Convergence Results for Inertial Regularized Bilevel Variational Inequality Problems

by

Kalu Okam Okorie

Dissertation submitted in fulfilment of the requirements for the degree
of
Master of Science

The University of the Witwatersrand

School of Mathematics
The University of the Witwatersrand, Johannesburg, South Africa.

January, 2024.

Convergence Results for Inertial Regularized Bilevel Variational Inequality Problems

by

Kalu Okam Okorie

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

Supervisor: Dr Chibueze Christian Okeke

. . . .

Co-supervisor: Dr Chinedu Izuchukwu

. . . .

Abstract

In this dissertation, we introduce and study the inertial forward-reflected-backward method for approximating a solution of bilevel variational inequality problems. Our proposed method involves a single projection onto a feasible set, one functional evaluation and adopts the inertial extrapolation term. These features make our algorithm cost-effective and efficient, which is desirable when the cost operator and the feasible set have a complex structure. We incorporate the regularization technique in our method and establish that the sequences generated by our method converge strongly to a solution of the bilevel variational inequality problem studied in this work; furthermore, we modified our method by replacing the stepsizes and projection onto a feasible set with a self-adaptive non-monotonic stepsizes and projection onto a constructive halfspace, respectively. The non-monotonic stepsizes ensure that our method performs without the previous detail of the Lipschitz constant, and the projection onto a constructive halfspace is cheap since its computation is through an explicit formula. These adjustments in our method ensure an improved performance, cheap computation and easy implementation of our method. We show the strong convergence result of the iterative sequences. Lastly, we give numerical experiments comparing the performance of the proposed methods with existing methods.

Acknowledgements

My sincere gratitude goes to Almighty God, who, in his infinite mercy, made my journey to obtaining an M.Sc. Degree possible and successful. I thank Almighty God for life, good health, knowledge and understanding throughout the period of this program.

My sincere appreciation goes to my parents, Mr Okorie Kalu and Mrs Mercy Okorie, for their love, care, sound advice, words of encouragement, prayers and emotional support. I thank my siblings, Valentine, Thomas, Amarachi, Chukwuebuka, Chigozie and Oluchukwu, for their love, care, words of encouragement and emotional support. I would not forget to thank my good friend, Chidi Nwakpa, for words of encouragement and emotional support.

Special thanks to my supervisors, Dr C. C. Okeke and C. Izuchukwu, for their brotherly love, financial support, words of encouragement and for always being available, even in their most challenging times, to guide, criticise, correct, direct and provide insightful materials; these acts of kindness made this demanding journey manageable. I am indebted to Dr C. C. Okeke, his wife, Mrs Ogechukwu Okeke, and Dr C. Izuchukwu for their overwhelming love and care and for going out of their way to ensure I am comfortable. Also, I must not forget to thank Professor Jerry Ezeora, Dr Theresa Efor and every staff of the department of Industrial mathematics of Ebonyi State University, Nigeria, for giving me a solid foundation during my undergraduate studies and being very supportive over the years.

Many thanks to my uncle, Chief Kenneth Agwu and his beautiful wife, Dr(Mrs) Prisca Kenneth, for their financial support, sound advice and words of encouragement. I am indebted to Mr Innocent Umoru for his financial support and assistance during the visa process. I must not forget to thank Mr Inyang Ekpenyong for his financial support and Mrs Ebere Ogudu for her motherly love, care and words of encouragement.

I acknowledge the University of the Witwatersrand, Johannesburg, for its financial support through the Postgraduate Merit Award (PMA). My heartfelt gratitude goes to Mr Shashilan Singh, the administrator of the School of Mathematics, the University of the Witwatersrand and every staff of this school for their support, hospitality and kindness.

Dedication

To Almighty God and my Guardian Angels: Mr Okorie Kalu and Mrs Mercy Okorie.

Contents

Declaration

This dissertation, titled Convergence Results for Inertial Regularized Bilevel Variational Inequality Problems, has not been presented to this or any other institution in fulfilment for the award of a degree. It portrays the author's original work and references the work of other authors anywhere it is used in this dissertation.

Kalu Okam Okorie

11th March, 2024.

Submitted papers from this dissertation.

1. Kalu Okam Okorie, Chinedu Izuchukwu, Chibueze Christian Okeke, Inertial Forward-Reflected-Backward Method for solving bilevel variational inequality problem, Computational and Applied Mathematics.

2. Kalu Okam Okorie, Chinedu Izuchukwu, Chibueze Christian Okeke, A Modified Forward-Reflected-Backward Algorithm for Solving Bilevel Variational inequality in a real Hilbert space, Networks and Spatial Economics.

Chapter 1

Introduction

1.1 Background of study

The Bilevel Program (shortly, BP) is a mathematical program that has received extensive attention in recent years. BP refers to hierarchical optimization problems consisting of two constraints: the upper-level and the lower-level problem. Here, the graph of the solution set mapping of the lower-level problem restricts the upper-level problem. Mathematically, we express the problem as follows:

$$\begin{cases} \text{Search } z^* \in \Psi \subset H \text{ that solves ULP,} \\ \text{where } \Psi \text{ is the solution set of the problem,} \\ \text{search } x^* \in C \subset H \text{ that solves LLP,} \end{cases}$$

where ULP, LLP and C represent the upper-level problem, lower-level problem and a nonempty closed convex subset of the (real) Hilbert space H, respectively.

The history of bilevel programming was first traced back to Von Stackelberg [84]. He discovered a hierarchical game of two players known as the Stackelberg game. However, Bracken and McGill [16] were the first to formulate problems in bilevel programming. Later, Candler and Norton [17] gave this area of mathematics its present name – "bilevel programming." Although bilevel programming started around the 1970s, it was not until the early 1980s that its usefulness in modelling hierarchical decision processes and engineering design problems led to an in-depth study of bilevel programs.

The applications of bilevel programming are diverse and extensive, it finds its use in transportation (network design, optimal pricing), engineering (optimal chemical equilibria, optimal design), economics (Stackelberg games, principal-agent problem, taxation, policy decisions), and management (network facility location, coordination of multi-divisional firms), see [26, 29, Chapters 6 and 7] and [68, Chapter 6]. The flexibility of the bilevel program makes it possible to obtain various models of inverse problems by choosing different inverse problems for the lower and upper levels problem, see [26, Sect. 7] . This approach leads to the bilevel variational inequality problem (BVIP), which results from choosing the lower and upper levels of problems in the BP as variational inequality problems.

Let C be a nonempty closed convex subset of H, where H is a (real) Hilbert space with inner product $\langle \cdot, \cdot \rangle$. Let $\mathcal{A} : H \to H$ and $\mathcal{T}$: H $\to$ H be two single-valued operators; the Bilevel Variational Inequality Problem (BVIP) is as follows:

$$\text{Search } z^* \in VI(C, \mathcal{A}) \ \text{ such that } \langle \mathcal{T} z^*, z - z^* \rangle \geq 0 \ \ \forall z \in VI(C, \mathcal{A}), \tag{1.1.1}$$

where VI(C,$\mathcal{A}$) represents the solution set of the Variational Inequality Problem (shortly, (VIP)) defined by:

$$\text{Search } x^* \in C \ \text{ such that } \langle \mathcal{A} x^*, y - x^* \rangle \geq 0 \ \ \forall\, y \in C. \tag{1.1.2}$$

Clearly, we see that from (1.1.1)-(1.1.2), BVIP incorporates VIP as constraints. Many areas of mathematics such as mathematical programs with equilibrium constraints [63], variational inequalities [25, 35, 58], bilevel minimization problems [82], bilevel convex programming models [95], minimum-norm problems with the solution set of variational inequalities [106, 108], and bilevel linear programming [37] are different forms of BVIP.

Many practical applications of BVIP (1.1.1)-(1.1.2) have motivated and attracted the attention of researchers who strive to construct fast and efficient algorithms for approximating the solutions of BVIP (1.1.1)-(1.1.2). As a result, many algorithms were introduced (see [5, 46, 73, 87, 91, 92, 93] and references therein). However, algorithms for approximating the solution of BVIP (1.1.1)-(1.1.2) with one functional evaluation of $\mathcal{A}$ and a single projection onto the feasible set are scarce.

It is clear that BVIP (1.1.1)-(1.1.2) is applicable in various fields; however, its imperfection stems from the fact that, sometimes, it can be ill-posed. Ill-posed BVIPs are sensitive to perturbations, leading to difficulty in approximating a reliable and stable solution. Several methods for solving such problems are available in the literature. Some of these methods are as follows: regularization methods, numerical algorithms, hybrid approaches, metaheuristic algorithms, problem reformulation, etc. The regularization method is considered a helpful technique for solving ill-posed problems. Recently, this technique has found application in approximating solutions of monotone VIPs [2]. The regularization technique works by adding to the monotone operator of the original VIP a strongly monotone operator depending on a regularization parameter to obtain a new strongly monotone VIP. The resulting regularized problem becomes relatively easy to solve, and its solution relies on the regularization parameter. The resulting solutions can converge asymptotically or finitely to some solution of the original VIP as the regularization parameter tends to an appropriate limit. However, it is evident in [69] that the Tikhonov regularization technique [27, 28, 69] approximates solutions that are far away from the optimal solution.

In this dissertation, we adopt the hybrid approach– a combination of a projection-type algorithm and a regularization technique. These iterative regularization-projection algorithms feature a single projection onto a feasible set or halfspace and one functional evaluation and obtain the approximate solution of BVIP (1.1.1)-(1.1.2) even when it is ill-posed.

1.2 Research motivation

Malitsky and Tam [67] introduced and studied the forward-reflected-backward splitting method for finding a zero of the sum of two monotone operators. Unique features of this method are its simple and elegant structure and the ability to execute a single projection onto a feasible set and one functional evaluation of $\mathcal{A}$; however, the iterative sequences generated by the algorithm in [67] converge weakly to the solution of the monotone inclusion problem studied in [67]. The authors also showed in [67] that the projected reflected gradient method [65] for VIs is a specific case of their proposed method. Furthermore, the authors proposed a variant of their proposed algorithm; however, this method uses a line search approach. Line-search procedures involve additional computations and can be time-consuming and expensive.

Recently, Izuchukwu et al.[50] incorporated inertial term and anchor term and replaced the line search criterion in [67, Algorithm 3.1.] with a self-adaptive non-monotonic stepsize and obtain the strong convergence of their algorithm used in finding a zero of the sum of two monotone mappings in a (real) Hilbert space H; we remark that the inertial term and the self-adaptive non-monotonic stepsize in [50, Algorithm 3.2.] accelerates the speed of convergence and improves the performance of this method, respectively. Also, very recently, Taiwo and Reich [86] modified [67, Algorithm 3.1.] by adding a regularization term, which ensures the strong convergence of their method used in the analysis for a common solution of monotone inclusion and variational inequality problems. These versions of [67, Algorithm 3.1.] can be used in the convergence analysis of VIPs when the backward operator is a normal cone to C. (see [67] and [65]).

Recently, Heiu and Moudafi [46] introduced a regularization projection method for solving BVIP (1.1.1)-(1.1.2). Also, Ogbuisi et al. [73] proposed a relaxed single projection method for solving BVIP (1.1.1)-(1.1.2). These methods are motivated by the modified subgradient extragradient and Tseng's extragradient methods, respectively, and they employ self-adaptive monotonic stepsizes. Since these stepsizes do not increase, it affects the performance of these algorithms. Many algorithms for solving BVIP (1.1.1)-(1.1.2) have been proposed motivated by different methods for solving VIPs; however, they either feature two functional evaluations of $\mathcal{A}$ or multiple projections, which affects the computational performance of these algorithms. To our knowledge, BVIP (1.1.1)-(1.1.2) has not been solved by a method inspired by the forward-reflected-backward method.

Motivated and inspired by the proliferation of studies towards this direction, we propose and study the iterative regularization-projection algorithms for solving BVIP (1.1.1)-(1.1.2), and these methods are motivated by the forward-reflected-backward method, the inertial method and the regularization technique. Our algorithms take advantage of the fine features of the methods in [46, 50, 65, 67, 73, 86].

1.3 Statement of problem

In this dissertation, we introduce and study two new methods for approximating a solution of BVIP (1.1.1)-(1.1.2) where $\mathcal{A}$ is monotone and Lipschitz continuous, $\mathcal{T}$ is strongly monotone and Lipschitz continuous and VI(C,$\mathcal{A}$) is nonempty. We study these methods in this order:

1.

Algorithm 1.3.1. *Inertial regularised forward-reflected-backward method with constant inertial parameter*

Step 0: $\vartheta \in [0,1)$, $\epsilon \in \left(0, \frac{1}{4}\right)$, *and choose sequence* $\{\sigma_n\}$ *in* $(0,1)$. *For arbitrary* x_0, $x_1 \in H$. *Set* $n := 1$.

Step 1: *Compute*

$$x_{n+1} = P_C(m_n - \delta_n \mathcal{A} x_n - k_n - \delta_n \sigma_n \mathcal{T} x_n) \quad \forall n \geq 1,$$

where

$$\begin{cases} m_n = x_n + \vartheta(x_n - x_{n-1}) \\ k_n = \delta_{n-1}(1 - \gamma \delta_n \sigma_n)(\mathcal{A} x_n - \mathcal{A} x_{n-1}) \end{cases}$$

and

$$\delta_n \in \left[\varpi, \min\left\{\frac{1-2\varpi}{2L_{\mathcal{A}}}, \frac{1}{\gamma}\right\}\right] \text{ for some } \varpi > \quad 0.$$

Step 2: *Set* $n := n+1$ *and go to* **Step 1.**

Variable inertial parameter version

Algorithm 1.3.2. *Inertial regularised forward-reflected-backward method with variable inertial parameter*

Step 0: $\epsilon \in \left(0, \frac{1}{4}\right)$, *and choose sequence* $\{\sigma_n\}$ *in* $(0,1)$. *For arbitrary* x_0, $x_1 \in H$. *Set* $n := 1$.

Step 1: *Compute*

$$x_{n+1} = P_C(m_n - \delta_n \mathcal{A} x_n - k_n - \delta_n \sigma_n \mathcal{T} x_n) \quad \forall n \geq 1,$$

where

$$\begin{cases} m_n = x_n + \vartheta_n(x_n - x_{n-1}) \\ k_n = \delta_{n-1}(1 - \gamma \delta_n \sigma_n)(\mathcal{A} x_n - \mathcal{A} x_{n-1}) \end{cases}$$

and

$$\delta_n \in \left[\varpi, \min\left\{\frac{1-2\varpi}{2L_{\mathcal{A}}}, \frac{1}{\gamma}\right\}\right] \text{ for some } \varpi > \quad 0.$$

Step 2: *Set $n := n+1$ and go to* **Step 1.**

---

2. We obtained our second method by replacing the stepsizes and projection onto a feasible set in Algorithm 1.3.2 by a self-adaptive non-monotonic stepsizes and projection onto a constructive halfspace, respectively.

Remark 1.3.3. *Our method involves a single projection onto a feasible set, one functional evaluation of $\mathcal{A}$ and adopts an inertial extrapolation term. These features ensure a high computational performance and outperform most algorithms in the literature. Moreover, we adjusted Algorithm 1.3.2 by replacing the stepsizes and projection onto a feasible set by a self-adaptive non-monotonic stepsize and projection onto a constructive halfspace, respectively. These adjustments further improve its performance and ensure a cheap and easy computation and, thus, more implementable and more efficient than existing ones in the literature.*

1.4 Objectives

The central objectives of this study are to:

(i) survey some fundamental results on BVIPs and VIPs in real Hilbert space,

(ii) propose and analyse methods for approximating a solution of BVIP (1.1.1)-(1.1.2),

(iii) prove the strong convergence of the iterative sequences generated by our methods,

(iv) illustrate, through numerical experiments, the performance of the proposed methods with other existing methods.

The remainder of this dissertation is in this order: Chapter Two captures the basic definitions, fundamental lemmas, concepts, and propositions instrumental in establishing our main results. We presented a detailed review of VIPs and BVIPs. Furthermore, we give significant examples to serve as a support for proper understanding of these ideas. In Chapter Three, we propose and study the regularized forward-reflected-backward method for solving BVIP (1.1.1)-(1.1.2). We show how to integrate the regularization term in the forward-reflected-backward method and establish the strong convergence result of the sequences iteratively generated by our method under moderate conditions. This chapter also includes the performance of our algorithms through numerical experiments. In Chapter Four, we modified the regularised forward-reflected-backward method for solving BVIP (1.1.1)-(1.1.2) presented in Chapter Three. This modification results from replacing the stepsizes and single projection onto a feasible set by self-adaptive non-monotonic stepsizes and a single projection onto a constructive halfspace, respectively. We show that the sequences generated by this algorithm converge strongly. Again, the performance of the obtained algorithm in comparison with other algorithms was illustrated through numerical experiments. Chapter Five captures our conclusion. Moreover, we highlight the contributions of our study to existing knowledge and suggest possible future research.

Chapter 2

Literature Review

In this chapter, we review some fundamental concepts that are relevant to this work. Also, we give basic definitions and propositions that are instrumental and serve as a bedrock in achieving our results.

2.1 Hilbert spaces

Hilbert spaces are Banach spaces with a norm derived from inner product. They have an extra feature which makes them unique from other Banach spaces. The introduction of Hilbert spaces is in connection to David Hilbert and Erhard Schmidt [81] during their thorough study of integral equations (between 1862 and 1943). However, John von Neumann gave it the name–Hilbert space. Their study revealed that two square-integrable real-valued functions $\mathcal{A}$ and $\mathcal{T}$ on an interval [a, b] have an inner product.

$$\langle \mathcal{A}, \mathcal{T} \rangle = \int_a^b \mathcal{A}(x)\mathcal{T}(x)dx$$

which shares similar properties of the Euclidean dot product. In fact, Hilbert spaces are the closest generalization of the Euclidean spaces to infinite dimensional spaces. Hilbert spaces are indispensable tools in quantum mechanics, probability theory, fluid mechanics, Fourier analysis (these include applications to signal processing and heat transfer), PDEs, etc. In this dissertation, we explore the fine properties of Hilbert spaces as a support in establishing our main results.

Definition 2.1.1. *An inner product on H, where H denotes a nonempty set, is a function* $\langle \cdot,\cdot \rangle$ *defined on* $H \times H$ *with values in* $K = \mathbb{R}$ *or* $\mathbb{C}$ *such that the following conditions are satisfied:*

I_1. $\langle aq_1 + bq_2, r \rangle = a\langle q_1, r \rangle + b\langle q_2, r \rangle \quad \forall\, q_1, q_2, r \in H$ *and* $a, b \in K$;

I_2. $\langle q, r \rangle = \overline{\langle r, q \rangle}$ *(the bar represents the complex conjugate);*

I_3. $\langle q, q \rangle \geq 0$ *and* $\langle q, q \rangle = 0 \iff q = 0$.

Remark 2.1.2. *(H, $\langle\cdot,\cdot\rangle$) is called a complex inner product space. We give the following observation:*

(i) combining I_1 and I_2, we have
$\langle q,\, cr_1 + dr_2\rangle = \overline{\langle cr_1 + dr_2, q\rangle} = \overline{c\langle r_1, q\rangle + d\langle r_2, q\rangle} = \bar{c}\langle q,\, r_1\rangle + \bar{d}\langle q,\, r_2\rangle$;

(ii) by induction, I_1 and I_2, we obtain the generalization formula:
$\left\langle \sum_i a_i q_i, \sum_j b_j r_j \right\rangle = \sum_i \sum_j a_i \overline{b_j} \langle q_i, r_j\rangle$;

(iii) $I_1 \implies \langle 0,\, 0\rangle = \langle 0r,\, 0\rangle = 0\langle r,\, 0\rangle = 0$.
Hence, I_1, I_2, I_3 are equivalent to I_1, I_2, and if I_3 is replaced by I_3' (where I_3' is I_3 but with $q \neq 0$) then $\langle q,\, q\rangle$ is positive. A function satisfying I_1, I_2, I_3' is a complex inner product on H.

Remark 2.1.3. *If $\langle r,\, q\rangle = \overline{\langle r, q\rangle}$, then ($H$, $\langle\cdot,\cdot\rangle$) is called a real inner product space. We give the following observation:*

(i) combining I_1 and I_2, we have
$\langle q,\, cr_1 + dr_2\rangle = \langle cr_1 + dr_2,\, q\rangle = c\langle r_1,\, q\rangle + d\langle r_2,\, q\rangle = c\langle q,\, r_1\rangle + d\langle q,\, r_2\rangle$;

(ii) by induction, I_1 and I_2, we obtain the generalization formula:
$\left\langle \sum_i a_i q_i, \sum_j b_j r_j \right\rangle = \sum_i \sum_j a_i b_j \langle q_i, r_j\rangle$;

(iii) for real inner product space, a similar property as Remark 2.1.2(iii) is obtainable.

Remark 2.1.4. *A direct consequence of I_3 is that the square root of $\langle q,\, q\rangle$ exists since $\langle q,\, q\rangle$ is nonnegative. We denote it as*

$$\|q\| = \sqrt{\langle q, q\rangle}.$$

$\|q\|$ is called the norm or length of q.

Definition 2.1.5. *An inner product space denoted as (H, $\langle\cdot,\cdot\rangle$) is called complete if every Cauchy sequence in H converges to a point in H. A Hilbert space is a complete inner product space.*

Proposition 2.1.6. *(Cauchy Schwartz inequality) Let an inner product space be denoted as (H, $\langle\cdot,\cdot\rangle$). For q, $r \in H$*

$$|\langle q, r\rangle|^2 \leq \langle q, q\rangle\langle r, r\rangle \quad \text{or} \quad |\langle q, r\rangle| \leq \|q\|\|r\|.$$

Remark 2.1.7. *For $q \neq 0$, $r \neq 0$, $\langle q,\, r\rangle = \|q\|\|r\| \iff q = \beta r$ or $r = \beta q$ for some $\beta \in \mathbb{R}_+$.*

2.1.1 Examples of Hilbert spaces

(i) Let H = $\mathbb{R}^n$ be a space endowed with the inner product defined by

$$\langle q, r\rangle = q_1 r_1 + q_1 r_1 + \cdots + q_n r_n = \sum_{i=1}^{n} q_i r_i$$

where

$$q = (q_1, q_2, \cdots, q_n) \quad \text{and} \quad r = (r_1, r_2, \cdots, r_n) \quad \text{are in} \quad \mathbb{R}^n,$$

then $(H, \langle\cdot\rangle)$ is a (real) Hilbert space.

(ii) The spaces $W^{s,2}(\Omega)$ or $H^s(\Omega)$ (Sobolev spaces) are Hilbert spaces with inner product

$$\langle \mathcal{A}, \mathcal{T}\rangle = \int_{\Omega} \mathcal{A}(q)\overline{\mathcal{T}(q)}dq + \int_{\Omega} D\mathcal{A}(q) \cdot D\overline{\mathcal{T}(q)}dq + \cdots + \int_{\Omega} D^s\mathcal{A}(q) \cdot D^s\overline{\mathcal{T}(q)}dq$$

where s is a non-negative integer, $\Omega \subset \mathbb{R}^n$, and $\mathcal{A}$ and $\mathcal{T}$ are $L^2(\Omega)$ functions whose weak derivatives of order up to s are also $L^2(\Omega)$. The dot indicates the dot product in the Euclidean space of partial derivatives of each order.

(iii) The space C[a, b] of all complex-valued continuous function on a closed interval with inner product

$$\langle \mathcal{A}, \mathcal{T}\rangle = \int_a^b \overline{\mathcal{A}(q)}\mathcal{T}(q)dq$$

where $\mathcal{A}, \mathcal{T}$: [a, b] $\to \mathbb{C}$ are continuous is not a Hilbert space; consider for example, for the interval [-1, 1] the sequence of continuous "step" functions, $\{\mathcal{A}_m\}_m$, defined by

$$\mathcal{A}_m(q) = \begin{cases} 0 & q \in [-1, 0], \\ 1 & q \in \left[\frac{1}{m}, 1\right], \\ mq & q \in (0, \frac{1}{m}). \end{cases}$$

This sequence is a Cauchy sequence for the norm induced by the preceeding inner product which does not converge to a continuous function. However, the completion of C[a, b] with respect to the associated norm

$$\|\mathcal{A}\| = \left(\int_a^b |\mathcal{A}(q)|^2 dq\right)^{\frac{1}{2}}$$

is a Hilbert space and it is denoted by $L^2([a, b])$.

Remark 2.1.8. *The spaces $L^p([a, b])$ are Banach spaces but they are not Hilbert spaces when $p \neq 2$.*

2.1.2 Geometric properties of Hilbert spaces

Hilbert spaces have one of the finest geometric structures. As a result, it makes most problems posed in this space more manageable compared to other Banach spaces [24]. The availability of the inner product, the fact that the proximity map defined on a (real) Hilbert space H onto a closed convex subset C of H is Lipschitzian with constant one, and some exciting equalities and inequalities defined on H play a crucial role in the development of our main results. In what follows, we state some of these relations.

$$\|q+r\|^2 = \|q\|^2 + 2\langle q, r\rangle + \|r\|^2 \tag{2.1.1}$$

and

$$\|\beta q + (1-\beta)r\|^2 = \beta\|q\|^2 + (1-\beta)\|r\|^2 - \beta(1-\beta)\|q-r\|^2 \tag{2.1.2}$$

which hold for all q, $r \in$ H and $\beta \in (0,1)$. We remark that (2.1.1) can be written as

$$2\langle q, r\rangle = \|q\|^2 + \|r\|^2 - \|q+r\|^2 = \|q+r\|^2 - \|q\|^2 - \|r\|^2 \quad \forall\, q, r \in H \tag{2.1.3}$$

and

$$2\|q\|^2 + 2\|r\|^2 = \|q+r\|^2 + \|q-r\|^2, \tag{2.1.4}$$

where (2.1.4) is the well-known parallelogram identity.

2.1.3 Basic Identities and Inequalities in Hilbert spaces

Lemma 2.1.9 ([15]). *Let q, r, and s be in real Hilbert space H. Then we have the following:*

(i)
$$\|q \pm r\|^2 = \|q\|^2 \pm 2\langle q, r\rangle + \|r\|^2 \tag{2.1.5}$$

(ii)
$$2\|q\|^2 + 2\|r\|^2 = \|q+r\|^2 + \|q-r\|^2, \tag{2.1.6}$$

(iii)
$$4\langle q, r\rangle = \|q+r\|^2 - \|q+r\|^2 \tag{2.1.7}$$

(iv)
$$\|q-r\|^2 = 2\|s-q\|^2 + 2\|s-r\|^2 - 4\left\|s - \left(\frac{q+r}{2}\right)\right\|^2 \tag{2.1.8}$$

(v)
$$\|q+r\|^2 \le \|q\|^2 + 2\langle r, q+r\rangle \tag{2.1.9}$$

Proof. (i)

$$\begin{aligned}
\|q \pm r\|^2 &= \langle q \pm r, q \pm r\rangle \\
&= \langle q, q\rangle \pm 2\langle q, r\rangle + \langle r, r\rangle \\
&= \|q\|^2 \pm 2\langle q, r\rangle + \|r\|^2.
\end{aligned} \tag{2.1.10}$$

(ii) From (2.1.5), we have

$$\|q+r\|^2 = \|q\|^2 + 2\langle q, r\rangle + \|r\|^2 \tag{2.1.11}$$

and

$$\|q-r\|^2 = \|q\|^2 - 2\langle q, r\rangle + \|r\|^2 \tag{2.1.12}$$

addition of (2.1.11) and (2.1.12) gives (2.1.6).

(iii) Subtraction of (2.1.12) from (2.1.11) gives (2.1.7).

(iv) Applying (2.1.6) to these points $\frac{s-q}{2}$ and $\frac{s-r}{2}$, we have

$$\begin{aligned} 2\left\|\frac{s-q}{2}\right\|^2 + 2\left\|\frac{s-r}{2}\right\|^2 &= \left\|\frac{s-q}{2} + \frac{s-r}{2}\right\|^2 + \left\|\frac{s-q}{2} - \frac{s-r}{2}\right\|^2 \\ \Longrightarrow 4\left\|s - \left(\frac{q+r}{2}\right)\right\|^2 &= 2\left\|s-q\right\|^2 + 2\left\|s-r\right\|^2 - \left\|q-r\right\|^2 \end{aligned}$$

(v)

$$\begin{aligned} \|q+r\|^2 &= \langle q+r, q+r\rangle \\ &= \langle q, q\rangle + 2\langle q, r\rangle + \langle r, r\rangle \\ &\le \langle q, q\rangle + 2\langle q, r\rangle + \langle r, r\rangle + \langle r, r\rangle \\ &= \|q\|^2 + 2\langle r, q+r\rangle. \end{aligned}$$

□

Lemma 2.1.10. *Let q, $r \in H$ and $\beta \in [0, 1]$, we have*

$$\|\beta q + (1-\beta)r\|^2 = \beta\|q\|^2 + (1-\beta)\|r\|^2 - \beta(1-\beta)\|q-r\|^2$$

Proof.

$$\begin{aligned} \|\beta q + (1-\beta)r\|^2 &= \langle \beta q + (1-\beta)r, \beta q + (1-\beta)r\rangle \\ &= \beta^2\langle q, q\rangle + \beta(1-\beta)\langle q, r\rangle + \beta(1-\beta)\langle r, q\rangle + (1-\beta)^2\langle r, r\rangle \\ &= \beta^2\langle q, q\rangle + 2\beta(1-\beta)\langle q, r\rangle + (1-\beta)^2\langle r, r\rangle \qquad (2.1.13) \end{aligned}$$

applying (2.1.11) to (2.1.13) gives

$$\|\beta q + (1-\beta)r\|^2 = \beta\|q\|^2 + (1-\beta)\|r\|^2 - \beta(1-\beta)\|q-r\|^2.$$

□

Lemma 2.1.11 ([57])**.** *For s, t, q, $r \in H$, we have*

$$2\langle s-t, q-r\rangle = \|s-r\|^2 + \|t-q\|^2 - \|s-q\|^2 - \|t-r\|^2.$$

We also have

$$\begin{aligned} \|s-t+q-r\|^2 &= \|s-t\|^2 + \|q-r\|^2 + 2\langle s-t, q-r\rangle \\ &= \|s-t\|^2 + \|q-r\|^2 + \|s-r\|^2 + \|t-q\|^2 - \|s-q\|^2 - \|t-r\|^2. \end{aligned}$$

2.1.4 Some nonlinear single-valued mappings in Hilbert spaces

Definition 2.1.12. *An operator $\mathcal{A} : H \to H$ is called*

- *$L_{\mathcal{A}}$-Lipschitzian if there exists $L_{\mathcal{A}} > 0$ such that*

$$||\mathcal{A}t - \mathcal{A}q|| \leq L_{\mathcal{A}}||t - q|| \ \forall t, q \in H;$$

  if $L_{\mathcal{A}} = 1$, then $\mathcal{A}$ is said to be nonexpansive while $\mathcal{A}$ is referred to as a contraction if $L_{\mathcal{A}} \in$ (0, 1),

- *γ-strongly monotone, if there exists $\gamma > 0$ such that*

$$\langle \mathcal{A}t - \mathcal{A}q, t - q\rangle \geq \gamma||t - q||^2 \ \forall t, q \in H;$$

- *γ-inverse strongly monotone (γ-cocoercive), if there exists $\gamma > 0$ such that*

$$\langle \mathcal{A}t - \mathcal{A}q, t - q\rangle \geq \gamma||\mathcal{A}t - \mathcal{A}q||^2 \ \forall t, q \in H;$$

  if $\gamma = 1$, then $\mathcal{A}$ is firmly nonexpansive,

- *monotone, if*

$$\langle \mathcal{A}t - \mathcal{A}q, t - q\rangle \geq 0 \ \forall t, q \in H;$$

- *pseudomonotone, if*

$$\langle \mathcal{A}t, q - t\rangle \geq 0 \implies \langle \mathcal{A}q, q - t\rangle \geq 0 \ \forall t, q \in H;$$

- *strictly pseudomonotone, if*

$$\langle \mathcal{A}t, q - t\rangle \geq 0 \implies \langle \mathcal{A}q, q - t\rangle > 0 \ \forall t, q \in H;$$

- *quasimonotone, if*

$$\langle \mathcal{A}t, q - t\rangle > 0 \implies \langle \mathcal{A}q, q - t\rangle \geq 0 \ \forall q, t \in H;$$

- *strictly quasimonotone, if it is monotone and for all t, $q \in H$, $t \neq q$, there exists some $r \in$ (t, q) such that*

$$\langle \mathcal{A}r, q - t\rangle \neq 0;$$

Remark 2.1.13. *We observe that the following implications are true:*

- *Strongly monotone operator $\implies$ monotone operator $\implies$ strictly pseudomonotone operator $\implies$ pseudomonotone operator $\implies$ quasimonotone operator,*
- *strictly pseudomonotone operator $\implies$ strictly quasimonotone operator $\implies$ quasimonotone operator.*

We note that the converse of these implications may not be true, see the following examples:

Example 2.1.14. *Let $H = \mathbb{R}$ be a Hilbert space with the usual metric and $\mathcal{A}$:(-∞, 1] $\rightarrow \mathbb{R}$ be defined by*

$$\mathcal{A}(q) = \begin{cases} -1 & q \in (-\infty, 0) \cup (0, 1), \\ 1, & q \geq 1, \\ 0, & q = 0. \end{cases}$$

$\mathcal{A}$ is a quasimonotone operator but not a pseudomonotone operator.

Example 2.1.15 ([54])**.** *Let $H = \mathbb{R}$ be a Hilbert space with the usual metric and $\mathcal{A}$:(0, ∞) $\rightarrow$ (0, ∞) be defined by*

$$\mathcal{A}(q) = \frac{1}{1+q}, \quad q \in (0, \infty).$$

$\mathcal{A}$ is a pseudomonotone operator but it is not a monotone operator.

Remark 2.1.16. *(i) It is well-known that any operator that is γ-inverse strongly monotone is monotone and $\frac{1}{\gamma}$-Lipschitz continuous.*

(ii) Any operator that is $L_{\mathcal{A}}$-Lipschitz continuous is $\frac{2}{L_{\mathcal{A}}}$-inverse strongly monotone operator.

(iii) If $\mathcal{A}$ is γ-strongly monotone and $L_{\mathcal{A}}$-Lipschitz continuous, then $\mathcal{A}$ is $\gamma/L_{\mathcal{A}}^2$ -inverse strongly monotone.

2.1.5 The metric projection on Hilbert spaces

Definition 2.1.17. *The operator $P_C : H \rightarrow C$ called the metric projection (or nearest point mapping) onto C where C is a nonempty, closed and convex subset of H, is a mapping which assigns to each $r \in H$ the unique point $P_C r$ in C such that*

$$||r - P_C r|| = \inf\{||r - q|| \ : \ q \in C\}.$$

Example 2.1.18 ([15])**.** *Let $q \in H$, let a_1 and $a_2 \in \mathbb{R}$ and set*

$$C = \{r \in H : a_1 \leq \langle r, q\rangle \leq a_2\}$$

where $q \neq 0$ and $a_1 \leq a_2$ then

$$P_C r = \begin{cases} r + \frac{a_1 - \langle r,q\rangle}{\|q\|^2} q, & if\ \langle r, q\rangle < a_1; \\ r, & if\ a_1 \leq \langle r, q\rangle \leq a_2; \\ r + \frac{a_2 - \langle r,q\rangle}{\|q\|^2} q, & if\ \langle r, q\rangle > a_2, \qquad \forall\, r \in H; \end{cases} ,$$

is the metric projection onto C.

Example 2.1.19 ([15]). *Let q be a nonzero vector in H and $a_1 \in H$, set $C = \{r \in H : \langle r, q\rangle = a_1\}$, then*

$$P_C r = r + \frac{a_1 - \langle r, q\rangle}{\|q\|^2} q \quad \forall\, r \in H,$$

is the metric projection onto C.

Example 2.1.20 ([15]). *Let q be a nonzero vector in H and $a_1 \in H$, set $C = \{r \in H : \langle r, q\rangle \le a_1\}$, then*

$$P_C r = \begin{cases} r, & \text{if } \langle r, q\rangle \le a_1, \\ r + \frac{a_1 - \langle r,q\rangle}{\|q\|^2} q, & \text{if } \langle r, q\rangle > a_1, \end{cases}$$

is the metric projection onto C.

Remark 2.1.21. *Observe that Example 2.1.19 is a special case of Example 2.1.20.*

Proposition 2.1.22 ([14]). *Let $r \in H$, then*

$$\langle r - P_C r, t - P_C r\rangle \le 0, \ \forall t \in C \tag{2.1.14}$$

is the characterization of the metric projection P_C.

The following results are consequences of Proposition 2.1.14:

(i) The metric projection is firmly nonexpansive, that is

$$||P_C r - P_C t||^2 \le \langle r - t, P_C r - P_C t\rangle \ \forall r, t \in H, \tag{2.1.15}$$

(ii)

$$||r - P_C r||^2 \le ||r - t||^2 - ||t - P_C r||^2 \quad \forall r \in H \text{ and } t \in C, \tag{2.1.16}$$

(iii)

$$\|P_C r - P_C t\| \le \|r - t\| \qquad \forall r, t \in H. \tag{2.1.17}$$

2.2 Variational inequality problems

VIP (1.1.2) *serves as a mathematical model which unifies several problems in science and engineering. It is evident from [11] that many problems can be converted to VIP* (1.1.2), *for example, problems arising from economics, mathematical programming, engineering mechanics, transportation, and other sciences. Stampacchia [85] (and also Fichera [33]) introduced the idea of VIP as an efficient way of modelling non-stationary problems emerging from structural analysis, pure and applied sciences, and other areas (for details, see [8, 18, 72] and references therein). Due to its vast applicability to other fields, VIP* (1.1.2) *has drawn the attention of many scholars. At this point, we give some examples of VIP and suggest two monographic books [31, 56] for broader familiarity with this subject.*

2.2.1 Examples of variational inequality problems

Example 2.2.1. *Consider a beam of length l with firmly fixed edges. Let $M(x)$ and $E_m(x)$ represent the moment of inertia of a section and an elastic modulus, respectively, and $P(x) = E_m(x)M(x)$, $x \in [0, l]$. Suppose the axis OX is along the beam's axis, and OY is up through the beam's left end. The beam bends as a result of the force produced by the fixed rigid body pressing down on it, with a smooth contact surface represented by the equation $y = \eta(x)$; points lying beyond this surface satisfy the inequality $y \leq \eta(x)$. The equilibrium state of the beam of this contact problem of deformable bodies with an ideally smooth boundary is modelled by a differential equation as follows:*

$$\frac{d^2}{dx^2}\left[P(x)\frac{d^2b(x)}{dx^2}\right] = -r(x), \tag{2.2.1}$$

where $r(x)$ and $b(x)$ are rigid body reaction and beam deformation, respectively, with the boundary conditions

$$b(0) = b(l) = b'(0) = b'(l) = 0. \tag{2.2.2}$$

Also, we have the following:

- *the one sided-restriction*

$$b(x) \leq \eta(x) \tag{2.2.3}$$

  which characterize the non-penetration of beam points into the rigid body;

- *the condition*

$$r(x) \geq 0 \tag{2.2.4}$$

  determining the direction of the rigid body reaction;

- *the equation*

$$(b(x) - \eta(x))r(x) = 0, \quad x \in (0, l), \tag{2.2.5}$$

which implies that either (2.2.3) or (2.2.4) turns into a strict equality at any point $x \in (0, l)$.

Introduce the space

$$V = \{\vartheta : \vartheta = \vartheta(x), x \in [0, l], \vartheta \in W_2^2[0, l], \vartheta(0) = \vartheta(l) = \vartheta'(0) = \vartheta'(l) = 0\}$$

and define

$$\|\vartheta\| = \left(\int_0^l |\vartheta''(x)|^2 dx\right)^{\frac{1}{2}}.$$

Next, we assume that any solution $b(x)$ of the given problem has derivatives up to the fourth order. By (2.2.1)

$$\int_0^l \frac{d^2}{dx^2}\left[P(x)\frac{d^2b(x)}{dx^2}\right]\vartheta(x)dx = -\int_0^l r(x)\vartheta(x)dx \;\; \forall\, \vartheta \in V,\; b \in V. \tag{2.2.6}$$

That the obtained equality is equivalent to (2.2.1) results from the basic lemma of the calculus of variations [34, Lemma 1, p.9]. Apply twice the formula of integration by parts to the left-hand side of (2.2.6). Taking into account the boundary conditions (2.2.2) we deduce

$$\int_0^l P(x)b^{''}(x)\vartheta^{''}(x)dx = -\int_0^l r(x)\vartheta(x)dx \;\; \forall\vartheta \in V,\; b \in V. \tag{2.2.7}$$

The function b(x) satisfying (2.2.7) is called a weak solution of the problem (2.2.1)-(2.2.5). Suppose that in (2.2.7) $\vartheta(x) = \tau(x) - b(x)$*, where* $\tau \in V$*, and analyze the sign of the function*

$$S(x) = r(x)(\tau(x) - b(x)).$$

If there is no contact between the beam and rigid body at a point x ∈ (0, l), then r(x)= 0 and S(x) = 0. If there is contact at that point then we assert that $\tau(x) \le b(x)$ *and, hence, S(x) ≤ 0. Consequently, solution b(x) of the problem (2.2.1)-(2.2.5) satisfies the variational inequality*

$$\int_0^l P(x)b^{''}(x)(\tau^{''}(x) - b^{''}(x))dx \ge 0 \;\; \forall\vartheta \in \Omega,\; b \in \Omega. \tag{2.2.8}$$

where

$$\Omega = \{\vartheta : \vartheta = \vartheta(x) \in V, \vartheta(x) \le \eta(x) \textit{ for almost all } x \in [0, l]\}$$

is a convex closed subset in V.

Let b(x) now be a solution of (2.2.8). If there is no contact between the beam and rigid body at a point x ∈ (0, 1), then difference $\tau(x) - b(x)$ *may take both positive and negative values. Integrating the left-hand side of (2.2.8) by parts twice and taking into account again the boundary conditions (2.2.2), we come to the inequality*

$$\int_0^l \frac{d^2}{dx^2}\left[P(x)\frac{d^2b(x)}{dx^2}\right](\tau(x) - b(x))dx \ge 0 \;\; \forall\tau \in \Omega, \;\; b \in \Omega. \tag{2.2.9}$$

Then by repeating the reasoning of the proof of the basic lemma of the calculus of variations, we establish that

$$\frac{d^2}{dx^2}\left[P(x)\frac{d^2b(x)}{dx^2}\right] = 0. \tag{2.2.10}$$

Let x ∈ (0, l) be a contact point of the beam and rigid body. Then $\tau(x) - b(x) \le 0$*. Using (2.2.9) and the proof by contradiction, we deduce the inequality*

$$\frac{d^2}{dx^2}\left[P(x)\frac{d^2b(x)}{dx^2}\right] = -r(x) \le 0. \tag{2.2.11}$$

By virtue of (2.2.10) *and* (2.2.11), *b(x) satisfies the equation* (2.2.1) *with condition* (2.2.4). *Thus, the equivalence of the problem* (2.2.1)-(2.2.5) *and variational inequality* (2.2.8) *has been established.*

We introduce the operator $\mathcal{A} : V \to V^*$ *by the equality*

$$\langle \mathcal{A}b, \vartheta \rangle = \int_0^l P(x) b^{''}(x) \vartheta^{''}(x) dx \quad \forall \vartheta, b \in V. \tag{2.2.12}$$

One can confirm that $\mathcal{A}$ *is a continuous, monotone, linear and potential map [32, 60, 76]. Using the natural assumption: There exists a constant* $K > 0$ *such that*

$$|P(x)| \leq K \quad \forall x \in [0, l].$$

The boundedness of $\mathcal{A}$ *has been established.* (2.2.8) *can be rewritten, by the definition of* $\mathcal{A}$*, as follows:*

$$\langle \mathcal{A}b, \tau - b \rangle \geq 0 \ \ \forall \tau \in \Omega, \ \ b \in \Omega \quad (\text{ see } [2]).$$

Example 2.2.2. *Let the operator* $\mathcal{A}$ *and the feasible set* C *be given as follows:*

$$\mathcal{A}(t) = \begin{bmatrix} 3t_1 - \frac{1}{t_1} + 3t_2 - 2 \\ 3t_1 + 3t_2 \\ 4t_3 + 4t_4 \\ 4t_3 + 4t_4 - \frac{1}{t_4} - 3 \end{bmatrix} \quad \text{and}$$

$C = \{t \in \mathbb{R}^n \mid t_1 + t_2 = 1, \ t_3 + t_4 \geq 0, \ r \leq t \leq q\}$, *where* $q = (10, 10, 10, 10)^T$ *and* $r = (0.1, 0, 0, 1)^T$. *Then the nonlinear VIP above has a unique solution* $t^* = (1, 0, 0, 1)^T$ *(see [107]).*

Several authors have been motivated by the vast application of VIP (1.1.2) *to obtain not only theoretical results on the existence and the stability of solutions of the VIP* (1.1.2) *but also to construct efficient algorithms for approximating the solution of this problem. As a result, many methods for the solution of VIP* (1.1.2) *have been proposed, such as the projection-type methods and regularization methods. Following that, we review some of the previous works in these areas:*

2.2.2 Projection-type methods

VIP (1.1.2) *corresponds to the fixed point problem of searching for a point*

$$x^* \in C \quad \text{such that} \quad x^* = P_C(x^* - \delta \mathcal{A} x^*), \tag{2.2.13}$$

where $\delta > 0$ *and* P_C *is the metric projection from* H *onto* C*. Recently, some authors have introduced and examined ways to solve the VIP* (1.1.2) *through this approach. Gradient Projection Method (GPM) is the easiest and oldest algorithm for solving VIP* (1.1.2), *and it reads:*

$$x_{n+1} = P_C(x_n - \delta \mathcal{A} x_n), \quad \forall n \geq 1. \tag{2.2.14}$$

This method involves only one functional evaluation of $\mathcal{A}$ and a single projection of elements of H onto set C. The sequence generated by (2.2.14) is convergent to elements of VI(C,$\mathcal{A}$) when $\mathcal{A}$ is $L_{\mathcal{A}}$-Lipschitz continuous, γ-strongly monotone and $\delta \in (0, \frac{2\gamma}{L_{\mathcal{A}}})$. As noted in [42], relaxation of the strong monotonicity assumption to monotonicity may cause the sequences generated by (2.2.14) to diverge. Korpelevich in [59] introduced the extragradient method (EGM) in his attempt to address this shortcoming. The algorithm reads:

$$y_n = P_C(x_n - \delta \mathcal{A} x_n), \qquad x_{n+1} = P_C(x_n - \delta \mathcal{A} y_n), \quad \forall n \geq 1. \tag{2.2.15}$$

The algorithm is convergent to elements of VI(C, $\mathcal{A}$) when $\mathcal{A}$ is monotone and $L_{\mathcal{A}}$-Lipschitz continuous, $\delta \in (0, \frac{1}{L_{\mathcal{A}}})$, and $L_{\mathcal{A}} > 0$ provided that VI(C,$\mathcal{A}$) is nonempty. The relaxation to the monotonicity assumption was a remarkable improvement and has attracted the attention of researchers who improved it in many ways (see [22, 49, 55, 83]). EGM is deficient since it requires the computation of two functional evaluations of $\mathcal{A}$ and two projections onto feasible set C per iteration. As a result, it is costly when the mapping $\mathcal{A}$ and feasible set C have a complex feature. Popov's method [75] partly addressed this challenge. His modification of (2.2.15) results in a single evaluation of $\mathcal{A}$ under a similar assumption. Popov's method is as follows:

$$\begin{cases} x_{n+1} = P_C(x_n - \delta \mathcal{A} y_n), \\ y_{n+1} = P_C(x_{n+1} - \delta \mathcal{A} y_n), \ \ n \geq 1, \end{cases} \tag{2.2.16}$$

where $\mathcal{A}$ is $L_{\mathcal{A}}$-Lipschitz continuous and monotone, $\delta \in (0, \frac{1}{3L_{\mathcal{A}}})$, and $L_{\mathcal{A}} > 0$. We note that the first method to overcomes the challenge of double projection onto C is the projection and contraction method introduce by He [42]. We state the algorithm as follows:

$$\begin{cases} y_n = P_C(x_n - \delta \mathcal{A} x_n), \\ d(x_n, y_n) = (x_n - y_n) - \gamma(\mathcal{A} x_n - \mathcal{A} y_n), \\ x_{n+1} = x_n - \alpha \beta_n d(x_n, y_n) \mathcal{A} y_n), \ \ n \geq 1, \\ \ \textit{where } \alpha \in (0, 2) \ \textit{where } \beta_n = \frac{\langle x_n - y_n, d(x_n, y_n)\rangle}{\|d(x_n, y_n)\|^2}, \end{cases} \tag{2.2.17}$$

where $\mathcal{A}$ is $L_{\mathcal{A}}$-Lipschitz continuous and monotone, $\gamma \in (0, \frac{1}{L_{\mathcal{A}}})$ and $L_{\mathcal{A}} > 0$. Clearly, this method involves only a single projection but feaures double evalutions of $\mathcal{A}$. The subgradient extragradient method (SEGM) [19] and the forward-backward-forward method (FBFM) [96] were also proposed to address the shortcomings of (2.2.15). SEGM is as follows:

$$\begin{cases} y_n = P_C(x_n - \delta \mathcal{A} x_n), \\ T_n := \{w \in H : \langle x_n - \delta \mathcal{A} x_n - y_n, w - y_n\rangle \leq 0\}, \\ x_{n+1} = P_{T_n}(x_n - \delta \mathcal{A} y_n), \ \ n \geq 1, \end{cases} \tag{2.2.18}$$

where $\mathcal{A}$ is $L_{\mathcal{A}}$-Lipschitz continuous and monotone, $\delta \in (0, \frac{1}{L_{\mathcal{A}}})$, and $L_{\mathcal{A}} > 0$. SEGM is more beneficial over EGM because the second projection onto the closed convex set C in EGM was replaced by a constructive half-space, T_n, and requires an explicit formula to compute this projection. However, it features two functional evaluations of $\mathcal{A}$ and provides weak convergence in this setting. In the same vein, FBFM is the following algorithm:

$$\begin{cases} y_n = P_C(x_n - \delta \mathcal{A} x_n) \\ x_{n+1} = y_n + \delta(\mathcal{A} x_n - \mathcal{A} y_n), \ \ n \geq 1, \end{cases} \tag{2.2.19}$$

where $\mathcal{A}$ is $L_{\mathcal{A}}$-Lipschitz continuous and monotone, $\delta \in (0, \frac{1}{L_{\mathcal{A}}})$, and $L_{\mathcal{A}} > 0$. It is clear that algorithm (2.2.19) features two evaluations of $\mathcal{A}$ and one projection onto feasible set C in each step.

Motivated by the works of Popov [75] and Censor et al. [19], Malitsky and Semenov [66] proposed the following algorithm:

$$\begin{cases} T_n := \{w \in H : \langle x_n - \delta \mathcal{A} y_{n-1} - y_n, w - y_n \rangle \leq 0\}, \\ x_{n+1} = P_{T_n}(x_n - \delta \mathcal{A} y_n), \\ y_{n+1} = P_C(x_{n+1} - \delta \mathcal{A} y_n), \ n \geq 1, \end{cases} \tag{2.2.20}$$

where $\mathcal{A}$ is $L_{\mathcal{A}}$-Lipschitz continuous and monotone, $\delta \in (0, \frac{1}{3L_{\mathcal{A}}})$, and $L_{\mathcal{A}} > 0$. We observe that (2.2.20) involves only a single projection onto C (similar to (2.2.14), (2.2.18), (2.2.19)) and a single functional evaluation of $\mathcal{A}$ per iteration. The latter makes method (2.2.20) very impressive where computations of operator $\mathcal{A}$ are costly. A practical example of such a situation is in a huge-scale VIP or a VIP that arises from optimal control.

Very recently, Malitsky [65] proposed the Projected reflected gradient method:

$$x_{n+1} = P_C(x_n - \delta \mathcal{A}(2x_n - x_{n-1})), \tag{2.2.21}$$

where $\mathcal{A}$ is $L_{\mathcal{A}}$-Lipschitz continuous and monotone, $\delta \in (0, \frac{\sqrt{2}-1}{L_{\mathcal{A}}})$, and $L_{\mathcal{A}} > 0$. We note that Algorithm (2.2.21) is similar to (2.2.14); however, it computes the value of a gradient at the point that is a reflection of x_{n-1} in x_n.

Also, Malitsky and Tam [67] proposed the following forward-reflected-backward splitting method:

$$x_{n+1} = J_{\delta \mathcal{S}}(x_n - 2\delta \mathcal{A} x_n + \delta \mathcal{A} x_{n-1}), \tag{2.2.22}$$

where $\mathcal{A}$ is single-valued, monotone and $L_{\mathcal{A}}$-Lipschitz continuous, $\delta \in (0, \frac{1}{2L_{\mathcal{A}}})$, $L_{\mathcal{A}} > 0$ and $J_{\delta \mathcal{S}}$ is the resolvent of $\mathcal{S}$ where $\mathcal{S}$ is a set-valued mapping and also monotone. We note that if $\mathcal{S} = N_C$ is the normal cone to set C, (2.2.22) becomes

$$x_{n+1} = P_C(x_n - \delta \mathcal{A} x_n - \delta(\mathcal{A} x_n - \mathcal{A} x_{n-1})). \tag{2.2.23}$$

The algorithm in (2.2.23) has a lot of benefits over algorithms (2.2.14)-(2.2.19) by virtue that it requires to computes only a single projection of elements of H onto the feasible set C and one evaluation of $\mathcal{A}$ per step. We remark that algorithms (2.2.20), (2.2.21) and (2.2.23) have the same computational complexity for each iteration; moreover, algorithms (2.2.21) and (2.2.23) have a more simple and elegant structure. These algorithms have drawn the attention of some researchers who have extended and modified it to get strong and weak convergence.

2.2.3 Regularization methods

The advancement in science and technology of the last decades ushered in practical problems which are ill-posed by their nature. Finding the solution to such problems turns out essential; therefore, developing techniques for that purpose evolved into an active area of *research in the intersection of theoretical mathematics with the applied sciences. Addressing the challenges in solving these ill-posed problems, A. N. Tikhonov [94] initiated the regularization technique. As a motivation for regularization theory and, in particular, for convergence theory, however, the difficulty in obtaining a stable solution for ill-posed problems seems to be the most central issue. It was evident that to have any chance of computing a meaningful solution, the problem needs to be approximated by well-posed ones, usually a family parametrized by the regularization parameter. This idea birthed the regularization method. The regularization technique is a helpful way of solving ill-posed problems in convex optimization. Recently, several authors have applied this method in approximating the solution of monotone VIPs (see [2]). The regularization technique works by adding to the monotone operator of the original VIP a strongly monotone operator depending on a regularization parameter to obtain a new strongly monotone VIP*

$$\text{find } x^* \in C \quad \text{such that } \langle \mathcal{A}x^* + \sigma\Phi x^*, x - x^*\rangle \geq 0 \quad \forall x \in C, \tag{2.2.24}$$

where Φ and σ are the operator and regularization parameter, respectively. We note that Φx^* can be taken as $\Phi x^* := \mathcal{T}x^*$ (upper-level operator) or $\Phi x^* := \|x^*\|^2$ (see [27, 28, 30, 69]). Under suitable assumptions, this perturbation then forces the lower-level problem to approximately generate a unique optimal solution as σ tends to the appropriate limit. This result enables the Tikhonov regularization technique to approximate the solution of bilevel monotone variational inequalities. The regularized subproblem has more useful properties than the original problem; however, getting a regularization solution for each parameter requires an additional iterative procedure. This technique can be time-consuming and computationally expensive, and thus, it may affect the performance of this method. Again, It is evident in [69], and to some extent in the other related references, that (2.2.24) *often leads to solutions that are far away from a real optimal solution. To address this problem, several authors have combined the regularization technique and the projection-type methods since the projection methods are valuable in approximating the solution of VIPs. This approach is known as the iterative regularization–projection method.*

Recently, Hieu et al. [47] proposed an iterative regularization–projection method for approximating a solution for a monotone VIP (1.1.2) *and proved the strong convergence of this method. Their method is a combination of Tseng's extragradient method and regularization technique. Their method is as follows: Choose the sequences $\{\lambda_n\}$ and $\{\alpha_n\}$ with the following properties:*

$$\text{(C0) } \lambda_n \in [a,b] \subset \left(0, \frac{1}{L_{\mathcal{A}}}\right), \text{ (C1) } \lim_{n\to\infty} \alpha_n = 0, \text{ (C2) } \sum_{n=1}^{\infty} \alpha_n = +\infty, \quad \text{(C3) } \lim_{n\to\infty} \frac{\alpha_n - \alpha_{n+1}}{\alpha_n^2}.$$

Algorithm 2.2.3. ***Initialization:*** *Choose $u_0 \in H$, two numbers $\lambda_0 \in (0, +\infty)$, $\mu \in (0, +\infty)$ and the sequences $\lambda_0 > 0$, $\mu \in (0, 1)$.*
Iterative Steps:

Step 1: *Compute:* $v_n = P_C(u_n - \lambda_n \mathcal{A}^{\mathcal{T}}_{\alpha_n} u_n)$,

Step 2: *Compute:* $u_{n+1} = v_n + \lambda_n(\mathcal{A}u_n - \mathcal{A}v_n)$.

Step 3: *Update:* λ*: if* $\lambda_n\|\mathcal{A}v_n - \mathcal{A}u_n\| \leq \mu\|v_n - u_n\|$ *then* $\lambda_{n+1} = \lambda_n$*, otherwise, set*

$$\lambda_{n+1} = \frac{\mu\|v_n - u_n\|}{\|\mathcal{A}v_n - \mathcal{A}u_n\|}.$$

Very recently, some authors have proposed iterative regularization-projection methods for approximating a solution for BVIP (1.1.1)-(1.1.2)*; for example, Hieu and Moudafi [46] introduced an algorithm based on iterative regularization-projection method. Their method combines the modified subgradient extragradient method and a regularization technique and establishes its strong convergence. Also, Ogbuisi et al. [73] proposed a relaxed single projection method and proved its strong convergence result. Their iterative algorithm combines Tseng's extragradient method and regularization technique.*

2.3 Bilevel Variational Inequality Problems

Bilevel variational inequality problems (BVIPs) are bilevel programs with variational inequality problems in the upper-level and lower-level constraints. They are unique classes of quasi-variational inequalities (see [7, 12, 31, 102]) and of equilibrium with equilibrium constraints (see [70]). BVIP can be well-posed or ill-posed; well-posed BVIPs give advantages to the upper-level player (leader); however, this type of BVIP does not often model real-life problems where the reaction of the lower-level player (follower) reflects their most profitable interest, whether it aligns with the interest of the leader or not. Moreover, if the lower-level player chooses options in favour of the leader, the model results in optimistic BVIPs. Otherwise, the resulting model is called the pessimistic BVIPs. BVIPs are ill-posed when the lower-level problem has multiple solutions, and a well-posed BVIP if it has a single solution.

BVIPs are powerful mathematical models applicable in bilevel models; in particular, BVIPs model problems emerging from optimization, economics, operations research and transportation. Consider, for example, a supplier who does supply his best goods to a business owner. Since both partners intend to excel in business, the supplier gives goods still in the best state to the business owner who also desires to excel in business. Their transactions are successful if they maximize their profit or minimize their loss. This is an example of supply chain management, and it represents a bilevel model where the supplier and the business owner are the lower-level and the upper-level decision makers, respectively. Some of these practical applications of this model have attracted significant attention from researchers.

2.3.1 Bilevel variational inequality problems and related problems

We mentioned earlier that bilevel variational inequality problems cover a class of mathematical models. In what follows, we show its relationship with some of these models.

Let Γ: $H \to \mathbb{R}$ be a convex and differentiable operator on VI(C, $\mathcal{A}$), then z^* is a solution to

$$\min \{\Gamma z : z \in VI(C, \mathcal{A})\}$$

if and only if z^* is a solution to the bilevel variational inequality VI($\nabla\Gamma$, VI(C, $\mathcal{A}$)), where $\nabla\Gamma$ is a gradient of Γ. Then we write BVIP (1.1.1)-(1.1.2) *in the form of mathematical programs with equilibrium constraints [6] as follows:*

$$\begin{cases} \min \Gamma z, \\ z \in \{x^* : \langle \mathcal{A}x^*, m - x^* \rangle \geq 0, \forall\, m \ \in C\}. \end{cases} \tag{2.3.1}$$

If Γ, ξ are two convex and differentiable functions and $\mathcal{T} := \nabla\Gamma$, $\mathcal{A} := \nabla\xi$, then BVIP (1.1.1)-(1.1.2) *becomes the following minimization problem (see [82]):*

$$\begin{cases} \min \Gamma z, \\ z \in \arg\min\{\xi z : z \in C\}. \end{cases} \tag{2.3.2}$$

In a special case $\mathcal{T}z = z \ \forall z \in C$, BVIP (1.1.1)-(1.1.2) *becomes the minimum-norm problems of the solution set of variational inequalities as follows:*

$$\textit{Find } z^* \in C \textit{ such that } z^* = P_{VI(C,\mathcal{A})}(0), \tag{2.3.3}$$

where $P_{VI(C,\mathcal{A})}(0)$ is the projection of 0 onto $VI(C, \mathcal{A})$. The least-squares solution to the constrained linear inverse problem in [79] is a good example of (2.3.3).

2.3.2 Examples of bilevel variational inequality problems

Example 2.3.1. *Suppose that $H = L^2([0, 1])$ with the inner product*

$$\langle x, y \rangle = \int_0^1 x(t)y(t)dt, \ \ \forall\, x, y \in H,$$

and the induced norm

$$\|x\| = \left(\int_0^1 |x(t)|^2 dt\right)^{\frac{1}{2}}.$$

Let $\mathcal{T}x(t) = \frac{2x(t)}{3}$, $\mathcal{A}x(t) = \max\{0, x(t)\} = \frac{x(t)+|x(t)|}{2}$ and $C = \{x \in H : \|x\| \leq 1\}$. The solution set of this problem is a singleton set containing 0. Hence $x^*(t) = 0$ (see [93]).

Example 2.3.2. *Let $C = \mathbb{R}^2$,*

$$\mathcal{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \mathcal{T} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The solution set of the problem is VI(C, $\mathcal{A}$) = $\{0\}$. Hence, the solution set of the bilevel variational inequality problem is $\{0\}$ (see [93]).

Example 2.3.3. *Let $\mathbb{R}^m = \mathbb{R}^n = \mathbb{R} = (-\infty, +\infty)$, $C = [-2, -1]$ and $N(x) = [-2, x]$ for all $x \in C$, $r, q, x, y \in \mathbb{R}$.*
Let

$$\begin{aligned} \langle \mathcal{A}r, r - q\rangle &= \langle -(r+1)(2r+3)(3r+4)e^r, r - q\rangle \\ &= -(r+1)(2r+3)(3r+4)e^r(r-q) \end{aligned}$$

and

$$\begin{aligned} \langle \mathcal{T}(r,x), x - y\rangle &= \langle -(r+2)(r+1)xe^r, x - y\rangle \\ &= -(r+2)(r+1)xe^r(x-y). \end{aligned}$$

The BVIP is defined as follows: find $x \in C$ such that

$$-(r+2)(r+1)xe^r(x-y) \leq 0, \quad \forall y \in C \tag{2.3.4}$$

where r is a solution of the following variational inequality: find $r \in N(x)$ such that

$$-(r+1)(2r+3)(3r+4)e^r(r-q) \leq 0, \quad \forall q \in N(x). \tag{2.3.5}$$

A simple computation shows that the solution set of lower-level variational inequality is

$$M(x) = \begin{cases} \{-2\}, & if \;\; x \in [-2, -\frac{3}{2}), \\ \{-\frac{3}{2}\}, & if \;\; x \in [-\frac{3}{2}, -\frac{4}{3}), \\ \{-\frac{3}{2}, -\frac{4}{3}\}, & if \;\; x \in [-\frac{4}{3}, -1), \\ \{-2, -\frac{3}{2}, -\frac{4}{3}\}, & if \;\; x = -1. \end{cases}$$

Observe that M(x) is not a singleton set if $x \in [-\frac{4}{3}, -1)$ or $x = -1$, hence it is called an ill-posed BVIP (see [61]). Moreover, $\{(-2, x) : x \in [-2, -\frac{3}{2})$ or $x = -1\}$ is the solution set of BVIP (2.3.4)-(2.3.5).

Remark 2.3.4. *We note that if $N(x) \equiv C \;\; \forall x \in C$ then $\mathcal{T}(r,x) \equiv \mathcal{T}(x)$ and BVIP (2.3.4)-(2.3.5) reduces to quasi-bilevel variational inequality (1.1.1)-(1.1.2) [97].*

2.3.3 Recent works on bilevel variational inequality problems

Recently, Anh et al. [5] presented the extragradient method for solving problem BVIP (1.1.1)-(1.1.2) in the Euclidean space $\mathbb{R}^n$. The scheme is as follows:

Algorithm 2.3.5. ***Initialization:*** *Choose $u \in \mathbb{R}^n$, $x^0 \in C$, $k = 0$, $0 < \lambda < \frac{2\beta}{L_{\mathcal{T}}^2}$, positive sequences $\{\lambda_k\}$, $\{\delta_k\}$, $\{\alpha_k\}$, $\{\beta_k\}$, $\{\gamma_k\}$ and $\{\bar{\epsilon_k}\}$ such that*

$$\begin{cases} \lim\limits_{k\to\infty} \delta_k = 0, \; \lim\limits_{k\to\infty} \bar{\epsilon_k} = 0, \\ \alpha_k + \beta_k + \gamma_k = 1 \quad \forall k \geq 0, \quad \sum_{k=1}^{\infty} \alpha_k = \infty, \\ \lim\limits_{k\to\infty} \alpha_k = 0, \; \lim\limits_{k\to\infty} \beta_k = \eta \in (0, \frac{1}{2}], \; \lim\limits_{k\to\infty} \lambda_k = 0, \; \lambda_k \leq \frac{1}{\mathcal{L}_{\mathcal{A}}} \;\; \forall k \geq 0. \end{cases} \tag{2.3.6}$$

Step 1: *Compute* $y^k := P_C(x^k - \lambda_k \mathcal{A}x^k)$ *and* $z^k := P_C(x^k - \lambda_k \mathcal{A}y^k)$.

Step 2: *Inner loop* $j = 0, 1, \cdots$. *Compute*

$$\begin{cases} x^{k,0} := z^k - \lambda \mathcal{T} z^k, \\ y^{k,j} := P_C(x^{k,j} - \delta_j \mathcal{A}x^{k,j}), \\ x^{k,j+1} := \alpha_j x^{k,0} + \beta_j x^{k,j} + \gamma_j P_C(x^{k,j} - \delta_j \mathcal{A}y^{k,j}). \\ if \|x^{k,j+1} - P_{VI(\mathcal{A},C)}(x^{k,0})\| \leq \bar{\epsilon}_k \text{ then set } h^k := x^{k,j+1} \text{ and go to Step 3.} \\ \text{ Otherwise, increase } j \text{ by } 1 \text{ and repeat the inner loop Step } 2. \end{cases} \tag{2.3.7}$$

Step 3: *Set* $x^{k+1} := \alpha_k u + \beta_k x^k + \gamma_k h^k$. *Then increase* k *by* 1 *and go to Step* 1.

One feature of this algorithm is that it adopts the line search procedure. However, a method with a line search approach involves an inner loop running until some finite stopping conditions are satisfied for each iteration. Therefore, line search techniques require additional iterative procedure and can be time-consuming and computationally expensive. Also, it is clear that Algorithm (2.3.5) involves more than one projection onto C *and more than one functional evaluation of* $\mathcal{A}$*, which further impacts negatively on the performance of the algorithm.*

Thong and Hieu [91] proposed a modified subgradient extragradient algorithm for solving BVIP (1.1.1)-(1.1.2) in their attempt to address some of the shortcomings present in Algorithm (2.3.5). Their method reads: Choose a sequence $\{\alpha_n\}$ *in* $(0, 1)$ *with the following properties:*

$$\text{(C1)} \lim_{n\to\infty} \alpha_n = 0, \quad \text{(C2)} \sum_{n=1}^{\infty} \alpha_n = +\infty.$$

Algorithm 2.3.6. ***Initialization:*** *Given* $\lambda \in (0, \frac{1}{L_\mathcal{A}})$, $\mu \in (0, 1)$, $\alpha \in (0, 2)$ *and* $0 < \beta < \frac{2\gamma}{L_\mathcal{T}^2}$ *(*$L_\mathcal{A}$ *is the Lipschitz constant of* $\mathcal{A}$*,* γ *is the modulus of the strong monotonicity of* $\mathcal{T}$ *and* $L_\mathcal{T}$ *is the Lipschitz constant of* $\mathcal{T}$*). Let* $x_0 \in H$ *be arbitrary.*
Iterative steps*: Calculate* x_{n+1} *as follows*

Step 1: *Compute*

$$y_n = P_C(x_n - \lambda \mathcal{A}x_n).$$

Step 2: *Compute*

$$z_n = P_{T_n}(x_n - \alpha\eta_n\lambda d_n \mathcal{A}y_n),$$

where

$$T_n = \{x \in H : \langle x_n - \lambda \mathcal{A}x_n - y_n, x - y_n\rangle \leq 0\},$$
$$d_n := x_n - y_n - \lambda(\mathcal{A}x_n - \mathcal{A}y_n)$$

and

$$\eta_n = \begin{cases} \frac{\langle x_n - y_n, d_n\rangle}{\|d_n\|^2}, & if \;\; d_n \neq 0 \\ 0, & if \; d_n = 0. \end{cases}$$

Step 3: *Compute*

$$x_{n+1} = z_n - \alpha_n \beta \mathcal{T} z_n, \quad \forall n \geq 0.$$

Set $n = n+1$ *and go back to* **Step 1.**

Clearly, this algorithm involves two functional evaluations of $\mathcal{A}$*, one projection onto the feasible set* C *and the other projection onto a halfspace. The projection onto a halfspace is easy to compute since it has an explicit formula. We also note that this method extends Algorithm* (2.3.5) *from the Euclidean space* $\mathbb{R}^n$ *to the general real Hilbert space. Also, Thong et al.* [93] *proposed the projection and contraction method for solving BVIP* (1.1.1)-(1.1.2) *as follows: Choose a sequence* $\alpha_n \in (0,1)$ *with the following properties:*

$$\text{(C1)} \lim_{n\to\infty} \alpha_n = 0, \quad \text{(C2)} \sum_{n=1}^{\infty} \alpha_n = +\infty.$$

Algorithm 2.3.7. ***Initialization:*** *Let* $\lambda_0 > 0$*,* $\mu \in (0,1)$*,* $\alpha \in (0,2)$ *and* $0 < \beta < \frac{2\gamma}{L_{\mathcal{T}}^2}$ *(*$L_{\mathcal{A}}$ *is the Lipschitz constant of* $\mathcal{A}$*,* γ *is the modulus of the strong monotonicity of* $\mathcal{T}$ *and* $L_{\mathcal{T}}$ *is the Lipschitz constant of* $\mathcal{T}$*). Let* $x_0 \in H$ *be arbitrary.*
Iterative steps: *Calculate* x_{n+1} *as follows*

Step 1: *Compute*

$$y_n = P_C(x_n - \lambda_n \mathcal{A} x_n), \quad \forall n \geq 0.$$

Step 2: *Compute*

$$z_n = x_n - \alpha \eta_n d_n, \quad \forall n \geq 0,$$

where

$$d_n := x_n - y_n - \lambda_n(\mathcal{A}x_n - \mathcal{A}y_n) \quad \forall n \geq 0$$

and

$$\eta_n = \begin{cases} \frac{\langle x_n - y_n, d_n \rangle}{\|d_n\|^2}, & if \;\; d_n \neq 0 \\ 0, & if \;\; d_n = 0. \end{cases}$$

Step 3: *Compute*

$$x_{n+1} = z_n - \alpha_n \beta \mathcal{T} z_n, \quad \forall n \geq 0.$$

Update

$$\lambda_{n+1} = \begin{cases} \min\left\{ \frac{\mu \|x_n - y_n\|}{\|\mathcal{A}x_n - \mathcal{A}y_n\|}, \lambda_n \right\}, & \text{if } \mathcal{A}x_n \neq \mathcal{A}y_n \\ \lambda_n, & \text{otherwise.} \end{cases}$$

Set $n = n+1$ *and go back to Step 1.*

This method requires a single projection onto the feasible set, two functional evaluations of $\mathcal{A}$ and uses self-adaptive stepsizes. Recently, Hieu and Moudafi [46] introduced a regularization projection method for solving BVIP (1.1.1)-(1.1.2). *Their algorithm is as follows: Choose the sequence $\{\alpha_n\} \in (0, +\infty)$ with the following properties:*

$$(C1)\ \lim_{n\to\infty} \alpha_n = 0, \quad (C2)\ \sum_{n=1}^{\infty} \alpha_n = +\infty; \quad (C3)\ \lim_{n\to\infty} \frac{\alpha_n - \alpha_{n+1}}{\alpha_n^2} = 0.$$

Algorithm 2.3.8. ***Initialization:*** *x_0, $y_0 \in H$ and $\lambda_0 = \lambda_1 > 0$, $\mu \in (0, \sqrt{2} - 1)$. Compute $x_1 = P_C(x_0 - \lambda_0(\mathcal{A}y_0 + \alpha_0 \mathcal{T} x_0))$, $y_1 = P_C(x_1 - \lambda_1(\mathcal{A}y_0 + \alpha_1 \mathcal{T} x_1))$.*
Iterative steps: *For each $n \geq 1$:*

1 *Compute:* $x_{n+1} = P_{T_n}(x_n - \lambda_n(\mathcal{A}y_n + \alpha_n \mathcal{T} x_n))$,
where
$T_n = \{\ x \in H : \langle x_n - \lambda_n(\mathcal{A}y_{n-1} + \alpha_n \mathcal{T} x_n) - y_n, x - y_n\rangle \leq 0\}$.

2 *Update:* $\lambda_{n+1} = \min\left\{\frac{\mu\|y_n - y_{n-1}\|}{\|\mathcal{A}y_n - \mathcal{A}y_{n-1}\|}, \lambda_n\right\}$.

3 *Compute:* $y_{n+1} = P_C(x_{n+1} - \lambda_{n+1}(\mathcal{A}y_n + \alpha_{n+1} \mathcal{T} x_{n+1}))$.

Their method features one projection onto a feasible set C and the other projection onto a halfspace, one functional evaluation of $\mathcal{A}$ and self-adaptive stepsizes. Very recently, Ogbuisi et al. [73] proposed a relaxed single projection method for solving BVIP (1.1.1)-(1.1.2) *as follows: Choose the sequence $\{\alpha_n\} \in (0, +\infty)$ with the following properties:*

$$(C1)\ \lim_{n\to\infty} \alpha_n = 0, \quad (C2)\ \sum_{n=1}^{\infty} \alpha_n = +\infty, \quad (C3)\ \lim_{n\to\infty} \frac{\alpha_n - \alpha_{n+1}}{\alpha_n^2} = \lim_{n\to\infty} \frac{\alpha_n - \alpha_{n+1}}{\alpha_{n+1}^2} = 0.$$

Algorithm 2.3.9. ***Initialization:*** *Choose $x_0 \in H$ and $\lambda_0 > 0$, $\mu \in (0, 1)$, $\beta \in (0, 1]$.*
Iterative Steps: *For each $n \geq 1$:*

Step 1: *Compute:* $y_n = P_C(x_n - \lambda_n(\mathcal{A}x_n + \alpha_n \mathcal{T} x_n))$,

Step 2: $x_{n+1} = (1 - \beta)x_n + \beta y_n + \beta\lambda_n(\mathcal{A}x_n - \mathcal{A}y_n)$, *and update:*

$$\lambda_{n+1} = \begin{cases} \min\left\{\frac{\mu\|x_n - y_n\|}{\|\mathcal{A}x_n - \mathcal{A}y_n\|}, \lambda_n\right\}, & \text{if } \mathcal{A}x_n \neq \mathcal{A}y_n \\ \lambda_n, & \text{otherwise.} \end{cases}$$

Set $n = n + 1$ and go back to **Step 1.**

Their method features a single projection onto a feasible set, double functional evaluations of $\mathcal{A}$ and also incorporates self-adaptive stepsizes (see also [4, 21, 87, 90, 92]). We note that methods with self-adaptive stepsizes can automatically update the stepsize in each iteration by using some previously known information to perform a simple calculation. However, we reiterate that the stepsize of these methods does not increase, which may affect the efficiency of these algorithms since they rely on the selection of the initial stepsize.

Again, the major disadvantage of some of these methods is that they either involve two functional evaluations of $\mathcal{A}$ or two projections, which may be computationally costly and may affect the efficiency of the methods.

Several authors have justified inertial effects in algorithms as a means of accelerating optimization algorithms. The inertial techniques result from the discretization of a second-order dynamical system for minimizing a smooth convex function known as the heavy ball method:

$$\frac{d^2x}{dt^2}(t) + \gamma\frac{dx}{dt}(t) + \nabla f(x(t)) = 0,$$

where $\gamma > 0$ is a damping or friction parameter. The inertial extrapolation $x_n + \vartheta(x_n - x_{n-1})$ is the most popular means for achieving this purpose. The heavy ball method was first traced to Polyak [74]. Alvarez and Attouch[3] modified and applied it to a general maximal monotone operator.

Very recently, several authors have incorporated the inertial method in their proposed methods to achieve fast algorithms for approximating the solution of BVIP (1.1.1)- (1.1.2). For example, Tan et al. [88] modified Algorithm (2.3.7) by incorporating inertial method and non-monotonic stepsizes. Their algorithm is as follows:

Algorithm 2.3.10. ***Initialization:*** *Take $\theta > 0$, $\mu_1 > 0$, $\beta \in (0,1)$, $\phi \in (0,2)$ and $0 < \gamma < \frac{2\psi}{L_{\mathcal{T}}^2}$. Let x_0, $x_1 \in H$.*
Iterative steps: *Take the iterates x_{n-1} and $x_n (n \geq 1)$. Calculate x_{n+1} as follows*

Step 1: *Compute $u_n = x_n + \theta_n(x_n - x_{n-1})$, where*

$$\theta_n = \begin{cases} \min\left\{\frac{\epsilon_n}{\|x_n - x_{n-1}\|}, \theta\right\}, & \text{if } x_n \neq x_{n-1} \\ \theta, & \text{otherwise.} \end{cases}$$

Step 2: *Compute $y_n = P_C(u_n - \mu_n \mathcal{A} u_n)$.*

Step 3: *Compute $z_n = u_n - \phi\delta_n d_n$, where*

$$d_n := u_n - y_n - \mu_n(\mathcal{A}u_n - \mathcal{A}y_n) \quad \forall n \geq 0$$

and

$$\delta_n = \begin{cases} \frac{\langle u_n - y_n, d_n\rangle}{\|d_n\|^2}, & if \;\; d_n \neq 0 \\ 0, & if \; d_n = 0. \end{cases}$$

Step 4: *Compute*

$$x_{n+1} = z_n - \sigma_n \gamma \mathcal{T} z_n \quad \forall n \geq 0.$$

Update

$$\mu_{n+1} = \begin{cases} \min\left\{\frac{\beta\|u_n - y_n\|}{\|\mathcal{A}u_n - \mathcal{A}y_n\|}, \mu_n + \eta_n\right\}, & \text{if } \mathcal{A}u_n \neq \mathcal{A}y_n \\ \mu_n + \eta_n, & \text{otherwise.} \end{cases}$$

(See also [87, Algorithm 3.1 and Algorithm 3.2]).

In this dissertation, we introduce methods for finding a solution of BVIP (1.1.1)-(1.1.2)*. To achieve this, we pair the γ-strongly monotone and $L_{\mathcal{T}}$-Lipschitz continuous mapping $\mathcal{T}$ with the upper-level problem and the operator $\mathcal{A}$, which is monotone and Lipschitz continuous on C, to the lower-level problem under the condition that $VI(C, \mathcal{A})$, the solution set of VIP* (1.1.2)*, is nonempty. We note that the condition of the operator $\mathcal{A}$ on C guarantees that $VI(C, \mathcal{A})$ is closed and convex, and together with the assumption that $VI(C, \mathcal{A})$ is nonempty guarantees the existence of the unique solution of BVIP* (1.1.1)-(1.1.2) *(see [71, Proposition 1]). Our methods are based on the projection-type methods, regularization technique and inertial method; the proposed methods are motivated by the forward-reflected-backward method [67], Tikhonov regularization [94], and the heavy ball method [74]. Following that, we study new iterative schemes called the inertial regularization-projection methods for approximating the solution of BVIP* (1.1.1)-(1.1.2).

Chapter 3

3.1 Introduction

In this chapter, we introduce and study the inertial forward-reflected-backward method for approximating the solution of BVIP (1.1.1)-(1.1.2). The proposed method combines the forward-reflected-backward method, regularization technique and inertial extrapolation term. This method involves a single projection onto feasible set C and one functional evaluation. We show that the sequences generated iteratively by our method converge strongly to the solution of BVIP (1.1.1)-(1.1.2). Also, we illustrate the performance of our method in comparison with other recent results in the literature through numerical experiments.

3.2 Preliminaries

In this section, we state some important results that will be instrumental in the development of the main results of this chapter.

Lemma 3.2.1. *If u and v are non-negative numbers, then*

(i) $uv \leq \frac{u^2}{2p} + \frac{pv^2}{2} \quad \forall p \geq 0,$ **(Young's inequality)**

(ii) $(u+v)^2 \leq (2+\sqrt{2})u^2 + \sqrt{2}v^2.$

Lemma 3.2.2 ([80])**.** *Let $\{u_n\}$ be a sequence of non-negative real numbers, $\{e_n\}$ be a sequence of numbers in (0,1) such that $\sum_{n=1}^{\infty} e_n = \infty$, and let $\{v_n\}$ be a sequence of real numbers. Assume that*

$$u_{n+1} \leq (1-e_n)u_n + e_n v_n \quad \forall \quad n \geq 1,$$

If $\limsup_{i\to\infty} v_{n_i} \leq 0$ for every subsequence $\{u_{n_i}\}$ of $\{u_n\}$ satisfying $\liminf_{i\to\infty}(u_{n_i+1} - u_{n_i}) \geq 0$, then $\lim_{n\to\infty} u_n = 0$.

3.3 Main Results

Assumption 3.3.1. *The following assumptions will be used in the convergence analysis.*

(a) $\sigma_n \in (0,1), \quad \lim_{n\to\infty} \sigma_n = 0 \quad$ *and* $\quad \sum_{n=1}^{\infty} \sigma_n = \infty.$

(b)

$$\lim_{n\to\infty} \frac{|\sigma_{n+1} - \sigma_n|}{\sigma_n^2} = 0$$

(c) $\mathcal{A}$ *is monotone and* $L_{\mathcal{A}}$*-Lipschitz continuous,* $\mathcal{T}$ *is* γ*-strongly monotone and* $L_{\mathcal{T}}$*-Lipschitz continuous.*

(d) $VI(C, \mathcal{A}\ )$ *is nonempty.*

We give some examples of sequence $\{\sigma_n\}$ *that satisfies Assumption 3.3.1(a) and (b) as follows:*

$$\sigma_n = (n+1)^{-r}, \quad \sigma_n = \frac{1}{n^r} \log^{-q}(n+2) \quad \textit{and} \quad \sigma_n = \frac{1}{n^r}$$

where $q \leq 1$ *and* $r \in (0,1)$ *(see [46, 47]).*

Consider the following Regularized Variational Inequality Problem (RVIP) of (1.1.1)-(1.1.2)*:*

$$\textit{find } \ x^* \in C \quad \textit{such that } \langle \mathcal{A}x^* + \sigma\mathcal{T}x^*, x - x^* \rangle \geq 0 \qquad \forall x \in C, \tag{3.3.1}$$

where $\sigma > 0$ *is a regularization parameter. We remark that* (3.3.1) *has a unique solution, it follows from the fact that* $\mathcal{A}$ *is monotone and Lipschitz continuous, and* $\mathcal{T}$ *is strongly monotone and Lipschitz continuous. We denote this unique solution of* (3.3.1) *by* x_σ *for each* $\sigma > 0$*. Since* $VI(C, \mathcal{A})$ *is nonempty, closed and convex then the unique solution of BVIP* (1.1.1)-(1.1.2) *is guaranteed. Let this unique solution be* z^**.*

Proposition 3.3.2 ([46])**.** *Let* $\{x_\sigma\}_{\sigma\in(0,1)}$ *be the solution net of* (3.3.1)*, then*

(i) $\|x_\sigma\| \leq \|x^*\| + \frac{\|\mathcal{T}x^*\|}{\gamma}$ *for each* $x^* \in VI(C, \mathcal{A})$*, where* γ *is the modulus of strong monotonicity* $\mathcal{T}$*;*

(ii) *Let* $x_{\sigma_1}, x_{\sigma_2} \in \{x_\sigma\}_{\sigma\in(0,1)}$*, then* $\|x_{\sigma_1} - x_{\sigma_2}\| \leq \frac{|\sigma_1 - \sigma_2|}{\sigma_1} M$*, where* $M > 0$ *is a constant;*

(iii) $\|x_\sigma - z^*\| \to 0$ *as* $\sigma \to \infty$*.*

Algorithm 3.3.3. *Inertial regularised forward-reflected-backward method with constant inertial parameter*

Step 0: $\vartheta \in [0,1)$, $\epsilon \in \left(0, \frac{1}{4}\right)$*, and choose sequence* $\{\sigma_n\}$ *in* $(0,1)$*. For arbitrary* $x_0, x_1 \in H$*. Set* $n := 1$*.*

Step 1: *Compute*

$$x_{n+1} = P_C(m_n - \delta_n \mathcal{A} x_n - k_n - \delta_n \sigma_n \mathcal{T} x_n) \quad \forall n \geq 1,$$

where

$$\begin{cases} m_n = x_n + \vartheta(x_n - x_{n-1}) \\ k_n = \delta_{n-1}(1 - \gamma\delta_n\sigma_n)(\mathcal{A}x_n - \mathcal{A}x_{n-1}) \end{cases}$$

and

$$\delta_n \in \left[\varpi, \min\left\{\frac{1-2\varpi}{2L_{\mathcal{A}}}, \frac{1}{\gamma}\right\}\right] \text{ for some } \varpi > \ 0.$$

Step 2: *Set* $n = n+1$ *and go back to* **Step 1.**

Lemma 3.3.4. *Let* $\{x_n\}$ *be the sequence generated by Algorithm 3.3.3. If* $0 \leq \vartheta < \min\left\{\frac{\varpi}{6}, \frac{\frac{1}{2}-\epsilon}{2}\right\}$, *then there exists* $n_2 \in \mathbb{N}$ *such that*

$$\begin{aligned}
&\|x_{n+1} - x_{\sigma_n}\|^2 + 2\delta_n\langle \mathcal{A}x_{n+1} - \mathcal{A}x_n, x_{\sigma_n} - x_{n+1}\rangle - \vartheta\|x_n - x_{\sigma_n}\|^2 + \frac{1}{2}\|x_n - x_{n+1}\|^2 \\
&\leq \left(1 - \frac{3\delta_n\sigma_n\gamma}{2}\right)\|x_n - x_{\sigma_n}\|^2 + 2\delta_{n-1}(1 - \delta_n\sigma_n\gamma)\langle \mathcal{A}x_{n-1} - \mathcal{A}x_n, x_n - x_{\sigma_n}\rangle \\
&\quad -\vartheta(1 - \delta_n\sigma_n\gamma)\|x_{n-1} - x_{\sigma_n}\|^2 \\
&\quad + \frac{1}{2}(1 - \delta_n\sigma_n\gamma)\|x_{n-1} - x_n\|^2 + \left(-\varpi + ((1+\sqrt{2})\vartheta)\right)\|x_{n+1} - x_n\|^2 \quad \forall n \geq n_2,
\end{aligned}$$

where x_{σ_n} *is the solution to the RVIP (3.3.1) with* $\sigma = \sigma_n$.

Proof. From Algorithm 3.3.3, (2.1.16) and (2.1.5), we have

$$\begin{aligned}
\|x_{n+1} - x_{\sigma_n}\|^2 &\leq \|m_n - \delta_n\mathcal{A}x_n - k_n - \delta_n\sigma_n\mathcal{T}x_n - x_{\sigma_n}\|^2 \\
&\quad -\|m_n - \delta_n\mathcal{A}x_n - k_n - \delta_n\sigma_n\mathcal{T}x_n - x_{n+1}\|^2 \\
&= \|m_n - \delta_n\mathcal{A}x_n - \delta_n\sigma_n\mathcal{T}x_n - x_{\sigma_n}\|^2 + \|k_n\|^2 \\
&\quad -2\langle m_n - \delta_n\mathcal{A}x_n - \delta_n\sigma_n\mathcal{T}x_n - x_{\sigma_n}, k_n\rangle \\
&\quad -\|m_n - \delta_n\mathcal{A}x_n - \delta_n\sigma_n\mathcal{T}x_n - x_{n+1}\|^2 - \|k_n\|^2 \\
&\quad +2\langle m_n - \delta_n\mathcal{A}x_n - \delta_n\sigma_n\mathcal{T}x_n - x_{n+1}, k_n\rangle \\
&= \|m_n - \delta_n\mathcal{A}x_n - \delta_n\sigma_n\mathcal{T}x_n - x_{\sigma_n}\|^2 \\
&\quad -\|m_n - \delta_n\mathcal{A}x_n - \delta_n\sigma_n\mathcal{T}x_n - x_{n+1}\|^2 \\
&\quad -2\delta_{n-1}(1 - \gamma\delta_n\sigma_n)\langle \mathcal{A}x_n - \mathcal{A}x_{n-1}, x_{n+1} - x_{\sigma_n}\rangle. \qquad (3.3.2)
\end{aligned}$$

From (3.3.2) and by (2.1.5), we have that

$$\begin{aligned}
\|m_n - x_{\sigma_n} - (\delta_n\mathcal{A}x_n + \delta_n\sigma_n\mathcal{T}x_n)\|^2 &= \|m_n - x_{\sigma_n}\|^2 + \delta_n^2\|\mathcal{A}x_n + \sigma_n\mathcal{T}x_n\|^2 \\
&\quad -2\delta_n\langle m_n - x_{\sigma_n}, \mathcal{A}x_n + \sigma_n\mathcal{T}x_n\rangle \qquad (3.3.3)
\end{aligned}$$

and

$$\begin{aligned}\|m_n - x_{n+1} - (\delta_n \mathcal{A}x_n + \delta_n\sigma_n \mathcal{T}x_n)\|^2 &= \|m_n - x_{n+1}\|^2 + \delta_n^2\|\mathcal{A}x_n + \sigma_n\mathcal{T}x_n\|^2 \\ &\quad -2\delta_n\langle m_n - x_{n+1}, \mathcal{A}x_n + \sigma_n\mathcal{T}x_n\rangle. \end{aligned} \tag{3.3.4}$$

Combining (3.3.2), (3.3.3) and (3.3.4), we obtain

$$\begin{aligned}\|x_{n+1} - x_{\sigma_n}\|^2 &\leq \|m_n - x_{\sigma_n}\|^2 + \delta_n^2\|\mathcal{A}x_n + \sigma_n\mathcal{T}x_n\|^2 - 2\delta_n\langle m_n - x_{\sigma_n}, \mathcal{A}x_n + \sigma_n\mathcal{T}x_n\rangle \\ &\quad -\|m_n - x_{n+1}\|^2 - \delta_n^2\|\mathcal{A}x_n + \sigma_n\mathcal{T}x_n\|^2 \\ &\quad +2\delta_n\langle m_n - x_{n+1}, \mathcal{A}x_n + \sigma_n\mathcal{T}x_n\rangle \\ &\quad -2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_n - \mathcal{A}x_{n-1}, x_{n+1} - x_{\sigma_n}\rangle \\ &= \|m_n - x_{\sigma_n}\|^2 - \|m_n - x_{n+1}\|^2 + 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{\sigma_n} - x_{n+1}\rangle \\ &\quad +2\delta_n\langle \mathcal{A}x_{n+1} + \sigma_n\mathcal{T}x_n, x_{\sigma_n} - x_{n+1}\rangle \\ &\quad -2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_n - \mathcal{A}x_{n-1}, x_{n+1} - x_n\rangle \\ &\quad -2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_n - \mathcal{A}x_{n-1}, x_n - x_{\sigma_n}\rangle. \end{aligned} \tag{3.3.5}$$

By the monotonicity of $\mathcal{A}$, we have

$$\langle \mathcal{A}x_{\sigma_n} - \mathcal{A}x_{n+1}, x_{\sigma_n} - x_{n+1}\rangle \geq 0,$$

which implies

$$\langle \mathcal{A}x_{n+1}, x_{\sigma_n} - x_{n+1}\rangle \leq \langle \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1}\rangle. \tag{3.3.6}$$

Recall that x_{σ_n} is the solution of RVIP (3.3.1), where $\sigma = \sigma_n$. Thus,

$$\langle \mathcal{A}x_{\sigma_n} + \sigma_n\mathcal{T}x_{\sigma_n}, x_{n+1} - x_{\sigma_n}\rangle \geq 0.$$

Hence,

$$\langle \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1}\rangle \leq \sigma_n\langle \mathcal{T}x_{\sigma_n}, x_{n+1} - x_{\sigma_n}\rangle. \tag{3.3.7}$$

From (3.3.6) and (3.3.7) we obtain the following:

$$\langle \mathcal{A}x_{n+1}, x_{\sigma_n} - x_{n+1}\rangle \leq \sigma_n\langle \mathcal{T}x_{\sigma_n}, x_{n+1} - x_{\sigma_n}\rangle. \tag{3.3.8}$$

Adding $\sigma_n\langle \mathcal{T}x_n, x_{\sigma_n} - x_{n+1}\rangle$ to both sides of (3.3.8) gives;

$$\langle \mathcal{A}x_{n+1} + \sigma_n\mathcal{T}x_n, x_{\sigma_n} - x_{n+1}\rangle \leq \sigma_n\langle \mathcal{T}x_{\sigma_n} - \mathcal{T}x_n, x_{n+1} - x_{\sigma_n}\rangle.$$

So,

$$\begin{aligned}2\delta_n\langle \mathcal{A}x_{n+1} + \sigma_n\mathcal{T}x_n, x_{\sigma_n} - x_{n+1}\rangle &\leq -2\delta_n\sigma_n\langle \mathcal{T}x_n - \mathcal{T}x_{\sigma_n}, x_{n+1} - x_n\rangle \\ &\quad -2\delta_n\sigma_n\langle \mathcal{T}x_n - \mathcal{T}x_{\sigma_n}, x_n - x_{\sigma_n}\rangle. \end{aligned} \tag{3.3.9}$$

Using (3.3.9) in (3.3.5) we get the following inequality:

$$\begin{aligned}
\|x_{n+1}-x_{\sigma_n}\|^2 &\leq \|m_n-x_{\sigma_n}\|^2-\|m_n-x_{n+1}\|^2-2\delta_n\sigma_n\langle \mathcal{T}x_n-\mathcal{T}x_{\sigma_n}, x_{n+1}-x_n\rangle \\
&\quad -2\delta_n\sigma_n\langle \mathcal{T}x_n-\mathcal{T}x_{\sigma_n}, x_n-x_{\sigma_n}\rangle \\
&\quad -2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_n-\mathcal{A}x_{n-1}, x_{n+1}-x_n\rangle \\
&\quad -2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_n-\mathcal{A}x_{n-1}, x_n-x_{\sigma_n}\rangle \\
&\quad -2\delta_n\langle \mathcal{A}x_{n+1}-\mathcal{A}x_n, x_{\sigma_n}-x_{n+1}\rangle \\
&= \|m_n-x_{\sigma_n}\|^2-\|m_n-x_{n+1}\|^2+2\delta_n\sigma_n\langle \mathcal{T}x_{\sigma_n}-\mathcal{T}x_n, x_{n+1}-x_n\rangle \\
&\quad -2\delta_n\sigma_n\langle \mathcal{T}x_n-\mathcal{T}x_{\sigma_n}, x_n-x_{\sigma_n}\rangle \\
&\quad +2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_{n-1}-\mathcal{A}x_n, x_{n+1}-x_n\rangle \\
&\quad +2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_{n-1}-\mathcal{A}x_n, x_n-x_{\sigma_n}\rangle \\
&\quad -2\delta_n\langle \mathcal{A}x_{n+1}-\mathcal{A}x_n, x_{\sigma_n}-x_{n+1}\rangle. \qquad (3.3.10)
\end{aligned}$$

By the strong monotonicity and Lipschitz continuity of $\mathcal{T}$, Lipschitz continuity of $\mathcal{A}$, and by Lemma 3.2.1(i), we have

$$\begin{aligned}
\|x_{n+1}-x_{\sigma_n}\|^2 &\leq \|m_n-x_{\sigma_n}\|^2-\|m_n-x_{n+1}\|^2 \\
&\quad +\delta_n\sigma_n L_{\mathcal{T}}\left(\frac{\|x_{\sigma_n}-x_n\|^2}{p^*}+\|x_{n+1}-x_n\|^2 p^*\right) \\
&\quad -2\delta_n\sigma_n\gamma\|x_n-x_{\sigma_n}\|^2 \\
&\quad +\delta_{n-1}L_{\mathcal{A}}(1-\gamma\delta_n\sigma_n)(\|x_{n-1}-x_n\|^2+\|x_{n+1}-x_n\|^2) \\
&\quad +2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_{n-1}-\mathcal{A}x_n, x_n-x_{\sigma_n}\rangle \\
&\quad -2\delta_n\langle \mathcal{A}x_{n+1}-\mathcal{A}x_n, x_{\sigma_n}-x_{n+1}\rangle. \qquad (3.3.11)
\end{aligned}$$

Let $p^*=\frac{2L_{\mathcal{T}}}{\gamma}$. Hence,

$$\begin{aligned}
\|x_{n+1}-x_{\sigma_n}\|^2 &\leq \|m_n-x_{\sigma_n}\|^2-\|m_n-x_{n+1}\|^2+\frac{\delta_n\sigma_n\gamma}{2}\|x_{\sigma_n}-x_n\|^2 \\
&\quad +\frac{2\delta_n\sigma_n L_{\mathcal{T}}^2}{\gamma}\|x_{n+1}-x_n\|^2-2\delta_n\sigma_n\gamma\|x_n-x_{\sigma_n}\|^2 \\
&\quad +\delta_{n-1}L_{\mathcal{A}}(1-\gamma\delta_n\sigma_n)(\|x_{n-1}-x_n\|^2+\|x_{n+1}-x_n\|^2) \\
&\quad +2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_{n-1}-\mathcal{A}x_n, x_n-x_{\sigma_n}\rangle \\
&\quad -2\delta_n\langle \mathcal{A}x_{n+1}-\mathcal{A}x_n, x_{\sigma_n}-x_{n+1}\rangle \\
&\leq \|m_n-x_{\sigma_n}\|^2-\|m_n-x_{n+1}\|^2-\frac{3\delta_n\sigma_n\gamma}{2}\|x_{\sigma_n}-x_n\|^2 \\
&\quad +\frac{2\delta_n\sigma_n L_{\mathcal{T}}^2}{\gamma}\|x_{n+1}-x_n\|^2+(1-\gamma\delta_n\sigma_n)\left(\frac{1}{2}-\varpi\right)\|x_{n-1}-x_n\|^2 \\
&\quad +\delta_{n-1}L_{\mathcal{A}}(1-\gamma\delta_n\sigma_n)\|x_{n+1}-x_n\|^2 \\
&\quad +2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_{n-1}-\mathcal{A}x_n, x_n-x_{\sigma_n}\rangle \\
&\quad -2\delta_n\langle \mathcal{A}x_{n+1}-\mathcal{A}x_n, x_{\sigma_n}-x_{n+1}\rangle. \qquad (3.3.12)
\end{aligned}$$

Now, by applying (2.1.5) and (2.1.3), we get

$$\begin{aligned}\|m_n - x_{\sigma_n}\|^2 &= \|(x_n - x_{\sigma_n}) + \vartheta(x_n - x_{n-1})\|^2 \\ &= \|x_n - x_{\sigma_n}\|^2 + \vartheta^2\|x_n - x_{n-1}\|^2 + 2\vartheta\langle x_n - x_{\sigma_n}, x_n - x_{n-1}\rangle \\ &= \|x_n - x_{\sigma_n}\|^2 + \vartheta^2\|x_n - x_{n-1}\|^2 \\ &\quad -\vartheta(\|x_{\sigma_n} - x_{n-1}\|^2 - \|x_{\sigma_n} - x_n\|^2 - \|x_n - x_{n-1}\|^2) \\ &= \|x_n - x_{\sigma_n}\|^2 + \vartheta^2\|x_n - x_{n-1}\|^2 \\ &\quad -\vartheta\|x_{\sigma_n} - x_{n-1}\|^2 + \vartheta\|x_{\sigma_n} - x_n\|^2 + \vartheta\|x_n - x_{n-1}\|^2. \end{aligned} \tag{3.3.13}$$

Replacing x_{σ_n} by x_{n+1} in (3.3.13), we obtain

$$\begin{aligned}\|m_n - x_{n+1}\|^2 &= \|x_n - x_{n+1}\|^2 + \vartheta^2\|x_n - x_{n-1}\|^2 - \vartheta\|x_{n+1} - x_{n-1}\|^2 \\ &\quad +\vartheta\|x_{n+1} - x_n\|^2 + \vartheta\|x_n - x_{n-1}\|^2. \end{aligned} \tag{3.3.14}$$

Subtracting (3.3.14) from (3.3.13) gives

$$\begin{aligned}\|m_n - x_{\sigma_n}\|^2 - \|m_n - x_{n+1}\|^2 &= (1+\vartheta)\|x_n - x_{\sigma_n}\|^2 - (1+\vartheta)\|x_n - x_{n+1}\|^2 \\ &\quad -\vartheta(1 - \gamma\delta_n\sigma_n)\|x_{n-1} - x_{\sigma_n}\|^2 \\ &\quad +\vartheta\|x_{n+1} - x_{n-1}\|^2 \\ &\quad -\gamma\delta_n\sigma_n\vartheta\|x_{n-1} - x_{\sigma_n}\|^2. \end{aligned} \tag{3.3.15}$$

Substituting equation (3.3.15) in (3.3.12), we obtain the following inequality

$$\begin{aligned}\|x_{n+1} - x_{\sigma_n}\|^2 &\leq (1+\vartheta)\|x_n - x_{\sigma_n}\|^2 - (1+\vartheta)\|x_n - x_{n+1}\|^2 \\ &\quad -2\delta_n\langle \mathcal{A}x_{n+1} - \mathcal{A}x_n, x_{\sigma_n} - x_{n+1}\rangle - \vartheta(1 - \gamma\delta_n\sigma_n)\|x_{n-1} - x_{\sigma_n}\|^2 \\ &\quad +\vartheta\|x_{n+1} - x_{n-1}\|^2 - \gamma\delta_n\sigma_n\vartheta\|x_{n-1} - x_{\sigma_n}\|^2 \\ &\quad +(1 - \gamma\delta_n\sigma_n)\left(\frac{1}{2} - \varpi\right)\|x_{n-1} - x_n\|^2 \\ &\quad +\delta_{n-1}L_{\mathcal{A}}(1 - \gamma\delta_n\sigma_n)\|x_{n+1} - x_n\|^2 \\ &\quad +2\delta_{n-1}(1 - \gamma\delta_n\sigma_n)\langle \mathcal{A}x_{n-1} - \mathcal{A}x_n, x_n - x_{\sigma_n}\rangle \\ &\quad -\frac{3\delta_n\sigma_n\gamma}{2}\|x_{\sigma_n} - x_n\|^2 + \frac{2\delta_n\sigma_n L_{\mathcal{T}}^2}{\gamma}\|x_{n+1} - x_n\|^2 \\ &= \left(1 - \frac{3\delta_n\sigma_n\gamma}{2}\right)\|x_n - x_{\sigma_n}\|^2 - \vartheta(1 - \gamma\delta_n\sigma_n)\|x_{n-1} - x_{\sigma_n}\|^2 \\ &\quad +2\delta_{n-1}(1 - \gamma\delta_n\sigma_n)\langle \mathcal{A}x_{n-1} - \mathcal{A}x_n, x_n - x_{\sigma_n}\rangle - \vartheta\|x_n - x_{n+1}\|^2 \\ &\quad +\vartheta\|x_{n+1} - x_{n-1}\|^2 - \gamma\delta_n\sigma_n\vartheta\|x_{n-1} - x_{\sigma_n}\|^2 \\ &\quad +\frac{1}{2}(1 - \gamma\delta_n\sigma_n)\|x_{n-1} - x_n\|^2 - \varpi(1 - \gamma\delta_n\sigma_n)\|x_{n-1} - x_n\|^2 \\ &\quad -2\delta_n\langle \mathcal{A}x_{n+1} - \mathcal{A}x_n, x_{\sigma_n} - x_{n+1}\rangle + \vartheta\|x_n - x_{\sigma_n}\|^2 \\ &\quad -\left(1 - \delta_{n-1}L_{\mathcal{A}}(1 - \gamma\delta_n\sigma_n) - \frac{2\delta_n\sigma_n L_{\mathcal{T}}^2}{\gamma}\right)\|x_n - x_{n+1}\|^2. \end{aligned} \tag{3.3.16}$$

Since

$$\begin{aligned}\lim_{n\to\infty}\left(1-\delta_{n-1}L_{\mathcal{A}}(1-\gamma\delta_n\sigma_n)-\frac{2\delta_n\sigma_n L_{\mathcal{T}}^2}{\gamma}\right) &\geq \lim_{n\to\infty}\left(1-(\frac{1}{2}-\varpi)(1-\gamma\delta_n\sigma_n)-\frac{2\delta_n\sigma_n L_{\mathcal{T}}^2}{\gamma}\right)\\ &= \frac{1}{2}+\varpi,\end{aligned}$$

there exists $n_1\in\mathbb{N}$, $n_1\geq n_0$ such that $1-\delta_{n-1}L_{\mathcal{A}}(1-\gamma\delta_n\sigma_n)-\frac{2\delta_n\sigma_n L_{\mathcal{T}}^2}{\gamma}>\frac{1}{2}+\varpi$ $\forall n\geq n_1$. Thus, (3.3.16) becomes

$$\begin{aligned}\|x_{n+1}-x_{\sigma_n}\|^2 \leq& \left(1-\frac{3\delta_n\sigma_n\gamma}{2}\right)\|x_n-x_{\sigma_n}\|^2-\vartheta(1-\gamma\delta_n\sigma_n)\|x_{n-1}-x_{\sigma_n}\|^2\\ &+2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle\mathcal{A}x_{n-1}-\mathcal{A}x_n,x_n-x_{\sigma_n}\rangle\\ &-\vartheta\|x_n-x_{n+1}\|^2+\vartheta\|x_{n+1}-x_{n-1}\|^2-\gamma\delta_n\sigma_n\vartheta\|x_{n-1}-x_{\sigma_n}\|^2\\ &+\frac{1}{2}(1-\gamma\delta_n\sigma_n)\|x_{n-1}-x_n\|^2-\varpi\|x_n-x_{n+1}\|^2\\ &-\varpi(1-\gamma\delta_n\sigma_n)\|x_{n-1}-x_n\|^2-2\delta_n\langle\mathcal{A}x_{n+1}-\mathcal{A}x_n,x_{\sigma_n}-x_{n+1}\rangle\\ &+\vartheta\|x_n-x_{\sigma_n}\|^2-\frac{1}{2}\|x_n-x_{n+1}\|^2 \quad\forall\ \ n\geq n_1.\end{aligned}\tag{3.3.17}$$

Now, by Lemma 3.2.1(ii), we get

$$\begin{aligned}\vartheta\|x_{n+1}-x_{n-1}\|^2 &= \vartheta\|(x_{n+1}-x_n)+(x_n-x_{n-1})\|^2\\ &\leq (2+\sqrt{2})\vartheta\|x_{n+1}-x_n\|^2+\sqrt{2}\vartheta\|x_n-x_{n-1}\|^2.\end{aligned}\tag{3.3.18}$$

Using (3.3.18) in (3.3.17), we get

$$\begin{aligned}\|x_{n+1}-x_{\sigma_n}\|^2 \leq& \left(1-\frac{3\delta_n\sigma_n\gamma}{2}\right)\|x_n-x_{\sigma_n}\|^2\\ &+2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle\mathcal{A}x_{n-1}-\mathcal{A}x_n,x_n-x_{\sigma_n}\rangle\\ &-\vartheta(1-\gamma\delta_n\sigma_n)\|x_{n-1}-x_{\sigma_n}\|^2+\frac{1}{2}(1-\gamma\delta_n\sigma_n)\|x_{n-1}-x_n\|^2\\ &+(-\varpi+((1+\sqrt{2})\vartheta))\|x_{n+1}-x_n\|^2\\ &+(-\varpi(1-\gamma\delta_n\sigma_n)+\sqrt{2}\vartheta)\|x_{n-1}-x_n\|^2\\ &-\gamma\delta_n\sigma_n\vartheta\|x_{n-1}-x_{\sigma_n}\|^2-2\delta_n\langle\mathcal{A}x_{n+1}-\mathcal{A}x_n,x_{\sigma_n}-x_{n+1}\rangle\\ &+\vartheta\|x_n-x_{\sigma_n}\|^2-\frac{1}{2}\|x_n-x_{n+1}\|^2.\end{aligned}\tag{3.3.19}$$

Since $\lim_{n\to\infty}(-\varpi(1-\gamma\delta_n\sigma_n)+\sqrt{2}\vartheta)=-\varpi+\sqrt{2}\vartheta<-\varpi+2\vartheta<-\varpi+2\times\frac{\varpi}{6}=\frac{-2\varpi}{3}<0$, there exists $n_2\in\mathbb{N}$, $n_2\geq n_1$ such that $(-\varpi(1-\gamma\delta_n\sigma_n)+\sqrt{2}\vartheta)<0\quad\forall\ n\geq n_2$. Also, γ

$\delta_n\sigma_n\vartheta > 0$. Therefore, it follows from (3.3.19) that

$$
\begin{aligned}
&\|x_{n+1}-x_{\sigma_n}\|^2 + 2\delta_n\langle \mathcal{A}x_{n+1}-\mathcal{A}x_n, x_{\sigma_n}-x_{n+1}\rangle - \vartheta\|x_n - x_{\sigma_n}\|^2 + \frac{1}{2}\|x_n - x_{n+1}\|^2 \\
&\leq \left(1-\frac{3\delta_n\sigma_n\gamma}{2}\right)\|x_n - x_{\sigma_n}\|^2 + 2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_{n-1}-\mathcal{A}x_n, x_n - x_{\sigma_n}\rangle \\
&\quad -\vartheta(1-\gamma\delta_n\sigma_n)\|x_{n-1}-x_{\sigma_n}\|^2 + \frac{1}{2}(1-\gamma\delta_n\sigma_n)\|x_{n-1}-x_n\|^2 \\
&\quad +(-\varpi + ((1+\sqrt{2})\vartheta))\|x_{n+1}-x_n\|^2, \quad \forall \ n \geq n_2.
\end{aligned} \tag{3.3.20}
$$

□

Theorem 3.3.5. *Let $\{x_n\}$ be the sequence generated by Algorithm 3.3.3. If $0 \leq \vartheta < \min\left\{\frac{\varpi}{6}, \frac{\frac{1}{2}-\epsilon}{2}\right\}$, then $\{x_n\}$ converges strongly to z^*.*

Proof. By Proposition 3.3.2 (ii), we obtain

$$
\begin{aligned}
\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1}-x_{\sigma_{n+1}}\rangle &= \langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1}-x_{\sigma_n}\rangle \\
&\quad +\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{\sigma_n}-x_{\sigma_{n+1}}\rangle \\
&\leq \langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1}-x_{\sigma_n}\rangle \\
&\quad +L_{\mathcal{A}}\|x_n - x_{n+1}\|\|x_{\sigma_n}-x_{\sigma_{n+1}}\| \\
&\leq \langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1}-x_{\sigma_n}\rangle \\
&\quad +L_{\mathcal{A}}\|x_n - x_{n+1}\|\frac{|\sigma_{n+1}-\sigma_n|}{\sigma_n}M.
\end{aligned} \tag{3.3.21}
$$

Let $p_n \in (0,1)$, then Inequality (3.3.21) becomes

$$
\begin{aligned}
2\delta_n(1-p_n)\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1}-x_{\sigma_{n+1}}\rangle &\leq 2\delta_n(1-p_n)\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1}-x_{\sigma_n}\rangle \\
&\quad +2\delta_n(1-p_n)L_{\mathcal{A}}\|x_n - x_{n+1}\| \\
&\quad \times\frac{|\sigma_{n+1}-\sigma_n|}{\sigma_n}M.
\end{aligned} \tag{3.3.22}
$$

Combining Equation (2.1.5), Proposition 3.3.2 (ii), Lemma 3.2.1(i), and for any $p_n \in (0,1)$,

$$
\begin{aligned}
\|x_{n+1}-x_{\sigma_n}\|^2 &= \|x_{n+1}-x_{\sigma_{n+1}}\|^2 + \|x_{\sigma_{n+1}}-x_{\sigma_n}\|^2 + 2\langle x_{n+1}-x_{\sigma_{n+1}}, x_{\sigma_{n+1}}-x_{\sigma_n}\rangle \\
&\geq \|x_{n+1}-x_{\sigma_{n+1}}\|^2 + \|x_{\sigma_{n+1}}-x_{\sigma_n}\|^2 \\
&\quad -2\left(\frac{p_n\|x_{\sigma_{n+1}}-x_{n+1}\|^2}{2} + \frac{\|x_{\sigma_{n+1}}-x_{\sigma_n}\|^2}{2p_n}\right) \\
&= (1-p_n)\|x_{n+1}-x_{\sigma_{n+1}}\|^2 - \left(\frac{1-p_n}{p_n}\right)\|x_{\sigma_{n+1}}-x_{\sigma_n}\|^2 \\
&\geq (1-p_n)\|x_{n+1}-x_{\sigma_{n+1}}\|^2 - \left(\frac{1-p_n}{p_n}\right)\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2 M^2.
\end{aligned} \tag{3.3.23}
$$

We combine (3.3.22) and (3.3.23) to obtain the following:

$$
\begin{aligned}
&(1-p_n)(\|x_{n+1}-x_{\sigma_{n+1}}\|^2+2\delta_n\langle \mathcal{A}x_n-\mathcal{A}x_{n+1},x_{n+1}-x_{\sigma_{n+1}}\rangle\\
&\ -\vartheta\|x_n-x_{\sigma_n}\|^2+\frac{1}{2}\|x_n-x_{n+1}\|^2)\\
&\leq \|x_{n+1}-x_{\sigma_n}\|^2+2\delta_n\langle \mathcal{A}x_n-\mathcal{A}x_{n+1},x_{n+1}-x_{\sigma_n}\rangle\\
&\ -\vartheta\|x_n-x_{\sigma_n}\|^2+\frac{1}{2}\|x_n-x_{n+1}\|^2\\
&\ +(1-p_n)(1-2\varpi)\|x_n-x_{n+1}\|\frac{|\sigma_{n+1}-\sigma_n|}{\sigma_n}M\\
&\ +\left(\frac{1-p_n}{p_n}\right)\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2M^2\\
&\ -2\delta_np_n\langle \mathcal{A}x_n-\mathcal{A}x_{n+1},x_{n+1}-x_{\sigma_n}\rangle\\
&\ +p_n\vartheta\|x_n-x_{\sigma_n}\|^2-\frac{p_n}{2}\|x_n-x_{n+1}\|^2. \qquad (3.3.24)
\end{aligned}
$$

We estimate $(1-p_n)(1-2\varpi)\|x_n-x_{n+1}\|\frac{|\sigma_{n+1}-\sigma_n|}{\sigma_n}M$ using Lemma 3.2.1(i)

$$
\begin{aligned}
&(1-p_n)(1-2\varpi)\|x_n-x_{n+1}\|\frac{|\sigma_{n+1}-\sigma_n|}{\sigma_n}M\\
&=(1-p_n)\|x_n-x_{n+1}\|\frac{|\sigma_{n+1}-\sigma_n|}{\sigma_n}M\\
&\ \ -2\varpi(1-p_n)\|x_n-x_{n+1}\|\frac{|\sigma_{n+1}-\sigma_n|}{\sigma_n}M\\
&\leq (1-p_n)\left(\frac{p}{4}\|x_n-x_{n+1}\|^2+\frac{1}{p}\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2M^2\right). \qquad (3.3.25)
\end{aligned}
$$

We also estimate $-2\delta_np_n\langle \mathcal{A}x_n-\mathcal{A}x_{n+1},x_{n+1}-x_{\sigma_n}\rangle$.

$$
\begin{aligned}
&-2\delta_np_n\langle \mathcal{A}x_n-\mathcal{A}x_{n+1},x_{n+1}-x_{\sigma_n}\rangle\\
&=-2\delta_np_n\langle \mathcal{A}x_n-\mathcal{A}x_{n+1},x_{n+1}-x_n\rangle-2\delta_np_n\langle \mathcal{A}x_n-\mathcal{A}x_{n+1},x_n-x_{\sigma_n}\rangle\\
&\leq 2\delta_np_nL_{\mathcal{A}}\|x_{n+1}-x_n\|^2+2\delta_np_nL_{\mathcal{A}}\|x_{n+1}-x_n\|\|x_n-x_{\sigma_n}\|\\
&\leq p_n(1-2\varpi)\|x_{n+1}-x_n\|^2\\
&\ +p_n(1-2\varpi)\left(\frac{\|x_{n+1}-x_n\|^2}{4}+\|x_n-x_{\sigma_n}\|^2\right). \qquad (3.3.26)
\end{aligned}
$$

Using (3.3.24), (3.3.25), (3.3.26) and the condition $\vartheta \in [0, 1)$, we obtain

$$
\begin{aligned}
&(1-p_n)(\|x_{n+1}-x_{\sigma_{n+1}}\|^2 + 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1}-x_{\sigma_{n+1}}\rangle \\
&\quad -\vartheta\|x_n - x_{\sigma_n}\|^2 + \frac{1}{2}\|x_n - x_{n+1}\|^2) \\
&\le \|x_{n+1}-x_{\sigma_n}\|^2 + 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1}-x_{\sigma_n}\rangle \\
&\quad -\vartheta\|x_n - x_{\sigma_n}\|^2 + \frac{1}{2}\|x_n - x_{n+1}\|^2 \\
&\quad +(1-p_n)\left(\frac{p}{4}\|x_n - x_{n+1}\|^2 + \frac{1}{p}\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2 M^2\right) \\
&\quad +\left(\frac{1-p_n}{p_n}\right)\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2 M^2 \\
&\quad +p_n(1-2\varpi)\|x_{n+1}-x_n\|^2 + p_n\vartheta\|x_n - x_{\sigma_n}\|^2 - \frac{p_n}{2}\|x_n - x_{n+1}\|^2 \\
&\quad +p_n(1-2\varpi)\frac{\|x_{n+1}-x_n\|^2}{4} + p_n\|x_n - x_{\sigma_n}\|^2 - 2p_n\varpi\|x_n - x_{\sigma_n}\|^2 \\
&\le \|x_{n+1}-x_{\sigma_n}\|^2 + 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1}-x_{\sigma_n}\rangle - \vartheta\|x_n - x_{\sigma_n}\|^2 \\
&\quad +\frac{1}{2}\|x_n - x_{n+1}\|^2 + \frac{(1-p_n)}{p}\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2 M^2 \\
&\quad +\left(\frac{1-p_n}{p_n}\right)\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2 M^2 + 2p_n\|x_n - x_{\sigma_n}\|^2 \\
&\quad +\left(\frac{(1-p_n)p}{4} + p_n(1-2\varpi) + \frac{p_n(1-2\varpi)}{4} - \frac{p_n}{2}\right)\|x_n - x_{n+1}\|^2.
\end{aligned}
\tag{3.3.27}
$$

Combining (3.3.20) and (3.3.27) we obtain the following inequality:

$$
\begin{aligned}
&(1-p_n)(\|x_{n+1}-x_{\sigma_{n+1}}\|^2 + 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1}-x_{\sigma_{n+1}}\rangle \\
&\quad -\vartheta\|x_n - x_{\sigma_n}\|^2 + \frac{1}{2}\|x_n - x_{n+1}\|^2) \\
&\le \left(1 - \frac{3\delta_n\sigma_n\gamma}{2} + 2p_n\right)\|x_n - x_{\sigma_n}\|^2 \\
&\quad +2\delta_{n-1}(1-\delta_n\sigma_n\gamma)\langle \mathcal{A}x_{n-1} - \mathcal{A}x_n, x_n - x_{\sigma_n}\rangle \\
&\quad -\vartheta(1-\gamma\delta_n\sigma_n)\|x_{n-1}-x_{\sigma_n}\|^2 + \frac{1}{2}(1-\delta_n\sigma_n\gamma)\|x_{n-1}-x_n\|^2 \\
&\quad +\frac{(1-p_n)}{pp_n}(p_n+p)\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2 M^2 \\
&\quad +\left(\frac{(1-p_n)p}{4} + p_n(1-2\varpi) + \frac{p_n(1-2\varpi)}{4} - \frac{p_n}{2}\right. \\
&\quad \left. +(-\varpi + (1+\sqrt{2})\vartheta)\right)\|x_n - x_{n+1}\|^2.
\end{aligned}
\tag{3.3.28}
$$

$$
\text{Take} \quad p_n = 0.25\gamma\delta_n\sigma_n \quad \text{then} \quad \left(1 - \frac{3\gamma\sigma_n\delta_n}{2} + 2p_n\right) = (1-\gamma\sigma_n\delta_n). \tag{3.3.29}
$$

From (3.3.28) and (3.3.29), we have

$$\begin{aligned}
&(1-p_n)(\|x_{n+1}-x_{\sigma_{n+1}}\|^2+2\delta_n\langle \mathcal{A}x_n-\mathcal{A}x_{n+1}, x_{n+1}-x_{\sigma_{n+1}}\rangle \\
&\ -\vartheta\|x_n-x_{\sigma_n}\|^2+\frac{1}{2}\|x_n-x_{n+1}\|^2) \\
&\le (1-\gamma\sigma_n\delta_n)(\|x_n-x_{\sigma_n}\|^2+2\delta_{n-1}\langle \mathcal{A}x_{n-1}-\mathcal{A}x_n, x_n-x_{\sigma_n}\rangle \\
&\ -\vartheta\|x_{n-1}-x_{\sigma_n}\|^2+\frac{1}{2}\|x_{n-1}-x_n\|^2) \\
&\ +3p_n\left(\frac{(1-p_n)}{3pp_n^2}(p_n+p)\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2M^2\right) \\
&\ +\left(\frac{(1-p_n)p}{4}+p_n(1-2\varpi)+\frac{p_n(1-2\varpi)}{4}-\frac{p_n}{2}\right. \\
&\ \left.+(-\varpi+(1+\sqrt{2})\vartheta)\right)\|x_n-x_{n+1}\|^2.
\end{aligned} \tag{3.3.30}$$

Let

$$\begin{cases}
w_n=\left(\frac{(1-p_n)p}{4}+p_n(1-2\varpi)+\frac{p_n(1-2\varpi)}{4}-\frac{p_n}{2}+\left(-\varpi+((1+\sqrt{2})\vartheta)\right)\right)\\
v_n=\frac{(1-p_n)}{3pp_n^2}(p_n+p)\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2M^2\\
u_n=\|x_n-x_{\sigma_n}\|^2+2\delta_{n-1}\langle \mathcal{A}x_{n-1}-\mathcal{A}x_n, x_n-x_{\sigma_n}\rangle-\vartheta\|x_{n-1}-x_{\sigma_n}\|^2+\frac{1}{2}\|x_{n-1}-x_n\|^2;
\end{cases}$$

then (3.3.30) reduces to

$$(1-p_n)u_{n+1} \quad\le\quad (1-4p_n)\,u_n+3p_nv_n+w_n\|x_n-x_{n+1}\|^2 \tag{3.3.31}$$

Now,

$$\begin{aligned}
\lim_{n\to\infty} w_n &= \lim_{n\to\infty}\left(\frac{(1-p_n)p}{4}+p_n(1-2\varpi)+\frac{p_n(1-2\varpi)}{4}-\frac{p_n}{2}+\left(-\varpi+((1+\sqrt{2})\vartheta)\right)\right)\\
&= \frac{p}{4}+\left(-\varpi+((1+\sqrt{2})\vartheta)\right),\\
&\quad < \frac{p}{4}-\varpi+3\vartheta\le\frac{p}{4}-\varpi+3\times\frac{\varpi}{6}=\frac{p}{4}-\frac{\varpi}{2}.
\end{aligned}$$

Take $p=\varpi$ then $\frac{p}{4}-\frac{\varpi}{2}=\frac{-\varpi}{4}<\frac{-\varpi}{8}$, thus there exists $n_3\in\mathbb{N}, n_3\ge n_2\ge n_1\ge n_0$, such that $w_n<\frac{-\varpi}{8}\ \ \forall n\ge n_3$. From Inequality (3.3.31), we obtain

$$(1-p_n)u_{n+1}\le(1-4p_n)\,u_n+3p_nv_n-\frac{\varpi}{8}\|x_n-x_{n+1}\|^2 \ \text{ for all } n\ge n_3. \tag{3.3.32}$$

Dividing through (3.3.32) by the coefficient of u_{n+1} and take $e_n=\frac{3p_n}{(1-p_n)}$ yields the following:

$$u_{n+1}\le(1-e_n)u_n+e_nv_n-\frac{\varpi}{8(1-p_n)}\|x_n-x_{n+1}\|^2 \ \ \forall \ \ n\ge n_3. \tag{3.3.33}$$

By Asuumption 3.3.1(b) and the definition of δ_n, we estimate the limit of $\{v_n\}$ as follows:

$$\begin{aligned}\lim_{n\to\infty} v_n &= \lim_{n\to\infty}\frac{(1-p_n)}{3pp_n^2}(p_n+p)\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2 M^2\\ &= \lim_{n\to\infty}\frac{(1-0.25\sigma_n)}{3p(0.25)^2}(0.25\sigma_n+p)\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n^2}\right)^2 M^2\\ &= 0.\end{aligned}$$

Hence, $\limsup v_{n_i}=0$.

Next, we show that $\{u_n\}$ is nonnegative and bounded.

$$\begin{aligned}u_n &= \|x_n-x_{\sigma_n}\|^2+2\delta_{n-1}\langle \mathcal{A}x_{n-1}-\mathcal{A}x_n, x_n-x_{\sigma_n}\rangle-\vartheta\|x_{n-1}-x_{\sigma_n}\|^2\\ &\quad +\frac{1}{2}\|x_{n-1}-x_n\|^2\\ &\geq \|x_n-x_{\sigma_n}\|^2-\delta_{n-1}L_{\mathcal{A}}\left(\|x_n-x_{n-1}\|^2+\|x_n-x_{\sigma_n}\|^2\right)-\vartheta\|x_{n-1}-x_{\sigma_n}\|^2\\ &\quad +\frac{1}{2}\|x_{n-1}-x_n\|^2\\ &\geq (1-\delta_{n-1}L_{\mathcal{A}})\|x_n-x_{\sigma_n}\|^2+(\frac{1}{2}-\delta_{n-1}L_{\mathcal{A}})\|x_{n-1}-x_n\|^2\\ &\quad -\vartheta\|x_{n-1}-x_{\sigma_n}\|^2.\end{aligned} \tag{3.3.34}$$

By (2.1.3), we have

$$\begin{aligned}\|x_{n-1}-x_{\sigma_n}\|^2 &= \|(x_{n-1}-x_n)+(x_n-x_{\sigma_n})\|^2\\ &= \|x_{n-1}-x_n\|^2+\|x_n-x_{\sigma_n}\|^2+2\langle x_{n-1}-x_n, x_n-x_{\sigma_n}\rangle\\ &\leq 2\|x_{n-1}-x_n\|^2+2\|x_n-x_{\sigma_n}\|^2.\end{aligned} \tag{3.3.35}$$

Now, we substitue (3.3.35) in (3.3.34)

$$\begin{aligned}u_n &\geq (1-\delta_{n-1}L_{\mathcal{A}})\|x_n-x_{\sigma_n}\|^2+(\frac{1}{2}-\delta_{n-1}L_{\mathcal{A}})\|x_{n-1}-x_n\|^2\\ &\quad -2\vartheta(\|x_{n-1}-x_n\|^2+\|x_n-x_{\sigma_n}\|^2)\\ &\geq (\frac{1}{2}-\delta_{n-1}L_{\mathcal{A}}-2\vartheta)\left(\|x_{n-1}-x_n\|^2+\|x_n-x_{\sigma_n}\|^2\right)\\ &\geq (\frac{1}{2}-\frac{1}{2}+\varpi-2\times\frac{\varpi}{6})\left(\|x_{n-1}-x_n\|^2+\|x_n-x_{\sigma_n}\|^2\right)\\ &= \frac{2\varpi}{3}\left(\|x_{n-1}-x_n\|^2+\|x_n-x_{\sigma_n}\|^2\right)>0.\end{aligned} \tag{3.3.36}$$

It is clear that $\{v_n\}$ is bounded. Thus, there exists $N>0$ such that

$$u_{n+1}\leq(1-e_n)u_n+e_nN \quad \forall\, n\geq n_3,$$

which by [64, Lemma 3.1] implies that $\{u_n\}$ is bounded. Lemma 3.2.2 demands that $\limsup v_{n_i}\leq 0$ for every subsequence $\{u_{n_i}\}$ of $\{u_n\}$ satisfying $\liminf_{i\to\infty}(u_{n_i+1}-u_{n_i})\geq 0$. To fulfill this demand, let $\{u_{n_i}\}$ be a subsequence of $\{u_n\}$ satisfying $\liminf_{i\to\infty}(u_{n_i+1}-u_{n_i})\geq 0$. Then, since $\lim_{n\to\infty} e_n=0$, we get from (3.3.33) that

$$
\begin{aligned}
\limsup_{i\to\infty} \frac{\varpi}{8\left(1-0.25\gamma\sigma_{n_i}\delta_{n_i}\right)}\|x_{n_i}-x_{n_i+1}\|^2 &\leq \limsup_{i\to\infty}\left((u_{n_i}-u_{n_i+1})+e_{n_i}(v_{n_i}-u_{n_i})\right),\\
&= -\liminf_{i\to\infty}((u_{n_i}-u_{n_i+1})),\\
&\leq 0.
\end{aligned}
\tag{3.3.37}
$$

$$
\text{But,}\quad \lim_{i\to\infty}\frac{\varpi}{8\left(1-0.25\gamma\sigma_{n_i}\delta_{n_i}\right)} = \frac{\varpi}{8} > 0.
$$

It follows from (3.3.37) that

$$
\lim_{i\to\infty}\|x_{n_i}-x_{n_i+1}\| = 0. \tag{3.3.38}
$$

It is obvious that $e_n \in (0,1)$ and that $\sum_{n=1}^{\infty} e_n = \infty$. So using Lemma 3.2.2 in (3.3.33), we conclude that

$$
\lim_{n\to\infty} u_n = 0. \tag{3.3.39}
$$

Using (3.3.36), we get

$$
\lim_{n\to\infty}\|x_n - x_{\sigma_n}\| = 0. \tag{3.3.40}
$$

Using Assumption 3.3.1 (a), Proposition 3.3.2(iii) we derive that $\|x_{\sigma_n}-z^*\| \to 0$ as $n\to\infty$, where z^* solves BVIP (1.1.1) - (1.1.2). Therefore, it follows from (3.3.40) that

$$
\|x_n - z^*\| \leq \|x_n - x_{\sigma_n}\| + \|x_{\sigma_n} - z^*\| \to 0,\quad n\to\infty.
$$

We conclude that $\{x_n\}$ converges strongly to z^* as $n\to\infty$.

□

Next, we propose an algorithm with variable inertial parameter and the corresponding strong convergence theorem.

Algorithm 3.3.6. *Inertial regularised forward-reflected-backward method with variable inertial parameter*

Step 0: *$\epsilon \in \left(0,\frac{1}{4}\right)$, and choose sequence $\{\sigma_n\}$ in $(0,1)$. For arbitrary $x_0, x_1 \in H$. Set $n := 1$.*

Step 1: *Compute*

$$
x_{n+1} = P_C(m_n - \delta_n\mathcal{A}x_n - k_n - \delta_n\sigma_n\mathcal{T}x_n)\quad \forall n\geq 1,
$$

where

$$
\begin{cases}
m_n = x_n + \vartheta_n(x_n - x_{n-1})\\
k_n = \delta_{n-1}(1-\gamma\delta_n\sigma_n)(\mathcal{A}x_n - \mathcal{A}x_{n-1})
\end{cases}
$$

and

$$
\delta_n \in \left[\varpi, \min\left\{\frac{1-2\varpi}{2L_{\mathcal{A}}},\frac{1}{\gamma}\right\}\right] \text{ for some } \varpi > \quad 0.
$$

Step 2: *Set $n := n + 1$ and go to* **Step 1.**

Theorem 3.3.7. *Let $\{x_n\}$ be the sequence generated by Algorithm 3.3.6. If $0 \leq \vartheta_n \leq \vartheta_{n+1} \leq \vartheta < \min\left\{\frac{\varpi}{6}, \frac{\frac{1}{2}-\epsilon}{2}\right\}$, then x_n converges strongly to z^*.*

Proof. Using ϑ_n instead of ϑ in the proof of Lemma 3.3.4 up to (3.3.19), we get

$$\begin{aligned}
\|x_{n+1} - x_{\sigma_n}\|^2 \quad \leq \quad & \left(1 - \frac{3\delta_n\sigma_n\gamma}{2}\right)\|x_n - x_{\sigma_n}\|^2 \\
& +2\delta_{n-1}(1 - \gamma\delta_n\sigma_n)\langle \mathcal{A}x_{n-1} - \mathcal{A}x_n, x_n - x_{\sigma_n}\rangle \\
& -\vartheta_n(1 - \gamma\delta_n\sigma_n)\|x_{n-1} - x_{\sigma_n}\|^2 + \frac{1}{2}(1 - \gamma\delta_n\sigma_n)\|x_{n-1} - x_n\|^2 \\
& +(-\varpi + ((1 + \sqrt{2})\vartheta_n))\|x_{n+1} - x_n\|^2 \\
& +(-\varpi(1 - \gamma\delta_n\sigma_n) + \sqrt{2}\vartheta_n)\|x_{n-1} - x_n\|^2 \\
& -\gamma\delta_n\sigma_n\vartheta_n\|x_{n-1} - x_{\sigma_n}\|^2 - 2\delta_n\langle \mathcal{A}x_{n+1} - \mathcal{A}x_n, x_{\sigma_n} - x_{n+1}\rangle \\
& +\vartheta_n\|x_n - x_{\sigma_n}\|^2 - \frac{1}{2}\|x_n - x_{n+1}\|^2. \qquad (3.3.41)
\end{aligned}$$

Since $\vartheta_n \leq \vartheta_{n+1}$, we have

$$\begin{aligned}
\|x_{n+1} - x_{\sigma_n}\|^2 \quad \leq \quad & \left(1 - \frac{3\delta_n\sigma_n\gamma}{2}\right)\|x_n - x_{\sigma_n}\|^2 \\
& +2\delta_{n-1}(1 - \gamma\delta_n\sigma_n)\langle \mathcal{A}x_{n-1} - \mathcal{A}x_n, x_n - x_{\sigma_n}\rangle \\
& -\vartheta_n(1 - \gamma\delta_n\sigma_n)\|x_{n-1} - x_{\sigma_n}\|^2 + \frac{1}{2}(1 - \gamma\delta_n\sigma_n)\|x_{n-1} - x_n\|^2 \\
& +(-\varpi + ((1 + \sqrt{2})\vartheta_n))\|x_{n+1} - x_n\|^2 \\
& +(-\varpi(1 - \gamma\delta_n\sigma_n) + \sqrt{2}\vartheta_n)\|x_{n-1} - x_n\|^2 \\
& -\gamma\delta_n\sigma_n\vartheta_n\|x_{n-1} - x_{\sigma_n}\|^2 - 2\delta_n\langle \mathcal{A}x_{n+1} - \mathcal{A}x_n, x_{\sigma_n} - x_{n+1}\rangle \\
& +\vartheta_{n+1}\|x_n - x_{\sigma_n}\|^2 - \frac{1}{2}\|x_n - x_{n+1}\|^2. \qquad (3.3.42)
\end{aligned}$$

Since

$$\begin{aligned}
\lim_{n\to\infty}(-\varpi(1 - \gamma\delta_n\sigma_n) + \sqrt{2}\vartheta_n) \quad & \leq \quad \lim_{n\to\infty}(-\varpi(1 - \gamma\delta_n\sigma_n) + \sqrt{2}\vartheta) \\
& = \quad -\varpi + \sqrt{2}\vartheta \\
& < \quad -\varpi + 2\vartheta < -\varpi + 2 \times \frac{\varpi}{6} \\
& = \quad \frac{-2\varpi}{3} < 0,
\end{aligned}$$

thus there exists $n_4 \in \mathbb{N}$, $n_4 \geq n_3$ such that $(-\varpi(1 - \gamma\delta_n\sigma_n) + \sqrt{2}\vartheta_n) < 0 \quad \forall\, n \geq n_4$, also

$\gamma\delta_n\sigma_n\vartheta > 0$. Therefore, it follows from (3.3.42) that

$$\begin{aligned}
&\|x_{n+1} - x_{\sigma_n}\|^2 + 2\delta_n\langle \mathcal{A}x_{n+1} - \mathcal{A}x_n, x_{\sigma_n} - x_{n+1}\rangle - \vartheta_{n+1}\|x_n - x_{\sigma_n}\|^2 \\
&+\frac{1}{2}\|x_n - x_{n+1}\|^2 \\
&\leq \left(1 - \frac{3\delta_n\sigma_n\gamma}{2}\right)\|x_n - x_{\sigma_n}\|^2 + 2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_{n-1} - \mathcal{A}x_n, x_n - x_{\sigma_n}\rangle \\
&-\vartheta_n(1-\gamma\delta_n\sigma_n)\|x_{n-1} - x_{\sigma_n}\|^2 + \frac{1}{2}(1-\gamma\delta_n\sigma_n)\|x_{n-1} - x_n\|^2 \\
&+(-\varpi + ((1+\sqrt{2})\vartheta_n))\|x_{n+1} - x_n\|^2, \quad \forall n \geq n_4. \qquad (3.3.43)
\end{aligned}$$

Inequality (3.3.43) is a variant of Inequality (3.3.20). By following the arguments from (3.3.20) up to (3.3.27), we have

$$\begin{aligned}
&(1-p_n)(\|x_{n+1} - x_{\sigma_{n+1}}\|^2 + 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1} - x_{\sigma_{n+1}}\rangle \\
&-\vartheta_{n+1}\|x_n - x_{\sigma_n}\|^2 + \frac{1}{2}\|x_n - x_{n+1}\|^2) \\
&\leq \|x_{n+1} - x_{\sigma_n}\|^2 + 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle \\
&-\vartheta_n\|x_n - x_{\sigma_n}\|^2 + \frac{1}{2}\|x_n - x_{n+1}\|^2 \\
&+\frac{(1-p_n)}{p}\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2 M^2 + \left(\frac{1-p_n}{p_n}\right)\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2 M^2 \\
&+\left(\frac{(1-p_n)p}{4} + p_n(1-2\varpi) + \frac{p_n(1-2\varpi)}{4} - \frac{p_n}{2}\right)\|x_n - x_{n+1}\|^2 \\
&+p_n\|x_n - x_{\sigma_n}\|^2 + p_n\vartheta_n\|x_n - x_{\sigma_n}\|^2. \qquad (3.3.44)
\end{aligned}$$

We estimate $p_n\vartheta_n\|x_n - x_{\sigma_n}\|^2$ having in mind that $p_n\epsilon > 0$;

$$\begin{aligned}
p_n\vartheta_n\|x_n - x_{\sigma_n}\|^2 &\leq \frac{1}{2}p_n\left(\frac{1}{2} - \epsilon\right)\|x_n - x_{\sigma_n}\|^2 \\
&\leq p_n\|x_n - x_{\sigma_n}\|^2. \qquad (3.3.45)
\end{aligned}$$

Put (3.3.45) in (3.3.44) to obtain the following inequality:

$$\begin{aligned}
&(1-p_n)(\|x_{n+1} - x_{\sigma_{n+1}}\|^2 + 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1} - x_{\sigma_{n+1}}\rangle \\
&-\vartheta_{n+1}\|x_n - x_{\sigma_n}\|^2 + \frac{1}{2}\|x_n - x_{n+1}\|^2) \\
&\leq \|x_{n+1} - x_{\sigma_n}\|^2 + 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle - \vartheta_n\|x_n - x_{\sigma_n}\|^2 \\
&+\frac{1}{2}\|x_n - x_{n+1}\|^2 + \frac{(1-p_n)}{p}\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2 M^2 \\
&+\left(\frac{1-p_n}{p_n}\right)\left(\frac{\sigma_{n+1}-\sigma_n}{\sigma_n}\right)^2 M^2 + 2p_n\|x_n - x_{\sigma_n}\|^2 \\
&+\left(\frac{(1-p_n)p}{4} + p_n(1-2\varpi) + \frac{p_n(1-2\varpi)}{4} - \frac{p_n}{2}\right)\|x_n - x_{n+1}\|^2. \qquad (3.3.46)
\end{aligned}$$

Using similar argument from (3.3.27) to the end of the proof of Theorem 3.3.5, we have that $\{x_n\}$ converges strongly to z^*. $\square$

3.4 Numerical Examples

In this section, we discuss the numerical behavior of Algorithm 1.3.2 in both finite and infinite dimensional Hilbert spaces. We compare our algorithm with Algorithms 2.3.7, 2.3.8 and 2.3.9. All codes are written in Matlab 2023 (b) and performed on a personal computer with an Intel(R) Core(TM) i5-10210U CPU at 1.60GHz and 8.00 Gb-RAM. In Tables 1-2, Iter. means the number of iterations while CPU means the CPU time in seconds. In our computations, we choose $\epsilon = 0.1$ *and* $\delta_n = \min\left\{\frac{1-2\varpi}{2L_{\mathcal{A}}}, \frac{1}{\gamma}\right\}$ *with* $\varpi = 0.1$. *Also, we take* $\sigma_n = \frac{1}{\sqrt{n+1}}$ *and* $\vartheta_n = \vartheta = 0.9\min\left\{\frac{\varpi}{6}, \frac{\frac{1}{2}-\epsilon}{2}\right\}$.

Example 3.4.1. *Let* $H = \mathbb{R}^N$. *Define* $\mathcal{A} : \mathbb{R}^N \to \mathbb{R}^N$ *by* $\mathcal{A}(x) = Mx + q$, *where the matrix* M *is formed as:* $M = V\sum V'$, *where* $V = I - \frac{2vv'}{\|v\|^2}$ *and* $\sum = diag(\sigma_{11}, \sigma_{12}, \cdots, \sigma_{1N})$ *are the Householder and the diagonal matrix, and*

$$\sigma_{1j} = \cos\frac{j\pi}{N+1} + 1 + \frac{\cos\frac{\pi}{N+1} + 1 - \widehat{C}(\cos\frac{N\pi}{N+1} + 1)}{\widehat{C} - 1}, \qquad j = 1, 2, \cdots, N,$$

with $\widehat{C}$ *been the present condition number of* M *([43, Example 5.2]). In the numerical computation, we choose* $\widehat{C} = 10^4$, $q = 0$ *and uniformly take the vector* $v \in \mathbb{R}^N$ *in* $(-1, 1)$. *Thus,* $\mathcal{A}$ *is monotone and Lipschitz continuous with* $L_{\mathcal{A}} = ||M||$ *(see [43]). We set* $\mathcal{Q} = \{x \in \mathbb{R}^N : ||x|| \leq 1\}$. *The projection onto* $\mathcal{Q}$ *is effectively computed in Matlab. We define* $\mathcal{T} : \mathbb{R}^N \to \mathbb{R}^N$ *by* $\mathcal{T}x = Dx + p$, *where* D *is a symmetric and positive-definite matrix of size* $N \times N$ *and* p *is a vector in* $\mathbb{R}^N$. *Clearly,* $\mathcal{T}$ *is* $||D||$*-Lipschitz continuous. Moreover, since* D *is symmetric and positive-definite, we have that* $\langle Dx - Dy\rangle = \langle D(x-y)\rangle \geq ||D||^{-1}||x-y||^2$. *Hence,* $\mathcal{T}$ *is strongly monotone.*

Since the solution of the problem is unknown, we define the sequence $TOL_n := ||x_{n+1} - x_n||^2$, *and use the stopping criterion* $TOL_n < \varepsilon$ *for the iterative processes, where* ε *is the predetermined error. Moreover, we consider different scenarios of the problem's dimensions. That is,* $N = 100, 300, 500, 1000$, *with starting points* $x_1 = (1, 1, \ldots, 1)'$ *and* $x_0 = (0, 0, \ldots, 0)'$. *For this example, we take* $\varepsilon = 10^{-5}$ *as the stopping criterion and obtain the numerical results reported in Table 1 and Figure 3.1.*

Table 1: Example 4.4.1: Comparison of algorithms with $\epsilon = 10^{-5}$

Algorithms	$N = 100$		$N = 300$		$N = 500$		$N = 1000$	
	CPU	Iter	CPU	Iter	CPU	Iter	CPU	Iter
Algorithm 3.3.6	5.0279	992	5.0365	2574	4.0435	2657	3.0259	769
Algorithm 2.3.7	13.2116	16093	13.2387	16130	12.1624	12583	10.1297	11724
Algorithm 2.3.8	10.1873	13376	10.3471	19527	9.3787	20073	8.1941	12948
Algorithm 2.3.9	9.1787	13051	9.2925	19735	8.2901	19435	8.1623	12922

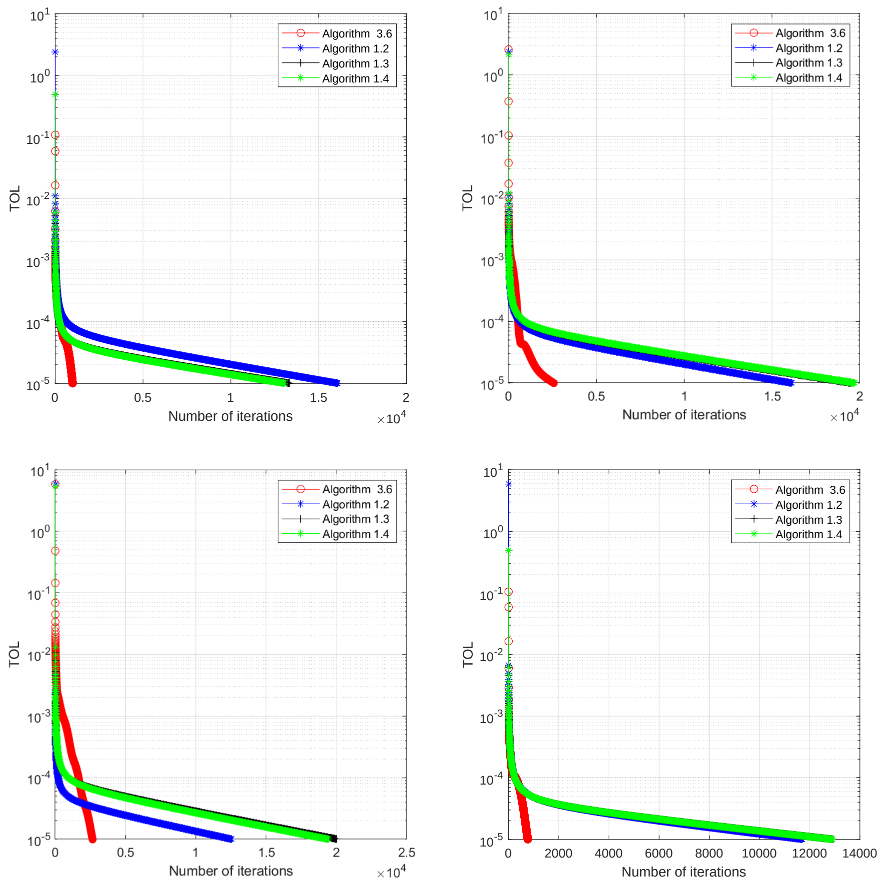

Figure 3.1: The behavior of TOL_n with $\varepsilon = 10^{-5}$ for Example 4.4.1: Top Left: $N = 100$; Top Right: $N = 300$; Bottom Left: $N = 500$; Bottom Right: $N = 1000$.

Example 3.4.2. *Let* $H = (l_2(\mathbb{R}), \ ||.||_{l_2})$, *where* $l_2(\mathbb{R}) := \{x = (x_1, x_2, x_3, \dots), \ x_i \in \mathbb{R} : \sum_{i=1}^{\infty} |x_i|^2 < \infty\}$ *and* $||x||_{l_2} := \left(\sum_{i=1}^{\infty} |x_i|^2\right)^{\frac{1}{2}}, \quad \forall x \in l_2(\mathbb{R})$.

Let $\mathcal{Q} = \{x \in l_2(\mathbb{R}) : |x_i| \leq \frac{1}{i}, \ i = 1, 2, 3, \dots\}$. *Thus, we have explicit formula for* $P_{\mathcal{Q}}$. *Now, define the operator* $\mathcal{A} : l_2(\mathbb{R}) \to l_2(\mathbb{R})$ *by*

$$\mathcal{A}x := \left(\frac{x_1 + |x_1|}{2}, \frac{x_2 + |x_2|}{2}, \dots, \frac{x_i + |x_i|}{2}, \dots\right), \ \forall x \in \ell_2(\mathbb{R}).$$

Then $\mathcal{A}$ *is Lipschitz continuous and monotone with Lipschitz constants* $L_{\mathcal{A}} = 1$. *Furthermore, define the mapping* $\mathcal{T} : l_2(\mathbb{R}) \to l_2(\mathbb{R})$ *by* $\mathcal{T}x = x - x_0$. *Then,* $\mathcal{T}$ *is strongly monotone and Lipschitz continuous.*
For this example, we take $\varepsilon = 10^{-8}$ *as the stopping criterion and choose the starting points as follows:*

Case 1: *Take* $x_1 = (1, \frac{1}{2}, \frac{1}{3}, \cdots)$ *and* $x_0 = (\frac{1}{2}, \frac{1}{5}, \frac{1}{10}, \cdots)$.

Case 2: *Take* $x_1 = (\frac{1}{2}, \frac{1}{5}, \frac{1}{10}, \cdots)$ *and* $x_0 = (1, \frac{1}{2}, \frac{1}{3}, \cdots)$.

Case 3: *Take* $x_1 = (1, \frac{1}{4}, \frac{1}{9}, \cdots)$ *and* $x_0 = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \cdots)$.

Case 4: *Take* $x_1 = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \cdots)$ *and* $x_0 = (1, \frac{1}{4}, \frac{1}{9}, \cdots)$.
The numerical results are reported in Table 2 and Figure 3.2.

Table 2: Example 4.4.2: Comparison of algorithms with $\epsilon = 10^{-8}$

Algorithms	Case 1		Case 2		Case 3		Case 4	
	CPU	Iter	CPU	Iter	CPU	Iter	CPU	Iter
Algorithm 3.3.6	1.0260	333	1.0171	537	1.0212	640	1.0184	806
Algorithm 2.3.7	4.0104	668	4.0125	1076	4.0142	1280	4.0182	1612
Algorithm 2.3.8	4.0145	673	4.0187	1047	4.0189	1245	4.0252	1569
Algorithm 2.3.9	3.0113	582	3.0130	973	2.0153	1170	3.0173	1490

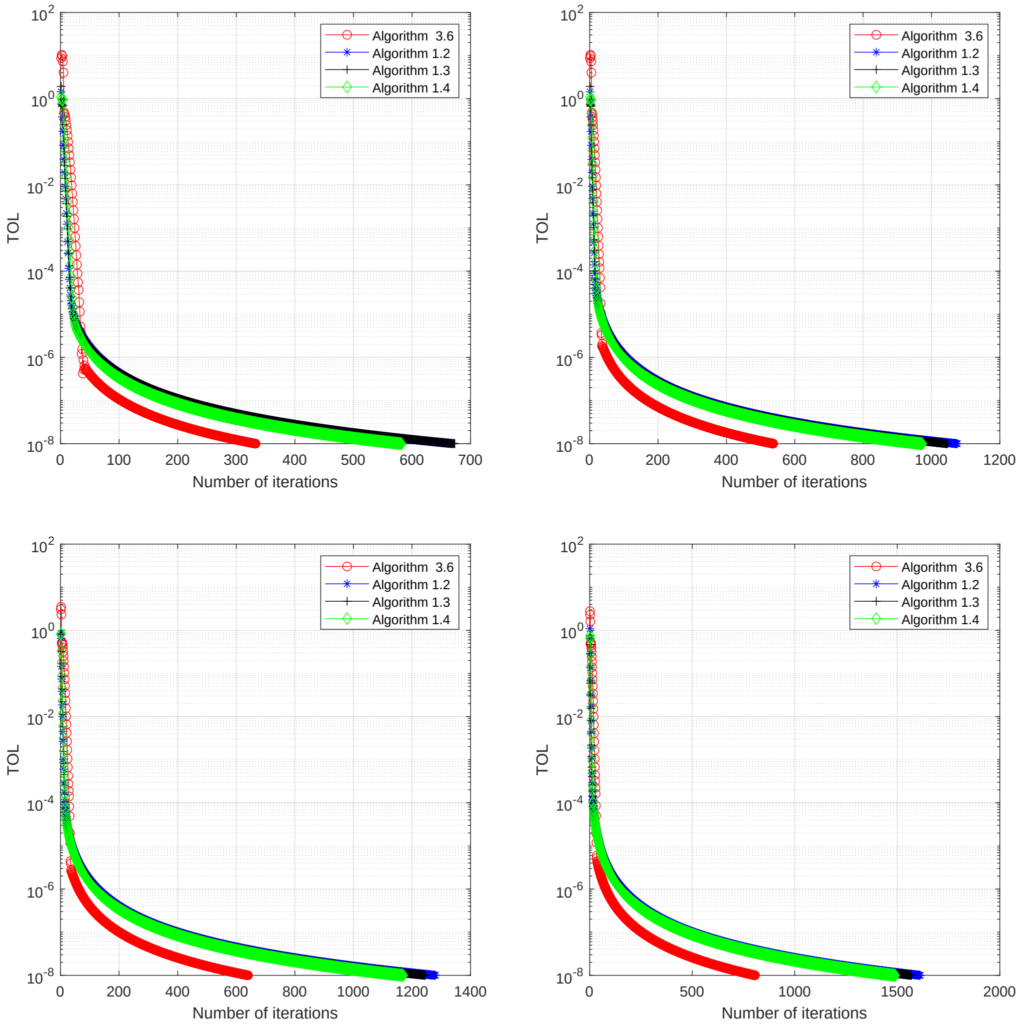

Figure 3.2: The behavior of TOL_n with $\varepsilon = 10^{-8}$ for Example 4.4.2: Top Left: **Case 1**; Top Right: **Case 2**; Bottom Left: **Case 3**; Bottom Right: **Case 4**.

3.5 Conclusion

We have studied two inertial forward-reflected-backward methods for solving BVIP (1.1.1) - (1.1.2)*. The first method uses a constant inertial parameter while the second one uses a variable inertial parameter. We proved that both methods converge strongly to a solution of the BVIP* (1.1.1) - (1.1.2)*. Our methods are modifications of the forward-reflected-backward splitting method in [67]. As in [67], they require only a single projection onto the feasible set* $\mathcal{Q}$ *and one functional evaluation of* $\mathcal{A}$*. Our numerical experiments show that our methods perform better than related methods in the literature.*

Chapter 4

4.1 Introduction

In this chapter, we introduce and study the modified forward-reflected-backward algorithm for approximating a solution of BVIP (1.1.1)-(1.1.2)*. The proposed method is a modification of the method studied in Chapter Three. We remark that the method in Chapter Three works with prior knowledge of the Lipschitz constant, which is often difficult to estimate when it is unknown. The method studied here replaces the stepsizes and the single projection onto a feasible set of the method in Chapter Three with a self-adaptive non-monotonic stepsize and a single projection onto a constructive half-space, respectively. This simple adjustment leads to a remarkable improvement in the performance of this method. We also show that the sequences generated iteratively by this method converge strongly to the solution of BVIP* (1.1.1)-(1.1.2)*. Moreover, we provide the numerical experiments of this method as it compares with other recent methods in the literature.*

4.2 Preliminaries

This section captures definitions, lemmas, propositions and ideas that are paramount in the establishment of the main results of this chapter; we start with some basic definitions.

Definition 4.2.1. *An operator $\mathcal{Q} : H \to \mathbb{R}$ is called convex, if for all $\lambda \in [0, 1]$ and x, $y \in H$,*

$$\mathcal{Q}(\lambda x + (1 - \lambda)y) \leq \lambda \mathcal{Q}(x) + (1 - \lambda)\mathcal{Q}(y)$$

Definition 4.2.2. *A convex function $\mathcal{Q} : H \to \mathbb{R}$ is subdifferentiable at a point $x \in H$ if the set*

$$\partial \mathcal{Q}(x) = \{v \in H : \mathcal{Q}(v) \geq \mathcal{Q}(x) + \langle v, y - x\rangle, \forall y \in H\} \tag{4.2.1}$$

is nonempty, where each element in $\partial \mathcal{Q}(x)$ is called a subgradient of $\mathcal{Q}$ at x, $\partial \mathcal{Q}(x)$ is called the subdifferential of $\mathcal{Q}$ at x and the inequality in (4.2.1) *is called the subdifferential inequality of $\mathcal{Q}$ at x. We say that $\mathcal{Q}$ is subdifferentiable on H if $\mathcal{Q}$ is subdifferentiable at each $x \in H$.*

Lemma 4.2.3. *If u and v are non-negative numbers, then*

(i) $uv \leq \frac{u^2}{2p} + \frac{pv^2}{2} \quad \forall p > 0$ ***(Young's Inequality)***

(ii) $(u+v)^2 \leq (2+\sqrt{2})u^2 + \sqrt{2}v^2$

Lemma 4.2.4 ([100])**.** *Let $\{e_n\}$ be a sequence of non-negative real numbers. Suppose that*

$$e_{n+1} \leq (1-v_n)e_n + w_n \quad \forall \quad n \geq 0,$$

where $\{v_n\}$ be a sequence of numbers in (0,1) and $w_n \in \mathbb{R}$ satisfy the conditions: $\lim_{n\to\infty} v_n = 0$, $\sum_{n=1}^{\infty} v_n = \infty$, *and* $\limsup_{n\to\infty} \frac{w_n}{v_n} \leq 0$. *Then* $\lim_{n\to\infty} e_n = 0$.

Lemma 4.2.5 ([44], Theorem 4.1)**.** *Assume that the solution set VI($\mathcal{A}$, C) of the classical VIP (1.1.2) is nonempty and C is defined as $C := \{x \in H : \mathcal{Q}(x) \leq 0\}$, where $\mathcal{Q} : H \to \mathbb{R}$ is a continuously differentiable and convex function. Let y^* be in C. Then $y^* \in V(\mathcal{A}, C)$ if and only if either*

(i) $\mathcal{A}y^* = 0$, *or*

(ii) $y^* \in \partial C$ *and there exists $\rho > 0$ such that $\mathcal{A}y^* = -\rho\mathcal{Q}'(y^*)$, where ∂C denotes the boundary of C.*

4.3 Main Results

Assumption 4.3.1. *The following assumptions will be used in the convergence analysis.*

(a) $\sigma_n \in (0,1)$, $\lim_{n\to\infty} \sigma_n = 0$ *and* $\sum_{n=1}^{\infty} \sigma_n = \infty$.

(b)
$$\lim_{n\to\infty} \frac{|\sigma_{n+1} - \sigma_n|}{\sigma_n^2} = 0.$$

(c) *$\mathcal{A}$ is monotone and $L_{\mathcal{A}}$-Lipschitz continuous, $\mathcal{T}$ is γ-strongly monotone and $L_{\mathcal{T}}$-Lipschitz continuous.*

(d) *The feasible set C is defined as*

$$C = \{x \in H : \mathcal{Q}(x) \leq 0\},$$

where $\mathcal{Q}$: $H \to \mathbb{R}$ is continuously differentiable and convex such that $\mathcal{Q}'(.)$ is Lipschitz continuous with a constant $K > 0$.

(e) *VI(C, $\mathcal{A}$) is nonempty.*

Consider the following Regularized Variational Inequality Problem (RVIP) (1.1.1)- (1.1.2):

$$\text{find } \ x^* \in C \quad \text{such that } \langle \mathcal{A}x^* + \sigma \mathcal{T}x^*, x - x^* \rangle \geq 0 \qquad \forall x \in C, \tag{4.3.1}$$

where $\sigma > 0$ is a regularization parameter. We remark that (3.3.1) has a unique solution, it follows from the fact that $\mathcal{A}$ is monotone and Lipschitz continuous, and $\mathcal{T}$ is strongly monotone and Lipschitz continuous. We denote this unique solution of (3.3.1) by x_σ for each $\sigma > 0$. Since $VI(C, \mathcal{A})$ is nonempty, closed and convex then the unique solution of BVIP (1.1.1) - (1.1.2) is guaranteed. Let this unique solution be z^*.

Proposition 4.3.2 ([46]). *Let $\{x_\sigma\}_{\sigma \in (0,1)}$ be the solution net of (3.3.1), then*

(i) $\|x_\sigma\| \leq \|x^*\| + \frac{\|\mathcal{T}x^*\|}{\gamma}$ *for each $x^* \in VI(C, \mathcal{A})$, where γ is the modulus of strong monotonicity $\mathcal{T}$;*

(ii) Let x_{σ_1}, $x_{\sigma_2} \in \{x_\sigma\}_{\sigma \in (0,1)}$, then $\|x_{\sigma_1} - x_{\sigma_2}\| \leq \frac{|\sigma_1 - \sigma_2|}{\sigma_1} M$, where $M > 0$ is a constant;

(iii) $\|x_\sigma - z^*\| \to 0$ *as* $\sigma \to \infty$.

Algorithm 4.3.3. *Modified forward-reflected-backward algorithm.*

Step 0: *Let $\delta_0, \delta_1 > 0$, $\alpha \in (\epsilon, \frac{1-2\epsilon}{2})$ with $\epsilon \in (0, \frac{1}{4})$. Choose sequence $\{\sigma_n\}$ in $(0, 1)$ and $\{a_n\}$ in $[0, \infty)$ such that $\sum_{n=1}^{\infty} a_n < \infty$. Let x_0, $x_1 \in H$ and $0 \leq \vartheta_n \leq \vartheta_{n+1} \leq \vartheta < \min\left\{\frac{\frac{1}{2}-\alpha}{2}, \frac{1-\beta-\frac{\rho_1 \beta}{3}}{3}\right\}$ for some $\beta, \rho_1 > 0$. Set $n := 1$.*

Step 1: *Construct the half-space*

$$C_n = \{x \in H : \mathcal{Q}x_n + \langle \mathcal{Q}'x_n, x - x_n \rangle \leq 0\}.$$

Then compute

$$x_{n+1} = P_{C_n}(m_n - \delta_n \mathcal{A}x_n - k_n - \delta_n \sigma_n \mathcal{T}x_n), \quad \forall n \geq 1,$$

where

$$\begin{cases} m_n = x_n + \vartheta_n(x_n - x_{n-1}) \\ k_n = \delta_{n-1}(1 - \gamma\delta_n\sigma_n)(\mathcal{A}x_n - \mathcal{A}x_{n-1}) \end{cases}$$

and

$$\delta_{n+1} = \begin{cases} \min\left\{\frac{\alpha\|x_{n+1} - x_n\|}{\|\mathcal{A}x_{n+1} - \mathcal{A}x_n\| + \|\mathcal{Q}'x_{n+1} - \mathcal{Q}'x_n\|}, \delta_n + a_n\right\}, & \text{if } \mathcal{A}x_n \neq \mathcal{A}x_{n+1} \\ \delta_n + a_n, & \text{otherwise.} \end{cases}$$

Step 2: *Set $n := n + 1$ and go to 1.*

Remark 4.3.4. *[104] Note that* $\lim_{n\to\infty}\delta_n = \bar{\delta}$, *where* $\bar{\delta} \in [\min\{\alpha(L_{\mathcal{A}}+k)^{-1}, \delta_1\}, \delta_1+\bar{a}]$ *with* $\bar{a} = \sum_{n=1}^{\infty} a_n$.

Lemma 4.3.5. *If* $x_{n+1} = x_n = x_{n-1}$ *for some* $n \geq M$, *then* x_M *is the solution to* (1.1.1).

Proof. If $x_{n+1} = x_n = x_{n-1}$, we obtain from Algorithm 4.3.3 that

$$\begin{aligned} x_n &= P_{C_n}(x_n - \delta_n \mathcal{A}x_n - \delta_n\sigma_n\mathcal{T}x_n) \\ &\iff \langle x_n - x_n + \delta_n\mathcal{A}x_n + \delta_n\sigma_n\mathcal{T}x_n, y - x_n\rangle \geq 0 \quad \forall y \in C \\ &\iff \langle \mathcal{A}x_n + \sigma_n\mathcal{T}x_n, y - x_n\rangle \geq 0. \end{aligned}$$

Therefore,

$$x_n = x_{\sigma_n} = x_{n+1} = x_{\sigma_{n+1}} = x_{\sigma_{n+2}} = \cdots$$

solve (3.3.1). If n tends to ∞, we have that $\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} x_{\sigma_n} = z^*$. This implies $x_M = z^*$ and by Proposition 3.3.2 iii, x_M is the solution (1.1.1) (see also [86]). $\square$

Lemma 4.3.6. *Let* $\{x_n\}$ *be a sequence generated by Algorithm 4.3.3. Then there exists* $n_0 \in \mathbb{N}$ *such that*

$$\begin{aligned} \|x_{n+1} - x_{\sigma_n}\|^2 \leq\; & (1 + \vartheta_n - \delta_n\sigma_n(2\gamma - aL_{\mathcal{T}}) + 3\alpha)\|x_n - x_{\sigma_n}\|^2 \\ & + \left(2\vartheta_n + 2\frac{\delta_{n-1}}{\delta_n}\alpha(1 - \gamma\delta_n\sigma_n)\right)\|x_n - x_{n-1}\|^2 - \vartheta_n\|x_{\sigma_n} - x_{n-1}\|^2 \\ & - \left(1 - \vartheta_n - \frac{\delta_n\sigma_n L_{\mathcal{T}}}{a} - 3\alpha\frac{\delta_n}{\delta_{n+1}} - \frac{\delta_{n-1}}{\delta_n}\alpha(1 - \gamma\delta_n\sigma_n)\right. \\ & \left. -2\delta_n\frac{\rho_1\alpha}{\delta_{n+1}}\right)\|x_{n+1} - x_n\|^2. \end{aligned} \tag{4.3.2}$$

where x_{σ_n} *is the solution to the RVIP* (4.3.1) *with* $\sigma = \sigma_n$.

Proof. By Proposition 2.1.22, we have

$$2\langle m_n - \delta_n\mathcal{A}x_n - k_n - \delta_n\sigma_n\mathcal{T}x_n - x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle \geq 0,$$

which gives

$$\begin{aligned} 0 \leq 2 & \langle m_n - x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle - 2\langle k_n, x_{n+1} - x_{\sigma_n}\rangle - 2\delta_n\sigma_n\langle\mathcal{T}x_n, x_{n+1} - x_{\sigma_n}\rangle \\ & -2\delta_n\langle\mathcal{A}x_n, x_{n+1} - x_{\sigma_n}\rangle. \end{aligned}$$

which implies

$$\begin{aligned} 0 \leq\; & 2\langle m_n - x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle - 2\langle k_n, x_{n+1} - x_{\sigma_n}\rangle - 2\delta_n\sigma_n\langle\mathcal{T}x_n, x_{n+1} - x_{\sigma_n}\rangle \\ & +2\delta_n\langle\mathcal{A}x_n - \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1}\rangle + 2\delta_n\langle\mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1}\rangle. \end{aligned} \tag{4.3.3}$$

Clearly x_{σ_n} and $x_{n+1} \in C_n$, so we have

$$\mathcal{Q}x_{\sigma_n} + \langle \mathcal{Q}'x_{\sigma_n}, x_{n+1} - x_{\sigma_n} \rangle \leq 0 \tag{4.3.4}$$

Since x_{σ_n} is a solution of RVIP (3.3.1), we get from Lemma 4.2.5 that $x_{\sigma_n} \in \partial C$ and there exists $\rho_1 > 0$ such that

$$\mathcal{Q}'x_{\sigma_n} = -\frac{\mathcal{A}x_{\sigma_n}}{\rho_1}. \tag{4.3.5}$$

Since $x_{\sigma_n} \in \partial C$, we have that

$$\mathcal{Q}x_{\sigma_n} = 0. \tag{4.3.6}$$

Using (4.3.5) and (4.3.6) in (4.3.4), we obtain

$$\langle \mathcal{A}x_{\sigma_n}, x_{n+1} - x_{\sigma_n} \rangle \geq 0,$$

implying that $x_{n+1} \in C$.
Hence,

$$2\delta_n \langle \mathcal{A}x_{\sigma_n} + \sigma_n \mathcal{T}x_{\sigma_n}, x_{n+1} - x_{\sigma_n} \rangle \geq 0. \tag{4.3.7}$$

Adding inequality (4.3.7) to the right of (4.3.3) gives

$$\begin{aligned} 0 \leq\ & 2\langle m_n - x_{n+1}, x_{n+1} - x_{\sigma_n} \rangle - 2\langle k_n, x_{n+1} - x_{\sigma_n} \rangle \\ & +2\delta_n \sigma_n \langle \mathcal{T}x_{\sigma_n} - \mathcal{T}x_n, x_{n+1} - x_{\sigma_n} \rangle \\ & +2\delta_n \langle \mathcal{A}x_n - \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1} \rangle + 2\delta_n \langle \mathcal{A}x_{\sigma_n}, x_{n+1} - x_{\sigma_n} \rangle \\ & +2\delta_n \langle \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1} \rangle. \end{aligned} \tag{4.3.8}$$

Now if $\mathcal{A}x_{\sigma_n} = 0$, it is very easy to obtain (4.3.2). However, if $\mathcal{A}x_{\sigma_n} \neq 0$, then we get from Lemma 4.2.5 that $x_{\sigma_n} \in \partial C$ and there exists $\rho_1 > 0$ such that $\mathcal{A}x_{\sigma_n} = -\rho_1 \mathcal{Q}'x_{\sigma_n}$. Since $x_{\sigma_n} \in \partial C$, $\mathcal{Q}x_{\sigma_n} = 0$.

Since $\mathcal{Q}$ is a differentiable convex function, it follows from (4.2.1) that

$$\mathcal{Q}x_{n+1} \geq \mathcal{Q}x_{\sigma_n} + \langle \mathcal{Q}'x_{\sigma_n}, x_{n+1} - x_{\sigma_n} \rangle,$$

which implies

$$\mathcal{Q}x_{n+1} \geq \frac{-1}{\rho_1} \langle Ax_{\sigma_n}, x_{n+1} - x_{\sigma_n} \rangle,$$

which can be rewritten as

$$\langle Ax_{\sigma_n}, x_{\sigma_n} - x_{n+1} \rangle \leq \rho_1 \mathcal{Q}x_{n+1}. \tag{4.3.9}$$

Since $x_{n+1} \in C_n$, we have

$$\mathcal{Q}x_n + \langle \mathcal{Q}'x_n, x_{n+1} - x_n \rangle \leq 0. \tag{4.3.10}$$

Again, since $\mathcal{Q}$ is differentiable convex function, we have

$$\langle \mathcal{Q}'x_{n+1}, x_n - x_{n+1} \rangle + \mathcal{Q}x_{n+1} \leq \mathcal{Q}x_n. \tag{4.3.11}$$

From (4.3.10) and (4.3.11), we have

$$\mathcal{Q}x_{n+1} \le \langle \mathcal{Q}'x_n - \mathcal{Q}'x_{n+1}, x_n - x_{n+1}\rangle \le \frac{\alpha}{\delta_{n+1}}\|x_n - x_{n+1}\|^2. \tag{4.3.12}$$

Combining (4.3.9) and (4.3.12), we have that

$$\langle \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1}\rangle \le \frac{\rho_1\alpha}{\delta_{n+1}}\|x_n - x_{n+1}\|^2. \tag{4.3.13}$$

Let $x_{\sigma_n} \in C \subset C_n$, we get

$$\mathcal{Q}x_{n+1} + \langle \mathcal{Q}'x_{n+1}, x_{\sigma_n} - x_{n+1}\rangle \le 0, \tag{4.3.14}$$

From (4.3.14), we obtain for ρ_2, $\delta_n > 0$

$$-2\delta_n\rho_2\mathcal{Q}x_{n+1} + 2\delta_n\rho_2\langle \mathcal{Q}'x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle \ge 0. \tag{4.3.15}$$

Add (4.3.15) to the right of (4.3.8) to obtain the following

$$\begin{aligned} 0 \;\le\; & 2\langle m_n - x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle - 2\langle k_n, x_{n+1} - x_{\sigma_n}\rangle \\ & + 2\delta_n\sigma_n\langle \mathcal{T}x_{\sigma_n} - \mathcal{T}x_n, x_{n+1} - x_{\sigma_n}\rangle \\ & + 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1}\rangle + 2\delta_n\langle \mathcal{A}x_{\sigma_n}, x_{n+1} - x_{\sigma_n}\rangle \\ & + 2\delta_n\langle \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1}\rangle \\ & - 2\delta_n\rho_2\mathcal{Q}x_{n+1} + 2\delta_n\rho_2\langle \mathcal{Q}'x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle. \end{aligned} \tag{4.3.16}$$

By Lemma 4.3.5, x_{n+1} is a solution and $\mathcal{A}x_{n+1} \ne 0$, then we get from Lemma 4.2.5 that $x_{n+1} \in \partial C$ and there exists $\rho_2 > 0$ such that $\mathcal{A}x_{n+1} = -\rho_2\mathcal{Q}'x_{n+1}$. So we have

$$\mathcal{Q}'x_{n+1} = -\frac{\mathcal{A}x_{n+1}}{\rho_2}. \tag{4.3.17}$$

Also, since $x_{n+1} \in \partial C$, $\mathcal{Q}x_{n+1} = 0$. Now, substituting (4.3.17) in (4.3.16), we obtain

$$\begin{aligned} 0 \;\le\; & 2\langle m_n - x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle - 2\langle k_n, x_{n+1} - x_{\sigma_n}\rangle \\ & + 2\delta_n\sigma_n\langle \mathcal{T}x_{\sigma_n} - \mathcal{T}x_n, x_{n+1} - x_{\sigma_n}\rangle \\ & + 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1}\rangle + 2\delta_n\langle \mathcal{A}x_{\sigma_n} - \mathcal{A}x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle \\ & + 2\delta_n\langle \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1}\rangle. \end{aligned} \tag{4.3.18}$$

By montonocity of $\mathcal{A}$, the second to the last term in the right of (4.3.18) drops. Also substituting (4.3.13) in (4.3.18), we have

$$\begin{aligned} 0 \;\le\; & 2\langle m_n - x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle - 2\langle k_n, x_{n+1} - x_{\sigma_n}\rangle \\ & + 2\delta_n\sigma_n\langle \mathcal{T}x_{\sigma_n} - \mathcal{T}x_n, x_{n+1} - x_{\sigma_n}\rangle \\ & + 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1}\rangle + 2\delta_n\frac{\rho_1\alpha}{\delta_{n+1}}\|x_n - x_{n+1}\|^2 \end{aligned} \tag{4.3.19}$$

We estimate the first term in the right of (4.3.19) as follows:

$$2\langle m_n - x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle = \|m_n - x_{\sigma_n}\|^2 - \|m_n - x_{n+1}\|^2 - \|x_{n+1} - x_{\sigma_n}\|^2. \tag{4.3.20}$$

Now, by applying (2.1.5) and (2.1.3)

$$
\begin{aligned}
\|m_n - x_{\sigma_n}\|^2 &= \|(x_n - x_{\sigma_n}) + \vartheta_n(x_n - x_{n-1})\|^2 \\
&= \|x_n - x_{\sigma_n}\|^2 + \vartheta^2\|x_n - x_{n-1}\|^2 + 2\vartheta\langle x_n - x_{\sigma_n}, x_n - x_{n-1}\rangle \\
&= \|x_n - x_{\sigma_n}\|^2 + \vartheta_n^2\|x_n - x_{n-1}\|^2 \\
&\quad -\vartheta_n(\|x_{\sigma_n} - x_{n-1}\|^2 - \|x_{\sigma_n} - x_n\|^2 - \|x_n - x_{n-1}\|^2) \\
&= \|x_n - x_{\sigma_n}\|^2 + \vartheta_n^2\|x_n - x_{n-1}\|^2 - \vartheta_n\|x_{\sigma_n} - x_{n-1}\|^2 \\
&\quad +\vartheta_n\|x_{\sigma_n} - x_n\|^2 + \vartheta_n\|x_n - x_{n-1}\|^2. \qquad (4.3.21)
\end{aligned}
$$

Replace x_{α_n} by x_{n+1} in (4.3.21), we obtain

$$
\begin{aligned}
\|m_n - x_{n+1}\|^2 &= \|x_n - x_{n+1}\|^2 + \vartheta_n^2\|x_n - x_{n-1}\|^2 - \vartheta_n\|x_{n+1} - x_{n-1}\|^2 \\
&\quad +\vartheta_n\|x_{n+1} - x_n\|^2 + \vartheta_n\|x_n - x_{n-1}\|^2. \qquad (4.3.22)
\end{aligned}
$$

Subtracting (4.3.22) from (4.3.21) gives

$$
\begin{aligned}
&\|m_n - x_{\sigma_n}\|^2 - \|m_n - x_{n+1}\|^2 \\
&= (1+\vartheta_n)\|x_n - x_{\sigma_n}\|^2 - (1+\vartheta_n)\|x_n - x_{n+1}\|^2 \\
&\quad -\vartheta_n\|x_{n-1} - x_{\sigma_n}\|^2 + \vartheta_n\|x_{n+1} - x_{n-1}\|^2. \qquad (4.3.23)
\end{aligned}
$$

By (2.1.3), we get

$$
\begin{aligned}
\vartheta_n\|x_{n+1} - x_{n-1}\|^2 &= \vartheta_n(\|x_{n+1} - x_n + x_n - x_{n-1}\|)^2 \\
&= \vartheta_n\|x_{n+1} - x_n\|^2 + \vartheta_n\|x_n - x_{n-1}\|^2 + 2\vartheta_n\langle x_{n+1} - x_n, x_n - x_{n-1}\rangle \\
&\leq 2\vartheta_n\|x_{n+1} - x_n\|^2 + 2\vartheta_n\|x_n - x_{n-1}\|^2. \qquad (4.3.24)
\end{aligned}
$$

Combining (4.3.20), (4.3.23) and (4.3.24), we obtain the following inequality:

$$
\begin{aligned}
2\langle m_n - x_{n+1}, x_{n+1} - x_{\sigma_n}\rangle &\leq (1+\vartheta_n)\|x_n - x_{\sigma_n}\|^2 - (1-\vartheta_n)\|x_n - x_{n+1}\|^2 \\
&\quad -\vartheta_n\|x_{n-1} - x_{\sigma_n}\|^2 + 2\vartheta_n\|x_n - x_{n-1}\|^2 \\
&\quad -\|x_{n+1} - x_{\sigma_n}\|^2. \qquad (4.3.25)
\end{aligned}
$$

Next, we estimate the second term in the right of (4.3.19)

$$
\begin{aligned}
-2\langle k_n, x_{n+1} - x_{\sigma_n}\rangle &= 2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_n - \mathcal{A}x_{n-1}, x_{\sigma_n} - x_{n+1}\rangle \\
&\leq 2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_n - \mathcal{A}x_{n-1}, x_{\sigma_n} - x_n\rangle \\
&\quad +2\delta_{n-1}(1-\gamma\delta_n\sigma_n)\langle \mathcal{A}x_n - \mathcal{A}x_{n-1}, x_n - x_{n+1}\rangle \\
&\leq \frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)(\|x_n - x_{n-1}\|^2 + \|x_{\sigma_n} - x_n\|^2) \\
&\quad +\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n) \\
&\quad \times(\|x_n - x_{n-1}\|^2 + \|x_n - x_{n+1}\|^2). \qquad (4.3.26)
\end{aligned}
$$

Also, using Lemma 4.2.3(i) and strong monotonicity of $\mathcal{T}$, we obtain

$$\begin{aligned}
2\delta_n\sigma_n\langle \mathcal{T}x_{\sigma_n} - \mathcal{T}x_n, x_{n+1} - x_{\sigma_n}\rangle &= 2\delta_n\sigma_n\langle \mathcal{T}x_{\sigma_n} - \mathcal{T}x_n, x_{n+1} - x_n\rangle \\
&\quad +2\delta_n\sigma_n\langle \mathcal{T}x_{\sigma_n} - \mathcal{T}x_n, x_n - x_{\sigma_n}\rangle \\
&\leq 2\delta_n\sigma_n L_{\mathcal{T}}\|x_{\sigma_n} - x_n\|\|x_{n+1} - x_{\sigma_n}\| \\
&\quad -2\delta_n\sigma_n\gamma\|x_n - x_{\sigma_n}\|^2 \\
&\leq 2\delta_n\sigma_n L_{\mathcal{T}}\left(\frac{a}{2}\|x_n - x_{\sigma_n}\|^2 + \frac{\|x_n - x_{n+1}\|^2}{2a}\right) \\
&\quad -2\delta_n\sigma_n\gamma\|x_n - x_{\sigma_n}\|^2 \\
&= -2\delta_n\sigma_n\left(\gamma - \frac{aL_{\mathcal{T}}}{2}\right)\|x_n - x_{\sigma_n}\|^2 \\
&\quad +\frac{\delta_n\sigma_n L_{\mathcal{T}}}{a}\|x_n - x_{n+1}\|^2. \qquad (4.3.27)
\end{aligned}$$

We estimate the fourth term in the right of (4.3.19) using the monotonicity of $\mathcal{A}$

$$\begin{aligned}
2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1}\rangle &= 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{\sigma_n} - x_{n+1}\rangle + 2\delta_n\langle \mathcal{A}x_{n+1} - \mathcal{A}x_{\sigma_n}, x_{\sigma_n} - x_{n+1}\rangle \\
&\leq 2\delta_n\langle \mathcal{A}x_n - \mathcal{A}x_{n+1}, x_{\sigma_n} - x_{n+1}\rangle \\
&\leq \frac{\delta_n}{\delta_{n+1}}\alpha(\|x_n - x_{n+1}\|^2 + \|x_{\sigma_n} - x_{n+1}\|^2) \\
&\leq \frac{\delta_n}{\delta_{n+1}}\alpha(\|x_n - x_{n+1}\|^2 + (2\|x_{\sigma_n} - x_n\|^2 + 2\|x_n - x_{n+1}\|^2)) \\
&\leq 3\alpha\frac{\delta_n}{\delta_{n+1}}\|x_n - x_{n+1}\|^2 + 2\alpha\frac{\delta_n}{\delta_{n+1}}\|x_{\sigma_n} - x_n\|^2 \\
&= 3\alpha\frac{\delta_n}{\delta_{n+1}}\|x_n - x_{n+1}\|^2 + 2\alpha(1 - \frac{a_n}{\delta_{n+1}})\|x_{\sigma_n} - x_n\|^2 \\
&\leq 3\alpha\frac{\delta_n}{\delta_{n+1}}\|x_n - x_{n+1}\|^2 + 2\alpha\|x_{\sigma_n} - x_n\|^2. \qquad (4.3.28)
\end{aligned}$$

Combining (4.3.19), (4.3.25), (4.3.26), (4.3.27) and (4.3.28) we have

$$\begin{aligned}
\|x_{n+1} - x_{\sigma_n}\|^2 &\leq \Bigg(1 + \vartheta_n - 2\delta_n\sigma_n\left(\gamma - \frac{aL_{\mathcal{T}}}{2}\right) \\
&\quad +2\alpha + \frac{\delta_{n-1}}{\delta_n}\alpha(1 - \gamma\delta_n\sigma_n)\Bigg)\|x_n - x_{\sigma_n}\|^2 \\
&\quad +\left(2\vartheta_n + 2\frac{\delta_{n-1}}{\delta_n}\alpha(1 - \gamma\delta_n\sigma_n)\right)\|x_n - x_{n-1}\|^2 \\
&\quad -\vartheta_n\|x_{\sigma_n} - x_{n-1}\|^2 \\
&\quad -\Bigg(1 - \vartheta_n - \frac{\delta_n\sigma_n L_{\mathcal{T}}}{a} - 3\alpha\frac{\delta_n}{\delta_{n+1}} \\
&\quad -\frac{\delta_{n-1}}{\delta_n}\alpha(1 - \gamma\delta_n\sigma_n) - 2\delta_n\rho_1\frac{\alpha}{\delta_{n+1}}\Bigg)\|x_{n+1} - x_n\|^2. \qquad (4.3.29)
\end{aligned}$$

Equivalently,

$$
\begin{aligned}
\|x_{n+1}-x_{\sigma_n}\|^2 \;\le\; & (1+\vartheta_n-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}}) \\
& +2\alpha+(1-\frac{a_n}{\delta_n})\alpha(1-\gamma\delta_n\sigma_n))\|x_n-x_{\sigma_n}\|^2 \\
& +\left(2\vartheta_n+2\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)\right)\|x_n-x_{n-1}\|^2 \\
& -\vartheta_n\|x_{\sigma_n}-x_{n-1}\|^2 \\
& -\left(1-\vartheta_n-\frac{\delta_n\sigma_nL_{\mathcal{T}}}{a}-3\alpha\frac{\delta_n}{\delta_{n+1}}\right. \\
& \left.-\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)-2\delta_n\rho_1\frac{\alpha}{\delta_{n+1}}\right)\|x_{n+1}-x_n\|^2. \qquad (4.3.30)
\end{aligned}
$$

So (4.3.30) becomes

$$
\begin{aligned}
\|x_{n+1}-x_{\sigma_n}\|^2 \;\le\; & (1+\vartheta_n-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})+3\alpha)\|x_n-x_{\sigma_n}\|^2 \\
& +\left(2\vartheta_n+2\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)\right)\|x_n-x_{n-1}\|^2 \\
& -\vartheta_n\|x_{\sigma_n}-x_{n-1}\|^2 \\
& -\left(1-\vartheta_n-\frac{\delta_n\sigma_nL_{\mathcal{T}}}{a}-3\alpha\frac{\delta_n}{\delta_{n+1}}-\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)-2\delta_n\frac{\rho_1\alpha}{\delta_{n+1}}\right) \\
& \times\|x_{n+1}-x_n\|^2.
\end{aligned}
$$

□

Theorem 4.3.7. *Let Assumption 4.3.1 holds and $\{x_n\}$ be the sequence iteratively generated by Algorithm 4.3.3, then $\{x_n\}$ converges strongly to the unique solution z^* of BVIP (1.1.1)-(1.1.2).*

Proof. Since $\mathcal{T}$ is $L_{\mathcal{T}}$-Lipschitz continuous and γ-strongly monotone, there exists two positive numbers a, b such that

$$2aL_{\mathcal{T}}+b<4\gamma, \qquad \text{(see [46])}. \tag{4.3.31}$$

We fix these numbers a, b in order to prove Theorem 4.3.7. So we have from (4.3.31) that $2\gamma-aL_{\mathcal{T}}>\frac{b}{2}>0$. We estimate $\|x_{n+1}-x_{\sigma_n}\|^2$ using (2.1.5) and Lemma 4.2.3(i) with $u=\|x_{\sigma_{n+1}}-x_{\sigma_n}\|$, $v=\|x_{n+1}-x_{\sigma_{n+1}}\|$ and $p_n=\frac{s_n}{2}$.

$$
\begin{aligned}
\|x_{n+1}-x_{\sigma_n}\|^2 &= \|x_{n+1}-x_{\sigma_{n+1}}\|^2+\|x_{\sigma_{n+1}}-x_{\sigma_n}\|^2 \\
&\quad +2\langle x_{n+1}-x_{\sigma_{n+1}},x_{\sigma_{n+1}}-x_{\sigma_n}\rangle \\
&\ge \|x_{n+1}-x_{\sigma_{n+1}}\|^2+\|x_{\sigma_{n+1}}-x_{\sigma_n}\|^2 \\
&\quad -2\|x_{n+1}-x_{\sigma_{n+1}}\|\|x_{\sigma_{n+1}}-x_{\sigma_n}\|. \qquad (4.3.32)
\end{aligned}
$$

$$2\|x_{n+1}-x_{\sigma_{n+1}}\|\|x_{\sigma_{n+1}}-x_{\sigma_n}\| \;\le\; \frac{s_n}{2}\|x_{n+1}-x_{\sigma_{n+1}}\|^2+\frac{2}{s_n}\|x_{\sigma_{n+1}}-x_{\sigma_n}\|^2. \tag{4.3.33}$$

Substitute (4.3.33) in (4.3.32)

$$\begin{aligned}\|x_{n+1}-x_{\sigma_n}\|^2 &\geq \frac{2-s_n}{2}\|x_{n+1}-x_{\sigma_{n+1}}\|^2 - \frac{s_n-2}{s_n}\|x_{\sigma_{n+1}}-x_{\sigma_n}\|^2\\ &= \frac{2-s_n}{2}\|x_{n+1}-x_{\sigma_{n+1}}\|^2 - \frac{2-s_n}{s_n}\|x_{\sigma_{n+1}}-x_{\sigma_n}\|^2. \qquad (4.3.34)\end{aligned}$$

Take $s_n := b\delta_n\sigma_n - \beta$, for some $b, \beta > 0$ and by applying Proposition 4.3.2(ii) to (4.3.34), we obtain

$$\begin{aligned}\|x_{n+1}-x_{\sigma_n}\|^2 \geq\ & \frac{2+\beta-b\delta_n\sigma_n}{2}\|x_{n+1}-x_{\sigma_{n+1}}\|^2\\ &-\frac{2+\beta-b\delta_n\sigma_n}{b\delta_n\sigma_n-\beta}\frac{(\sigma_{n+1}-\sigma_n)^2}{\sigma_n^2}M^2. \qquad (4.3.35)\end{aligned}$$

Combining (4.3.2) and (4.3.35), we obtain

$$\begin{aligned}&\frac{2+\beta-b\delta_n\sigma_n}{2}\|x_{n+1}-x_{\sigma_{n+1}}\|^2 \leq \left(1+\vartheta_n-\delta_n\sigma_n\left(2\gamma-aL_{\mathcal{T}}\right)+3\alpha\right)\|x_n-x_{\sigma_n}\|^2\\ &+\left(2\vartheta_n+2\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)\right)\|x_n-x_{n-1}\|^2-\vartheta_n\|x_{\sigma_n}-x_{n-1}\|^2\\ &+\frac{2+\beta-b\delta_n\sigma_n}{b\delta_n\sigma_n-\beta}\frac{(\sigma_{n+1}-\sigma_n)^2}{\sigma_n^2}M^2\\ &-\left(1-\vartheta_n-\frac{\delta_n\sigma_nL_{\mathcal{T}}}{a}-3\alpha\frac{\delta_n}{\delta_{n+1}}-\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)-2\delta_n\frac{\rho_1\alpha}{\delta_{n+1}}\right)\|x_{n+1}-x_n\|^2.\end{aligned}$$

which is equivalent to

$$\begin{aligned}\|x_{n+1}-x_{\sigma_{n+1}}\|^2 \leq\ & \frac{2+2\vartheta_n-2\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})+6\alpha}{2+\beta-b\delta_n\sigma_n}\|x_n-x_{\sigma_n}\|^2\\ &+\frac{2}{2+\beta-b\delta_n\sigma_n}\left(2\vartheta_n+2\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)\right)\|x_n-x_{n-1}\|^2\\ &-\frac{2\vartheta_n}{2+\beta-b\delta_n\sigma_n}\|x_{n-1}-x_{\sigma_n}\|^2+\frac{2\sigma_n}{b\delta_n\sigma_n-\beta}\frac{(\sigma_{n+1}-\sigma_n)^2}{\sigma_n^3}M^2\\ &-\frac{2}{2+\beta-b\delta_n\sigma_n}\left(1-\vartheta_n-\frac{\delta_n\sigma_nL_{\mathcal{T}}}{a}-3\alpha\frac{\delta_n}{\delta_{n+1}}\right.\\ &\left.-\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)-2\delta_n\frac{\rho_1\alpha}{\delta_{n+1}}\right)\|x_{n+1}-x_n\|^2. \qquad (4.3.36)\end{aligned}$$

Since $\vartheta_n \leq \vartheta_{n+1}$, $2+\beta-b\delta_n\sigma_n \geq 2+\beta-2b\delta_n\sigma_n > 0$ for some $n_1 \in \mathbb{N}$, $\forall n \geq n_1$. Take $b = 2\gamma - aL_{\mathcal{T}}$ and we have the resulting inequality

$$
\begin{aligned}
\|x_{n+1}-x_{\sigma_{n+1}}\|^2 \leq\ & \frac{2-2\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})+6\alpha}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\|x_n-x_{\sigma_n}\|^2 \\
& +\frac{2\vartheta_n}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\|x_n-x_{\sigma_n}\|^2 \\
& +\frac{2}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\left(2\vartheta_n+2\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)\right)\|x_n-x_{n-1}\|^2 \\
& -\frac{2\vartheta_n}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\|x_{n-1}-x_{\sigma_n}\|^2 \\
& +\frac{2\sigma_n}{\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})-\beta}\frac{(\sigma_{n+1}-\sigma_n)^2}{\sigma_n^3}M^2-\frac{2}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})} \\
& \times\left(1-\vartheta_n-\frac{\delta_n\sigma_nL_{\mathcal{T}}}{a}-3\alpha\frac{\delta_n}{\delta_{n+1}}-\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)-2\delta_n\frac{\rho_1\alpha}{\delta_{n+1}}\right)\|x_{n+1}-x_n\|^2 \\
\leq\ & \frac{2-2\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})+6\alpha}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\|x_n-x_{\sigma_n}\|^2 \\
& +\frac{2\vartheta_n}{2+\beta-2\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\|x_n-x_{\sigma_n}\|^2 \\
& +\frac{2}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\left(2\vartheta_n+2\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)\right)\|x_n-x_{n-1}\|^2 \\
& -\frac{2\vartheta_n}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\|x_{n-1}-x_{\sigma_n}\|^2 \\
& +\frac{2\sigma_n}{\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})-\beta}\frac{(\sigma_{n+1}-\sigma_n)^2}{\sigma_n^3}M^2 \\
& -\frac{2}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})} \\
& \times\left(1-\vartheta_n-\frac{\delta_n\sigma_nL_{\mathcal{T}}}{a}-3\alpha\frac{\delta_n}{\delta_{n+1}}-\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)-2\delta_n\frac{\rho_1\alpha}{\delta_{n+1}}\right)\|x_{n+1}-x_n\|^2.
\end{aligned}
$$

Equivalently,

$$
\begin{aligned}
& \|x_{n+1}-x_{\sigma_{n+1}}\|^2-\frac{\vartheta_{n+1}}{1+\frac{\beta}{2}-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\|x_n-x_{\sigma_n}\|^2 \\
& \leq\frac{2+6\alpha-2\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\|x_n-x_{\sigma_n}\|^2 \\
& \quad -\frac{2\vartheta_n}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\|x_{n-1}-x_{\sigma_n}\|^2 \\
& \quad +\frac{2}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\left(2\vartheta_n+2\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)\right)\|x_n-x_{n-1}\|^2 \\
& \quad +\frac{2\sigma_n}{\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})-\beta}\frac{(\sigma_{n+1}-\sigma_n)^2}{\sigma_n^3}M^2-\frac{2}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})} \\
& \quad \times\left(1-\vartheta_n-\frac{\delta_n\sigma_nL_{\mathcal{T}}}{a}-3\alpha\frac{\delta_n}{\delta_{n+1}}-\frac{\delta_{n-1}}{\delta_n}\alpha(1-\gamma\delta_n\sigma_n)-2\delta_n\frac{\rho_1\alpha}{\delta_{n+1}}\right) \\
& \quad \times\|x_{n+1}-x_n\|^2. \qquad (4.3.37)
\end{aligned}
$$

Adding $\frac{2\vartheta_{n+1}+2\frac{\delta_n}{\delta_{n+1}}\alpha(1-\gamma\delta_{n+1}\sigma_{n+1})}{1+2\alpha-\delta_{n+1}\sigma_{n+1}(2\gamma-aL_{\mathcal{T}})}\|x_{n+1}-x_n\|^2$ to both sides of (4.3.37) and taking $\beta = 6\alpha$ gives

$$
\begin{aligned}
&\|x_{n+1}-x_{\sigma_{n+1}}\|^2 - \frac{\vartheta_{n+1}}{1+\frac{\beta}{2}-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\|x_n-x_{\sigma_n}\|^2 \\
&+\frac{2\vartheta_{n+1}+\frac{2\delta_n}{\delta_{n+1}}\frac{\beta}{6}(1-\gamma\delta_{n+1}\sigma_{n+1})}{1+\frac{\beta}{2}-\delta_{n+1}\sigma_{n+1}(2\gamma-aL_{\mathcal{T}})}\|x_{n+1}-x_n\|^2 \\
&\leq \frac{2+\beta-2\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\left(\|x_n-x_{\sigma_n}\|^2 - \frac{\vartheta_n}{1+\frac{\beta}{2}-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\|x_{n-1}-x_{\sigma_n}\|^2\right. \\
&\left.+\frac{2\vartheta_n+2\frac{\delta_{n-1}}{\delta_n}\frac{\beta}{6}(1-\gamma\delta_n\sigma_n)}{1+\frac{\beta}{2}-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}\|x_n-x_{n-1}\|^2\right) + \frac{2\sigma_n}{\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})-\beta}\frac{(\sigma_{n+1}-\sigma_n)^2}{\sigma_n^3}M^2 \\
&-\left(\frac{2}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}(1-\vartheta_n-\frac{\delta_n\sigma_nL_{\mathcal{T}}}{a}-\frac{\beta}{2}\frac{\delta_n}{\delta_{n+1}}\right. \\
&\left.-\frac{\delta_{n-1}}{\delta_n}\frac{\beta}{6}(1-\gamma\delta_n\sigma_n)-2\delta_n\rho_1\frac{\beta}{6\delta_{n+1}}) - \frac{2\vartheta_{n+1}+2\frac{\delta_n}{\delta_{n+1}}\frac{\beta}{6}(1-\gamma\delta_{n+1}\sigma_{n+1})}{1+\frac{\beta}{2}-\delta_{n+1}\sigma_{n+1}(2\gamma-aL_{\mathcal{T}})}\right) \\
&\times\|x_{n+1}-x_n\|^2. \qquad (4.3.38)
\end{aligned}
$$

So,

$$
\begin{aligned}
&\lim_{n\to\infty}\left(\frac{2}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}(1-\vartheta_n-\frac{\delta_n\sigma_nL_{\mathcal{T}}}{a}-\frac{\beta}{2}\frac{\delta_n}{\delta_{n+1}}\right. \\
&\left.-\frac{\delta_{n-1}}{\delta_n}\frac{\beta}{6}(1-\gamma\delta_n\sigma_n)-2\delta_n\rho_1\frac{\beta}{6\delta_{n+1}}) - \frac{2\vartheta_{n+1}+2\frac{\delta_n}{\delta_{n+1}}\frac{\beta}{6}(1-\gamma\delta_{n+1}\sigma_{n+1})}{1+\frac{\beta}{2}-\delta_{n+1}\sigma_{n+1}(2\gamma-aL_{\mathcal{T}})}\right) \\
&\geq \lim_{n\to\infty}\left(\frac{2}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}(1-\vartheta-\frac{\delta_n\sigma_nL_{\mathcal{T}}}{a}-\frac{\beta}{2}\frac{\delta_n}{\delta_{n+1}}\right. \\
&\left.-\frac{\delta_{n-1}}{\delta_n}\frac{\beta}{6}(1-\gamma\delta_n\sigma_n)-2\delta_n\rho_1\frac{\beta}{6\delta_{n+1}}) + \frac{-2\vartheta-2\frac{\delta_n}{\delta_{n+1}}\frac{\beta}{6}(1-\gamma\delta_{n+1}\sigma_{n+1})}{1+\frac{\beta}{2}-\delta_{n+1}\sigma_{n+1}(2\gamma-aL_{\mathcal{T}})}\right) \\
&= \frac{2}{2+\beta}\left(1-\vartheta-\frac{\beta}{2}-\rho_1\frac{\beta}{3}+\frac{-2\vartheta-\frac{\beta}{3}}{1+\frac{\beta}{2}}\right) \\
&\geq \left(\frac{2}{2+\beta}(1-3\vartheta-\beta-\frac{1}{3}\rho_1\beta)\right) > 0.
\end{aligned}
$$

Thus, there exist $n_2 \geq n_1$ such that

$$
\begin{aligned}
&\left(\frac{2}{2+\beta-\delta_n\sigma_n(2\gamma-aL_{\mathcal{T}})}(1-\vartheta_n-\frac{\delta_n\sigma_nL_{\mathcal{T}}}{a}-\frac{\beta}{2}\frac{\delta_n}{\delta_{n+1}}\right. \\
&\left.-\frac{\delta_{n-1}}{\delta_n}\frac{\beta}{6}(1-\gamma\delta_n\sigma_n)-2\delta_n\rho_1\frac{\beta}{6\delta_{n+1}}) - \frac{2\vartheta_{n+1}+2\frac{\delta_n}{\delta_{n+1}}\frac{\beta}{6}(1-\gamma\delta_{n+1}\sigma_{n+1})}{1+\frac{\beta}{2}-\delta_{n+1}\sigma_{n+1}(2\gamma-aL_{\mathcal{T}})}\right) \\
&> 0, \quad \forall n \geq n_2.
\end{aligned}
$$

Equation (4.3.38) becomes

$$\begin{aligned}
&\|x_{n+1} - x_{\sigma_{n+1}}\|^2 - \frac{\vartheta_{n+1}}{1 + \frac{\beta}{2} - \delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}\|x_n - x_{\sigma_n}\|^2 \\
&+ \frac{2\vartheta_{n+1} + \frac{2\delta_n}{\delta_{n+1}}\frac{\beta}{6}(1 - \gamma\delta_{n+1}\sigma_{n+1})}{1 + \frac{\beta}{2} - \delta_{n+1}\sigma_{n+1}(2\gamma - aL_{\mathcal{T}})}\|x_{n+1} - x_n\|^2 \\
&\leq \frac{2 + \beta - 2\delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}{2 + \beta - \delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}\left(\|x_n - x_{\sigma_n}\|^2 - \frac{\vartheta_n}{1 + \frac{\beta}{2} - \delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}\|x_{n-1} - x_{\sigma_n}\|^2\right. \\
&\left.+ \frac{2\vartheta_n + 2\frac{\delta_{n-1}}{\delta_n}\frac{\beta}{6}(1 - \gamma\delta_n\sigma_n)}{1 + \frac{\beta}{2} - \delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}\|x_n - x_{n-1}\|^2\right) \\
&\frac{2\sigma_n}{\delta_n\sigma_n(2\gamma - aL_{\mathcal{T}}) - \beta}\frac{(\sigma_{n+1} - \sigma_n)^2}{\sigma_n^3}M^2
\end{aligned} \tag{4.3.39}$$

Also, there exist $n_3 \in \mathbb{N}$, $n_3 > n_2 > n_1 > n_0$, such that $\forall$ n $\geq n_3$, and $b = 2\gamma - aL_{\mathcal{T}}$.

$$2 + \beta - b\delta_n\sigma_n > 0,$$

$$0 \ < \ \frac{2 + \beta - 2\delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}{2 + \beta - b\delta_n\sigma_n} \ < \ 1,$$

$$0 < \frac{\delta_n\sigma_n(4\gamma - 2aL_{\mathcal{T}} - b)}{2 + \beta - b\delta_n\sigma_n} < 1, \tag{4.3.40}$$

$$1 - v_n = \frac{2 + \beta - 2\delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}{2 + \beta - b\delta_n\sigma_n}, \qquad v_n = \frac{\delta_n\sigma_n(4\gamma - 2aL_{\mathcal{T}} - b)}{2 + \beta - b\delta_n\sigma_n},$$

take

$$\begin{aligned}
e_n \ = \ & \|x_n - x_{\sigma_n}\|^2 - \frac{\vartheta_n}{1 + \frac{\beta}{2} - \delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}\|x_{n-1} - x_{\sigma_n}\|^2 \\
& + \frac{2\vartheta_n + 2\frac{\delta_{n-1}}{\delta_n}\frac{\beta}{6}(1 - \gamma\delta_n\sigma_n)}{1 + \frac{\beta}{2} - \delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}\|x_n - x_{n-1}\|^2.
\end{aligned}$$

There exist $n_4 \in \mathbb{N}$, $n_4 > n_3$ such that $\forall$ n $\geq n_4$,

$$\begin{aligned}
e_n &\geq \|x_n - x_{\sigma_n}\|^2 - \frac{\vartheta_n}{1+\frac{\beta}{2}-\delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}2(\|x_{n-1}-x_n\|^2 + \|x_n - x_{\sigma_n}\|^2) \\
&\quad + \frac{2\vartheta_n + 2\frac{\delta_{n-1}}{\delta_n}\frac{\beta}{6}(1-\gamma\delta_n\sigma_n)}{1+\frac{\beta}{2}-\delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}\|x_n - x_{n-1}\|^2 \\
&\geq \|x_n - x_{\sigma_n}\|^2 - \frac{2\vartheta_n}{1+\frac{\beta}{2}-\delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}\|x_n - x_{\sigma_n}\|^2 \\
&\quad + \frac{2\frac{\delta_{n-1}}{\delta_n}\frac{\beta}{6}(1-\gamma\delta_n\sigma_n)}{1+\frac{\beta}{2}-\delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}\|x_n - x_{n-1}\|^2 \\
&\geq \|x_n - x_{\sigma_n}\|^2 - 2\vartheta_n\|x_n - x_{\sigma_n}\|^2 + \frac{2\frac{\delta_{n-1}}{\delta_n}\frac{\beta}{6}(1-\gamma\delta_n\sigma_n)}{1+\frac{\beta}{2}-\delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}\|x_n - x_{n-1}\|^2 \\
&\geq (\frac{1}{2}+\frac{\beta}{6})\|x_n - x_{\sigma_n}\|^2 + \frac{2\frac{\delta_{n-1}}{\delta_n}\frac{\beta}{6}(1-\gamma\delta_n\sigma_n)}{1+\frac{\beta}{2}-\delta_n\sigma_n(2\gamma - aL_{\mathcal{T}})}\|x_n - x_{n-1}\|^2 > 0. \qquad (4.3.41)
\end{aligned}$$

And,

$$w_n = \frac{2\sigma_n}{b\delta_n\sigma_n - \beta}\frac{(\sigma_{n+1}-\sigma_n)^2}{\sigma_n^3}M^2,$$

So (4.3.39) reduces to

$$e_{n+1} \leq (1-v_n)e_n + w_n \quad \forall\ n \geq n_4,$$

also

$$\begin{aligned}
\frac{w_n}{v_n} &= \frac{2\sigma_n}{b\delta_n\sigma_n - \beta}\frac{(\sigma_{n+1}-\sigma_n)^2}{\sigma_n^3}M^2 \times \frac{2+\beta - b\delta_n\sigma_n}{\delta_n\sigma_n(4\gamma - 2aL_{\mathcal{T}} - b)} \\
&= \frac{2\sigma_n(2+\beta-b\delta_n\sigma_n)}{(b\delta_n\sigma_n - \beta)\delta_n(4\gamma - 2aL_{\mathcal{T}} - b)}\frac{(\sigma_{n+1}-\sigma_n)^2}{\sigma_n^4}M^2
\end{aligned}$$

By Assumption 4.3.1 (b) it is very easy to see that $\lim\limits_{n\to\infty}\frac{w_n}{v_n} = 0$. From (4.3.40) and the definition of v_n, it is obvious that $v_n \in (0,1)$ for all $n \geq n_3$. Also, since $\sigma_n \to 0$, and $\delta_n \to \bar{\delta} > 0$, we observe that $v_n \to 0$ as $n \to \infty$. Furthermore, from the definition of v_n and Assumption 4.3.1 (a), we obtain that $\sum_{n=n_3}^{\infty} v_n = +\infty$.

Therefore by applying Lemma 4.2.4 to (4.3.39), we have that $e_n \to 0$ as $n \to \infty$. It follows from (4.3.41) that

$$\lim_{n\to\infty}\|x_n - x_{\sigma_n}\|^2 = \lim_{n\to\infty}\|x_n - x_{n-1}\|^2 = 0.$$

Using Proposition 4.3.2 and Assumption 4.3.1, we derive that $\|x_{\sigma_n} - z^*\| \to 0$ as $n \to \infty$, where z^* solves BVIP (1.1.1)-(1.1.2). Therefore,

$$\|x_n - z^*\| \leq \|x_n - x_{\sigma_n}\| + \|x_{\sigma_n} - z^*\| \to 0, \quad n \to \infty.$$

We have that $\{x_n\}$ converges strongly to z^*. □

4.4 Numerical Examples

In this section, we discuss the numerical behavior of Algorithm 4.3.3 in both finite and infinite dimensional Hilbert spaces. We compare our algorithm with Algorithms 2.3.7, 2.3.8 and 2.3.9. All codes are written in Matlab 2023 (b) and performed on a personal computer with an Intel(R) Core(TM) i5-10210U CPU at 1.60GHz and 8.00 Gb-RAM. In Tables 1-2, Iter. means the number of iterations while CPU means the CPU time in seconds. In our computations, we choose $\delta_0 = 1$, $\delta_1 = 0.5$, $\epsilon = 0.1$ *and* $\alpha = 0.9(\frac{1-2\epsilon}{2})$*. Also, we take* $\sigma_n = \frac{1}{\sqrt{n+1}}$, $a_n = \frac{1}{(n+100)^2}$ *and* $\vartheta_n = \vartheta = 0.9 \min\left\{\frac{\frac{1}{2}-\alpha}{2}, \frac{1-\beta-\frac{\rho_1\beta}{3}}{3}\right\}$ *with* $\beta = 6\alpha$.

Example 4.4.1. *Let* $H = \mathbb{R}^N$*. Define* $\mathcal{A} : \mathbb{R}^N \to \mathbb{R}^N$ *by* $\mathcal{A}(x) = Mx + q$*, where the matrix* M *is formed as:* $M = V \sum V'$*, where* $V = I - \frac{2vv'}{||v||^2}$ *and* $\sum = diag(\sigma_{11}, \sigma_{12}, \cdots, \sigma_{1N})$ *are the Householder and the diagonal matrix, and*

$$\sigma_{1j} = \cos\frac{j\pi}{N+1} + 1 + \frac{\cos\frac{\pi}{N+1} + 1 - \widehat{C}(\cos\frac{N\pi}{N+1} + 1)}{\widehat{C} - 1}, \qquad j = 1, 2, \cdots, N,$$

with $\widehat{C}$ *been the present condition number of* M *([43, Example 5.2]). In the numerical computation, we choose* $\widehat{C} = 10^4$, $q = 0$ *and uniformly take the vector* $v \in \mathbb{R}^N$ *in* $(-1, 1)$*. Thus,* $\mathcal{A}$ *is monotone and Lipschitz continuous with* $L_{\mathcal{A}} = ||M||$ *(see [43]). We define* $\mathcal{T} : \mathbb{R}^N \to \mathbb{R}^N$ *by* $\mathcal{T}x = Dx + p$*, where* D *is a symmetric and positive-definite matrix of size* $N \times N$ *and* p *is a vector in* $\mathbb{R}^N$*. Clearly,* $\mathcal{T}$ *is* $||D||$*-Lipschitz continuous. Moreover, since* D *is symmetric and positive-definite, we have that* $\langle Dx - Dy\rangle = \langle D(x-y)\rangle \geq ||D||^{-1}||x-y||^2$*. Hence,* $\mathcal{T}$ *is strongly monotone.*

Now, consider an ellipsoid in $\mathbb{R}^N$ *as the feasible set* C*. That is,* $C = \{x \in \mathbb{R}^N : (x - d)'P(x-d) \leq r^2\}$*, where* P *is a positive definite matrix,* $d \in \mathbb{R}^N$ *and* $r > 0$*. Then, the projection onto* C *is difficult to compute (see, for example [41]). Consider the convex function* $\mathcal{Q} : \mathbb{R}^N \to \mathbb{R}$ *defined by* $\mathcal{Q}(x) = \frac{1}{2}\left[(x-d)'P(x-d) - r^2\right]$, $\forall x \in \mathbb{R}^N$*. Then,* C *is a level set of* q*. That is,* $C := \{x \in \mathbb{R}^N : \mathcal{Q}(x) \leq 0\}$*. Also, it is easy to see that* $\mathcal{Q}$ *is differentiable on* $\mathbb{R}^N$*. In fact,* $\mathcal{Q}'(x) = P(x-d)$, $\forall x \in \mathbb{R}^N$*. Hence,* Q' *is* K*-Lipschitz continuous with* $K = ||P||$*.*

Now, let $\lambda_{\max}$ *and* $\lambda_{\min}$ *be maximum and minimum eigenvalues of* P*, respectively. Suppose* $d = 0$*, since* $\frac{||\mathcal{A}(x)||}{||\mathcal{Q}'(x)||} = \frac{||Mx||}{||Px||} \leq \frac{||M||||x||}{\lambda_{\min}||x||} = \frac{||M||}{\lambda_{\min}}$ *holds for arbitrary* $x \in \mathbb{R}^N$*,* $x \neq 0$*, we have that* $\rho_1 = \frac{||M||}{\lambda_{\min}}$ *(see [41]). On the other hand, suppose* $d \neq 0$*, since* $\frac{||\mathcal{A}(x)||}{||\mathcal{Q}'(x)||} = \frac{||Mx||}{||P(x-d)||} \leq \frac{||M||(||x-d||+||d||)}{\lambda_{\min}||x-d||} \leq \frac{||M||\sqrt{\lambda_{\max}}\left(r+||d||\sqrt{\lambda_{\min}}\right)}{\lambda_{\min}^{\frac{3}{2}} r}$ *holds for all* $x \in \partial C$*, we have that* $\rho_1 = \frac{||M||\sqrt{\lambda_{\max}}\left(r+||d||\sqrt{\lambda_{\min}}\right)}{\lambda_{\min}^{\frac{3}{2}} r}$*.*

Since the solution of the problem is unknown, we define the sequence $TOL_n := ||x_{n+1} - x_n||^2$*, and use the stopping criterion* $TOL_n < \varepsilon$ *for the iterative processes, where* ε *is the predetermined error. Moreover, we consider different scenarios of the problem's dimensions. That is,* $N = 100, 500, 1000$*, with starting points* $x_1 = (1, 1, \ldots, 1)'$ *and* $x_0 = (0, 0, \ldots, 0)'$*. For this example, we take* $\varepsilon = 10^{-9}$ *as the stopping criterion and obtain the numerical results reported in Table 1 and Figure 4.1.*

Table 1. Numerical results for Example 4.4.1 with $\epsilon = 10^{-9}$.

	N=100		*N=500*		*N=1000*	
Algorithms	*CPU Time*	*Iter.*	*CPU Time*	*Iter.*	*CPU Time*	*Iter.*
Algorithm 4.3.3	*2.1099*	*41*	*2.0390*	*45*	*2.0053*	*48*
Algorithm 2.3.7	*4.0274*	*54*	*4.0260*	*122*	*4.0033*	*65*
Algorithm 2.3.8	*8.0563*	*170*	*8.0585*	*306*	*8.0170*	*195*
Algorithm 2.3.9	*7.0350*	*122*	*7.0171*	*220*	*7.0021*	*122*

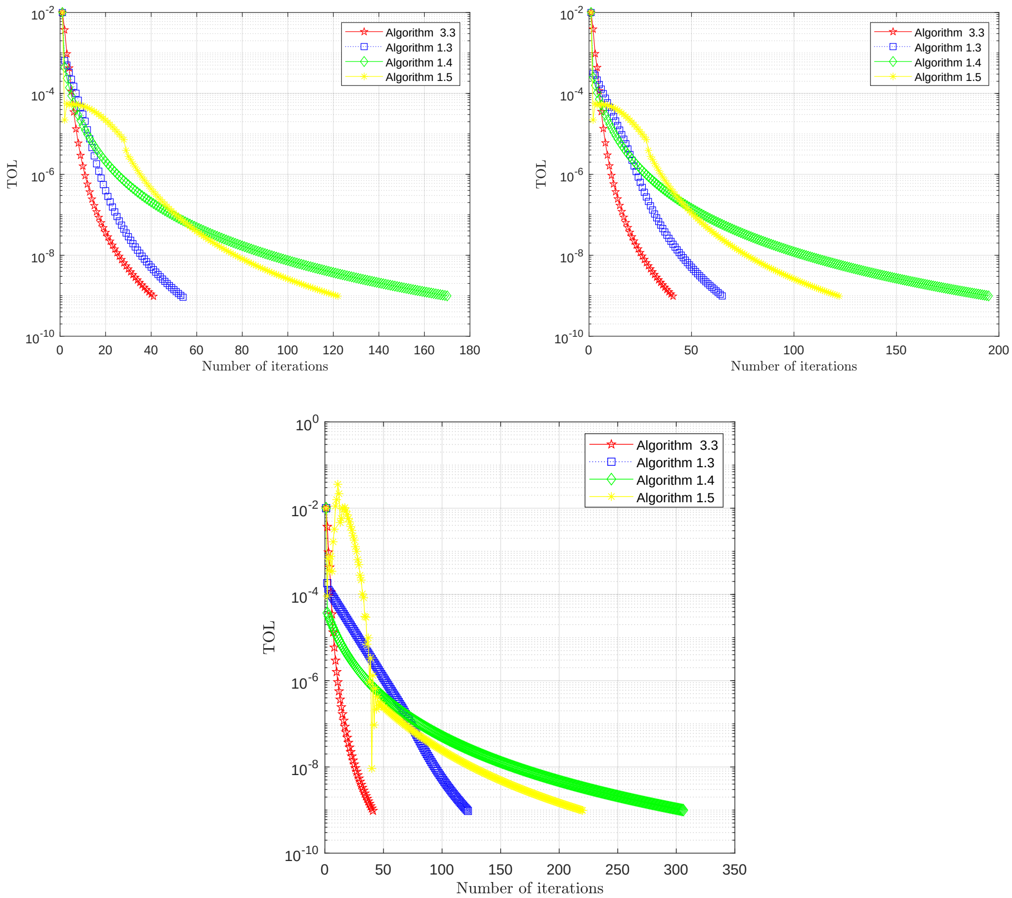

Figure 4.1: The behavior of TOL_n with $\epsilon = 10^{-9}$: Top Left: $N = 100$; Top Right: $N = 500$; Bottom: $N = 1000$.

Example 4.4.2. *Let* $C = \{x \in \mathbb{R}^2 : \mathcal{Q}(x) := x_1^2 - x_2 \leq 0\}$, *where* $x = (x_1, x_2)$. *Define* $\mathcal{A} : \mathbb{R}^2 \to \mathbb{R}^2$ *by*

$$\mathcal{A}(x) = (3\Psi(x_1), 2x_1 + x_2),$$

where

$$\Psi(u) = \begin{cases} y(u-1) + y, & \text{if } u > 1, \\ y^u, & \text{if } -1 \leq u \leq 1, \\ y^{-1}(u+1) + y^{-1}, & \text{if } u < -1. \end{cases}$$

Then, one can easily verify that $L_{\mathcal{A}} = \sqrt{9y^2 + 5}$. *Also, it is easy to see that* $\mathcal{Q}$ *is a continuously differentiable convex function with Lipschitz constant* $K = 2$. *Moreover, we have that* $\rho_1 = 3\sqrt{y^2 + 1}$ *(see for example, [41]). Furthermore, we choose the same parameters as in Example 4.4.1 and consider the following cases for the numerical experiments.*

Case 1: *Take* $x_1 = (0.5,\ 1)'$ *and* $x_0 = (1,\ 2)'$.

Case 2: *Take* $x_1 = (2,\ 1)'$ *and* $x_0 = (0.5,\ -1)'$.

Case 3: *Take* $x_1 = (1,\ -1)'$ *and* $x_0 = (2,\ -2)'$.

The numerical results are reported in Table 2 and Figure 4.2.

Table 2. Numerical results for Example 4.4.2 with $\epsilon = 10^{-6}$**.**

	Case 1		*Case 2*		*Case 3*	
Algorithms	*CPU Time*	*Iter.*	*CPU Time*	*Iter.*	*CPU Time*	*Iter.*
Algorithm 4.3.3	*1.0064*	*8*	*1.0063*	*4*	*1.0373*	*6*
Algorithm 2.3.7	*3.0050*	*14*	*2.0045*	*11*	*3.0309*	*15*
Algorithm 2.3.8	*3.0041*	*14*	*4.0027*	*19*	*4.0379*	*37*
Algorithm 2.3.9	*4.0047*	*29*	*4.0069*	*14*	*4.0254*	*12*

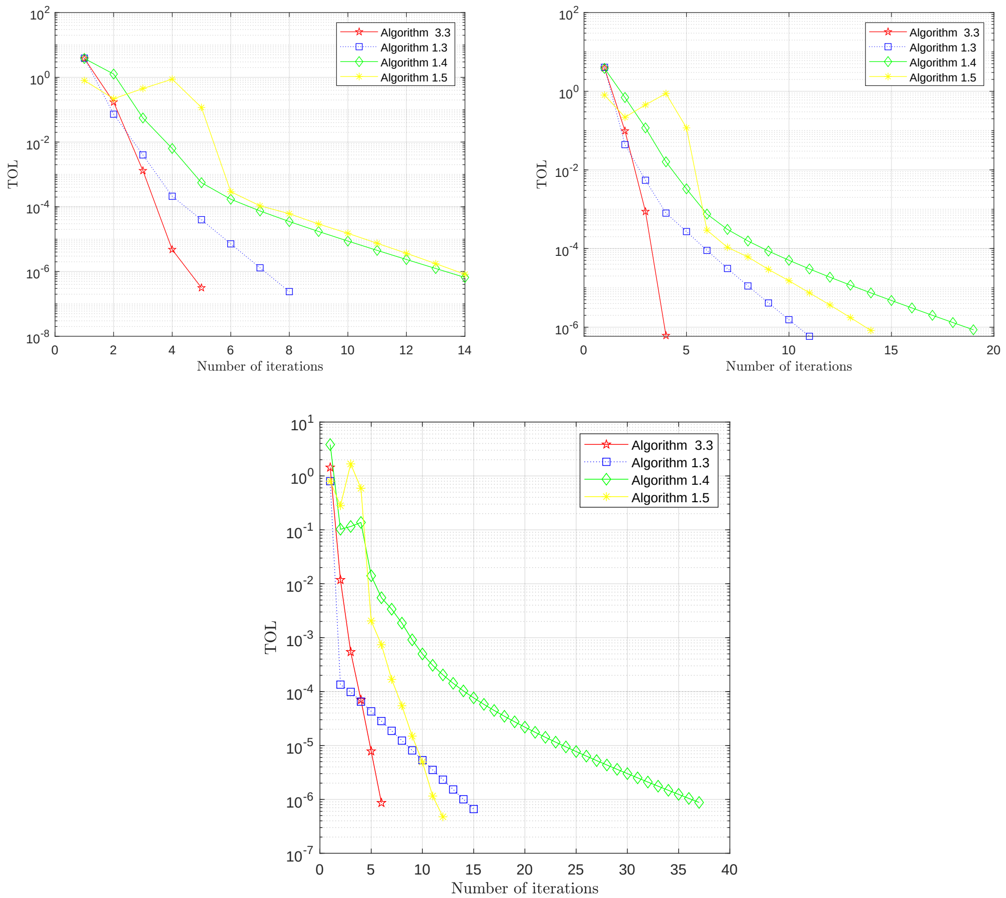

Figure 4.2: The behavior of TOL_n with $\epsilon = 10^{-6}$: Top Left: **Case 1**; Top Right: **Case 2**; Bottom: **Case 3**.

Example 4.4.3. *Let $H = \mathcal{L}_2([0,1])$ be endowed with inner product $\langle x, y\rangle = \int_0^1 x(t)y(t)dt, \ \forall\, x, y \in \mathcal{L}_2([0,1])$ and norm $||x|| := \Big(\int_0^1 |x(t)|^2 dt\Big)^{\frac{1}{2}}, \ \ \forall\, x, y \in \mathcal{L}_2([0,1])$. Let φ be an arbitrary fixed element in $C([0,1])$ (the space of all real-valued continuous functions on $[0,1]$) and $\mathcal{C} = \{x \in \mathcal{L}_2([0,1]) : \|\varphi x\| \le 1\}$, where $\varphi(t) = e^{-t}$ for all $t \in [0,1]$. Then, $\mathcal{C}$ is a nonempty closed and convex subset of $\mathcal{L}_2([0,1])$. However, we know from [41] that the explicit formula for projection onto $\mathcal{C}$ may be difficult to find.*
Now, let $\mathcal{Q} : \mathcal{L}_2([0,1]) \to \mathbb{R}$ be defined by $\mathcal{Q}(x) = \frac{1}{2}\Big(\|\varphi x\|^2 - 1\Big), \ \forall x \in \mathcal{L}_2([0,1])$. Then, $\mathcal{Q}$ is a convex function with $\mathcal{C}$ a level set of $\mathcal{Q}$. Also, $\mathcal{Q}$ is differentiable on $\mathcal{L}_2([0,1])$. In fact, $\mathcal{Q}'(x) = \varphi^2 x, \ \forall x \in \mathcal{L}_2([0,1])$. Furthermore, $\mathcal{Q}'$ is Lipschitz continuous.
Now, define $\mathcal{A} : \mathcal{L}_2([0,1]) \to \mathcal{L}_2([0,1])$ by $\mathcal{A}x(t) = \int_0^t x(s)ds, \ \forall x \in \mathcal{L}_2([0,1]), \ t \in [0,1]$. Then, $\mathcal{A}$ is monotone and Lipschitz continuous. Let $m\varphi = min_{t\in[0,1]}|\varphi(t)| = e^{-1}$. Then,

$$\frac{\|\mathcal{A}x\|}{\|\mathcal{Q}'(x)\|} = \frac{\|x_+\|}{\|\varphi^2 x\|} \le \frac{\|x\|}{m^2\varphi\|x\|} = e^2$$

holds for all $x \in \mathcal{L}_2([0,1]), \ x \ne 0$ (see [41]). Thus, we have that $\rho_1 = e^2$. Furthermore, define the mapping $\mathcal{T} : \mathcal{L}_2([0,1]) \to \mathcal{L}_2([0,1])$ by $\mathcal{T}x = x - x_0$. Then, $\mathcal{T}$ is strongly monotone and Lipschitz continuous.

We consider the following cases for the numerical experiments of this example.
Case 1: *Take $x_1(t) = 1 + t^2$ and $x_0(t) = t + 5$.*

Case 2: *Take $x_1(t) = \sin(t)$ and $x_0(t) = t + 1$.*

Case 3: *Take $x_1(t) = t + 1$ and $x_0(t) = t + t^3$.*

The numerical results are reported in Table 3 and Figure 4.3.

Table 3. Numerical results for Example 4.4.3 with $\epsilon = 10^{-12}$.

	Case 1		*Case 2*		*Case 3*	
Algorithms	*CPU Time*	*Iter.*	*CPU Time*	*Iter.*	*CPU Time*	*Iter.*
Algorithm 4.3.3	*4.0341*	*7*	*4.0313*	*22*	*4.0304*	*15*
Algorithm 2.3.7	*8.0164*	*17*	*5.0261*	*28*	*7.0174*	*51*
Algorithm 2.3.8	*8.0292*	*30*	*8.0316*	*57*	*8.0393*	*102*
Algorithm 2.3.9	*9.0162*	*32*	*8.0159*	*33*	*8.0249*	*90*

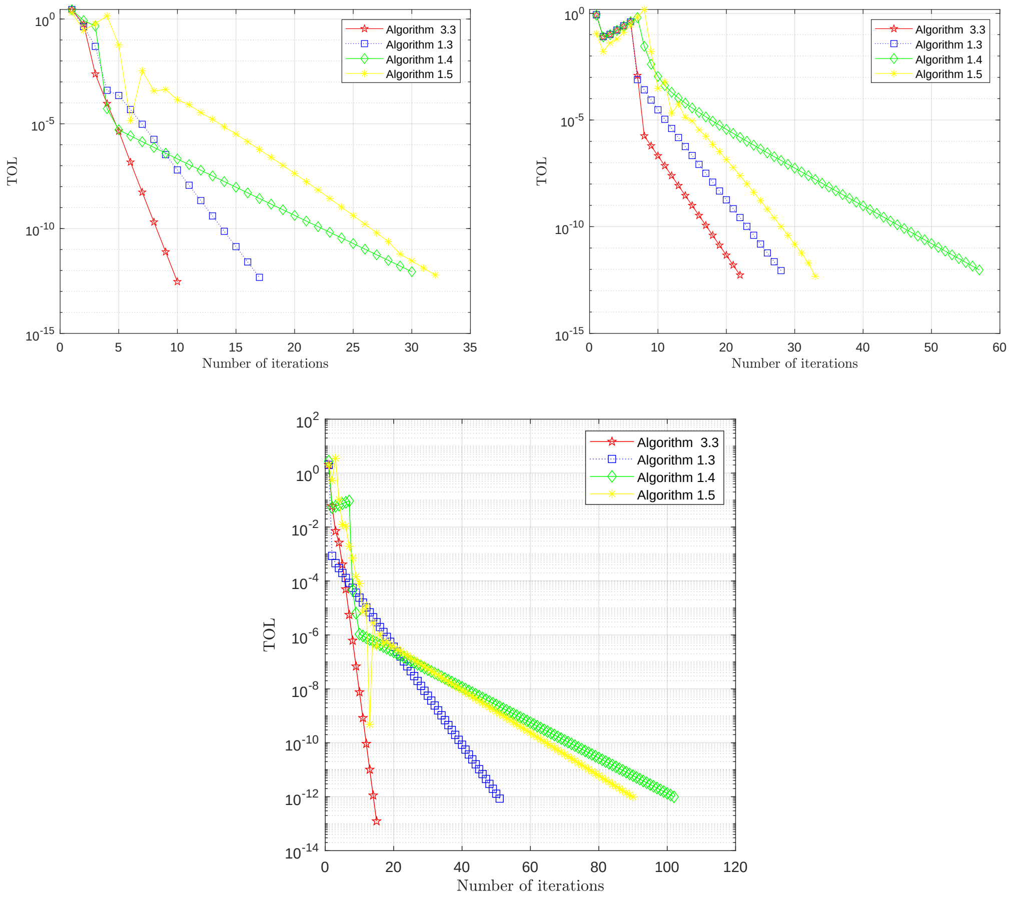

Figure 4.3: The behavior of TOL_n with $\epsilon = 10^{-12}$: Top Left: **Case 1**; Top Right: **Case 2**; Bottom: **Case 3**.

4.5 Conclusion

We studied an inertial forward-reflected-backward method for solving BVIP (1.1.1) - (1.1.2)*. The proposed method uses a variable inertial parameter and we proved that it converges strongly to a solution of the BVIP* (1.1.1) - (1.1.2)*. The method is a modification of the forward-reflected-backward splitting method in [67]. As in [67], it requires only a single projection onto a half-space* C_n *and one functional evaluation of* $\mathcal{A}$*. Our numerical experiments show that our method performs better than related methods in the literature.*

Chapter 5

Conclusion, Contribution to Knowledge and Future Research

5.1 Conclusion

In this dissertation, we gave a brief background of our proposed study, our motivation towards this study and a review of outstanding contributions in the area of our study. Also, we presented fundamental results in the form of lemmas, propositions and definitions that were instrumental in achieving the results in Chapter Three and Chapter Four. In Chapter Three, we introduced and studied the iterative regularization-projection method for approximating a solution of BVIP (1.1.1)-(1.1.2)*. Our method is motivated by the forward-reflected-backward method, the Tikhonov regularization method and the inertial method. The introduced method involves a single projection onto feasible set C and one evaluation of $\mathcal{A}$, which is desirable in large-scale optimization problems such as optimal control. Also, the inertial extrapolation term in our method improves its speed of convergence. Finally, we established the strong convergence result of the proposed method; moreover, we gave numerical experiments to illustrate the performance of the proposed method in comparison with other existing methods in the literature. In Chapter Four, we modified the method presented in Chapter Three by replacing the stepsizes and single projection onto feasible set C by a non-monotonic stepsizes and a projection onto a constructive half space, respectively. The proposed method performs without prior knowledge of the Lipschitz constant since it uses the non-monotonic stepsizes. This simple adjustment ensures that the obtained algorithm is cheaper and easy to implement. We also prove the strong convergence result of this algorithm. Lastly, we gave the numerical experiments illustrating the performance of this method.*

5.2 Contribution to knowledge

1. *The method presented in Chapter Three involves one projection onto a feasible set unlike methods in [5, 46, 87, 90, 91, 92] and one evaluation of $\mathcal{A}$ unlike algorithms in [5, 73, 87, 88, 90, 91, 92, 93] in each iteration. Therefore, our method is desirable*

when the feasible set and the operator has a complex structure, thus our method performs better than these algorithms.

2. *The method introduced in Chapter Four involves a single projection onto a half-space only, without any other projection onto feasible sets. This makes our method entirely different from all existing methods for solving BVIP* (1.1.1)-(1.1.2)*. Moreover, our method ensures a cheap and easy computation and, thus, is more implementable and efficient than existing ones in the literature.*

3. *Method studied in Chapter Four incorporate non-monotonic stepsizes which are better than methods in [46, 73, 87, 92, 93] that adopt non-increasing stepsizes, methods in [5, 90] which employ Armijo-type stepsizes, and method in [91] that uses the knowledge of Lipschitz constant. Moreover, our methods in Chapters Three and Four achieve the same strong convergence results as in [6] but when operator $\mathcal{A}$ is monotone instead of inverse strongly monotone assumption on $\mathcal{A}$ used in [6] which is rigid.*

4. *Observe that if $\vartheta_n = 0$ and the backward operator, say S, is neither a zero operator nor normal cone to C, our algorithms coincide with the regularized forward-reflected-backward splitting method [86], and if $\sigma_n = 0$, $\delta_n = \delta$, our algorithms become the forward-reflected-backward method [67] for finding the zero of monotone inclusion problem; furthermore, if $\mathcal{A} = 0$ our algorithms further reduce to the proximal point algorithm [78]. Directly from our algorithms, if $\mathcal{A}$ is an affine operator, $\sigma_n = 0$, $C_n = C$, $\delta_n = \delta$, then our algorithms reduce to the projected reflected gradient method [65] for VIs. Hence, our methods are generalizations and extensions of these methods to BVIP* (1.1.1)*, and our method of analysis in Chapter Four is different from the methods of proof in [65, 67, 78, 86].*

5. *The methods presented in Chapters Three and Four have a more simple and elegant structure as compared to algorithms in [5, 73, 46, 87, 88, 90, 91, 92, 93]. Hence, it is relatively easy to implement and modifiable for solving problems in the future.*

5.3 Future research

1. *VIP* (1.1.2) *can alternatively be called a Deterministic Variational Inequality Problem (DVIP); important extensions of DVIP are the stochastic VIPs and the BVIPs studied in this dissertation. As mentioned earlier, VIPs and BVIPs are applicable in modelling practical problems emerging from optimization, finance, economics and engineering. However, practical problems arising from transportation problems and supply chain problems include large of uncertain data, which VIPs and BVIPs cannot handle. For solving the uncertain problems, the stochastic VIP (SVIP) was proposed by interested researchers (see [51, 62, 77]). Let ω be a random vector defined for some probabilistic space and $\Gamma(\cdot,\, \omega)$: $\mathbb{R}^n \to \mathbb{R}^n$ be a nonlinear mapping with the expectation $\mathcal{A}x = \mathbb{E}[\Gamma(x, \omega)]$. The SVIP is defined as follows:*

$$\textit{Find } x^* \in X \textit{ such that } \langle \mathcal{A}x^*, x - x^* \rangle \ \forall x \in X. \tag{5.3.1}$$

Let the solution set of (5.3.1) *be represented as SVIP($\mathcal{A}$, X).*
In practice, $\mathcal{A}x$ is not easy to calculate because of the unknown distribution of ω; as a result, we cannot directly apply the available methods for solving VIPs to SVIPs. However, there are two methods in the literature for solving SVIPs – the stochastic average approximation (SAA) method (see [45, 98, 101]) and the stochastic approximation (SA) method (see [53, 104, 105]).

Recently, Jiang and Xu [52] introduced the SA-based single-projection algorithm for solving (5.3.1)*. Their algorithm is as follows: $x_0 \in X$, and*

$$x_{k+1} = \Pi_X(x_k - \alpha_k \Gamma(x_k, \omega_k)), \; \forall k \geq 1 \tag{5.3.2}$$

where Π_X is the Euclidean projection from $\mathbb{R}^n$ onto X, ω_k is an independent identity distributed sample of ω, and $\{\alpha_k\}$ is a positive sequence. The authors proved when $\mathcal{A}$ is Lipschitz continuous and strongly monotone and $\{\alpha_k\}$ satisfies the conditions that $\sum_{k=0}^{\infty} \alpha_k = \infty$ and $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$ the sequence $\{x_k\}$ almost surely (a.s.) converges to the unique solution of the SVIP (5.3.1)*. The method* (5.3.2) *only needs a sample at each iteration and, hence, is performed relatively easily compared to the SAA method. This method has attracted the attention of many researchers who have modified it in various ways. For example, Kannan and Shanbhag [53] proposed an SA-based extragradient method for solving the pseudomonotone SVIP. Also, Yang et al. [104] considered an SA-based backward-forward method for the SVIP where they used a weaker condition on $\mathcal{A}$ than pseudomonotonicity. Very recently, Wang et al. [99] also presented an SA-based subgradient extragradient method for solving a pseudomonotone SVIP. In our future research, we shall construct and study an SA-based regularization projection algorithm, an extension of our methods, for approximating the solution of stochastic BVIPs. These stochastic BVIPs are natural extensions of the BVIP model studied in this dissertation.*

2. *In this dissertation, we used the monotonicity assumption for the cost operator $\mathcal{A}$; however, the monotonicity assumption seems too rigid and is not always the case in practical application. There are many practical problems whose cost operators do not meet the monotonicity assumption. On the other hand, the Tikhonov regularization method works by adding to the cost operator an operator that is strongly monotone and depends on a regularization parameter to obtain a parameterized strongly monotone VIP, which is uniquely solved. However, when the cost operator $\mathcal{A}$ is quasi-monotone rather than monotone, the monotonicity of the regularized problem may fail to hold; this is because the sum of a strongly monotone operator and a quasimonotone operator, in general, is not quasimonotone. Therefore, the iterative projection-regularized method that requires some monotonicity assumption cannot be applied to solve the regularized subvariational inequalities. Hence, we plan to study the method in Chapter Three under the quasimonotone assumption of the cost operator $\mathcal{A}$ with a method of proof different from the one we used here. We note that the algorithm in Chapter Four may not work under the quasi-monotone assumption of the cost operator $\mathcal{A}$.*

Bibliography

[1] J. A. Abuchu, A. E. Ofem, G. C. Ugwunnadi, O. K. Narain, A. Hussain, *Hybrid alternated inertial projection and contraction algorithm for solving bilevel variational inequality problems, J. Math., https://doi.org/10.1155/2023/318574.*

[2] Y. I. Alber, I. Ryazantseva, *Nonlinear Ill-Posed Problems of Monotone Type, Springer, Dordrecht, (2006).*

[3] F. Alvarez, H. Attouch, *An Inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, Set-Valued Anal.,* **9,** *(2001) 3-11.*

[4] T. V. Anh, *Linesearch methods for bilevel split pseudomonotone variational inequality problems, Numer. algorithms.,* **81,** *(2019) 1067–1087, https://doi.org/10.1007/s11075-018-0583-2.*

[5] P. N. Anh, J. K. Kim, L. D. Muu, *An extragradient algorithm for solving bilevel pseudomonotone variational inequalities, J. Glob. Optim.,* **52,** *(2012) 627–639.*

[6] T.T. Anh, L. B. Long, T. V. Anh, *A projection method for bilevel variational inequalities, J. Inequal. Appl.,* **1,** *(2014) 205.*

[7] P. N. Anh, L. D. Muu, N. V. Hien, J. J. Strodiot, *Using the Banach contraction principle to implement the proximal point method for multivalued monotone variational inequalities, J. Optim. Theory Appl.,* **124,** *(2005) 285-306.*

[8] Q. H. Ansari, J. C. Yao, *A fixed point theorem and its application to a system of variational inequalities, Bull. Aust. Math. Soc.,* **59,** *(1999) 433-442.*

[9] H. Attouch, M. O. Czarnecki, *Asymptotic control and stabilization of nonlinear oscillators with nonisolated equilibria, J. Differ. Equ.,* **179,** *(2002) 278–310.*

[10] H. Attouch, X. Goudon, P. Redont *The heavy ball with friction. I. The continuous dynamical system, Commun. Contemp. Math.,* **2,** *(2000) 1–34.*

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

[12] C. Baiocchi, A. Capelo, *Variational and quasivariational inequalities: Applications to free boundary problems, Wiley, New York, (1984).*

[13] A. B. Bakushinskii, *Methods for the solution of monotone variational inequalities that are based on the principle of iterative regularization, Zh. Vychisl. Mat. Mat. Fiz.,* **17,** *(1977) 1350-1362.*

[14] H. H. Bauschke, J. M. Borwein, *On projection algorithms for solving convex feasibility problems, SIAM,* **38,** *(1996), 367-426.*

[15] H. H. Bauschke, P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces , Springer, ISBN 978-3-319-48310-8, (2017).*

[16] J. Bracken, J. T. McGill, *Mathematical programs with optimization problems in the constraints, Oper. Res.,* **21,** *(1973) 37-44.*

[17] W. Candler, R. Norton, *Multilevel programming and development policy, Tech. Rep. 258, World Bank Staff, Washington, USA, (1977).*

[18] Y. Censor, T. Elfving, *A multiplication algorithm using Bregman projections in a product space, Numer. Algorithms,* **8,** *(1994) 221-239.*

[19] Y. Censor, A. Gabali, S. Reich, *Strong convergenceof subgradient extragradient methods for the variational inequality problem in Hilbert space, Optim. Methods Softw.,* **26***, (2011) 827-845.*

[20] L. C. Ceng,. N. Hadjisavvas, N. C. Wong, *Strong convergence theorem by a hybrid extragradient-like approximation method for variational inequalities and fixed point problems, J. of Glob. Optim.,* **46***, (2010) 635-646.*

[21] L. C. Ceng, *Two inertial linesearch extragradient algorithms for the bilevel split pseudomonotone variational inequality with constraints, J. Appl. Numer. Optim.,* **2***, (2020) 213-233.*

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

[23] Y. Censor, A. Gabali, S. Reich, *Algorithms for the split variational inequality problem, Numer. Algorithms,* **59***, (2012) 301-323.*

[24] 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).*

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

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

[27] S. Dempe, *bundle algorithm applied to bilevel programming problems with non-unique lower level solutions, A Comput. Optim. Appl.,* **15***, (2000) 145–166.*

[28] S. Dempe, J. F. Bard, *Bundle trust-region algorithm for bilinear bilevel programming, J. Optim. Theory Appl.* **110**, *(2001) 265–288.*

[29] S. Dempe, V. Kalashnikov, G. A. Pérez-Valdés, N. Kalashnykova, *Bilevel Programming Problems, Springer, Berlin,* **52**, *(2015) 333–359.*

[30] S. Dempe, H. Schmidt, *On an algorithm solving two-level programming problems with nonunique lower level solutions, Comput. Optim. Appl.,* **6**, *(1996) 227–249.*

[31] F. Facchinei, J. S. Pang, *Finite - Dimensional Variational Inequalities and Complementarity Problems, Berlin: Springer, (2003).*

[32] X. Gaewskli, K. Greger, K. Zacharias, *Nichtlineare Operatorgleihungen and Operatordifferentialgleihungen, Mathematishe Monographien, Band 38, Academie-Verlag, Berlin, (1974).*

[33] G. Fichera, *Sul pproblem elastostatico di signorini con ambigue contorno, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.,* **34**, *(1963) 138-142.*

[34] I.M. Gel′fand, S.V. Fomin, *Calculus of Variations, Englewood Cliffs, N.J., Prentice-Hall, (1963).*

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

[36] A. Gibali, D.V. Hieu, *A new inertial double-projection method for solving variational inequalities, J. Fixed Point Theory Appl.,* **21** *(2019) 97.*

[37] J. Glackin, J. G. Ecker, M. Kupferschmid, *Solving bilevel linear programs using multiple objective linear programming, J Optim Theory Appl,* **140**, *(2009) 197-212.*

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

[39] N. M. Hai, L. H. Van, T. V. Anh, *An algorithm for a class of bilevel variational inequalities with split variational inequality and fixed point problem constriants, Acta Math. Vietnam.,* **46**, *(2021) 515-530.*

[40] B. S. He, *A class of projection and contraction methods for monotone variational inequalities, Appl. Math. Optim.,* **35**, *(1997) 69-76.*

[41] S. He, Q-L. Dong, H. Tian *Relaxed projection and contraction methods for solving Lipschitz continuous monotone variational inequalities, RACSAM,* **113** *(2019), 2763-2781.*

[42] B. S. He, L. Z. Liao, *Improvements of some projection methods for monotone nonlinear variational inequalities, J Optim Theory Appl,* **112**, *(2002) 111-128.*

[43] H. He, C. Ling, H.K. Xu, A relaxed projection method for split variational inequalities, *J. Optim. Theory Appl.,* **166** *(2015), 213-233.*

[44] S. He, H. K. Xu, *Uniqueness of supporting hyperplanes and an alternative to solutions of variational inequalities, J. Global Optim.,* **57**, *(2013) 1375-1384.*

[45] S.X. He, P. Zhang, X. Hu, R. Hu, *A sample average approximation method based on a D-gap function for stochastic variational inequality problems, J. Indust. Manag. Optim.,* **10**, *(2014) 977–987.*

[46] D.V. Hieu, A. Moudafi, *Regularization projection method for solving bilevel variational inequality problem, Optim. Lett.,* **15**, *(2020) 205–229.*

[47] D. V. Hieu, L. D. Muu, H. N. Duong, B. H. Thai, *Strong convergence of new iterative projection methods with regularization for solving monotone variational inequalities in Hilbert spaces., Math. Methods Appl. Sci.,* **43**, *(2020) 9745–9765.*

[48] H. Iiduka, I. Yamada, *A subgradient-type method for the equilibrium problem over the fixed point set and its applications, Optimization,* **58**, *(2009) 251-261.*

[49] A. N. Iusem, B. F. Svaiter, *A variant of Korpelevich's method for variational inequalities with a new search strategy, Optimization,* **42**, *(1997) 309–321.*

[50] C. Izuchukwu, S. Reich, Y. Shehu, A. Taiwo, *Strong Convergence of Forward–Reflected–Backward Splitting Methods for Solving Monotone Inclusions with Applications to Image Restoration and Optimal Control., J. Sci. Comput.,* **94**, *(2023) 73.*

[51] B. Jadamba, F. Raciti, *Variational inequality approach to stochastic Nash equilibrium problems with an application to Cournot oligopoly, J. Optim. Theory Appl.,* **165**, *(2015) 1050–1070.*

[52] H. Jiang, H. Xu, *Stochastic approximation approaches to the stochastic variational inequality problem, IEEE Trans. Automat. Control,* **53**, *(2008) 1462–1475.*

[53] A. Kannan, U.V. Shanbhag, *The pseudomonotone stochastic variational inequality problem: analytical statements and stochastic extragradient schemes, In: The 2014 Conference on American Control Conference,* **53**, *(2014) 2930–2935.*

[54] P. D. Khanh and P. T. Vuong, *Modified projection method for strongly pseudomonotone variational inequalities, J. Global Optim.,* **58**, *(2014), 341-350.*

[55] E. N. Khobotov, *Modification of the extragradient method for solving variational inequalities and certain optimization problems, USSR Comput. Math. Math. Phys.,* **27** *(1989), 120–127.*

[56] D. Kinderlehrer, G. M. Korpelevich, *An introduction to variational inequalities and their applications, New York: Academic Press, (1980).*

[57] I. V. Konnov, W. Takahashi, *Strong convergence theorem for an infinite family of demimetric mappings in a Hilbert space, J. Convex Anal.,* **24** *(2017).*

[58] I. V. Konnov, *Combined relaxation methods for variational inequalities, Berlin: Springer, (2001).*

[59] G. M. Korpelevich, *The extragradient method for finding saddle points and other problems, Ekonomika i Mathematicheskie Metody,* **12**, *(1976) 747-756.*

[60] A.S. Kravchyk, *Variational and Quasi-variational Inequalities in Mechanics, MGAPI, Moscow, (1997).*

[61] G. Li, Z. Wan, J. Chan, X. Zhao, *Existence of solution and algorithms for a class of bilevel variational inequalities with hierarchical nesting structure, J. Fixed Point Theory Appl.,* **1**, *(2016) 1-21, https://doi.org/10.1186/s13663-016-0524-5.*

[62] G.H. Lin, M. Fukushima, *Stochastic equilibrium problems and stochastic mathematical programs with equilibrium constraints: A survey, Pac. J. Optim.,* **6**, *(2010) 455–482.*

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

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

[65] Y. Malistky, *Projected reflected gradient methods for monotone variational inequalities, SIAM J. Optim.* **25,** *(2015) 502-520.*

[66] Y. V. Malitsky, V. V. Semenov, *An extragradient algorithm for monotone variational inequalities, Cybernet. Systems Anal.* **50,** *(2014), 271–277.*

[67] Y. Malistky, M. K. Tam, *A forward-backward splitting method for monotone inclusioons without copcoercivity, SIAM J. Optim.* **30,** *(2020) 1451-1472.*

[68] B. Mordukhovich, *Variational Analysis and Applications, Springer, Cham (2018).*

[69] J. Morgan, F. Patrone, *Stackelberg problems: Subgame perfect equilibria via Tikhonov regularization. In: Haurie, A., Muto, S., Petrosjan, L.A., Raghavan, T.E.S. (eds.) Advances in dynamic games,* **8,** *(2020) 209–221.*

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

[71] L. D. Muu, N. V. Quy, *On Existence and Solution Methods for Strongly Pseudomonotone Equilibrium Problems, J. Math.* **43,** *(2015) 229–238.*

[72] 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.*

[73] F. U. Ogbuisi, Y. Shehu, J. C. Yao, *Relaxed single projection methods for solving bilevel variational inequality problems in Hilbert Spaces, Netw. Spat. Econ., (2023), https://doi.org/10.1007/s11067-023- 09594-z.*

[74] B.T. Polyak, *Some methods of speeding up the convergence of iterates methods, U.S.S.R Comput. Math. Phys.,* **4**, *(1964) 1-17.*

[75] L. D. Popov, *A modification of the Arrow-Hurwicz method for finding saddle points, Math. Notes,* **28**, *(1980), 845–848.*

[76] J.L. Lions, *Quelques methodes de resolution des problems aux limites non lin´eaires, Dunod, Paris, (1969).*

[77] U. Ravat, U.V. Shanbhag, *On the characterization of solution sets of smooth and nonsmooth convex stochastic Nash games, SIAM J. Optim.,* **21**, *(2011) 1168–1199.*

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

[79] A. Sabharwal, L. C. Potter, *Convexly constrained linear inverse problems: iterative least-squares and regularization, IEEE Trans. Signal Process.,* **46**, *(1998) 2345-2352.*

[80] S. Saejung, P. Yotkaew, *Approximation of zeros of inverse strongly monotone operators in Banach spaces, Nonlinear Anal.,* **75**, *(2012) 742–750.*

[81] E. Schmidt, *Über die Auflösung linearer Gleichungen mit unendlich vielen Unbekannten, Rend. Circ. Mat. Palermo,* **25**, *(1908) 63–77. doi:10.1007/BF03029116.*

[82] M. V. Solodov, *An explicit descent method for convex optimization, J. Convex Anal.,* **14**, *(2007) 227-237.*

[83] M. V. Solodov, B. F. Svaiter, *A new projection method for variational inequality problems, SIAM J. Control Optim.,* **37**, *(1999), 765–776.*

[84] H. V. Stackelberg, *Marktform und Gleichgewicht, Julius Springer, Vienna, Austria, (1934).*

[85] G. Stampacchia, *Variational inequalities in theory and applications of momtone operators, Vanice, Italy, (1998).*

[86] A. Taiwo, S. Reich, *Bounded perturbation resilience of a regularized forward-reflected-backward splitting method for solving variational inclusion problems with applications, Optimization, https://doi.org/: 10.1080/02331934.2023.2187664.*

[87] B. Tan, L. Liu, X. Qin, *Self adaptive inertial extragradient algorithms for solving bilevel pseudomonotone variational inequality problems, Jpn. J. Ind. Appl. Math.,* **38,** *(2021) 519-543.*

[88] B. Tan, S. Li, X. Qin, *An accelerated extragradient algorithm for bilevel pseudomonotone variational inequality problems with application to optimal control problems, RACSAM REV R ACAD A,* **115,** *(2021) 174.*

[89] B. Tan, X. Qin, S. Y. Cho, *Revisiting subgradient extragradient methods for solving variational inequalities, Numer. Algorithms,* **90,** *(2022) 1593-1615.*

[90] B. Tan, X. Qin, J. Yao, *Two modified inertial projection algorithms for bilevel pseudomontone variational inequalities with applications to optimal control problems, Numer. Algorithms,* **88,** *(2021) 1756-1786.*

[91] D.V. Thong, D.V. Hieu, *A strong convergence of modified subgradient extragradient method for solving bilevel pseudomonotone variational inequality problems, Optimization,* **69,** *(2020) 1313-1334.*

[92] D.V. Thong, N.A. Triet, X.H. Li, , Q. L. Dong, *Strong convergence of extragradient methods for solving bilevel pseudo-monotone variational inequality problems, Numer. Algorithms,* **83,** *(2020) 1123–1143.*

[93] D. V. Thong, X. H. Li, Q. L. Dong, Y. J. Cho, T. M. Rassias, *A projection and contraction method with adaptive step sizes for solving bilevel pseudo-monotone variational inequality problems, Optimization, (2020) https://doi.org/10.1080/02331934.2020.1849206.*

[94] A.N. Tikhonov, *On the stability of inverse problems, Acad. Sci. URSS,* **39**, *(1943), 176-179.*

[95] C. R. Trujillo, S. Zlobec, *Bilevel convex programming models, Optimization,* **58**, *(2009) 1009-1028.*

[96] P. Tseng, *A modified forward-backward spliting method for maximal monotone mappings, SIAM J. Control Optim.,* **38**, *(2000) 431-446.*

[97] Z. Wan, J. W. Chen, *On Bilevel Variational Inequalities, J. Oper. Res. Soc. China,* **1**, *(2013) 483–510, https://doi.org/10.1007/s40305-013-0036-5.*

[98] M.Z. Wang, G.H. Lin, Y.L. Gao, M.M. Ali, *Sample average approximation method for a class of stochastic variational inequality problems, J. Syst. Sci. Comp.,* **24**, *(2011) 1143–1153.*

[99] S.H. Wang, H.Y. Tao, R.G. Lin, Y.J. Cho, *A self-adaptive stochastic subgradient extragradient algorithm for the stochastic pseudomonotone variational inequality problem with application, Z. Angew. Math. Phys.,* **73**, *(2022) 164.*

[100] H. K. Xu, *Another control condition in an iterative method for nonexpansive mappings, Bull. Aust. Math. Soc.,* **65**, *(2002) 109–113.*

[101] H. K. Xu, *Sample average approximation methods for a class of stochastic variational inequality problems, Asia-Pac. J. Oper. Res.,* **27**, *(2010) 103–119.*

[102] H. Xu, M. Li, C. C. Yang *Neural networks for a class of bilevel variational inequalities, J. Glob. Optim.,* **44**, *(2009) 535-552.*

[103] I. Yamada, *The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings, Inherently Parallel Algorithms Feasibility Optim. Their Appl.,* **8**, *(2001) 473–504.*

[104] Z. P. Yang, Y.L. Wang, G.H. Lin, *Variance-based modified backward-forward algorithm with line search for stochastic variational inequality problems and its applications, Asia-Pac J. Oper. Res.,* **37**, *(2020) 2050011.*

[105] Z. P. Yang, J. Zhang, Y.L. Wang, G.H. Lin, *Variance-based subgradient extragradient method for stochastic variational inequality problems, J. Sci. Comput.,* **89***, (2021) 1–28.*

[106] Y. Yao, G. Marino, L. Muglia, *A modified korpelevich's method convergent to the minimum-norm solution of a variational inequality, Optimization,* **63***, (2014) 559-569.*

[107] M. Yashtini, A. Malek, *Solving complementarity and variational inequalities using neural networks, Appl. Math. and Comp.,* **190***, (2007), 216-230.*

[108] H. Zegeye, N. Shahzad, Y. Yao, *Minimum-norm solution of variation inequality and fixed point problem in Banach spaces, Optimization,* **64***, (2015) 453-471.*