilibrium problem. Besides their unification aspect, bilevel equilibrium problems have proved over the past two decades, very good applicability in different fields covering mechanics, engineering sciences and economy, see [11] and references therein.

More attention was then given to this class of problems with regard to, either the existence of solutions via dynamic and algorithmic approaches, or of the parametric stability point of view. The interested reader can consult the following recent contributions [12,13,14,15,16,17] and the references therein. In recent years, algorithmic solving procedures have been widely studied to solve **(BEP)**. For instance, Moudafi [16] used a penalty method [3] and introduced the regularized proximal point method **(RPPM)**. This algorithm is described as follows: from a starting point $x_0 \in \mathbf{K}$, for each $n \geq 0$, the next iterate x_{n+1} is defined by the proximal iteration $x_{n+1} := J_{\lambda_n}^{f+\beta_n g}(x_n)$, i.e.,

$$f(x_{n+1}, y) + \beta_n g(x_{n+1}, y) + \frac{1}{\lambda_n} \langle x_{n+1} - x_n, y - x_{n+1} \rangle \geq 0, \quad \forall y \in K, \quad (3)$$

where $\{\beta_n\}$ and $\{\lambda_n\}$ are two sequences of nonnegative reals. More precisely, under suitable assumptions on the bifunctions f and g , Moudafi proved that the sequence $\{x_n\}$ generated by the algorithm (3) converges weakly to a solution of **(BEP)** provided that

$$\liminf_{n \rightarrow +\infty} \lambda_n > 0, \quad \sum_{n=0}^{+\infty} \lambda_n \beta_n < +\infty \quad \text{and} \quad \|x_{n+1} - x_n\| = o(\beta_n). \quad (4)$$

The drawback of the last assumption is the difficulty to choose such a control sequence (β_n) because we do not know the convergence rate of $\|x_{n+1} - x_n\|$. This led Moudafi [16] to conjecture that this limiting assumption can be removed via the introduction of a conditioning notion for equilibrium bifunctions.

Later on Chbani et al. [15] considered an alternate proximal scheme, which generates the next iterates x_{n+1} by solving the regularized problem $x_{n+1} := J_{\lambda_n}^{\beta_n f + g}(x_n)$, i.e.,

$$\beta_n f(x_{n+1}, y) + g(x_{n+1}, y) + \frac{1}{\lambda_n} \langle x_{n+1} - x_n, y - x_{n+1} \rangle, \quad \forall y \in \mathbf{K}. \quad (5)$$

In this paper, the difficulty of the method **(RPPM)** mentioned in [16] has been solved. Following [18] and under a similar geometric assumption formulated in terms of the Fenchel conjugate of the bifunction f , they analyzed the weak and strong convergence of their algorithm to a solution of **(BEP)**. More recently, in [19], the authors proposed a forward-forward algorithm and a forward-backward algorithm for solving **(BEP)** under quite mild conditions supposing that the bifunctions of the two level equilibrium problems are pseudomonotone.

As a continuation of the studies of equilibrium problems by means of proximal iterative methods, we propose an inertial proximal method for solving **(BEP)**. It is well known that the inertial proximal iteration, where the next iterate is defined by

making use of the previous two iterates, may be interpreted as a discretization of differential systems of second order in time. The presence of inertial terms improves the convergence behavior of the generated sequences. We emphasize that the origin of these methods dates back to [20] as part of the approach to a solution of an abstract inclusion of the form: find $\bar{x} \in \mathbb{H}$ such that

$$0 \in A(\bar{x}), \quad (6)$$

where $A : \mathbb{H} \rightrightarrows \mathbb{H}$ is a *maximally monotone operator*, and the solution set $A^{-1}(\{0\})$ is assumed to be nonempty. In this regard, giving two sequences of nonnegative numbers $\{\alpha_n\}$ and $\{\lambda_n\}$, the authors in [20] considered the following iterative scheme:

$$x_{n+1} - x_n - \alpha_n(x_n - x_{n-1}) + \lambda_n A(x_{n+1}) \ni 0, \quad (7)$$

and proved the weak convergence of the sequence $\{x_n\}$ generated by the above algorithm towards a solution of (6) under appropriate conditions on the parameters $\{\alpha_n\}$ and $\{\lambda_n\}$ whenever the restrictive assumption $\sum_{n=1}^{+\infty} \alpha_n \|x_n - x_{n-1}\|^2 < +\infty$ holds.

Let us consider the second-order differential system in time

$$\ddot{x}(t) + \alpha \dot{x}(t) \in \mathbf{P}_{Tu}(-\partial(g + \beta f)(x(t), \cdot)x(t)), \quad (8)$$

where $Tu = \text{cl}(\mathbb{R}_+(C - u))$ is the Bouligand tangent cone to C at $u \in C$ and $\mathbf{P}_{\mathbf{K}}$ is the orthogonal projection onto a closed convex set \mathbf{K} and defined as $\mathbf{P}_{\mathbf{K}}(x) = \inf_{u \in \mathbf{K}} \|x - u\|$.

For solving the problem (**BEP**), we propose two approximate inertial proximal schemes. Each of them is derived from a different discretization of (8). This approach is inspired from the results presented in [6] and the contributions [15,16].

- The first scheme (Algorithm (**IPA**)) is expressed as an inertial prox-penalization algorithm. It is considered in order to reach a solution to the two-level equilibrium problem (**BEP**) in its general form (i.e., the two associated bifunctions f and g are arbitrary).
- The second scheme (Algorithm (**IFBA**)) is expressed as an inertial forward-backward algorithm. It is specifically designed to handle the situation where $f(x, y) = \langle Bx, y - x \rangle$, and $B : \mathbb{H} \rightarrow \mathbb{H}$ is a single-valued and cocoercive operator.

Algorithm 1: (Inertial proximal algorithm (**IPA**)).

Initialization: Choose positive sequences $\{\beta_n\}$, $\{\lambda_n\}$, and a nonnegative real number $\alpha \in (0, \frac{1}{3})$. Take arbitrary $x_0, x_1 \in \mathbf{K}$.

Iterative step: For every $n \geq 1$ and given current iterates $x_{n-1}, x_n \in \mathbf{K}$ set $y_n := x_n + \alpha(x_n - x_{n-1})$ and define $x_{n+1} \in \mathbf{K}$ by $x_{n+1} := J_{\lambda_n}^{\beta_n f + g}(y_n)$, i.e.,

$$\beta_n f(x_{n+1}, y) + g(x_{n+1}, y) + \frac{1}{\lambda_n} \langle x_{n+1} - y_n, y - x_{n+1} \rangle \geq 0, \quad \forall y \in \mathbf{K}. \quad (9)$$

Due to the significant difficulty in computing the resolvent $J^{\beta_n f + g}$ of $\beta_n f + g$ compared to computing the resolvent of f and g separately, an alternative approach is proposed. This is the object of the algorithm **(IFBA)** presented below.

Algorithm 2: (Inertial forward-backward algorithm **(IFBA)**).

Initialization: Choose positive sequences $\{\beta_n\}$, $\{\lambda_n\}$, and a nonnegative real number $\alpha \in (0, \frac{1}{3})$. Take arbitrary $x_0, x_1 \in \mathbf{K}$.

Iterative step: For every $n \geq 1$ and given current iterates $x_{n-1}, x_n \in \mathbf{K}$ set $y_n := x_n + \alpha(x_n - x_{n-1})$ and define $x_{n+1} \in \mathbf{K}$ by $x_{n+1} := J_{\lambda_n}^g(y_n - \beta_n \lambda_n Bx_n)$, i.e.,

$$\beta_n \langle Bx_n, y - x_{n+1} \rangle + g(x_{n+1}, y) + \frac{1}{\lambda_n} \langle x_{n+1} - y_n, y - x_{n+1} \rangle \geq 0, \quad \forall y \in \mathbf{K}. \quad (10)$$

In the above algorithms, $\{\lambda_n\}$ denotes the sequence of step sizes, $\{\beta_n\}$ the sequence of penalization parameters, and $\alpha \in (0, \frac{1}{3})$ the parameter that controls the inertial terms. The proposed numerical scheme **(IPA)** recovers, when $\alpha = 0$, the algorithm investigated in [15], and if in addition $f = 0$, the one suggested in [5]. The *Fitzpatrick transform* of the bifunction f will be a key ingredient in our convergence analysis. Indeed, we provide conditions under which the sequence generated by the algorithm **(IPA)** weakly or strongly converges to a solution of **(BEP)**. More precisely, under a discrete counterpart (12) of the geometric condition used in [14] and formulated in terms of the Fitzpatrick transform of the bifunction f , we first prove that (see Theorem 2.3) the sequence generated by **(IPA)** weakly converges to a solution of **(BEP)** provided that $0 \leq \alpha < \frac{1}{3}$, $\liminf_{n \rightarrow +\infty} \lambda_n > 0$ and $\lim_{n \rightarrow +\infty} \beta_n = 0$. Afterwards, by strengthening the monotonicity assumption on the upper level bifunction g , and whenever $0 \leq \alpha < \frac{1}{3}$ and $\sum_{n=1}^{\infty} \lambda_n = +\infty$, we show (see Theorem 2.4) the strong convergence of the trajectories generated by the algorithm **(IPA)** to the unique solution of **(BEP)**. Then, we show (see Theorem 2.5) that, without the need of the geometric assumption (12), the sequence converges strongly to the unique solution of **(BEP)** when the parameters λ_n and β_n satisfy additionally the conditions: $\lim_{n \rightarrow +\infty} \lambda_n = 0$, $\lim_{n \rightarrow +\infty} \beta_n = +\infty$ and $\liminf_{n \rightarrow +\infty} \lambda_n \beta_n > 0$. In the subsequent step of the work, we state and prove (see Theorem 3.3 and Theorem 3.4) the main results concerning the weak and the strong convergence of the sequence generated by **(IFBA)**. The main advantage of our approach is that it provides convergence without any restrictive assumption on the trajectories. The convergence results of **(IPA)** can be seen as an extension to the second order counterparts of the ones given in [15,16]. To our knowledge, such inertial proximal schemes have been studied only for the first level equilibrium problem **(EP)**, see for instance [21,22] and the references therein. As applications, we discuss the hierarchical convex minimization case and equilibrium problems under a saddle point constraint. A numerical experiment is thereafter given to illustrate our theoretical results. We end up the paper by concluding comments.

1. Background material and technical lemmata

In this section, we give some preliminary results and definitions that will be used in the sequel. Throughout this paper, unless stated otherwise, let \mathbf{K} be a nonempty closed and convex subset of a real Hilbert space \mathbb{H} . We first recall some well known concepts on monotonicity and continuity of real bifunctions and operators.

Definition 1.1. A bifunction $f : \mathbf{K} \times \mathbf{K} \rightarrow \mathbb{R}$ is called:

- (i) monotone if $f(x, y) + f(y, x) \leq 0$ for all $x, y \in \mathbf{K}$;
- (ii) γ -strongly monotone, if there exists $\gamma > 0$ such that

$$f(x, y) + f(y, x) \leq -\gamma\|x - y\|^2 \text{ for all } x, y \in \mathbf{K};$$

- (iii) upper hemicontinuous, if

$$\lim_{t \searrow 0} f(tz + (1-t)x, y) \leq f(x, y) \text{ for all } x, y, z \in \mathbf{K};$$

- (iv) lower semicontinuous at y with respect to the second argument on \mathbf{K} , if

$$f(x, y) \leq \liminf_{w \rightarrow y} f(x, w) \text{ for all } x \in \mathbf{K};$$

- (v) an equilibrium bifunction, if for each $x \in \mathbf{K}$, $f(x, x) = 0$ and $f(x, \cdot)$ is convex and lower semicontinuous.

Definition 1.2. An operator $B : \mathbb{H} \rightarrow \mathbb{H}$ is said to be:

- (i) monotone if $\langle Bx - By, x - y \rangle \geq 0$ for all $x, y \in \mathbb{H}$;
- (ii) θ -inverse strongly monotone (or θ -cocoercive) if there exists $\theta > 0$ such that

$$\langle Bx - By, x - y \rangle \geq \theta\|Bx - By\|^2 \text{ for all } x, y \in \mathbb{H};$$

- (iii) monotone if for every $(x, y) \in \mathbb{H} \times \mathbb{H}, v \in Ax, v' \in Ay$ it holds

$$\langle x - y, v - v' \rangle \geq 0;$$

- (iv) maximally monotone, if it is monotone and if it has no monotone extension.

Remark 1. Let $B : \mathbb{H} \rightarrow \mathbb{H}$ be θ -cocoercive, with $\theta > 0$. Then B is maximally monotone.

The *dual equilibrium problem* associated with the bifunction f on \mathbf{K} is stated as follows: find $\bar{x} \in \mathbf{K}$ such that

$$f(y, \bar{x}) \leq 0, \quad \forall y \in \mathbf{K}. \tag{DEP}$$

The set of solutions to (DEP) is called the *Minty solution set*. The following result gives the link between Minty equilibria and the standard ones.

Lemma 1.3 (Minty's Lemma, [1]). (i) *Whenever f is monotone, every solution of (EP) is a solution of (DEP);*

(ii) Conversely, if f is upper hemicontinuous and an equilibrium bifunction, then each solution of **(DEP)** is a solution of **(EP)**.

The next lemma introduces the notion of *resolvent* associated with a bifunction. This concept is crucial in our approach for solving **(BEP)**.

Lemma 1.4. [23] Suppose that $f : \mathbf{K} \times \mathbf{K} \rightarrow \mathbb{R}$ is a monotone equilibrium bifunction. Then the following are equivalent:

- (i) f is maximal: $(x, u) \in \mathbf{K} \times \mathbb{H}$ and $f(x, y) \leq \langle u, x - y \rangle$, $\forall y \in \mathbf{K}$ imply that $f(x, y) + \langle u, x - y \rangle \geq 0 \forall y \in \mathbf{K}$;
- (ii) for each $x \in \mathbb{H}$ and $\lambda > 0$, there exists a unique $z_\lambda = J_\lambda^f(x) \in \mathbf{K}$, called the resolvent of f at x , such that

$$\lambda f(z_\lambda, y) + \langle y - z_\lambda, z_\lambda - x \rangle \geq 0, \quad \forall y \in \mathbf{K}. \quad (11)$$

Moreover, $\bar{x} \in \mathbf{S}_f$ if, and only if, $\bar{x} = J_\lambda^f(\bar{x})$ for every $\lambda > 0$ if, and only if, $\bar{x} = J_\lambda^f(\bar{x})$ for some $\lambda > 0$.

We will make use of the two following useful lemmata for establishing the main results on the convergence of the sequence generated by the algorithms (9) and (10).

Lemma 1.5 (discrete Opial's Lemma, [24]). Let C be a nonempty subset of \mathbb{H} and $\{x_k\}_{k \geq 0}$ be a sequence in \mathbb{H} such that the following two conditions hold:

- (i) For every $x \in C$, $\lim_{k \rightarrow +\infty} \|x_k - x\|$ exists;
- (ii) Every weak sequential cluster point of $\{x_k\}_{k \geq 0}$ is in C .

Then, $\{x_k\}_{k \geq 0}$ converges weakly to an element in C .

Lemma 1.6. Let $0 < p \leq 1$ and $\{b_k\}_{k \geq 0}, \{w_k\}_{k \geq 0}$ be two sequences of nonnegative numbers such that, for all $k \geq 0$,

$$b_{k+1} \leq pb_k + w_k.$$

If $\sum_{k=0}^{+\infty} w_k < +\infty$, then $\lim_{k \rightarrow +\infty} b_k$ exists. Further, if $p < 1$ then $\sum_{k=0}^{+\infty} b_k < +\infty$.

Proof. Since $b_k \geq 0$ for all k and since $\sum_{k=0}^{+\infty} w_k < +\infty$, then the sequence $\{b_{k+1} - \sum_{i=0}^k w_i\}$ is bounded from below. We also have

$$b_{k+1} - \sum_{i=0}^k w_i \leq pb_k - \sum_{i=0}^{k-1} w_i \leq b_k - \sum_{i=0}^{k-1} w_i,$$

which implies that the sequence $\{b_{k+1} - \sum_{i=0}^k w_i\}$ is nonincreasing, hence convergent. It follows that $\lim_{k \rightarrow +\infty} b_k$ exists.

Observe that

$$(1 - p)b_k \leq b_k - b_{k+1} + w_k.$$

Summing up from $k = 0$ to n , we get

$$\begin{aligned}
(1-p) \sum_{k=0}^n b_k &\leq \sum_{k=0}^n (b_k - b_{k+1}) + \sum_{k=0}^n w_k \\
&= b_0 - b_{n+1} + \sum_{k=0}^n w_k \\
&\leq b_0 + \sum_{k=0}^n w_k.
\end{aligned}$$

Since $1-p > 0$ and $\sum_{k=0}^{+\infty} w_k < +\infty$, we conclude that $\sum_{k=0}^{+\infty} b_k < +\infty$. \square

We also need the two following technical lemmata.

Lemma 1.7. [25] *For all $x, y \in \mathbb{H}$ and $\beta \in \mathbb{R}$, the following equality holds,*

$$\|\beta x + (1-\beta)y\|^2 = \beta\|x\|^2 + (1-\beta)\|y\|^2 - \beta(1-\beta)\|x-y\|^2.$$

Lemma 1.8. [15] *Let $\{a_n\}$ be a sequence of real numbers that does not decrease at infinity, in the sense that there exists a subsequence $\{a_{n_k}\}_{k \geq 0}$ of $\{a_n\}$ which satisfies*

$$a_{n_k} < a_{n_{k+1}} \quad \text{for all } k \geq 0.$$

Then, the sequence of integers $\{\sigma_n\}_{n \geq n_0}$ defined by $\sigma_n := \max\{k \leq n : a_k < a_{k+1}\}$ is a nondecreasing sequence verifying $\lim_{n \rightarrow +\infty} \sigma_n = \infty$ and, for all $n \geq n_0$

$$a_{\sigma_n} < a_{\sigma_{n+1}} \quad \text{and} \quad a_n \leq a_{\sigma_{n+1}}.$$

In the rest of this section we recall some standard definitions and tools from convex analysis. For an extended real-valued function $\varphi : \mathbb{H} \rightarrow \mathbb{R} \cup \{+\infty\}$, we denote by $\text{dom } \varphi = \{x \in \mathbb{H} : \varphi(x) < +\infty\}$ the *effective domain* of φ and we say that φ is *proper*, if $\text{dom } \varphi \neq \emptyset$. We also denote by $\min \varphi := \inf_{x \in \mathbb{H}} \varphi(x)$ the optimal value of the objective function φ and by $\text{argmin } \varphi := \{x \in \mathbb{H} : \varphi(x) = \min \varphi\}$ the set of its global minima.

By $\Gamma_0(\mathbb{H})$, we mean the set of extended real-valued proper lower semicontinuous convex function $\varphi : \mathbb{H} \rightarrow \mathbb{R} \cup \{+\infty\}$. Given $\varphi \in \Gamma_0(\mathbb{H})$ and $x \in \mathbb{H}$, we recall that the *Fenchel conjugate* of φ is the function $\varphi^* : \mathbb{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by $\varphi^*(x) := \sup_{y \in \mathbb{H}} \{\langle x, y \rangle - \varphi(y)\}$. If $\varphi = \delta_{\mathbf{K}}$ is the *indicator function* of $\mathbf{K} \subset \mathbb{H}$, i.e., when $\delta_{\mathbf{K}}(x) = 0$ if $x \in \mathbf{K}$ and $+\infty$ otherwise, we remind that $\delta_{\mathbf{K}}^*$ is the *support function* of \mathbf{K} . It is denoted by $\sigma_{\mathbf{K}}$, and it is defined as $\delta_{\mathbf{K}}^*(x^*) = \sigma_{\mathbf{K}}(x^*) = \sup_{y \in \mathbf{K}} \langle x^*, y \rangle$. The (convex)

subdifferential of φ at $x \in \mathbb{H}$ is the set

$$\partial\varphi(x) := \{v \in \mathbb{H} : \varphi(y) \geq \varphi(x) + \langle v, y - x \rangle, \forall y \in \mathbb{H}\} \quad \text{and} \quad \partial\varphi(x) := \emptyset \quad \text{if } \varphi(x) = +\infty,$$

while the *normal cone* to $\mathbf{K} \subset \mathbb{H}$ at $x \in \mathbb{H}$ is the set

$$N_{\mathbf{K}}(x) = \begin{cases} \{x^* \in \mathbb{H} : \langle x^*, u - x \rangle \leq 0, \forall u \in \mathbf{K}\} & \text{if } x \in \mathbf{K} \\ \emptyset & \text{otherwise.} \end{cases}$$

We mention that $N_{\mathbf{K}} = \partial\delta_{\mathbf{K}}$, and that $x^* \in N_{\mathbf{K}}(x)$ if, and only if, $\sigma_{\mathbf{K}}(x^*) = \langle x^*, x \rangle$. For every $u \in \mathbf{K}$, we denote by f_u the function defined on \mathbb{H} by $f_u(x) = f(u, x)$ if $x \in \mathbf{K}$ and $f_u(x) = +\infty$ otherwise. For an equilibrium bifunction $f : \mathbf{K} \times \mathbf{K} \rightarrow \mathbb{R}$, the associated operator $A^f : \mathbb{H} \rightrightarrows \mathbb{H}$ defined by

$$A^f(x) := \partial f_x(x) = \begin{cases} \{z \in \mathbb{H} : f(x, y) + \langle z, x - y \rangle \geq 0, \forall y \in \mathbf{K}\} & \text{if } x \in \mathbf{K} \\ \emptyset & \text{otherwise,} \end{cases}$$

is monotone when f is monotone and satisfies $f(x, x) = 0$. Following Alizadeh & Hadjisavvas [26,27], the *Fitzpatrick transform* $\mathcal{F}_f : \mathbf{K} \times \mathbb{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ associated with a bifunction f is defined by

$$\mathcal{F}_f(x, u) = \sup_{y \in \mathbf{K}} \{\langle u, y \rangle + f(y, x)\}.$$

Given its continuity and convexity properties, the function \mathcal{F}_f has proven to be an important tool when studying the asymptotic properties of dynamical equilibrium systems, see [14] for a detailed presentation of these elements. This section concludes with the following auxiliary result needed for establishing our results.

Proposition 1.9. [26,14]

- Suppose that $f(x, y) = \varphi(y) - \varphi(x)$ where $\varphi : \mathbb{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ is convex and lower semicontinuous with $\text{dom } \varphi \subset \mathbf{K}$. Then for every $(x, u) \in \mathbf{K} \times \mathbb{H}$, $\mathcal{F}_f(x, u) = \varphi(x) + \varphi^*(u)$.
- Let $B : \mathbb{H} \rightarrow \mathbb{H}$ be a monotone operator and $f(x, y) := \langle Bx, y - x \rangle$. Then, the Fitzpatrick function \mathcal{F}_f is exactly the Fitzpatrick function of the operator B denoted by \mathcal{F}_B which was first introduced by Fitzpatrick in [28]. In other words, for every $(x, u) \in \mathbb{H} \times \mathbb{H}$, we have:

$$\mathcal{F}_B(x, u) = \mathcal{F}_f(x, u) = \sup_{y \in \mathbb{H}} \{\langle u, y \rangle + \langle By, x - y \rangle\}.$$

Assumptions

In the remaining part of the paper, we suppose that

- ▶ f and g are two monotone and upper hemicontinuous bifunctions satisfying assumptions (iv) and (v) of Definition 1.1;
- ▶ for each $y \in \mathbf{K}$, $\partial g_y(y) \neq \emptyset$ (i.e., $\text{dom}(A^g) = \mathbf{K}$) and that $\mathbf{K} \cap \mathbf{S}_f \neq \emptyset$;
- ▶ $\mathbb{R}_+(\mathbf{K} - \mathbf{S}_f)$ is a closed linear subspace of \mathbb{H} . In this case, the operator $g_x + \delta_{\mathbf{S}_f}$ is maximally monotone, see [29,30], and the subdifferential sum formula $\partial(g_x + \delta_{\mathbf{S}_f}) = \partial g_x + N_{\mathbf{S}_f}$ holds;
- ▶ the following geometric assumption: $\forall u \in \mathbf{S}_f$, for all $p \in N_{\mathbf{S}_f}(u)$,

$$\sum_{n=1}^{+\infty} \lambda_n \beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right] < +\infty. \quad (12)$$

This assumption will be also needed and considered as a key tool in our treatment of the convergence analysis. Let us mention that hypothesis (12) is the discrete

counterpart of the condition introduced in [14] in the context of continuous-time dynamical equilibrium systems. Note also that it is a natural extension of similar assumptions known in the literature for the convergence analysis of variational inequalities expressed as monotone inclusion problems and for constrained convex optimization problems, see [18,31,32] and references therein for further useful comments on these assumptions.

2. Convergence of the inertial proximal algorithm (IPA)

Consider the sequence $\{x_n\}$ defined by the algorithm (9), where $x_{n+1} := J_{\lambda_n}^{\beta_n f+g}(y_n)$. Our objective is to analyze the convergence behavior of the sequence $\{x_n\}$.

2.1. Weak convergence analysis

In this subsection, under natural conditions, we obtain a weak convergence result for the trajectory generated by (9) to a solution of (BEP). We first prove the following preliminary estimation.

Lemma 2.1. *Let $\{x_n\}$ be a sequence generated by the algorithm (9). Take $u \in \mathbf{S}$ and set $a_n := \|x_n - u\|^2$. Then, there exists $p \in N_{\mathbf{S}_f}(u)$ such that for each $n \geq 1$ the following inequality holds:*

$$\begin{aligned} & a_{n+1} - a_n - \alpha(a_n - a_{n-1}) + \lambda_n \beta_n f(u, x_{n+1}) \\ & \leq (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 + \lambda_n \beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right]. \end{aligned} \quad (13)$$

Proof. Since $\{x_n\}$ is generated by the algorithm (9), we have for each $x \in \mathbf{K}$

$$\begin{aligned} 0 & \leq \lambda_n \beta_n f(x_{n+1}, x) + \lambda_n g(x_{n+1}, x) \\ & \quad + \frac{1}{2} (\|y_n - x\|^2 - \|x_{n+1} - x\|^2 - \|x_{n+1} - y_n\|^2). \end{aligned} \quad (14)$$

By Lemma 1.7, we have for all $n \geq 1$

$$\begin{aligned} \|y_n - x\|^2 & = \|x_n + \alpha(x_n - x_{n-1}) - x\|^2 \\ & = \|(1 + \alpha)(x_n - x) - \alpha(x_{n-1} - x)\|^2 \\ & = (1 + \alpha)\|x_n - x\|^2 - \alpha\|x_{n-1} - x\|^2 + \alpha(1 + \alpha)\|x_n - x_{n-1}\|^2. \end{aligned} \quad (15)$$

Also, we have

$$\begin{aligned} \|x_{n+1} - y_n\|^2 & = \|x_{n+1} - x_n - \alpha(x_n - x_{n-1})\|^2 \\ & = \|x_{n+1} - x_n\|^2 + \alpha^2\|x_n - x_{n-1}\|^2 - 2\alpha\langle x_{n+1} - x_n, x_n - x_{n-1} \rangle \\ & \geq (1 - \alpha)\|x_{n+1} - x_n\|^2 + (\alpha^2 - \alpha)\|x_n - x_{n-1}\|^2. \end{aligned} \quad (16)$$

Combining (15) and (16) with (14), we get for every $x \in \mathbf{K}$

$$\begin{aligned} & \|x_{n+1} - x\|^2 - (1 + \alpha)\|x_n - x\|^2 + \alpha\|x_{n-1} - x\|^2 \\ & \leq (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 + 2\lambda_n\beta_n f(x_{n+1}, x) + 2\lambda_n g(x_{n+1}, x). \end{aligned} \quad (17)$$

Since $u \in \mathbf{S}$, according to the first-order optimality condition, we have

$$0 \in \partial(g_u + \delta_{\mathbf{S}_f})(u) = A^g(u) + N_{\mathbf{S}_f}(u).$$

Let $p \in N_{\mathbf{S}_f}(u)$ be such that $-p \in A^g(u)$, we have for every $n \geq 1$

$$\lambda_n g(u, x_{n+1}) + \lambda_n \langle -p, u - x_{n+1} \rangle \geq 0, \quad (18)$$

and by taking $x = u$ and $a_n = \|x_n - u\|^2$ in (17), we also have

$$\begin{aligned} a_{n+1} - (1 + \alpha)a_n + \alpha a_{n-1} & \leq (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 \\ & \quad + 2\lambda_n\beta_n f(x_{n+1}, u) + 2\lambda_n g(x_{n+1}, u). \end{aligned} \quad (19)$$

Summing up the above inequalities and using the monotonicity of g , we get

$$\begin{aligned} a_{n+1} - a_n - \alpha(a_n - a_{n-1}) & \leq (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 \\ & \quad + 2\lambda_n\beta_n f(x_{n+1}, u) + 2\lambda_n \langle -p, u - x_{n+1} \rangle. \end{aligned}$$

Using the monotonicity of f , we obtain

$$\begin{aligned} & a_{n+1} - a_n - \alpha(a_n - a_{n-1}) + \lambda_n\beta_n f(u, x_{n+1}) \\ & \leq (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 + \lambda_n\beta_n f(x_{n+1}, u) + 2\lambda_n \langle -p, u - x_{n+1} \rangle \\ & = (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 \\ & \quad + \lambda_n\beta_n \left[\left\langle \frac{2p}{\beta_n}, x_{n+1} \right\rangle + f(x_{n+1}, u) - \left\langle \frac{2p}{\beta_n}, u \right\rangle \right]. \end{aligned}$$

Finally, using the fact that $p \in N_{\mathbf{S}_f}(u)$, i.e., $\delta_{\mathbf{S}_f}(u) + \sigma_{\mathbf{S}_f}(p) = \langle p, u \rangle$, we obtain

$$\begin{aligned} & a_{n+1} - a_n - \alpha(a_n - a_{n-1}) + \lambda_n\beta_n f(u, x_{n+1}) \\ & \leq (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 \\ & \quad + \lambda_n\beta_n \left[\sup_{x \in \mathbf{K}} \left\{ \left\langle \frac{2p}{\beta_n}, x \right\rangle + f(x, u) \right\} - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right] \\ & = (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 \\ & \quad + \lambda_n\beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right]. \end{aligned}$$

The proof is complete. \square

Remark 2. We can continue our analysis assuming that $\sum_{n=1}^{+\infty} \|x_n - x_{n-1}\|^2 < +\infty$; however this condition involves the trajectory $\{x_n\}$ which is unknown. In the next corollary we prove that the above condition holds under a suitable control of the parameter α .

Corollary 2.2. *Under hypothesis (12) and by assuming that $0 \leq \alpha < \frac{1}{3}$, we have*

- (i) $\sum_{n=1}^{+\infty} \|x_n - x_{n-1}\|^2 < +\infty;$
(ii) $\sum_{n=1}^{+\infty} \lambda_n \beta_n f(u, x_{n+1}) < +\infty,$ for each $u \in \mathbf{S}.$

Proof. (i) First we simplify the writing of the estimation (13) given in Lemma 2.1. Since $u \in \mathbf{S}_f$ and $\lambda_n \beta_n \geq 0,$ we have $\lambda_n \beta_n f(u, x_{n+1}) \geq 0.$ Setting $\delta_n = \|x_n - x_{n-1}\|^2,$ then inequality (13) gives

$$a_{n+1} - a_n - \alpha(a_n - a_{n-1}) + (1 - \alpha)\delta_{n+1} - 2\alpha\delta_n \leq \lambda_n \beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right]. \quad (20)$$

In order to simplify its summation we rewrite (20) as

$$\begin{aligned} & a_{n+1} - a_n - \alpha(a_n - a_{n-1}) + (1 - \alpha)(\delta_{n+1} - \delta_n) + (1 - 3\alpha)\delta_n \\ & \leq \lambda_n \beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right]. \end{aligned} \quad (21)$$

Let be $N \in \mathbb{N}^*.$ Summing up from $n = 1$ to N the inequality (21), we get

$$\begin{aligned} & (a_{N+1} - a_1) - \alpha(a_N - a_0) + (1 - \alpha)(\delta_{N+1} - \delta_1) + (1 - 3\alpha) \sum_{n=1}^N \delta_n \\ & \leq \sum_{n=1}^N \lambda_n \beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right]. \end{aligned} \quad (22)$$

Assumption (12), infers that

$$(a_{N+1} - \alpha a_N) + (1 - \alpha)\delta_{N+1} + (1 - 3\alpha) \sum_{n=1}^N \delta_n \leq C, \quad (23)$$

where $C = a_1 - \alpha a_0 + (1 - \alpha)\delta_1 + \sum_{n \geq 1} \lambda_n \beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right] \in \mathbb{R}.$

Since $\alpha < \frac{1}{3}$ yields $1 - 3\alpha > 0$ and $1 - \alpha > 0,$ then inequality (23) implies that for all $n \geq 1$

$$a_{N+1} \leq \alpha a_N + C, \quad \text{with } C \in \mathbb{R}. \quad (24)$$

Recursively we obtain for all $N \geq n_0 \geq 1$

$$\begin{aligned} a_{N+1} & \leq \alpha^{N-n_0+1} a_{n_0} + C(1 + \alpha + \alpha^2 + \dots + \alpha^{N-n_0}) \\ & = \alpha^{N-n_0+1} a_{n_0} + C \frac{1 - \alpha^{N-n_0+1}}{1 - \alpha}. \end{aligned}$$

Therefore, the sequence $\{x_N\}$ is bounded and since

$$\sup_N \|x_{N+1} - x_N\| \leq 2 \sup_N \|x_N\| < +\infty, \quad (25)$$

the sequence $\{\delta_n\}$ is also bounded. Combining (25) with (23) and noticing that $1 - 3\alpha >$

0, yields

$$\sum_{n=1}^{+\infty} \delta_n < +\infty,$$

ensuring (i).

Returning to inequality (13), we have

$$\begin{aligned} a_{n+1} - a_n - \alpha(a_n - a_{n-1}) + \lambda_n \beta_n f(u, x_{n+1}) &\leq \lambda_n \beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right] \\ &\quad + \underbrace{(\alpha - 1) \delta_{n+1} + 2\alpha \delta_n}_{\leq 0} \\ &\leq \lambda_n \beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right] \\ &\quad + 2\alpha \delta_n. \end{aligned}$$

Summing up from $n = 1$ to $+\infty$, we obtain

$$\begin{aligned} \sum_{n=1}^{+\infty} \lambda_n \beta_n f(u, x_{n+1}) &\leq a_1 - \alpha a_0 + \sum_{n=1}^{+\infty} \lambda_n \beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right] \\ &\quad + 2\alpha \sum_{n=1}^{+\infty} \|x_n - x_{n-1}\|^2. \end{aligned}$$

Then, assumptions (12) and (i) ensure (ii). \square

In order to further proceed with the convergence analysis, we have to choose the sequences $\{\lambda_n\}$ and $\{\beta_n\}$ such that $\liminf_{n \rightarrow +\infty} \lambda_n > 0$ and $\beta_n \rightarrow +\infty$. We are now able to state and prove the first main result of this section.

Theorem 2.3. *Suppose given monotone and upper hemicontinuous bifunctions f and g . Let $\{x_n\}$ be a sequence generated by the algorithm (9). Under hypothesis (12) and by assuming that*

$$0 \leq \alpha < \frac{1}{3}, \quad \liminf_{n \rightarrow +\infty} \lambda_n > 0 \quad \text{and} \quad \lim_{n \rightarrow +\infty} \beta_n = +\infty,$$

the sequence $\{x_n\}$ weakly converges to some $\bar{x} \in \mathbf{S}$.

Proof. The proof relies on the discrete Opial's Lemma. To this end we will prove that the conditions (i) and (ii) in Lemma 1.5 for $C = S$ are satisfied.

Returning to inequality (13), since $u \in \mathbf{S}_f$ and $\lambda_n \beta_n \geq 0$, we have $\lambda_n \beta_n f(u, x_{n+1}) \geq 0$, and then

$$a_{n+1} - a_n \leq \alpha(a_n - a_{n-1}) + 2\alpha \|x_n - x_{n-1}\|^2 + \lambda_n \beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right].$$

Taking the positive part, we immediately deduce that

$$[a_{n+1} - a_n]_+ \leq \alpha [a_n - a_{n-1}]_+ + 2\alpha \|x_n - x_{n-1}\|^2 + \lambda_n \beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right].$$

Using assumption (12) together with the fact that $\sum_{n=1}^{+\infty} \|x_n - x_{n-1}\|^2 < +\infty$ and applying Lemma 1.6 with

$$b_n = [a_n - a_{n-1}]_+ \text{ and } w_n = 2\alpha \|x_n - x_{n-1}\|^2 + \lambda_n \beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right],$$

we obtain

$$\sum_{n=1}^{+\infty} [a_n - a_{n-1}]_+ < +\infty.$$

Since a_n is nonnegative, this implies the existence of $\lim_{n \rightarrow +\infty} a_n$ and the one of

$$\lim_{n \rightarrow +\infty} \|x_n - u\|.$$

It remains to show that every weak cluster point \bar{x} of the sequence $\{x_n\}$ lies in \mathbf{S} . Let $n_k \rightarrow +\infty$ as $k \rightarrow +\infty$ such that $x_{n_k} \rightharpoonup \bar{x}$. We want to show that $\bar{x} \in \mathbf{S}$. Thanks to the monotonicity of f and g , inequality (9) ensures that for all $y \in \mathbf{K}$ and for all k large enough

$$f(y, x_{n_k+1}) \leq -\frac{1}{\beta_{n_k}} g(y, x_{n_k+1}) + \frac{1}{\lambda_{n_k} \beta_{n_k}} \langle x_{n_k+1} - y_{n_k}, y - x_{n_k+1} \rangle. \quad (26)$$

Since $\partial g_y(y) \neq \emptyset$, pick $x^*(y) \in \mathbb{H}$ such that for every $z \in \mathbf{K}$

$$g(y, z) \geq \langle x^*(y), z - y \rangle \geq -\|x^*(y)\| \cdot \|y - z\|.$$

Thus there exists $\gamma(y) := \|x^*(y)\| > 0$ such that for every $z \in \mathbf{K}$

$$-g(y, z) \leq \gamma(y) \cdot \|y - z\|. \quad (27)$$

Returning to (26), we can write

$$f(y, x_{n_k+1}) \leq \frac{\gamma(y)}{\beta_{n_k}} \|y - x_{n_k+1}\| + \frac{1}{\lambda_{n_k} \beta_{n_k}} \|x_{n_k+1} - y_{n_k}\| \cdot \|y - x_{n_k+1}\|.$$

Passing to the limit, and using the facts that $\{x_{n_k}\}$ is bounded, $\{\beta_{n_k}\} \rightarrow +\infty$, $\liminf_{k \rightarrow +\infty} \lambda_{n_k} > 0$ and $\|x_{n_k+1} - y_{n_k}\| \rightarrow 0$, we deduce that $f(y, \bar{x}) \leq 0$ for all $y \in \mathbf{K}$.

Lemma 1.3 leads to $\bar{x} \in \mathbf{S}_f$.

Using (9) and the monotonicity of f and g , we have for every $u \in \mathbf{S}_f$,

$$\lambda_n \beta_n f(u, x_{n+1}) + \lambda_n g(u, x_{n+1}) \leq \langle y_n - x_{n+1}, x_{n+1} - u \rangle.$$

Exploiting the fact that $\lim_{n \rightarrow +\infty} \|x_n - u\|$ exists and thanks to (i) of Corollary 2.2, we deduce that

$$\lim_{n \rightarrow +\infty} \langle y_n - x_{n+1}, x_{n+1} - u \rangle = 0.$$

Using Corollary 2.2[(ii)], we obtain that $\limsup_{n \rightarrow +\infty} \lambda_n g(u, x_{n+1}) \leq 0$. Since $g(u, \cdot)$ is lower semicontinuous, from the assumption $\liminf_{n \rightarrow +\infty} \lambda_n > 0$ we derive that $g(u, \bar{x}) \leq 0$. Lemma 1.3 allows us to conclude that

$$g(\bar{x}, u) \geq 0, \forall u \in \mathbf{S}_f,$$

establishing the proof. \square

2.2. Strong convergence analysis

In this subsection, under an additional assumption on the monotonicity of the bifunction of the upper level g , we ensure the strong convergence of the trajectory in (9).

2.2.1. Strong convergence under assumption (12)

Theorem 2.4. *Suppose that the bifunctions f and g are monotone and upper hemi-continuous. Under hypothesis (12), if the bifunction g is ρ -strongly monotone, and if*

$$0 \leq \alpha < \frac{1}{3} \quad \text{and} \quad \sum_{n=1}^{\infty} \lambda_n = +\infty,$$

the sequence $\{x_n\}$ generated by the algorithm (9) strongly converges to a unique solution $u \in \mathbf{S}$.

Proof. Uniqueness of the solution for (BEP) follows from the strong monotonicity of g . For the existence, see [3, Theorem 4.3].

Using inequalities (18) and (17), with $x = u$, summing up and using the ρ -strong monotonicity of g , we get for $a_n := \|x_n - u\|^2$

$$\begin{aligned} a_{n+1} - a_n - \alpha(a_n - a_{n-1}) &\leq -2\rho\lambda_n a_{n+1} + (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 \\ &\quad + 2\lambda_n\beta_n f(x_{n+1}, u) + 2\lambda_n \langle -p, u - x_{n+1} \rangle. \end{aligned}$$

Following the arguments used in the proof of Lemma 2.1 we obtain

$$\begin{aligned} a_{n+1} - a_n - \alpha(a_n - a_{n-1}) + 2\rho\lambda_n a_{n+1} &\leq (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 \\ &\quad + \lambda_n\beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right]. \end{aligned}$$

Then, summing up from $n = 1$ to $+\infty$ yields

$$\begin{aligned} 2\rho \sum_{n=1}^{+\infty} \lambda_n \|x_{n+1} - u\|^2 &\leq a_1 - \alpha a_0 + \sum_{n=1}^{+\infty} \lambda_n\beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right] \\ &\quad + 2\alpha \sum_{n=1}^{+\infty} \|x_n - x_{n-1}\|^2. \end{aligned}$$

Using condition (12) and assumption (i) of Corollary 2.2, we deduce that

$$\sum_{n=1}^{+\infty} \lambda_n \|x_{n+1} - u\|^2 < +\infty.$$

Since $\lim_{n \rightarrow +\infty} \|x_n - u\|$ exists and $\sum_{n=1}^{\infty} \lambda_n = +\infty$, we conclude that $\lim_{n \rightarrow +\infty} \|x_n - u\| = 0$, which guarantees the strong convergence of the whole sequence $\{x_n\}$ to u . \square

2.2.2. Strong convergence without assumption (12)

We will show that in this case, the algorithm strongly converges without the need of the geometric hypothesis (12).

Theorem 2.5. *Suppose that the bifunctions f and g are monotone and upper hemicontinuous with $\mathbf{S}_f \neq \emptyset$ and g is ρ -strongly monotone. Suppose moreover that*

$$0 \leq \alpha < \frac{1}{3}, \quad \lim_{n \rightarrow +\infty} \lambda_n = 0, \quad \sum_{n=0}^{+\infty} \lambda_n = +\infty, \quad \lim_{n \rightarrow +\infty} \beta_n = +\infty \quad \text{and} \quad \liminf_{n \rightarrow +\infty} \lambda_n \beta_n > 0.$$

Then, the sequence $\{x_n\}$ generated by the algorithm (9) converges strongly to the unique solution u of (BEP).

Proof. Under assumptions on the two bifunctions f and g , we get the unique solution denoted by \bar{x} of the bilevel equilibrium problem (BEP).

Step 1: We show that $\{x_n\}$ is bounded.

Since $\{x_n\}$ is generated by the algorithm (9), then by (17), we have for each $x \in \mathbf{K}$

$$\begin{aligned} & \|x_{n+1} - x\|^2 - (1 + \alpha)\|x_n - x\|^2 + \alpha\|x_{n-1} - x\|^2 \\ & \leq (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 + 2\lambda_n \beta_n f(x_{n+1}, x) + 2\lambda_n g(x_{n+1}, x). \end{aligned} \quad (28)$$

Fix $x \in \mathbf{S}_f$, and set $a_n(x) = \|x_n - x\|^2$ and $\delta_n = \|x_n - x_{n-1}\|^2$. Thanks to the monotonicity of f , then for each $n \geq 0$,

$$\begin{aligned} & a_{n+1}(x) - \alpha a_n(x) + 2\alpha \delta_{n+1} \\ & \leq (a_n(x) - \alpha a_{n-1}(x) + 2\alpha \delta_n) + (3\alpha - 1)\delta_{n+1} + 2\lambda_n g(x_{n+1}, x). \end{aligned} \quad (29)$$

Setting $b_n(x) = a_n(x) - \alpha a_{n-1}(x) + 2\alpha \delta_n$, we obtain, for $n \geq 1$,

$$b_{n+1}(x) \leq b_n(x) + (3\alpha - 1)\delta_{n+1} + 2\lambda_n g(x_{n+1}, x). \quad (30)$$

- If there is $n_0 \in \mathbb{N}$ such that $\{b_n(x)\}$ is decreasing for all $n \geq n_0$, then $b_{n+1}(x) \leq b_{n_0}(x)$, which infers that

$$a_{n+1}(x) \leq \alpha a_n(x) + b_{n_0}(x) \quad \text{for all } n \geq n_0.$$

Recursively, we obtain for all $n \geq n_0 \geq 1$

$$a_{n+1}(x) \leq \alpha^{n-n_0} a_{n_0}(x) + b_{n_0}(x) \frac{1 - \alpha^{n-n_0}}{1 - \alpha},$$

and the boundedness of the sequence $\{a_n(x)\}$.

- Otherwise there exists an increasing sequence $\{k_n\}$ such that for every $n \geq 0$, $b_{k_{n+1}}(x) > b_{k_n}(x)$. By Lemma 1.8, there exists a nondecreasing sequence $\{\sigma_n\}$ and $n_0 > 0$ such that $\lim_{n \rightarrow +\infty} \sigma_n = \infty$, and for all $n \geq n_0$, $b_{\sigma_n}(x) < b_{\sigma_{n+1}}(x)$ and $b_n(x) \leq b_{\sigma_{n+1}}(x)$. For $n = \sigma_n$ in (30), we get

$$0 < b_{\sigma_{n+1}}(x) - b_{\sigma_n}(x) \leq (3\alpha - 1)\delta_{\sigma_{n+1}} + 2\lambda_{\sigma_n}g(x_{\sigma_{n+1}}, x). \quad (31)$$

Using the ρ -strong monotonicity of g and relation (27), we deduce that for $n \geq n_0$

$$-2\lambda_{\sigma_n}\gamma(x)\sqrt{a_{\sigma_{n+1}}(x)} \leq 2\lambda_{\sigma_n}g(x, x_{\sigma_{n+1}}) \leq (3\alpha - 1)\delta_{\sigma_{n+1}} - 2\lambda_{\sigma_n}\rho a_{\sigma_{n+1}}(x). \quad (32)$$

Since $3\alpha - 1 < 0$, we conclude that for $n \geq n_0$

$$a_{\sigma_{n+1}}(x) \leq \left(\frac{\gamma(x)}{\rho}\right)^2 \quad \text{and} \quad \delta_{\sigma_{n+1}} \leq \frac{2\gamma^2(x)\lambda_{\sigma_n}}{\rho(1-3\alpha)}. \quad (33)$$

Hence, $\{a_{\sigma_{n+1}}(x)\}$ is bounded. Since $\{\lambda_{\sigma_n}\}$ is bounded too, then $\{\delta_{\sigma_{n+1}}\}$ is bounded, which means that $\{b_{\sigma_{n+1}}(x)\}$ also is bounded. So, there exists $M > 0$, such that for all $n \geq n_0$, we have

$$\begin{aligned} a_n(x) &\leq \alpha a_{n-1}(x) + b_n(x) \\ &\leq \alpha a_{n-1}(x) + b_{\sigma_{n+1}}(x) \\ &\leq \alpha a_{n-1}(x) + M \\ &\leq \alpha^{n-n_0} a_{n_0}(x) + M \frac{1-\alpha^{n-n_0}}{1-\alpha}. \end{aligned}$$

Therefore the sequence $\{a_n(x)\}$ is bounded, ensuring the boundedness of $\{x_n\}$.

Step 2: We show that the sequence $\{x_n\}$ strongly converges to \bar{x} , the unique solution of (BEP).

Let us consider two cases:

Case 1: There exists n_0 such that $\{b_n(\bar{x})\} := a_n(\bar{x}) - \alpha a_{n-1}(\bar{x}) + 2\alpha\delta_n$ is decreasing for $n \geq n_0$.

Then, the limit of the sequence $\{b_n(\bar{x})\}$ exists and $\lim_{n \rightarrow +\infty} (b_n(\bar{x}) - b_{n+1}(\bar{x})) = 0$. For $n \geq n_0$, we have

$$a_{n+1}(\bar{x}) - \alpha a_n(\bar{x}) + 2\alpha\delta_{n+1} \leq a_n(\bar{x}) - \alpha a_{n-1}(\bar{x}) + 2\alpha\delta_n,$$

which implies that

$$a_{n+1}(\bar{x}) \leq a_n(\bar{x}) + \alpha (a_n(\bar{x}) - a_{n-1}(\bar{x})) + 2\alpha(\delta_n - \delta_{n+1}).$$

Since $\sum_{n=n_0}^{+\infty} (a_n(\bar{x}) - a_{n-1}(\bar{x})) \leq \sup_{n>n_0} a_n(\bar{x})$ and due to the fact that $a_n(\bar{x})$ is bounded

(by step 1), it follows that $\sum_{n=n_0}^{+\infty} (a_n(\bar{x}) - a_{n-1}(\bar{x})) < +\infty$. By Lemma 1.6, and since

$\sum_{n=n_0}^{+\infty} [\alpha (a_n(\bar{x}) - a_{n-1}(\bar{x})) + 2\alpha(\delta_n - \delta_{n+1})] < +\infty$, then the limit of $\{a_n(\bar{x})\}$ exists.

Therefore, it suffices to show that $\liminf_{n \rightarrow \infty} a_n(\bar{x}) = 0$. Since g is ρ -strongly monotone, then we have

$$\begin{aligned} \lim_{n \rightarrow +\infty} a_{n+1}(\bar{x}) &= \liminf_{n \rightarrow +\infty} \|x_{n+1} - \bar{x}\|^2 \\ &\leq \frac{1}{\rho} \liminf_{n \rightarrow +\infty} (-g(x_{n+1}, \bar{x})) + \frac{1}{\rho} \limsup_{n \rightarrow +\infty} (-g(\bar{x}, x_{n+1})). \end{aligned} \quad (34)$$

Thus, it suffices to prove that $\liminf_{n \rightarrow +\infty} (-g(x_{n+1}, \bar{x})) \leq 0$ and $\liminf_{n \rightarrow +\infty} g(\bar{x}, x_{n+1}) \geq 0$. Taking into account that $\bar{x} \in \mathbf{S}_f$, we derive from (30) that

$$b_{n+1}(\bar{x}) \leq b_n(\bar{x}) + (3\alpha - 1)\delta_{n+1} + 2\lambda_n g(x_{n+1}, \bar{x}).$$

Hence, since $3\alpha - 1 < 0$, then

$$-\lambda_n g(x_{n+1}, \bar{x}) \leq \frac{1}{2} (b_n(\bar{x}) - b_{n+1}(\bar{x})). \quad (35)$$

Summing up inequality (35) from 1 to $+\infty$, we deduce that

$$\sum_{n=1}^{+\infty} -\lambda_n g(x_{n+1}, \bar{x}) \leq \frac{1}{2} \left(b_1(\bar{x}) - \lim_{n \rightarrow +\infty} b_n(\bar{x}) \right) < +\infty,$$

which in combination with $\sum_{n=0}^{+\infty} \lambda_n = +\infty$ leads to $\liminf_{n \rightarrow \infty} (-g(x_{n+1}, \bar{x})) \leq 0$.

Let us prove that $\liminf_{n \rightarrow +\infty} g(\bar{x}, x_{n+1}) \geq 0$. The sequence $\{x_n\}$ is bounded, so there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ that converges weakly to some $x \in \mathbf{K}$. By using the weak lower semicontinuity of $g(\bar{x}, \cdot)$ we have

$$g(\bar{x}, x) \leq \liminf_{k \rightarrow +\infty} g(\bar{x}, x_{n_k+1}).$$

Since \bar{x} is the unique solution of (BEP), we need just to check that $x \in \mathbf{S}_f$. In doing so, by (27) and (17), we have for every $y \in \mathbf{K}$,

$$f(y, x_{n+1}) \leq -\frac{1}{2\lambda_n \beta_n} (b_{n+1}(y) - b_n(y)) + \frac{1}{2\beta_n} \gamma(y) \sqrt{a_n(y)}. \quad (36)$$

We have

$$\begin{aligned} &b_n(y) - b_{n+1}(y) \\ &= (a_n(y) - \alpha a_{n-1}(y) + 2\alpha \delta_n) - (a_{n+1}(y) - \alpha a_n(y) + 2\alpha \delta_{n+1}) \\ &= (a_n(y) - a_{n+1}(y)) + \alpha (a_n(y) - a_{n-1}(y)) + 2\alpha (\delta_n - \delta_{n+1}) \\ &= (a_n(\bar{x}) - a_{n+1}(\bar{x}) + 2\langle x_n - x_{n+1}, \bar{x} - y \rangle) \\ &\quad - \alpha (a_{n-1}(\bar{x}) - a_n(\bar{x}) + 2\langle x_n - x_{n-1}, \bar{x} - y \rangle) + 2\alpha (\delta_n - \delta_{n+1}) \\ &= b_n(\bar{x}) - b_{n+1}(\bar{x}) + 2\langle x_n - x_{n+1}, \bar{x} - y \rangle + 2\alpha \langle x_n - x_{n-1}, \bar{x} - y \rangle. \end{aligned}$$

Since $\lim_{n \rightarrow +\infty} (b_n(\bar{x}) - b_{n+1}(\bar{x})) = 0$ and $\lim_{n \rightarrow +\infty} \|x_{n+1} - x_n\| = 0$, then $\lim_{n \rightarrow +\infty} (b_n(y) - b_{n+1}(y)) = 0$.

By using the weak lower semicontinuity of $f(y, \cdot)$ and the fact that $\{x_n\}$ is bounded, $\lim_{n \rightarrow +\infty} \lambda_n = 0$, $\liminf_{n \rightarrow +\infty} \lambda_n \beta_n > 0$ and $\lim_{n \rightarrow +\infty} \beta_n = +\infty$, we conclude from (36) that for every $y \in \mathbf{K}$

$$f(y, x) \leq \liminf_{n \rightarrow +\infty} f(y, x_{n+1}) \leq 0.$$

Hence, by using Minty's lemma we deduce that $x \in \mathbf{S}_f$ and therefore,

$$0 \leq g(\bar{x}, x) \leq \liminf_{n \rightarrow +\infty} g(\bar{x}, x_{n+1}).$$

Hence, by (34),

$$\lim_{n \rightarrow +\infty} a_{n+1}(\bar{x}) \leq -\frac{1}{\rho} \liminf_{n \rightarrow +\infty} g(\bar{x}, x_{n+1}) \leq 0,$$

and so $\lim_{n \rightarrow +\infty} a_n(\bar{x}) = 0$.

Case 2: There exists a subsequence $\{x_{n_j}\}$ of $\{x_n\}$ such that $b_{n_j}(\bar{x}) \leq b_{n_j+1}(\bar{x})$ for all $j \in \mathbb{N}$.

By Lemma 1.8, the sequence $\sigma(n) := \max\{k \leq n : b_k(\bar{x}) < b_{k+1}(\bar{x})\}$ is nondecreasing, $\lim_{n \rightarrow +\infty} \sigma(n) = \infty$ and, for all $n \geq n_0$

$$b_{\sigma(n)}(\bar{x}) < b_{\sigma(n)+1}(\bar{x}) \quad \text{and} \quad b_n(\bar{x}) \leq b_{\sigma(n)+1}(\bar{x}).$$

Let us take $n = \sigma(n)$ and $x = \bar{x}$ in (30). We have

$$0 < b_{\sigma(n)+1}(\bar{x}) - b_{\sigma(n)}(\bar{x}) \leq 2\lambda_{\sigma(n)}g(x_{\sigma(n)+1}, \bar{x}), \quad (37)$$

which yields $g(x_{\sigma(n)+1}, \bar{x}) \geq 0$, and thus $\limsup_{n \rightarrow +\infty} g(x_{\sigma(n)+1}, \bar{x}) \geq 0$.

Using again the ρ -strong monotonicity of g and passing to the limit we have

$$\begin{aligned} \limsup_{n \rightarrow +\infty} a_{\sigma(n)+1}(\bar{x}) &\leq \frac{1}{\rho} \limsup_{n \rightarrow +\infty} -g(x_{\sigma(n)+1}, \bar{x}) + \frac{1}{\rho} \limsup_{n \rightarrow +\infty} -g(\bar{x}, x_{\sigma(n)+1}) \\ &\leq \underbrace{\limsup_{n \rightarrow +\infty} -g(x_{\sigma(n)+1}, \bar{x})}_{\leq 0} + \frac{1}{\rho} \limsup_{n \rightarrow +\infty} -g(\bar{x}, x_{\sigma(n)+1}) \\ &\leq -\frac{1}{\rho} \liminf_{n \rightarrow +\infty} g(\bar{x}, x_{\sigma(n)+1}). \end{aligned} \quad (38)$$

Under the boundedness of $\{x_n\}$, and similarly to the case 1, one can show that

$$\liminf_{n \rightarrow +\infty} g(\bar{x}, x_{\sigma(n)+1}) \geq 0.$$

Hence, by (34), we conclude that

$$\lim_{n \rightarrow +\infty} a_{\sigma(n)+1}(\bar{x}) = 0. \quad (39)$$

Since $b_n(\bar{x}) \leq b_{\sigma(n)+1}(\bar{x})$ for each $n \geq n_0$, we derive that

$$\begin{aligned} \lim_{n \rightarrow +\infty} a_n(\bar{x}) &\leq \lim_{n \rightarrow +\infty} b_n(\bar{x}) \leq \lim_{n \rightarrow +\infty} b_{\sigma(n)+1}(\bar{x}) \\ &\leq \lim_{n \rightarrow +\infty} (a_{\sigma(n)}(\bar{x}) + 2\alpha\delta_{\sigma(n)}). \end{aligned}$$

Since $\delta_{\sigma(n)} = \|x_{\sigma(n)} - x_{\sigma(n-1)}\|^2 \leq 2a_{\sigma(n)}(\bar{x}) + 2a_{\sigma(n)-1}(\bar{x})$, then

$$\begin{aligned} \lim_{n \rightarrow +\infty} a_n(\bar{x}) &\leq (1 + 4\alpha) \underbrace{\lim_{n \rightarrow +\infty} a_{\sigma(n)}(\bar{x})}_{=0} + 4\alpha \underbrace{\lim_{n \rightarrow +\infty} a_{\sigma(n)-1}(\bar{x})}_{=0} \\ &= 0, \end{aligned}$$

thus guaranteeing the strong convergence of the whole sequence $\{x_n\}$ to \bar{x} . \square

3. Convergence of inertial forward-backward algorithm (IFBA)

3.1. Weak convergence analysis

In this subsection, we study the weak convergence of the sequence $\{x_n\}$ generated by the algorithm (IFBA), given by:

$$g(x_{n+1}, y) + \beta_n \langle Bx_n, y - x_{n+1} \rangle + \frac{1}{\lambda_n} \langle x_{n+1} - y_n, y - x_{n+1} \rangle \geq 0, \quad \forall y \in \mathbf{K},$$

where $g : \mathbf{K} \times \mathbf{K} \rightarrow \mathbb{R}$ is a bifunction satisfying (i) – (iii) – (iv) – (v) of Definition 1.1 and B is an θ -inverse strongly monotone mapping from \mathbb{H} into \mathbb{H} .

Let us underline that it is necessary to include the geometric condition (12). Here, we have $f(x, y) = \langle Bx, y - x \rangle$ for all $x, y \in \mathbb{H}$, and $\mathbf{S}_f = B^{-1}(0)$. In this case, the condition (12) can be expressed in a simplified form as: $\forall u \in B^{-1}(0)$, for all $p \in N_{B^{-1}(0)}(u)$,

$$\sum_{n=1}^{+\infty} \lambda_n \beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right] < +\infty, \quad (40)$$

where $\mathcal{F}_B(u, x) = \sup_{y \in \mathbb{H}} \{ \langle x, y \rangle + \langle By, u - y \rangle \}$.

Similar to what was done in Lemma 2.1, we initiate our study by proving the following estimation:

Lemma 3.1. *Let $\{x_n\}$ be a sequence generated by the algorithm (10). Take $u \in \mathbf{S}$ and set $a_n := \|x_n - u\|^2$. Then, there exists $p \in N_{B^{-1}(0)}(u)$ such that for each $\rho > 0$ and $n \geq 1$, the following inequality holds:*

$$\begin{aligned} a_{n+1} - a_n - \alpha(a_n - a_{n-1}) &\leq ((1 + \rho)\alpha - 1) \|x_{n+1} - x_n\|^2 + 2\alpha \|x_n - x_{n-1}\|^2 \\ &\quad + \frac{1}{2\rho\alpha} \lambda_n^2 \|p\|^2 + \lambda_n \beta_n \left(\frac{2\lambda_n \beta_n}{\rho\alpha} - \theta \right) \|Bx_n\|^2 \\ &\quad + \lambda_n \beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right]. \end{aligned} \quad (41)$$

Proof. Take $u \in \mathbf{S}$. Let $p \in N_{\mathbf{S}_f}(u)$ be such that $-p \in A^g(u)$, we have for every $n \geq 1$

$$\lambda_n g(u, x_{n+1}) + \lambda_n \langle -p, u - x_{n+1} \rangle \geq 0. \quad (42)$$

Given that $\{x_n\}$ is generated by the algorithm (10), then it easy to check that

$$0 \leq \lambda_n \beta_n \langle Bx_n, u - x_{n+1} \rangle + \lambda_n g(x_{n+1}, u) + \frac{1}{2} (\|y_n - u\|^2 - \|x_{n+1} - u\|^2 - \|x_{n+1} - y_n\|^2). \quad (43)$$

Using the relations (15) and (16) (see the proof of Lemma 2.1), along with the two last inequalities, we get

$$a_{n+1} - a_n - \alpha(a_n - a_{n-1}) \leq (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 + 2\lambda_n \beta_n \langle Bx_n, u - x_{n+1} \rangle + 2\lambda_n \langle -p, u - x_{n+1} \rangle. \quad (44)$$

For each $\rho > 0$, on one hand, we have

$$2\lambda_n \langle -p, u - x_{n+1} \rangle = 2\lambda_n \langle -p, u - x_n \rangle + 2\lambda_n \langle -p, x_n - x_{n+1} \rangle \leq \lambda_n \langle -p, u - x_n \rangle + \frac{2}{\rho\alpha} \lambda_n^2 \|p\|^2 + \frac{\rho\alpha}{2} \|x_n - x_{n+1}\|^2. \quad (45)$$

On the other hand, we have

$$2\lambda_n \beta_n \langle Bx_n, u - x_{n+1} \rangle = 2\lambda_n \beta_n \langle Bx_n, u - x_n \rangle + 2\lambda_n \beta_n \langle Bx_n, x_n - x_{n+1} \rangle. \quad (46)$$

Due to the fact that B is θ -inverse strongly monotone, we have

$$\langle Bx_n, u - x_n \rangle \leq -\theta \|Bx_n\|^2.$$

We also have

$$2\lambda_n \beta_n \langle Bx_n, x_n - x_{n+1} \rangle \leq \frac{2\lambda_n^2 \beta_n^2}{\rho\alpha} \|Bx_n\|^2 + \frac{\rho\alpha}{2} \|x_n - x_{n+1}\|^2.$$

The two last inequalities together with (46), lead to

$$\begin{aligned} 2\lambda_n \beta_n \langle Bx_n, u - x_{n+1} \rangle &\leq -\lambda_n \beta_n \theta \|Bx_n\|^2 + \lambda_n \beta_n \langle Bx_n, u - x_n \rangle \\ &\quad + \frac{2\lambda_n^2 \beta_n^2}{\rho\alpha} \|Bx_n\|^2 + \frac{\rho\alpha}{2} \|x_n - x_{n+1}\|^2 \\ &= \lambda_n \beta_n \left(\frac{2\lambda_n \beta_n}{\rho\alpha} - \theta \right) \|Bx_n\|^2 + \lambda_n \beta_n \langle Bx_n, u - x_n \rangle \\ &\quad + \frac{\rho\alpha}{2} \|x_n - x_{n+1}\|^2. \end{aligned} \quad (47)$$

Combining this inequality with (44) and (45), we obtain

$$\begin{aligned} a_{n+1} - a_n - \alpha(a_n - a_{n-1}) &\leq ((1 + \rho)\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 \\ &\quad + \frac{1}{2\rho\alpha} \lambda_n^2 \|p\|^2 + \lambda_n \beta_n \left(\frac{2\lambda_n \beta_n}{\rho\alpha} - \theta \right) \|Bx_n\|^2 \\ &\quad + \lambda_n \beta_n \langle Bx_n, u - x_n \rangle + 2\lambda_n \langle -p, u - x_n \rangle. \end{aligned} \quad (48)$$

Since,

$$\begin{aligned}
& \lambda_n \beta_n \langle Bx_n, u - x_n \rangle + 2\lambda_n \langle -p, u - x_n \rangle \\
&= \lambda_n \beta_n \left(\langle Bx_n, u - x_n \rangle + \left\langle -\frac{2p}{\beta_n}, u - x_n \right\rangle \right) \\
&= \lambda_n \beta_n \left[\left\langle \frac{2p}{\beta_n}, x_n \right\rangle + \langle Bx_n, u - x_n \rangle - \left\langle \frac{2p}{\beta_n}, u \right\rangle \right] \\
&\leq \lambda_n \beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right]. \quad (49)
\end{aligned}$$

It follows that for every $n \geq 1$,

$$\begin{aligned}
& a_{n+1} - a_n - \alpha(a_n - a_{n-1}) \\
&\leq ((1 + \rho)\alpha - 1) \|x_{n+1} - x_n\|^2 + 2\alpha \|x_n - x_{n-1}\|^2 \\
&\quad + \frac{1}{2\rho\alpha} \lambda_n^2 \|p\|^2 + \lambda_n \beta_n \left(\frac{2\lambda_n \beta_n}{\rho\alpha} - \theta \right) \|Bx_n\|^2 \\
&\quad + \lambda_n \beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right].
\end{aligned}$$

The proof is complete. \square

Corollary 3.2. *Under hypothesis (40) and by assuming that there exists $\rho > 0$ such that*

$$0 \leq \alpha < \frac{1}{3 + \rho}, \quad (\lambda_n) \in \ell^2 \text{ and } \limsup_{n \rightarrow +\infty} \lambda_n \beta_n < \frac{\rho\alpha\theta}{2},$$

we have

$$\begin{aligned}
& (i) \sum_{n=1}^{+\infty} \|x_n - x_{n-1}\|^2 < +\infty; \\
& (ii) \sum_{n=1}^{+\infty} \lambda_n \beta_n \|Bx_n\|^2 < +\infty.
\end{aligned}$$

Proof. Let us take $\delta_n = \|x_{n+1} - x_n\|^2$, then the inequality (41) becomes

$$\begin{aligned}
a_{n+1} - a_n - \alpha(a_n - a_{n-1}) &\leq ((1 + \rho)\alpha - 1)\delta_{n+1} + 2\alpha\delta_n \\
&\quad + \frac{1}{2\rho\alpha} \lambda_n^2 \|p\|^2 + \lambda_n \beta_n \left(\frac{2\lambda_n \beta_n}{\rho\alpha} - \theta \right) \|Bx_n\|^2 \\
&\quad + \lambda_n \beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right]. \quad (50)
\end{aligned}$$

Since $\limsup_{n \rightarrow +\infty} \lambda_n \beta_n < \frac{\rho\alpha\theta}{2}$, then

$$\begin{aligned}
& a_{n+1} - a_n - \alpha(a_n - a_{n-1}) + (1 - \alpha)(\delta_{n+1} - \delta_n) + (1 - 3(\alpha + \rho))\delta_n \\
&\leq \frac{1}{2\rho\alpha} \lambda_n^2 \|p\|^2 + \lambda_n \beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right]. \quad (51)
\end{aligned}$$

Following the same argument as (22) in the proof of Corollary 2.2, one can obtain for $N \in \mathbb{N}^*$

$$(a_{N+1} - \alpha a_N) + (1 - \alpha)\delta_{N+1} + (1 - 3(\alpha + \rho)) \sum_{n=1}^N \delta_n \leq D, \quad (52)$$

where

$$D = a_1 - \alpha a_0 + (1 - \alpha)\delta_1 + \frac{1}{2\rho\alpha} \|p\|^2 \sum_{n \geq 1} \lambda_n^2 + \sum_{n \geq 1} \lambda_n \beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right] \in \mathbb{R}.$$

By hypothesis, we have $1 - 3(\alpha + \rho) > 0$ and $1 - \alpha > 0$, then for all $n \geq 1$, the inequality (52) leads to

$$a_{N+1} \leq \alpha a_N + D, \quad \text{with } D \in \mathbb{R}, \quad (53)$$

which is identical to (24) in the proof of Corollary 2.2. Therefore, one can conclude that the sequence $\{x_N\}$ is bounded and hence $\{\delta_n\}$ is bounded too. Using again (52) with the fact that $\{x_n\}$ is bounded and $1 - 3(\alpha + \rho) > 0$, we obtain

$$\sum_{n=1}^{+\infty} \delta_n < +\infty,$$

ensuring (i).

Returning to inequality (50), we have

$$\begin{aligned} & a_{n+1} - a_n - \alpha(a_n - a_{n-1}) + \lambda_n \beta_n \left(\theta - \frac{2\lambda_n \beta_n}{\rho\alpha} \right) \|Bx_n\|^2 \\ & \leq \lambda_n \beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right] + \frac{1}{2\rho\alpha} \lambda_n^2 \|p\|^2 + \underbrace{((1 + \rho)\alpha - 1)}_{<0} \delta_{n+1} + 2\alpha\delta_n \\ & \leq \lambda_n \beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right] + \frac{1}{2\rho\alpha} \lambda_n^2 \|p\|^2 + 2\alpha\delta_n. \end{aligned} \quad (54)$$

Then, since $\limsup_{n \rightarrow +\infty} \lambda_n \beta_n < \frac{\rho\alpha\theta}{2}$, there exists $\epsilon > 0$ such that $\frac{2\lambda_n \beta_n}{\rho\alpha} - \theta \leq -\epsilon$. Hence, summing up from $n = 1$ to $+\infty$ the inequality (54), we conclude that

$$\begin{aligned} \epsilon \sum_{n=1}^{+\infty} \lambda_n \beta_n \|Bx_n\|^2 & \leq a_1 - \alpha a_0 + \sum_{n=1}^{+\infty} \lambda_n \beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right] \\ & \quad + 2\alpha \sum_{n=1}^{+\infty} \|x_n - x_{n-1}\|^2. \end{aligned}$$

Then, assumptions (40) and (i) ensure (ii). \square

Theorem 3.3. Let $g : \mathbf{K} \times \mathbf{K} \rightarrow \mathbb{R}$ be a bifunction satisfying (i) – (iii) – (iv) – (v) of Definition 1.1 and $B : \mathbb{H} \rightarrow \mathbb{H}$ is θ -inverse strongly monotone for some $\theta > 0$. Under hypothesis (40) and by supposing there exists $\rho > 0$ such that

$$0 \leq \alpha < \frac{1}{3+\rho}, \quad (\lambda_n) \in \ell^2 \setminus \ell^1, \quad \lim_{n \rightarrow +\infty} \beta_n = +\infty, \quad \limsup_{n \rightarrow +\infty} \lambda_n \beta_n < \frac{\rho \alpha \theta}{2} \quad \text{and} \\ \liminf_{n \rightarrow +\infty} \lambda_n \beta_n > 0.$$

Then, the sequence $\{x_n\}$ generated by algorithm (10) weakly converges to some $\bar{x} \in \mathbf{S}$.

Proof. In order to reach the desired conclusion, we shall check the hypotheses of Opial's Lemma. Let us take $u \in \mathbf{S}$. By Lemma 3.1, there exists $p \in N_{B^{-1}(0)}(u)$ such that for each $\rho > 0$ and $n \geq 1$, we have

$$a_{n+1} - a_n - \alpha(a_n - a_{n-1}) \leq ((1 + \rho)\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 \\ + \frac{1}{2\rho\alpha}\lambda_n^2\|p\|^2 + \lambda_n\beta_n \left(\frac{2\lambda_n\beta_n}{\rho\alpha} - \theta \right) \|Bx_n\|^2 \\ + \lambda_n\beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right].$$

By taking $(3 + \rho)\alpha < 1$, and $\limsup_{n \rightarrow +\infty} \lambda_n\beta_n < \frac{\rho\alpha\theta}{2}$, we get

$$a_{n+1} - a_n - \alpha(a_n - a_{n-1}) \\ \leq 2\alpha\|x_n - x_{n-1}\|^2 + \frac{1}{2\rho\alpha}\lambda_n^2\|p\|^2 + \lambda_n\beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right].$$

Similar to the proof of Theorem 2.3, to prove that $\lim_{n \rightarrow +\infty} a_n$ exists, using the fact that

$$\sum_{n \geq 1} \frac{1}{2\rho\alpha}\lambda_n^2\|p\|^2 + \lambda_n\beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right] < +\infty, \text{ it suffices to replace}$$

$$\lambda_n\beta_n \left[\mathcal{F}_f \left(u, \frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n} \right) \right]$$

by

$$\frac{1}{2\rho\alpha}\lambda_n^2\|p\|^2 + \lambda_n\beta_n \left[\mathcal{F}_B \left(u, \frac{2p}{\beta_n} \right) - \sigma_{B^{-1}(0)} \left(\frac{2p}{\beta_n} \right) \right].$$

Now, we need to verify that each weak cluster point of $\{x_n\}$ belongs to \mathbf{S} . Let $n_k \rightarrow +\infty$ as $k \rightarrow +\infty$ such that $x_{n_k} \rightharpoonup \bar{x}$. By the inequality (10), we have for all k large enough and for each $y \in \mathbf{K}$

$$\beta_{n_k} \langle Bx_{n_k}, x_{n_k} - x_{n_k+1} \rangle + \beta_{n_k} \langle Bx_{n_k}, y - x_{n_k} \rangle \\ + g(x_{n_k+1}, y) + \frac{1}{\lambda_{n_k}} \langle x_{n_k+1} - y_{n_k}, y - x_{n_k+1} \rangle \\ \geq 0.$$

Using the fact that g and B are both monotone, we obtain for each $y \in \mathbf{K}$

$$\begin{aligned} \langle By, x_{n_k} - y \rangle &\leq \langle Bx_{n_k}, x_{n_k} - x_{n_k+1} \rangle - \frac{1}{\beta_{n_k}} g(y, x_{n_k+1}) \\ &\quad + \frac{1}{\lambda_{n_k} \beta_{n_k}} \langle x_{n_k+1} - y_{n_k}, y - x_{n_k+1} \rangle. \end{aligned} \quad (55)$$

Using the inequality (27) used in the proof of Theorem 2.3 leads us to the following one, valid for each $y \in \mathbf{K}$

$$\begin{aligned} \langle By, x_{n_k} - y \rangle &\leq \langle Bx_{n_k}, x_{n_k} - x_{n_k+1} \rangle \\ &\quad + \frac{\gamma(y)}{\beta_{n_k}} \|y - x_{n_k+1}\| + \frac{1}{\lambda_{n_k} \beta_{n_k}} \langle x_{n_k+1} - y_{n_k}, y - x_{n_k+1} \rangle. \end{aligned}$$

Hence for all $y \in \mathbf{K}$ we have

$$\begin{aligned} \langle By, x_{n_k} - y \rangle &\leq \|Bx_{n_k}\| \|x_{n_k} - x_{n_k+1}\| + \frac{\gamma(y)}{\beta_{n_k}} \|y - x_{n_k+1}\| \\ &\quad + \frac{1}{\lambda_{n_k} \beta_{n_k}} \|x_{n_k+1} - y_{n_k}\| \|y - x_{n_k+1}\|. \end{aligned}$$

So, now by passing to the limit, and using the facts that $\{x_{n_k}\}$ is bounded, $\lim \beta_{n_k} = +\infty$, $\liminf_{k \rightarrow +\infty} \lambda_{n_k} \beta_{n_k} > 0$ and $\|x_{n_k+1} - y_{n_k}\| \rightarrow 0$, we deduce that $\langle By, \bar{x} - y \rangle \leq 0$ for all $y \in \mathbf{K}$. Hence, by the maximality of B (see Remark 1), we conclude that $\bar{x} \in B^{-1}(0)$.

Returning to the inequality (43), for every $x = u \in B^{-1}(0)$, we have

$$\begin{aligned} &\lambda_n \beta_n \langle Bx_n, u - x_{n+1} \rangle + \lambda_n g(x_{n+1}, u) \\ &+ \frac{1}{2} (\|y_n - u\|^2 - \|x_{n+1} - u\|^2 - \|x_{n+1} - y_n\|^2) \\ &\geq 0. \end{aligned} \quad (56)$$

Using the monotonicity of B and g , and the fact that $u \in B^{-1}(0)$, we get

$$\begin{aligned} &\lambda_n \beta_n \langle Bx_n, x_n - x_{n+1} \rangle \\ &+ \frac{1}{2} (\|y_n - u\|^2 - \|x_{n+1} - u\|^2 - \|x_{n+1} - y_n\|^2) \\ &\geq \lambda_n g(u, x_{n+1}). \end{aligned}$$

Using again the inequalities (15) and (16) utilized in the proof of Lemma 2.1, we obtain

$$\begin{aligned}
2\lambda_n g(u, x_{n+1}) &\leq 2\lambda_n \beta_n \langle Bx_n, x_n - x_{n+1} \rangle + (1 + \alpha) \|x_n - u\|^2 - \|x_{n+1} - u\|^2 \\
&\quad - \alpha \|x_{n-1} - u\|^2 + 2\alpha \|x_n - x_{n-1}\|^2 + (\alpha - 1) \|x_{n+1} - x_n\|^2 \\
&\leq \lambda_n \beta_n \|Bx_n\|^2 + \lambda_n \beta_n \|x_n - x_{n+1}\|^2 + (\|x_n - u\|^2 - \|x_{n+1} - u\|^2) \\
&\quad + \alpha (\|x_n - u\|^2 - \|x_{n-1} - u\|^2) \\
&\quad + (1 - \alpha) [\|x_n - x_{n-1}\|^2 - \|x_{n+1} - x_n\|^2] + \underbrace{(3\alpha - 1)}_{\leq 0} \|x_n - x_{n-1}\|^2.
\end{aligned}$$

Let a fixed $N > 1$. Summing up the latter relation from $n = 1$ to $n = N$ and letting $N \rightarrow +\infty$, we get

$$\begin{aligned}
2 \sum_{n \geq 1} \lambda_n g(u, x_{n+1}) &\leq \sum_{n \geq 1} \lambda_n \beta_n \|Bx_n\|^2 + \sum_{n \geq 1} \lambda_n \beta_n \|x_n - x_{n+1}\|^2 + \|x_1 - u\|^2 \\
&\quad + \alpha \lim_{N \rightarrow +\infty} \|x_N - u\|^2 + (1 - \alpha) \|x_1 - x_0\|^2.
\end{aligned}$$

Corollary 3.2 yields $\sum_{n \geq 1} \lambda_n \beta_n \|Bx_n\|^2 < +\infty$ and $\sum_{n \geq 1} \|x_n - x_{n+1}\|^2 < +\infty$. Since

$\lim_{n \rightarrow +\infty} x_n$ exists and $\limsup_{n \rightarrow +\infty} \lambda_n \beta_n < \frac{\rho \alpha \theta}{2}$, then

$$\sum_{n \geq 1} \lambda_n g(u, x_{n+1}) < +\infty.$$

The fact that $\sum_{n \geq 1} \lambda_n = +\infty$ gives $\liminf_{n \rightarrow +\infty} g(u, x_{n+1}) \leq 0$. Since $g(u, \cdot)$ is lower semicontinuous, we derive that $g(u, \bar{x}) \leq 0$. Lemma 1.3 allows us to conclude that

$$g(\bar{x}, u) \geq 0, \quad \forall u \in \mathbf{S}_f,$$

establishing the proof. □

3.2. Strong convergence without assumption (40)

Similar to the first algorithm, we demonstrate the strong convergence without relying on the assumption (40).

Theorem 3.4. *Let $g : \mathbf{K} \times \mathbf{K} \rightarrow \mathbb{R}$ be a bifunction satisfying the conditions (i) – (iii) – (iv) – (v) of Definition 1.1. Suppose that $B : \mathbb{H} \rightarrow \mathbb{H}$ is θ -inverse strongly monotone for some $\theta > 0$. Suppose moreover that g is strongly monotone and there exists $\rho > 0$ such that*

$$\begin{aligned}
0 \leq \alpha < \frac{1}{3+\rho}, \quad \sum_{n=0}^{+\infty} \lambda_n = +\infty, \quad \lim_{n \rightarrow +\infty} \beta_n = +\infty, \quad \limsup_{n \rightarrow +\infty} \lambda_n \beta_n < \frac{\rho \alpha \theta}{2} \quad \text{and} \\
\liminf_{n \rightarrow +\infty} \lambda_n \beta_n > 0.
\end{aligned}$$

Then, the sequence $\{x_n\}$ generated by the algorithm (10) converges strongly to the unique solution u of (BEP).

Proof. Similar to the proof of Theorem 2.5, since g is strongly monotone, then the bilevel equilibrium problem (**BEP**) has a unique solution denoted by \bar{x} . Firstly, we show that the sequence $\{x_n\}$ is bounded. Since $\{x_n\}$ is generated by the algorithm (10), then combining (43) with the relations (15) and (16), (see the proof of Lemma 2.1), it holds for each $x \in \mathbf{K}$,

$$\begin{aligned} & \|x_{n+1} - x\|^2 - (1 + \alpha)\|x_n - x\|^2 + \alpha\|x_{n-1} - x\|^2 \\ & \leq (\alpha - 1)\|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 \\ & \quad + 2\lambda_n\beta_n\langle Bx_n, x - x_{n+1}\rangle + 2\lambda_n g(x_{n+1}, x). \end{aligned} \quad (57)$$

Using again the inequality (47), we have, for $\rho > 0$

$$\begin{aligned} 2\lambda_n\beta_n\langle Bx_n, x - x_{n+1}\rangle & = \lambda_n\beta_n \left(\frac{2\lambda_n\beta_n}{\rho\alpha} - \theta \right) \|Bx_n\|^2 + \lambda_n\beta_n\langle Bx_n, x - x_n \rangle \\ & \quad + \frac{\rho\alpha}{2}\|x_n - x_{n+1}\|^2, \end{aligned}$$

and so

$$\begin{aligned} & \|x_{n+1} - x\|^2 - (1 + \alpha)\|x_n - x\|^2 + \alpha\|x_{n-1} - x\|^2 \\ & \leq \left(\alpha \left(1 + \frac{\rho}{2} \right) - 1 \right) \|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 \\ & \quad + \lambda_n\beta_n \left(\frac{2\lambda_n\beta_n}{\rho\alpha} - \theta \right) \|Bx_n\|^2 + \lambda_n\beta_n\langle Bx_n, x - x_n \rangle \\ & \quad + 2\lambda_n g(x_{n+1}, x). \end{aligned} \quad (58)$$

Since $\limsup_{n \rightarrow +\infty} \lambda_n\beta_n < \frac{\rho\alpha\theta}{2}$, then

$$\begin{aligned} & \|x_{n+1} - x\|^2 - (1 + \alpha)\|x_n - x\|^2 + \alpha\|x_{n-1} - x\|^2 \\ & \leq \left(\alpha \left(1 + \frac{\rho}{2} \right) - 1 \right) \|x_{n+1} - x_n\|^2 + 2\alpha\|x_n - x_{n-1}\|^2 \\ & \quad + \lambda_n\beta_n\langle Bx_n, x - x_n \rangle + 2\lambda_n g(x_{n+1}, x). \end{aligned} \quad (59)$$

Let us fix $x \in B^{-1}(0)$, and set $a_n(x) := \|x_n - x\|^2$ and $\delta_n = \|x_n - x_{n-1}\|^2$. Then, using the monotonicity of B , we get for each $n \geq 0$,

$$\begin{aligned} & a_{n+1}(x) - \alpha a_n(x) + 2\alpha\delta_{n+1} \\ & \leq (a_n(x) - \alpha a_{n-1}(x) + 2\alpha\delta_n) + \left(\left(3 + \frac{\rho}{2} \right) \alpha - 1 \right) \delta_{n+1} + 2\lambda_n g(x_{n+1}, x). \end{aligned} \quad (60)$$

Setting $b_n(x) = a_n(x) - \alpha a_{n-1}(x) + 2\alpha\delta_n$, we obtain, for $n \geq 1$,

$$b_{n+1}(x) \leq b_n(x) + \left(\left(3 + \frac{\rho}{2} \right) \alpha - 1 \right) \delta_{n+1} + 2\lambda_n g(x_{n+1}, x). \quad (61)$$

Finally, one can conclude the boundedness of the sequence $\{a_n(x)\}$ by using the same argument as in step 1, in the proof of Theorem 2.5.

Secondly, similar to the step 2 in the proof of Theorem 2.5, by using the inequality (61), one can show the existence of the limit of $\{a_n(\bar{x})\}$ and that $\lim_{n \rightarrow +\infty} a_n(\bar{x}) = 0$. \square

4. Application to optimization and saddle point problems

In this section, we give two examples of particular bifunctions, for which our main weak and strong convergence theorems apply.

4.1. Hierarchical minimization

Our contribution in this subsection concerns the hierarchical minimization problem:

$$\min_{x \in \underset{\mathbf{K}}{\operatorname{argmin}} \psi} \varphi(x), \quad (\mathbf{HMP})$$

where ψ and φ belong to $\Gamma_0(\mathbb{H})$ with $\mathbf{K} = \operatorname{dom} \varphi = \operatorname{dom} \psi$ a closed subset of \mathbb{H} . The above problem can be equivalently expressed as:

Find $\bar{x} \in \underset{\mathbf{K}}{\operatorname{argmin}} \psi$ such that

$$\varphi(\bar{x}) \leq \varphi(y), \quad \forall y \in \underset{\mathbf{K}}{\operatorname{argmin}} \psi. \quad (62)$$

Clearly, (62) can be viewed as a bilevel equilibrium problem (**BEP**) such that the associated bifunctions are defined for all $x, y \in \mathbf{K}$ by $f(x, y) = \psi(y) - \psi(x)$ and $g(x, y) = \varphi(y) - \varphi(x)$. In this case the bifunctions f and g are obviously monotone and upper hemicontinuous. Hence the theorem on the weak convergence applies whenever (12) is satisfied.

- **Weak convergence:** Without any loss of generality we assume that $\min_{\mathbf{K}} \psi = 0$. Set $M = \underset{\mathbf{K}}{\operatorname{argmin}} \psi$, and consider $\bar{\psi}(x) = \psi(x)$ if $x \in \mathbf{K}$, and $\bar{\psi}(x) = +\infty$ if $x \notin \mathbf{K}$; then $\bar{\psi}(x) \leq \delta_M(x)$ for all $x \in \mathbb{H}$. Using the reverse inequality for their Fenchel conjugates, we deduce $\bar{\psi}^*(p) \geq \sigma_M(p)$ for all $p \in \mathbb{H}$, and in view of Proposition 1.9, condition (12) becomes: $\forall u \in M$, for all $p \in N_M(u)$,

$$\sum_{n=1}^{+\infty} \lambda_n \beta_n \left[\bar{\psi}^* \left(\frac{2p}{\beta_n} \right) - \sigma_M \left(\frac{2p}{\beta_n} \right) \right] < +\infty. \quad (63)$$

Applying Theorem 2.3, and supposing that M is nonempty, $\liminf_{n \rightarrow +\infty} \lambda_n > 0$, $\lim_{n \rightarrow +\infty} \beta_n = +\infty$ and $0 \leq \alpha < \frac{1}{3}$, then the whole sequence $\{x_n\}$ generated by the algorithm (9) weakly converges to a point \bar{x} solution of (**HMP**).

Consider the particular case $\psi(x) = \frac{1}{2} \mathbf{d}(x, M)^2$, where $M \subset \mathbf{K}$ is a nonempty closed convex set and $\mathbf{d}(x, M) = \inf_{y \in M} \|x - y\|$. Then, $\bar{\psi}^*(p) - \sigma_M(p) = \frac{1}{2} \|p\|^2$ for all $p \in \mathbb{H}$.

Here, M is the minimum set of ψ , and then condition (63) is equivalent to $\sum_{n=1}^{+\infty} \frac{\lambda_n}{\beta_n} <$

$+\infty$.

Remark 3. We note that the condition (63) is simply the assumption from [18] in the framework of solving a variational inequality of the form

$$Ax + N_C(x) \ni 0,$$

where $A : \mathbb{H} \rightrightarrows \mathbb{H}$ is a maximally monotone operator and $C \subset \mathbb{H}$ is a closed convex set. For this problem, the authors in [18] obtained solutions by means of the convergence analysis of the trajectories of the following prox-penalization algorithm

$$x_n = (I + \lambda_n(A + \beta_n \partial \psi))^{-1} x_{n-1},$$

where $\{\beta_n\}$ and $\{\lambda_n\}$ are two sequences of nonnegative reals and $\psi : \mathbb{H} \rightarrow \mathbb{R} \cup \{+\infty\}$ acts as an external penalization function with respect to the constraint $x \in C$. Indeed, several ergodic and non ergodic convergence results have been justified for $\{x_n\}$ under the key assumption: for all $p \in R(N_C)$,

$$\sum_{n=1}^{+\infty} \lambda_n \beta_n \left[\psi^* \left(\frac{p}{\beta_n} \right) - \sigma_C \left(\frac{p}{\beta_n} \right) \right] < +\infty,$$

where $R(N_C)$ denotes the range of N_C .

- **Strong convergence:** To deduce the strong convergence of the algorithm (IPA) to a solution of (HMP), we'll have to add a strong monotonicity condition on the function g . However, when we set $g(x, y) = \varphi(y) - \varphi(x)$, the strong monotonicity of g is not ensured, so that we suppose in addition φ to be differentiable and strongly convex on \mathbf{K} , i.e., for some $\kappa > 0$ and for all $x, y \in \mathbf{K}$ and all $t \in [0, 1]$

$$\varphi(tx + (1-t)y) \leq t\varphi(x) + (1-t)\varphi(y) - \kappa t(1-t)\|x - y\|^2,$$

and we take $g(x, y) = \langle \nabla \varphi(x), y - x \rangle$, where $\nabla \varphi$ is the gradient of φ (we identify φ with $\bar{\varphi}(x) = \varphi(x)$ if $x \in \mathbf{K}$, and $\bar{\varphi}(x) = +\infty$ if $x \notin \mathbf{K}$). In this case g is strongly monotone and our inertial proximal scheme associated with the problem (62) is the following: $y_n := x_n + \alpha(x_n - x_{n-1})$ and $x_{n+1} \in \mathbf{K}$ such that

$$\beta_n(\psi(y) - \psi(x_{n+1})) + \langle \nabla \varphi(x_{n+1}), y - x_{n+1} \rangle + \frac{1}{\lambda_n} \langle x_{n+1} - y_n, y - x_{n+1} \rangle \geq 0, \quad \forall y \in \mathbf{K}. \quad (64)$$

Thus, the conclusion of Theorem 2.4 is valid whenever $\sum_{n=1}^{\infty} \lambda_n = +\infty$ and $0 \leq \alpha < \frac{1}{3}$.

4.2. Equilibrium problem under a saddle point constraint

Let $\mathbb{H}_1, \mathbb{H}_2$ be two real Hilbert spaces, $U \subset \mathbb{H}_1$ and $V \subset \mathbb{H}_2$ be nonempty closed convex sets, and let $L : U \times V \rightarrow \mathbb{R}$ be closed and convex-concave, i.e., for each $(u, v) \in U \times V$ the real-valued functions $L(\cdot, v)$ and $-L(u, \cdot)$ are convex and lower semicontinuous.

We consider the saddle point problem: find $(\bar{u}, \bar{v}) \in U \times V$ such that

$$L(\bar{u}, v) \leq L(\bar{u}, \bar{v}) \leq L(u, \bar{v}) \text{ for every } (u, v) \in U \times V, \quad (\mathbf{SP})$$

which is equivalent, see [33], to

$$\max_{v \in V} \inf_{u \in U} L(u, v) = \min_{u \in U} \sup_{v \in V} L(u, v) = L(\bar{u}, \bar{v}).$$

Setting $\mathbb{H} = \mathbb{H}_1 \times \mathbb{H}_2$, $\mathbf{K} = U \times V$, we define the bifunction $f : \mathbf{K} \times \mathbf{K} \rightarrow \mathbb{R}$ by:

$$f((u_1, v_1), (u_2, v_2)) := L(u_2, v_1) - L(u_1, v_2), \text{ for each } (u_1, v_1), (u_2, v_2) \in \mathbf{K}.$$

Let us observe that problems (\mathbf{SP}) and (\mathbf{EP}) are equivalent and we denote the solution set of (\mathbf{SP}) by S_L .

Using the definition of the Fitzpatrick transform \mathcal{F}_f , for all $(u_1, v_1), (u_2, v_2) \in \mathbf{K}$ we have:

$$\begin{aligned} \mathcal{F}_f((u_1, v_1), (u_2, v_2)) &= \sup_{(x, y) \in \mathbf{K}} \{ \langle u_2, x \rangle + \langle v_2, y \rangle + f((x, y), (u_1, v_1)) \} \\ &= \sup_{(x, y) \in \mathbf{K}} \{ \langle v_2, y \rangle + L(u_1, y) - L(x, v_1) + \langle u_2, x \rangle \} \\ &= \sup_{y \in V} \{ \langle v_2, y \rangle - (-L(u_1, y)) \} + \sup_{x \in U} \{ \langle u_2, x \rangle - L(x, v_1) \} \\ &= (-L(u_1, \cdot))^*(v_2) + (L(\cdot, v_1))^*(u_2). \end{aligned}$$

Therefore the condition (12) is satisfied when for all pairs $(u, v) \in \mathbf{S}_f$ and $(p, q) \in N_{\mathbf{S}_f}(u, v)$,

$$\sum_{n=1}^{+\infty} \lambda_n \beta_n \left[(-L(u, \cdot))^* \left(\frac{2q}{\beta_n} \right) + (L(\cdot, v))^* \left(\frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n}, \frac{2q}{\beta_n} \right) \right] < +\infty. \quad (65)$$

We consider two single-valued monotone operators A and B such that $\mathbf{K} \subset \text{dom } A \times \text{dom } B$ and $A \times B + N_{S_L}$ is a maximally monotone operator (see [34,35]). Furthermore we suppose that the solution set S_{VL} of $0 \in A\bar{x} \times B\bar{y} + N_{S_L}(\bar{x}, \bar{y})$ is nonempty. By $A \times B$, we mean the operator defined for $(u, v) \in \mathbb{H} = \mathbb{H}_1 \times \mathbb{H}_2$ by $(A \times B)(u, v) = Au \times Bv$. When the monotone operator $A \times B + N_{S_L}$ is maximally monotone, then

$$(\bar{x}, \bar{y}) \in S_{VL} \iff (\bar{x}, \bar{y}) \in S_L \text{ and } \langle A\bar{x}, u - \bar{x} \rangle + \langle B\bar{y}, v - \bar{y} \rangle \geq 0, \forall (u, v) \in S_L.$$

For each $(u_1, v_1), (u_2, v_2) \in \mathbf{K}$, let us set

$$g((u_1, v_1), (u_2, v_2)) := \langle Au_1, u_2 - u_1 \rangle + \langle Bv_1, v_2 - v_1 \rangle.$$

Then, our inertial proximal algorithm (\mathbf{IPA}) used for approaching a solution to the problem (\mathbf{BEP}) associated with the above bifunctions f and g , i.e., the problem of finding a solution in S_{VL} , takes the following form: for every $n \geq 1$, given current

iterates $(x_{n-1}^i, x_n^i) \in \mathbf{K}$, $i = 1, 2$, set $y_n^i = x_n^i + \alpha(x_n^i - x_{n-1}^i)$ and define $(x_{n+1}^1, x_{n+1}^2) \in \mathbf{K}$ in this way:

$$\left\{ \begin{array}{l} \text{for all } (u, v) \in U \times V, \\ \frac{1}{\lambda_n} \langle (x_{n+1}^1, x_{n+1}^2) - (y_n^1, y_n^2), (u, v) - (x_{n+1}^1, x_{n+1}^2) \rangle + \langle Ax_{n+1}^1, u - x_{n+1}^1 \rangle \\ \quad + \langle Bx_{n+1}^2, v - x_{n+1}^2 \rangle + \beta_n (L(u, x_{n+1}^2) - L(x_{n+1}^1, v)) \geq 0. \end{array} \right. \quad (66)$$

In this case, Theorems 2.3 and 2.4 can be summarized as follows:

Corollary 4.1. *Let $\{x_n^1, x_n^2\}$ be the sequence generated by (66). Under the hypothesis (65) and whenever $0 \leq \alpha < \frac{1}{3}$, $\liminf_{n \rightarrow +\infty} \lambda_n > 0$ and $\beta_n \rightarrow +\infty$, the weak convergence of $\{x_n^1, x_n^2\}$ to a solution of S_{VL} is ensured. Also, the strong convergence of $\{x_n^1, x_n^2\}$ to the unique element of S_{VL} is ensured when $0 \leq \alpha < \frac{1}{3}$, $\sum_{n=1}^{\infty} \lambda_n = +\infty$ and $A \times B$ is strongly monotone on \mathbf{K} .*

Next, let us give an example where the condition (65) is verified.

Example 4.2. Take $\mathbf{K} = [0, 1] \times [0, 1]$ and L the closed convex-concave function defined on \mathbf{K} by $L(u, v) = u^2(1 + v)$. Then, the set of saddle points of L , which is also the solution set \mathbf{S}_f , is $\mathbf{S}_f = \{0\} \times [0, 1]$.

We also have

$$\begin{aligned} (p, q) \in N_{(\{0\} \times [0, 1])}(0, v) &\Leftrightarrow (p, q)(s, t - v) \leq 0, \quad \forall (s, t) \in \{0\} \times [0, 1] \\ &\Leftrightarrow q(t - v) \leq 0, \quad \forall t \in [0, 1]; \end{aligned}$$

and then

$$N_{\mathbf{S}_f}(0, 0) = \mathbb{R} \times \mathbb{R}_-, \quad N_{\mathbf{S}_f}(0, 1) = \mathbb{R} \times \mathbb{R}_+ \quad \text{and} \quad N_{\mathbf{S}_f}(0, v) = \mathbb{R} \times \{0\} \quad \text{for every } v \in]0, 1[.$$

To ensure (65), we check

$$\begin{aligned} \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n}, \frac{2q}{\beta_n} \right) &= \sup_{v \in [0, 1]} \left\{ v \frac{2q}{\beta_n} \right\} = \begin{cases} \frac{2q}{\beta_n} & \text{if } q > 0 \\ 0 & \text{if } q \leq 0, \end{cases} \\ (-L(0, \cdot))^* \left(\frac{2q}{\beta_n} \right) &= \sup_{0 \leq s \leq 1} \left\{ \frac{2q}{\beta_n} s \right\} = \begin{cases} \frac{2q}{\beta_n} & \text{if } q > 0 \\ 0 & \text{if } q \leq 0, \end{cases} \end{aligned}$$

and

$$(L(\cdot, v))^* \left(\frac{2p}{\beta_n} \right) = \sup_{0 \leq s \leq 1} \left\{ \frac{2p}{\beta_n} s - (1 + v)s^2 \right\} = \frac{p^2}{(1 + v)\beta_n^2}.$$

Thus

$$\begin{aligned} &\sum_{n=1}^{+\infty} \lambda_n \beta_n \left[(-L(u, \cdot))^* \left(\frac{2q}{\beta_n} \right) + (L(\cdot, v))^* \left(\frac{2p}{\beta_n} \right) - \sigma_{\mathbf{S}_f} \left(\frac{2p}{\beta_n}, \frac{2q}{\beta_n} \right) \right] \\ &= \frac{p^2}{(1 + v)} \sum_{n=1}^{+\infty} \frac{\lambda_n}{\beta_n}, \end{aligned}$$

and then (65) is satisfied, whenever $\sum_{n=1}^{+\infty} \frac{\lambda_n}{\beta_n} < +\infty$.

5. Numerical experiment

In this section, we present a numerical experiment to illustrate the convergence of the proposed algorithm. Let us consider the constrained minimization problem (**HMP**), with

$$\mathbf{K} = \mathbb{R}^2, \quad \varphi(x) = \frac{1}{4}(x_1 - x_2 - 2)^2 \quad \text{and} \quad \psi(x) = \frac{1}{4}(x_1 + x_2 - 4)^2.$$

Since ψ is convex and smooth, the minimum set of ψ is $M = \underset{\mathbf{K}}{\operatorname{argmin}} \psi = \nabla \psi^{-1}(0, 0) = \{x = (x_1, x_2) \in \mathbb{R}^2 : x_2 = 4 - x_1\}$ and the solution set of the hierarchical problem $\min_M \varphi$ is $S = \{\bar{x}\} = \{(3, 1)\}$.

Let us evaluate $\frac{1}{2}\mathbf{d}(x, M)^2$ where

$$\mathbf{d}(x, M) = \inf_{y \in M} \|y - x\|_2 \quad \text{and} \quad \|(x_1, x_2)\|_2 := \sqrt{x_1^2 + x_2^2}.$$

For $x = (x_1, x_2) \in \mathbb{R}^2$, we have

$$\mathbf{d}(x, M)^2 = \inf_{y_1 \in \mathbb{R}} ((y_1 - x_1)^2 + (y_1 + x_2 - 4)^2) = \inf_{y_1 \in \mathbb{R}} \alpha(y_1).$$

Since the function $t \rightarrow \alpha(t) = (t - x_1)^2 + (t + x_2 - 4)^2$ is strongly convex and

$$\alpha'(\bar{y}_1) = 2(2\bar{y}_1 + x_2 - 4 - x_1) = 0 \Leftrightarrow \bar{y}_1 = \frac{1}{2}(x_1 - x_2 + 4),$$

we get

$$\mathbf{d}(x, M)^2 = \alpha(\bar{y}_1) = (\bar{y}_1 - x_1)^2 + (\bar{y}_1 + x_2 - 4)^2 = 2\psi(x),$$

which yields $\psi(x) = \frac{1}{2}\mathbf{d}(x, M)^2$. Thus condition (63) is equivalent to $\sum_{n=1}^{+\infty} \frac{\lambda_n}{\beta_n} < +\infty$.

Note that the associated bifunctions are defined for all $x, y \in \mathbf{K}$ by $f(x, y) = \psi(y) - \psi(x)$ and $g(x, y) = \varphi(y) - \varphi(x)$, that f and g are monotone and that weak and strong convergences coincide in finite dimension.

By using the proximal operator of $\varphi + \beta_n \psi$, the drawing in Figure 1 displays the asymptotic behavior of the trajectories $x_n = (y_n, z_n)$ from the initial values $(y_0, z_0) = (0, 0.5)$ and $(y_1, z_1) = (0, 0.5)$ with $\alpha = 0.1$, $\lambda_n = \frac{1}{n}$ and different values of β_n . We also use the iterate error $\|x_n - \bar{x}\|_2$ as a measure to describe the computational performance of our algorithm. The numerical results in Figure 2 illustrate the rate of convergence of $\|x_n - \bar{x}\|_2$ for different choices of β_n and $\alpha = 0.1$, while Figure 3 displays the convergence rate of $\|x_n - \bar{x}\|_2$ for different choices of α and $\beta_n = (1 + n)$.

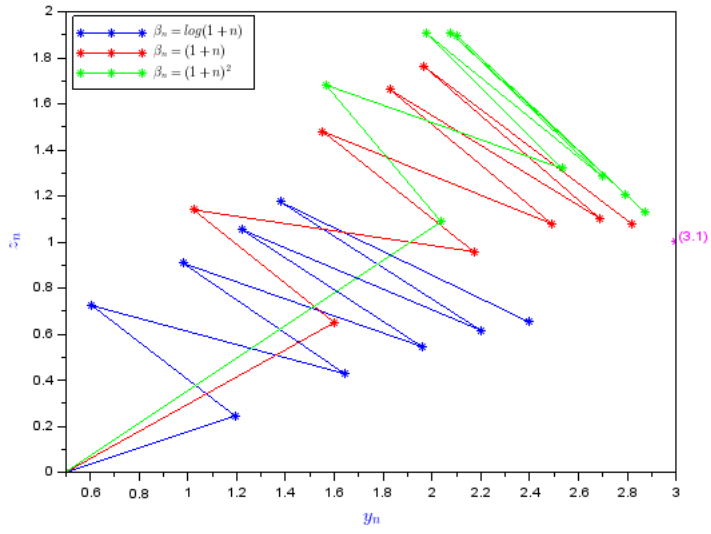


Figure 1. The asymptotic behavior of the trajectories $x_n = (y_n, z_n)$.

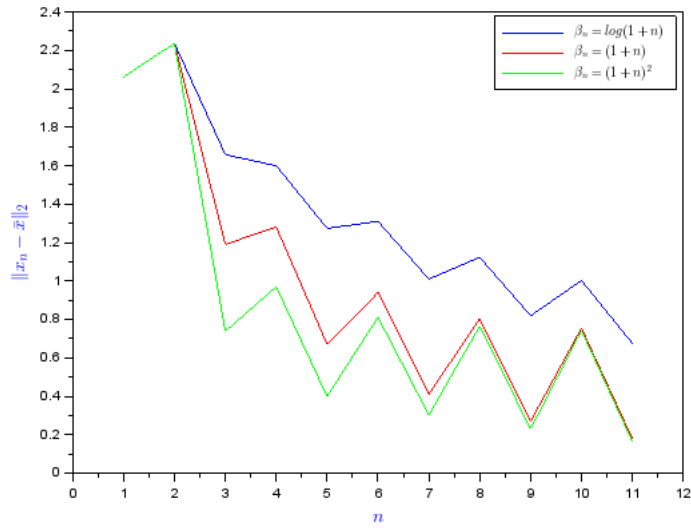


Figure 2. The rate of convergence of $\|x_n - \bar{x}\|_2$ for $\alpha = 0.1$.

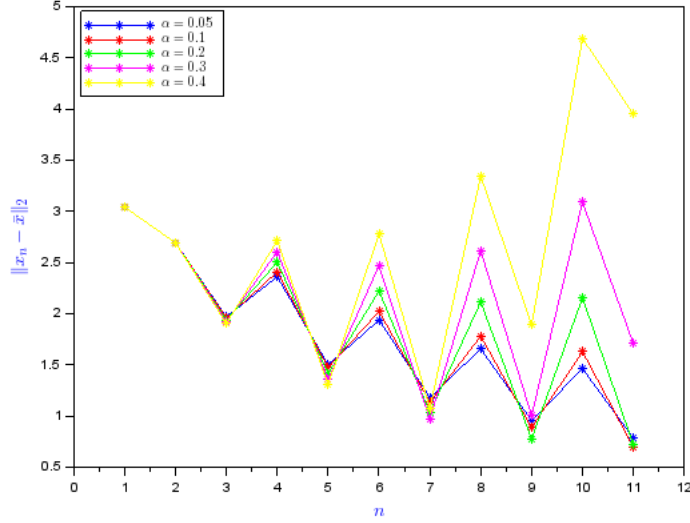


Figure 3. The convergence rate of $\|x_n - \bar{x}\|_2$ for $\beta_n = (1 + n)$.

We note in Figure 2, that when β_n increases then the rate of convergence of $\|x_n - \bar{x}\|_2$ rapidly increases to 0, while in Figure 3, the constant coefficient α acts inversely on the speed of convergence of $\|x_n - \bar{x}\|_2$, (the convergence gets worst as the values of α exceed $\frac{1}{3}$), which confirms the importance of taking $\alpha < \frac{1}{3}$ in our theoretical results.

We note that all codes in this digital test are written in SCILAB-6.1.

6. Concluding Remark

In this paper, we presented two inertial type methods for solving bilevel monotone equilibrium problems in Hilbert spaces. Our analysis provides interesting convergence results of the trajectory generated by the proposed algorithms under natural assumptions. The results obtained can be seen as an extension and an improvement of some known results in the literature. In particular, the geometric assumption (12) shows that, as conjectured in [16], the restrictive assumption $\|x_{n+1} - x_n\| = o(\epsilon_n)$ may be removed via the introduction of a notion of conditioning for equilibrium bifunctions. We illustrate this assumption with two concrete particular cases and conclude this work by a numerical experiment, which shows that, with a suitable choice of the parameters, the convergence conditions are satisfied and the proposed iterative method succeeds in approximating a solution to bilevel equilibrium problems.

Finally, we note that, to the best of our knowledge, our algorithm (**IPA**) seems to be the first introduced inertial proximal scheme for solving (**BEP**) and then several extensions of our main results may be analyzed. In particular, an interesting direction of future research will be to obtain the above weak convergence result without condition (12) and also to develop new splitting inertial proximal algorithms for solving bilevel equilibrium problems.

7. Funding

Research of Aïcha Balhag was supported by the FEDER and the EIPHI Graduate School (contract ANR-17-EURE-0002) and research of Michel Théra was supported by a public grant as part of the Investissement d’avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH.

References

- [1] Blum E, Oettli W. From optimization and variational inequalities to equilibrium problems. *Math Student*. 1994;63(1-4):123–145.
- [2] Brézis H, Nirenberg L, Stampacchia G. A remark on Ky Fan’s minimax principle. *Boll Un Mat Ital* (4). 1972;6:293–300.
- [3] Chadli O, Chbani Z, Riahi H. Equilibrium problems with generalized monotone bifunctions and applications to variational inequalities. *J Optim Theory Appl*. 2000;105(2):299–323. Available from: <https://doi.org/10.1023/A:1004657817758>.
- [4] Mosco U. Implicit variational problems and quasi variational inequalities. In: *Nonlinear operators and the calculus of variations (Summer School, Univ. Libre Bruxelles, Brussels, 1975)*. Springer, Berlin; 1976. *Lecture Notes in Math.*, Vol. 543; p. 83–156.
- [5] Moudafi A. Proximal point algorithm extended to equilibrium problems. *J Nat Geom*. 1999;15(1-2):91–100.
- [6] Moudafi A. Second-order differential proximal methods for equilibrium problems. *JIPAM J Inequal Pure Appl Math*. 2003;4(1):Article 18, 7.
- [7] Antipin AS. Convergence and estimates for the rate of convergence of proximal methods to fixed points of extremal mappings. *Zh Vychisl Mat i Mat Fiz*. 1995;35(5):688–704.
- [8] Burachik R, Kassay G. On a generalized proximal point method for solving equilibrium problems in Banach spaces. *Nonlinear Anal*. 2012;75(18):6456–6464. Available from: <https://doi.org/10.1016/j.na.2012.07.020>.
- [9] Moudafi A, Théra M. Proximal and dynamical approaches to equilibrium problems. In: *Ill-posed variational problems and regularization techniques (Trier, 1998)*. (Lecture Notes in Econom. and Math. Systems; Vol. 477). Springer, Berlin; 1999. p. 187–201. Available from: https://doi.org/10.1007/978-3-642-45780-7_12.
- [10] Cotrina J, Théra M, Zúñiga J. An existence result for quasi-equilibrium problems via Ekeland’s variational principle. *J Optim Theory Appl*. 2020;187(2):336–355. Available from: <https://doi.org/10.1007/s10957-020-01764-0>.
- [11] Dempe S. Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. *Optimization*. 2003;52(3):333–359. Available from: <https://doi.org/10.1080/0233193031000149894>.
- [12] Ait Mansour M, Mazgouri Z, Riahi H. A dynamical approach for the quantitative stability of parametric bilevel equilibrium problems and applications. *Optimization*. 2022; 71(5):1389–1408. Available from: <https://doi.org/10.1080/02331934.2021.1981892>.
- [13] Bento GC, Cruz Neto JX, Lopes JO, et al. Generalized proximal distances for bilevel equilibrium problems. *SIAM J Optim*. 2016;26(1):810–830. Available from: <https://doi.org/10.1137/140975589>.
- [14] Chbani Z, Mazgouri Z, Riahi H. From convergence of dynamical equilibrium systems to bilevel hierarchical Ky Fan minimax inequalities and applications. *Minimax Theory Appl*. 2019;4(2):231–270.
- [15] Chbani Z, Riahi H. Weak and strong convergence of prox-penalization and splitting algo-

- rithms for bilevel equilibrium problems. *Numer Algebra Control Optim.* 2013;3(2):353–366. Available from: <https://doi.org/10.3934/naco.2013.3.353>.
- [16] Moudafi A. Proximal methods for a class of bilevel monotone equilibrium problems. *J Global Optim.* 2010;47(2):287–292. Available from: <https://doi.org/10.1007/s10898-009-9476-1>.
- [17] Thuy LQ, Hai TN. A projected subgradient algorithm for bilevel equilibrium problems and applications. *J Optim Theory Appl.* 2017;175(2):411–431. Available from: <https://doi.org/10.1007/s10957-017-1176-2>.
- [18] Attouch H, Czarnecki M, Peypouquet J. Prox-penalization and splitting methods for constrained variational problems. *SIAM J Optim.* 2011;21(1):149–173. Available from: <https://doi.org/10.1137/100789464>.
- [19] Riahi H, Chbani Z, Loumi MT. Weak and strong convergences of the generalized penalty Forward-Forward and Forward-Backward splitting algorithms for solving bilevel hierarchical pseudomonotone equilibrium problems. *Optimization.* 2018;67(10):1745–1767. Available from: <https://doi.org/10.1080/02331934.2018.1490957>.
- [20] Alvarez F, Attouch H. An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. *Set-Valued Var Anal.* 2001;9(1-2):3–11. Available from: <https://doi.org/10.1023/A:1011253113155>.
- [21] Chbani Z, Riahi H. Weak and strong convergence of an inertial proximal method for solving Ky Fan minimax inequalities. *Optim Lett.* 2013;7(1):185–206. Available from: <https://doi.org/10.1007/s11590-011-0407-y>.
- [22] Hieu DV, Gibali A. Strong convergence of inertial algorithms for solving equilibrium problems. *Optim Lett.* 2020;14(7):1817–1843. Available from: <https://doi.org/10.1007/s11590-019-01479-w>.
- [23] Chbani Z, Riahi H. Variational principles for monotone and maximal bifunctions. *Serdica Math J.* 2003;29(2):159–166. Available from: <https://doi.org/10.4099/math1924.29.159>.
- [24] Opial Z. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull Amer Math Soc.* 1967;73:591–597. Available from: <https://doi.org/10.1090/S0002-9904-1967-11761-0>.
- [25] Bauschke HH, Combettes PL. *Convex analysis and monotone operator theory in Hilbert spaces.* Springer, New York; 2011. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC; Available from: <https://doi.org/10.1007/978-1-4419-9467-7>.
- [26] Alizadeh MH, Hadjisavvas N. On the Fitzpatrick transform of a monotone bifunction. *Optimization.* 2013;62(6):693–701. Available from: <https://doi.org/10.1080/02331934.2011.653975>.
- [27] Boţ RI, Grad S. Approaching the maximal monotonicity of bifunctions via representative functions. *J Convex Anal.* 2012;19(3):713–724.
- [28] Fitzpatrick S. Representing monotone operators by convex functions. In: *Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988).* (Proc. Centre Math. Anal. Austral. Nat. Univ.; Vol. 20). Austral. Nat. Univ., Canberra; 1988. p. 59–65.
- [29] Attouch H, Riahi H, Théra M. Somme ponctuelle d’opérateurs maximaux monotones. *Serdica Math J.* 1996;22(3):267–292.
- [30] Riahi H. On the maximality of the sum of two maximal monotone operators. *Publ Mat.* 1990;34(2):269–271. Available from: https://doi.org/10.5565/PUBLMAT_34290_05.
- [31] Boţ RI, Csetnek ER. Forward-backward and Tseng’s type penalty schemes for monotone inclusion problems. *Set-Valued Var Anal.* 2014;22(2):313–331. Available from: <https://doi.org/10.1007/s11228-014-0274-7>.
- [32] Boţ RI, Csetnek ER, Nimana N. An inertial proximal-gradient penalization scheme for constrained convex optimization problems. *Vietnam J Math.* 2018;46(1):53–71. Available from: <https://doi.org/10.1007/s10013-017-0256-9>.
- [33] Ekeland I, Temam R. *Convex analysis and variational problems.* North-Holland Publish-

- ing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York; 1976. Studies in Mathematics and its Applications, Vol. 1; translated from the French.
- [34] Rockafellar RT. On the maximal monotonicity of subdifferential mappings. Pacific J Math. 1970;33:209–216. Available from: <http://projecteuclid.org/euclid.pjm/1102977253>.
- [35] Rockafellar RT. On the maximality of sums of nonlinear monotone operators. Trans Amer Math Soc. 1970;149:75–88. Available from: <https://doi.org/10.2307/1995660>.