symmetry

MDPI

Article

A Subgradient-Type Extrapolation Cyclic Method for Solving an Equilibrium Problem over the Common Fixed-Point Sets

Porntip Promsinchai [1,2] and Nimit Nimana [3,*]

[1] Department of Mathematics, Faculty of Science, Naresuan University, Phitsanulok 65000, Thailand; porntippr56@nu.ac.th
[2] Center of Excellence in Nonlinear Analysis and Optimization, Faculty of Science, Naresuan University, Phitsanulok 65000, Thailand
[3] Department of Mathematics, Faculty of Science, Khon Kaen University, Khon Kaen 40002, Thailand
* Correspondence: nimitni@kku.ac.th

Abstract: In this paper, we consider the solving of an equilibrium problem over the common fixed set of cutter mappings in a real Hilbert space. To this end, we present a subgradient-type extrapolation cyclic method. The proposed method is generated based on the ideas of a subgradient method and an extrapolated cyclic cutter method. We prove a strong convergence of the method provided that some suitable assumptions of step-size sequences are assumed. We finally show the numerical behavior of the proposed method.

Keywords: convergence analysis; cutter; equilibrium problem; fixed point; monotone operator

check for updates

Citation: Promsinchai, P.; Nimana, N. A Subgradient-Type Extrapolation Cyclic Method for Solving an Equilibrium Problem over the Common Fixed-Point Sets. *Symmetry* **2022**, *14*, 992. https://doi.org/10.3390/sym14050992

Academic Editor: Eulalia Martínez Molada

Received: 11 April 2022
Accepted: 10 May 2022
Published: 12 May 2022

1. Introduction

Let $\mathcal{H}$ be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\|\cdot\|$. In this paper, we present an iterative method to the equilibrium problem over the intersection of fixed-point set:

Problem 1 (BEP). *Let $T_i : \mathcal{H} \to \mathcal{H}$, $i = 1, 2, \ldots, m$, be cutters with $\bigcap_{i=1}^{m} \operatorname{Fix} T_i \neq \emptyset$, and let $f : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ be a bifunction satisfying $f(x, x) = 0$ for all $x \in \mathcal{H}$. Then, our objective is to find a point $\bar{u} \in \bigcap_{i=1}^{m} \operatorname{Fix} T_i$ such that*

$$f(\bar{u}, y) \geq 0 \qquad \text{for all } y \in \bigcap_{i=1}^{m} \operatorname{Fix} T_i,$$

where $\operatorname{Fix} T_i := \{x \in \mathcal{H} : T_i x = x\}$ denotes the set of fixed points of T_i.

The equilibrium problem, which was first introduced by Fan [1], includes many problems as particular cases, for example, the fixed-point problem, the variational inequality, the optimization problem, the saddle point problem, the Nash equilibrium problem in non-cooperative games, and others; see, for instance, [2–5] and references therein.

The equilibrium problems over the fixed-point set have been considered in many articles; see, for instance, [6–10] and references therein. The computational algorithms for solving these kinds of problems have been studied and developed by using the idea of the methods for equilibrium problems and the iterative schemes for fixed-point problems. In particular, Iiduka and Yamada [6] considered the equilibrium problems over the fixed-point set of a firmly nonexpansive mapping and presented a subgradient-type method for solving the considered problems. They showed the convergence of their method and applied the method to the Nash equilibrium problems. After that, the equilibrium problems over the common fixed-point of nonexpansive mappings were considered by Duc and Muu in [7]. They proposed the splitting algorithm, which was updated based on the idea of the classical gradient method and the Krasnosel'skii–Mann method and presented the strong convergence of the presented algorithm. Recently, Thuy and Hai [8] considered the bilevel equilibrium problems and proposed the projected subgradient algorithm to solve the considered problem. They exhibited the strong convergence of the proposed method and applied it to the equilibrium problems over the fixed-point set of a nonexpansive mapping. We notice that the aforementioned literature is considered in the case of the equilibrium problems over the fixed-point set of nonexpansive mappings.

Let us focus on the constrained set of **BEP**. Now, let $T_i : \mathcal{H} \to \mathcal{H}$, $i = 1, 2, \dots, m$, be cutter operators. The common fixed-point problem is to find a point

$$x^* \in \bigcap_{i=1}^{m} \operatorname{Fix} T_i \neq \emptyset$$

The well-known methods of finding a point that belongs to the intersection of fixed-point sets are initially motivated by the cyclic projection method, which was introduced by Kaczmarz [11]. After that, the convergence of cyclic projection-type methods are investigated in several directions and their convergence results are guaranteed under the operators' assumptions, such as cutters or nonexpansive operators, see [12–16] and references therein. In particular, Bauschke and Combettes [17] proposed the cyclic cutter method and showed a weak convergence of the proposed method. In [18], Cegielski and Censor presented the extrapolated cyclic cutter method, which is an acceleration of the cyclic cutter method by imposing an appropriate step-size function to the method. Indeed, let $T_i : \mathcal{H} \to \mathcal{H}$, $i = 1, 2, \dots, m$, be cutters with $\bigcap_{i=1}^{m} \operatorname{Fix} T_i \neq \emptyset$, we define the step-size function $\sigma : \mathcal{H} \to (0, \infty)$ as follows:

$$\sigma(x) := \begin{cases} \dfrac{\sum_{i=1}^{m} \langle Tx - S_{i-1}x, S_i x - S_{i-1}x \rangle}{\|Tx - x\|^2}, & \text{for } x \notin \bigcap_{i=1}^{m} \operatorname{Fix} T_i, \\ 1, & \text{otherwise}, \end{cases} \tag{1}$$

where the operator T, S_0 and $S_i, i = 1, 2, \dots, m$ are defined as

$$T := T_m T_{m-1} \cdots T_1, \qquad S_0 := Id, \qquad \text{and} \qquad S_i := T_i T_{i-1} \cdots T_1. \tag{2}$$

It was shown that the extrapolated cyclic cutter method weakly converges provided that the cutter operators T_i satisfied the demi-closedness principle for all $i = 1, 2, \dots, m$. Note that, for some practical problem, the value of the extrapolation function may be huge, which lead to some numerical instabilities. To keep away from these instabilities, Cegielski and Nimana [19] proposed the modified extrapolated subgradient projection method for solving the convex feasibility problem, which is a particular situation of common fixed-point problem, and established the convergence of the proposed method as well as demonstrated the performance of the method by the numerical results. After that, the authors in [20] also utilized the idea of the extrapolated cyclic cutter method for dealing with the variational inequality problem with common fixed-point constraints. It can be noted that from [19,20], the iterative methods with extrapolated cyclic cutter terms achieve not only some numerical superiorities to utilizing the classical cyclic cutter scheme but also guarantee the boundedness of the generated sequence, see [20] for further discussions.

In this paper, we propose an iterative algorithm called the Subgradient-type extrapolation cyclic method for solving the equilibrium problems over the intersection of fixed-point sets of cutter operators. The proposed algorithm can be considered as a combination of the subgradient iterative scheme for equilibrium problems in [8] and the extrapolated cyclic cutter method for the intersection of fixed-point sets of cutter operators in [18]. Using the cutter operators and assumptions, we investigate the convergence of the presented algorithm. Moreover, we also present a numerical result of our presented method to illustrate the efficiency of the method.

This paper is organized as follows. In Section 2, we recall some definitions and tools which are needed for our convergence work. In Section 3, we present the *subgradient-type extrapolation cyclic method* for finding the solution of **BEP**. We subsequently present the convergence result in this section. In Section 4, the efficacy of the subgradient-type extrapolation cyclic method is illustrated by numerical experiments of the solving equilibrium problem governed by the positive definite symmetric matrices over the common fixed-point set. Finally, we give some concluding remarks in Section 5.

2. Preliminaries

In section, we collect some basic definitions, properties, and useful tools for our work. The readers can consult the books [16,21] for more details.

We denote by Id the identity operator on a real Hilbert space $\mathcal{H}$. For a sequence $\{x^n\}_{n=1}^{\infty}$, the strong and weak convergences of a sequence $\{x^n\}_{n=1}^{\infty}$ to a point $x \in \mathcal{H}$ are defined by the expression $x^n \to x$ and $x^n \rightharpoonup x$, respectively.

In what follows, we recall some definitions and properties of the operator that will be referred to in our analysis.

Definition 1 ([16])**.** *Let $T : \mathcal{H} \to \mathcal{H}$ be an operator having a fixed point. The operator T is called*

(i) *quasi-nonexpansive, if*
$$\|Tx - z\| \leq \|x - z\|,$$
for all $x \in \mathcal{H}$ and $z \in \operatorname{Fix} T$,

(ii) *η-strongly quasi-nonexpansive, if there exists $\eta \geq 0$,*
$$\|Tx - z\|^2 \leq \|x - z\|^2 - \eta\|Tx - x\|^2,$$
for all $x \in \mathcal{H}$ and $z \in \operatorname{Fix} T$,

(iii) *cutter , if*
$$\langle x - Tx, z - Tx\rangle \leq 0,$$
for all $x \in \mathcal{H}$ and $z \in \operatorname{Fix} T$.

Lemma 1 ([16] (Remark 2.1.31 and Theorem 2.1.39))**.** *Let $T : \mathcal{H} \to \mathcal{H}$ be an operator having a fixed point. Then the following statements are equivalent:*

(i) *T is cutter.*
(i) *$\langle Tx - x, z - x\rangle \geq \|Tx - x\|^2$ for all $x \in \mathcal{H}$ and for all $z \in \operatorname{Fix} T$.*
(ii) *T is 1-strongly quasi-nonexpansive.*

Definition 2 ([16] (Definition 3.2.6))**.** *Let $T : \mathcal{H} \to \mathcal{H}$ be an operator having a fixed point. The operator T is said to satisfy the demi-closedness (DC) principle if for every sequence $\{x^n\} \subset \mathcal{H}$, $x^n \rightharpoonup u \in \mathcal{H}$ and $\|Tx^n - x^n\| \to 0$, we have $u \in \operatorname{Fix} T$.*

Definition 3 ([16] (Definition 2.1.2))**.** *Let $T : \mathcal{H} \to \mathcal{H}$ be an operator and $\lambda \in [0, 2]$ be given. We define the relaxation of the operator T by*
$$T_\lambda := (1 - \lambda)Id + \lambda T,$$
and we call λ a relaxation parameter.

Next, we recall the definition of a generalization of relaxation of an operator.

Definition 4 ([16] (Definition 2.4.1))**.** *Let $T : \mathcal{H} \to \mathcal{H}$ be an operator, $\lambda \in [0, 2]$ and $\sigma : \mathcal{H} \to (0, \infty)$. We define the operator $T_{\sigma,\lambda} : \mathcal{H} \to \mathcal{H}$ by*

$$T_{\sigma,\lambda}x := x + \lambda\sigma(x)(Tx - x).$$

This operator T is called a generalized relaxation of the operator T, the value λ is called a relaxation parameter and the function σ is called a step-size function. The operator $T_{\sigma,\lambda}$ is called an extrapolation of T_λ if the function $\sigma(x) \geq 1$ for every $x \in \mathcal{H}$.

We notice that, if $\sigma(x) = 1$, for every $x \in \mathcal{H}$, then $T_{\sigma,\lambda} = T_\lambda$. Note that $T_\sigma := T_{\sigma,1}$. Then, for every $x \in \mathcal{H}$, the following relations hold

$$T_{\sigma,\lambda}x - x = \lambda\sigma(x)(Tx - x) = \lambda(T_\sigma x - x), \tag{3}$$

and for any $\lambda \neq 0$, we have

$$\mathrm{Fix}\, T_{\sigma,\lambda} = \mathrm{Fix}\, T_\sigma = \mathrm{Fix}\, T.$$

Next, we provide an important lemma of the step-size function for proving the convergence result.

Lemma 2 ([16] (Section 4.10))**.** *Let $T_i : \mathcal{H} \to \mathcal{H}, i = 1, 2, \ldots, m$, be cutters with $\bigcap_{i=1}^{m} \mathrm{Fix}\, T_i \neq \emptyset$ and denote the operator T, S_0 and $S_i, i = 1, 2, \ldots, m$ as in (2). Let the function $\sigma : \mathcal{H} \to (0, \infty)$ be given by (1). Then the following statements are true:*

(i) *For every $x \notin \mathrm{Fix}\, T$, it is true that*

$$\sigma(x) \geq \frac{\frac{1}{2}\sum_{i=1}^{m} \|S_i x - S_{i-1}x\|^2}{\|Tx - x\|^2} \geq \frac{1}{2m}.$$

(ii) *The operator T_σ is a cutter.*

Now, we recall a notion and some properties of a diagonal subdifferential which will be used in this work.

A function $h : \mathcal{H} \to \mathbb{R}$ is said to be subdifferentiable at $x^0 \in \mathcal{H}$ if there exists a vector $w \in \mathcal{H}$ such that

$$h(x) \geq h(x^0) + \langle w, x - x^0 \rangle, \qquad \forall x \in \mathcal{H}.$$

The vector w is called a subgradient of the function h at x^0. The collection of all such vectors constitute the subdifferential of h at x^0 and is denoted by $\partial h(x^0)$.

Let $f : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ be a bifunction which is convex in the second argument, that is, the function $f(x, \cdot) : \mathcal{H} \to \mathbb{R}$ is convex at x, for all $x \in \mathcal{H}$. Then, the set of all subgradient of $f(x, \cdot)$ at x is called the diagonal subdifferential and is denoted by $\partial_2 f(x, x) := \partial f(x, \cdot)(x)$. The reader can find the notion of the diagonal subdifferential in [22], for more detail.

We end this section by recalling some technical lemmas that are important tools in proving our convergence result.

Lemma 3 ([23] (Lemma 3.1))**.** *Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be sequences of nonnegative real numbers such that*

$$a_{n+1} \leq a_n + b_n.$$

If $\sum_{n=1}^{\infty} b_n < \infty$, then $\lim_{n\to\infty} a_n$ exists.

Lemma 4 ([24] (Lemma 3.1))**.** *Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of real numbers such that there exists a subsequence $\{a_{n_j}\}_{j=1}^{\infty}$ of $\{a_n\}_{n=1}^{\infty}$ with $a_{n_j} < a_{n_{j+1}}$ for all $j \in \mathbb{N}$. If, for all $n \geq n_0$, we define*

$$\nu(n) = \max\{k \in [n_0, n] : a_k < a_{k+1}\},$$

then the following hold:

(i) $\{\nu(n)\}_{n \geq n_0}$ *is non-decreasing.*
(ii) $\lim\limits_{n\to\infty} \nu(n) = \infty$.
(iii) $a_{\nu(n)} \leq a_{\nu(n)+1}$ *and* $a_n \leq a_{\nu(n)+1}$ *for every* $n \geq n_0$.

3. Algorithm and Its Convergence Result

In this section, we firstly propose the subgradient-type extrapolation cyclic method for solving **BEP**. Subsequently, we present useful lemmas and prove the main convergence theorem.

Remark 1. *(i) When the number* $m = 1$ *and* $\sigma(x^n) = 1$*, Algorithm 1 becomes Algorithm 2 considered in [8]. Moreover, it is worth noting that the class of operator considered in this work is different from [8]. In fact, we consider the cutter property of* T_i*, whereas the nonexpansiveness of* T *is assumed in [8].*

(ii) If the function $f(\cdot,\cdot) = 0$*, Algorithm 1 is reduced to*

$$x^{n+1} = x^n - \frac{\lambda_n}{\eta_n}\sigma(x^n)(x^n - Tx^n),$$

where $\eta_n = \max\{\mu, \|d^n\|\}$ *for all* $n \geq 1$*. In the case when* $\eta_n = 1$ *for all* $n \geq 1$*, this scheme is related to the extrapolated cyclic cutter proposed by [18]. Moreover, this scheme is also related to the work of Cegielski and Nimana [19] for solving a convex feasibility problem when the operator* T_m *is omitted in their paper.*

The following assumption relating to the convergence of Algorithm 1 is assumed throughout this work.

Algorithm 1: Subgradient-type extrapolation cyclic method

Initialization: Given the positive real sequences $\{\alpha_n\}_{n=1}^{\infty}$ and $\{\lambda_n\}_{n=1}^{\infty}$. Choose $\mu \in (0, +\infty)$ and $x^1 \in \mathcal{H}$ arbitrarily.

Step 1. For given $x^n \in \mathcal{H}$, compute the step size as

$$\sigma(x^n) := \begin{cases} \dfrac{\sum_{i=1}^{m}\langle Tx^n - S_{i-1}x^n, S_i x^n - S_{i-1}x^n\rangle}{\|Tx^n - x^n\|^2}, & \text{for } x^n \notin \bigcap\limits_{i=1}^{m} \operatorname{Fix} T_i, \\ 1, & \text{otherwise.} \end{cases}$$

Step 2. Update the next iterate x^{n+1} as

$$\begin{aligned} d^n &:= \sigma(x^n)(x^n - Tx^n) + \alpha_n w^n; \text{ where } w^n \in \partial_2 f(x^n, x^n), \\ \eta_n &:= \max\{\mu, \|d^n\|\}, \\ x^{n+1} &:= x^n - \frac{\lambda_n}{\eta_n} d^n. \end{aligned}$$

Put $n := n+1$ and go to **Step 1**.

Assumption 1. *Assume that*

(A1) The bifunction f *is* ρ*-strongly monotone on* $\mathcal{H}$*, that is, there exists a constant* $\rho > 0$ *satisfying*

$$f(x, y) + f(y, x) \leq -\rho\|x - y\|^2, \quad \forall x, y \in \mathcal{H}.$$

(A2) *For each $x \in \mathcal{H}$, the function $f(x,\cdot)$ is convex, subdifferentiable and lower semicontinuous on $\mathcal{H}$.*

(A3) *The function $x \mapsto \partial_2 f(x,x)$ is bounded on a bounded subset of $\mathcal{H}$.*

(A4) *The sequences $\{\lambda_n\}_{n=1}^{\infty}$ and $\{\alpha_n\}_{n=1}^{\infty}$ satisfy $\sum\limits_{n=1}^{\infty} \lambda_n = \infty$, $\sum\limits_{n=1}^{\infty} \lambda_n^2 < \infty$, $\sum\limits_{n=1}^{\infty} \alpha_n\lambda_n = \infty$, and $\lim\limits_{n\to\infty} \alpha_n = 0$.*

Remark 2. *(i) If the whole space $\mathcal{H}$ is finite dimensional, the assumption that, for all $x \in \mathcal{H}$, the function $f(x,\cdot)$ is subdifferentiable and weakly lower semicontinuous in (A2) can be omitted. This is because, in the finite dimensional setting, the convexity implies the continuity of a function.*

(ii) The convexity of the function $f(x,\cdot)$ implies that the lower semicontinuity is equivalent to the weak lower semicontinuity of the function $f(x,\cdot)$ for all $x \in \mathcal{H}$.

(iii) If the whole space $\mathcal{H}$ is finite dimensional, by invoking the assumption (A2), we have the diagonal subdifferential $\partial_2 f(x^n,x^n) := \partial f(x^n,\cdot)(x^n)$ is nonempty for all $n \in \mathbb{N}$. Moreover, in this case, the assumption (A3) can be omitted, see [21] (Proposition 16.20).

(iv) An example of the corresponding step-size sequences in (A4) is the positive real sequences $\{\alpha_n\}_{n=1}^{\infty}$ and $\{\lambda_n\}_{n=1}^{\infty}$ given by

$$\alpha_n := \frac{\alpha}{(n+1)^a} \quad \text{and} \quad \lambda_n := \frac{\lambda}{(n+1)^b},$$

where $\alpha,\lambda > 0$ and $a,b > 0$ with $b > 0.5$ and $a+b \leq 1$. In fact, since $0 < a+b \leq 1$ and $b > 0.5$, we have $0.5 < b < 1$ and then $\sum_{n=1}^{\infty} \lambda_n = \sum_{n=1}^{\infty} \frac{\lambda}{(n+1)^b} > \lambda \sum_{n=1}^{\infty} \frac{1}{(n+1)} = \infty$. Furthermore, since $1 < 2b < 2$, we have $\sum_{n=1}^{\infty} \lambda_n^2 = \sum_{n=1}^{\infty} \frac{\lambda^2}{(n+1)^{2b}} < \infty$. We note that $\sum_{n=1}^{\infty} \alpha_n\lambda_n = \sum_{n=1}^{\infty} \frac{\alpha}{(n+1)^a}\frac{\lambda}{(n+1)^b} = \alpha\lambda\sum_{n=1}^{\infty} \frac{1}{(n+1)^{a+b}} = \infty$. Moreover, we have that $\lim\limits_{n\to\infty} \alpha_n = \lim\limits_{n\to\infty} \frac{\alpha}{(n+1)^a} = 0$.

The following lemma states the important relation of the generated iterates.

Lemma 5. *Let $\{x^n\}_{n=1}^{\infty}$ be the sequence generated by Algorithm 1. Then, for every $n \in \mathbb{N}$ and $u \in \bigcap_{i=1}^{m} \operatorname{Fix} T_i$, it holds that*

$$\|x^{n+1}-u\|^2 \leq \|x^n-u\|^2 - \frac{\lambda_n}{4m\eta_n}\sum_{i=1}^{m}\|S_i x^n - S_{i-1}x^n\|^2 - \frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n-u, w^n\rangle + \lambda_n^2.$$

Proof. Let $n \in \mathbb{N}$ be fixed. Now, let us note that

$$\begin{aligned}
\|x^{n+1}-u\|^2 &= \left\|x^n - \frac{\lambda_n}{\eta_n}d^n - u\right\|^2 \\
&= \|x^n-u\|^2 - \frac{2\lambda_n}{\eta_n}\langle x^n-u, d^n\rangle + \left(\frac{\lambda_n}{\eta_n}\|d^n\|\right)^2 \\
&= \|x^n-u\|^2 - \frac{2\lambda_n}{\eta_n}\langle x^n-u, \sigma(x^n)(x^n-Tx^n)+\alpha_n w^n\rangle + \left(\frac{\lambda_n}{\eta_n}\|d^n\|\right)^2 \\
&= \|x^n-u\|^2 - \frac{2\lambda_n}{\eta_n}\langle x^n-u, \sigma(x^n)(x^n-Tx^n)\rangle - \frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n-u, w^n\rangle + \left(\frac{\lambda_n}{\eta_n}\|d^n\|\right)^2.
\end{aligned}$$

By using the properties of T_σ in (3), Lemma 1 and Lemma 2, we note that

$$\begin{aligned}
\|x^{n+1}-u\|^2 = &\ \|x^n-u\|^2 - \frac{2\lambda_n}{\eta_n}\langle u-x^n, T_\sigma x^n - x^n\rangle - \frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n-u, w^n\rangle + \left(\frac{\lambda_n}{\eta_n}\|d^n\|\right)^2 \\
\leq &\ \|x^n-u\|^2 - \frac{2\lambda_n}{\eta_n}\|T_\sigma x^n - x^n\|^2 - \frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n-u, w^n\rangle + \left(\frac{\lambda_n}{\eta_n}\|d^n\|\right)^2 \\
= &\ \|x^n-u\|^2 - \frac{2\lambda_n}{\eta_n}\sigma^2(x^n)\|Tx^n - x^n\|^2 - \frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n-u, w^n\rangle + \left(\frac{\lambda_n}{\eta_n}\|d^n\|\right)^2 \\
\leq &\ \|x^n-u\|^2 - \frac{\lambda_n}{2\eta_n}\frac{\left(\sum\limits_{i=1}^{m}\|S_ix^n - S_{i-1}x^n\|^2\right)^2}{\|Tx^n-x^n\|^2} - \frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n-u, w^n\rangle + \left(\frac{\lambda_n}{\eta_n}\|d^n\|\right)^2 \\
\leq &\ \|x^n-u\|^2 - \frac{\lambda_n}{4m\eta_n}\sum_{i=1}^{m}\|S_ix^n - S_{i-1}x^n\|^2 - \frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n-u, w^n\rangle + \left(\frac{\lambda_n}{\eta_n}\|d^n\|\right)^2.
\end{aligned}$$

Finally, by utilizing the fact that $\eta_n \geq \|d^n\|$, we obtain

$$\|x^{n+1}-u\|^2 \leq \ \|x^n-u\|^2 - \frac{\lambda_n}{4m\eta_n}\sum_{i=1}^{m}\|S_ix^n - S_{i-1}x^n\|^2 - \frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n-u, w^n\rangle + \lambda_n^2,$$

as desired. □

The following lemma guarantees the boundedness of the constructed sequence $\{x^n\}_{n=1}^\infty$.

Lemma 6. *The sequence $\{x^n\}_{n=1}^\infty$ generated by Algorithm 1 is bounded.*

Proof. Let $n \in \mathbb{N}$ and $u \in \bigcap\limits_{i=1}^{m} \operatorname{Fix} T_i$ be fixed. Let us notice that

$$\frac{\lambda_n}{4m\eta_n}\sum_{i=1}^{m}\|S_ix^n - S_{i-1}x^n\|^2 \geq 0,$$

which together with Lemma 5 yields

$$\|x^{n+1}-u\|^2 \leq \|x^n-u\|^2 - \frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n-u, w^n\rangle + \lambda_n^2. \tag{4}$$

Now, we set $A_n := \|x^n-u\|^2 - \sum\limits_{j=1}^{n-1}\lambda_j^2$ for all $n \in \mathbb{N}$. Thus, the relation (4) can be rewritten as

$$A_{n+1} - A_n + \frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n-u, w^n\rangle \leq 0. \tag{5}$$

To show that the sequence $\{x^n\}_{n=1}^\infty$ is bounded, we will consider the proof in two cases:

Case I: Suppose that there exists $n_0 \in \mathbb{N}$ such that the sequence $\{A_n\}_{n=1}^\infty$ is nonincreasing for all $n \geq n_0$. Then, we obtain that $\|x^n-u\|^2 - \sum\limits_{j=1}^{n-1}\lambda_j^2 \leq A_{n_0}$ for all $n \geq n_0$, which means that the sequence $\{\|x^n-u\|\}_{n=1}^\infty$ is bounded and, subsequently, $\{x^n\}_{n=1}^\infty$ is also a bounded sequence.

Case II: Suppose that there exists a subsequence $\{A_{n_k}\}_{k=1}^\infty$ of $\{A_n\}_{n=1}^\infty$ such that $A_{n_k} < A_{n_k+1}$ for all $k \in \mathbb{N}$, and let $\{\nu(n)\}_{n=1}^\infty$ be given in Lemma 4. This yields, for every $n \geq n_0$, that

$$A_{\nu(n)} \leq A_{\nu(n)+1} \tag{6}$$

and

$$A_n \leq A_{\nu(n)+1}. \tag{7}$$

Invoking the relation (6) in the inequaluty (5) and the positivity of the sequences $\{\alpha_n\}_{n=1}^{\infty}$, $\{\lambda_n\}_{n=1}^{\infty}$ and $\{\eta_n\}_{n=1}^{\infty}$, we obtain that

$$\langle x^{\nu(n)} - u, w^{\nu(n)} \rangle \leq 0. \tag{8}$$

On the other hand, by using the definition of $w^{\nu(n)} \in \partial_2 f(x^{\nu(n)}, x^{\nu(n)})$ and the fact that $f(x^{\nu(n)}, x^{\nu(n)}) = 0$, we get

$$\langle u - x^{\nu(n)}, w^{\nu(n)} \rangle \leq f(x^{\nu(n)}, u) - f(x^{\nu(n)}, x^{\nu(n)}) = f(x^{\nu(n)}, u).$$

This together with the inequality (8) yields that

$$f(x^{\nu(n)}, u) \geq 0.$$

Now, it follows from the ρ-strong monotonicity of f that

$$\rho\|x^{\nu(n)} - u\|^2 \leq -f(x^{\nu(n)}, u) - f(u, x^{\nu(n)}) \leq -f(u, x^{\nu(n)}). \tag{9}$$

On the other hand, for a fixed $\hat{u} \in \partial_2 f(u, u)$, we have

$$-f(u, x^{\nu(n)}) \leq \langle -\hat{u}, x^{\nu(n)} - u \rangle,$$

which together with the inequality (9) implies that

$$\rho\|x^{\nu(n)} - u\|^2 \leq \langle -\hat{u}, x^{\nu(n)} - u \rangle \leq \|\hat{u}\|\|x^{\nu(n)} - u\|,$$

and so

$$\|x^{\nu(n)} - u\| \leq \rho^{-1}\|\hat{u}\|.$$

This means that the sequence $\{\|x^{\nu(n)} - u\|\}_{n=1}^{\infty}$ is bounded. Now, since

$$A_{\nu(n)+1} = \|x^{\nu(n+1)} - u\|^2 - \sum_{j=1}^{\nu(n)} \lambda_j^2 \leq \|x^{\nu(n+1)} - u\|^2,$$

it follows that $\{A_{\nu(n)+1}\}_{n=1}^{\infty}$ is bounded above. Thus, by using (7), we get that $\{A_n\}_{n=1}^{\infty}$ is bounded and hence $\{x^n\}_{n=1}^{\infty}$ is also bounded. This completes the proof. □

The following lemma provides some important boundedness properties of the sequences $\{d^n\}_{n=1}^{\infty}$ and $\{\eta_n\}_{n=1}^{\infty}$.

Lemma 7. *The sequences $\{d^n\}_{n=1}^{\infty}$ and $\{\eta_n\}_{n=1}^{\infty}$ are bounded.*

Proof. Let $n \in \mathbb{N}$ and $u \in \bigcap_{i=1}^{m} \mathrm{Fix}\, T_i$ be fixed. Now, let note that

$$\begin{aligned} \|d^n\| &= \|\sigma(x^n)(x^n - Tx^n) + \alpha_n w^n\| \\ &\leq \|T_\sigma x^n - x^n\| + \alpha_n\|w^n\| \\ &\leq \|T_\sigma x^n - u\| + \|x^n - u\| + \alpha_n\|w^n\| \\ &\leq 2\|x^n - u\| + \alpha_n\|w^n\|, \end{aligned}$$

where the first inequality holds true by (3) and the last one holds true by the fact that T_σ is a cutter and consequently a quasi-nonexpansive operator.

As $w^n \in \partial_2 f(x^n, x^n)$, we have from Assumption (A3) and the boundedness of $\{x_n\}_{n=1}^\infty$ that the sequence $\{w_n\}_{n=1}^\infty$ is bounded which implies the sequence $\{d^n\}_{n=1}^\infty$ is also bounded. Consequently, from the definition of the sequence $\{\eta_n\}_{n=1}^\infty$, it can be seen that $\{\eta_n\}_{n=1}^\infty$ is also bounded. □

Now, we are in a position to present our main theorem.

Theorem 1. *Let $\{x^n\}_{n=1}^\infty$ be the sequence generated by Algorithm 1. Suppose that Assumption 1 is satisfied and the operators T_i, $i = 1, 2, \ldots, m$, satisfy the DC principle. Then the sequence $\{x^n\}_{n=1}^\infty$ converges strongly to the unique solution x^* of **BEP**.*

Proof. Let x^* be the unique solution of **BEP**. Firstly, we note from Lemma 5 with replacing $u = x^*$ that

$$\begin{aligned} \|x^{n+1} - x^*\|^2 &+ \sum_{j=1}^{n} \frac{\lambda_j}{8m\eta_j} \sum_{i=1}^{m} \|S_i x^j - S_{i-1} x^j\|^2 - \sum_{j=1}^{n} \lambda_j^2 \\ &\leq \|x^n - x^*\|^2 + \sum_{j=1}^{n-1} \frac{\lambda_j}{8m\eta_j} \sum_{i=1}^{m} \|S_i x^j - S_{i-1} x^j\|^2 - \sum_{j=1}^{n-1} \lambda_j^2 \\ &\quad - \frac{\lambda_n}{8m\eta_n} \sum_{i=1}^{m} \|S_i x^n - S_{i-1} x^n\|^2 - \frac{2\alpha_n \lambda_n}{\eta_n} \langle x^n - x^*, w^n \rangle. \end{aligned} \tag{10}$$

For simplicity, we denote $\Gamma_n := \|x^n - x^*\|^2 + \sum_{j=1}^{n-1} \frac{\lambda_j}{4m\eta_j} \sum_{i=1}^{m} \|S_i x^j - S_{i-1} x^j\|^2 - \sum_{j=1}^{n-1} \lambda_j^2$ for all $n \geq 2$. Then the inequality (10) is nothing else than

$$\Gamma_{n+1} \leq \Gamma_n - \frac{\lambda_n}{8m\eta_n} \sum_{i=1}^{m} \|S_i x^n - S_{i-1} x^n\|^2 - \frac{2\alpha_n \lambda_n}{\eta_n} \langle x^n - x^*, w^n \rangle. \tag{11}$$

To obtain the strong convergence of the generated sequence, we investigate the proof in two cases based on the behavior of the sequence $\{\Gamma_n\}_{n=1}^\infty$.

Case I: Suppose that there is $n_0 \in \mathbb{N}$ such that $\Gamma_{n+1} \leq \Gamma_n$ for every $n \geq n_0$. Thus, by using the definition of Γ_n, we note that

$$\|x^{n+1} - x^*\|^2 \leq \|x^n - x^*\|^2 - \frac{\lambda_n}{8m\eta_n} \sum_{i=1}^{m} \|S_i x^n - S_{i-1} x^n\|^2 + \lambda_n^2.$$

By utilizing Lemma 3 and the fact that $\sum_{n=1}^{\infty} \lambda_n^2 < \infty$, we obtain that the sequence $\{\|x^n - x^*\|\}_{n=1}^\infty$ is convergent and

$$\sum_{n=1}^{\infty} \frac{\lambda_n}{8m\eta_n} \sum_{i=1}^{m} \|S_i x^n - S_{i-1} x^n\|^2 < \infty.$$

Now, as $\sum_{n=1}^{\infty} \lambda_n = \infty$, we get that

$$\lim_{n\to\infty} \frac{1}{\eta_n} \sum_{i=1}^{m} \|S_i x^n - S_{i-1} x^n\|^2 = 0.$$

As the sequence $\{\eta_n\}_{n=1}^\infty$ is bounded, we have that, for all $i = 1, 2, \ldots, m$,

$$\lim_{n\to\infty} \|S_i x^n - S_{i-1} x^n\| = 0. \tag{12}$$

On the other hand, we note from Lemma 5 and the fact that $\frac{\lambda_n}{4m\eta_n}\sum_{i=1}^{m}\|S_i x^n - S_{i-1}x^n\|^2 \geq 0$, for all $n \in \mathbb{N}$, that

$$\frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n - x^*, w^n\rangle \leq \|x^n - x^*\|^2 - \|x^{n+1} - x^*\|^2 + \lambda_n^2.$$

By summing up this relation and the condition that $\sum_{n=1}^{\infty}\lambda_n^2 < \infty$, we obtain

$$\sum_{n=1}^{\infty}\frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n - x^*, w^n\rangle < \infty. \tag{13}$$

Now, since the sequence $\{\eta_n\}_{n=1}^{\infty}$ is bounded, there is a real number $M \geq 0$ such that $\eta_n \leq M$ for all $n \in \mathbb{N}$. This together with the assumption $\sum_{n=1}^{\infty}\alpha_n\lambda_n = \infty$ implies that

$$\sum_{n=1}^{\infty}\frac{\alpha_n\lambda_n}{\eta_n} \geq \sum_{n=1}^{\infty}\frac{\alpha_n\lambda_n}{M} = \infty. \tag{14}$$

Next, we show that $\liminf_{n\to\infty}\langle x^n - x^*, w^n\rangle \leq 0$. Suppose to the contrary that there exist $n_0 \in \mathbb{N}$ and $\delta > 0$ such that $\langle x^n - x^*, w^n\rangle \geq \delta$ for all $n \geq n_0$. Then,

$$\infty = \delta\sum_{n=n_0}^{\infty}\frac{2\alpha_n\lambda_n}{M} \leq \delta\sum_{n=n_0}^{\infty}\frac{2\alpha_n\lambda_n}{\eta_n} \leq \sum_{n=n_0}^{\infty}\frac{2\alpha_n\lambda_n}{\eta_n}\langle x^n - x^*, w^n\rangle < \infty,$$

which leads to a contradiction. Thus, we obtain

$$\liminf_{n\to\infty}\langle x^n - x^*, w^n\rangle \leq 0. \tag{15}$$

From the ρ-strongly monotonicity of f, it follows that

$$\rho\|x^n - x^*\|^2 \leq -f(x^n, x^*) - f(x^*, x^n) \leq \langle x^n - x^*, w^n\rangle - f(x^*, x^n).$$

Then,

$$\rho\|x^n - x^*\|^2 + f(x^*, x^n) \leq \langle x^n - x^*, w^n\rangle.$$

By taking the inferior limit, we have

$$\begin{aligned}\rho\liminf_{n\to\infty}\|x^n - x^*\|^2 + \liminf_{n\to\infty} f(x^*, x^n) &\leq \liminf_{n\to\infty}(\rho\|x^n - x^*\|^2 + f(x^*, x^n)) \\ &\leq \liminf_{n\to\infty}\langle x^n - x^*, w^n\rangle.\end{aligned}$$

Combining this and the inequality (15), we have

$$\liminf_{n\to\infty}\|x^n - x^*\|^2 \leq -\rho^{-1}\liminf_{n\to\infty} f(x^*, x^n). \tag{16}$$

Since the sequence $\{x^n\}_{n=1}^{\infty}$ is bounded, there exist a weakly cluster point $z \in \mathcal{H}$ and a subsequence $\{x^{n_k}\}_{k=1}^{\infty}$ of $\{x^n\}_{n=1}^{\infty}$ such that $x^{n_k} \rightharpoonup z \in \mathcal{H}$. We note from (12) that

$$\lim_{k\to\infty}\|(T_1 - Id)x^{n_k}\| = \lim_{k\to\infty}\|S_1 x^{n_k} - S_0 x^{n_k}\| = 0.$$

Thus, by using the DC principle of T_1, we have that $z \in \operatorname{Fix} T_1$. Furthermore, since $x^{n_k} \rightharpoonup z$ and it holds that

$$\lim_{k\to\infty}\|(T_1 x^{n_k} - T_1 z) - (x^{n_k} - z)\| = 0,$$

which imply that $T_1 x^{n_k} \rightharpoonup z$. Moreover, we note that

$$\lim_{k\to\infty} \|(T_2 - Id)T_1 x^{n_k}\| = \lim_{k\to\infty} \|S_2 x^{n_k} - S_1 x^{n_k}\| = 0.$$

By utilizing the DC principle of T_2, we have $z \in \mathrm{Fix}\, T_2$.

By processing the similar argument as above, we acquire that $z \in \mathrm{Fix}\, T_i$ for all $i = 1, 2, \ldots, m$, and hence $z \in \bigcap_{i=1}^{m} \mathrm{Fix}\, T_i$.

In virtue of the weak lower semicontinuity of $f(x^*, \cdot)$, we obtain

$$\liminf_{n\to\infty} f(x^*, x^n) = \lim_{k\to\infty} f(x^*, x^{n_k}) = \liminf_{k\to\infty} f(x^*, x^{n_k}) \geq f(x^*, z) \geq 0. \tag{17}$$

Combining the inequality (16) and (17), we have $\liminf_{n\to\infty} \|x^n - x^*\| = 0$. From the existence of $\lim_{n\to\infty} \|x^n - x^*\|$, we can conclude that

$$\lim_{n\to\infty} \|x^n - x^*\| = 0.$$

Case II: Suppose that there exists a subsequence $\{\Gamma_{n_k}\}_{k=1}^{\infty}$ of $\{\Gamma_n\}_{n=1}^{\infty}$ such that $\Gamma_{n_k} < \Gamma_{n_{k+1}}$ for all $k \in \mathbb{N}$. By Lemma 4, there exists a sequence of indices $\{\nu(n)\}_{n=1}^{\infty}$ such that, for all $n \geq n_0$,

$$\Gamma_{\nu(n)} \leq \Gamma_{\nu(n)+1}, \tag{18}$$

and

$$\Gamma_n \leq \Gamma_{\nu(n)+1}. \tag{19}$$

By using the inequalities (11) and (18), we have

$$0 \leq \Gamma_{\nu(n)+1} - \Gamma_{\nu(n)} \leq -\frac{\lambda_{\nu(n)}}{4m\eta_{\nu(n)}} \sum_{i=1}^{m} \|S_i x^{\nu(n)} - S_{i-1} x^{\nu(n)}\|^2 - \frac{2\alpha_{\nu(n)}\lambda_{\nu(n)}}{\eta_{\nu(n)}} \langle x^{\nu(n)} - x^*, w^{\nu(n)} \rangle.$$

Then,

$$\sum_{i=1}^{m} \|S_i x^{\nu(n)} - S_{i-1} x^{\nu(n)}\|^2 \leq -8m\alpha_{\nu(n)} \langle x^{\nu(n)} - x^*, w^{\nu(n)} \rangle. \tag{20}$$

By using the definition of $w^{\nu(n)} \in \partial_2 f(x^{\nu(n)}, x^{\nu(n)})$ and the fact that $f(x^{\nu(n)}, x^{\nu(n)}) = 0$, we get

$$\langle x^* - x^{\nu(n)}, w^{\nu(n)} \rangle \leq f(x^{\nu(n)}, x^*) - f(x^{\nu(n)}, x^{\nu(n)}) = f(x^{\nu(n)}, x^*), \tag{21}$$

which implies that

$$\sum_{i=1}^{m} \|S_i x^{\nu(n)} - S_{i-1} x^{\nu(n)}\|^2 \leq 8m\alpha_{\nu(n)} f(x^{\nu(n)}, x^*).$$

Now, using the ρ-strongly monotonicity of f that

$$f(x^{\nu(n)}, x^*) \leq -\rho \|x^{\nu(n)} - x^*\|^2 - f(x^*, x^{\nu(n)}),$$

and for a fixed $w^* \in \partial_2 f(x^*, x^*)$ such that

$$-f(x^*, x^{\nu(n)}) \leq \langle -w^*, x^{\nu(n)} - x^* \rangle,$$

we obtain

$$\sum_{i=1}^{m} \|S_i x^{\nu(n)} - S_{i-1} x^{\nu(n)}\|^2 \leq -8m\alpha_{\nu(n)}\rho \|x^{\nu(n)} - x^*\|^2 - 8m\alpha_{\nu(n)} \langle w^*, x^{\nu(n)} - x^* \rangle.$$

By using the boundedness of $\{x^n\}$ and $\lim_{n\to\infty} \alpha_n = 0$, we obtain

$$\lim_{n\to\infty} \sum_{i=1}^{m} \|S_i x^{\nu(n)} - S_{i-1} x^{\nu(n)}\|^2 = 0.$$

This implies that

$$\lim_{n\to\infty} \|S_i x^{\nu(n)} - S_{i-1} x^{\nu(n)}\|^2 = 0. \tag{22}$$

On the other hand, by using the ρ-strongly monotonicity of f and the inequality (21), we have

$$\rho\|x^{\nu(n)} - x^*\|^2 \leq -f(x^{\nu(n)}, x^*) - f(x^*, x^{\nu(n)}) \leq \langle x^{\nu(n)} - x^*, w^{\nu(n)}\rangle - f(x^*, x^{\nu(n)}).$$

By means of the fact that $\sum_{i=1}^{m} \|S_i x^{\nu(n)} - S_{i-1} x^{\nu(n)}\|^2 \geq 0$ in (20), it follows that

$$\langle x^{\nu(n)} - x^*, w^{\nu(n)}\rangle \leq 0.$$

Combining this and the above inequality, we obtain

$$\|x^{\nu(n)} - x^*\|^2 \leq -\rho^{-1} f(x^*, x^{\nu(n)}).$$

By taking the superior limit, we have

$$\limsup_{n\to\infty} \|x^{\nu(n)} - x^*\|^2 \leq -\rho^{-1} \limsup_{n\to\infty} f(x^*, x^{\nu(n)}). \tag{23}$$

As the sequence $\{x^{\nu(n)}\}_{n=1}^{\infty}$ is bounded, there exist a weakly cluster point $z \in \mathcal{H}$ and a subsequence $\{x^{\nu(n_k)}\}_{k=1}^{\infty}$ of $\{x^{\nu(n)}\}_{n=1}^{\infty}$ such that $x^{\nu(n_k)} \rightharpoonup z \in \mathcal{H}$. By following the argument as used in **Case I** together with the fact (22) and the DC principle of each T_i, we obtain that, for any subsequence $\{x^{\nu(n_k)}\}_{k=1}^{\infty}$ of $\{x^{\nu(n)}\}_{n=1}^{\infty}$, $x^{\nu(n_k)} \rightharpoonup z \in \bigcap_{i=1}^{m} \mathrm{Fix}\, T_i$.
By using the weak lower semicontinuity of $f(x^*, \cdot)$, we obtain

$$\lim_{k\to\infty} f(x^*, x^{\nu(n_k)}) = \liminf_{k\to\infty} f(x^*, x^{\nu(n_k)}) \geq f(x^*, z) \geq 0.$$

It follows from the inequality (23) that

$$\limsup_{n\to\infty} \|x^{\nu(n)} - x^*\|^2 \leq -\rho^{-1} \limsup_{n\to\infty} f(x^*, x^{\nu(n)}) \leq 0.$$

Then, we obtain

$$\lim_{n\to\infty} \|x^{\nu(n)} - x^*\|^2 = 0. \tag{24}$$

Note that from the definition of $x^{\nu(n)+1}$ and using the fact that $\eta_{\nu(n)} \geq \|d^{\nu(n)}\|$, we have

$$\|x^{\nu(n)+1} - x^{\nu(n)}\| = \|x^{\nu(n)} + \frac{\lambda_{\nu(n)}}{\eta_{\nu(n)}} d^{\nu(n)} - x^{\nu(n)}\| = \frac{\lambda_{\nu(n)}}{\eta_{\nu(n)}} \|d^{\nu(n)}\| \leq \lambda_{\nu(n)}.$$

Combining this and using the triangle inequality, we have

$$\begin{aligned} \|x^{\nu(n)+1} - x^*\| &\leq \|x^{\nu(n)+1} - x^{\nu(n)}\| + \|x^{\nu(n)} - x^*\| \\ &\leq \lambda_{\nu(n)} + \|x^{\nu(n)} - x^*\|. \end{aligned}$$

By using the inequality (24) and the fact that $\lim_{n\to\infty} \lambda_n = 0$, we obtain

$$\lim_{n\to\infty} \|x^{\nu(n)+1} - x^*\| = 0. \tag{25}$$

Next, using the inequality (19) and the fact that $\nu(n) \leq n$, we have

$$\begin{aligned}
\|x^n - x^*\|^2 \leq & \|x^{\nu(n)+1} - x^*\|^2 - \sum_{j=1}^{n-1} \frac{\lambda_j}{4m\eta_j} \sum_{i=1}^{m} \|S_i x^j - S_{i-1} x^j\|^2 \\
& + \sum_{j=1}^{\nu(n)} \frac{\lambda_j}{4m\eta_j} \sum_{i=1}^{m} \|S_i x^j - S_{i-1} x^j\|^2 + \sum_{j=1}^{n-1} \lambda_j^2 - \sum_{j=1}^{\nu(n)} \lambda_j^2 \\
\leq & \|x^{\nu(n)+1} - x^*\|^2 - \sum_{j=\nu(n)}^{n} \frac{\lambda_j}{4m\eta_j} \sum_{i=1}^{m} \|S_i x^j - S_{i-1} x^j\|^2 + \sum_{j=\nu(n)}^{n} \lambda_j^2 \\
\leq & \|x^{\nu(n)+1} - x^*\|^2 + \sum_{j=\nu(n)}^{n} \lambda_j^2.
\end{aligned} \tag{26}$$

Finally, by using the inequality (25), and the fact $\lim_{n\to\infty} \sum_{j=\nu(n)}^{n} \lambda_j^2 = 0$, we obtain

$$\lim_{n\to\infty} \|x^n - x^*\|^2 = 0.$$

This completes the proof. $\square$

Remark 3. *The DC principle assumption, which is assumed in the Theorem 1 holds true when the operators T_i, $i = 1, \ldots, m$, are nonexpansive. Actually, the metric projection onto closed convex sets and the subgradient projections of a continuous convex function, which is Lipschitz continuous on every bounded subset also satisfy the DC principle, see [16] further details.*

Remark 4. *It can be noted that the convergence result obtained in Theorem 1 holds true without any boundedness assumption of the generated sequence as in the previous works, for instance [20]. This underlines the convergence improvements accomplished in this work.*

4. A Numerical Example

In this section, we present a numerical example for solving the equilibrium problem over a finite number of half-space constraints. Let A, B be $n \times n$ matrices, $c_i \in \mathbb{R}^n$, and $d_i \geq 0$ be given for all $i = 1, 2, \ldots, m$, we consider the following equilibrium problem: find a point $\bar{u} \in \bigcap_{i=1}^{m} \mathrm{Fix}\, T_i$ such that

$$\langle A\bar{u} + By, y - \bar{u} \rangle \geq 0 \qquad \text{for all } y \in \bigcap_{i=1}^{m} \mathrm{Fix}\, T_i, \tag{27}$$

where the constrained set is

$$\mathrm{Fix}\, T_i = C_i := \{x \in \mathbb{R}^n : \langle c_i, x \rangle \leq d_i\}, i = 1, 2, \ldots, m.$$

We consider the operator T_i in two cases. In the first case, we put T_i to be the subgradient projection defined by

$$P_{g_i}(x) = \begin{cases} x - \frac{g_i(x)}{\|\nabla g_i(x)\|^2} \nabla g_i(x) & \text{if } g_i(x) \neq 0, \\ x & \text{otherwise,} \end{cases}$$

where $g_i(x) = \frac{1}{2}\text{dist}(x, C_i)^2$ with the distance function is given by $\text{dist}(x, C_i) := \inf_{z \in C_i} \|z - x\|$. In the second case, we put $T_i := P_{C_i}$, the metric projection onto C_i, for all $i = 1, 2, \ldots, m$. Note that it is known that the operators P_{g_i} and P_{C_i} are cutters and satisfy the DC principle with $\text{Fix}T_i = C_i$. We consider positive definite symmetric matrices A and B defined by $B := N^\top N + nI_n$, $A := B + M^\top M + nI_n$, where the $n \times n$ matrices N, M are randomly generated in $(0, 1)$, and I_n is the identity $n \times n$ matrix. Note that the bifunction $f(x, y) := \langle Ax + By, y - x \rangle$ is strongly monotone on $\mathcal{H}$, and for fixed $x \in \mathcal{H}$, we have $f(x, \cdot)$ is convex on $\mathcal{H}$. Moreover, we note that the diagonal subdifferential $\partial_2 f(x, x) = \{(A + B)x\}$, and we also know that the function $x \mapsto \partial_2 f(x, x)$ is bounded on a bounded subset of $\mathcal{H}$. These mean that the assumptions (A1)–(A3) are now satisfied. In this case, the problem (27) is the particular case of Problem 1 so that the sequence generated by Algorithm 1 can be applied to solve the problem.

We consider behavior of the sequence $\{x^n\}_{n=1}^\infty$ generated by Algorithm 1 for various positive real sequences $\{\alpha_n\}_{n=1}^\infty$ and $\{\lambda_n\}_{n=1}^\infty$ in the forms of Remark 3. We choose $\mu = 1$, and generate a vector c_i in $\mathbb{R}^n$ by uniformly distributed random generating between $(0, 1)$ and a scalar $d_i = 0$, for all $i = 1, 2, \ldots, m$. We choose the starting point of Algorithm 1 to be a vector whose coordinates are one. We terminate Algorithm 1 by the stopping criterions

$$\frac{\|x^{n+1} - x^n\|}{\|x^n\| + 1} \leq \varepsilon.$$

In the first experiment, we fix the parameters $a = 0.40, b = 0.60$ and $\varepsilon = 10^{-6}$. We perform 10 independent tests for any collections of parameters $\alpha = 0.10, 0.20, 0.30, 0.40$, and 0.50 and $\lambda = 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80$, and 0.90 when utilizing the operator $T_i := P_{g_i}$ and $T_i := P_{C_i}$ and the results are presented respectively in Tables 1 and 2, where the average number of iterations and the average computational runtime for any collection of α and λ are presented.

Table 1. Influence of parameters α and λ when using the subgradient projection operator $T_i = P_{g_i}$ where $a = 0.40$ and $b = 0.60$.

λ	$\alpha = 0.10$		$\alpha = 0.20$		$\alpha = 0.30$		$\alpha = 0.40$		$\alpha = 0.50$	
	Iter	**Time**	**Iter**	**Time**	**Iter**	**Time**	**Iter**	**Time**	**Iter**	**Time**
0.10	642	0.14	394	0.06	531	0.08	682	0.11	842	0.13
0.20	434	0.07	714	0.10	1013	0.16	1327	0.20	1644	0.25
0.30	594	0.09	1037	0.17	1492	0.23	1964	0.28	2439	0.37
0.40	754	0.11	1351	0.26	1970	0.29	2598	0.39	3223	0.48
0.50	914	0.14	1674	0.25	2443	0.36	3228	0.54	4013	0.60
0.60	1075	0.19	1987	0.33	2919	0.43	3861	0.65	4805	0.77
0.70	1232	0.18	2300	0.35	3396	0.51	4489	0.68	5571	0.84
0.80	1393	0.23	2618	0.40	3869	0.58	5119	0.77	6362	0.93
0.90	1550	0.23	2928	0.44	4342	0.64	5735	0.85	7157	1.05

Table 2. Influence of parameters α and λ when using the metric projection $T_i = P_{C_i}$ where $a = 0.40$ and $b = 0.60$.

λ	$\alpha = 0.10$		$\alpha = 0.20$		$\alpha = 0.30$		$\alpha = 0.40$		$\alpha = 0.50$	
	Iter	**Time**	**Iter**	**Time**	**Iter**	**Time**	**Iter**	**Time**	**Iter**	**Time**
0.10	640	0.16	412	0.06	553	0.07	698	0.09	846	0.13
0.20	458	0.07	735	0.09	1033	0.12	1335	0.17	1644	0.24
0.30	617	0.08	1056	0.13	1510	0.18	1967	0.23	2440	0.49
0.40	775	0.10	1371	0.17	1984	0.23	2599	0.30	3230	0.52
0.50	936	0.11	1686	0.21	2454	0.29	3225	0.39	4030	0.49
0.60	1091	0.16	2003	0.25	2923	0.35	3851	0.47	4802	0.59
0.70	1252	0.15	2315	0.30	3390	0.41	4493	0.56	5588	0.67
0.80	1407	0.17	2634	0.32	3868	0.48	5117	0.62	6364	0.79
0.90	1568	0.21	2951	0.37	4332	0.52	5751	0.70	7148	0.86

In Table 1, we presented the number of iterations (k) (Iter), the computational time (Time) in seconds when the stopping criteria of Algorithm 1 was met. Note that the larger $\lambda \in [0.20, 0.90]$ requires a larger number of iterations and computational runtime.

Furthermore, the best choice of the involved parameters for both cases is $\alpha = 0.20$ and $\lambda = 0.10$.

In a similar fashion with Table 1, we also presented in Table 2 the number of iterations (k) (Iter), the computational time (Time) in seconds, when the stopping criterions of Algorithm 1 when using the operator $T_i = P_{C_i}$ was met. The experimented results are in the same direction with Table 1 where the best choice of the involved parameters for both cases is $\alpha = 0.20$ and $\lambda = 0.10$.

In the next experiment, we consider the influence of parameters a and b by fixing the best parameters $\alpha = 0.20, \lambda = 0.10$ and $\varepsilon = 10^{-6}$. We performed 10 independent tests for any collections of parameters $a = 0.10, 0.15, 0.20, 0.25, 0.30, 0.35$ to 0.40 and $b = 0.55, 0.60$, $0.65, 0.70, 0.75, 0.80, 0.85$, and 0.90 when utilizing the operator $T_i := P_{g_i}$ and $T_i := P_{C_i}$ and the results of the average number of iterations and the average computational runtime for any collection of a and b are presented in Tables 3 and 4, respectively. We omit the combinations that do not satisfy the assumption in Theorem 1 and label it by -.

Table 3. Influence of parameters a and b when using the subgradient projection $T_i = P_{g_i}$ where $\alpha = 0.20$ and $\lambda = 0.10$.

b	$a = 0.10$		$a = 0.15$		$a = 0.20$		$a = 0.25$		$a = 0.30$		$a = 0.35$		$a = 0.40$	
	Iter	**Time**	**Iter**	**Time**	**Iter**	**Time**	**Iter**	**Time**	**Iter**	**Time**	Iter	**Time**	**Iter**	**Time**
0.55	6968	1.08	3759	0.59	2216	0.33	1396	0.22	932	0.14	663	0.10	505	0.09
0.60	3765	0.58	2209	0.33	1391	0.21	929	0.15	649	0.10	489	0.08	394	0.06
0.65	2200	0.35	1391	0.22	926	0.14	649	0.10	473	0.07	450	0.07	-	-
0.70	1387	0.22	924	0.15	648	0.10	534	0.09	619	0.10	-	-	-	-
0.75	923	0.16	726	0.14	844	0.15	1055	0.17	-	-	-	-	-	-
0.80	1405	0.25	1649	0.27	2101	0.34	-	-	-	-	-	-	-	-
0.85	4000	0.67	5199	0.85	-	-	-	-	-	-	-	-	-	-
0.90	18640	3.03	-	-	-	-	-	-	-	-	-	-	-	-

In Table 3, we see that the numbers of iterations as well as computational running time decrease when the value a increases. The the best result is obtained for the combination of $a = 0.40$ and $b = 0.60$.

Table 4. Influence of parameters a and b when using the metric projection $T_i = P_{C_i}$ where $\alpha = 0.20$ and $\lambda = 0.10$.

b	$a = 0.10$		$a = 0.15$		$a = 0.20$		$a = 0.25$		$a = 0.30$		$a = 0.35$		$a = 0.40$	
	Iter	**Time**	**Iter**	**Time**	**Iter**	**Time**	**Iter**	**Time**	**Iter**	**Time**	**Iter**	**Time**	**Iter**	**Time**
0.55	6962	0.85	3769	0.48	2212	0.27	1395	0.17	941	0.14	682	0.09	526	0.07
0.60	3770	0.46	2207	0.33	1386	0.17	928	0.11	667	0.09	509	0.07	412	0.06
0.65	2206	0.31	1390	0.17	929	0.12	649	0.09	491	0.06	450	0.06	-	-
0.70	1382	0.19	924	0.12	649	0.10	534	0.08	619	0.09	-	-	-	-
0.75	923	0.13	726	0.10	842	0.12	1060	0.15	-	-	-	-	-	-
0.80	1405	0.19	1649	0.23	2099	0.30	-	-	-	-	-	-	-	-
0.85	4014	0.55	5200	0.70	-	-	-	-	-	-	-	-	-	-
0.90	18636	2.55	-	-	-	-	-	-	-	-	-	-	-	-

In the same direction as the results in Table 3, it can be seen from Table 4 that the numbers of iterations and the computational running time is decreases when the values a grow up. The the best result is acquired for the combination of $a = 0.40$ and $b = 0.60$.

From these all above experiment, we observe that the choice of corresponding parameters $\alpha = 0.20$, $\lambda = 0.10$, $a = 0.40$ and $b = 0.60$ yields the best performance of both considered cases.

In the next experiment, we consider the behavior of Algorithm 1 for various values n and m by fixing the corresponding parameters as the above best choice. We also terminate Algorithm 1 when the error tolerance $\varepsilon = 10^{-6}$ was met, and the results are presented in Table 5.

Table 5. Comparisons between the using of the subgradient projection $T_i = P_{g_i}$ and the metric projection $T_i = P_{C_i}$ for different sizes of n and m.

n	m	$T_i := P_{g_i}$		$T_i := P_{C_i}$	
		Iter	**Time**	**Iter**	**Time**
200	100	393	0.15	412	0.21
	200	394	0.25	411	0.35
	300	392	0.36	413	0.50
	400	396	0.51	411	0.53
	500	393	0.56	413	0.62
	1000	392	1.03	411	1.16
300	100	780	0.37	810	0.34
	200	781	0.56	810	0.54
	300	780	0.78	810	0.79
	400	781	1.10	810	1.09
	500	781	1.20	809	1.22
	1000	781	2.21	809	2.14
400	100	1317	1.16	1353	0.97
	200	1319	1.72	1354	1.59
	300	1318	2.19	1354	2.11
	400	1320	2.65	1356	2.57
	500	1318	2.95	1354	2.96
	1000	1318	5.23	1354	5.34
500	100	2009	2.51	2046	2.25
	200	2008	3.63	2047	3.36
	300	2007	4.38	2048	4.34
	400	2008	5.24	2047	5.19
	500	2009	5.91	2046	5.95
	1000	2007	9.70	2046	10.30
1000	100	7755	33.31	7751	32.04
	200	7749	38.88	7749	37.39
	300	7750	43.94	7747	44.26
	400	7749	49.75	7748	48.42
	500	7750	55.72	7752	53.64
	1000	7747	81.83	7751	80.63

It is observed from Table 5 that for the values $n = 200, 300, 400$, and 500, the using of the subgradient projection is more efficient than using the metric projection in the sense that the first one requires less computation than the second one in the average number of iterations for all values m. In the case of $n = 1000$, we observe that there is no difference on these two cases. One notable behavior is that for each value n, we observe that even if the value m increases, the average numbers of iterations are almost the same, whereas the average computational runtime is increasing.

Finally, we present the comparison of the use of the subgradient projection and the metric projection for various optimality tolerances ε. We set $n = 500$ and $m = 50$ and choose the corresponding parameters in the same manner as above, the average numbers of iterations with respect to the optimality tolerances are presented in Figure 1.

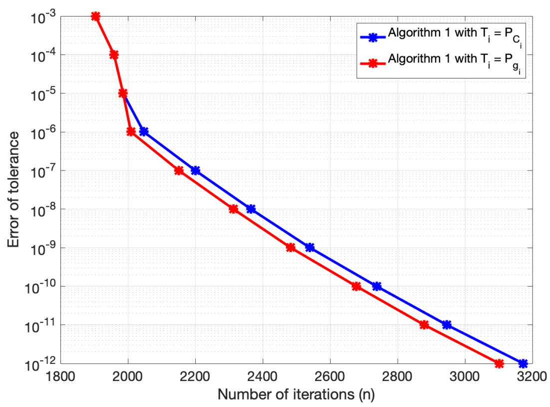

Figure 1. Comparison between the using of the subgradient projection $T_i = P_{g_i}$ and the metric projection $T_i = P_{C_i}$ with different errors of tolerance.

The plots in Figure 1 show that using the subgradient projection is more efficient than the metric projection for all the optimality tolerances. This emphasizes the superiority of using the subgradient projection when performing Algorithm 1.

5. Conclusions

In this work, we consider the solving of the bilevel equilibrium problem governed by a strongly monotone bifunction over the intersection of fixed-point sets of cutter operators. We associated with it the so-called subgradient-type extrapolation cyclic method. We present that the generated sequence generated by the proposed method converges to the unique solution to the problem. Our numerical experiment showed that using appropriate operators can yield a better convergence behavior to the proposed method.

It can be seen that the proposed subgradient-type extrapolation cyclic method (Algorithm 1) allows us to compute the operator $T_i, i = 1, \ldots, m$, sequentially. The main advantage of our method is that the computing machine is not necessary to store information while computing. Notwithstanding, the nature of the cyclic method, it is well-known that to compute S_i,, the cyclic method needs to have the estimate S_{i-1} in hand. This means that there has a waiting process while performing the method. In this case, one may consider the simultaneous extrapolation method [25] when dealing with the common fixed-point constrained of **BEP**.

Author Contributions: Conceptualization, P.P. and N.N.; methodology, P.P. and N.N.; software, N.N.; validation, P.P. and N.N.; convergence analysis, P.P. and N.N.; investigation, P.P. and N.N.; writing—original draft preparation, P.P. and N.N.; writing—review and editing, P.P. and N.N.; visualization, P.P. and N.N.; supervision, N.N.; project administration, N.N.; funding acquisition, N.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Fundamental Fund of Khon Kaen University. This research has received funding support from the National Science, Research and Innovation Fund or NSRF.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors are thankful to the Editor and three anonymous referees for comments and remarks which improved the quality and presentation of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fan, K. A minimax inequality and applications. In *Inequality III*; Shisha, O., Ed.; Academic Press: New York, NY, USA, 1972; pp. 103–113.
2. Blum, E.; Oettli, W. From optimization and variational inequalities to equilibrium problems. *Math. Student* **1994**, *63*, 123–145.
3. Muu, L.D.; Oettli, W. Convergence of an adaptive penalty scheme for finding constrained equilibria. *Nonlinear Anal.* **1992**, *18*, 1159–1166. [CrossRef]
4. Mastroeni, G. On auxiliary principle for equilibrium problems. In *Equilibrium Problems and Variational Models*; Daniele, P., Giannessi, F., Maugeri, A., Eds.; Kluwer Academic: Dordrecht, The Netherlands, 2003; pp. 289–298.
5. Nguyen, T.T.V.; Strodiot, J.J.; Nguyen, V.H. The interior proximal extragradient method for solving equilibrium problems. *J. Glob. Optim.* **2009**, *44*, 175–192. [CrossRef]
6. Iiduka, H.; Yamada, I. A subgradient-type method for the equilibrium problem over the fixed point set and its applications. *Optimization* **2009**, *58*, 251–261. [CrossRef]
7. Duc, P.M.; Muu, L.D. A splitting algorithm for a class of bilevel equilibrium problems involving nonexpansive mappings. *Optimization* **2016**, *65*, 1855–1866. [CrossRef]
8. Thuy, L.Q.; Hai, T.N. A Projected Subgradient Algorithm for Bilevel Equilibrium Problems and Applications. *J. Optim. Theory Appl.* **2017**, *175*, 411–431. [CrossRef]
9. Anh, P.N.; Van Hong, N. New projection methods for equilibrium problems over fixed point sets. *Optim. Lett.* **2021**, *15*, 627–648. [CrossRef]

10. Kim, J.K.; Anh, P.N.; Hai, T.N. The Bruck's ergodic iteration method for the Ky Fan inequality over the fixed point set. *Int. J. Comput. Math.* **2017**, *94*, 2466–2480.
11. Kaczmarz, S. Angenäherte Auflösung von Systemen linearer Gleichungen. *Bull. Int.* **1937**, *A35*, 355–357; reprinted in *Int. J. Control* **1993**, *57*, 1269–1271.
12. Opial, Z. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Am. Math. Soc.* **1967**, *73*, 591–597. [CrossRef]
13. Cegielski, A.; Censor, Y. Opial-Type Theorems and the Common Fixed Point Problem. In *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*; Bauschke, H., Burachik, R., Combettes, P., Elser, V., Luke, D., Wolkowicz, H., Eds.; Springer Optimization and Its Applications; Springer: New York, NY, USA, 2011; Volume 49.
14. Liu, C. An acceleration scheme for row projection methods. *J. Comput. Appl.* **1995**, *57*, 363–391. [CrossRef]
15. Cegielski, A. Generalized relaxations of nonexpansive operators and convex feasibility problems. *Contemp. Math.* **2010**, *513*, 111–123.
16. Cegielski, A. *Iterative Methods for Fixed Point Problems in Hilbert Spaces*; Lecture Notes in Mathematics, 2057; Springer: Berlin/Heidelberg, Germany, 2012.
17. Bauschke, H.H.; Combettes, P.L. A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces. *Math. Oper. Res.* **2001**, *26*, 248–264. [CrossRef]
18. Cegielski, A.; Censor, Y. Extrapolation and local acceleration of an iterative process for common fixed point problems. *J. Math. Anal. Appl.* **2012**, *394*, 809–818. [CrossRef]
19. Cegielski, A.; Nimana, N. Extrapolated cyclic subgradient projection methods for the convex feasibility problems and their numerical behaviour. *Optimization* **2019**, *68*, 145–161. [CrossRef]
20. Prangprakhon, M.; Nimana, N. xtrapolated sequential constraint method for variational inequality over the intersection of fixed-point sets. *Numer. Algorithms* **2021**, *88*, 1051–1075. [CrossRef]
21. Bauschke, H.H.; Combettes, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed.; CMS Books in Mathematics; Springer: New York, NY, USA, 2017.
22. Iusem, A.N. On the maximal monotonicity of diagonal subdifferential operators. *J. Convex Anal.* **2011**, *18*, 489–503.
23. Combettes, P.L. Quasi-Fejérian analysis of some optimization algorithms. In *Inherently Parallel Algorithms for Feasibility and Optimization*; Butnariu, D., Censor, Y., Reich, S., Eds.; Elsevier: New York, NY, USA, 2001; pp. 115–152.
24. Maingé, P.E. Strong convergence of projected subgradient methods for nonsmooth and nonstrictly convex minimization. *Set-Valued Anal.* **2008**, *16*, 899–912. [CrossRef]
25. Butnariu, D.; Censor, Y.; Gurfil, P.; and Hadar, E. On the behavior of subgradient projections methods for convex feasibility problems in Euclidean spaces *SIAM J. Optim.* **2008**, *19*, 786–807. [CrossRef]