

https://dspace.org/

Relaxed Inertial Algorithm for Solving Equilibrium Problems

Elijah, Nwakpa Chidi

2024
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/42246
Elijah, Nwakpa Chidi. (2024). Relaxed Inertial Algorithm for Solving Equilibrium Problems [Master's dissertation, University of the Witwatersrand, Johannesburg]. WireDSpace.

Downloaded from a DSpace repository. https://dspace.org/

Relaxed Inertial Algorithm for Solving Equilibrium Problems

by

Nwakpa Chidi Elijah

DISSERTATION
Submitted in Fulfilment of the
Requirements for the Degree
of
Master of Science

University of The Witwatersrand



UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

School of Mathematical Science
University of The Witwatersrand, Johannesburg, South Africa.

January, 2024.

Relaxed Inertial Algorithm for Solving Equilibrium Problems

by

Nwakpa Chidi Elijah

As the candidate's supervisors, we have approved this dissertation for submission.

Dr. C. C. Okeke

.....

Dr. C. Izuchukwu

.....

Abstract

In this dissertation, we propose and study two relaxed inertial methods for solving equilibrium problems. In our first proposed method, we establish that the generated sequence of our proposed method weakly converges to a solution of the equilibrium problems. We apply this proposed method to variational inequality and fixed point problems. Furthermore, a modification of the first method leads us to our second iterative method. Again, we established that the sequence generated by this method converges strongly to a solution of the equilibrium problems. Our proposed methods involve self-adaptive stepsizes and hence, do not require the fore knowledge of the Lipschitz constants for implementation. In each of our proposed methods, the convergence is established when the associated cost bifunction is pseudomonotone and satisfies the Lipschitz-type condition.

Acknowledgments

I am mostly grateful to the Supreme God for the goodness and favor he has been providing me with before and during this program.

My sincere gratitude goes to my Dad, and my wonderful and lovely siblings—Celestine, Blessing, Chioma and Judith, and every other member of my family for the unending encouragement and support. In a most special way, I am grateful to my beloved late mum who longed to witness my academic progress, but could not as she took the final break from this word.

My heartfelt gratitude goes to Mr. Moses Iloke and family, Chief Odumba and family for the financial support throughout my academic pursuit.

In an immeasurable special way, I am extending my sincere gratitude to my best supervisors Dr. C. C. Okeke and Dr. C. Izuchuckwu who have gone beyond being academic supervisors to taking me as part of their families. I am verily grateful for all they have done for me (encouragement, financial, material and moral support), and as well seeing that this program came into fruition by putting their time in reviewing and constructively criticizing some concepts in this work.

I am heartfully thankful to Mrs. Obioma, Dr. (Mrs.) Lizzy Eweputanna, Mrs. Confidence Martins, Dr. (Mrs.) Onwuezobe and Prof. (Mrs.) Rose for their financial support and pieces of advice that have been my propelling force.

I sincerely acknowledge the University of the Witwatersrand (Wits), for her financial support and warm reception throughout this study program. My special thanks to Mr. Singh, the administrator, School of Mathematics, Wits, and all staff members of the school for their neighbourliness, too.

I also wish to thank Dr. C. N. Okoro, Dr. (Mrs.) E. T. Efor and every staff member of the Department of Industrial Mathematics and Applied Statistics, Ebonyi State University, Nigeria.

My earnest thanks to my best friend, Kalu for always being a kind friend. To all my colleagues and friends, Thembinkosi, Miriam, Jabulile, Shallwell, Godfrey, Mpoh, Kubiat, Blessed Chinedu, Oluchukwu and others, it has been a great time knowing you all. Thank you all.

Dedication

To my *beloved mum* ♡ who took her final break from this world during the time of this study program.

Contents

Title page ii

Abstract iii

Acknowledgements iv

Dedication v

Declaration viii

Submitted Papers From This Dissertation ix

1 Introduction **1**

1.1 Background of the Study 1

1.2 Research Motivation 2

1.3 Statement of Problem 4

1.4 Objectives 6

2 Literature Review **7**

2.1 Preliminaries and Definitions 7

2.1.1 Space of Interest 7

2.1.2 Set Convexity 8

2.1.3 Bifunctions 9

2.1.4 Differentiability of Convex Bifunctions 13

2.1.5 Metric Projection on Hilbert spaces 17

2.1.6 Basic Identities and Inequalities in Real Hilbert Spaces. 19

2.2 Equilibrium Problems 20

2.2.1 Examples of Equilibrium Problems 20

2.2.2 Previous Works on Equilibrium Problems in Real Hilbert Spaces . . 21

2.2.3 Equilibrium Problem as a Generalization of Other Convex Problems 25

2.2.4 Existence of Solutions for Equilibrium Problems 29

3 Relaxed Inertial Algorithm for Solving Equilibrium Problems **31**

3.1 Introduction 31

3.2 Preliminaries 31

3.3 Main Results 34

3.4 Applications 46

3.4.1 Variational Inequality Problem 46

3.4.2 Fixed Points Problems 47

4 Relaxed Inertial Sub-gradient Extra-gradient Method for Solving Equilibrium Problems. **49**

4.1 Introduction 49

4.2 Preliminaries 49

4.3 Main Results 50

5 Conclusion, Contribution to Knowledge and Future Research **59**

5.1 Conclusion 59

5.2 Contribution to knowledge 59

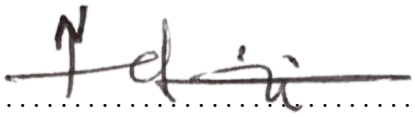
5.3 Future Research 60

Bibliography 61

Declaration

The work embodied in this dissertation is original and has not been submitted in part or full for any other diploma or degree of this or any other university, and proper referencing is provided for the usage of others' works.

Nwakpa Chidi Elijah



19th March, 2024.

Submitted Papers From This Dissertation

1. **Chidi Elijah Nwakpa**, Chinedu Izuchukwu, Chibueze Christian Okeke, Relaxed Inertial Algorithm for Solving Equilibrium Problems, *Computational and Applied Mathematics.*

2. **Chidi Elijah Nwakpa**, Chinedu Izuchukwu, Chibueze Christian Okeke, Relaxed Inertial Sub-gradient Extra-gradient Method for Solving Equilibrium Problems, *Networks and Spatial Economics.*

Chapter 1

Introduction

1.1 Background of the Study

Throughout this dissertation, we will let $\mathcal{H}$ denote a real Hilbert space furnished with the inner product function, $\langle \cdot, \cdot \rangle$, and induced norm $\| \cdot \|$. We will also let $\mathcal{C}$ denote a closed convex nonempty subset of $\mathcal{H}$. Suppose $f(\cdot, \cdot) : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{R}$ is a bifunction in such a way $f(s, s) = 0$, $\forall s \in \mathcal{C}$.
The equilibrium problem (briefly, EP) for the bifunction f is defined as:

$$\text{find } \bar{r} \in \mathcal{C} \quad \text{such} \quad \text{that } f(\bar{r}, s) \not< 0 \quad \forall s \in \mathcal{C}. \tag{1.1.1}$$

Mathematical programming problems can be viewed as variational inequality problems and optimization problems, but interestingly, both models are special realizations of an abstract equilibrium problems. In the algorithms that solve the mathematical programming, one thing stands out—the convexity of the objective function. This is in no any way disregarded for bifunctions of equilibrium problems as monotonicity correspondingly replaces the classic convexity assumption, though this classical assumptions are sufficient but not necessary [8]. This shows, as we shall see subsequently, that equilibrium problems have earlier been in existence.

In the early study, in 1969 and 1972, Zukhovitskii *et al.* [95] and Brezis *et al.* [15] respectively studied the general model, that is, inequality (1.1.1). Also, in 1972, Ky Fan [28] in his work made some great remarkable contributions to inequality (1.1.1), hence, this inequality is some times referred to as Ky Fan inequality. In 1991, at the 57th Annual Conference of the Indian Mathematical Society held at Aligarh, Blum and Oettli [14] presented their work, "From Optimization and Variational Inequalities to Equilibrium Problems," and the term "Equilibrium Problems" was mentioned. They extensively explained that equilibrium problem is a unification of variational inequality, fixed points, Nash equilibrium, complementarity, saddle points, optimizations, convex differentiable optimization problems. It is, of course, important to also note that EP(1.1.1) generalizes problems like image restoration, financial analysis in Economics and other problems in Physics, thus, making equilibrium problem an interesting aspect of Mathematics.

Earlier before the publication of the work of Blum and Oettli, equilibrium problem was not a well known path in mathematical programming and optimization, but after the publication, it began gaining the real interest of many researchers. Thus, many authors have devoted efforts in studying EP(1.1.1), thereby developing and analyzing several effective iterative methods of approximating its equilibrium point. In fact, nowadays, there is a good theory for equilibria and a rapidly increasing number of methods for finding them. These methods, not in any order, include: the projection method [61, 66, 85]; The proximal point methods (PPM) [43, 62, 63, 64]; The extra-gradient methods (EGM) with or without line search (such as: The hybrid projection-type extra-gradient method (HPEGM), The shrinking projection-type extra-gradient method (SPEGM), The extra-gradient-viscosity method(EGVM) [54], The Halpern-type exra-gradient method (HEGM)) [41, 45, 70, 77, 91, 93]; The methods using the Bregman distance [48, 80]; The gap function methods [13, 60, 76]; The sub-gradient extra-gradient methods [18, 35, 37, 86, 88]; The descent methods [10, 12]; and The inertial type methods [22, 36, 44, 47, 69, 78, 86]. Each method is formulated for a class of equilibrium problems. In all these methods, the methods of relaxed inertial type sub-gradient extra-gradient for approximating the solutions of EP(1.1.1) are our total focus in this dissertation and will be handled in details in chapters 3 and 4.

1.2 Research Motivation

In the early years, methods like conjugate gradients, Abramov's, Lysternik's, Faddeyev's general three-term iteration process, etc. were widely used in Linear Algebra for speeding up the convergence of iterative methods for obtaining the solutions to the Linear equations, but an intriguing thing is that these methods involved the extrapolation term. There was a strong desire to improve this extrapolation term which led Frankel [30] to observe that there would be an improvement in convergence rate of iterative process of solving the second order differential equations using the second-order Richardson method over the Liebmann extrapolation method. This improvement was as a result of one not only using $\phi^n_{j',k'}$ (j', k' are points all through the neighbourhood of j, k), but as well use $\phi^{n-1}_{j,k}$ in the calculation of $\phi^{n+1}_{j,k}$. This is a modification brought by the formal equivalence between the Richardson method and the solution of the time-dependent equation, $\dfrac{\partial \phi}{\partial t} = L\phi$, where the modified iteration method is similar to the solution of the second order differential equation in time t,

$$\frac{\partial^2 \phi}{\partial^2 t} + A\frac{\partial \phi}{\partial t} - L\phi = 0,$$

which can be written as

$$\begin{array}{ll} \phi^{n+1}_{j,k} = \phi^n_{j,k} + \alpha(\phi^n_{j,k} - \phi^{n-1}_{j,k}) + \beta L\phi^n_{j,k} & (\text{ interior points}), \\ b_{j,k} & (\text{boundary points}). \end{array} \tag{1.2.1}$$

And the solution of the equation $L\phi_{j,k} = 0$ satisfying the included boundary condition is denoted by $\phi_{j,k}$. Polyak [74] replicated (1.2.1) by studying the iterative method

$$x_{k+1} = x_k + \alpha(x_k - x_{k-1}) - \beta p(x_k), \tag{1.2.2}$$

whose continuous analogue form is

$$\ddot{x}(t) = c_1\dot{x}(t) + c_2 p(x(t)) \tag{1.2.3}$$

and observed that the two-step methods in relation to the corresponding one-step methods indeed speed up the rate of convergence. He was the first to call (1.2.2) the inertial type extrapolation algorithm.

Attouch and Cabot [4] assumed $c_1 = \vartheta(t)$ as a function of t and used an explicit finite difference discretization of the dynamical system (1.2.3) in time t with a time step size $h_k > 0$ and by setting $t_k := \sum_j^k h_j$, $x_k := x(t_k)$, $\vartheta_k := \vartheta(t_k)$ obtained

$$\frac{1}{h_k^2}(x_{k+1} - 2x_k + x_{k-1}) + \frac{\vartheta_k}{h_k}(x_k - x_{k-1}) + \gamma p(w_k) = 0, \tag{1.2.4}$$

where, for each $k \geq 1$, w_k is a point belonging to the line passing through the points x_{k-1} and x_k.
Expanding (1.2.4) and setting $\theta_k = 1 - \vartheta_k h_k \geq 0$, gives

$$x_{k+1} = x_k + \theta_k(x_k - x_{k-1}) - \gamma h_k^2 p(w_k), \tag{1.2.5}$$

where θ_k is called the inertial coefficient.

(1.2.5) which is a modification of (1.2.1) that uses the constants α and β is more practicable since it involves a variable inertial coefficient, θ_k. Nesterov [68] obtained faster convergence when he introduced the first formal inertial algorithm in which his extrapolation choice is $w_k = x_k + \theta_k(x_k - x_{k-1})$. This is a part of what our algorithms seek to incorporate. Moudafi [64] proposed a method that combines the relaxed proximal point method and the extrapolation term.

Attouch and Cabot in their work [3], obtained weak convergence for the proximal algorithm for maximally monotone operators through the relaxed inertial technique.

Shehu *et al.* [86] studied the sub-gradient extra-gradient type of EPs and achieved strong convergence through the introduction of inertial and relaxation parameters.

Thong *et al.* in [89] introduced a relaxed inertial Tseng's method for finding the solutions set of a variational inequality and fixed point problem, and used the on-line rule and the assumptions that f satisfies the Lipschitzness and pseudomonotonicity conditions to obtain the strong convergence of their scheme.

Izuchuckwu *et al.* [44] examined the split monotone variational inclusion problems in two real Hilbert Spaces and combined the inertial and relaxation techniques with the proximal contraction algorithm to obtain the strong convergence of their proposed algorithms.

Going in the same direction, this work is motivated by the works of Attouch and Cabot [3], Izuchuckwu *et al.* [44], Shehu *et al.* [86] and Thong *et al.* [89]. We study the sub-gradient extra-gradient equilibrium problems by introducing the extrapolation effect and the relaxation techniques.

1.3 Statement of Problem

Let $\mathcal{H}$ be a real Hilbert space, and $\mathcal{C} \subset \mathcal{H}$ a closed convex nonempty subset of $\mathcal{H}$. Let $f(\cdot,\cdot) : \mathcal{C}\times\mathcal{C} \longrightarrow \mathbb{R}$ denote a pseudomonotone bifunction in such a way $f(s,s) = 0, \forall s \in \mathcal{C}$. Then, the general model of our problems is

$$\text{find } \bar{r} \in \mathcal{C} \ \text{ such that } \ f(\bar{r}, s) \not\geq 0 \ \ \forall s \in \mathcal{C}. \tag{1.3.1}$$

Where $EP(f,\mathcal{C}) := \{\bar{r} \in \mathcal{C} \mid f(\bar{r}, s) \geq 0 \ \ \forall s \in \mathcal{H}\}$ is the solutions set.
Solving for a solution point ($\bar{r} \in EP(f,\mathcal{C})$) of (1.3.1), we use the following methods for evaluating x_{n+1} :

Algorithm 1.3.1.
Condition on the inertial and relaxation factors.
Suppose that $\theta_n \in [0,1)$, $\alpha_n \in (0,1)$ *and* $\liminf_{n\to\infty} \alpha_n > 0$. *Assume for* n *sufficiently large enough, there is* $\epsilon \in (0,1)$ *for which*

$$\begin{aligned}&(1-\epsilon)\frac{1-\alpha_{n-1}}{2\alpha_{n-1}}(1-\theta_{n-1})\\ &\geq \theta_n t_{n+1}\left(1+\theta_n+\left[\frac{1-\alpha_n}{2\alpha_n}(1-\theta_n)+\frac{1-\alpha_{n-1}}{2\alpha_{n-1}}(1-\theta_{n-1})\right]_+\right).\end{aligned}$$

Let $x_0,\ x_1 \in \mathcal{H},\ \lambda_1 > 0,\ 0 < \mu < 1$.
STEP 1*: Start with the present iterates* x_{n-1}, x_n *and evaluate*

$$\begin{cases} w_n = x_n + \theta_n(x_n - x_{n-1}) \\ \\ y_n = arg\min\limits_{y\in\mathcal{C}} \left\{\lambda_n f(w_n, y) + \dfrac{1}{2}\|w_n - y\|^2\right\}. \end{cases}$$

Stop the computation if $y_n = w_n$, *and* y_n *becomes a solution, otherwise, move to step 2.*
STEP 2*:*
Choose $v_n \in \partial f(w_n, .)(y_n)$ *in such a way there is* $q_n \in N_{\mathcal{C}}(y_n)$ *satisfying* $q_n = w_n - \lambda_n v_n - y_n$, *and construct a half space* $T_n = \{z \in \mathcal{H} : \langle w_n - \lambda_n v_n - y_n, z - y_n\rangle \leq 0\}$.

Evaluate
$z_n = arg\min\limits_{y\in T_n} \left\{\lambda_n f(y_n, y) + \frac{1}{2}\|w_n - y\|^2\right\}$.

STEP 3*: Compute*

$x_{n+1} = (1-\alpha_n)w_n + \alpha_n z_n$, *and*

$$\lambda_{n+1} = \begin{cases} \min\left\{\dfrac{\mu}{2}\dfrac{\|w_n - y_n\|^2 + \|z_n - y_n\|^2}{f(w_n,z_n) - f(w_n,y_n) - f(y_n,z_n)},\ \ \lambda_n\right\}, & \text{if } f(w_n,z_n) - f(w_n,y_n) \\ & \quad -f(y_n,z_n) > 0, \\ \lambda_n, \quad \text{otherwise.} \end{cases}$$

Set $n+1 \leftarrow n$ *and continue again from* ***step 1***.

Algorithm 1.3.2.
Let $\alpha_n \in (0,1)$ *and* $\liminf_{n\to\infty} \alpha_n > 0$. *Assume that* $\{\xi_n\}$ *is any positive sequence such that* $\xi_n = o(\sigma_n)$, *that is,* $\lim_{n\to\infty} \dfrac{\xi_n}{\sigma_n} = 0$, *where* $\{\sigma_n\} \subset (0,1)$ *in a way*

$$\lim_{n\to\infty} \sigma_n = 0, \quad \sum_{n=1}^{\infty} \sigma_n = \infty.$$

Let $x_0,\ x_1 \in \mathcal{H},\ \lambda_1 > 0,\ \mu \in]0,1[$.
STEP 1*:*
Given the current iterates x_{n-1} *and* x_n, *choose* θ_n *such that* $0 \le \theta_n \le \bar{\theta}_n$, *where*

$$\bar{\theta}_n = \begin{cases} \min\left\{\theta, \dfrac{\xi_n}{\|x_n - x_{n-1}\|}\right\} & \text{if } x_n \neq x_{n-1}, \\ \theta \qquad\qquad \text{otherwise.} \end{cases}$$

STEP 2*:*
Set $w_n = (1-\sigma_n)(x_n + \theta_n(x_n - x_{n-1}))$ *and calculate*
$y_n = \underset{y\in\mathcal{C}}{arg\min}\{\lambda_n f(w_n, y) + \frac{1}{2}\|w_n - y\|^2\}$.
Stop the computation if $y_n = w_n$, *and* y_n *becomes a solution, otherwise, move to step 3.*
STEP 3*:*
Choose $v_n \in \partial f(w_n, .)(y_n)$ *in a way there is*
$q_n \in N_{\mathcal{C}}(y_n)$ *satisfying*
$q_n = w_n - \lambda_n v_n - y_n$, *and construct a half space*
$T_n = \{z \in \mathcal{H} : \langle w_n - \lambda_n v_n - y_n, y - y_n\rangle \le 0\}$.
Evaluate
$z_n = \underset{y\in T_n}{arg\min}\{\lambda_n f(y_n, y) + \frac{1}{2}\|w_n - y\|^2\}$.

STEP 4*: Compute*

$x_{n+1} = (1-\alpha_n)w_n + \alpha_n z_n$, *and*

$$\lambda_{n+1} = \begin{cases} \min\left\{\dfrac{\mu}{2}\dfrac{\|w_n - y_n\|^2 + \|z_n - y_n\|^2}{f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n)}, \quad \lambda_n\right\}, & \text{if } f(w_n, z_n) - f(w_n, y_n) \\ & -f(y_n, z_n) > 0, \\ \lambda_n, \quad \text{otherwise.} \end{cases}$$

Set $n+1 \leftarrow n$ *and continue again from* ***step 1***.

In this dissertation, these are our proposed methods for solving problem (1.3.1).

1.4 Objectives

The major aims of this dissertation are to:

(i) review some already existing and important outcomes on equilibrium problems,

(ii) introduce and study iterative methods for finding the approximate solutions of the EPs,

(iii) establish weak and strong convergent results for EP using our suggested schemes,

(iv) use our achieved results in some mathematical programming problems.

The remaining of this dissertation is presented thus:

- Chapter 2: We revisit some important fundamental concepts, definitions, propositions and theorems that played vital roles in the formulation and the analysis of our main results. This chapter also captures some good examples to aid in better understanding of some concepts. Extensive literature review on equilibrium problems is as well presented. More so, comprehensive relationships between EP and other convex problems are introduced and finally, the chapter ended with existence of solution theorem for EPs in $\mathcal{H}$.
- Chapter 3: We presented and analyzed our first proposed method for finding a real approximate solution of an EP with applications to variational inequality and fixed points problems.
- Chapter 4: Our second iterative method for solving Ep was analyzed in this chapter.
- Chapter 5: In this chapter, we present our conclusion, contribution to knowledge and possible time ahead research.

Chapter 2

Literature Review

We aim to highlight some important fundamental concepts, definitions, propositions and theorems that play vital roles in the formulation and the analysis of our main results. This chapter also captures some good examples to aid in better understanding of some concepts. Extensive literature review on equilibrium problems is as well presented. More so, comprehensive relationships between EP and other convex problems are introduced and finally, the chapter ended with existence of solution theorem for EPs in $\mathcal{H}$.

2.1 Preliminaries and Definitions

2.1.1 Space of Interest

Our space of interest is a special type of Banach spaces called the real Hilbert space, and this is because of its simple geometric structure when compared to other Banach spaces. Some unique features make working in Hilbert spaces relatively easy which we are going to explore in this chapter and thereafter use them in the subsequent chapters.

Now, we give definition of Hilbert spaces and our adapted notations.

Definition 2.1.1. *Suppose $\mathcal{V}$ is a linear space, the single-valued function $\langle\cdot,\cdot\rangle : \mathcal{V}\times\mathcal{V} \longrightarrow \mathcal{K}$ defined from $\mathcal{V}\times\mathcal{V}$ to $\mathcal{K} = \mathbb{R}$ or $\mathbb{C}$ is called an inner product on $\mathcal{V}$ if for each $s, w, z \in \mathcal{V}$ and $\gamma, \delta \in \mathcal{K}$ the following are true:*

$C_1 :$ $\langle s, s\rangle \geq 0$ *and* $\langle s, s\rangle = 0$ *if and only if* $s = 0$;

$C_2 :$ $\langle s, w\rangle = \overline{\langle w, s\rangle}$, *where* $\overline{z}$ *indicates complex conjugation of* z;

$C_3 :$ $\langle \gamma s + \delta w, z\rangle = \gamma\langle s, z\rangle + \delta\langle w, z\rangle$.

If the pair $(\mathcal{V}, \langle\cdot,\cdot\rangle)$ called an inner product space is complete, then it becomes a Hilbert space. Keeping in mind that $(\mathcal{V}, \langle\cdot,\cdot\rangle)$ is complete, we therefore define $\mathcal{H} = (\mathcal{V}, \langle\cdot,\cdot\rangle)$. Hence, as earlier mentioned in chapter 1 except otherwise stated, $\mathcal{H}$ denotes a real Hilbert space endowed with the inner product $\langle\cdot\ ,\ \cdot\rangle$ and induced norm $\|\cdot\|$. We will let $\{x_n\} \to \bar{r}$

and $\{x_n\} \rightharpoonup \bar{r}$ to respectively mean strong and weak convergence of $x_n \in \mathcal{H}$. While $\overline{\mathbb{R}}$ represents the extension of real numbers, $\mathbb{R}$ to $+\infty$.

2.1.2 Set Convexity

Definition 2.1.2. *Let $s, z \in \mathcal{H}$, then, the set of points of the form $[s, z] = (1-\lambda)s+\lambda z = s+\lambda(z-s), \ \ \forall \lambda \in \mathbb{R}$ is called the line segment, while a set $\mathcal{E}$ is called a flat (or an affine set) if $(1-\lambda)s + \lambda z \in \mathcal{E} \ \ \forall s, z \in \mathcal{E}$ and $\lambda \in \mathbb{R}$.*

Definition 2.1.3. *Let $s_1, ..., s_n \in \mathcal{E}$; a vector $s = \lambda_1 s_1 + ... + \lambda_n s_n$ with $\sum_{i=1}^{n} \lambda_i = 1$ is an affine combination of $s_i, \ \ i = 1, 2, ..., n$. Thus an affine set $\mathcal{E}$ contains affine combination of any set of points from $\mathcal{E}$.*
Geometrically, affine sets which contain the origin are subspaces of $\mathcal{H}$.

Definition 2.1.4. *A subset $\mathcal{C}$ of $\mathcal{H}$ is called a convex set if the line segment $[s, z] = (1-\lambda)s + \lambda z$ is in $\mathcal{C}$ whenever s and z are in $\mathcal{C}$ and $0 < \lambda < 1$.*
This shows that all affine sets (including $\emptyset$ and $\mathcal{H}$ itself) are convex but conversely this is not generally true.

Example 2.1.5. *Solid ellipsoid and cubes in $\mathcal{H} = \mathbb{R}^3$ are convex but not affine.*

Definition 2.1.6. *A nonempty convex set $\mathcal{T}$ is said to be a halfspace if for any non-zero $s \in \mathcal{H}$ and any $\delta \in \mathbb{R}$*

$$\begin{aligned} \mathcal{T} = \{s : \langle s, k\rangle \leq \delta\}, \qquad & \{s : \langle s, k\rangle \geq \delta\}, \\ \{s : \langle s, k\rangle < \delta\} \ \ \textit{or} \ \ & \{s : \langle s, k\rangle > \delta\}. \end{aligned}$$

Definition 2.1.7. *A point $z_0 \in \mathcal{C}$ is known as an algebraic interior point of $\mathcal{C}$ if every straight line through z_0 has an open segment centered at z_0, while the collections of algebraic interior points of $\mathcal{C}$ is known as the algebraic kernel $Ai(\mathcal{C})$ of $\mathcal{C}$.*

It is important we note that whenever $\mathcal{H}$ is finite dimensional, then every convex subset of $\mathcal{H}$ contains algebraic interior points. It is not always true if $\mathcal{H}$ is infinite dimensional, however, a particularly important class of convex bodies called the convex $\alpha-$bodies or convex algebraic bodies are formed by those infinite dimensional spaces which have at least one algebraic interior.

Definition 2.1.8. *A set $\mathcal{C}$ is a cone whenever we have $\lambda s \in \mathcal{C}$ for each $s \in \mathcal{C}$ and $\lambda \geq 0$, and it is a convex cone only when it is convex and a cone. A point $s = \sum_{i=1}^{n} \lambda_i s_i$ is called a conic combination or weighted sum of $s_1, ..., s_n$ if $\lambda_i \geq 0$ for each $i \geq 1$. Hence, a set becomes a convex cone if and only if it contains all conic combinations of its elements.*

Definition 2.1.9. *Let $\mathcal{C}$ denote a convex cone, then the conjugate cone is given by*

$$\mathcal{C}^* = \{v \in \mathcal{H} : \langle v, s\rangle \not> 0 \ \ \textit{for each} \ \ s \in \mathcal{C}\}.$$

Example 2.1.10. *Given a halfspace $\mathcal{C} = \{s \in \mathbb{R}^n : \langle a, s\rangle \not> 0\}, \ \ a \neq 0$; obviously $\mathcal{C}$ is a convex cone , and its conjugate cone is given by*

$$\mathcal{C}^* = \{v \in \mathbb{R}^n : v = \lambda a, \ \ \lambda \geq 0\}.$$

Definition 2.1.11. *Suppose $\mathcal{C} \subset \mathcal{H}$ is a convex set, with $z \in \mathcal{C}$, then*

$$N_{\mathcal{C}}(z) := \{u \in \mathcal{H} : \langle u, s - z\rangle \leq 0 \ \ \textit{for any} \ \ s \in \mathcal{C}\}$$

is called the normal cone to $\mathcal{C}$ at z.

Remark 2.1.12. *It is easily verified that the normal cone $N_{\mathcal{C}}(z)$ is a closed convex cone, and also, given $w \in int(\mathcal{C})$, then $N_{\mathcal{C}}(w) = 0$.*

2.1.3 Bifunctions

Definition 2.1.13. *Let $f : \mathcal{C} \longrightarrow \overline{\mathbb{R}}$ be a mapping. Then*

a. f is called a convex function if:

i. $\mathcal{C}$ is a convex set, and

ii. $\forall \lambda \in \mathbb{R}$ for which $0 \leq \lambda \leq 1$ and for each $s, w \in \mathcal{C}$, then

$$\begin{aligned} f(w_\lambda) = f((1-\lambda)w + \lambda s) &\leq (1-\lambda)f(w) + \lambda f(s) \\ &= \lambda(f(s) - f(w)) + f(w); \end{aligned}$$

b. f is concave if:

i. $\mathcal{C}$ is a convex set, and

ii. $\forall \lambda \in \mathbb{R}$ for which $0 \leq \lambda \leq 1$ and for each $s, w \in \mathcal{C}$, then

$$\begin{aligned} f(w_\lambda) = f((1-\lambda)w + \lambda s) &\geq (1-\lambda)f(w) + \lambda f(s) \\ &= \lambda(f(s) - f(w)) + f(w); \end{aligned}$$

c. f is said to be proper if there is at least one $w_0 \in \mathcal{C}$ for which $f(w_0) \neq +\infty$;

d f is a strictly convex function if $\leq$ is replaced with $<$.

Remark 2.1.14.

i. The intuitive geometric interpretation of definition (a.) is that the graph of f lies below the chord that joins the points $(w, f(w))$ and $(s, f(s))$.

ii. The function $f : \mathcal{H} \to \mathbb{R}$ is convex provided the function $\varphi(v) = f(w + vd)$ is convex $\forall w, d \in \mathcal{H}$.

iii. A function f is concave whenever $-f$ is convex.

iv. Convex functions are continuous at every interior point, satisfy Lipschitz conditions and are differentiable at any given direction.

v. A strictly convex function has a unique minimum. In fact, a strictly convex function defined on a convex set has at most one global minimizer, and the only local minimizer or stationary point (if the function is differentiable) is, in fact, the global minimizer if it exists.

Example 2.1.15. *The quadratic function* $f(w) = \frac{1}{2}\langle w, Tw\rangle + \langle w, b\rangle + \alpha$, *where* T *is an* $n-$*square symmetric matrix, is convex on* $\mathcal{H}$ *if and only if* T *is positive semi-definite, i.e.,* $\langle z, Tz\rangle \geq 0$ *for each* $z \in \mathcal{H}$.

Definition 2.1.16. *Suppose* $\mathcal{C} \subset \mathcal{H}$ *is a closed convex nonempty subset of* $\mathcal{H}$. *For any given function* $f : \mathcal{C} \subset \mathcal{H} \longrightarrow \overline{\mathbb{R}}$, *its effective domain is described by*

$$dom\ f := \{s \in \mathcal{C} : f(s) < +\infty\}$$

and function f *is proper if dom* $f \neq \emptyset$.

Definition 2.1.17. *The epigraph of the function* f *is the set denoted by*

$$epi(f) := \{(s, \beta) \in \mathcal{H} \times \mathbb{R} \mid s \in dom\ f \text{ and } f(s) \leq \beta\}.$$

Definition 2.1.18. *If* $\mathcal{H}$ *denotes a real normed space and* $N(s_0)$ *the set of all neighbourhoods of* s_0, *then a function* $f : \mathcal{H} \longrightarrow \overline{\mathbb{R}}$ *is said to be lower semi-continuous function at* s_0 *if for any* $z < f(s_0)$, *there exists* $N(s_0)$ *such that* $f(s) > z$ *for each* $s \in N(s_0)$.

Equivalently, f *is lower semi-continuous at* s_0 *if and only if*

$$f(s_0) \leq \liminf_{s \to s_0} f(s).$$

It follows that if f is lower semi-continuous at $s_0 \in \mathcal{H}$ and given a sequence $\{s_n\} \subset \mathcal{H}$ for which $s_n \to s_0$, then

$$\liminf_{n \to \infty} f(s_n) \geq f(s_0).$$

Definition 2.1.19. *Let* $f : \mathcal{H} \longrightarrow \overline{\mathbb{R}}$, *then* f *is said to be upper semi-continuous at a point* $s_0 \in \mathcal{H}$ *if for all* $z > f(s_0)$, *there is a neighbourhood* N *of* s_0 *for which* $f(s) < z$ *for every* $s \in N(s_0)$. *Or* f *is upper semi-continuous at* s_0 *if and only if*

$$\limsup_{s \to s_0} \leq f(s_0).$$

We quickly point out that the function f is upper semi-continuous if and only if $-f$ is lower semi-continuous. Also, a function is continuous if and only if it is both upper and lower semi-continuous.

Definition 2.1.20. *According to Blum and Oettli [14], a function of a pair of values,* $f(\cdot, \cdot) : \mathcal{C} \times \mathcal{C} \subset \mathcal{H} \times \mathcal{H} \longrightarrow \mathbb{R}$, *is called a bifunction if* $f(z, z) = 0$ *for each* $z \in \mathcal{C}$.

Definition 2.1.21. *As a set valued function, Rockafellar [81] defined a bifunction from* $\mathbb{R}^m$ *to* $\mathbb{R}^n$ *to be a mapping* G *that assigns to each* $s \in \mathbb{R}^m$ *a function* $G(s)$ *on* $\mathbb{R}^n$ *with values in* $\overline{\mathbb{R}} \cup \{-\infty\}$, *where the value of* $G(s)$ *at a point* $z \in \mathbb{R}^n$ *is described by* $(Gs)(z) = f(s, z)$.

However, a bifunction can therefore be simply regarded as the first stage of a function broken down into two stages:

$$G : s \longrightarrow G(s) : z \longrightarrow (Gs)(z) = f(s, z).$$

Definition 2.1.22. *A convex bifunction is said to be closed or proper whenever its graph function, $(s,z) \longrightarrow (Gs)(z)$ for each $(s,z) \in \mathbb{R}^{m+n}$ is either closed or proper, respectively.*

Example 2.1.23. *[81] Let $G : \mathbb{R}^2 \longrightarrow \mathbb{R}^n$ be given by*

$$(Gz)(w) = \begin{cases} \left(\dfrac{1}{1+z_1}\right)\langle w, Tw\rangle + \langle a, w\rangle \quad \text{if} \quad z_1 > -1 \quad \text{and} \quad w \in B + z_2 e, \\ \\ 0 \quad \text{if} \quad z_1 = -1 \quad \text{and} \quad w \in B + z_2 e, \\ \\ +\infty \quad \text{otherwise,} \end{cases}$$

where $z = (z_1, z_2)$, T is an $n-$positive semi-definite symmetric matrix, B is the Euclidean unit ball of $\mathbb{R}^n$ and $a, e \in \mathbb{R}^n$ with $\|e\| = 1$. Therefore, G is a closed proper convex bifunction.

Definition 2.1.24. *For any bifunction G from $\mathbb{R}^m \longrightarrow \mathbb{R}^n$, $G : z \longrightarrow Gz : w \longrightarrow (Gz)(w)$, the adjoint of G is defined as the bifunction $G^* : w^* \longrightarrow G^*w^* : z^* \longrightarrow (G^*w^*)(z^*)$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ given by:*

$$(G^*w^*)(z^*) = \begin{cases} \inf\limits_{z,w}\{(Gz)(w) - \langle w, w^*\rangle + \langle z, z^*\rangle\} \quad \text{if } G \text{ is convex and,} \\ \\ \sup\limits_{z,w}\{(Gz)(w) - \langle w, w^*\rangle + \langle z, z^*\rangle\} \quad \text{if } G \text{ is concave.} \end{cases}$$

Remark 2.1.25. *In order to express the dependence of the objective function in a minimization problem on a vector z corresponding to a perturbation, the concept "bifunction" is conveniently introduced as a generalization of a multivalued mapping.*

Definition 2.1.26. *Let $\mathcal{C} \subset \mathcal{H}$ be a closed convex nonempty subset of $\mathcal{H}$ and f be a bifunction defined from $\mathcal{C}$ to $\mathbb{R}$, $f : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{R}$. Then, for each $s, w, z \in \mathcal{C}$, f is said to be:*

i. monotone on $\mathcal{C}$ whenever

$$0 \geq f(s,z) + f(z,s);$$

ii. strongly pseudomonotone on $\mathcal{C}$ whenever there is a positive constant μ for which

$$f(s,z) \geq 0 \quad \text{implies} \quad f(z,s) + \mu\|s - z\|^2 \leq 0;$$

iii. pseudomonotone on $\mathcal{C}$ whenever

$$f(s,z) \not< 0 \quad \text{implies} \quad f(z,s) \not> 0;$$

iv. satisfying Lipschitz-type condition on $\mathcal{C}$ whenever there are two constants $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$f(s,z) - f(s,w) - f(w,z) \leq \delta_1\|s - w\|^2 + \delta_2\|w - z\|^2;$$

v. strictly pseudomonotone on $\mathcal{C}$ whenever for each $s, z \in \mathcal{C}, s \neq z$,

$$f(s,z) > 0 \implies f(z,s) < 0;$$

vi. quasimonotone on $\mathcal{C}$ whenever

$$f(s,z) > 0 \implies f(z,s) \leq 0;$$

vii. strictly quasimonotone on $\mathcal{C}$ whenever f is quasimonotone and, for each $s, z \in \mathcal{C}$, $s \neq z$, there is w in the open line segment between s and z such that either

$$f(w,z) \neq 0 \quad or \quad f(w,s) \neq 0.$$

Remark 2.1.27.

i. The convexity of a function is analogous to its monotonicity.

ii. The pseudomonotonicity of f automatically implies that $f(z,z) = 0 \ \forall z \in \mathcal{C}$.

iii. Clearly, one can easily obtain the following relationships: monotonicity implies pseudomonotonicity; strong pseudomonotonicity implies pseudomonotonicity; strict pseudomonotonicity implies pseudomonotonicity; strict quasimonotonicity implies quasimonotonicity; pseudomonotonicity implies quasimonotonicity; monotonicity implies quasimonotonicity; strong and strict pseudomonotonicities imply quasimonotonicity. But in general, the converse statements are not always true, as can be seen as follows:

Example 2.1.28. *Let $f : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ be the bifunction given by*

$$f(v,z) = \begin{cases} v^2(z-v)^3, & v \neq 0 \\ \\ 0, & v = 0. \end{cases}$$

Clearly, one can see that $f(v,z)$ is quasimonotone in $\mathbb{R} \times \mathbb{R}$ but not pseudomonotone. In fact, $f(v,z) > 0$ implies $v \neq 0$, $z > v$, and then

$$f(z,v) = z^2(v-z)^3 \leq 0 \quad \forall z \in \mathbb{R},$$

and thus, f is quasimonotone. However, f is not pseudomonotone, since

$$f(0,-1) = 0 \quad and \quad f(-1,0) = 1 > 0.$$

Example 2.1.29. *Suppose $\mathcal{H} = \mathbb{R}$, $\mathcal{C} = [\frac{1}{2}, 1]$ and $f : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{R}$ is given by*

$$f(s,z) = s(s-z) \quad for \ each \quad s, z \in \mathcal{C}.$$

Then f is pseudomonotone but not monotone.
To see this, let $f(s,z) = s(s-z) \geq 0$.
But $s \geq \frac{1}{2}$, this implies that $(s-z) \geq 0$.
Similarly, $z \in \mathcal{C} \implies z \geq \frac{1}{2}$, thus, $f(z,s) = z(z-s) \leq 0$, showing that f is therefore pseudomonotone.
However, f is not monotone since for $s \neq z$, $f(s,z) + f(z,s) = (s-z)^2 > 0$.

2.1.4 Differentiability of Convex Bifunctions

Non-smooth convex functions are very crucial in optimization, and it is important we note that every convex function has a sub-gradient at each $z \in domf$. The sub-gradient of the non-smooth (non-differentiable) functions is not unique while the sub-gradient of a smooth (differentiable) convex function at each point $z \in domf$ is then the gradient of f, ∇f, which is unique. We shall see more of sub-gradient in the subsequent chapters as it forms a basis of our study.
In this section, we present a few basic definitions of differentiability of functions which in specific terms are applicable to convex bifunctions.

Definition 2.1.30. *A scalar function $f : \mathcal{H} = \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable at $z \in domf$ if there is a vector $\nabla f(z) \in \mathcal{H}$, known as the gradient of the function f at z, such that for each $w \in \mathcal{H}$, then*

$$f(w) = f(z) + \langle \nabla f(z), w - z \rangle + o(\|w - z\|), \tag{2.1.1}$$

where $\nabla f(z) = \left(\dfrac{\partial f}{z_1}, ..., \dfrac{\partial f}{z_n} \right)$.

In other words, f is differentiable at z whenever we can find a linear approximation

$$\begin{aligned} \tilde{f}(z) &= f(z) + \langle \nabla f(z), w - z \rangle \text{ for which} \\ |f(w) - \tilde{f}(w)| &= o(\|w - z)\|. \end{aligned} \tag{2.1.2}$$

Remark 2.1.31. *A function is only differentiable at z_0 if it can be approximated at the neighbourhood of z_0 by an affine function.*

Definition 2.1.32. *A vector function $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is said to be differentiable at $z \in \mathbb{R}^n$ whenever there is a matrix $\nabla g(z) \in \mathbb{R}^{m \times n}$, known as the Jacobian of the vector g at $z \in \mathbb{R}^n$, such that for each $w \in \mathbb{R}^n$, we have*

$$g(w) = g(z) + \nabla g(z)(w - z) + o(\|w - z\|), \tag{2.1.3}$$

and $\nabla g(z)_{i,j} = \dfrac{\partial g_i(z)}{\partial z_j}$, $g(z) = \big(g_1(z), ..., g_m(z)\big)^T$.

In other words, a vector function g is differentiable at z whenever we can find the following linear approximation at z :

$$\|g(w) - g(z) - \nabla g(z)(w - z)\| \leq o(\|w - z\|). \tag{2.1.4}$$

Definition 2.1.33. *A scalar function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is said to be twice differentiable at $z \in \mathbb{R}^n$ whenever there is a symmetric matrix $\nabla^2 f(z) \in \mathbb{R}^{n \times n}$, known as the Hessian of f at $z \in \mathbb{R}^n$, such that for each $w \in \mathbb{R}^n$, we have*

$$f(w) = f(z) + \langle \nabla f(z), w - z \rangle + \frac{1}{2} \langle \nabla^2 f(z)(w - z), w - z \rangle + o(\|w - z\|^2), \tag{2.1.5}$$

or f is twice differentiable at $z \in \mathbb{R}^n$ whenever it admits the following quadratic approximation:

$$\tilde{f}(w) = f(z) + \langle \nabla f(z), w - z \rangle + \frac{1}{2}\langle \nabla^2 f(z)(w-z), w - z \rangle,$$
$$\text{so that } \; |f(w) - \tilde{f}(w)| = o(\|w - z\|^2). \tag{2.1.6}$$

Definition 2.1.34. *Suppose $s \in int\ domf$, then the bifunction $f : \mathcal{H} \times \mathcal{H} \longrightarrow \mathcal{H}$ is differentiable at $s = (z, y)$ in the direction of $d = (d_1, d_2)$ if the following limit*

$$\nabla f(s; d) = \lim_{t \to 0} \frac{1}{t}[f(z + d_1 t, y + d_2 t) - f(z, y)] = \lim_{t \to 0} \frac{1}{t}[f(s + td) - f(s) \tag{2.1.7}$$

exists, where the value $\nabla f(s; d)$ is known as the directional derivative of f at s in the direction d and it always exists for convex function, though it may be $+\infty$ or $-\infty$,
$domf = \{u \in \mathcal{H} : f(u) < +\infty\}$.

Remark 2.1.35. *We note here that the limit in the definition above always exists, to see this, we make use of the fact that the ratio $\frac{1}{t}[f(s+td) - f(s)]$ is non-decreasing in t. Thus, for $0 < t_1 \leq t_2$, we have $0 \leq \lambda \leq 1$, where $\lambda = \frac{t_1}{t_2}$, and*

$$\begin{aligned} \frac{f(s + t_1 d) - f(s)}{t_1} &= \frac{f[\lambda(s + t_2 d) + (1 - \lambda)s] - f(s)}{t_1} \\ &\leq \frac{\lambda f(s + t_2 d)}{t_1} + \frac{(1 - \lambda)f(s) - f(s)}{t_1} \\ &= \frac{f(s + t_2 d) - f(s)}{t_2}, \end{aligned} \tag{2.1.8}$$

hence, the limit exists.

Definition 2.1.36. *A vector g is called a sub-gradient of a convex function f at $v_0 \in int\ domf$ whenever*

$$f(v) \geq f(v_0) + \langle g, v - v_0 \rangle \;\; \textit{for each} \;\; v \in domf. \tag{2.1.9}$$

The set of sub-gradients g of f at v_0 is known as the sub-differential $\partial f(v_0)$ of f described by:

$$\partial f(v_0) = \{g : f(v) \geq f(v_0) + \langle g, v - v_0 \rangle \;\; \forall v \in domf\}.$$

Thus, the sub-differential of a function f at v_0 is the set valued mapping $\partial f : \mathcal{H} \longrightarrow 2^{\mathcal{H}}$ described by

$$\partial f(v_0) = \{g \in \mathcal{H} : f(v) - f(v_0) \geq \langle v - v_0, g \rangle \;\; \forall v \in domf\}. \tag{2.1.10}$$

A function is said to be sub-differentiable at v whenever there is at least one sub-gradient of f at v.

Remark 2.1.37. *The sub-differential of f at any point $v_0 \in int\ dom f$ is a convex nonempty bounded set.*
This is because sub-differential is the intersection of an infinite set of halfspaces

$$\partial f(v_0) = \bigcap_{v \in dom f} \{g : f(v) \geq f(v_0) + \langle g, v - v_0 \rangle\}, \tag{2.1.11}$$

and since halfspaces are closed convex set, hence, $\partial f(v_0)$ is also a closed convex set even when f is not convex. More so, suppose f is continuous at v_0, we have that the sub-differential $\partial f(v_0)$ is bounded.

Example 2.1.38. *Suppose $f : \mathbb{R} \longrightarrow \mathbb{R}$ is described by*

$$f(v) = \begin{cases} v^2, & v \leq 0, \\ \\ v, & v \geq 0. \end{cases} \tag{2.1.12}$$

Then, at $v = 0$, the right derivative is 1 while the left derivative is 0. So, $\partial f(0) = [0, 1]$.

Example 2.1.39. *Define $f : \mathbb{R} \longrightarrow \mathbb{R}$ by*

$$f(x) = \max(x, -x) = |x|, \tag{2.1.13}$$

then, at $x = 0$, the right and left derivatives are 1 and -1 respectively. Therefore, at $x = 0$, the sub-differential is defined by the inequality $|u| \geq gu$ for all u which is satisfied if and only if $g \in [-1, 1]$, and so, $\partial f(0) = [-1, 1]$.

Example 2.1.40. *Suppose $f : \mathcal{H} \longrightarrow \mathcal{H}$ is a convex function and $h(s) = f(As + b)$, then*

$$\partial h(s) = A^T \partial f(As + b). \tag{2.1.14}$$

Example 2.1.41. *[81] Consider the function f defined by*

$$f(w) = \begin{cases} |w| - 2(1 - w)^{\frac{1}{2}} & if \quad -3 \leq w \leq 1, \\ \\ +\infty & otherwise, \end{cases} \tag{2.1.15}$$

then, f is a closed proper convex function on $\mathbb{R}$ while

$$\partial f(w) = \begin{cases} \emptyset \quad if \quad w \geq 1, \\ \\ \{1 + (1 - w)^{-\frac{1}{2}}\} \quad if \quad 0 < w < 1, \\ \\ [0, 2] \quad if \quad w = 0, \\ \\ \{-1 + (1 - w)^{-\frac{1}{2}}\} \quad if \quad -3 < w < 0, \\ \\ (-\infty, -\frac{1}{2}] \quad if \quad w = -3, \\ \\ \emptyset \quad if \quad w < -3. \end{cases} \tag{2.1.16}$$

Definition 2.1.42. *Suppose $f : \mathcal{E} \longrightarrow \mathcal{U}$ is a function. Let $h \neq 0$ and v be vectors in $\mathcal{E}$. The Gâteaux differential $df(v;h)$ of f at $v \in \mathcal{E}$ in the direction $h \in \mathcal{E}$ is given by*

$$d(v;h) = \lim_{\varrho \to 0} \frac{f(v+\varrho h) - f(v)}{\varrho} = \frac{d}{d\varrho} f(v+\varrho h)|_{\varrho=0}, \tag{2.1.17}$$

or equivalently

$$f(v+\varrho h) = f(v) + \varrho df(v;h) + o(\varrho). \tag{2.1.18}$$

Remark 2.1.43.

i. The Gâteaux differential is the generalization of the classical directional derivative.

ii. A discontinuous linear functional in an infinite dimensional space $\mathcal{E}$ is Gâteaux differentiable, but at the origin O, its Gâteaux differential is linear and discontinuous.

iii. The theorem on differentiation of a composite function is not valid for the Gâteaux derivative if the derivative is allowed to be discontinuous.

iv. A function is Gâteaux differentiable provided the sub-differential $\partial f(v)$ contains only one element $f'(v) = \nabla f(v)$, [21].

Definition 2.1.44. *Suppose $\mathcal{A}$ and $\mathcal{B}$ are normed vector spaces. $f : \mathcal{A} \longrightarrow \mathcal{B}$ is called a Fréchet differentiable function at $w \in \mathcal{B}$ whenever it is Gâteaux differentiable and there is a bounded linear mapping $Df : \mathcal{A} \longrightarrow \mathcal{B}$ for which*

$$\lim_{\|k\| \to 0} \frac{\|f(w+k) - f(w) - (Df)k\|_{\mathcal{B}}}{\|k\|_{\mathcal{A}}} = 0, \tag{2.1.19}$$

or equivalently,

$$f(w+k) = f(w) + (Df)k + o(\|k\|). \tag{2.1.20}$$

Remark 2.1.45.

i. If $k = \varrho h$, then we have $f(w+\varrho h) = f(w)+\varrho(Df)h+o(\varrho)h$, and with $\varrho \to 0$, we have $(Df)h = df(w;h)$ which is the Gâteaux differential, and $Df = T$ if $df(w;h)$ has the form Th. Therefore, any Fréchet differentiable function is also Gâteaux differentiable, but the converse is not generally true, since the Gâteaux derivative may fail to be linear or continuous which is analogous to the fact that the existence of all directional derivatives at a point does not guarantee total continuity or differentiability at that point.

ii If a function f is continuously Fréchet differentiable in a neighbourhood of a point v_0, and if the Fréchet derivative $Df(v_0)$ at v_0 is a homeomorphism of the Banach spaces $\mathcal{A}$ and $\mathcal{B}$, then the inverse mapping theorem holds.

iii. The Fréchet derivative defines a tangent hyperplane.

2.1.5 Metric Projection on Hilbert spaces

Definition 2.1.46. *Suppose $s \in \mathcal{H}$, and $\mathcal{C} \subset \mathcal{H}$ is a closed convex nonempty set. The metric projection or the proximal mapping of s onto $\mathcal{C}$ is the closest point in $\mathcal{C}$ to s defined by*

$$P_{\mathcal{C}}(s) = \underset{z \in \mathcal{C}}{arg\min} \left\{ \frac{1}{2}\|s - z\|^2 \right\} \tag{2.1.21}$$

and satisfies

$$\|s - P_{\mathcal{C}}(s)\| = \min_{z \in \mathcal{C}} \|s - z\|. \tag{2.1.22}$$

$P_{\mathcal{C}}(s)$ always exists and is unique.
Notice that the function $f(s) = \frac{1}{2}\|s - z\|^2$ is continuous in z and is strongly convex, hence the minimizer $P_{\mathcal{C}}(s)$ is unique. Again, whenever $s \in \mathcal{C}$, we have $P_{\mathcal{C}}(s) = s$.

We give some examples on metric projection as follows:

Example 2.1.47.

(a) Suppose that $z \in \mathcal{H}$ such that $z \neq 0$ and $\alpha \in \mathbb{R}$. Define $\mathcal{C}$ by $\mathcal{C} := \{s \in \mathcal{H} : \langle s, z\rangle = \alpha\}$. Then for each $s \in \mathcal{H}$, the projection of s onto $\mathcal{C}$ is

$$P_{\mathcal{C}}(s) = s + \frac{\alpha - \langle s, z\rangle}{\|z\|^2} z. \tag{2.1.23}$$

Observe that if we define $\mathcal{C} = \{s \in \mathcal{H} : \langle s, z\rangle \leq \alpha\}$ and $z \in \mathcal{H}$ has no restriction, thus, exactly one of these statements is true:

i $z = 0$ and $\alpha \geq 0$, then $\mathcal{C} = \mathcal{H}$ and $P_{\mathcal{C}} = Id$.

ii. $z = 0$ and $\alpha < 0$, then $\mathcal{C} = \emptyset$.

iii. $z \neq 0$, then $\mathcal{C} \neq \emptyset$ and for each $s \in \mathcal{H}$,

$$P_{\mathcal{C}}(s) = \begin{cases} s, & \text{if } \langle s, z\rangle \leq \alpha \\ \\ s + \dfrac{\alpha - \langle s, z\rangle}{\|z\|^2} z, & \text{if } \langle s, z\rangle > \alpha. \end{cases} \tag{2.1.24}$$

(b) Given that $z \in \mathcal{H}$ with $z \neq 0$, set $\mathcal{C} := \mathbb{R}_+ z$. Then for each $s \in \mathcal{H}$,

$$P_{\mathcal{C}} s = \begin{cases} \dfrac{\langle s, z\rangle}{\|z\|^2} z, & \text{if } \langle s, z\rangle \geq 0 \\ \\ 0, & \text{otherwise.} \end{cases} \tag{2.1.25}$$

(c) Let $\mathcal{C} = B(0, 1)$, then for all $s \in \mathcal{H}$

$$P_{\mathcal{C}} s = \frac{1}{\max\{\|s\|, 1\}} s. \tag{2.1.26}$$

Proposition 2.1.48. *[6] Suppose $\mathcal{C}$ is a closed nonempty convex subset of $\mathcal{H}$. The projection $P_{\mathcal{C}}s$ of a point $s \notin \mathcal{C}$ is characterized by*

$$P_{\mathcal{C}}s \in \mathcal{C} \quad \text{and} \quad \langle z - P_{\mathcal{C}}s, s - P_{\mathcal{C}}s\rangle \leq 0 \quad \forall z \in \mathcal{C}, \tag{2.1.27}$$

or equivalently by

$$s - P_{\mathcal{C}}s \in N_{\mathcal{C}}(P_{\mathcal{C}}s). \tag{2.1.28}$$

From proposition (2.1.48), we obtain the following consequences:

i. $\|P_{\mathcal{C}}s - P_{\mathcal{C}}z\|^2 \leq \langle s - z, P_{\mathcal{C}}s - P_{\mathcal{C}}z\rangle$ for all $s, z \in \mathcal{H}$; that is, the metric projection is firmly non-expansive map.

ii. $\|s - P_{\mathcal{C}}s\|^2 + \|z - P_{\mathcal{C}}s\|^2 \leq \|s - z\|^2$ $\forall s \in \mathcal{H}$ and $z \in \mathcal{C}$.

iii. If $\mathcal{C}$ is a closed subspace, then $P_{\mathcal{C}}$ coincides with the orthogonal projection from $\mathcal{H}$ onto $\mathcal{C}$. That is, for any $s \in \mathcal{H}$ and $z \in \mathcal{C}$, $\langle z, s - P_{\mathcal{C}}s\rangle = 0$.

It is very crucial to note that both closeness and convexity of $\mathcal{C}$ cannot be dropped in proposition (2.1.48).

Proposition 2.1.49. *[6]*

Let $\mathcal{C}$ be a closed affine subspace of $\mathcal{H}$. then the following holds:

a. The projection $P_{\mathcal{C}}$ onto $\mathcal{C}$ is a weakly continuous affine operator.

b. $s - P_{\mathcal{C}}s$ is orthogonal to $\mathcal{C}$, that is, for any $s \in \mathcal{H}$ and $w, z \in \mathcal{C}$, $\langle s - P_{\mathcal{C}}s, w - z\rangle = 0$.

c. For every $s \in \mathcal{H}$ and $z \in \mathcal{C}$, $\|s - P_{\mathcal{C}}s\|^2 = \langle s - z, s - P_{\mathcal{C}}s.\rangle$.

Remark 2.1.50. *Weak continuity of projectors may fail to exist as can be seen in a case where $\mathcal{H}$ is an infinite-dimensional space and $\mathcal{C} = B_1(0, 1)$, a unit ball centered at the origin.*

Proposition 2.1.51. *Let $\mathcal{C}$ be a nonempty closed convex cone in $\mathcal{H}$, then*

$$P_{\mathcal{C}}(\delta s) = \delta P_{\mathcal{C}}s \quad \forall s \in \mathcal{H} \quad \text{and} \quad \delta \in \mathbb{R}_+. \tag{2.1.29}$$

The following are basic properties inherited by projectors:

Proposition 2.1.52. *[6] Suppose $\mathcal{C} \subset \mathcal{H}$ is a nonempty closed convex subset of $\mathcal{H}$ and $z \in \mathcal{H}$. Then the following are true:*
If:

i. $\mathcal{K} = s + \mathcal{C}$, where $s \in \mathcal{H}$, then $P_{\mathcal{K}}z = s + P_{\mathcal{C}}(z - s)$.

ii. $\mathcal{K} = \delta\mathcal{C}$, where $\delta \in \mathbb{R} - \{0\}$, then $P_{\mathcal{K}}z = \delta P_{\mathcal{C}}\left(\frac{z}{\delta}\right)$.

2.1.6 Basic Identities and Inequalities in Real Hilbert Spaces.

Let $s, v, z \in \mathcal{H}$, then the following are true:

(a)

$$|\langle s, v\rangle| \leq \|s\|\|v\| \quad \text{(Cauchy-Schwarz inequality).} \tag{2.1.30}$$

Moreover, $\langle s, v\rangle = \|s\|\|v\|$ provided there is $\alpha \in \mathbb{R}_+$ such that $s = \alpha v$ or $v = \alpha s$.

(b)

$$\|s + v\|^2 = \|s\|^2 + 2\langle s, v\rangle + \|v\|^2, \tag{2.1.31}$$

thus,

$$\|s - \alpha v\|^2 - \|s\|^2 = \alpha(\alpha\|v\|^2 - 2\langle s, v\rangle) \quad \forall \alpha \in \mathbb{R}. \tag{2.1.32}$$

(c)

$$\|s + v\|^2 + \|s - v\|^2 = 2\|s\|^2 + 2\|v\|^2 \quad \text{(Parallelogram identity).} \tag{2.1.33}$$

(d)

$$\langle s, v\rangle = \frac{1}{4}\left[\|s + v\|^2 - \|s - v\|^2\right] \quad \text{(Polarization identity).} \tag{2.1.34}$$

(e)

$$\langle s + v, s - v\rangle = \|s\|^2 - \|v\|^2. \tag{2.1.35}$$

(f) $\|s - v\|^2 = 2\|z - s\|^2 + 2\|z - v\|^2 - 4\left\|z - \dfrac{(s + v)}{2}\right\|^2$ (Apollonius's identity).

(g) $\|\beta s + (1 - \beta)v\|^2 = \beta\|s\|^2 + (1 - \beta)\|v\|^2 - \beta(1 - \beta)\|s - v\|^2 \quad \forall \beta \in \mathbb{R}.$

(h) $\|\beta s - (1 - \beta)v\|^2 = \beta\|s\|^2 + (1 - \beta)\|v\|^2 - \beta(1 - \beta)\|s + v\|^2 \quad \forall \beta \in \mathbb{R}.$

Proof.

(a) It is trivial.

(b)

$$\begin{aligned} \|s + v\|^2 &= \langle s + v, s + v\rangle \\ &= \langle s, s + v\rangle + \langle v, s + v\rangle \\ &= \langle s, s\rangle + \langle s, v\rangle + \langle v, s\rangle + \langle v, v\rangle \\ &= \|s\|^2 + 2\langle s, v\rangle + \|v\|^2. \end{aligned}$$

Hence, $\forall \alpha \in \mathbb{R}$,

$$\|s - \alpha v\|^2 = \|s\|^2 - 2\alpha\langle s, v\rangle + \alpha^2\|v\|^2.$$

This implies

$$\|s - \alpha v\|^2 - \|s\|^2 = \alpha(\alpha\|v\|^2 - 2\langle s, v\rangle).$$

(c) Adding together the expansions of $\|s+v\|^2$ and $\|s-v\|^2$ gives the result.

(d) Subtracting the expansion of $\|s-v\|^2$ from the expansion of $\|s+v\|^2$ gives the required result.

(e) Using the points $s = x+y$ and $z = x-y$ in (2.1.34) gives the desired result.

(f) Applying (2.1.33) to the points $\left(\frac{z-s}{2}\right)$ and $\left(\frac{z-v}{2}\right)$ gives the result.

(g) From (2.1.32), take $\alpha = +1$ and use the result in the expansion of $\|\beta s + (1-\beta)v\|^2$.

(h) Again, from (2.1.32) take $\alpha = -1$ and use the result in the expansion of $\|\beta s - (1-\beta)v\|^2$.

□

2.2 Equilibrium Problems

The concept of equilibrium problem (shortly EP), as mentioned earlier, is a unification of several mathematical problems such as variational inequalities, fixed point problem, optimization problems and a lot of other nonlinear convex problems which can be found in sciences and engineering (example, see [11, 14, 39, 53, 65] and references quoted therein). Recall that an EP in $\mathcal{H}$ is defined thus:

Suppose $f(\cdot,\cdot) : C \times C \to \mathbb{R}$ is a bifunction such that $f(z,z) = 0$ for all $z \in \mathcal{C}$.
An equilibrium problem for the bifunction f, is defined by:

$$\text{obtain } \bar{r} \in \mathcal{C} \quad \text{for} \quad \text{which } f(\bar{r}, z) \not\leq 0 \quad \forall z \in \mathcal{C}. \tag{2.2.1}$$

Where $EP(f,\mathcal{C}) := \{\bar{r} \in \mathcal{C} \mid f(\bar{r}, z) \geq 0 \quad \forall z \in \mathcal{H}\}$ is the solutions set.

Now, in what that follows, we give some examples of EP, thus:

2.2.1 Examples of Equilibrium Problems

Example 2.2.1. *Suppose $\mathcal{C} = [-1,1]$, and for each $s, z \in \mathcal{C}$, define the equilibrium bifunction $f : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{R}$ by*

$$f(s,z) = |s_1 + s_2|\left[(z_1^2 - s_1^2) + (z_2^2 - s_2^2)\right],$$

where $s = (s_1, s_2)$ and $z = (z_1, z_2)$.
Then, the set of solutions of EP(2.2.1) is

$$EP(f,\mathcal{C}) = \{\alpha(1,-1) \mid -1 \leq \alpha \leq 1\}.$$

Observe that the bifunction f is not pseudomonotone.

Example 2.2.2. *Find $s \in \mathcal{H} = \mathbb{R}^2$ such that $f(s,z) := \langle \mathcal{B}s, z-s\rangle \geq 0$ for each $z \in \mathcal{H}$, where*

$$\mathcal{B} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Now, for each $s, z \in \mathcal{H}$, we have

$$\begin{aligned} f(s,z) + f(z,s) &= \langle \mathcal{B}(s-z), z-s\rangle \\ &= (s_2 - z_2)(z_1 - s_1) - (s_1 - z_1)(z_2 - s_2) \\ &= 0. \end{aligned}$$

Thus, f is a monotone bifunction on $\mathcal{H}$. Also, it is easily verified that $\bar{r} = (0,0)^t$ is the only unique solution to the equilibrium problem since $f(\bar{r}, z) = 0$ for any $z \in \mathcal{H}$.

Example 2.2.3. *[50] Suppose $f : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{R}$ is an equilibrium bifunction defined by*

$$f(s,z) = \sum_{j=1}^{5} (z_j - s_j)\|s\|, \quad \text{for each } s, z \in \mathbb{R}^5,$$

where $\mathcal{C} \subset \mathbb{R}^5$ is defined by $\mathcal{C} = \{(s_1, s_2, ..., s_5) : s_1 \geq -1,\ s_j \geq 1, \quad j = 2, ..., 5\}$, $s = (s_1, ..., s_5)$ and $z = (z_1, ...z_5)$.

Then,

$$EP(f, \mathcal{C}) = \{(s_1, 1, 1, 1, 1) : s_1 \geq -1\}.$$

Notice that f is pseudomonotone and satisfies the Lipschitz-type conditions with the Lipschitz constants $\delta_1 = \delta_2 = 2$.

Example 2.2.4. *Consider $\mathcal{H} = \mathbb{R}$ and let $\mathcal{C} = [-10, 10]$. Define an equilibrium bifunction $f : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{R}$ by*

$$f(s,z) = 4z^2 + sz - 5s^2, \quad \forall s, z \in \mathcal{C}.$$

Then, $f(s,z) + f(z,s) = -(s-z)^2 \leq 0$ for each $s, z \in \mathcal{C}$, thus f is a monotone bifunction, and $EP(f, \mathcal{C}) = \{0\}$.

2.2.2 Previous Works on Equilibrium Problems in Real Hilbert Spaces

Much progress has been made in the study of equilibrium problems both in finite and infinite dimensional spaces by numerous researchers but has yet to be fully exploited. Thus, equilibrium theory has yet to take advantage of some of the recent progress in optimization, where much has be learned about how to handle very general classes of set $\mathcal{C}$ directly, without necessarily relying on excursions through convex analysis.

Equilibrium problem became a well known path in nonlinear convex analysis after the work of Blum and Oettli [14]. Since then, EP(2.2.1) has been studied by many authors

who have developed and analyzed several effective schemes for finding an approximate solution in the solutions set, EP(f , $\mathcal{C}$), we refer you to [2, 37, 63, 84, 77]. Early in the study of EP(2.2.1), proximal point method (PPM) gained more interest as it was the first successful scheme for finding an approximate solution to the EP. Moudafi in [63] proposed the Proximal Point Method (PPM) defined by $x_{n+1} = \mathcal{T}_{f_\lambda}(x_n)$, $n \in \mathbb{N}$, $\lambda > 0$, $x_0 \in \mathcal{C}$. The map $\mathcal{T}_{f_\lambda} : \mathcal{H} \to \mathcal{C}$ is described by

$$\mathcal{T}_{f_\lambda}(x) = \{s \in \mathcal{C} : f(s,u) + \frac{1}{\lambda}\langle u - s, s - x\rangle \geq 0, \quad \forall\ u \in \mathcal{C}\}. \tag{2.2.2}$$

One more effective scheme is the extragradient method (EGM), [29]. This method helps to overcome the challenges faced in computing the value of $\mathcal{T}_{f_\lambda}$ in PPM. The EGM involves the solving of two optimization programs over each iteration. Thus, the algorithm is simply defined by:

$$\begin{cases} y_n = arg\min\limits_{s\in\mathcal{C}} \left\{\lambda f(x_n, s) + \dfrac{\|s - x_n\|^2}{2}\right\}, \\ x_{n+1} = arg\min\limits_{s\in\mathcal{C}} \left\{\lambda f(y_n, s) + \dfrac{\|s - x_n\|^2}{2}\right\}, \text{ with } x_0 \in \mathcal{C} \text{ and } \lambda > 0. \end{cases} \tag{2.2.3}$$

This method was further studied and extended to bifunctions that are pseudomonotone and satisfy the Lipschitz-type condition, see [77]. The authors in [17] modified and improved the Korpelevich's extagradient methods, also known as the proximal-like algorithm, by proposing the subgradient extragradient method. According to them, if given the current iterate $\{x_n\}$ and a starting point $\{x_0\} \in \mathcal{C}$, with $\lambda > 0$, the next iterate $\{x_{n+1}\}$ is computed as follows:

$$\begin{cases} y_n = arg\min\limits_{s\in\mathcal{C}} \left\{\lambda f(x_n, s) + \dfrac{\|s - x_n\|^2}{2}\right\}, \\ x_{n+1} = arg\min\limits_{s\in T_n} \left\{\lambda f(y_n, s) + \dfrac{\|s - x_n\|^2}{2}\right\}, \\ \text{where } T_n = \{u \in \mathcal{H} : \langle x_n - \lambda v_n - y_n, u - y_n\rangle \leq 0\}. \end{cases} \tag{2.2.4}$$

Under the conditions that the function f is monotone and Lipschitz continuous, they obtained weak convergence. It is notable here that the stepsize used throughout their work is a constant stepsize that depends on the Lipschitz constant. Hieu [37] in his scheme avoided the challenge of constant stepsizes by replacing it with variable stepsizes and obtained strong convergence under the assumptions that the cost bifunction satisfies the Lipschitz-type condition and as well is pseudomonotonic.

It is worth noting that without additional conditions, EGM and PPM give weak convergence. In order to have strong convergence of (2.2.2) and (2.2.3), Vuong *et al.* in 2012 [93, Algorithm 1] proposed a successful algorithm which is the hybrid projection-type extragradient method (HPEGM). The sequence $\{x_n\}$ generated by this method converges

strongly to $P_{EP(f,\,\mathcal{C})}(x_0)$. Their algorithm is of the form:

$$\begin{cases} q_n = arg\min\left\{\lambda_n f(x_n, s) + \dfrac{\|s - x_n\|^2}{2} \quad : s \in \mathcal{C}\right\}, \\ z_n = arg\min\left\{\lambda_n f(q_n, s) + \dfrac{\|s - x_n\|^2}{2} \quad : s \in \mathcal{C}\right\}, \\ B_n = \{v \in \mathcal{C} : \|z_n - v\| \leq \|x_n - v\|\}, \\ T_n = \{v \in \mathcal{C} : \langle v - x_n, x_0 - x_n\rangle \leq 0\}, \\ x_{n+1} = P_{Bn} \cap T_n(x_0), \end{cases} \tag{2.2.5}$$

where $x_0 \in \mathcal{C},\ 0 < \lambda_{min} \leq \lambda_{max} < min\left\{\dfrac{1}{2\delta_1}, \dfrac{1}{2\delta_2}\right\}$, and $\delta_1 > 0$, $\delta_2 > 0$ are Lipschitz-type constants of f.

Nguyen *et al.* [70, Algorithm 1] proposed another projection-type scheme for solving EP(2.2.1) called the shrinking projection-type extragradient method (SPEGM). These authors proved that the generated sequence by the scheme converges strongly. Also, the authors in [94, Alorithm 1] proposed an extragradient-viscosity method (EGVM) and proved that it converges strongly, and to overcome the tough challenge of computing the Lipschitz-type constants, the authors incorporated the Armijo-backtracking linesearch in the extragradient iteration. This is another setback since many computations are required in order to obtain a nearly optimal step-size. The algorithm for solving EP(2.2.1) as proposed by the authors states, thus:
$\{\theta_n\} \subset (0, 1), \quad \{\alpha_n\} \subset [0, 1), \quad \{\lambda_n\} \subset (0, \infty), \quad x_0 \in \mathcal{H}.$

$$\begin{cases} u_n = arg\min\left\{\lambda_n f(x_n, u) + \dfrac{\|u - x_n\|^2}{2} \quad : u \in \mathcal{C}\right\}, \\ z_n = arg\min\left\{\lambda_n f(u_n, u) + \dfrac{\|u - x_n\|^2}{2} \quad : u \in \mathcal{C}\right\}, \\ x_{n+1} = (1 - \theta_n)t_n + \theta_n S t_n, \\ t_n = z_n \alpha_n F z_n, \end{cases} \tag{2.2.6}$$

where $F : \mathcal{C} \longrightarrow \mathcal{H}$ is a given strongly monotone and Lipschitz continuous operator. It is important to note that the advantage of EGVM over HPEGM and SPEGM is the easier computation of the next iterate, x_{n+1}. Hieu *et al.* in [41, Section 1] remarked that in a unique special situation where $F = I - x_0$, with x_0 a given point in $\mathcal{C}$, EGVM becomes the Halpern-type extragradient method (HEGM) and the sequence $\{x_n\}$ generated by HEGM strongly converges to $P_{EP(f,\mathcal{C})}(x_0)$.

Recently, the quest to obtain faster convergent algorithms of EP(2.2.1) with mild conditions has been of great interest, researchers discovered that the introduction of inertial and relaxation factors to algorithms accelerates their convergence. On this note, Moudafi [64, Algorithm 1.5] proposed an inertial-proximal method for solving EP(2.2.1), thus:

Given the initial iterates $x_{n-1},\ x_n \in \mathcal{C}$ and $\alpha_n \in [0,1),\quad \lambda_n > 0,$

$$f(x_{n+1},v) + \frac{1}{\lambda_n}\langle x_{n+1} - x_n - \alpha_n(v - x_{n-1}), v - x_{n+1}\rangle \geq -\xi_n,$$
$$\forall v \in \mathcal{C}, \quad \xi_n \geq 0, \quad \forall n \in \mathbb{N}. \tag{2.2.7}$$

Also, Attouch and Cabot in their work [3], proposed a relaxed inertial proximal algorithm for maximally monotone operators. In a Similar way, Shehu *et al.* [86] constructed a relax inertial modified subgradient extragradient method for EP(2.2.1). According to them, x_{n+1} is computed thus:

$$\begin{cases} w_n = \alpha_n x_0 + (1-\alpha_n)x_n + \theta_n(x_n - x_{n-1}), \\ y_n = arg\min\limits_{y\in\mathcal{C}}\left\{\lambda_n f(w_n,y) + \frac{1}{2}\|y - w_n\|^2\right\}, \\ T_n = \{z \in \mathcal{H} \mid \langle w_n - \lambda_n v_n - y_n, z - y_n\rangle \leq 0\}, \\ z_n = arg\min\limits_{y\in T_n}\left\{\lambda_n f(y_n,y) + \frac{1}{2}\|y - w_n\|^2\right\}, \\ \\ x_{n+1} = (1-\tau)w_n + \tau z_n, \\ \\ \lambda_{n+1} = \begin{cases} \min\left\{\dfrac{\mu(\|w_n - y_n\|^2 + \|z_n - y_n\|^2)}{2(f(w_n,z_n) - f(w_n,y_n) - f(y_n,z_n))}, \lambda_n\right\}, & \text{if } f(w_n,z_n) \\ & -f(w_n,y_n) - f(y_n,z_n) > 0, \\ \lambda_n & \text{otherwise}, \end{cases} \\ \text{where } \tau \in (0,\frac{1}{2}], \quad \theta_n \in [0,\frac{1}{3}) \quad \mu,\alpha_n \in (0,1) \quad x_0,x_1 \in \mathcal{H}, \quad \lambda_1 > 0. \end{cases} \tag{2.2.8}$$

More so, Thong *et al.* in [89, Algorithm 1] proposed a relaxed inertial Tseng's method-based algorithm for finding the solutions set of a variational inequality and fixed point problem as follows:
Given, for $n \geq 1$, the current iterates x_{n-1}, x_n, the next iterate x_{n+1} is computed thus,

$$\begin{cases} w_n = (1-\alpha_n)(x_n + \theta_n(x_n - x_{n-1})), \\ y_n = arg\min\limits_{y\in\mathcal{C}}\left\{\tau_n f(w_n,y) + \frac{1}{2}\|w_n - y\|^2\right\}, \\ \\ x_{n+1} = (1-\beta_n)z_n + \beta_n S z_n, \\ \\ \tau_{n+1} = \begin{cases} \min\left\{\dfrac{\mu\|w_n - y_n\|}{\|f(y_n,y) - f(w_n,y)\|}, \tau_n + \lambda_n\right\} & \text{if } f(y_n,y) - f(w_n,y) \neq 0, \\ \\ \tau_n + \lambda_n & \text{otherwise}, \end{cases} \\ \\ \text{where } z_n = y_n - \tau_n(f(y_n,y) - f(w_n,y)) \\ \text{and } \theta > 0, \quad \tau_1 > 0, \quad x_0,x_1 \in \mathcal{H}, \quad \mu \in (0,1). \end{cases} \tag{2.2.9}$$

The authors using the on-line rule

$$\bar{\theta}_n = \begin{cases} \min\left\{\theta, \dfrac{\varepsilon_n}{\|x_n - x_{n-1}\|}\right\} & \text{if } \|x_n - x_{n-1}\| \neq 0, \\ \theta & \text{otherwise,} \end{cases}$$

and the assumptions that f satisfies the Lipschitzness and pseudomonotonicity conditions, were able to establish the strong convergence of their method.

Several other inertial type algorithms for solving equilibrium problem (2.2.1) have been constructed and analyzed by many different authors, (see, [34, 36, 38, 44, 69, 86], and the quoted references therein). In this dissertation, as we shall see in the subsequent chapters, we present two new methods of approximating the real solutions of EP(2.2.1).

2.2.3 Equilibrium Problem as a Generalization of Other Convex Problems

As earlier mentioned that equilibrium problem unifies other mathematical programming problems, in here, we present the relationship between equilibrium problem and other related problems (for more details, please check [14]). For $\mathcal{C} \subset \mathcal{X}$, a nonempty convex subset of the Banach space, $\mathcal{X}$, we have the following:

1. *Optimization problem:*
   Let $\tau : \mathcal{C} \longrightarrow \mathbb{R}$. Optimization problem is defined as:

   $$\text{find } s^* \in \mathcal{C} \text{ such that } \tau(s^*) \leq \tau(z) \quad \forall z \in \mathcal{C}, \tag{2.2.10}$$

   written as $\min\{\tau(s) : s \in \mathcal{C}\}$.

   Set $f(s, z) := \tau(z) - \tau(s)$, then (2.2.10) coincides with EP(2.2.1), and f is a monotone bifunction.

2. *Saddle point problem:*
   Let $\Gamma : \mathcal{D}_1 \times \mathcal{D}_2 \longrightarrow \mathbb{R}$. Then $s^* = (s_1^*, s_2^*)$ is said to be a saddle point of Γ if and only if

   $$(s_1^*, s_2^*) \in \mathcal{D}_1 \times \mathcal{D}_2, \ \Gamma(s_1^*, z_2) \leq \Gamma(z_1, s_2^*) \ \ \forall (z_1, z_2) \in \mathcal{D}_1 \times \mathcal{D}_2. \tag{2.2.11}$$

   Set $\mathcal{C} := \mathcal{D}_1 \times \mathcal{D}_2$ and define $f : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{R}$ by

   $$f(s, z) := \Gamma(z_1, s_2) - \Gamma(s_1, z_2),$$

   where $s = (s_1, s_2)$, $z = (z_1, z_2)$. Then $s^* = (s_1^*, s_2^*)$ is a solution of EP(2.2.1) if and only if (s_1^*, s_2^*) satisfies (2.2.11). Notice that f is monotone.

3. *Nash equilibria in non-cooperative games problem:*
Let I be a finite index set denoting the set of players. For every $i \in I$, let $\mathcal{C}_i$, the strategy set of the $i-th$ player, be given. Let $\mathcal{C} := \prod_{i\in I} \mathcal{C}_i$. For each $i \in I$, let the function $f_i : \mathcal{C} \longrightarrow \mathbb{R}$, which is the loss function of the $i-th$ player depending on the strategy of all players, be given. For $s = (s_i)_{i\in I} \in \mathcal{C}$, we define $s^i := (s_j)_{j\in I,\ j\neq i}$. The point $s^* = (s_i^*)_{i\in I} \in \mathcal{C}$ is called a Nash equilibrium if and only if

$$f_i(s^*) \leq f_i(s^{*i}, z_i) \tag{2.2.12}$$

holds for every $i \in I,\ z_i \in \mathcal{C}$. (I.e., no player can reduce his loss by varying his strategy alone).
Define $f : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{R}$ by

$$f(s, z) := \sum_{i\in I} (f_i(s^i, z_i) - f_i(s)).$$

Then, $s^* \in \mathcal{C}$ is a Nash equilibrium if and only if s^* satisfies EP(2.2.1). This is true, since EP(2.2.1) is fulfilled whenever (2.2.12) holds for each $i \in I$. Also, if for some $i \in I$ we select $z \in \mathcal{C}$ in such a way that $s^{*i} = z^i$, therefore, we have

$$0 \leq f(s^*, z) = f_i(s^{*i}, z_i) - f_i(s^*).$$

Thus, EP(2.2.1) implies (2.2.12) for any $i \in I$. Notice that f is not automatically monotone.

4. *Convex differentiable optimization problem:*
Suppose $\Gamma : \mathcal{X} \longrightarrow \mathbb{R}$ is a convex and Gâteaux differentiable function with Gâteaux differential at $s \in \mathcal{C}$ denoted by $d\Gamma(s; h) \in \mathcal{X}^*$.

Consider the optimization problem

$$\min\{\Gamma(s) : s \in \mathcal{C}\}. \tag{2.2.13}$$

Recall from convex analysis that s^* is a solution of (2.2.13), if and only if s^* satisfies the variational inequality of finding $s^* \in \mathcal{C}$, such that $\langle d\Gamma(s^*; h), z - s^*\rangle \geq 0\ \forall z \in \mathcal{C}$. Set $f(s, z) := \langle d\Gamma(s; h), z - s\rangle$, then the problem is now an EP(2.2.1).
Notice that $d\Gamma$ is monotone. That is,

$$\begin{aligned}
& \langle d\Gamma(z; h) - d\Gamma(s; h), z - s\rangle \geq 0 \quad \forall s, z \in \mathcal{C} \\
\Longrightarrow\quad & -\langle d\Gamma(z; h), s - z\rangle - \langle d\Gamma(s; h), z - s\rangle \geq 0 \\
\Longrightarrow\quad & -f(z, s) - f(s, z) \geq 0 \\
\Longrightarrow\quad & f(s, z) + f(z, s) \leq 0.
\end{aligned}$$

Hence, f inherits the monotonicity of $d\Gamma$.

5. *Fixed points problem:*
Let $\mathcal{X} = \mathcal{X}^*$ be a Hilbert space and $\mathcal{T} : \mathcal{C} \longrightarrow \mathcal{C}$ be a given mapping. The fixed point problem is described as:

$$\text{obtain } s^* \in \mathcal{C} \text{ such that } s^* = \mathcal{T}s^*. \tag{2.2.14}$$

Set $f(s,z) := \langle s - \mathcal{T}s, z - s\rangle$, therefore s^* solves EP(2.2.1) if and only if s^* is a solution of (2.2.14).
It is obviously clear that (2.2.14) implies EP(2.2.1). Also, suppose that EP(2.2.1) is fulfilled, then select $z := \mathcal{T}s^*$ to have

$$\begin{aligned} 0 \leq f(s^*, z) &= \langle s^* - \mathcal{T}s^*, \mathcal{T}s^* - s^*\rangle \\ &= -\|s^* - \mathcal{T}s^*\|^2 \\ &\leq 0, \end{aligned}$$

hence, $s^* = \mathcal{T}s^*$. Therefore, EP(2.2.1) implies (2.2.14). In this case, f is a monotone bifunction if and only if $\mathcal{T}$ is non-expansive, that is

$$\langle \mathcal{T}(s-z), s-z\rangle \leq \|s-z\|^2, \quad \forall s, z \in \mathcal{C}.$$

Definition 2.2.5. *The mapping $\mathcal{T} : \mathcal{C} \subset \mathcal{H} \longrightarrow \mathcal{C} \subset \mathcal{H}$ is said to be $\alpha-$strict pseudo-contractive if there is a constant $\alpha \in (0,1)$ for which*

$$\|\mathcal{T}s - \mathcal{T}z\|^2 \leq \|s-z\|^2 + \alpha\|(\boldsymbol{I} - \mathcal{T})(s-z)\|^2, \quad \forall s, z \in \mathcal{C}, \tag{2.2.15}$$

or, equivalently,

$$\langle \mathcal{T}(s-z), s-z\rangle \leq \|s-z\|^2 - \frac{1-\alpha}{2}\|(\boldsymbol{I} - \mathcal{T})(s-z)\|^2, \quad \forall s, z \in \mathcal{C}. \tag{2.2.16}$$

Theorem 2.2.6. *Let $\mathcal{C}$ be a nonempty closed convex subset of a real Hilbert space $\mathcal{H}$. If $\mathcal{T} : \mathcal{C} \longrightarrow \mathcal{C}$ is $\alpha-$strict pseudo-contractive, where $\alpha \in (0,1)$, then $f : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{R}$, defined by $f(s,z) := \langle s - \mathcal{T}s, z - s\rangle \ \ \forall s, z \in \mathcal{C}$, is a pseudomonotone bifunction.*

Proof.
Given that $f(s,z) = \langle s - \mathcal{T}s, z - s\rangle \ \ \forall s, z \in \mathcal{C}$.
Using (2.2.16) we have

$$\begin{aligned} f(z,s) &= \langle z - \mathcal{T}z, s - z\rangle \\ &= \langle z - s + s - \mathcal{T}s + \mathcal{T}s - \mathcal{T}z, s - z\rangle \\ &= -\|s-z\|^2 + \langle s - \mathcal{T}s, s - z\rangle + \langle \mathcal{T}(s-z), s - z\rangle \\ &\leq -\|s-z\|^2 + \langle \mathcal{T}(s-z), s - z\rangle \quad \forall s, z \in \mathcal{C} \\ &= -\|s-z\|^2 + \|s-z\|^2 - \frac{1-\alpha}{2}\|(\mathbf{I} - \mathcal{T})(s-z)\|^2 \\ &= -\frac{1-\alpha}{2}\|(\mathbf{I} - \mathcal{T})(s-z)\|^2 \\ &\leq 0. \end{aligned}$$

Thus, f is pseudomonotone. □

Moreover, using Cauchy-Schwarz and triangle inequalities, and letting

$\mathcal{T}$ be $\dfrac{2-\alpha}{1-\alpha}$−Lipschitz continuous, we have that,

$$\begin{aligned} f(s,z) &= \langle s-\mathcal{T}s, z-s\rangle \\ &= \langle s-\mathcal{T}s, w-s\rangle + \langle w-\mathcal{T}w, z-w\rangle + \langle s-\mathcal{T}s+\mathcal{T}w-w, z-w\rangle \\ &= f(s,w)+f(w,z)+\langle s-w+\mathcal{T}w-\mathcal{T}s, z-w\rangle \\ &\leq f(s,w)+f(w,z)+(\|s-w\|+\|\mathcal{T}w-\mathcal{T}s\|)\|z-w\| \\ &\leq f(s,w)+f(w,z)+\|s-w\|\|z-w\|+\frac{2-\alpha}{1-\alpha}+\|s-w\|\|z-w\| \\ &= f(s,w)+f(w,z)+\frac{3-2\alpha}{1-\alpha}\|s-w\|\|z-w\| \\ &\leq f(s,w)+f(w,z)+\frac{3-2\alpha}{2(1-\alpha)}\left(\|s-w\|^2+\|z-w\|^2\right). \end{aligned}$$

Hence, f satisfies Lipschitz-type condition on $\mathcal{C}$ with constants $\delta_1=\delta_2=\dfrac{3-2\alpha}{2(1-\alpha)}$, where $\alpha\in(0,1)$.

6. *Variational inequality problem:*
Let $\mathcal{B}:\mathcal{C}\longrightarrow\mathcal{X}^*$ be a given mapping. A variational inequality problem is a convex problem defined as follows:

$$\text{find } s^*\in\mathcal{C} \text{ so that}$$

$$\langle \mathcal{B}s^*, z-s^*\rangle \geq 0 \quad \forall z\in\mathcal{C}. \tag{2.2.17}$$

Set $f(s,z):=\langle \mathcal{B}s, z-s\rangle$, then, it is easily verified that s^* is a solution of EP(2.2.1) if and only if it is a solution of (2.2.17).

If $\mathcal{B}:\mathcal{C}\rightrightarrows\mathcal{X}^*$ is a multivalued mapping with $\mathcal{B}(s)$ compact, convex and nonempty for each $s\in\mathcal{C}$. Then (2.2.17) becomes a varitional inequality with multivalued mapping problem defined as:

$$\text{obtain } s^*\in\mathcal{C},\ r^*\in Ts^* \text{ such that}$$

$$\langle r^*, z-s^*\rangle \geq 0\ \forall z\in\mathcal{C}. \tag{2.2.18}$$

If we set $f(s,z):=\min\limits_{r\in Ts}\langle r, z-s\rangle$, then s^* solves EP(2.2.1) if and only if together with some $r^*\in\mathcal{C}$ it is a solution of (2.2.18).

7. *Complementarity prolems:*
Complementarity problem is a special case of variational inequality problem. Suppose $\mathcal{C}$ is a convex cone such that

$$\mathcal{C}^*:=\{s^*\in\mathcal{X}^*:\langle s^*, z\rangle\geq 0 \quad \forall z\in\mathcal{C}\}$$

denotes its polar cone. Let $T : \mathcal{C} \longrightarrow \mathcal{X}^*$ be any given operator. The complementarity problem is asked as:

$$\text{obtain all } \bar{r} \in \mathcal{C} \text{ for which}$$

$$T\bar{r} \in \mathcal{C}^* \text{ and}$$

$$\langle T\bar{r}, \bar{r} \rangle = 0. \tag{2.2.19}$$

Observe that (2.2.19) is equivalent to (2.2.17). In fact, it is obvious that (2.2.19) implies (2.2.17). Suppose (2.2.17) is true, then putting $z = 2\bar{r}$ and $z = 0$ respectively into (2.2.17) gives that $\langle T\bar{r}, \bar{r} \rangle = 0$, so that $\langle T\bar{r}, z \rangle \geq 0$ for each $z \in \mathcal{C}$. Hence, (2.2.17) implies (2.2.19), thus $\bar{r}$ solves EP(2.2.1) if and only if it is a solution to (2.2.19).

2.2.4 Existence of Solutions for Equilibrium Problems

An ill-posed equilibrium problem may cease to have equilibria, hence, it is crucial to know when an equilibrium problem has solutions. We dedicate this part of our work to provide with some important facts that guarantee the existence and characteristics of solutions to EP(2.2.1).

Suppose $\mathcal{C}$ is a closed convex nonempty subset of a Banach space $\mathcal{X}$ and $f : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{R}$ is a bifunction defined such that $f(s, s) = 0$ for each $s \in \mathcal{C}$.

$$\text{Define } \; EP((f, \mathcal{C}))(z) := \{z \in \mathcal{C} \mid f(s, z) \geq 0\}.$$

Assume

i. $f(\cdot, z)$ is hemicontinuous for every $z \in \mathcal{C}$,

ii. $f(s, \cdot)$ is semi-strictly quasiconvex and lower semi-continuous for every $s \in \mathcal{C}$,

iii. $f(s, z)$ is quasimonotone on $\mathcal{C} \times \mathcal{C}$,

iv. $f(s, z) = 0, \; f(s, \bar{z}) > 0 \implies f(s, z_t) > 0 \;\; \forall z_t = t\bar{z} + (1 - t)z, \quad 0 < t < 1,$

v. $\mathcal{C}$ is a convex $\alpha-$body, that is, the algebraic kernel of $\mathcal{C}$ $Ai(\mathcal{C})$ is nonempty.

Furthermore, if $\mathcal{C}$ is not compact, we assume the following coercivity condition:

vi. there exists a compact subset $D \subset \mathcal{X}$ and a vector $z_0 \in \mathcal{C} \cap D$ such that $f(s, z_0) < 0$ for every $s \in \mathcal{C} \setminus D$.

Theorem 2.2.7. *If Assumptions (i)—(vi) hold, then a vector $\bar{r} \in \mathcal{C}$ solves the EP(2.2.1) for all $z \in \mathcal{C}$, if and only if $\bar{r} \in \bigcap_{z \in \mathcal{C}} EP(f, \mathcal{C})(z)$, i.e., $\bigcap_{z \in \mathcal{C}} EP(f, \mathcal{C})(z) \neq \emptyset$.*

A detailed proof of this Theorem 2.2.7 can be found in [9]

Remark 2.2.8.

a. *If Assumptions (i), (ii), (vi) hold and f is a pseudomonotonic bifunction, then the solutions set $EP(f, \mathcal{C})$ is nonempty, convex and compact.*

b. *If Assumptions (i), (ii), (vi) are satisfied and f is strictly pseudomonotonic, then the solution of EP(2.2.1) is unique.*

c. *Pseudomonotonicity is a sufficient condition for the existence of a solution of EP.*

Chapter 3

Relaxed Inertial Algorithm for Solving Equilibrium Problems

3.1 Introduction

In this chapter, we study the equilibrium problems defined as follows:
Let $\mathcal{H}$ be a real Hilbert space, and $\mathcal{C}$ a convex nonempty subset of $\mathcal{H}$. Let $f : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{R}$ be an equilibrium bifunction defined such that $f(z, z) = 0, \ \forall z \in \mathcal{C}$. It is required to

$$\text{obtain } \bar{r} \in \mathcal{C} \text{ for which } f(\bar{r}, z) \not< 0, \quad \forall z \in \mathcal{C}. \tag{3.1.1}$$

Let $EP(f, \mathcal{C}) := \{s \in \mathcal{C} \mid f(s, z) \geq 0, \quad \forall z \in \mathcal{C}\}$ denote the solutions set.

In what that follows, we present a new iterative scheme of obtaining $\bar{r} \in \mathcal{C}$ for which $\bar{r} \in EP(f, \mathcal{C})$ and achieve weak convergent result. Finally, the applications of our result will be extended to some variational inequality and fixed point problems.

3.2 Preliminaries

Now, we introduce some fundamental definitions and lemmas that are necessary for the convergence analysis of our proposed method. The notations "$\rightarrow$" and "$\rightharpoonup$" will be used respectively to mean the strong and weak convergence.

Suppose that for each $n \geq 1$, the sequence $\{x_n\}$ recursively satisfies $x_{n+1} = x_n + \theta_n(x_n - x_{n-1})$, (see [3]).

Then, for $n = 1$,

$$x_2 \quad = \quad x_1 + \theta_1(x_1 - x_0).$$

This implies

$$x_2 - x_1 \quad = \quad \theta_1(x_1 - x_0).$$

For $n = 2$

$$\begin{aligned} x_3 - x_2 &= \theta_2(x_2 - x_1) \\ &= \theta_2 \cdot \theta_1(x_1 - x_0). \end{aligned}$$

For $n = 3$

$$\begin{aligned} x_4 - x_3 &= \theta_3(x_3 - x_2) \\ &= \theta_3 \cdot \theta_2 \cdot \theta_1(x_1 - x_0) \\ \vdots &= \vdots \end{aligned}$$

$$x_{n+1} - x_n = \left(\prod_{j=1}^{n} \theta_j\right)(x_1 - x_0),$$

and gives

$$x_n = x_1 + \left(\sum_{l=1}^{n-1} \prod_{j=1}^{l} \theta_j\right)(x_1 - x_0). \tag{3.2.1}$$

Hence, the sequence $\{x_n\}$ will converge if and only if $x_1 = x_0$ or whenever $\sum_{l=1}^{\infty} \prod_{j=1}^{l} \theta_j < \infty$. Thus, for the convergence analysis of our scheme, we will subsequently assume that

$$\sum_{l=i}^{\infty} \left(\prod_{j=i}^{l} \theta_j\right) < \infty \quad \text{for each} \quad i \geq 1, \tag{3.2.2}$$

and we define the real sequence $\{t_i\}$ by

$$t_i := \sum_{l=i-1}^{\infty} \left(\prod_{j=i}^{l} \theta_j\right) = 1 + \sum_{l=i}^{\infty} \left(\prod_{j=i}^{l} \theta_j\right). \tag{3.2.3}$$

Let $\{t_{i,n}\}$ be a sequence defined, for each $i, n \geq 1$, by

$$t_{i,n} := \begin{cases} \sum_{l=i-1}^{n-1} \left(\prod_{j=i}^{l} \theta_j\right) = 1 + \sum_{l=i}^{n-1} \left(\prod_{j=i}^{l} \theta_j\right), & i \leq n \\ 0 & \text{otherwise,} \end{cases} \tag{3.2.4}$$

conventionally, for $l = i - 1$, we take $\prod_{j=i}^{i-1} \theta_j = 1$.
One can obtain from (3.2.4) that

$$t_{i,n} = 1 + \theta_i t_{i+1,n} \quad \text{for each} \quad i \geq 1, n \geq i + 1. \tag{3.2.5}$$

For every $i \geq 1$, the sequence $\{t_{i,n}\}$ converges increasingly to $\{t_i\}$. That is, $t_i = \lim_{n\to\infty} t_{i,n}$, hence, (3.2.2) guarantees that we can have from (3.2.5) a well defined sequence,

$$t_i = 1 + \theta_i t_{i+1} \quad \text{for each} \quad i \geq 1. \tag{3.2.6}$$

Using '$\sim$' to mean the asymptotic equivalence, the proposition below provides a condition for ensuring that assumption (3.2.2) is true.

Proposition 3.2.1. ***([3, Proposition 3.1], [44, Proposition 2.3])*** *Let $\{\theta_n\}$ be a real sequence such that $\theta_n \in [0,1)$ for each $n \geq 1$. Suppose, for some $c \in [0,1)$, that*

$$\lim_{n\to\infty}\left(\frac{1}{1-\theta_{n+1}} - \frac{1}{1-\theta_n}\right) = c.$$

Then

(i) the assumption defined in (3.2.2) is true, and

$$t_{n+1} \sim \frac{1}{(1-c)(1-\theta_n)} \quad as \quad n \to \infty,$$

(ii) the equivalence $1-\theta_n \sim 1-\theta_{n+1}$ holds true as $n \to \infty$. Hence, $t_{n+1} \sim t_{n+2}$ as $n \to \infty$.

Remark 3.2.2. ***(see also [44])***

One can see that the sequence $\theta_n = 1 - \frac{\bar{\theta}}{n}$, $\bar{\theta} > 1$, is a good example of a sequence satisfying assumption (3.2.2). In fact, we can get that

$$\begin{aligned}\lim_{n\to\infty}\left(\frac{1}{1-\theta_{n+1}} - \frac{1}{1-\theta_n}\right) &= \lim_{n\to\infty}\left(\frac{1}{\bar{\theta}}(n+1) - \frac{1}{\bar{\theta}}n\right)\\ &= \frac{1}{\bar{\theta}} \in [0,1),\end{aligned}$$

satisfies the assumption of Proposition 3.2.1. Therefore, by proposition 3.2.1(i), assumption (3.2.2) indeed holds.

Definition 3.2.3.

i. Let $\partial f(s,t)$ denote the sub-differential of $f(s,.)$ at t, then,

$$\begin{aligned}\partial f(s,t) &:= \{u \in \mathcal{H} : f(s,v) \geq f(s,t) + \langle u, v-t\rangle \quad \forall v \in \mathcal{H}\}. \text{ Thus,}\\ \partial f(s,s) &:= \{u \in \mathcal{H} : f(s,v) \geq \langle u, v-s\rangle \quad \forall v \in \mathcal{H}\}.\end{aligned}$$

ii. The normal cone $\mathcal{N}_C$ of $\mathcal{C}$ at s is defined by:

$$\mathcal{N}_C(s) = \{v \in \mathcal{H} : \langle v, t-s\rangle \leq 0 \quad \forall t \in \mathcal{C}\}.$$

Lemma 3.2.4. ***([3, Lemma B.1])*** *Let $\{c_n\}$, $\{d_n\}$ and $\{\theta_n\}$ be sequences of real numbers satisfying*

$$c_{n+1} \leq \theta_n c_n + d_n \quad for\ each \quad n \geq 1,$$

suppose $\theta_n \geq 0$ $\forall n \geq 1$.

(i) For each $n \geq 1$, we have

$$\sum_{i=1}^{n} c_i \leq t_{1,n}c_1 + \sum_{i=1}^{n-1} t_{i+1,n}d_i$$

where the double sequence $\{t_{i,n}\}$ is defined by (3.2.4);

(ii) Under assumption (3.2.2), suppose the sequence $\{t_i\}$ defined by (3.2.3) satisfies $\sum_{i=1}^{\infty} t_{i+1}[d_i]_+ < \infty$. Then the series $\sum_{i=1}^{\infty}[c_i]_+$ converges, and

$$\sum_{i=1}^{\infty}[c_i]_+ \leq t_1[c_1]_+ + \sum_{i=1}^{\infty} t_{i+1}[d_i]_+,$$

where $[t]_+ := \max\{t, 0\}\ \forall t \in \mathbb{R}$.

Lemma 3.2.5. *[6] For each $x \in \mathcal{H}$, $y \in \mathcal{C}$ and positive constant λ, the following inequality is true*

$$\lambda(g(y) - g(prox_{\lambda g}(x))) \geq \langle x - prox_{\lambda g}(x), y - prox_{\lambda g}(x)\rangle,$$

where $prox_{\lambda g}(x) := \underset{y\in\mathcal{C}}{arg\,min}\left\{\lambda g(y) + \frac{\|x-y\|^2}{2}\right\}$.

Lemma 3.2.6. *[81] Let $g : \mathcal{C} \longrightarrow \mathbb{R}$ be a convex, lower semicontinuous and subdifferentiable function on $\mathcal{C}$. Suppose $int(\mathcal{C}) \neq 0$, then $\bar{r}$ is a solution of the convex problem, $min\{g(x) : x \in \mathcal{C}\}$ if and only if $\partial g(\bar{r}) + N_{\mathcal{C}}(\bar{r}) \ni 0$, where $\partial g(\bar{r})$ and $N_{\mathcal{C}}(\bar{r})$ respectively denote the subdifferential of g and normal cone of $\mathcal{C}$ at $\bar{r}$.*

Lemma 3.2.7. *The following is true for each $x, y \in \mathcal{H}$ and $\beta \in \mathbb{R}$;*

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

Lemma 3.2.8. *(Opial 1967)[73] Let $\mathcal{C} \subset \mathcal{H}$ be nonempty and, $\{x_n\}$ a sequence in $\mathcal{H}$ such that:*

(i) $\lim_{n\to\infty}\|x_n - \bar{r}\|$ exists for each $\bar{r} \in \mathcal{C}$;

(ii) every sequentially weak cluster point of $\{x_n\}$ is found in $\mathcal{C}$.

Then the sequence $\{x_n\}$ weakly converges to a point in $\mathcal{C}$.

3.3 Main Results

We will assume the following conditions are true for the bifunction f :

Assumption 3.3.1.

(A1) *For each* $s \in \mathcal{C}$, $f(s,s) = 0$ *and* f *is pseudomonotone on* $\mathcal{C}$;

(A2) *with two positive constants* δ_1 *and* δ_2, f *satisfies the Lipschitz-type conditions on* $\mathcal{H}$;

(A3) $f(s,.)$ *is convex, lower semicontinuous and subdifferentiable on* $\mathcal{C}$, $\forall s \in \mathcal{C}$;

(A4) f *is weakly and jointly continuous on* $\mathcal{C} \times \mathcal{C}$. *That is, whenever* $s, z \in \mathcal{C}$ *and* $s_n \rightharpoonup s$, $z_n \rightharpoonup z$ *as* $n \to \infty$, *then we have* $\limsup_{n\to\infty} f(s_n, z_n) \to f(s,z)$. *Particularly, in other words,* $\limsup_{n\to\infty} f(s_n, z_n) \leq f(s,z)$ *for each fixed* $s, z \in \mathcal{C}$ *and* $\{s_n\}, \{z_n\} \in \mathcal{C}$;

(A5) $EP(f, \mathcal{C}) \neq \emptyset$.

Now, we present our proposed method as follows:

Assumption 3.3.2. *Condition on the inertial and relaxation factors.*

Suppose $\theta_n \in [0,1)$, $\alpha_n \in (0,1)$ *and* $\liminf_{n\to\infty} \alpha_n > 0$. *Assume for* n *sufficiently large enough, there is* $\epsilon \in (0,1)$ *for which*

$$\begin{aligned} & (1-\epsilon)\frac{1-\alpha_{n-1}}{2\alpha_{n-1}}(1-\theta_{n-1}) \\ \geq\ & \theta_n t_{n+1}\left(1+\theta_n+\left[\frac{1-\alpha_n}{2\alpha_n}(1-\theta_n)+\frac{1-\alpha_{n-1}}{2\alpha_{n-1}}(1-\theta_{n-1})\right]_+\right). \end{aligned} \tag{3.3.1}$$

Algorithm 3.3.3.
Let $x_0,\ x_1 \in \mathcal{H},\ \lambda_1 > 0,\ \mu \in]0,1[$.
STEP 1*: Start with the present iterates* x_{n-1} *and* x_n,

$$\textit{compute} \quad \begin{cases} w_n = x_n + \theta_n(x_n - x_{n-1}) \\ \\ y_n = \underset{y\in\mathcal{C}}{argmin}\left\{\lambda_n f(w_n, y) + \frac{1}{2}\|w_n - y\|^2\right\}. \end{cases}$$

Stop the computation if $y_n = w_n$, *and* y_n *becomes a solution, else, move to the next step which is step 2.*
STEP 2*:*
Choose $v_n \in \partial f(w_n, .)(y_n)$ *for which there is*
$q_n \in N_{\mathcal{C}}(y_n)$ *satisfying*
$q_n = w_n - \lambda_n v_n - y_n$, *and construct a half space*
$T_n = \{z \in \mathcal{H} : \langle w_n - \lambda_n v_n - y_n, z - y_n\rangle \leq 0\}$.

Evaluate
$z_n = \underset{y\in T_n}{argmin}\left\{\lambda_n f(y_n, y) + \frac{1}{2}\|w_n - y\|^2\right\}$.

STEP 3*: Compute*

$x_{n+1} = (1-\alpha_n)w_n + \alpha_n z_n$, *and*

$$\lambda_{n+1} = \begin{cases} \min\left\{\dfrac{\mu}{2}\dfrac{\|w_n - y_n\|^2 + \|z_n - y_n\|^2}{f(w_n,z_n) - f(w_n,y_n) - f(y_n,z_n)},\ \ \lambda_n\right\}, & \text{if } f(w_n,z_n) - f(w_n,y_n) \\ & \quad -f(y_n,z_n) > 0, \\ \lambda_n, & \text{otherwise.} \end{cases}$$

Set $n+1 \longleftarrow n$ *and continue again from* ***step 1***.

The following proposition provides validation for Assumptions (3.2.2) and 3.3.2.

Proposition 3.3.4. ***(see [3, Proposition 3.3], [44, Propostion 3.8])*** *Let* $\alpha_n \in (0,1)$ *and* $\theta_n \in [0,1)$ $\forall n \geq 1$. *Suppose there is* $c \in [0,1)$ *and* $c^* \in [c,1)$ *for which*

$$\lim_{n\to\infty}\left(\frac{1}{1-\theta_{n+1}} - \frac{1}{1-\theta_n}\right) = c,$$
$$\lim_{n\to\infty}\frac{\psi_{n+1}-\psi_n}{\psi_n(1-\theta_n)} = c^*$$

and

$$\liminf_{n\to\infty}\psi_n(1-\theta_n)^2 > \limsup_{n\to\infty}\frac{\theta_n(1+\theta_n)}{1-c^*},$$

where $\psi_n = \dfrac{1-\alpha_n}{2\alpha_n}$. *Then Assumptions* (3.2.2) *and 3.3.2 are satisfied.*

Lemma 3.3.5. *Suppose* $\{x_n\}$ *is a sequence in* $\mathcal{H}$, *and* $\{\theta_n\}$ *a real positive sequence. Given* $\bar{r} \in \mathcal{H}$, *define the sequence* $\{\Phi_n\}$ *by* $\Phi_n := \frac{1}{2}\|x_n - \bar{r}\|^2$ *and* $w_n = x_n + \theta_n(x_n - x_{n-1})$. *Then,*

$$\begin{aligned}\Phi_{n+1} - \Phi_n - \theta_n(\Phi_n - \Phi_{n-1}) &= \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 \\ &\quad + \frac{1}{2}\|x_{n+1} - \bar{r}\|^2 - \frac{1}{2}\|w_n - \bar{r}\|^2. \end{aligned} \tag{3.3.2}$$

Proof.
Using the definition of $w_n = x_n + \theta_n(x_n - x_{n-1})$, we have,

$$\begin{aligned}\|w_n - \bar{r}\|^2 &= \|x_n + \theta_n(x_n - x_{n-1}) - \bar{r}\|^2 \\ &= \|x_n - \bar{r} + \theta_n(x_n - x_{n-1})\|^2 \\ &= \|x_n - \bar{r}\|^2 + \theta_n^2\|x_n - x_{n-1}\|^2 + 2\theta_n\langle x_n - \bar{r}, x_n - x_{n-1}\rangle \\ &= \|x_n - \bar{r}\|^2 + (\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 + \theta_n(\|x_n - \bar{r}\|^2 - \|x_{n-1} - \bar{r}\|^2) \\ &= 2\Phi_n + (\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 + 2(\Phi_n + \Phi_{n-1}). \end{aligned} \tag{3.3.3}$$

Rearranging and adding $\Phi_{n+1} = \frac{1}{2}\|x_{n+1} - \bar{r}\|^2$ to both sides we have,

$$\begin{aligned}\Phi_{n+1} - \Phi_n - \theta_n(\Phi_n - \Phi_{n-1}) &= \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 + \frac{1}{2}\|x_{n+1} - \bar{r}\|^2 \\ &\quad - \frac{1}{2}\|w_n - \bar{r}\|^2. \end{aligned} \tag{3.3.4}$$

□

Lemma 3.3.6. *Let $\{x_n\}$ be a sequence formed by Algorithm 3.3.3 ssuch that Assumption 3.3.1 holds. Then, for $\bar{r} \in EP(f, \mathcal{C})$, we have*

$$\begin{aligned}\|x_{n+1} - \bar{r}\|^2 \quad &\leq \quad \|w_n - \bar{r}\|^2 - \alpha_n \left(1 - \mu \frac{\lambda_n}{\lambda_{n+1}}\right) (\|w_n - y_n\|^2 + \|z_n - y_n\|^2) \\ &\quad - \frac{1 - \alpha_n}{\alpha_n} \|x_{n+1} - w_n\|^2 \qquad \forall n \geq 1.\end{aligned}$$

Proof.
Using the definition of y_n and Lemma 3.2.6, we have that

$$0 \in \partial \left(\lambda_n f(y_n, y) + \frac{1}{2}\|y - y_n\|^2\right)(y_n) + \mathcal{N}_{\mathcal{C}}(y_n).$$

From Step 2, we obtain

$$\begin{aligned}\langle w_n - y_n, y - y_n\rangle \quad &= \quad \langle \lambda_n v_n + q_n, y - y_n\rangle \\ &= \quad \lambda_n \langle v_n, y - y_n\rangle + \langle q_n, y - y_n\rangle. \end{aligned} \tag{3.3.5}$$

But $q_n \in \mathcal{N}_{\mathcal{C}}(y_n)$ implies that $\langle q_n, y - y_n\rangle \leq 0 \quad \forall y \in \mathcal{C}$.
Thus we obtain from (3.3.5) that

$$\langle w_n - y_n, y - y_n\rangle \leq \lambda_n \langle v_n, y - y_n\rangle \quad \forall y \in T_n. \tag{3.3.6}$$

Also, $v_n \in \partial f(w_n, y_n)$ implies that

$$f(w_n, y) - f(w_n, y_n) \geq \langle v_n, y - y_n\rangle \quad \forall y \in \mathcal{C}. \tag{3.3.7}$$

Using (3.3.6) and (3.3.7), we have

$$\lambda_n (f(w_n, y) - f(w_n, y_n)) \geq \langle w_n - y_n, y - y_n\rangle \quad \forall y \in \mathcal{C}. \tag{3.3.8}$$

Putting $z_n = y$ in (3.3.8), we see that

$$\lambda_n (f(w_n, z_n) - f(w_n, y_n)) \geq \langle w_n - y_n, z_n - y_n\rangle. \tag{3.3.9}$$

Again, using the definition of z_n and Lemma 3.2.5 give us

$$\lambda_n (f(y_n, y) - f(y_n, z_n)) \geq \langle w_n - z_n, y - z_n\rangle \quad \forall y \in \mathcal{C}. \tag{3.3.10}$$

Let $y = \bar{r} \in EP(f, \mathcal{C})$.
Then,

$$\lambda_n (f(y_n, \bar{r}) - f(y_n, z_n)) \geq \langle w_n - z_n, \bar{r} - z_n\rangle \quad \forall \bar{r} \in EP(f, \mathcal{C}). \tag{3.3.11}$$

But $\bar{r} \in EP(f, \mathcal{C})$, implies $f(\bar{r}, y_n) \geq 0$. This further implies by pseudomonotonicity of f that

$$f(y_n, \bar{r}) \leq 0.$$

Therefore, (3.3.11) becomes,

$$-\lambda_n f(y_n, z_n) \geq \langle w_n - z_n, \bar{r} - z_n \rangle. \tag{3.3.12}$$

Combining (3.3.9) and (3.3.12) implies that

$$\begin{aligned} 2\lambda_n(f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n)) \quad &\geq \quad 2\langle w_n - y_n, z_n - y_n \rangle \\ &\quad +2\langle w_n - z_n, \bar{r} - z_n \rangle. \end{aligned} \tag{3.3.13}$$

Note that:

$$\langle w_n - y_n, z_n - y_n \rangle = \frac{1}{2}(\|w_n - y_n\|^2 + \|z_n - y_n\|^2 - \|w_n - z_n\|^2). \tag{3.3.14}$$

Also,

$$\langle w_n - z_n, \bar{r} - z_n \rangle = \frac{1}{2}(\|w_n - z_n\|^2 + \|z_n - \bar{r}\|^2 - \|w_n - \bar{r}\|^2). \tag{3.3.15}$$

Using (3.3.14) and (3.3.15) in (3.3.13), we get

$$\begin{aligned} 2\lambda_n(f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n)) \quad &\geq \quad \|w_n - y_n\|^2 + \|z_n - y_n\|^2 \\ &\quad +\|z_n - \bar{r}\|^2 - \|w_n - \bar{r}\|^2. \end{aligned} \tag{3.3.16}$$

By the definition of λ_n, we have

$$2(f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n)) \leq \frac{\mu(\|w_n - y_n\|^2 + \|z_n - y_n\|^2)}{\lambda_{n+1}}. \tag{3.3.17}$$

Using (3.3.17) in (3.3.16), we obtain

$$\begin{aligned} \mu\frac{\lambda_n}{\lambda_{n+1}}(\|w_n - y_n\|^2 + \|z_n - y_n\|^2) \quad &\geq \quad \|w_n - y_n\|^2 + \|z_n - y_n\|^2 \\ &\quad +\|z_n - \bar{r}\|^2 - \|w_n - \bar{r}\|^2. \end{aligned} \tag{3.3.18}$$

This further implies

$$\begin{aligned} \|z_n - \bar{r}\|^2 \quad &\leq \quad \|w_n - \bar{r}\|^2 - \left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)\|w_n - y_n\|^2 \\ &\quad - \left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)\|z_n - y_n\|^2. \end{aligned} \tag{3.3.19}$$

Moreover, with $\alpha_n \in (0, 1)$, we have

$$\|x_{n+1} - \bar{r}\|^2 = (1 - \alpha_n)\|w_n - \bar{r}\|^2 + \alpha_n\|z_n - \bar{r}\|^2 - (1 - \alpha_n)\alpha_n\|z_n - w_n\|^2. \tag{3.3.20}$$

Using (3.3.19) and the definition of $\{x_{n+1}\}$, that is $z_n - w_n = \frac{1}{\alpha_n}(x_{n+1} - w_n)$ in (3.3.20), we establish the desired result. □

Lemma 3.3.7. *Let $\{\Phi_n\}$ be as defined in Lemma 3.3.5, and $\{x_n\}$ a sequence formed by Algorithm 3.3.3. Therefore, $\forall \bar{r} \in EP(f, \mathcal{C})$, we have*

$$\begin{aligned}\Phi_{n+1} - \Phi_n - \theta_n(\Phi_n - \Phi_{n-1}) \;\leq\; & \frac{1}{2}((\theta_n + \theta_n^2) - \psi_n(1-\theta_n)^2)\|x_n - x_{n-1}\|^2 \\ & -\frac{\alpha_n}{2}\left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)(\|w_n - y_n\|^2 + \|z_n - y_n\|^2) \\ & -\frac{1}{2}\psi_n(1-\theta_n)(\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2),\end{aligned}$$

where $\psi_n = \dfrac{1-\alpha_n}{2\alpha_n}$.

Proof.
Using the definition of $w_n = x_n + \theta_n(x_n - x_{n-1})$, we have

$$\begin{aligned}\|x_{n+1} - w_n\|^2 &= \|x_{n+1} - x_n - \theta_n(x_n - x_{n-1})\|^2 \\ &= \|x_{n+1} - x_n - (x_n - x_{n-1}) + (1-\theta_n)(x_n - x_{n-1})\|^2 \\ &= \|x_{n+1} - 2x_n + x_{n-1}\|^2 + (1-\theta_n)^2\|x_n - x_{n-1}\|^2 \\ &\quad +2(1-\theta_n)\langle x_{n+1} - 2x_n + x_{n-1}, x_n - x_{n-1}\rangle. \qquad (3.3.21)\end{aligned}$$

But

$$\begin{aligned}2\langle x_{n+1} - 2x_n + x_{n-1}, x_n - x_{n-1}\rangle &= \|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2 \\ &\quad -\|x_{n+1} - 2x_n + x_{n-1}\|^2. \qquad (3.3.22)\end{aligned}$$

Putting (3.3.22) into (3.3.21) therefore gives

$$\begin{aligned}\|x_{n+1} - w_n\|^2 &= \theta_n\|x_{n+1} - 2x_n + x_{n-1}\|^2 + (1-\theta_n)^2\|x_n - x_{n-1}\|^2 \\ &\quad +(1-\theta_n)(\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2) \\ &\geq (1-\theta_n)^2\|x_n - x_{n-1}\|^2 \\ &\quad +(1-\theta_n)(\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2). \qquad (3.3.23)\end{aligned}$$

Notice that $\dfrac{1-\alpha_n}{2\alpha_n} \geq \dfrac{1-\alpha_n}{4\alpha_n}, \quad \forall n \geq 1.$
Thus, we have from (3.3.23) that

$$\begin{aligned}-\frac{1-\alpha_n}{2\alpha_n}\|x_{n+1} - w_n\|^2 \;\leq\; & -\frac{1-\alpha_n}{4\alpha_n}(1-\theta_n)^2\|x_n - x_{n-1}\|^2 \\ & -\frac{1-\alpha_n}{4\alpha_n}(1-\theta_n)(\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2).\end{aligned}$$

That is,

$$\begin{aligned}-\psi_n\|x_{n+1} - w_n\|^2 \;\leq\; & -\frac{1}{2}\psi_n(1-\theta_n)^2\|x_n - x_{n-1}\|^2 \\ & -\frac{1}{2}\psi_n(1-\theta_n)(\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2). \qquad (3.3.24)\end{aligned}$$

Hence, the desired result follows immediately from Lemmas 3.3.5, 3.3.6, and inequality 3.3.24. □

Lemma 3.3.8. *Suppose that assumption (3.2.2) holds and the sequence $\{t_i\}$ is as defined by (3.2.3). Let $\psi_n = \dfrac{1-\alpha_n}{2\alpha_n}$ for $\{\alpha_n\} \subset (0,1)$. Given a sequence $\{x_n\}$ generated by Algorithm 3.3.3, then for each $n \geq 1$, we have*

$$\begin{aligned} & \sum_{i=1}^{n-1} \left(\psi_{i-1}(1-\theta_{i-1}) - \theta_i t_{i+1}(1+\theta_i + [\psi_i(1-\theta_i) - \psi_{i-1}(1-\theta_{i-1})]_+)\right) \|x_i - x_{i-1}\|^2 \\ \leq \;& 2\Phi_0 + 2t_1|\Phi_1 - \Phi_0| + t_1\psi_0(1-\theta_0)\|x_1 - x_0\|^2. \end{aligned}$$

Proof.
From Lemma 3.3.7, we have

$$\begin{aligned} \Phi_{n+1} - \Phi_n &\leq \theta_n(\Phi_n - \Phi_{n-1}) + \frac{1}{2}\left((\theta_n + \theta_n^2) - \psi_n(1-\theta_n)^2\right)\|x_n - x_{n-1}\|^2 \\ &- \frac{1}{2}\psi_n(1-\theta_n)(\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2). \end{aligned} \tag{3.3.25}$$

Let $c_n := \Phi_n - \Phi_{n-1}$, $d_n := \frac{1}{2}((\theta_n + \theta_n^2) - \psi_n(1-\theta_n)^2)\|x_n - x_{n-1}\|^2 - \frac{1}{2}\psi_n(1-\theta_n)(\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2)$.

Thus, (3.3.25) now shows the form $c_{n+1} \leq \theta_n c_n + d_n$, and applying Lemma 3.2.4(i), we obtain

$$\begin{aligned} \Phi_n - \Phi_0 &= \sum_{i=1}^{n} c_i \\ &\leq t_1, n(\Phi_1 - \Phi_0) + \sum_{i=1}^{n-1} t_{i+1,n}\Big(\frac{1}{2}((\theta_i + \theta_i^2) - \psi_i(1-\theta_i)^2)\|x_i - x_{i-1}\|^2 \\ &- \frac{1}{2}\psi_i(1-\theta_i)(\|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2)\Big) \\ &= t_1, n(\Phi_1 - \Phi_0) - \sum_{i=1}^{n-1} t_{i+1,n}\left(\frac{1}{2}(\psi_i(1-\theta_i)^2 - (\theta_i + \theta_i^2))\|x_i - x_{i-1}\|^2\right) \\ &- \sum_{i=1}^{n-1} t_{i+1,n}\left(\frac{1}{2}\psi_i(1-\theta_i)(\|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2)\right). \end{aligned}$$

This implies that

$$\begin{aligned} & \frac{1}{2}\Bigg[\sum_{i=1}^{n-1} t_{i+1,n}(\psi_i(1-\theta_i)^2 - (\theta_i + \theta_i^2))\|x_i - x_{i-1}\|^2 \\ & + \sum_{i=1}^{n-1} t_{i+1,n}\psi_i(1-\theta_i)(\|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2)\Bigg] \\ & \leq t_1, n(\Phi_1 - \Phi_0) + \Phi_0 - \Phi_n. \end{aligned} \tag{3.3.26}$$

Note that $\Phi_n \geq 0, \quad t_{1,n} \leq t_1 \quad \forall n \in \mathbb{N}$. Thus

$$\begin{aligned}
&\sum_{i=1}^{n-1} t_{i+1,n} \left(\psi_i(1-\theta_i)^2 - (\theta_i + \theta_i^2)\right) \|x_i - x_{i-1}\|^2 \\
&+ \sum_{i=1}^{n-1} t_{i+1,n}\psi_i(1-\theta_i) \left(\|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2\right) \\
&\leq 2t_1, n(\Phi_1 - \Phi_0) + 2\Phi_0 \\
&\leq 2t_1|\Phi_1 - \Phi_0| + 2\Phi_0. \qquad (3.3.27)
\end{aligned}$$

Again, notice that

$$\begin{aligned}
&\sum_{i=1}^{n-1} t_{i+1,n}\psi_i(1-\theta_i) \left(\|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2\right) \\
&= \sum_{i=1}^{n-1} t_{i+1,n}\psi_i(1-\theta_i)\|x_{i+1} - x_i\|^2 - \sum_{i=1}^{n-1} t_{i+1,n}\psi_i(1-\theta_i)\|x_i - x_{i-1}\|^2 \\
&\quad + \sum_{i=1}^{n-1} t_{i,n}\psi_{i-1}(1-\theta_{i-1})\|x_i - x_{i-1}\|^2 - \sum_{i=1}^{n-1} t_{i,n}\psi_{i-1}(1-\theta_{i-1})\|x_i - x_{i-1}\|^2 \\
&= \sum_{i=1}^{n-1} \left(t_{i,n}\psi_{i-1}(1-\theta_{i-1}) - t_{i+1,n}\psi_i(1-\theta_i)\right) \|x_i - x_{i-1}\|^2 + t_{n,n}\psi_n(1-\theta_{n-1})\|x_n - x_{n-1}\|^2 \\
&\quad - t_{1,n}\psi_0(1-\theta_0)\|x_1 - x_0\|^2.
\end{aligned}$$

Since $t_{n,n}\psi_n(1-\theta_{n-1})\|x_n - x_{n-1}\|^2 \geq 0$ and $t_{1,n} \leq t_1 \quad \forall n > 0$, it follows that

$$\begin{aligned}
&\sum_{i=1}^{n-1} t_{i+1,n}\psi_i(1-\theta_i) \left(\|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2\right) \\
&\geq \sum_{i=1}^{n-1} \left(t_{i,n}\psi_{i-1}(1-\theta_{i-1}) - t_{i+1,n}\psi_i(1-\theta_i)\right) \|x_i - x_{i-1}\|^2 \\
&\quad - t_1\psi_0(1-\theta_0)\|x_1 - x_0\|^2. \qquad (3.3.28)
\end{aligned}$$

Putting (3.3.27) and (3.3.28) together gives us

$$\begin{aligned}
&\sum_{i=1}^{n-1} \left[t_{i+1,n}(\psi_i(1-\theta_i)^2 - (\theta_i + \theta_i^2)) + t_{i,n}\psi_{i-1}(1-\theta_{i-1}) - t_{i+1,n}\psi_i(1-\theta_i)\right] \|x_i - x_{i-1}\|^2 \\
&\leq 2t_1|\Phi_1 - \Phi_0| + 2\Phi_0 + t_1\psi_0(1-\theta_0)\|x_1 - x_0\|^2. \qquad (3.3.29)
\end{aligned}$$

On the other hand, from the definition of $t_{i,n} = 1 + \theta_i t_{i+1,n}$ for each $i \geq 1$ and $n \geq i+1$ in (3.2.5), we have

$$\begin{aligned}
&t_{i+1,n}\left(\psi_i(1-\theta_i)^2 - (\theta_i + \theta_i^2)\right) + t_{i,n}\psi_{i-1}(1-\theta_{i-1}) - t_{i+1,n}\psi_i(1-\theta_i) \\
&= t_{i+1,n}(\psi_i(1-\theta_i)^2 - (\theta_i + \theta_i^2)) + \psi_{i-1}(1 + t_{i+1,n})(1-\theta_{i-1}) - t_{i+1,n}\psi_i(1-\theta_i) \\
&= \psi_{i-1}(1-\theta_{i-1}) - \theta_i t_{i+1,n}(1 + \theta_i + \psi_i(1-\theta_i) - \psi_{i-1}(1-\theta_{i-1})). \qquad (3.3.30)
\end{aligned}$$

But $t_{i,n} \leq t_i \quad \forall i \geq 1, \quad n \geq i+1$ and $[t]_+ \geq t \quad \forall t \in \mathcal{H}$. It follows that

$$\begin{aligned} & t_{i+1,n}(\psi_i(1-\theta_i)^2 - (\theta_i + \theta_i^2)) + t_{i,n}\psi_{i-1}(1-\theta_{i-1}) - t_{i+1,n}\psi_i(1-\theta_i) \\ & \geq \psi_{i-1}(1-\theta_{i-1}) - \theta_i t_{i+1}(1+\theta_i + [\psi_i(1-\theta_i) - \psi_{i-1}(1-\theta_{i-1})]_+). \end{aligned} \tag{3.3.31}$$

The combination of (3.3.31) and (3.3.29), therefore, establishes the desired result. □

Lemma 3.3.9.
Assume that Assumptions (3.2.2), 3.3.1 and 3.3.2 hold. Let $\{t_i\}$ be the sequence defined by (3.2.3) and $\{x_n\}$ the sequence formed by Algorithm 3.3.3. Then,

a. *(i.)* $\lim_{n\to\infty} \sum_{i=1}^{n-1} \frac{1-\alpha_i}{2\alpha_i}(1-\theta_{i-1})\|x_i - x_{i-1}\|^2 < \infty$,
and as such $\lim_{n\to\infty} \sum_{i=1}^{n-1} \theta_i t_{i+1}\|x_i - x_{i-1}\|^2 < \infty$;

(ii.) For each $\bar{r} \in EP(f, \mathcal{C})$, $\lim_{n\to\infty} \|x_n - \bar{r}\|$ *exists, and hence* $\{x_n\}$ *is bounded;*

b. *(i.)* $\lim_{n\to\infty} \|w_n - y_n\| = 0$, $\lim_{n\to\infty} \|z_n - y_n\| = 0$;

(ii.) $\lim_{n\to\infty} \|x_n - w_n\| = 0$;

(iii.) $\lim_{n\to\infty} \|w_n - z_n\| = 0$.

Proof.

a. (i.) Assumption 3.3.2 is true for n sufficiently large enough, thus, we may assume without any loss of generality that it is also true for $n \geq 1$.
Now, using inequality (3.3.1) and Lemma 3.3.8, noting that $\psi_n = \dfrac{1-\alpha_n}{2\alpha_n}$, we have that:

$$(\epsilon - 1)\psi_{i-1}(1-\theta_{i-1}) \leq -\theta_i t_{i+1}(1+\theta_i + [\psi_i(1-\theta_i) + \psi_{i-1}(1-\theta_{i-1})]_+).$$

This implies

$$\begin{aligned} & \sum_{i=1}^{n-1} \epsilon\psi_{i-1}(1-\theta_{i-1})\|x_i - x_{i-1}\|^2 \\ & \leq \sum_{i=1}^{n-1} (\psi_{i-1}(1-\theta_{i-1}) - \theta_i t_{i+1}(1+\theta_i + [\psi_i(1-\theta_i) + \psi_{i-1}(1-\theta_{i-1})]_+)\|x_i - x_{i-1}\|^2 \\ & \leq 2\Phi_0 + 2t_1|\Phi_1 - \Phi_0| + t_1\psi_0(1-\theta_0)\|x_1 - x_0\|^2. \end{aligned}$$

This implies that

$$\begin{aligned} \sum_{i=1}^{n-1} \psi_{i-1}(1-\theta_{i-1})\|x_i - x_{i-1}\|^2 \quad \leq \quad & \frac{1}{\epsilon}\left(2\Phi_0 + 2t_1|\Phi_1 - \Phi_0|\right. \\ & \left. + t_1\psi_0(1-\theta_0)\|x_1 - x_0\|^2\right). \end{aligned} \tag{3.3.32}$$

Taking the limit of (3.3.32) as $n \to \infty$, we arrive at

$$\sum_{i=1}^{\infty} \psi_{i-1}(1-\theta_{i-1})\|x_i - x_{i-1}\|^2 < \infty. \tag{3.3.33}$$

Since $\liminf_{n\to\infty} \alpha_n > 0$, (3.3.33) further implies that
$\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0$.

More so, using inequality (3.3.1) again, we have
$\theta_i t_{i+1} \leq \psi_{i-1}(1-\theta_{i-1}) - \theta_i t_{i+1}(\theta_i + [\psi_i(1-\theta_i) + \psi_{i-1}(1-\theta_{i-1})]_+)$.

This implies that

$$\begin{aligned}\sum_{i=1}^{n-1} \theta_i t_{i+1}\|x_i - x_{i-1}\|^2 &\leq \sum_{i=1}^{n-1}(\psi_{i-1}(1-\theta_{i-1}) - \theta_i t_{i+1}(\theta_i + [\psi_i(1-\theta_i) \\ &\quad +\psi_{i-1}(1-\theta_{i-1})]_+))\|x_i - x_{i-1}\|^2 \\ &\leq \sum_{i=1}^{n-1} \psi_{i-1}(1-\theta_{i-1})\|x_i - x_{i-1}\|^2. \end{aligned} \tag{3.3.34}$$

Taking the limit of (3.3.34) and using (3.3.33), we conclude that:

$\sum_{i=1}^{n-1} \theta_i t_{i+1}\|x_i - x_{i-1}\|^2 < \infty$. Thus, completing item a. (i.)

(ii.) Using Lemmas 3.3.5 and 3.3.6 gives us that

$$\Phi_{n+1} - \Phi_n - \theta_n(\Phi_n - \Phi_{n-1}) \leq \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 - \frac{1-\alpha_n}{2\alpha_n}\|x_{n+1} - w_n\|^2$$

Note that $\theta_n^2 \leq \theta_n, \quad \alpha_n > 0, \quad \forall n \geq 1$.
Thus,

$$\Phi_{n+1} - \Phi_n \leq \theta_n(\Phi_n - \Phi_{n-1}) + \theta_n\|x_n - x_{n-1}\|^2. \tag{3.3.35}$$

Let $c_n := \Phi_n - \Phi_{n-1}, \quad d_n := \theta_n\|x_n - x_{n-1}\|^2$, then, making use of Lemma 3.2.4(ii) and item a. (i.) of Lemma 3.3.9, we obtain

$$\sum_{n=1}^{\infty}[\Phi_n - \Phi_{n-1}]_+ \leq t_1[\Phi_1 - \Phi_0]_+ + \sum_{n=1}^{\infty} t_{n+1}[\theta_n\|x_n - x_{n-1}\|^2]_+ < \infty.$$

That is, $\sum_{n=1}^{\infty}[\Phi_n - \Phi_{n-1}]_+ < \infty$. In addition, this implies that for each $\bar{r} \in \mathcal{C}$, $\lim_{n\to\infty} \Phi_n = \lim_{n\to\infty} \|x_n - \bar{r}\|$ exists. Thus, as claimed, $\{x_n\}$ is indeed bounded.

b. (i.) Using Lemma 3.3.7,
define $\delta_n := \frac{\alpha_n}{2}\left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)(\|w_n - y_n\|^2 + \|z_n - y_n\|^2) - \frac{1}{2}\psi_n(1-\theta_n)(\|x_{n+1} - x_n\|^2 -$

$\|x_n - x_{n-1}\|^2)$,
then we have

$$\begin{aligned}\Phi_{n+1} - \Phi_n - \theta_n(\Phi_n - \Phi_{n-1}) &\leq \frac{1}{2}((\theta_n + \theta_n^2) - \psi_n(1-\theta_n)^2)\|x_n - x_{n-1}\|^2 - \delta_n \\ &\leq \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 - \delta_n.\end{aligned}$$

Therefore,

$$\Phi_{n+1} - \Phi_n \leq \theta_n(\Phi_n - \Phi_{n-1}) + \theta_n\|x_n - x_{n-1}\|^2 - \delta_n. \tag{3.3.36}$$

Again, applying Lemma 3.2.4(i) to (3.3.36), then

$$\begin{aligned}\Phi_n - \Phi_0 &= \sum_{i=1}^{n}(\Phi_i - \Phi_{i-1}) \\ &\leq t_{1,n}(\Phi_1 - \Phi_0) + \sum_{i=1}^{n-1} t_{i+1,n}(\theta_i\|x_i - x_{i-1}\|^2 - \delta_i).\end{aligned}$$

This implies that

$$\sum_{i=1}^{n-1} t_{i+1,n}\delta_i \leq \Phi_0 + t_1(\Phi_1 - \Phi_0) + \sum_{i=1}^{n-1} t_{i+1}\theta_i\|x_i - x_{i-1}\|^2. \tag{3.3.37}$$

Taking the limit as $n \to \infty$ and by Lemma 3.3.9 item a. (i.), this implies that

$$\sum_{i=1}^{\infty} t_{i+1,n}\delta_i \leq \Phi_0 + t_1(\Phi_1 - \Phi_0) + \sum_{i=1}^{\infty} t_{i+1}\theta_i\|x_i - x_{i-1}\|^2 < \infty. \tag{3.3.38}$$

But $\{t_{i,n}\}$ tends increasingly to t_i, as $n \to \infty$, and by monotone convergence theorem $\sum_{i=1}^{\infty} t_{i+1}\delta_i < \infty$.

This implies that $\lim_{n\to\infty} t_{n+1}\delta_n = 0$.
But $t_n \geq 1\ \forall n \geq 1$,
$\delta_n := \frac{\alpha_n}{2}\left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)(\|w_n - y_n\|^2 + \|z_n - y_n\|^2) - \frac{1}{2}\psi_n(1-\theta_n)(\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2)$.

Thus, $\lim_{n\to\infty} \frac{\alpha_n}{2}\left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)(\|w_n - y_n\|^2 + \|z_n - y_n\|^2) - \frac{1}{2}\psi_n(1-\theta_n)(\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2) = 0$.

Using the facts that $\liminf_{n\to\infty}\alpha_n > 0$, $\lim_{n\to\infty}\left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right) = 1 - \mu > 0$, and $\lim_{n\to\infty}\|x_n - x_{n-1}\| = 0$, we conclude that
$\lim_{n\to\infty}\|w_n - y_n\| = 0$ and $\lim_{n\to\infty}\|z_n - y_n\| = 0$.

(ii.) With the definition of $w_n = x_n + \theta_n(x_n - x_{n-1})$ and Lemma 3.3.9 item a. (i.), therefore:

$$\begin{aligned} \|x_n - w_n\|^2 &= \|\theta_n(x_n - x_{n-1})\|^2 \\ &= \theta_n^2\|x_n - x_{n-1}\|^2 \\ &\leq \theta_n\|x_n - x_{n-1}\|^2. \end{aligned} \tag{3.3.39}$$

This implies that

$$\sum_{n=1}^{\infty} t_{n+1}\|x_n - w_n\|^2 \leq \sum_{n=1}^{\infty} t_{n+1}\theta_n\|x_n - x_{n-1}\|^2 < \infty.$$

Hence, $\sum_{n=1}^{\infty} t_{n+1}\|x_n - w_n\|^2 < \infty$,

giving us that
$\lim_{n\to\infty} t_{n+1}\|x_n - w_n\|^2 = 0.$

Since $t_n \geq 1 \ \ \forall n \geq 1$, then $\lim_{n\to\infty} \|x_n - w_n\| = 0$, establishing (ii.).

(iii.) Recall that $x_{n+1} = (1 - \alpha_n)w_n + \alpha_n z_n$.
This implies

$$\begin{aligned} \alpha_n\|z_n - w_n\| &= \|x_{n+1} - w_n\| \\ &\leq \|x_n - w_n\| + \|x_{n+1} - x_n\|. \end{aligned} \tag{3.3.40}$$

That is, $0 \leq \alpha_n\|z_n - w_n\| \leq \|x_n - w_n\| + \|x_{n+1} - x_n\|$.

Taking limit as $n \to \infty$ and, using items a. (i.), b. (ii.) of Lemma 3.3.9 and the fact that $\liminf_{n\to\infty} \alpha_n > 0$, we conclude that $\lim_{n\to\infty} \|z_n - w_n\| = 0$.

□

Theorem 3.3.10. *Assume that Assumptions (3.2.2), 3.3.1, 3.3.2 hold. Then, the sequence $\{x_n\}$ generated by Algorithm 3.3.3 converges weakly to to an element of $EP(f, \mathcal{C})$.*

Proof.
Since $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_j}\}$ of $\{x_n\}$ that converges weakly to $\bar{r} \in \mathcal{C}$.
From items b. (i.) and (ii.) of Lemma 3.3.9, it follows that $\|x_n - y_n\| \longrightarrow 0, \quad n \longrightarrow \infty$.
This further implies that $y_{n_j} \rightharpoonup \bar{r} \in \mathcal{C}$.
Moreso, we have that $\{w_{n_j}\}$ converges weakly to $\bar{r} \in \mathcal{C}$.

Now, from (3.3.8) and since $\liminf_{n\to\infty} \lambda_n > 0$, we have

$$f(w_{n_j}, y) \geq f(w_{n_j}, y_{n_j}) + \frac{1}{\lambda_{n_j}}\langle w_{n_j} - y_{n_j}, y - y_{n_j}\rangle \ \ \forall y \in \mathcal{C}. \tag{3.3.41}$$

Taking the limit of (3.3.41) as $j \to +\infty$, and using conditions (A1), (A4), we obtain that $f(\bar{r}, y) \geq 0 \quad \forall y \in \mathcal{C}$. Hence, $\bar{r} \in EP(f, \mathcal{C})$. Therefore, by Lemma 3.2.8, we conclude that $\{x_n\}$ weakly converges to $\bar{r} \in EP(f, \mathcal{C})$. □

Remark 3.3.11.
Note that the main role of Assumption 3.3.2 is to guarantee the summability condition

$$\sum_{i=1}^{\infty} \theta_i t_{i+1} \|x_i - x_{i-1}\|^2 < \infty \tag{3.3.42}$$

obtained in Lemma 3.3.9 a. (i.). Hence, Lemma 3.3.9 holds true undoubtedly if we assume (3.3.42) directly.

3.4 Applications

3.4.1 Variational Inequality Problem

In relation to the variational inequality problem defined in (2.2.17) where the cost bifunction is $f(s, z) = \langle \mathcal{B}s, z - s \rangle$, $\forall s, z \in \mathcal{C}$ and the operator $\mathcal{B} : \mathcal{C} \longrightarrow \mathcal{H}$ has the Lipschitz constant $L = 2\delta_1 = 2\delta_2$. The sequences $\{y_n\}$ and $\{z_n\}$ of Algorithm 3.3.3 can be respectively translated into

$$y_n = \mathcal{P}_{\mathcal{C}}(w_n - \lambda_n \mathcal{B} w_n)$$
$$z_n = \mathcal{P}_{T_n}(w_n - \lambda_n \mathcal{B} y_n).$$

Assumption 3.4.1. *Assumption on the operator $\mathcal{B}$.*

B1. For some $L > 0$, $\mathcal{B}$ is $L-$Lipschitz continuous, that is $\|\mathcal{B}s - \mathcal{B}z\| \leq L\|s - z\|$ $\forall s, z \in \mathcal{C}$.

B2. $\mathcal{B}$ is pseudomonotone on $\mathcal{C}$, meaning $\langle \mathcal{B}s, z - s \rangle \geq 0 \implies \langle \mathcal{B}z, s - z \rangle \leq 0$ $\forall s, z \in \mathcal{C}$; and the solutions set $\mathcal{VI}(\mathcal{B}, \mathcal{C})$ is nonempty.

B3. $\limsup_{n \to \infty} \langle \mathcal{B}s_n, z - s_n \rangle \leq \langle \mathcal{B}\bar{r}, z - \bar{r} \rangle$ for each $z \in \mathcal{C}$ and $\{s_n\} \subset \mathcal{C}$ satisfying $\{s_n\} \rightharpoonup \bar{r}$.

Theorem 3.4.2. *Let $\mathcal{B} : \mathcal{C} \longrightarrow \mathcal{H}$ satisfy Assumption 3.4.1. Assume that Assumptions (3.2.2), 3.3.2 hold. Let $\{t_i\}$ be the sequence defined by (3.2.3). Iteratively generate the sequences $\{x_n\}$, $\{y_n\}$, $\{w_n\}$, $\{z_n\}$ as follows:*

Let $x_0, x_1 \in \mathcal{H}$, $\lambda_1 > 0$, $\mu \in]0, 1[$.
Start with the current iterates x_{n-1} and x_n, construct a half space

$$T_n = \{z \in \mathcal{H} : \langle w_n - \lambda_n v_n - y_n, z - y_n \rangle \leq 0\}.$$

$$\textit{Compute} \quad \begin{cases} w_n = x_n + \theta_n(x_n - x_{n-1}) \\ y_n = \mathcal{P}_{\mathcal{C}}(w_n - \lambda_n \mathcal{B} w_n) \\ z_n = \mathcal{P}_{T_n}(w_n - \lambda_n \mathcal{B} y_n) \\ x_{n+1} = (1 - \alpha_n) w_n + \alpha_n z_n. \end{cases}$$

Update the stepsize λ_{n+1} *as follows:*

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \dfrac{\mu}{2}\dfrac{\|w_n - y_n\|^2 + \|z_n - y_n\|^2}{\langle \mathcal{B}w_n - \mathcal{B}y_n, z_n - y_n\rangle}\right\} & \text{if } \langle \mathcal{B}w_n - \mathcal{B}y_n, z_n - y_n\rangle > 0, \\ \lambda_n & \text{otherwise.} \end{cases}$$

Therefore, the sequences $\{x_n\}$, $\{w_n\}$ *and* $\{y_n\}$ *weakly converge to* $\bar{r} \in \mathcal{VI}(\mathcal{B}, \mathcal{C})$.

3.4.2 Fixed Points Problems

From (2.2.14) where the cost bifunction $f(s,z) = \langle s - \mathcal{T}s, z - s\rangle$, the strongly convex programs one needs to solve in algorithm 3.3.3 are as follows:

$$\begin{aligned} y_n &= arg\min_{y\in\mathcal{C}}\left\{\frac{1}{2}\|y - w_n\|^2 + \lambda_n f(w_n, y)\right\} \\ &= arg\min_{y\in\mathcal{C}}\left\{\frac{1}{2}\|y - w_n\|^2 + \lambda_n\langle w_n - \mathcal{T}w_n, y - w_n\rangle\right\} \\ &= arg\min_{y\in\mathcal{C}}\left\{\frac{1}{2}\|y - (w_n - \lambda_n(w_n - \mathcal{T}w_n))\|^2\right\} \\ &= \mathcal{P}_{\mathcal{C}}((1-\lambda_n)w_n - \lambda_n\mathcal{T}w_n) \\ &= (1-\lambda_n)w_n + \lambda_n\mathcal{T}w_n. \end{aligned}$$

And,

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

which using similar argument as above gives

$$z_n = \mathcal{P}_{\Omega_n}(w_n - \lambda_n(y_n - \mathcal{T}y_n)).$$

Theorem 3.4.3. *Let* $\mathcal{C}$ *be a closed convex nonempty subset of* $\mathcal{H}$ *and* $\mathcal{T} : \mathcal{C} \longrightarrow \mathcal{C}$ *be weakly continuous* $\alpha-$*strict pseudo-contractive mapping with* $Fix(\mathcal{T}) \neq \emptyset$. *Suppose that Assumptions* (3.2.2) *and* 3.3.2 *hold. Let* $\{t_i\}$ *be the sequence defined by* (3.2.3). *Iteratively generate the sequences* $\{x_n\}$, $\{y_n\}$, $\{w_n\}$, $\{z_n\}$ *as follows:*

Let $x_0, x_1 \in \mathcal{H}$, $\lambda_1 > 0$, $\mu \in]0,1[$.
Start with the current iterates x_{n-1} *and* x_n, *construct a half space*

$$\Omega_n = \{z \in \mathcal{H} : \langle w_n - \lambda_n v_n - y_n, z - y_n\rangle \leq 0\}.$$

$$\text{Evaluate} \quad \begin{cases} w_n = x_n + \theta_n(x_n - x_{n-1}) \\ y_n = (1-\lambda_n)w_n + \lambda_n\mathcal{T}w_n \\ z_n = \mathcal{P}_{\Omega_n}(w_n - \lambda_n(y_n - \mathcal{T}y_n)) \\ x_{n+1} = (1-\alpha_n)w_n + \alpha_n\mathcal{P}_{\Omega_n}(u_n) \\ u_n = w_n - \lambda_n(y_n - \mathcal{T}y_n). \end{cases}$$

Update the stepsize λ_{n+1} *as follows:*

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \dfrac{\mu}{2}\dfrac{\|w_n - y_n\|^2 + \|z_n - y_n\|^2}{\langle(\boldsymbol{I} - \mathcal{T})(w_n - y_n), z_n - y_n\rangle}\right\} & \text{if } \langle(\boldsymbol{I} - \mathcal{T})(w_n - y_n), z_n - y_n\rangle > 0, \\ \lambda_n & \text{otherwise.} \end{cases}$$

Then the generated sequences $\{x_n\},\ \{w_n\},\ \{y_n\}$ *converge weakly to* $\bar{r} = \mathcal{P}_{Fix(\mathcal{T})}(0)$.

Chapter 4

Relaxed Inertial Sub-gradient Extra-gradient Method for Solving Equilibrium Problems.

4.1 Introduction

Again, throughout this chapter, our attentions will be focused on another new method of finding a real approximate solution of an equilibrium problem (3.1.1). This proposed method will seek to incorporate the on-line rule condition into the relaxed inertial technique to obtain a strong convergent result.

4.2 Preliminaries

It is important that we present in this section some definitions and lemmas that are crucial in the analysis of our proposed method. The notations "$\rightarrow$" and "$\rightharpoonup$" will be used to denote strong and weak convergence respectively.

For easy referencing, we present the following:

Definition 4.2.1.

i. Let $\partial f(s,z)$ denote the subdifferential of a convex function $f(s,.)$ at z, then,

$$\begin{aligned} \partial f(s,z) &:= \{u \in \mathcal{H} : f(s,w) \geq f(s,z) + \langle u, w-z\rangle \ \ \forall w \in \mathcal{H}\}. \ \ \textit{Thus,} \\ \partial f(s,s) &:= \{u \in \mathcal{H} : f(s,w) \geq \langle u, w-s\rangle \ \ \forall w \in \mathcal{H}\}. \end{aligned}$$

ii. The normal cone $\mathcal{N}_C$ of $\mathcal{C}$ at s is denoted by:

$$\mathcal{N}_C(s) = \{w \in \mathcal{H} : \langle w, z-s\rangle \leq 0 \ \ \forall z \in \mathcal{C}\}.$$

Lemma 4.2.2. *[6] For each $s \in \mathcal{H}$, $z \in \mathcal{C}$ and positive constant λ, the following inequality is true*

$\lambda(g(z) - g(prox_{\lambda g}(s))) \geq \langle s - prox_{\lambda g}(s), z - prox_{\lambda g}(s)\rangle$,

where $prox_{\lambda g}(s) := \underset{z\in\mathcal{C}}{arg\min}\left\{\lambda g(z) + \frac{\|s-z\|^2}{2}\right\}$.

Lemma 4.2.3. *[81] Let $g : \mathcal{C} \longrightarrow \mathbb{R}$ be a convex, lower semicontinuous and subdifferentiable function on $\mathcal{C}$. Suppose $int(\mathcal{C}) \neq 0$, then $\bar{r}$ is a solution of the convex problem, $\min\{g(s) : s \in \mathcal{C}\}$ if and only if $\partial g(\bar{r}) + N_{\mathcal{C}}(\bar{r}) \ni 0$, where $\partial g(\bar{r})$ and $N_{\mathcal{C}}(\bar{r})$ respectively denote the subdifferential of g and normal cone of $\mathcal{C}$ at $\bar{r}$.*

Lemma 4.2.4. *[82] Let $\{t_n\}$ be a nonnegative sequence in $\mathcal{H}$, $\{\sigma_n\} \subset (0,1)$ such that $\sum_{i=1}^{\infty}\sigma_n = \infty$, and $\{d_n\}$ be any sequence of real numbers.*
Assume that

$$t_{n+1} \leq (1-\sigma_n)t_n + \sigma_n d_n \quad \forall n \geq 1.$$

If $\limsup_{k\to\infty} d_{n_k} \leq 0$ *for each subsequence* $\{t_{n_k}\}$ *of* $\{t_n\}$ *satisfying* $\liminf_{k\to\infty}\left(t_{n_{k+1}} - t_{n_k}\right) \geq 0$, *then* $\lim_{n\to\infty} t_n = 0$.

Lemma 4.2.5. *The following is true for each $x, y \in \mathcal{H}$ and $\beta \in \mathbb{R}$:*

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

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

Lemma 4.2.6. *[59] Let $\{t_n\}$, $\{\delta_n\} \subset \mathbb{R}_+$ such that*
$t_{n+1} \leq (1-\delta_n)t_n + m_n + \gamma_n$, $\forall n \geq 1$,
where $\delta_n \in (0,1)$ and $\{m_n\}$ are real sequences. Suppose $\sum_{n=0}^{\infty}\gamma_n < \infty$ and $m_n \leq \delta_n\beta$ for some $\beta \geq 0$, therefore $\{t_n\}$ is a bounded sequence.

4.3 Main Results

Suppose the following conditions hold for the bifunction f :

Assumption 4.3.1.

(A1) For each $s \in \mathcal{C}$, $f(s,s) = 0$ and f is pseudomonotone on $\mathcal{C}$;

(A2) with two positive constants δ_1 and δ_2, f satisfies the Lipschitz-type condition on $\mathcal{H}$;

(A3) $f(s,.)$ is convex, lower semicontinuous and subdifferentiable on $\mathcal{C}$, $\forall s \in \mathcal{C}$;

(A4) f *is weakly and jointly continuous on* $\mathcal{C} \times \mathcal{C}$. *That is, if* $s, z \in \mathcal{C}$ *and* $s_n \rightharpoonup s$, $z_n \rightharpoonup z$ *as* $n \to \infty$, *then* $\limsup_{n\to\infty} f(s_n, z_n) \to f(s, z)$, *particularly, in other words,* $\limsup_{n\to\infty} f(s_n, z_n) \leq f(s, z)$ *for each fixed* $s, z \in \mathcal{C}$ *and* $s_n, z_n \in \mathcal{C}$;

(A5) $EP(f, \mathcal{C}) \neq \emptyset$.

Assumption 4.3.2.
Let $\alpha_n \in (0, 1)$ *and* $\liminf_{n\to\infty} \alpha_n > 0$. *Assume that* $\{\xi_n\}$ *is any positive sequence such that* $\xi_n = o(\sigma_n)$, *that is,* $\lim_{n\to\infty} \frac{\xi_n}{\sigma_n} = 0$, *where* $\{\sigma_n\} \subset (0, 1)$ *such that*

$$\lim_{n\to\infty} \sigma_n = 0, \quad \sum_{n=1}^{\infty} \sigma_n = \infty.$$

Now, we are ready to present and analyze our algorithm as follows:

Algorithm 4.3.3. *Let* $x_0,\ x_1 \in \mathcal{H},\ \lambda_1 > 0,\ \mu \in]0, 1[$.
STEP 1*:*
Given the current iterates x_{n-1} *and* x_n, *choose* θ_n *such that* $0 \leq \theta_n \leq \bar{\theta}_n$, *where*

$$\bar{\theta}_n = \begin{cases} \min\left\{\theta, \dfrac{\xi_n}{\|x_n - x_{n-1}\|}\right\} & \text{if } x_n \neq x_{n-1}, \\ \theta & \text{otherwise.} \end{cases} \tag{4.3.1}$$

STEP 2*:*
Set $w_n = (1 - \sigma_n)(x_n + \theta_n(x_n - x_{n-1}))$ *and evaluate*
$y_n = \underset{y\in\mathcal{C}}{arg\min} \left\{\lambda_n f(w_n, y) + \frac{1}{2}\|w_n - y\|^2\right\}$.
Stop the computation if $y_n = w_n$, *and* y_n *becomes a solution, else, move to the next step which is step 3.*

STEP 3*:*
Choose $v_n \in \partial f(w_n, .)(y_n)$ *in such a way there is*
$q_n \in N_{\mathcal{C}}(y_n)$ *satisfying*
$q_n = w_n - \lambda_n v_n - y_n$, *and construct a half space*
$T_n = \{z \in \mathcal{H} : \langle w_n - \lambda_n v_n - y_n, y - y_n\rangle \leq 0\}$.
Evaluate
$z_n = \underset{y\in T_n}{arg\min} \left\{\lambda_n f(y_n, y) + \frac{1}{2}\|w_n - y\|^2\right\}$.

STEP 4*: Compute*

$x_{n+1} = (1-\alpha_n)w_n + \alpha_n z_n$, *and*

$$\lambda_{n+1} = \begin{cases} \min\left\{\dfrac{\mu}{2}\dfrac{\|w_n - y_n\|^2 + \|z_n - y_n\|^2}{f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n)}, \quad \lambda_n\right\}, & \text{if } f(w_n, z_n) - f(w_n, y_n) \\ & -f(y_n, z_n) > 0, \\ \lambda_n, \quad \text{otherwise.} \end{cases}$$

Set $n+1 \longleftarrow n$ *and continue again from* ***step 1***.

Remark 4.3.4.
From (4.3.1)*, the line rule condition, one can easily see that*

$$\theta_n \leq \bar{\theta}_n \leq \frac{\xi_n}{\|x_n - x_{n-1}\|},$$

This implies that

$$0 \leq \frac{\theta_n}{\sigma_n}\|x_n - x_{n-1}\| \leq \frac{\xi_n}{\sigma_n}. \tag{4.3.2}$$

Also, since $\xi_n = o(\sigma_n)$*, there is* $\gamma > 0$ *such that*

$$\frac{\xi_n}{\sigma_n} \leq \gamma \quad \text{for each } n \geq 1. \tag{4.3.3}$$

Lemma 4.3.5. *Suppose the bifunction* f *satisfies Assumption* 4.3.1*, then the sequence* $\{x_n\}$ *produced by Algorithm* 4.3.3 *is bounded.*

Proof.
Using the definition of y_n and Lemma (4.2.3) gives us that

$$0 \in \partial\left(\lambda_n f(y_n, y) + \frac{1}{2}\|y - y_n\|^2\right)(y_n) + \mathcal{N}_{\mathcal{C}}(y_n).$$

From Step 2, we obtain

$$\begin{aligned} \langle w_n - y_n, y - y_n\rangle &= \langle \lambda_n v_n + q_n, y - y_n\rangle \\ &= \lambda_n\langle v_n, y - y_n\rangle + \langle q_n, y - y_n\rangle. \end{aligned} \tag{4.3.4}$$

But $q_n \in \mathcal{N}_{\mathcal{C}}(y_n)$ implies that $\langle q_n, y - y_n\rangle \leq 0 \quad \forall y \in \mathcal{C}$.
Thus we obtain from (4.3.4) that

$$\langle w_n - y_n, y - y_n\rangle \leq \lambda_n\langle v_n, y - y_n\rangle \quad \forall y \in T_n. \tag{4.3.5}$$

Also, $v_n \in \partial f(w_n, y_n)$ implies that

$$f(w_n, y) - f(w_n, y_n) \geq \langle v_n, y - y_n\rangle \quad \forall y \in \mathcal{C}. \tag{4.3.6}$$

Using (4.3.5) and (4.3.6), we have

$$\lambda_n(f(w_n, y) - f(w_n, y_n)) \geq \langle w_n - y_n, y - y_n\rangle \quad \forall y \in \mathcal{C}. \tag{4.3.7}$$

Putting $z_n = y$ in (4.3.7), we see that

$$\lambda_n(f(w_n, z_n) - f(w_n, y_n)) \geq \langle w_n - y_n, z_n - y_n\rangle. \tag{4.3.8}$$

Again, using the definition of z_n and Lemma 4.2.2, we obtain

$$\lambda_n(f(y_n, y) - f(y_n, z_n)) \geq \langle w_n - z_n, y - z_n\rangle \quad \forall y \in \mathcal{C}. \tag{4.3.9}$$

Let $y = \bar{r} \in EP(f, \mathcal{C})$.
Then,

$$\lambda_n(f(y_n, \bar{r}) - f(y_n, z_n)) \geq \langle w_n - z_n, \bar{r} - z_n\rangle \quad \forall \bar{r} \in EP(f, \mathcal{C}). \tag{4.3.10}$$

But $\bar{r} \in EP(f, \mathcal{C})$, implies $f(\bar{r}, y_n) \geq 0$. This further implies by pseudomonotonicity of f that

$$f(y_n, \bar{r}) \leq 0.$$

Therefore, (4.3.10) becomes,

$$-\lambda_n f(y_n, z_n) \geq \langle w_n - z_n, \bar{r} - z_n\rangle. \tag{4.3.11}$$

Combining (4.3.8) and (4.3.11) implies that

$$\begin{aligned} 2\lambda_n(f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n)) \quad &\geq \quad 2\langle w_n - y_n, z_n - y_n\rangle \\ &\quad +2\langle w_n - z_n, \bar{r} - z_n\rangle. \end{aligned} \tag{4.3.12}$$

Note that:

$$2\langle w_n - y_n, z_n - y_n\rangle = \|w_n - y_n\|^2 + \|z_n - y_n\|^2 - \|w_n - z_n\|^2. \tag{4.3.13}$$

Also,

$$2\langle w_n - z_n, \bar{r} - z_n\rangle = \|w_n - z_n\|^2 + \|z_n - \bar{r}\|^2 - \|w_n - \bar{r}\|^2. \tag{4.3.14}$$

Using (4.3.13) and (4.3.14) in (4.3.12), we get

$$\begin{aligned} 2\lambda_n(f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n)) \quad &\geq \quad \|w_n - y_n\|^2 + \|z_n - y_n\|^2 \\ &\quad +\|z_n - \bar{r}\|^2 - \|w_n - \bar{r}\|^2. \end{aligned} \tag{4.3.15}$$

By the definition of λ_n, we have

$$2(f(w_n, z_n) - f(w_n, y_n) - f(y_n, z_n)) \leq \frac{\mu(\|w_n - y_n\|^2 + \|z_n - y_n\|^2)}{\lambda_{n+1}}. \tag{4.3.16}$$

Using (4.3.16) in (4.3.15), we obtain

$$\begin{aligned}\mu\frac{\lambda_n}{\lambda_{n+1}}(\|w_n-y_n\|^2+\|z_n-y_n\|^2) &\geq \|w_n-y_n\|^2+\|z_n-y_n\|^2\\ &\quad +\|z_n-\bar{r}\|^2-\|w_n-\bar{r}\|^2. \qquad (4.3.17)\end{aligned}$$

This further implies that

$$\|z_n-\bar{r}\|^2\leq\|w_n-\bar{r}\|^2-(1-\mu\frac{\lambda_n}{\lambda_{n+1}})\|w_n-y_n\|^2-(1-\mu\frac{\lambda_n}{\lambda_{n+1}})\|z_n-y_n\|^2. \quad (4.3.18)$$

Moreover, with $\{\alpha_n\}\subset(0,1)$, we have

$$\|x_{n+1}-\bar{r}\|^2=(1-\alpha_n)\|w_n-\bar{r}\|^2+\alpha_n\|z_n-\bar{r}\|^2-(1-\alpha_n)\alpha_n\|z_n-w_n\|^2. \quad (4.3.19)$$

Using (4.3.18) and the definition of $\{x_{n+1}\}$, that is, $z_n-w_n=\dfrac{1}{\alpha_n}(x_{n+1}-w_n)$ in (4.3.19), we establish that

$$\begin{aligned}\|x_{n+1}-\bar{r}\|^2 &\leq \|w_n-\bar{r}\|^2-\alpha_n(1-\mu\frac{\lambda_n}{\lambda_{n+1}})(\|w_n-y_n\|^2+\|z_n-y_n\|^2)\\ &\quad -\frac{1-\alpha_n}{\alpha_n}\|x_{n+1}-w_n\|^2 \quad \forall n\geq 1. \qquad (4.3.20)\end{aligned}$$

Define $h_n:=x_n+\theta_n(x_n-x_{n-1})$.
From (4.3.20), using the facts that $\liminf\limits_{n\to\infty}\alpha_n>0$, and

$$\lim_{n\to\infty}\left(1-\mu\frac{\lambda_n}{\lambda_{n+1}}\right)=1-\mu>0,$$

so, there is $N_0\in\mathbb{N}$ for which

$$\lim_{n\to\infty}\alpha_n\left(1-\mu\frac{\lambda_n}{\lambda_{n+1}}\right)>0 \quad \forall n\geq N_0.$$

Therefore,

$$\begin{aligned}\|x_{n+1}-\bar{r}\| &\leq \|w_n-\bar{r}\|\\ &= \|(1-\sigma_n)h_n-\bar{r}\|\\ &= \|(1-\sigma_n)(h_n-\bar{r})-\sigma_n\bar{r}\|\\ &\leq (1-\sigma_n)\|h_n-\bar{r}\|+\sigma_n\|\bar{r}\| \quad \forall n\geq N_0. \qquad (4.3.21)\end{aligned}$$

But

$$\begin{aligned}\|h_n-\bar{r}\| &= \|x_n+\theta_n(x_n-x_{n-1})-\bar{r}\|\\ &= \|x_n-\bar{r}+\theta_n(x_n-x_{n-1})\|\\ &\leq \|x_n-\bar{r}\|+\theta_n\|x_n-x_{n-1}\|. \qquad (4.3.22)\end{aligned}$$

Putting (4.3.22) in (4.3.21), and using (4.3.2) and (4.3.3) of Remark 4.3.4 we have

$$\begin{aligned}\|x_{n+1}-\bar{r}\| &\leq (1-\sigma_n)\|x_n-\bar{r}\|+(1-\sigma_n)\theta_n\|x_n-x_{n-1}\|+\sigma_n\|\bar{r}\| \\ &\leq (1-\sigma_n)\|x_n-\bar{r}\|+\sigma_n\left((1-\sigma_n)\frac{\xi_n}{\sigma_n}+\|\bar{r}\|\right) \\ &\leq (1-\sigma_n)\|x_n-\bar{r}\|+\sigma_n(\gamma+\|\bar{r}\|) \\ &= (1-\sigma_n)\|x_n-\bar{r}\|+\sigma_n\beta, \end{aligned} \tag{4.3.23}$$

where $\beta=(\gamma+\|\bar{r}\|)>0$.
Therefore, by Lemma 4.2.6, we have that $\{\|x_n-\bar{r}\|\}$ is bounded which implies that the sequence $\{x_n\}$ is bounded. Thus, as a consequence, $\{w_n\}$, $\{y_n\}$ are also bounded. □

Lemma 4.3.6. *Let $\{\sigma_n\}$, $\{\alpha_n\}\subset(0,1)$, $\mu\in(0,1)$ and $\{\lambda_n\}$ be a real sequence. If $\{x_n\}$ is a sequence generated by Algorithm 4.3.3. Then for each $n\geq N_0$, we have*

$$\alpha_n\left(1-\mu\frac{\lambda_n}{\lambda_{n+1}}\right)(\|w_n-y_n\|^2+\|z_n-y_n\|^2)\leq\|x_n-\bar{r}\|^2-\|x_{n+1}-\bar{r}\|^2+\sigma_n\Phi_1.$$

Proof.

$$\begin{aligned}\|x_{n+1}-\bar{r}\|^2 &= \|(1-\alpha_n)w_n+\alpha_n z_n-\bar{r}\|^2 \\ &= \|(1-\alpha_n)(w_n-\bar{r})+\alpha_n(z_n-\bar{r})\|^2 \\ &= (1-\alpha_n)^2\|w_n-\bar{r}\|^2+\alpha_n^2\|z_n-\bar{r}\|^2+2\alpha_n(1-\alpha_n)\langle w_n-\bar{r},z_n-\bar{r}\rangle \\ &= (1-\alpha_n)^2\|w_n-\bar{r}\|^2+\alpha_n^2\|z_n-\bar{r}\|^2 \\ &\quad +\alpha_n(1-\alpha_n)(\|w_n-\bar{r}\|^2+\|z_n-\bar{r}\|^2-\|w_n-z_n\|^2) \\ &\leq (1-\alpha_n)\|w_n-\bar{r}\|^2+\alpha_n\|z_n-\bar{r}\|^2. \end{aligned} \tag{4.3.24}$$

On the other hand, from (4.3.21)—(4.3.23) also shows that
$\|w_n-\bar{r}\|\leq(1-\sigma_n)\|x_n-\bar{r}\|+\sigma_n\beta$.
This implies that

$$\begin{aligned}\|w_n-\bar{r}\|^2 &\leq (1-\sigma_n)^2\|x_n-\bar{r}\|^2+2\sigma_n\beta(1-\sigma_n)\|x_n-\bar{r}\|+\sigma_n^2\beta^2 \\ &= (1-\sigma_n)^2\|x_n-\bar{r}\|^2+\sigma_n[2\beta(1-\sigma_n)\|x_n-\bar{r}\|+\sigma_n\beta^2] \\ &\leq \|x_n-\bar{r}\|^2+\sigma_n\Phi_1, \end{aligned} \tag{4.3.25}$$

for some $\Phi_1>0$, where $\Phi_1:=\max_{n\in\mathbb{N}}\{\sigma_n 2\beta(1-\sigma_n)\|x_n-\bar{r}\|+\sigma_n\beta^2\}$.
Substituting (4.3.18), (4.3.25) in (4.3.24) gives

$$\|x_{n+1}-\bar{r}\|^2\leq\|x_n-\bar{r}\|^2-\alpha_n\left(1-\mu\frac{\lambda_n}{\lambda_{n+1}}\right)(\|w_n-y_n\|^2+\|z_n-y_n\|^2)+\sigma_n\Phi_1.$$

Therefore,

$$\begin{aligned}&\alpha_n\left(1-\mu\frac{\lambda_n}{\lambda_{n+1}}\right)(\|w_n-y_n\|^2+\|z_n-y_n\|^2) \\ &\leq\|x_n-\bar{r}\|^2-\|x_{n+1}-\bar{r}\|^2+\sigma_n\Phi_1 \end{aligned} \tag{4.3.26}$$

as required. □

Theorem 4.3.7. *Suppose Assumption 4.3.1 is true, therefore, the generated sequence* $\{x_n\}$ *of Algorithm 4.3.3 strongly converges to* $\bar{r} = P_{EP(f,\mathcal{C})}(0)$.

Proof.

$$\begin{aligned}
\|w_n - \bar{r}\|^2 &= \|(1-\sigma_n)(x_n + \theta_n(x_n - x_{n-1})) - \bar{r}\|^2 \\
&= \|(1-\sigma_n)(x_n - \bar{r}) + (1-\sigma_n)\theta_n(x_n - x_{n-1}) - \sigma_n \bar{r}\|^2 \\
&\leq \|(1-\sigma_n)(x_n - \bar{r}) + (1-\sigma_n)\theta_n(x_n - x_{n-1})\|^2 + 2\sigma_n\langle \bar{r}, \bar{r} - w_n\rangle \\
&\leq (1-\sigma_n)[\|x_n - \bar{r}\|^2 + \theta_n^2\|x_n - x_{n-1}\|^2 + 2\theta_n\|x_n - \bar{r}\|\|x_n - x_{n-1}\|] \\
&\quad +2\sigma_n\langle \bar{r}, \bar{r} - w_n\rangle \\
&\leq (1-\sigma_n)\|x_n - \bar{r}\|^2 + 2(1-\sigma_n)\theta_n\|x_n - \bar{r}\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2 \\
&\quad + 2\sigma_n\langle \bar{r}, x_{n+1}\rangle + 2\sigma_n\langle \bar{r}, \bar{r} - x_{n+1}\rangle \\
&\leq (1-\sigma_n)\|x_n - \bar{r}\|^2 + 2(1-\sigma_n)\theta_n\|x_n - \bar{r}\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2 \\
&\quad + 2\sigma_n\|\bar{r}\|\|w_n - x_{n+1}\| + 2\sigma_n\langle \bar{r}, \bar{r} - x_{n+1}\rangle \\
&\leq (1-\sigma_n)\|x_n - \bar{r}\|^2 + \sigma_n[2(1-\sigma_n)\|x_n - \bar{r}\|\frac{\theta_n}{\sigma_n}\|x_n - x_{n-1}\| \\
&\quad + \left(\frac{\theta_n}{\sigma_n}\|x_n - x_{n-1}\|\right)^2 + 2\|\bar{r}\|\|w_n - x_{n+1}\| \\
&\quad +\langle \bar{r}, \bar{r} - x_{n+1}\rangle] \qquad \text{(See[89], step 5).}
\end{aligned} \tag{4.3.27}$$

Define $d_n := 2(1-\sigma_n)\|x_n - \bar{r}\|\frac{\theta_n}{\sigma_n}\|x_n - x_{n-1}\| + \left(\frac{\theta_n}{\sigma_n}\|x_n - x_{n-1}\|\right)^2 + 2\|\bar{r}\|\|w_n - x_{n+1}\| + \langle \bar{r}, \bar{r} - x_{n+1}\rangle$.
Since $\lim_{n\to\infty}\alpha_n\left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right) > 0$ for each $n \geq N_0$, thus, neglecting the nonpositive terms in (4.3.20) and substituting 4.3.27, we arrive at

$$\|x_{n+1} - \bar{r}\|^2 \leq (1-\sigma_n)\|x_n - \bar{r}\|^2 + \sigma_n d_n \quad \text{for any} \quad n \geq N_0. \tag{4.3.28}$$

Now, we show that $\{\|x_n - \bar{r}\|\} \to 0$, and to do that, we use (4.3.28) and Lemma 4.2.4 and show that $\limsup_{k\to\infty} d_{n_k} \leq 0$, for any subsequence $\{\|x_{n_k} - \bar{r}\|\}$ of $\{\|x_n - \bar{r}\|\}$ which satisfies $\liminf_{k\to\infty}\left(\|x_{n_{k+1}} - \bar{r}\| - \|x_{n_k} - \bar{r}\|\right) \geq 0$.
For this purpose, let $\{\|x_{n_k} - \bar{r}\|\}$ be a subsequence of $\{\|x_n - \bar{r}\|\}$ for which $\liminf_{k\to\infty}\left(\|x_{n_{k+1}} - \bar{r}\| - \|x_{n_k} - \bar{r}\|\right) \geq 0$ holds.
Then,

$$\begin{aligned}
&\liminf_{k\to\infty}\left(\|x_{n_{k+1}} - \bar{r}\|^2 - \|x_{n_k} - \bar{r}\|^2\right) \\
&= \liminf_{k\to\infty}\left((\|x_{n_{k+1}} - \bar{r}\| - \|x_{n_k} - \bar{r}\|)(\|x_{n_{k+1}} - \bar{r}\| + \|x_{n_k} - \bar{r}\|)\right) \geq 0.
\end{aligned}$$

From Lemma 4.3.6, we have

$$\begin{aligned} 0 &\leq \limsup_{k\to\infty}\left[\alpha_{n_k}\left(1-\mu\frac{\lambda_{n_k}}{\lambda_{n_{k+1}}}\right)(\|w_{n_k}-y_{n_k}\|^2+\|z_{n_k}-y_{n_k}\|^2)\right] \\ &\leq \limsup_{k\to\infty}\left[\|x_{n_k}-\bar{r}\|^2-\|x_{n_{k+1}}-\bar{r}\|^2+\sigma_{n_k}\Phi_1\right] \\ &\leq \limsup_{k\to\infty}-\left[\|x_{n_{k+1}}-\bar{r}\|^2-\|x_{n_k}-\bar{r}\|^2\right]+\limsup_{k\to\infty}\sigma_{n_k}\Phi_1 \\ &= -\liminf_{k\to\infty}\left[\|x_{n_{k+1}}-\bar{r}\|^2-\|x_{n_k}-\bar{r}\|^2\right] \\ &\leq 0. \end{aligned}$$

But $\liminf_{n\to\infty}\alpha_n > 0$ and $\lim_{n\to\infty}\left(1-\mu\frac{\lambda_n}{\lambda_{n+1}}\right) = 1-\mu > 0$.
Thus, from the above observations, we have

$$\lim_{k\to\infty}\|w_{n_k}-y_{n_k}\| = 0, \quad \text{and as well} \quad \lim_{k\to\infty}\|z_{n_k}-y_{n_k}\| = 0. \tag{4.3.29}$$

This further implies that $\lim_{k\to\infty}\|w_{n_k}-z_{n_k}\| = 0$.
Using the definition of $x_{n+1} = (1-\alpha_n)w_n + \alpha_n z_n$, we obtain the following:

$$\lim_{k\to\infty}\|x_{n_{k+1}}-w_{n_k}\| = \lim_{k\to\infty}\alpha_{n_k}\|w_{n_k}-z_{n_k}\| = 0. \tag{4.3.30}$$

Again, the definition of w_n and Lemma 4.2.5ii give us that

$$\begin{aligned} \|w_n-x_n\|^2 &= \|\sigma_n x_n-(1-\sigma_n)\theta_n(x_n-x_{n-1})\|^2 \\ &= \sigma_n\|x_n\|^2+(1-\sigma_n)\theta_n^2\|x_n-x_{n-1}\|^2-\sigma_n(1-\sigma_n)\|h_n\|^2. \end{aligned}$$

Thus

$$\|w_{n_k}-x_{n_k}\|^2 \leq \sigma_{n_k}\|x_{n_k}\|^2+\sigma_{n_k}(1-\sigma_{n_k})\frac{\theta_{n_k}}{\sigma_{n_k}}\|x_{n_k}-x_{n_{k-1}}\|\theta_{n_k}\|x_{n_k}-x_{n_{k-1}}\|.$$

Taking the limit as $k\to\infty$, noting that $\lim_{k\to\infty}\sigma_{n_k} = 0$, and $\lim_{k\to\infty}\frac{\theta_{n_k}}{\sigma_{n_k}}\|x_{n_k}-x_{n_{k-1}}\| = 0$, we obtain

$$\|w_{n_k}-x_{n_k}\| \to 0, \quad k\to\infty. \tag{4.3.31}$$

Again,

$$0 \leq \|x_{n_{k+1}}-x_{n_k}\| \leq \|x_{n_{k+1}}-w_{n_k}\|+\|w_{n_k}-x_{n_k}\|.$$

Taking the limit as $k\to\infty$, we have

$$\|x_{n_{k+1}}-x_{n_k}\| \to 0, \quad k\to\infty.$$

Since the sequence $\{x_{n_k}\}$ is bounded, then there is a converging subsequence $\{x_{n_{k_j}}\} \subset \{x_{n_k}\}$ which converges weakly to some $r^* \in \mathcal{H}$, in such a way

$$\begin{aligned} \limsup_{k\to\infty}\langle\bar{r},\bar{r}-x_{n_k}\rangle &= \lim_{j\to\infty}\langle\bar{r},\bar{r}-x_{n_{k_j}}\rangle \\ &= \langle\bar{r},\bar{r}-r^*\rangle. \end{aligned}$$

(4.3.31) and (4.3.29) show that $w_{n_k} \rightharpoonup r^*$, $y_{n_k} \rightharpoonup r^*$, as $k\to\infty$.

Remark 4.3.8.
Recall that $\mathcal{C} \subset \mathcal{H}$ is closed and convex. This also means that it is weakly closed, hence, $r* \in \mathcal{C}$.

We now show that $r^* \in EP(f, \mathcal{C})$.
From (4.3.7) and since $\liminf_{n\to\infty} \lambda_n > 0$, we obtain

$$f(w_{n_k}, y) \geq f(w_{n_k}, y_{n_k}) + \frac{1}{\lambda_{n_k}} \langle w_{n_k} - y_{n_k}, y - y_{n_k} \rangle \quad \forall y \in \mathcal{C}. \tag{4.3.32}$$

Taking the limit of (4.3.32) as $k \to \infty$, and using conditions (A1), (A4), we obtain that $f(r^*, y) \geq 0 \ \forall y \in \mathcal{C}$. Thus, showing that $r^* \in EP(f, \mathcal{C})$.

By metric projection property, with $\bar{r} = P_{EP(f,\mathcal{C})}(0) \in \mathcal{C}$, and $r^* \in EP(f, \mathcal{C}) \subset \mathcal{C}$, we have

$$\begin{aligned} \limsup_{k\to\infty} \langle \bar{r} - 0, \bar{r} - x_{n_k} \rangle &= \langle \bar{r} - 0, \bar{r} - r^* \rangle \\ &\leq 0. \end{aligned} \tag{4.3.33}$$

Combining (4.3.30), (4.3.33) and using the fact that $\frac{\theta_{n_k}}{\sigma_{n_k}} \|x_{n_k} - x_{n_{k+1}}\| \to 0$, $k \to \infty$, we have $\limsup_{k\to\infty} d_{n_k} \leq 0$ for any subsequence $\{\|x_{n_k} - \bar{r}\|\} \subset \{\|x_n - \bar{r}\|\}$. Thus, by Lemma 4.2.4, $\lim_{n\to\infty} \|x_n - \bar{r}\| = 0$, that is, $x_n \to \bar{r}$, $n \to \infty$ as asserted. □

Chapter 5

Conclusion, Contribution to Knowledge and Future Research

5.1 Conclusion

Systematically, this dissertation introduced a comprehensive study of the methods of approximating a real solution of an EP in the real Hilbert spaces. The study was presented in a coherent manner, where firstly, we gave a brief knowledgeable background of our study. We recalled some important definitions, propositions, lemmas, theorems and remarks that form the backbone to the convergence results obtained in chapters 3 and 4, and thereafter, reviewed some great works that have been done in the literature. The inertial and relaxation technique introduced in our method in chapter 3 helped us to obtain weak convergence while the incorporation of the line rule condition to the method presented in chapter 4 allowed us to achieve a strong convergent result. The result of the method in chapter 3 was extended to the variational inequality and fixed point problems.

5.2 Contribution to knowledge

Our study done in this direction has contributed the following to the knowledge of equilibrium problems and convex analysis in finite and infinite dimensional Hilbert spaces:

- introduced and analyzed more manageable modified methods of solving equilibrium problem using the relaxation techniques and inertial parameters.
- involved self-adaptive step-sizes that update at every iteration and do not require the prior awareness of the Lipschitz-type constants nor use the Armijo linesearch to obtain them as was done in [44, Algorithm 3.3] [4, 24, 42, 56, 94]. Also, our self-adaptive step-sizes are more practically applicable when compared with the proposed scheme in [92] which uses constant stepsize.
- under mild condition, obtains weak and strong convergence algorithms to a solution of the EP(3.1.1) without the condition of strong pseudomonotonicity of the

bifunction f as it was done in [36, Theorem 4.1], [78, Algorithm 1].

- used the nonnegative sequence $\{\theta_n\}$ reflecting the inertial part to be in $[0, 1)$ which is quite different from what is obtainable in [86], where $\theta_n \in [0, \frac{1}{3})$.

5.3 Future Research

There is a good number of research works of many scholars that discussed the equilibrium problems in Hilbert spaces, though more works are still needed to be done in this direction as equilibrium problems have not been fully exploited.
In future study, it is our desire to extend some of the results achieved for the problem considered in this dissertation, and other convex problems to an interesting space known as the Hadamard space. Also, one may attempt to achieve similar results as they are in chapter 4 without the line rule condition. Let us briefly describe the Hadamard space:

Definition 5.3.1. *Let the pair $(\mathcal{E}, d)$ be a metric space with the metric d. For each $s, z \in \mathcal{E}$, an operator $\phi : [0, d(s, z)] \longrightarrow \mathcal{E}$ is called a geodesic with end endpoints s, z if $\phi(0) = s$, $\phi(d(s, z)) = z$ and $d(\phi(t), \phi(w)) = |t - w|$ for every $t, w \in [0, d(s, z)]$.*

The image of ϕ is known as a geodesic (or metric) segment joining s to z. If the geodesic is uniquely determined for each $s, z \in \mathcal{E}$, then the image of the geodesic segment or the geodesic with endpoints $s, z \in \mathcal{E}$ is denoted by $[s, z]$. $(\mathcal{E}, d)$ is called a geodesic space whenever there is a geodesic for each endpoints in $\mathcal{E}$, and $\mathcal{E}$ is said to be uniquely geodesic if there exists exactly one geodesic joining s and z for any $s, z \in \mathcal{E}$.

Definition 5.3.2. *Suppose $\mathcal{E}$ is a uniquely geodesic space and $\mathcal{C} \subset \mathcal{E}$. Then $\mathcal{C}$ is called a convex subset of $\mathcal{E}$ if $[s, z] \subset \mathcal{C}$ for every $s, z \in \mathcal{C}$.*

A geodesic triangle $\triangle(s_1, s_2, s_3)$ in a geodesic metric space $\mathcal{E}$ consists of three points $s_1, s_2, s_3 \in \mathcal{E}$ (the vertices of $\triangle$), and a geodesic segment between each pair of vertices (the edges of $\triangle$). A comparison triangle for $\triangle(s_1, s_2, s_3)$ in $\mathcal{E}$ is a triangle $\overline{\triangle}(s_1, s_2, s_3) = \triangle(\bar{s}_1, \bar{s}_2, \bar{s}_3)$ in the Euclidean plane $\mathbb{R}^2$ such that $d_{\mathbb{R}^2}(\bar{s}_i, \bar{z}_j) = d(s_i, z_j)$ for each $i, j \in \{1, 2, 3\}$, where $d_{\mathbb{R}^2}$ is the usual metric in $\mathbb{R}^2$. A geodesic metric space $\mathcal{E}$ is called a $CAT(0)$ space if all geodesic triangles satisfy the following axiom:

Let $\triangle$ be a geodesic triangle in $\mathcal{E}$, and let $\overline{\triangle}$ be its comparison triangle in $\mathbb{R}^2$. Then, $\triangle$ is said to satisfy $CAT(0)$ inequality if for each $s, z \in \triangle$ and all comparison points $\bar{s}, \bar{z} \in \overline{\triangle}$

$$d(s, z) = d_{\mathbb{R}^2}(\bar{s}, \bar{z}).$$

A complete $CAT(0)$ space is called a Hadamard space.

Definition 5.3.3. *A point $w \in CAT(0)$ denoted by $w = (1 - \lambda)s \oplus \lambda z$ is called a unique point in $CAT(0)$ if $d(s, w) = \lambda d(s, z)$ and $d(z, w) = (1 - \lambda)d(s, z)$ for all $\lambda \in [0, 1]$.*

An interesting thing that draws our attention to the Hadamard spaces is the fact that any real inner product space (not necessarily complete) is a $CAT(0)$ space and conversely, if

a real normed vector space is a $CAT(p)$ space for some real p, then it is an inner product space. Thus, in this space, we intend to study the equilibrium problem defined by

$$\text{obtain } \bar{r} \in \mathcal{C}, \text{ a convex subset of the geodesic metric space } \mathcal{E}, \\ \text{for which } f(\bar{r}, s) \not< 0, \quad \forall s \in \mathcal{C}.$$

Where $f : \mathcal{C} \subset \mathcal{E} \times \mathcal{C} \subset \mathcal{E} \longrightarrow \mathbb{R}$ is a bifunction such that $f(u, u) = 0$ for each $u \in \mathcal{C}$.

Bibliography

[1] Y. I. Alber, Metric and generalized projection operator in Banach spaces: Properties and applications, *Theory and Applications of Nonlinear Operators of Accretive and Monotone Type,* **178** of Lecture Notes in Pure and Applied Mathematics, New York: Dekker, (1996), 15-50.

[2] P. K. Anh, T.N. Hai, Novel self-adaptive algorithms for non-lipschitz equilibrium problems with applications, *J. Glob. Optim.,* **73** (2019) 637–657

[3] H. Attouch, A. Cabot, Convergence of a relaxed inertial proximal algorithm for maximally monotone operators, *Math Program.,* **184** (2020) 243–287.

[4] H. Attouch, A. Cabot, Convergence of a relaxed inertial forward–backward algorithm for structured monotone inclusions, *Appl Math Optim.,* **80** (2019) 547–598.

[5] J. P. Aubin, I. Ekeland, Appliead Nonlinear Analysis, *Wiley New York,* (1984).

[6] H. H. Bauschke, P. L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, *CMS-SMC,* (2011)

[7] L. P. Belluce, W. A. Kirk, E. F. Steiner, Normal structure in Banach spaces, *Pac. J. Math.,* **26**(3) (1968).

[8] M. Bianchi, S. Schaible, Equilibrium problems under generalized convexity and generalized monotonicity, *J. Glob. Optim.,* **30** (2004) 121-134.

[9] M. Bianchi, S. Schaible, Generalized Monotone Bifunctions and Equilibrium Problems, *J. Optim. Theory Appl.,* **90**(1) (1996) 31-43.

[10] G. Bigi, M. Castellani, M. Pappalardo, A new solution method for equilibrium problems, *Optim. Methods Softw.,* **24** (2009) 895–911.

[11] G. Bigi, M. Castellani, M. Pappalardo, Nonlinear programming techniques for equilibria, *Cham: Springer,* (2019).

[12] G. Bigi, M. Passacantando, Descent and penalization techniques for equilibrium problems with nonlinear constraints, *J. Optim. Theory Appl.* **164** (2015) 804–818.

[13] G. Bigi, M. Passacantando, Gap functions for quasi-equilibria, *J. Global Optim.,* **66** (2016) 791–810.

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

[15] H. Brezis, L. Niremerg, G. Stampacchia, A remark on Fan's minimax principle, *Bollettino della Unione Matematica Italiana,* **6** (1972) 293-300.

[16] F. E. Browder and W. V. Petryshyn, Construction of fixed point of nonlinear mappings in Hilbert spaces, *J. Math. Anal. Appl.,* **20**(1967), 197-228.

[17] Y. Censor, A. Gibali, S. Riech, Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space, *Optim. Methods Softw.,* **26**(4–5) (2011) 827–845.

[18] Y. Censor, A. Gibali, S. Riech, The subgradient extragradient method for solving variational inequalities in Hilbert space, *J. Optim. Theory Appl.,* **148**(2) (2011) 318–335.

[19] C. E. Chidume, Applicable functional analysis, *Ibadan University Press*, ISBN 978-978-8456-13-5, (2014).

[20] C. E. Chidume, Geometric properties of Banach spaces and nonlinear iterations, *Springer Verlag Series,* Lecture Notes in Mathematics, ISBN 978-1-84882-189-7, (2009).

[21] I. Cioranescu, Geometry of Banach spaces, duality mappings and nonlinear problems, *Kluwer Academic Publisher* (1990).

[22] L. Deng, R. Hu, Y.-P. Fang, Fast inertial extragradient algorithms for solving non-Lipschitzian equilibrium problems without monotonicity condition in real Hilbert spaces, *J. Comput. Appl. Math.,* **423** (2023) 114950.

[23] P. Daniele, F. Giannessi, A. Maugeri, Equilibruim problems and variational models, *Dordrecht: Kluwer Academic,* (2003).

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

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

[26] G. Fichera, Sul problema elastostatico di Signorini con ambigue condizionial contorno, *Atti Accad Naz Lincei VIII Ser Rend Cl Sci Fis Mat Nat.,* **34** (1963) 138-142.

[27] G. Fichera, Problemi elastostatici convincoli unilaterali: il problema di Signorini con ambigue condizionial contorno, *Atti Accad Naz Lincei Mem Cl Sci Fis Mat Nat Sez I VIII Ser,* **7** 91–140.

[28] K. Fan, A minimax inequality and applications, Inequalities III, *Academic Press, New York,* (1972)

[29] S. D. Flam, A. S. Antipin, Equilibrium programming and proximal-like algorithms, *Math. Program,* **78** (1997) 29–41.

[30] S. P. Frankel, Convergence Rates of Iterative Treatments of Partial Differential Equations, *AMS,* **4**(30) (1950) 65-75 doi:10.2307/2002770.

[31] F. Giannessi, A. Maugeri, P. M. Pardalos, Equilibruim problems: nonsmooth optimization and variational inequality models, *Dordrecht: Kluwer Academic,* (2004).

[32] A. A. Goldstein, Convex programming in Hilbert space, *Bull. Am. Math. Soc.,* **70** (1964) 709-710.

[33] D. V. Hieu, An extension of hybrid method without extrapolation step to equilibrium problems, *Manag. Optim.,* **13** (2017) 1723–1741.

[34] D. V. Hieu, An inertial-like proximal algorithm for equilibrium problems, *Math Meth Oper Res.,* **88** (2018) 399–415. https://doi.org/10.1007/s00186-018-0640-6

[35] D. V. Hieu, Halpern subgradient extragradient method extended to equilibrium problems, *Rev. R. Acad. Cienc. Exactas F´ı,s. Nat. Ser. A Math. RACSAM,* **111** (2017) 823–840.

[36] D. V. Hieu, New inertial algorithm for a class of equilibrium problems, *Numer Algor.* **80** (2018) 1413–1436 https://doi.org/10.1007/s11075-018-0532-0

[37] D. V. Hieu, New subgradient extragradient methods for common solutions to equilibrium problems, *Comput Optim Appl.,* **67** (2017) 571–594.

[38] D. V. Hieu, Y. J. Cho, Y. B. Xiao, Modified extragradient algorithms for solving equilibrium problems, *Optim.,* **69** (2018) 2003–2029. https://doi.org/10.1080/02331934.2018.1505886

[39] D. V. Hieu, Y. J. Cho, Y. B. Xiao, et al., Modified extragradient method for pseudomonotone variational inequalities in infinite dimensional hilbert spaces, *Vietnam J Math.,* (2020) https://doi.org/10.1007/s10013-020-00447-7

[40] D. V. Hieu, L. Muu, P. K. Anh, Parallel hybrid extragradient methods for pseudomonotone equilibrium problems and nonexpansive mappings, *Numer. Algor.,* **73** (2016) 197–217.

[41] D. V. Hieu, L. D. Muu, P. K. Quy, H. N. Duong, Regularization extragradient methods for equilibrium programming in Hilbert space, *Optim.,* (2021) https://doi.org/10.1080/02331934.2021.1873988

[42] D. V. Hieu, J. J. Strodiot, L. D. Muu, Strongly convergent algorithms by using new adaptive regularization parameter for equilibrium problems, *J. Comput. Appl. Math.,* **376** (2020) 112844 https://doi.org/:10.1016/j.cam.2020.112844

[43] A. N. Iusem, W. Sosa, On the proximal point method for equilibrium problem in Hilbert spaces, *Optim.,* **59** (2010) 1259–1274.

[44] C. Izuchukwu, S. Reich, Y. Shehu, Relaxed inertial methods for solving the split monotone variational inclusion problem beyond co-coerciveness, *Optim.,* (2021) https://doi.org/10.1080/02331934.2021.1981895

[45] C. Jaiboon, P. Kumam, A Hybrid Extragradient Viscosity Approximation Method for Solving Equilibrium Problems and Fixed Point Problems of Infinitely Many Nonexpansive Mappings. *Fixed Point Theory Appl.,* **1** (2009) 374815. doi:10.1155/2009/374815

[46] C. Jaiboon, P. Kumam, U. W. Humphries, Weak convergence theorem by an extragradient method for variational inequality, equilibrium and fixed point problems. *Bull. Malays. Math. Sci. Soc.,* **32**(2) (2009) 173–185.

[47] L. O. Jolaoso, M. Aphane, A self-adaptive inertial subgradient extragradient method for pseudomonotone equilibrium and common fixed point problems, *Fixed Point Theory Appl.,* (2020) https://doi.org/10.1186/s13663-020-00676-y.

[48] G. Kassay, S. Reich, S. Sabach, Iterative methods for solving systems of variational inequalities in reflexive Banach spaces, *SIAM J. Optim.,* **21** (2011) 1319–1344.

[49] H. Khatibzadeh, V. Mohebbi, Monotone and pseudo-monotone equilibrium problems in Hadamard spaces, *J. Aust. Math. Soc.,* **110** (2021) 220-242 doi:10.1017/S1446788719000041

[50] C. Khunpanuk, N. Pakkaranang, N. Pholasa, On solving pseudomonotone equilibrium problems via two new extragradient-type methods under convex constraints, *Demonstr. Math.,* **55** (2022) 297–314 https://doi.org/10.1515/dema-2022-0025

[51] D. Kinderlehrer, G. Stampcchia, An Introduction to Variational Inequalities and Their Applications, *Academic Press, New York,* (1980).

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

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

[54] P. Kumam, P. Katchang, A viscosity of extragradient approximation method for finding equilibrium problems and fixed point problems for nonexpansive mappings. *Nonlinear Anal. Hybrid Syst.,* **3** (2009) 475–486.

[55] W. Kumam, P. Kumam, Hybrid iterative scheme by relaxed extragradient method for solutions of equilibrium problems and a general system of variational inequalities with application to optimization, *Nonlinear Anal. Hybrid Syst.,* **3** (2009) 640–656.

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

[57] Z. Q. Luo, J. S. Pang, D. Ralph, Mathematical programs with equilibrium constraint, *Cambridge University press,* (1996).

[58] S. I. Lyashko, W. Semenov, Optimization and its applications in control and data sciences, *Cham: Springer,* (2016) p. 315–325.

[59] P. E. Mainge, Approximation methods for common fixed points of nonexpansive mappings in Hilbert spaces *J. Math. Anal. Appl.,* **325**(1) (2007) 469-479.

[60] G. Mastroeni, Gap functions for equilibrium problems, *J. Global Optim.,* **27** (2003) 411–426.

[61] G. Mastroeni, On auxiliary principle for equilibrium problems, *P. Daniele et al. (eds.), Equilibrium Problems and Variational Models, Kluwer Academic Publisher,* (2003) p. 289–298.

[62] A. Moudafi, On the convergence of splitting proximal methods for equilibrium problems in Hilbert spaces, *J. Math. Anal. Appl.,* **359** (2009) 508–513.

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

[64] A. Moudafi, Second-order differential proximal methods for equilibrium problems, *J. Inequal. Pure Appl. Math.,* **4** (2003) Article 18.

[65] L. D. Muu, W. Oettli, Convergence of an adaptive penalty scheme for finding constrained equilibria, *Nonlin Anal TMA.,* **18** (1992) 1159–1166.

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

[67] N. Nadezhkina,W. Takahashi, Weak convergence theorem by extragradient method for nonexpansive mappings and monotone mappings, *J. Optim. Theory and Appl.,* **128** (2006) 191-201.

[68] Y. Nesterov, A method of solving a convex programming with convergence rate $o(1/k^2)$, *Sov. Math. Dokl.,* **27**(2) (1983) 372-376.

[69] T. V. Nguyen, L. D. Muu, Inertial extragradient algorithms for solving equilibrium problems, *Acta Math. Vietnam.,* (2019) https://doi.org/10.1007/s40306-019-00338-1

[70] T. T. V. Nguyen, J. J. Strodiot, V. H. Nguyen, Hybrid methods for solving simultaneously an equilibrium problem and countably many fixed point problems in a Hilbert space, *J Optim Theory Appl.,* **160** (2014) 809–831.

[71] M. A. Noor, S. Zainab, Y. Yoa, Implicit methods for equilibrium problems on Hadamard manifolds, *J. Appl. Math.,* (2012) doi:10.1155/2012/437391

[72] P. R. Oliveira, P. Santos, A. N. Silva, A tikhonov-type regularization for equilibrium problems in Hilbert spaces, *Math Anal Appl.,* **401** (2013) 336–342.

[73] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Am. Math. Soc.,* **73**, (1967) 591–597.

[74] B. T. Polyak, Some methods of speeding up the convergence of iterative methods, *Zh. Vychisl. Mat. Mat. Fiz.,* **4**(5) (1964) 791-803.

[75] T. D. Quoc, P. N. Anh, L. D. Muu, Dual extragradient algorithms extended to equilibrium problems, *J. Global Optim.,* **52** (2012) 139–159.

[76] T. D. Quoc, L. D. Muu, Iterative methods for solving monotone equilibrium problems via dual gap functions, *Comput. Optim. Appl.,* **51** (2012) 709–728.

[77] T. D. Quoc, L. D. Muu, V. H. Nguyen, Extragradient algorithms extended to equilibrium problems, *Optim.,* **57** (2008) 749–776.

[78] H. Rehman, P. Kumam I. K. Argyros, et al, Inertial extra-gradient method for solving a family of strongly pseudomonotone equilibrium problems in real Hilbert spaces with application in variational inequality problem, *Symmetry,* **12** (2020) 503.

[79] H. Rehman, P. Kumam, I. K. Argyos, M. Shutaywi, Z. Shah, Optimization based methods for solving the equilibrium problems with applications in variational inequality problems and solution of Nash equilibrium models, *Mathematics,* **8** (2020) 822.

[80] S. Reich, S. Sabach, Three strong convergence theorems regarding iterative methods for solving equilibrium problems in reflexive Banach spaces, *Contemp. Math.,* **568** (2012) 225–240.

[81] R. T. Rockafellar, Convex analysis, *Princeton (NJ): Princeton University Press,* (1970).

[82] S. Saejung, P. Yotkaew, Approximations of zeros of inverse strongly monotone operators in Banach spaces, *Nonlinear Anal.,* **75** (2012) 742-750.

[83] D. R. Sahu, S. M. Kang, V. Sagar, Approximation of common fixed points of a sequence of nearly nonexpansive mappings and solutions of variational inequality problems, *J. Appl. Math.,* (2012) https://doi.org/10.1155/2012/902437

[84] P. Santos, S. Scheimberg, An inexact subgradient algorithm for equilibrium problems, *Comput Appl Math.,* **30** (2011) 91–107.

[85] S. Scheimberg, P. S. M. Santos, A relaxed projection method for finite dimensional equilibrium problems, *Optim.,* **60** (2011) 1193–1208.

[86] Y. Shehu, C. Izuchukwu, J. C. Yao, X. Qin, Strongly convergent inertial exragradient type methods for equilibrium problems,*Appl Anal.,* (2021) https://doi.org/10.1080/00036811.2021.2021187

[87] J. J. Strodiot, P. T. Vuong, T. I. V. Nguyen, A class of shrinking projection extragradient methods for solving non-monotone equilibrium problems in Hilbert spaces, *J. Glob. Optim.,* **64** (2016) 159–178

[88] B. Tan, S. Y. Cho, J. C. Yao, Accelerated inertial subgradient extragradient algorithms with non-monotonic step sizes for equilibrium problems and fixed point problems, *J. Nonlinear Var. Anal.,* **6** (2022) No. 1, pp. 89-122 https://doi.org/10.23952/jnva.6.2022.1.06

[89] D. V. Thong, L. L. Liu, Q. L. Dong, L. V. Long, P. A. Tuan, Fast relaxed inertial Tseng's method-based algorithm for solving variational inequality and fixed point problems in Hilbert spaces,*J. Comput. Appl. Math.* **418** (2023), 114739

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

[91] N. T. T. Van, J. J. Strodiot, V. H. Nguyen, The interior proximal extragradient method for solving equilibrium problems, *J. Global Optim.,* **44** (2009) 175–192.

[92] N. T. Vinh, L. D. Muu, Inertial extragradient algorithms for solving equilibrium problems, *Acta Math Vietnam.,* **44(3)** (2019) 639–663.

[93] P. T. Vuong, J. J. Strodiot, N. V. Hien, Extragradient methods and linesearch algorithms for solving Ky Fan inequalities and fixed point problems, *J Optim Theory Appl.,* **155** (2012) 605–627.

[94] P. T. Vuong, J. J. Strodiot, N. V. Hien, On extragradient-viscosity methods for solving equilibrium and fixed point problems in a Hilbert space, *Optim.,* **64** (2013) 429–451.

[95] S. I. Zukhovitskii, R. A. Polyak, M. E. Primak, Two methods of search for equilibrium points of n-person concave games, *Soviet Mathematics Doklady,* **10** (1969) 279-282.