beforestart english

NUMERICAL ALGORITHMS FOR OPTIMIZATION PROBLEMS AND APPLICATIONS

KUNRADA KANKAM

A Dissertation Submitted to University of Phayao

in Partial Fulfillment of the Requirements

for the Philosophy of Science Degree in Mathematics

May 2023

NUMERICAL ALGORITHMS FOR OPTIMIZATION PROBLEMS AND APPLICATIONS

KUNRADA KANKAM

A Dissertation Submitted to University of Phayao
in Partial Fulfillment of the Requirements
for the Philosophy of Science Degree in Mathematics
May 2023

Dissertation

Title

Numerical Algorithms for Optimization Problems and Applications

Submitted by Kunrada Kankam

Approved in partial fulfillment of the requirements for the

Philosophy of Science Degree in Mathematics

University of Phayao

....Chairman

(Professor Dr.Suthep Suantai)

....Advisor

(Associate Professor Dr.Prasit Cholamjiak)

....Internal Committee

(Associate Professor Dr.Tanakit Thianwan)

....Internal Committee

(Associate Professor Dr.Damrongsak Yambangwai)

....Internal Committee

(Associate Professor Dr.Watcharaporn Cholamjiak)

....External Committee

(Professor Dr. Poom Kumam)

Approved by

....

(Associate Professor Dr.Chayan Boonyarak)

Dean of School of Science

May 2023

ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude to my supervisor, Associate Professor Dr.Prasit Cholamjiak, for being an incredible source of inspiration and knowledge and for sharing all his fantastic scientific ideas openly with me while encouraging me to explore and implement new pictures of my own. Also, I would like to thank Associate Professor Dr.Watcharaporn Cholamjiak for your technical support in my study, encouragement and continued support. I am thankful to them for their help, warm relationships, and significant care until this study completion.

I greatly appreciate the project International Research Network in Digital Image Processing and Machine Learning (IRN62W007) for financial support during my Ph.D. study.

I wish to thank Assistant Professor Dr. Yann Savoye for their taking care, guidance and academic support while I was at Liverpool John Moores University, England. Moreover, I would like to thank Professor Dr.Choonkil Park and Professor Dr.Yeol Je Cho and family for taking care while I was in Korea.

A special thanks to my dissertation committees, Professor Dr.Suthep Suantai, Professor Dr. Poom Kumam, Associate Professor Dr.Tanakit Thianwan and Associate Professor Dr.Damrongsak Yambangwai motivated and introduced me when I was an undergraduate and master's student. It is my pleasure to express my thank to all of them for their generous assistance.

Also, I thank my family and friends for all their support and encouragement to go on with my Ph.D. study. Finally, I am grateful to myself for believing in myself.

Kunrada Kankam

เรื่อง: ขั้นตอนวิธีเชิงตัวเลขสำหรับปัญหาค่าเหมาะที่สุดและการประยุกต์
ผู้วิจัย: กุลดา ขันคำ, วิทยานิพนธ์: ปร.ด. (คณิตศาสตร์), มหาวิทยาลัยพะเยา, 2566
ประธานที่ปรึกษา: รองศาสตราจารย์ ดร.ประสิทธิ์ ช่อลำเจียก
กรรมการที่ปรึกษา: รองศาสตราจารย์ ดร.ธนกฤต เทียนหวาน, รองศาสตราจารย์ ดร.ดำรงศักดิ์ แย้มบางหวาย, รองศาสตราจารย์ ดร.วัชรภรณ์ ช่อลำเจียก
คำสำคัญ: วิธีการข้างหน้า–ข้างหลัง, เทคนิคแบบเฉื่อย, การหาค่าเหมาะสุดเชิงคอนเวกซ์, ปริภูมิฮิลเบิร์ต, การกู้คืนภาพ, การซ่อมแซมภาพ

บทคัดย่อ

ปัญหามากมายทางด้านวิทยาศาสตร์ประยุกต์ วิศวกรรมศาสตร์ และ เศรษฐศาสตร์ สามารถแปลงให้อยู่ในรูปแบบของปัญหาค่าต่ำสุดเชิงคอนเวกซ์ของผลรวมของสองฟังก์ชันกึ่งต่อเนื่องล่างโดยแท้และฟังก์ชันคอนเวกซ์ เพื่อที่จะแก้ปัญหานี้วิธีการแยกข้างหน้า–ข้างหลังได้ถูกนำมาใช้สำหรับการวิเคราะห์การลู่เข้า ซึ่งวิธีการนี้เป็นที่สนใจอย่างมากในด้านการหาค่าเหมาะที่สุดเพื่อแก้ปัญหาเชิงประยุกต์ต่างๆ เช่น การประมวลผลภาพและสัญญาณ การขนส่ง การวิเคราะห์ข้อมูลถดถอย การจำแนกข้อมูล เป็นต้น วัตถุประสงค์หลักของวิทยานิพนธ์นี้คือ การปรับปรุงและพัฒนาวิธีการแยกแบบใหม่สำหรับการแก้ปัญหาค่าต่ำสุดเชิงคอนเวกซ์โดยการปรับวิธีการที่มีอยู่และเพิ่มใช้เทคนิคต่างๆสำหรับเงื่อนไขแบบเฉื่อย ทฤษฎีบทการลู่เข้าได้ถูกสร้างขึ้นภายใต้เงื่อนไขที่เหมาะสมในขอบเขตของปริภูมิฮิลเบิร์ต สุดท้ายการทดลองเชิงตัวเลขแสดงให้เห็นถึงประสิทธิภาพของวิธีการที่ได้นำเสนอในการประมวลผลภาพ ผลลัพธ์เชิงตัวเลขแสดงให้เห็นว่าวิธีการที่ได้ถูกนำเสนอมีอัตราการลู่เข้าที่ดีกว่าวิธีการอื่นที่เกี่ยวข้อง

Title: NUMERICAL ALGORITHMS FOR OPTIMIZATION PROBLEMS AND APPLICATIONS
Author: Kunrada Kankam, Dissertation: Ph.d. (Mathematics), University of Phayao, 2023
Advisor: Associate Professor Dr.Prasit Cholamjiak
Committee: Associate Professor Dr. Tanakit Thianwan, Associate Professor Dr. Damrongsak Yambangwai, Associate Professor Dr. Watcharaporn Cholamjiak
Keywords: forward–backward method, inertial technique, convex minimization, Hilbert space, image deblurring, image inpainting

ABSTRACT

Many real–world problems in applied sciences, engineering and economics can be reformulated as the convex minimization problem of the sum of two proper lower–semicontinuous and convex functions. In order to solve this problem, the forward–backward splitting algorithm has been used for the convergence analysis. It has a great attention in optimization due to its broad application to many disciplines such as signal and image processing, transportation, data regression, and classification problems. The main objective of this dissertation is to improve and develop new splitting algorithms for solving convex minimization problems by adapting with another method and many techniques with inertial conditions. The convergence theorems are established under suitable conditions in the framework of Hilbert spaces. Finally, numerical experiments are given to show the efficiency of the proposed methods in image processing. Numerical results show that the proposed algorithms have a better convergence than some known algorithms.

LIST OF CONTENTS

LIST OF TABLES

Tables **Page**

LIST OF FIGURES

Figures **Page**

CHAPTER 1

INTRODUCTION

To solve basic optimization problems, researchers may use algorithm that terminat in finite number of steps, or heuristics that may provide approximate solutions to problem, one of classical techniques is the gradient method proposed by B.T. Polyak in 1963 [32] and it can be described by the fixed point equation in 1976 by J.C. Dunn [15] . The proximal point methods are widely used to deal with one maximal monotone operator. The proximal point algorithm initiated by B. Martinet in 1970 [23] and subsequently studied by R.T. Rockafellar in 1976 [34] is often referred. The splitting methods play a central role in the analysis and the numerical solution.Due to its applications, there have been several modifications and generalizations of these methods suggested and invented independently for solving the problem in many different contexts. This tool plays an important role in the analysis and the numerical solution of optimization problems. The main concept of technique is obtained by splitting in that the functions are used individually so as to yield an easily implementable algorithm. In 1979, P.L. Lions and B. Mercier [21] proposed splitting algorithm for the sum nonlinear operator, also sometime called forward-backward splitting algorithm. In 1956, J. Douglas and H.H. Rachford [13] proposed on the numerical solution of heat conduction problems in two and three space variables or called Douglas-Rachford splitting algorithms. The splitting methods is used to find a minimizer of the sum of two convex functions. Once the parameters for dispute model have been determined, a list of feasible solutions can be developed. In order to use optimization, a separate objective function could be developed for each party to represent its interest. In 1964, B.T. Polyak [33] introduced the algorithm that can speed up gradient descent and it modification was made immensely popular by Nesterov's accelerated gragient algorithm which was algorithm proposed by Y. Nesterov in

1983 [28]. This well-known method that has improved the convergence rate is know as the inertial iteration for operator.

Many problems have constraints. Each constraint limits the set of possible solutions and together the constraints define the feasible set. Feasible design points do not violate any constraints. In 1964, A.A. Goldstein [16] and in 1966, E.S. Levitin, B.T. Polyak [20] proposed gradients projection methods for minimizing a continuously differentiable mapping were originally. It is helpful to study the constrained minimization problem. Next, in 2000, P. Tseng [37] introduced the following modified forward-backward splitting method, also known as Tseng's splitting algorithm, for investigate them, a projected forward-backward algorithm for fine minimizer of the sum of two convex function. However, many proximal point method usually assumed that the gradient is Lipschitz continuous and the step size is bounded below and less than some constants related to the Lipschitz constant, which is some how not known in practice. For this reason, it is our purpose to study and develop the proximal mapping involving new techniques for solving minimization problems. In 1976, D.P. Bertsekas [5] was the first to propose a practical finite procedure to determine the step. In 2016, J.Y.B. Cruz and T.A. Nghia [4] proposed the proximal gradient algorithm using linesearch technique for solving the convex minimization problem in Hilbert spaces. The main advantage of the proposed method is that the Lipschitz condition on the gradient of functions is dropped in computing. As reviewed, it is therefore the main objective in this research to develop and modify the numerical algorithms by using the prvious technique for solving minimization problems and to establish some convergence theorems. The main results established in this research can improve and generalize the corresponding results in this area and, of course, can be applied to solve major problems existed in science, engineering, economics and other related branches. Finally, to give some applications of optimization problem including its numerical experiments. The main results can improve and extend the corresponding results in this area and, can be applied to solve major problems existed in science.

CHAPTER 2

REVIEW OF RELATED LITERATURE AND RESEARCH

In optimization theory, many problems can be formulated as the minimization problem which is modeled as the following basic form:

$$\min_{x \in H} f(x), \tag{2.1.1}$$

where f is a functional in the real Hilbert space H. Any value of x from among all points in Hilbert space H that minimizes the objective function is called a solution or minimizer. A particular solution is written as x_*.

In 1963, B.T. Polyak [32] proposed gradient method for the minimization of functionals, gradient minimizing method consist in constructing the sequence (x^n) according to the formula:

$$x^{n+1} = x^n - \alpha_n \nabla f(x^n), \tag{2.1.2}$$

where $\alpha_n > 0$ is the stepsize and ∇f is the gradient of f, then any limit point of (x^n) is a stationary point of problem (2.1.1).

In 1964, B.T. Polyak [33] introduced the following algorithm that can speed up gradient descent as follows:

Algorithm 2.1.1 Let f be a functional in the real Hilbert space H and $x^0 \in H$. Calculate

$$\begin{aligned} y^n &= x^n + \beta(x^n - x^{n-1}) \\ x^{n+1} &= y^n - \alpha \nabla f(x^n), \end{aligned}$$

where $\alpha, \beta > 0$. The parameter β is called the inertial term that accelerates the convergence.

This modification was made immensely popular by Nesterov's accelerated gradient algorithm.

Algorithm 2.1.2 Let T be an operator and let $x^0, x^1 \in H$, generate

$$\begin{aligned} y^n &= x^n + \beta_n(x^n - x^{n-1}) \\ x^{n+1} &= T(y^n), \end{aligned}$$

where $\beta_n > 0$ is the stepsize, which was proposed by Y. Nesterov in 1983 [28].

In 1976, J.C. Dunn [15] proposed convexity monotonicity and gradient processes in Hilbert space by zeros of ∇f are fixed point of $T = I - \nabla f$ and conversely. Furthermore (2.1.2) may be put into form

$$x^{n+1} = (1 - \alpha_n)x^n + \alpha_n T x^n,$$

where $\alpha_n \in [0, 1]$ generates a sequence (x^n) in nonempty closed convex subset of H. Furthermore, the associated defect norm sequence $(\|Tx^n - x^n\|)$ is monotone nonincreasing and therefore converges downward to a limit Lipschitz constant $L > 0$.

Many problems have constraints. Each constraint limits the set of possible solutions, and together the constraints define the feasible set Ω. For solving these problems usually restrict the change in the working subspace by only dropping or adding one constraint at each iteration.This apply to methods based on projected gradients.

Gradients projection methods for minimizing a continuously differentiable mapping were originally proposed by A.A. Goldstein (1964) [16] and E.S. Levitin, B.T. Polyak (1966) [20]. It is helpful to study the constrained minimization problem which is modeled as the following form:

$$\min_{x\in\Omega} f(x), \tag{2.1.3}$$

where Ω is a nonempty closed convex subset of H. The method assumes one can constructively project points onto a nonempty closed convex set and denote the projection of x into Ω by $P_\Omega(x)$. Given the projection P into Ω, the gradient projection algorithm is defined by

$$x^{n+1} = P_\Omega(x^n - \alpha_n \nabla f(x^n)),$$

where $\alpha_n > 0$ is the stepsize and ∇f is the gradient of f with respect to the inner product associated with the norm. The results show that if ∇f is Lipschitz continuous with $L-$Lipschitz constant and if $\varepsilon \le \alpha_n \le \frac{2}{L}(1-\varepsilon)$ for some $\varepsilon \in (0,1)$, then any limit point of (x^n) is a stationary point of problem (2.1.3).

In 1976, D.P. Bertsekas [5] was the first to propose a practical finite procedure to determine the step. The algorithm consists of the iteration:

Algorithm 2.1.3 Given $\sigma > 0, \theta \in (0,1)$ and $\delta \in (0,1)$. Let $x^0 \in \Omega$. For $n \ge 1$, calculate

$$x^{n+1} = P_\Omega(x^n - \alpha_n \nabla f(x^n)),$$

where $\alpha_n = \sigma\theta^{m_n}$ and m_n is the smallest m nonnegative integer such that

$$f(x^n(\sigma\theta^m)) \le f(x^n) - \delta\frac{|x^n - x^n(\sigma\theta^m)|^2}{\sigma\theta^m}.$$

Assuming that f is continuously differentiable and Ω is the convex set and limit points of (x^n) are stationary points of problem (2.1.3).

In 1976, R.T. Rockafellar [34] proposed monotone operator and the proximal point algorithm based on the fact of G.T. Minty (1962) [24] that for each $x \in H$ and $\alpha > 0$ there is unique x_* such that $x_* - x \in \alpha T(x)$ or in other word:

$$x_* \in (I + \alpha T)(x).$$

The operator $\text{prox}_{\alpha T} = (I+\alpha T)^{-1}$ is single-valued from all of H into H and prox_{α} is called the proximal mapping associated with αT, it following the terminology of J.J. Moreau (1965) [25] for the case of $T = \partial f$. The proximal point algorithm is generated by

$$x^{n+1} = \text{prox}_{\alpha_n f}(x^n),$$

where $\text{prox}_{\alpha_n f} = (I + \alpha_n \partial f)^{-1}$ and (α_n) is a sequence of positive real numbers.

In many problems, however, can be modeled in the following form:

$$\min_{x \in H} f(x) + g(x), \tag{2.1.4}$$

where $f : H \to (-\infty, +\infty]$ and $g : H \to \mathbb{R}$ are two proper lower semi-continuous convex function such that f is differentiable on H.

In 1979, P.L. Lions and B. Mercier [21] study two splitting algorithms for sum of two monotone operators and proposed splitting algorithm for the sum nonlinear operator. The algorithm consists of the iteration:

Algorithm 2.1.4 Given f and g are maximal monotone operators, generate

$$x^{n+1} = (I + \alpha \partial g)^{-1}(I - \alpha \nabla f)x^n,$$

where α is sufficiently small if f is Lipschitz continuous.

In 2003, A. Moudafi and M. Oliny [26] proposed the first inertial forward-backward under the setting of finding zeros of monotone inclusion problem. Specify the algorithm to the case of solving (2.1.4), they obtain the following iteration:

Algorithm 2.1.5 Given $f : H \to (-\infty, +\infty]$ is convex differentiable and $g : H \to (-\infty, +\infty]$ is proper convex and lower simi-continuous. Generate

$$\begin{aligned} y^n &= x^n + \theta_n(x^n - x^{n-1}), \\ x^{n+1} &= \operatorname{prox}_{\alpha g}(y^n - \alpha \nabla f(x^n)), \end{aligned}$$

where $\alpha_n \in (0, 2/L)$ and θ_n is the inertial parameter which controls the momentum $x^n - x^{n-1}$.

The above scheme recovers the heavy-ball method when $g = 0$ [29] and become the scheme of [22] if replace $\nabla f(x^n)$ with $\nabla f(y^n)$.

In 2005, P.L. Combettes and V.R. Wajs [10] study the properties of problems (2.1.4) and analyze the convergence of forward-backward splitting algorithm. The fixed point characterization provided

$$x = \operatorname{prox}_{\alpha g}(x - \alpha \nabla f(x)),$$

where $\operatorname{prox}_{\alpha g} = (I + \alpha \partial g)^{-1}$ is called the resolvent of the operator ∂g with a parameter $\alpha > 0$, suggests solving problems (2.1.4) via the fixed point iteration

$$x^{n+1} = \operatorname{prox}_{\alpha g}(x^n - \alpha \nabla f(x^n)),$$

where α is a suitable value. P.L. Combettes and V.R. Wajs [10] proposed main convergence result generated by

Algorithm 2.1.6 (RFB) Given $f : H \to (-\infty, +\infty]$ and $g : H \to \mathbb{R}$ are two proper lower semi-continuous convex functions such that f is differentiable on H with a $\frac{1}{L}-$Lipschitz continuous gradient for some $L \in (0, +\infty)$. Let (a_n) and (b_n) be sequences in H such that $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$ and $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$. Fix $x^0 \in H$, calculate

$$\begin{aligned} y^n &= x^n - \lambda_n(\nabla f(x^n) + b_n) \\ x^{n+1} &= x^n + \alpha_n(\text{prox}_{\alpha_n g}(y^n) + a_n - x^n), \end{aligned}$$

where $\lambda \in (0, \infty)$ such that $0 < \inf_{n\in\mathbb{N}} \lambda_n \leq \sup_{n\in\mathbb{N}} \lambda_n < 2L$ and $\alpha_n \in (0, 1]$ such that $\inf_{n\in\mathbb{N}} \alpha_n > 0$. The results show that (x^n) converges weakly to a stationary point of problem (2.1.4).

In 2011, P.L. Combettes and J.C. Pesquet [11] review the basic properties of proximity operators and present optimization methods based on these operators. These proximal splitting methods are shown to capture and extend several well-known algorithms in a unifying framework such as

Algorithm 2.1.7 (Forward-backward algorithm)
Fix $\varepsilon \in (0, \min(1, \frac{1}{L}))$ and $x^0 \in \mathbb{R}^N$, calculate

$$\begin{aligned} y^n &= x^n - \alpha_n \nabla f(x^n) \\ x^{n+1} &= x^n + \beta_n(\text{prox}_{\alpha_n g} y^n - x^n), \end{aligned}$$

where $\alpha_n \in [L, 2/L - \varepsilon]$ and $\beta_n \in [\varepsilon, 1]$. The results show that sequence (x^n) generated by Forward-backward algorithm converges to a solution to (2.1.4) where $f \in \Gamma_0(\mathbb{R})$, $\Gamma_0(\mathbb{R})$ is the class of lower semicontinuous convex functions from $\mathbb{R}^N$ and $g : \mathbb{R}^N \to \mathbb{R}$ are convex and differentiable with a $L-$Lipschitz continuous gradient ∇f.

Algorithm 2.1.8 (Beck-Teboulle proximal gradient algorithm)[3]
(FISTA-BT) Fix $x^0 \in \mathbb{R}^N$, set $z^0 = x^0$ and $t_0 = 1$, calculate

$$\begin{aligned} y^n &= z^n - L^{-1}\nabla f(z^n) \\ x^{n+1} &= \operatorname{prox}_{L^{-1}g} y^n \\ z^{n+1} &= x^n + \alpha_n(x^{n+1} - x^n), \end{aligned}$$

where $\alpha_n = 1 + \frac{t_n - 1}{t_{n+1}}$ and $t_{n+1} = \frac{1+\sqrt{4t_n^2+1}}{2}$.

While little is known about the actual convergence of sequences produced by Beck-Teboulle proximal gradient algorithm, the $O(\frac{1}{n^2})$ rate of convergence of the objective function they achieve is optimal.

Algorithm 2.1.9 (Douglas-Rachford algorithm)
Fix $\varepsilon \in (0,1)$ and $\alpha > 0$, let $y^0 \in \mathbb{R}^N$ and calculate

$$\begin{aligned} y^n &= \operatorname{prox}_{\alpha f} y^n \\ y^{n+1} &= y^n + \beta_n(\operatorname{prox}_{\alpha g}(2x^n - y^n) - x^n), \end{aligned}$$

where $\beta_n \in [\varepsilon, 2-\varepsilon]$. The results show that the sequence (x^n) generated by Douglas-Rachford algorithm converges to a solution to (2.1.4) where f and g are functions in the class of lower semi-continuous convex functions from $\mathbb{R}^N$.

In 2016, J.Y.B. Cruz and T.A. Nghia [4] proposed the proximal gradient algorithm using linesearch technique for solving the convex minimization problem in Hilbert spaces. The main advantage of the proposed method is that the Lipschitz condition on the gradient of functions is dropped in computing. The linesearch is defined as follows:

Algorithm 2.1.10 (FBS-CN) Given $\sigma > 0, \theta \in (0,1)$ and $\delta \in (0, \frac{1}{2})$. Let $x^0 \in$

dom$g$, calculate

$$x^{n+1} = \text{prox}_{\alpha_n g}(x^n - \alpha_n \nabla f(x^n)),$$

where $\alpha_n = \sigma\theta^{m_n}$ and m_n is the smallest nonnegative integer such that

$$\alpha_n \|\nabla f(x^{n+1}) - \nabla f(x^n)\| \leq \delta \|x^{n+1} - x^n\|. \tag{2.1.5}$$

The results show that the sequence (x^n) converges weakly to a solution of minimization problem (2.1.4).

In 2017, M. Verma and K.K. Schonlieb [39] proposed a new accelerated proximal gradient technique for regularized multitask learning framework as follows:

Algorithm 2.1.11 (NAGA) Given T_n is the forward-backward operator of f and g. Calculate

$$\begin{aligned} y^n &= x^n + \theta_n(x^n - x^{n-1}), \\ x^{n+1} &= T_n[(1-\alpha_n)y^n + \alpha_n T_n y^n], \end{aligned}$$

where $\alpha_n \in (0, 1)$. The results show that (x^n) converges strongly to a stationary point of problem (2.1.4).

In practical applications, many problems in real world can be modeled as a subproblem. To investigate them, a projected forward-backward algorithm for solving the constrained convex minimization problem modeled as follows:

$$\min_{x \in \Omega} f(x) + g(x), \tag{2.1.6}$$

where f and g are convex function on H, f is differentiable on H and Ω is a nonempty closed convex subset of H.

In 2000, P. Tseng [37] introduced the following modified forward-backward splitting method, also known as Tseng's splitting algorithm.

Algorithm 2.1.12 (FBFS) Given $\Omega \subseteq H$ is a closed and convex set which intersects the solution set of (2.1.4). Let $x^0 \in \Omega$. Calculate

$$\begin{aligned} y^n &= (I + \alpha_n \partial g)^{-1}(I - \alpha_n \nabla f)x^n \\ x^{n+1} &= P_\Omega(y^n - \alpha_n(\nabla f(y^n) - \nabla f(x^n))), \end{aligned}$$

where $\alpha_n \in (0, \infty)$, the results show that the sequence (x^n) converges weakly to a solution of problem (2.1.6).

In 2015, Z. Mu and Y. Peng [27] presented an alternated inertial proximal point algorithm for solving the maximal monotone inclusion problem as follows

Algorithm 2.1.13 Given T is a set-valued maximal monotone operator and $x^0, x^1 \in \mathbb{R}^N$. Calculate

$$y^n = \begin{cases} x^n & \text{if } n \text{ is even} \\ x^n + \alpha_n(x^n - x^{n-1}) & \text{if } n \text{ is odd} \end{cases}$$

and

$$x^{n+1} = (I + \beta T)^{-1}(y^n),$$

where $\beta > 0$. This form is a lot less popular than general inertia. However, it has pretty good convergence properties and performance.

In 2016, J.Y.B. Cruz and T.A. Nghia [4] proposed a fast multistep forward-backward method with Linesearch (2.1.5) as follows:

Algorithm 2.1.14 (FISTA-CN) Fix $x^0 \in \mathbb{R}^N$, set $x^0 = x^1 \in H$, $t_0 = 1$ and $\delta \in (0, 1/2)$, calculate

$$\begin{aligned} t_{n+1} &= \frac{1+\sqrt{4t_n^2+1}}{2} \\ y^n &= P_\Omega\left[x^n + \left(\frac{t_n - 1}{t_{n+1}}\right)(x^n - x^{n-1})\right] \\ x^{n+1} &= \text{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)), \end{aligned}$$

where $\alpha_n :=$ Linesearch (2.1.5). The results show that the sequence (x^n) converges weakly to a solution of problem (2.1.6).

In 2016, W. Jin, Y. Censor and M. Jiang [18] presented the projected scaled gradient (PSG) method with bounded perturbations in a finite dimensional setting for solving (2.1.3) where f is a continuously differentiable convex function. More precisely, the method is generated by

Algorithm 2.1.15 Given Ω is a nonempty closed convex subset of H and $x^0 \in \Omega$. Calculate

$$x^{n+1} = P_\Omega(x^n - \alpha_n D(x^n)\nabla f(x^n) + e(x^n)),$$

where $D(x^n)$ is a diagonal scaling matrix, then (x^n) converges to a solution of problem (2.1.3) under suitable conditions.

In 2020, P. Duan, Y. Zhang and Q. Bu [14] presented an alternated inertial proximal gradient algorithm for unconstrained convex optimization problem. Consider the following algorithm.

Algorithm 2.1.16 Let $x^0, x^1 \in H$ and compute

$$y^n = \begin{cases} x^n & \text{if } n \text{ is even} \\ x^n + \alpha_n(x^n - x^{n-1}) & \text{if } n \text{ is odd.} \end{cases}$$

Calculate the next iterate via

$$x^{n+1} = \text{prox}_{\alpha_n g}(y^n - \alpha_n D(y^n) + e(y^n)),$$

where D is a diagonal scaling matrix, then (x^n) converges weakly to a solution of problem (2.1.4) under suitable conditions.

For numerical examples and applications, the least absolute shrinkage and selection operator (lasso) received much attention from optimization community to many disciplines, such as optimal control, signal processing, system identification and image analysis. The lasso problem is

$$\min_{x\in\mathbb{R}^N} \quad (1/2)\|Ax - b\|_2^2 + \alpha\|x\|_1, \tag{2.1.7}$$

where $A \in \mathbb{R}^{M\times N}$, $b \in \mathbb{R}^M$ and $\alpha > 0$. Consider the splitting for problem (2.1.6),

$$f(x) = (1/2)\|Ax - b\|_2^2,\ g(x) = \alpha\|x\|_1$$

with $\nabla f(x) = A^T(Ax - b)$ and $\text{prox}_{\alpha g}(x) = S_\alpha(x)$, S_α is the solf-thresholding operator. Moreover, many real world problems can be formulated in the form of a minimization problem. In particular, if A is a bounded linear operator, x is a vector with to be recovered data and y is the observed or measured. It is known that to solve (2.1.7) can be seen as solving the signal recovery and image processing. In experiments on regression and classification problems, the main goal of extreme learning machine is to find

$$x = [x_1^T, ..., x_N^T]^T \text{ such that } Ax = b,$$

where A is hidden layer output matrix defied by

$$A = \begin{bmatrix} G(a_1x_1+b_1) & \cdots & G(a_Mx_1+b_M) \\ \vdots & \ddots & \vdots \\ G(a_1x_N+b_1) & \cdots & G(a_Mx_N+b_M) \end{bmatrix},$$

where G is an activate function, a_i and b_i are parameters of the i-th hidden node and $b = [b_1^T, ..., b_N^T]^T$ is the training data. The problem is formulated as the following convex minimization problem

$$\min_{x\in\mathbb{R}^N} \quad \|Ax-b\|_2^2 + \alpha\|x\|_1,$$

where α is a regularization parameter. In addition, the convex minimization problem problem is found solution by split feasibility problem.

Split feasibility problem. Consider the split feasibility problem

$$\text{find } x_* \in C \text{ such that } Ax_* \in Q,$$

where C and Q are nonempty closed convex subsets of real Hibert spaces H_1 and H_2, respectively. In case, $x \in C$ such that $Ax \in Q$ is found by

$$\min_{x\in C} \ (1/2)\|Ax-b\|_2^2 + \delta_C(x),$$

where $A : H_1 \to H_2$ is a bounded linear operator. $\delta_C(x)$ is indicator function, then split feasibility problem becomes the following problem (2.1.6).

CHAPTER 3

PRELIMINARIES

This chapter is to present the notation, definition and basic facts that are used throughout the dissertation.

3.1 Sequences of real numbers

In this section, we present definition and basic facts about sequences of real numbers. Let $\mathbb{N}$ be the set of natural numbers and let $\mathbb{R}$ be the set of real numbers. For each positive integer n, we are given a real number a^n. Then, the list of numbers,

$$a^1, a^2, \ldots, a^n, \ldots,$$

is called a *sequence* and this *ordered list* is usually written as

$$(a^1, a^2, \ldots, a^n, \ldots) \text{ or } (a^n) \text{ or } \{a^n\}.$$

Definition 3.1.1 (Convergent sequence)[35] A sequence (a^n) in $\mathbb{R}$ is said to *converge* to a real number a if for every $\varepsilon > 0$, there exists positive integer N (in general depending on ε) such that

$$|a^n - a| < \varepsilon$$

for all $n \geq N$. We denote that (a^n) converges to a by $a^n \to a$ or $\lim_{n\to\infty} a^n = a$. The number a is called a *limit* of the sequence (a^n) and (a^n) is called a *convergent sequence*.

Definition 3.1.2 (Bounded sequence)[35] A sequence (a^n) in $\mathbb{R}$ is said to *bounded* if there is $M > 0$ such that

$$|a^n| \leq M$$

for all $n \geq N$.

Remark 3.1.3 [35] Let (a^n) be a sequence in $\mathbb{R}$, if a sequence (a^n) converge to a real number a. Then (a^n) is bounded.

Definition 3.1.4 [35] Let (a^n) be a sequence of real numbers. Then, we called

· *Increasing* or *nondecreasing* if $a^n \leq a^{n+1}$ for all $n \in \mathbb{N}$.

· *Decreasing* or *nonincreasing* if $a^{n+1} \leq a^n$ for all $n \in \mathbb{N}$.

· *Strictly increasing* if $a^n < a^{n+1}$ for all $n \in \mathbb{N}$.

· *Strictly decreasing* if $a^{n+1} < a^n$ for all $n \in \mathbb{N}$.

Lemma 3.1.5 (Theorem 1.3.7 in [35]) *Let $(a^n), (b^n)$ and (c^n) be sequences of real numbers and $a^n \leq b^n \leq c^n$ for all $n \in \mathbb{N}$. If there exists $a \in \mathbb{R}$ such that $a^n \to a$ and $c^n \to a$, then $b^n \to a$.*

Definition 3.1.6 (Subsequence)[35] Let (a^n) be a sequence of real numbers and (n_i) be a sequence in $\mathbb{N}$ such that $n_i < n_{i+1}$ for all $i \in \mathbb{N}$. Then the sequence (a^{n_i}) given by

$$(a^{n_1}, a^{n_2}, \ldots, a^{n_i}, \ldots)$$

is called a *subsequence* of (a^n).

Remark 3.1.7 (Remark 2.1.5 in [35]) Let (a^n) be a sequence in $\mathbb{R}$, if a sequence (a^n) converge to a real number a. Then every subsequence (a^{n_i}) of (a^n) converge to a.

Lemma 3.1.8 (Theorem 1.2.3 in [35]) *Let (a^n) be sequences of nonnegative real numbers such that $a^n \to 0$. For $n \in \mathbb{N}$ define $\alpha_n = \sup\{a_k | k \geq n\}$, then $\alpha_n \to 0$.*

Definition 3.1.9 [35] Let (a^n) be a bounded sequence of real numbers. Then, we called

- The *limit inferior* of (a^n) defined by

$$\liminf_{n\to\infty} a^n = \lim_{n\to\infty} \alpha_n,$$

where $\alpha_n = \inf\{a^k | k \geq n\}$. It is denoted by $\liminf_{n\to\infty} a^n$.

- The *limit superior* of (a^n) defined by

$$\limsup_{n\to\infty} a^n = \lim_{n\to\infty} \alpha_n,$$

where $\alpha_n = \sup\{a^k | k \geq n\}$. It is denoted by $\limsup_{n\to\infty} a^n$.

Lemma 3.1.10 (Remark 1.4.4 in [35]) *Let (a^n) be a bounded sequence of real numbers. Then $\liminf_{n\to\infty} a^n$ and $\limsup_{n\to\infty} a^n$ exists.*

Lemma 3.1.11 (Theorem 1.4.5 in [35]) *Let (a^n) be a bounded sequence of real numbers. Then, the following hold:*

$$\inf_{n\to\infty} a^n \leq \liminf_{n\to\infty} a^n \leq \limsup_{n\to\infty} a^n \leq \sup_{n\to\infty} a^n.$$

Lemma 3.1.12 (Theorem 1.4.6 in [35]) *Let (a^n) be a bounded sequence of real numbers and $a \in \mathbb{R}$. Then, (a^n) converge to a if and only if*

$$\liminf_{n\to\infty} a^n = \limsup_{n\to\infty} a^n = a.$$

Lemma 3.1.13 (Lemma 1 in [36]) *Let $(a^n), (b^n)$ be sequences of nonnegative real numbers such that*

$$a^{n+1} \leq a^n + b^n, n \in \mathbb{N}.$$

If $\sum_{n=1}^{\infty} b^n < \infty$, *then* $\lim_{n\to\infty} a^n$ *exists.*

Lemma 3.1.14 (Lemma 1 in [30]) *Let* $(a^n), (b^n)$ *and* (δ_n) *be sequences of non-negative real numbers such that*

$$a^{n+1} \leq (1+\delta_n)a^n + b^n, n \in \mathbb{N}.$$

If $\sum_{n=1}^{\infty} b^n < \infty$ *and* $\sum_{n=1}^{\infty} \delta_n < \infty$, *then* (a^n) *is bounded and* $\lim_{n\to\infty} a^n$ *exists.*

In particular, if (a^n) *has a subsequence which converges strongly to zero, then* $\lim_{n\to\infty} a^n = 0$.

Lemma 3.1.15 (Lemma 5 in [17]) *Let* (a^n) *and* (θ_n) *be sequences of nonnegative real numbers such that*

$$a^{n+1} \leq (1+\theta_n)a^n + \theta_n a^{n-1}, \ n \in \mathbb{N}.$$

Then the following holds $a^{n+1} \leq K \cdot \prod_{j=1}^{n}(1+2\theta_j)$, *where* $K = \max\{a^1, a^2\}$. *Moreover, if* $\sum_{n=1}^{\infty} \theta_n < +\infty$, *then* (a^n) *is bounded.*

3.2 Hilbert space

Definition 3.2.1 (Metric space)[1] Let X be a nonempty set and $d : X \times X \to [0, \infty)$ be a function. Then d is called a *metric* on X if the following properties hold:

(M1) $d(x, y) \geq 0$ for all $x, y \in X$;

(M2) $d(x, y) = 0$ if and only if $x = y$ for all $x, y \in X$;

(M3) $d(x, y) = d(y, x)$ for all $x, y \in X$;

(M4) $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$.

The value of metric d at (x, y), we write $d(x, y)$, is called *distance* between x and

y, and the ordered pair (X, d) is called a *metric space.*

Example 3.2.2 (Examples of metric space)

1. In real line $\mathbb{R}$, define
$$d(x, y) = |x - y|$$
for all $x, y \in \mathbb{R}$. Then $(\mathbb{R}, d)$ is a metric space and the metric d is called the usual metric for $\mathbb{R}$.

2. In euclidean plane $\mathbb{R}^2$, define
$$d(x, y) = \sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2},$$
where $x = (x^1, x^2), y = (y^1, y^2) \in \mathbb{R}^2$. Then $(\mathbb{R}^2, d)$ is a metric space.

3. In euclidean space $\mathbb{R}^n$, define
$$d(x, y) = \sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2 + ... + (x^n - y^n)^2},$$
where $x = (x^1, x^2, ..., x^n), y = (y^1, y^2, ..., y^n) \in \mathbb{R}^n$. Then $(\mathbb{R}^n, d)$ is a metric space.

Definition 3.2.3 (Vector space)[1] A vector space or *linear space* X *over the field* F ($\mathbb{R}$ or $\mathbb{C}$) is a set X together with an internal binary operation $(+)$ called addition and a scalar multiplication carrying (α, x) in $F \times X$ to αx in X satisfying the following statements for all $x, y, z \in X$ and $\alpha, \beta \in F$:

(V1) $x + y = y + x$;

(V2) $(x + y) + z = x + (y + z)$;

(V3) there exists an element $0 \in X$ call the *zero vector* of X such that $x + 0 = x$ for all $x \in X$;

(V4) for every element $x \in X$, there exists an element $-x \in X$ called *the* *additive inverse* or *the negative* of x such that $x + (-x) = 0$;

(V5) $\alpha(x + y) = \alpha x + \alpha y$;

(V6) $(\alpha + \beta)x = \alpha x + \beta y$;

(V7) $(\alpha\beta)x = \alpha(\beta x)$;

(V8) $1 \cdot x = x$.

The scalar field is the real numbers the vector space is called a *real vector space*. If the scalar field is the complex numbers, then the vector space is called a *complex vector space*.

Definition 3.2.4 (Normed space)[1] Let X be a norm linear space over field F ($\mathbb{R}$ or $\mathbb{C}$) and $\|\cdot\| : X \to \mathbb{R}^+$ be a function. Then $\|\cdot\|$ is said to be a norm if the following properties hold:

(N1) $\|x\| \geq 0$, and $\|x\| = 0 \Leftrightarrow x = 0$;

(N2) $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in X$ and $\alpha \in F$;

(N3) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

The ordered pair $(X, \|\cdot\|)$ is called a *normed space*.

Definition 3.2.5 (Inner product space)[1] Let X be a linear space over the field F. An inner product on X is a function $\langle\cdot,\cdot\rangle : X \times X \to F$ with the following properties:

(IP1) $\langle x, x\rangle \geq 0$;

(IP2) $\langle x, x\rangle = 0 \Leftrightarrow x = 0$;

(IP3) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$;

(IP4) $\langle x, y\rangle = \overline{\langle y, x\rangle}$;

(IP5) $\langle x + y, z\rangle = \langle x, z\rangle + \langle z, y\rangle$.

The ordered pair $(X, \langle\cdot,\cdot\rangle)$ is called an *inner product space*.

Example 3.2.6 (Examples of inner product space)

1. The function $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$\langle x, y \rangle = \sum_{i=1}^{n} x^i y^i \text{ for all } x = (x^1, x^2, ..., x^n),\ y = (y^1, y^2, ..., y^n) \in \mathbb{R}^n$$

is an inner product on $\mathbb{R}^n$. In this case $\mathbb{R}^n$ with this inner product is called real Euclidean n-space.

2. Let $\mathbb{C}^n$ be the set of n-tuples of complex numbers. Then the function $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$\langle x, y \rangle = \sum_{i=1}^{n} x^i \overline{y^i} \text{ for all } x = (x^1, x^2, ..., x^n),\ y = (y^1, y^2, ..., y^n) \in \mathbb{C}^n$$

is an inner product on $\mathbb{C}^n$ In this case $\mathbb{C}^n$ with this inner product is called complex Euclidean n-space.

3. Let ℓ_2 be the set of all sequences of complex numbers $(a^1, a^2, \ldots, a^i, \ldots)$ with $\sum_{i=1}^{\infty} |a^i|^2 < \infty$. Then the function $\langle \cdot, \cdot \rangle : \ell_2 \times \ell_2 \to \mathbb{C}$ defined by

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x^i \overline{y^i} \text{ for all } x = \{x^i\}_{i=1}^{\infty}, y = \{y^i\}_{i=1}^{\infty} \in \ell_2$$

is an inner product on ℓ_2.

Proposition 3.2.7 (The Cauchy-Schwarz inequality))[1] Let X be an inner product space. Then the following holds:

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle \text{ for all } x, y \in X$$

i.e.,

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

for all $x, y \in X$.

Proposition 3.2.8 [Proposition 1.2.27 in [1]] Let X be an inner product space. Then the function $\|\cdot\|$:

$$\|x\| = \sqrt{\langle x, x\rangle}, \text{ for all } x \in X$$

is norm on X.

Proposition 3.2.9 (The parallelogram law)[1] Let X be an inner product space. Then

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

for all $x, y \in X$.

Definition 3.2.10 (Bounded sequence)[1] A sequence (x^n) in a normed space X is bounded if there exists $M > 0$ such that $\|x^n\| \leq M$ for all $n \in \mathbb{N}$.

Definition 3.2.11 (Convergent sequence)[1] A sequence (x^n) in a normed space X is said to be *convergent* if there is an $x \in X$ such that

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

In this case, we write $x^n \to x$ or $\lim\limits_{n\to\infty} x^n = x$. If (x^n) is not convergent, it is said to be *divergent*.

Definition 3.2.12 (Cauchy sequence)[1] A sequence (x^n) in a normed space X is said to be *Cauchy sequence* if $\lim\limits_{n,m\to\infty} \|x^m - x^n\| = 0$, *i.e.*, for $\varepsilon > 0$, there exists an integer $n^0 \in \mathbb{N}$ such that $\|x^m - x^n\| < \varepsilon$ for all $m, n \geq n^0$.

Definition 3.2.13 (Complete space)[1] The space X is said to be *complete space* if every Cauchy sequence in X converges.

Definition 3.2.14 (Hilbert space)[1] An inner product space which is complete with respect to the induced norm is called a *Hilbert space*.

Example 3.2.15 (Examples of Hilbert space)

1. The space $\mathbb{R}^n$ is a Hilbert space with inner product defined by

$$\langle x, y\rangle = x^1y^1 + x^2y^2 + ... + x^ny^n,$$

where $x = (x^1, x^2, ..., x^n),\ y = (y^1, y^2, ..., y^n) \in \mathbb{R}^n$.

2. The space $\mathbb{C}^n$ is a Hilbert space with inner product defined by

$$\langle x, y\rangle = x^1\bar{y}^1 + x^2\bar{y}^2 + ... + x^n\bar{y}^n,$$

where $x = (x^1, x^2, ..., x^n),\ y = (y^1, y^2, ..., y^n) \in \mathbb{C}^n$.

3. The space ℓ_2 of real sequence $x = (x^1, x^2, ...)$ which are squaresummable, *i.e.*, $\ell_2 = \{(x^n)_{n=1}^{\infty} | \sum_{n=1}^{\infty} |x^n|^2 < +\infty\}$ with the inner product

$$\langle x, y\rangle = \sum_{n=1}^{\infty} x^n\bar{y}^n$$

and with the norm

$$\|x\| = \left(\sum_{n=1}^{\infty} (x^n)^2\right)^{\frac{1}{2}}.$$

4. $L_2([a, b])$ where $a, b \in [-\infty, +\infty],\ a < b$. The space of (equivalence classes of) Lebesgue functions $f : \mathbb{R} \to \mathbb{R}$ which are square integrable, *i.e.*, $L_2(a, b) = \{f : \int_a^b f^2(x)dx < \infty\}$ with the inner product

$$\langle x, y\rangle = \int_a^b f(x)g(x)dx$$

and with the norm

$$\|f\| = \left(\int_a^b f(x)dx\right)^{\frac{1}{2}}.$$

Proposition 3.2.16 *Let H be a Hilbert space. Then*

1. $\|\alpha x+(1-\alpha y)\|^2 = \alpha\|x\|^2+(1-\alpha)\|y\|^2-\alpha(1-\alpha)\|x-y\|^2$ *for all* $x, y \in H$ *and* $\alpha \in \mathbb{R}$.

2. $\langle x-y, z-w\rangle = \|x-w\|^2+\|y-z\|^2-\|x-z\|^2-\|y-w\|^2$ *for all* $x, y, z, w \in H$.

3. $\|x+y\|^2 \leq \|x\|^2+2\langle y, x+y\rangle$ *for all* $x, y \in H$.

4. $\|y-z\| \leq \|x-z\|$ *if and only if* $\|y-x\|^2+2\langle y-x, x-z\rangle \leq 0$ *for all* $x, y, z \in H$.

Definition 3.2.17 (Strongly convergent)[1] A sequence (x^n) in a Hilbert space H is said to be *strongly convergent* to x if $\lim\limits_{n\to\infty}\|x^n-x\| = 0$ and is denoted by $x^n \to x$.

Definition 3.2.18 (Weakly convergent)[1] A sequence (x^n) in a Hilbert space H is said to be *weakly convergent* to x if $\lim\limits_{n\to\infty}|\langle x^n-x, y\rangle| = 0$ for all $y \in H$ and is denoted by $x^n \rightharpoonup x$.

If a subsequence $(x^{n_i}) \subseteq (x^n)$ converges weakly to x, then x is called a *weak cluster point* of the sequence (x^n). The set of all weak limit point of (x^n) denoted by

$$w_w(x^n) = \{x|\exists(x^{n_i}) \subset (x^n) \text{ such that} x^{n_i} \rightharpoonup x\}$$

Proposition 3.2.19 (Lemma 3.2.1 in [9]) *If* $x^n \rightharpoonup x \in H$, *then* $\|x\| \leq \liminf\limits_n \|x^n\|$.

Theorem 3.2.20 *(Opial's theorem)*[29] *If* $x^n \rightharpoonup x \in H$, *then the following inequality holds*

$$\liminf_{n\to\infty}\|x^n-x\| < \liminf_{n\to\infty}\|x^n-y\|$$

for all $y' \in H$ *and* $x \neq y$.

Theorem 3.2.21 (Lemma 3.2.4 in [9]) *If* $x^n \rightharpoonup x \in H$, *then the following inequality holds*

$$\limsup_{n\to\infty}\|x^n-x\| < \limsup_{n\to\infty}\|x^n-y'\|$$

for all $y \in H$ *and* $x \neq y$.

Definition 3.2.22 [9] Let $x^0 \in H$ and $r > 0$ be radius. Then,

· *Open ball* with a center x^0 is defined by the subset

$$B(x^0; r) = \{x \in H | \|x - x^0\| < r\}.$$

· *Closed ball* with a center x^0 is defined by the subset

$$B[x^0; r] = \{x \in H | \|x - x^0\| \leq r\}.$$

· *Sphere* with a center x^0 is defined by the subset

$$S(x^0; r) = \{x \in H | \|x - x^0\| = r\}.$$

C' denotes the complement of a subset $C \subset H$, *i.e.*, $C' = H \setminus C$.

Definition 3.2.23 (Minkowski sum)[9] Let H be a Hilbert space and $C, D \subseteq H$, we say that the *Minkowski sum* of C, D if

$$C + D = \{z \in H | z = x + y,\ x \in C \text{ and } y \in D\}.$$

Definition 3.2.24 (Segment)[9] Let H be a Hilbert space, we say that a *segment* with endpoints $x, y \in H$ if

$$[x, y] = \{z \in H | z = (1 - \alpha)x + \alpha y, \alpha \in [0, 1]\}.$$

Definition 3.2.25 (Subset of minimizers)[9] Let H be a Hilbert space and

$C \subseteq H$, we say that a *subset of minimizers* of f if

$$\operatorname*{Argmin}_{x\in C} f(x) = \{z \in H | f(z) \leq f(x) \text{ for all } x \in C\}$$

where $f : C \to \mathbb{R}$. An element of $\operatorname*{Argmin}_{x\in C} f(x)$ is called *minimizer* of f and is denoted by $\operatorname*{argmin}_{x\in C} f(x)$.

Definition 3.2.26 (Hyperplane)[9] Let H be a Hilbert space. Then *hyperplane* in H is defined by

$$H[x, \beta] = \{y \in H | \langle x, y\rangle = \beta\}$$

where $x \in H$, $x \neq 0$ and $\beta \in \mathbb{R}$. The hyperplane $H[x, \beta]$ is the boundary of two *half-spaces*

$$H(x, \beta)_+ = \{y \in H | \langle x, y\rangle \geq \beta\}$$

and

$$H(x, \beta)_- = \{y \in H | \langle x, y\rangle \leq \beta\}.$$

Definition 3.2.27 (Sublevel set)[9] Let H be a Hilbert space, $C \subseteq H$ and $f : C \to \mathbb{R}$. Then

$$S(f, \alpha) = \{x \in C | f(x) \leq \alpha\}$$

is called a *sublevel set* of f at a level $\alpha \in \mathbb{R}$.

Definition 3.2.28 [9]Let H be a Hilbert space, $C \subseteq H$ and $f : C \to \mathbb{R} \cup \{+\infty\}$. Then,

- *Domain* of f is

$$\text{dom} f = \{x \in C | f(x) < +\infty\}.$$

- *Epigraph* of f is

$$\text{epi} f = \{(x, \alpha) \in C \times \mathbb{R} | \alpha \geq f(x)\}.$$

· *Graph* of f is

$$\mathrm{gph}(f) = \{(x, \alpha) \in C \times \mathbb{R} | \alpha = f(x)\}.$$

The function f is *proper* if $-\infty \notin f(C)$ and $\mathrm{dom} f \neq \emptyset$.

Definition 3.2.29 (Semi-continuous function) Let H be a real Hilbert space. let $f : H \to (-\infty, +\infty]$ and let $x \in H$. Then

· f is *lower semi-continuous* at x if, for every $(x^n) \subseteq H$,

$$x^n \to x \quad \Rightarrow \quad f(x) \leq \liminf_{n\to\infty} f(x).$$

· f is *upper semi-continuous* at x if, for every $(x^n) \subseteq H$,

$$x^n \to x \quad \Rightarrow \quad f(x) \geq \limsup_{n\to\infty} f(x).$$

Definition 3.2.30 [9] Let H be a Hilbert space and $C \subseteq H$. Then,

· $f : C \to \mathbb{R}$ is *weakly lower semi-continuous* at $x \in C$ if

$$x^n \rightharpoonup x \quad \Rightarrow \quad f(x) \leq \liminf_{n\to\infty} f(x).$$

· $f : H \to \mathbb{R} \cup \{+\infty\}$ is *coercive* if $\lim_{\|x\|\to\infty} f(x) = +\infty$.

· $f : C \to \mathbb{R}$ attains the *global minimum* at a point $x_* \in C$ if $f(x_*) \leq f(x)$ for all $x \in C$.

· $f : C \to \mathbb{R}$ attains a *local minimum* at a point $x_* \in C$ if there exists $r > 0$ such that $f(x_*) \leq f(x)$ for all $x \in C \cap B(x_*, r)$.

Definition 3.2.31 (directional derivative)[9] Let H be a Hilbert space and

$f : H \to \mathbb{R}$, then *directional derivative* of f at x in the direction d is

$$f'(x, d) = \lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t}$$

which always exists.

Definition 3.2.32 (Fréchet-differentiable)[9] Let H be a Hilbert space and $f : H \to \mathbb{R}$, then f is *Fréchet-differentiable* or differentiable at $x \in H$ if there exists $y \in H$ such that

$$f(x + t) = f(x) + \langle y, t \rangle + o(\|t\|),$$

where $\lim_{t \to 0} \frac{o(t)}{t} = 0$. The element y is called a derivative or *differential* of f at x and denoted by $f'(x)$.

Definition 3.2.33 (Gâteaux-differentiable)[9] Let H be a Hilbert space and $f : H \to \mathbb{R}$, then f is *Gâteaux-differentiable* at a point $x \in H$ if it has directional derivative $f'(x, d)$ for all $d \in H$

$$f'(x, d) = \langle g, d \rangle$$

holds for some $g \in H$. The element g is called a Gâteaux-derivative or *Gâteaux-differential* of f at x and denoted by $Df(x)$.

Lemma 3.2.34 (Theorem 1.1.13 in [9]) *If a function $f : H \to \mathbb{R}$ is Fréchet-differentiable at $x \in H$, then it is Gâteaux-differentiable at x and $f'(x) = Df(x)$.*

Definition 3.2.35 [9] Let H, H_1, H_2 be Hilbert spaces. The operator $I : H \to H$ denotes the *identity*, *i.e.*, $Ix = x$ for all $x \in H$. Let $A : H_1 \to H_2$ be a bounded linear operator. Then,

- $A^* : H_2 \to H_1$ is an *adjoint operator* if $\langle Ax, y \rangle = \langle x, A^*y \rangle$ for all $x \in H_1$ and $y \in H_2$.

- $A^* : H_2 \to H_1$ is *unitary* if $A^*A = I$ holds.

- $A^* : H_2 \to H_1$ is *self-adjoint* if $A^* = A$.

Lemma 3.2.36 (Theorem 1.1.21 in [9]) *For any bounded linear operator $A : H_1 \to H_2$ it holds*

$$\|A\|^2 = \|A^*\|^2 = \|A^*A\| = \|AA^*\|.$$

Definition 3.2.37 (Fixed point)[9] A point $x \in C$ is called a *fixed point* of an operator $T : C \to C$, if $Tx = x$. The subset of all fixed point of the operator T is denoted by Fix$T$.

Definition 3.2.38 (Convex subset)[9] Let H be a Hilbert space and $C \subseteq H$. Then a subset C is said to be *convex* if

$$(1-\alpha)x + \alpha y \in C$$

for all $x, y \in C$ and $\alpha \in [0, 1]$.

Definition 3.2.39 (Convex functions) [9] Let H be a Hilbert space and $C \subseteq H$ be a convex subset. Let $f : C \to \mathbb{R}$. Then f is said to be

- *Convex* if

  $$f((1-\alpha)x + \alpha y) \leq (1-\alpha)f(x) + \alpha f(y)$$

  for all $x, y \in C$ and $\alpha \in [0, 1]$.

- *Strictly convex* if

  $$f((1-\alpha)x + \alpha y) < (1-\alpha)f(x) + \alpha f(y)$$

  for all $x, y \in C$, $x \neq y$ and $\alpha \in (0, 1)$.

· *α-strongly convex* or *strongly convex* if

$$f((1-\alpha)x+\alpha y) \leq \alpha f(y)+(1-\alpha)f(x)-\frac{1}{2}\lambda\alpha(1-\alpha)\|x-y\|^2$$

for all $x, y \in C$, $\alpha \in [0,1]$ and $\lambda > 0$.

· *Concave* if $-f$ is convex.

Properties of convex functions [9]

1. A convex function $f : \mathbb{R}^n \to \mathbb{R}$ is continuous.

2. A strongly convex function $f : H \to \mathbb{R}$ is coercive.

3. A convex function $f : H \to \mathbb{R}$ has directional derivatives $f'(\cdot, d)$ for all $d \in H$ and for any $x, d \in H$ it holds

$$f'(x,d) = \lim_{t\to 0^+} \frac{f(x+td)-f(x)}{t}.$$

4. If $f : H \to \mathbb{R}$ is convex function and $f'(x,d) \geq 0$ foe all $d \in H$, then f attains its minimum at $x \in H$.

5. Let $C \subseteq H$ be nonempty. A function $d(\cdot, C) : H \to \mathbb{R}$ defined by $d(x,C) = \inf_{y\in C}\|y-x\|$ is called a *distance function* to the subset C. If C is convex, then $d(\cdot, C)$ is a convex function.

Lemma 3.2.40 (Theorem 1.1.51 in [9]) *A lower semi-continuous convex function $f : H \to \mathbb{R}$ is weakly lower semi-continuous.*

Lemma 3.2.41 (Theorem 1.1.52 in [9]) *A continuous convex function $f : H \to \mathbb{R}$ defined on a closed bounded convex subset $C \subseteq H$ attains its global minimum.*

Definition 3.2.42 (Subdifferential)[9] Let $f : H \to \mathbb{R} \cup \{+\infty\}$ be convex. The subset

$$\partial f(x) = \{g \in H | \langle g, y - x\rangle \leq f(y) - f(x) \text{ for all } y \in H\}$$

is called a *subdifferential* of f at $x \in H$. The function f is said to be *subdifferentiable* at x if $\partial f(x) \neq \emptyset$. An element of the subdifferential $\partial f(x)$ is called a *subgradient* of f at x.

Example 3.2.43 1. The subdifferential of the absolute $|\cdot|$ has the form

$$\partial(|x|) = \begin{cases} -1 & \text{if } x > 0, \\ [-1, 1] & \text{if } x = 0, \\ 1 & \text{if } x < 0. \end{cases}$$

2. The subdifferential of the norm $\|\cdot\|$ has the form

$$\partial(\|x\|) = \begin{cases} (\frac{x}{\|x\|}) & \text{if } x \neq 0, \\ B(0, 1) & \text{if } x = 0. \end{cases}$$

Lemma 3.2.44 (Theorem 1.156 in [9]) *For any $x \in H$ the subdifferential $\partial f(x)$ of a continuous convex function $f : H \to \mathbb{R}$ is a nonempty, weakly closed and bounded convex set.*

Definition 3.2.45 (Metric projection)[9] Let $C \subseteq H$ be a nonempty subset and $x \in H$. If there exists a point $y \in C$ such that

$$\|y - x\| \leq \|z - x\|$$

for any $z \in C$, then y is called a *metric projection* of x on to C and denoted by $P_C x$.

If $P_C x$ exists and is uniquely determined for all $x \in H$, then the operator $P_C : H \to C$ is called the *metric projection* on to C. A projection of $x \in H$ on to a nonempty closed and convex subset C of H is defined by

$$P_C x = \operatorname*{argmin}_{y \in C} \|x - y\|^2.$$

Lemma 3.2.46 (Remark 1.2.2 in [9]) *Let $C \subseteq H$ be a nonempty closed convex subset. Then for any $x \in H$ there exists a metric projection $P_C x$ and it is uniquely determined.*

Lemma 3.2.47 ***(Characterization)****[9] Let $x \in H$, $C \subseteq H$ be a convex subset and $y \in C$. The following conditions are equivalent:*

1. $y = P_C x$,
2. $\langle x - y, z - y \rangle \leq 0$ *for all* $z \in C$.

Lemma 3.2.48 (Lemma 1.2.5 in [9]) *Let $x, y, z \in H$. The following conditions are equivalent:*

1. $\langle x - y, z - y \rangle \leq 0$
2. $\langle z - x, y - x \rangle \geq \|y - x\|^2$,
3. $\|z - y\|^2 \leq \|z - x\|^2 - \|y - x\|^2$,
4. $\langle z - x, z - y \rangle \geq 0$.

Definition 3.2.49 (Lipschitz continuous)[9] The operator $T : H \to H$ is Lipschitz continuous, if

$$\|Tx - Ty\| \leq L\|x - y\|$$

for a positive constant L. The constant L is called a *Lipschitz constant.* We also say that T is *L-Lipschitz continuous.*

Definition 3.2.50 (Nonexpansive operator)[9] We say that an operator $T : C \to H$ is:

· *Nonexpansive*, if

$$\|Tx - Ty\| \leq \|x - y\|$$

for all $x, y \in C$.

· *Strictly nonexpansive*, if

$$\|Tx - Ty\| < \alpha\|x - y\|$$

for all $x, y \in C$.

· An $\alpha-$*contraction* or *contraction*, where $\alpha \in (0, 1)$ if

$$\|Tx - Ty\| \leq \alpha\|x - y\|$$

for all $x, y \in C$ and $\alpha \in [0, 1]$.

Theorem 3.2.51 ***(Banach,1922)****[9] Let C be a complete metric space and $T : C \to C$ be contraction. Then T has exactly one fixed point $x_* \in C$. Furthermore, for any $x \in C$, the orbit $(T^n x)_{n=1}^{\infty}$ converges to x^* with a rate of geometric progression.*

Theorem 3.2.52 ***(Brouwer,1912)****[9] Let $C \subseteq \mathbb{R}^n$ be a nonempty compact and convex and $T : C \to C$ be continuous. Then T has a fixed point.*

Theorem 3.2.53 ***(Schauder,1930)****[9] Let C be a onempty compact and convex subset of a Banach space and $T : C \to C$ be continuous. Then T has a fixed point.*

Theorem 3.2.54 ***(Browder-Göhde-Kirk,1965)****[9] Let C be a nonempty closed convex and bounded subset of a uniformly convex Banach space (e.g., of a Hilbert space H) and $T : C \to C$ be nonexpansive. Then T has a fixed point.*

Proposition 3.2.55 (Proposition 2.1.11 in [9]) *Let C be nonempty closed convex subset of H and $T : C \to H$ be nonexpansive operator. Then* $\mathrm{Fix}T$ *is closed and convex.*

Lemma 3.2.56 ***(Demiclosedness principle)****[2] Let $T : C \to H$ be nonexpansive and $y \in C$ be a weak cluster point of a sequence $(x^n)_{n=0}^{\infty}$. If $\|Tx^n - x^n\| \to 0$, then $y \in \mathrm{Fix}T$.*

Definition 3.2.57 (Firmly nonexpansive operator)[9] The operator $T : C \to H$ is said to be *firmly nonexpansive*, if

$$\langle Tx - Ty, x - y\rangle \geq \|Tx - Ty\|^2$$

or

$$\|x - y\|^2 \geq \|Tx - Ty\|^2 + \|(\mathrm{I} - T)x - (\mathrm{I} - T)y\|^2$$

for all $x, y \in C$.

Lemma 3.2.58 (Lemma 2.2.2 in [9]) *Let $T : C \to H$ and $x, y \in C$. The following inequalities are equivalent:*

1. $\langle Tx - Ty, x - y\rangle \geq \|Tx - Ty\|^2$,
2. $\langle Tx - Ty, (x - Tx) - (y - Ty)\rangle \geq 0$,
3. $\langle Ty - Tx, x - Tx\rangle + \langle Tx - Ty, y - Ty\rangle \leq 0$,
4. $\langle Ty - x, Tx - x\rangle + \langle Tx - y, Ty - y\rangle \geq \|Tx - x\|^2 + \|Ty - y\|^2$.

Lemma 3.2.59 (Theorem 2.2.4 in [9]) *A firmly nonexpansive operator $T : C \to H$ is monotone and nonexpansive.*

Definition 3.2.60 (Monotone sequence)[9] Let $C \subseteq H$ be nonempty. We say that a sequence $(x^n)_{n=1}^{\infty} \subseteq H$ is *Fejér monotone*, if

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

for all $x \in C$ and $n \geq 0$.

Lemma 3.2.61 (Proposition 5.4 in [2]) *Let H be a real Hilbert space. Let C be a nonempty, closed and convex subset of H and let (x^n) be a sequence in H that satisfies:*

1. $\lim_{n\to\infty} \|x^n - x\|$ *exists for each $x \in C$,*
2. $w_w(x^n) \subset C$.

Then (x^n) weakly converges to an element of C.

Theorem 3.2.62 (Theorem 5.5 in [2]) *Let (x^n) be a sequence in H and $C \subseteq H$ be a nonempty subset of H. Suppose that (x^n) is Fejér monotone with respect to C and that every weak sequential cluster point of (x^n) belong to C. Then (x^n) converges weakly to a point in C.*

3.3 Proximal operator

Definition 3.3.1 (Proximal operator)[31] Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a closed proper convex function. The *proximal operator* of f is defined by

$$\text{prox}_f(y) = \underset{x}{\text{argmin}}(f(x) + (1/2)\|x - y\|_2^2)$$

and the proximal operator of the scalar function αf, where $\alpha > 0$, which can be expressed as

$$\text{prox}_{\alpha f}(y) = \underset{x}{\text{argmin}}(f(x) + (1/2\alpha)\|x - y\|_2^2)$$

then $\text{prox}_{\alpha f}$ is call the proximal operator of f with parameter α.

The proximal operator of f is characterized by

$$x_* = \text{prox}_f x \Leftrightarrow x - x_* \in \partial f(x_*)$$

which reduces to

$$x_* = \text{prox}_f x \Leftrightarrow x - x_* \in \nabla f(x_*)$$

if f is differentiable. The proximal operator $\text{prox}_g : H \to \text{dom}g$ with $\text{prox}_g(z) = (I + \partial g)^{-1}(z), z \in H$. It is well - known that the proximal operator is single - valued with full domain. It is also known that

$$\frac{z - \text{prox}_{\alpha g}(z)}{\alpha} \in \partial g(\text{prox}_{\alpha g}(z)) \text{ for all } z \in H,\ \alpha > 0. \qquad (3.3.1)$$

Lemma 3.3.2 ***(Fermat's rule)****[2] Let $f : H \to (-\infty, +\infty]$ be proper. Then*

$$\text{argmin} f = \{x \in H | 0 \in \partial f(x)\}.$$

Lemma 3.3.3 (Theorem 4.7.1 and Proposition 4.2.1(i) in [6]) *Let $h : H \to \mathbb{R} \cup \{+\infty\}$ be a closed proper lower-semicontinuous and convex function. The subdifferential operator ∂h is maximal monotone. Moreover, the graph of ∂h, $\text{gph}(\partial h) = \{(x, y) \in H \times H : y \in \partial h(x)\}$ is demiclosed, i.e., if the sequence $(x^n, y^n) \subset \text{gph}(\partial h)$ satisfies that $(x^n)_{n\in\mathbb{N}}$ converges weakly to x and $(y^n)_{n\in\mathbb{N}}$ converges strongly to y, then $(x, y) \in \text{gph}(\partial h)$.*

CHAPTER 4

MAIN RESULTS

This chapter is to present our results for convex optimization problems. We focus on the convergence analysis of the forward-backward algorithm for solving nonsmooth optimization problem in Hilbert spaces when the objective function is the sum of two convex functions. Also, we provide some applications in image restoration including the numerical experiments for supporting our main theorems.

4.1 Double inertial proximal gradient algorithms for convex optimization problems

In this section, we propose a double inertial forward-backward method for solving unconstrained convex minimization problem of the sum of two functions. We assume that $f : H \to \mathbb{R} \cup \{+\infty\}$ and $g : H \to \mathbb{R} \cup \{+\infty\}$ are proper, lower semi-continuous and convex functions that f is differentiable on H with the Lipschitz constant L of ∇f. The following is our algorithm.

Algorithm 4.1.1 Double inertial forward-backward algorithms (DIFB)
Initialization: Given α_n, θ_n and η_n be real positive sequences.
Iterative step: Let $x^0, x^1 \in H$ be arbitrarily and calculate x^{n+1} as follows:
Step1. Compute the inertial step:

$$w^n = x^n + \theta_n(x^n - x^{n-1}).$$

Step 2. Calculate:

$$v^n = \text{prox}_{\alpha_n g}(w^n - \alpha_n \nabla f(w^n))$$

and $$y^n = v^n + \eta_n(v^n - x^{n-1}).$$

Step 3. Calculate the next iteration via:

$$x^{n+1} = \operatorname{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)).$$

Set $n := n+1$ and return to **Step1.**

Theorem 4.1.2 *Let (x^n) be generated by Algorithm 4.1.1. Assume that* $0 < \liminf_{n\to\infty} \alpha_n \leq \limsup_{n\to\infty} \alpha_n < \frac{2}{L}$, $\sum_{n=1}^{\infty} \theta_n < +\infty$ *and* $\sum_{n=1}^{\infty} \eta_n < +\infty$. *Then we have*

1. *for each* $x_* \in \operatorname{argmin}(f+g)$, $\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1 + 2(\theta_j + \eta_j + \theta_j\eta_j))$, *where* $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$.
2. (x^n) *weakly converges to an element of* $\operatorname{argmin}(f+g)$.

Proof. Let $x_* \in \operatorname{argmin}(f+g)$. Since the proximal operator is firmly nonexpansive, it follows that

$$\begin{aligned}
& \|x^{n+1} - x_*\|^2 \\
= \ & \|\operatorname{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)) - \operatorname{prox}_{\alpha_n g}(x_* - \alpha_n \nabla f(x_*))\|^2 \\
\leq \ & \|(y^n - \alpha_n \nabla f(y^n)) - (x_* - \alpha_n \nabla f(x_*))\|^2 \\
& - \|(\mathrm{I} - \operatorname{prox}_{\alpha_n g})(y^n - \alpha_n \nabla f(y^n)) - (\mathrm{I} - \operatorname{prox}_{\alpha_n g})(x_* - \alpha_n \nabla f(x_*))\|^2 \\
\leq \ & \|y^n - x_*\|^2 - \frac{2\alpha_n}{L}\|\nabla f(y^n) - \nabla f(x_*)\|^2 + \alpha_n^2 \|\nabla f(y^n) - \nabla f(x_*)\|^2 \\
& - \|y^n - \operatorname{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)) - \alpha_n(\nabla f(y^n) - \nabla f(x_*))\|^2 \\
= \ & \|y^n - x_*\|^2 - \alpha_n(\frac{2}{L} - \alpha_n)\|\nabla f(y^n) - \nabla f(x_*)\|^2 \\
& - \|y^n - x^{n+1} - \alpha_n(\nabla f(y^n) - \nabla f(x_*))\|^2. \qquad (4.1.1)
\end{aligned}$$

Next, we consider

$$\begin{aligned}
\|y^n - x_*\|^2 &= \|v^n + \eta_n(v^n - x^{n-1}) - x_*\|^2 \\
&= \|v^n - x_*\|^2 + 2\eta_n\langle v^n - x_*, v^n - x^{n-1}\rangle + \eta_n^2\|v^n - x^{n-1}\|^2 \\
&\leq \|v^n - x_*\|^2 + 2\eta_n\|v^n - x_*\|\|v^n - x^{n-1}\| + \eta_n^2\|v^n - x^{n-1}\|^2.
\end{aligned} \tag{4.1.2}$$

Similarly, we see that

$$\begin{aligned}
&\|v^n - x_*\|^2 \\
=\ & \|\mathrm{prox}_{\alpha_n g}(w^n - \alpha_n\nabla f(w^n)) - \mathrm{prox}_{\alpha_n g}(x_* - \alpha_n\nabla f(x_*))\|^2 \\
\leq\ & \|(w^n - \alpha_n\nabla f(w^n)) - (x_* - \alpha_n\nabla f(x_*))\|^2 \\
& -\|(\mathrm{I} - \mathrm{prox}_{\alpha_n g})(w^n - \alpha_n\nabla f(w^n)) - (\mathrm{I} - \mathrm{prox}_{\alpha_n g})(x_* - \alpha_n\nabla f(x_*))\|^2 \\
\leq\ & \|w^n - x_*\|^2 - \frac{2\alpha_n}{L}\|\nabla f(w^n) - \nabla f(x_*)\|^2 + \alpha_n^2\|\nabla f(w^n) - \nabla f(x_*)\|^2 \\
& -\|w^n - \mathrm{prox}_{\alpha_n g}(w^n - \alpha_n\nabla f(w^n)) - \alpha_n(\nabla f(w^n) - \nabla f(x_*))\|^2 \\
=\ & \|w^n - x_*\|^2 - \alpha_n(\frac{2}{L} - \alpha_n)\|\nabla f(w^n) - \nabla f(x_*)\|^2 \\
& -\|w^n - v^n - \alpha_n(\nabla f(w^n) - \nabla f(x_*))\|^2.
\end{aligned} \tag{4.1.3}$$

By definition of w^n, we get

$$\begin{aligned}
\|w^n - x_*\|^2 &= \|x^n + \theta_n(x^n - x^{n-1}) - x_*\|^2 \\
&= \|x^n - x_*\|^2 + 2\theta_n\langle x^n - x_*, x^n - x^{n-1}\rangle + \theta_n^2\|x^n - x^{n-1}\|^2 \\
&\leq \|x^n - x_*\|^2 + 2\theta_n\|x^n - x_*\|\|x^n - x^{n-1}\| + \theta_n^2\|x^n - x^{n-1}\|^2.
\end{aligned} \tag{4.1.4}$$

From (4.1.1), we have

$$\|x^{n+1} - x_*\| \leq \|y^n - x_*\|$$

$$\begin{aligned} &= \|v^n + \eta_n(v^n - x^{n-1}) - x_*\| \\ &\leq \|v^n - x_*\| + \eta_n\|v^n - x^{n-1}\|. \end{aligned} \tag{4.1.5}$$

From (4.1.3) and (4.1.5), we get

$$\begin{aligned} \|x^{n+1} - x_*\| &\leq \|w^n - x_*\| + \eta_n\|v^n - x^{n-1}\| \\ &= \|x^n + \theta_n(x^n - x^{n-1}) - x_*\| + \eta_n\|v^n - x^{n-1}\| \\ &\leq \|x^n - x_*\| + \theta_n(\|x^n - x_*\| + \|x^{n-1} - x_*\|) \\ &\quad +\eta_n(\|x^n - x_*\| + \theta_n(\|x^n - x_*\| + \|x^{n-1} - x_*\|) + \|x^{n-1} - x_*\|) \\ &= (1 + \theta_n + \eta_n + \theta_n\eta_n)\|x^n - x_*\| + (\theta_n + \eta_n + \theta_n\eta_n)\|x^{n-1} - x_*\|. \end{aligned}$$

By Lemma 3.1.15, we have

$$\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1 + 2(\theta_j + \eta_j + \theta_j\eta_j))$$

where $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$. Since $\sum_{n=1}^{\infty}\theta_n < +\infty$ and $\sum_{n=1}^{\infty}\eta_n < +\infty$, the sequence (x^n) is bounded. By Lemma 3.1.14, it can be concluded that $\lim_{n\to\infty}\|x^n - x_*\|$ exists. From (4.1.1), (4.1.2) and (4.1.4), we obtain

$$\begin{aligned} &\|x^{n+1} - x_*\|^2 \\ \leq\ &\|x^n - x_*\|^2 + 2\theta_n\|x^n - x_*\|\|x^n - x^{n-1}\| + \theta_n^2\|x^n - x^{n-1}\|^2 \\ &-\alpha_n(\frac{2}{L} - \alpha_n)\|\nabla f(w^n) - \nabla f(x_*)\|^2 \\ &-\|w^n - v^n - \alpha_n(\nabla f(w^n) - \nabla f(x_*))\|^2 + 2\eta_n\|v^n - x_*\|\|v^n - x^{n-1}\| \\ &+\eta_n^2\|v^n - x^{n-1}\|^2 - \alpha_n(\frac{2}{L} - \alpha_n)\|\nabla f(y^n) - \nabla f(x_*)\|^2 \\ &-\|y^n - x^{n+1} - \alpha_n(\nabla f(y^n) - \nabla f(x_*))\|^2. \end{aligned}$$

Since $\lim_{n\to\infty}\|x^n - x_*\|$ exists, $\sum_{n=1}^{\infty}\theta_n < +\infty$ and $\sum_{n=1}^{\infty}\eta_n < +\infty$ and $0 < \liminf_{n\to\infty}\alpha_n \leq$

$\limsup_{n\to\infty} \alpha_n < \frac{2}{L}$, we have

$$\lim_{n\to\infty} \|\nabla f(w^n) - \nabla f(x_*)\| = 0 \text{ and } \lim_{n\to\infty} \|\nabla f(y^n) - \nabla f(x_*)\| = 0.$$

Moreover, we obtain

$$\lim_{n\to\infty} \|w^n - v^n - \alpha_n(\nabla f(w^n) - \nabla f(x_*))\| = 0$$

and

$$\lim_{n\to\infty} \|y^n - x^{n+1} - \alpha_n(\nabla f(y^n) - \nabla f(x_*))\| = 0.$$

We can check that $\lim_{n\to\infty} \|w^n - v^n\| = 0$ and $\lim_{n\to\infty} \|y^n - x^{n+1}\| = 0$. Since ∇f is uniform continuous, we obtain $\lim_{n\to\infty} \|\nabla f(w^n) - \nabla f(v^n)\| = 0$ and $\lim_{n\to\infty} \|\nabla f(y^n) - \nabla f(x^{n+1})\| = 0$. By definition of x^{n+1} and (3.3.1), we see that

$$\frac{y^n - \alpha_n \nabla f(y^n) - x^{n+1}}{\alpha_n} \in \partial g(x^{n+1}).$$

It follows that

$$\frac{y^n - x^{n+1}}{\alpha_n} + \nabla f(x^{n+1}) - \nabla f(y^n) \in \nabla f(x^{n+1}) + \partial g(x^{n+1}).$$

We see that

$$\begin{aligned} \|\frac{y^n - x^{n+1}}{\alpha_n} + \nabla f(x^{n+1}) - \nabla f(y^n)\| &\leq \frac{1}{\alpha_n}\|y^n - x^{n+1}\| + \|\nabla f(x^{n+1}) - \nabla f(y^n)\| \\ &\to 0 \text{ as } n \to \infty. \end{aligned}$$

Since (x^n) is bounded, the set of its weak limit points is nonempty. Take any weak limit point $\bar{x}$ of (x^n). By Lemma 3.3.3, we have $0 \in \nabla f(\bar{x}) + \partial g(\bar{x})$. Thus $\bar{x} \in \operatorname{argmin}(f + g)$. By Theorem 3.2.62, we can conclude that Theorem 4.1.2

holds. □

Next, we consider the projected version of Algorithm 4.1.1. Let Ω be a nonempty, closed and convex subset of H. Also, let $\operatorname{argmin}(f+g)\cap\Omega\neq\emptyset$.

Algorithm 4.1.3 Double inertial projected forward-backward algorithms (DIPFB)

Initialization: Given α_n, θ_n and η_n be real positive sequences.

Iterative step: Let $x^0, x^1 \in H$ be arbitrarily and calculate x^{n+1} as follows:

Step1. Compute the inertial step:

$$w^n = x^n + \theta_n(x^n - x^{n-1}).$$

Step 2. Calculate:

$$v^n = \operatorname{prox}_{\alpha_n g}(w^n - \alpha_n \nabla f(w^n))$$

and

$$y^n = v^n + \eta_n(v^n - x^{n-1}).$$

Step 3. Calculate:

$$z^n = \operatorname{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)).$$

Step 4. Calculate the next iteration via:

$$x^{n+1} = P_\Omega(z^n).$$

Set $n := n+1$ and return to **Step1.**

Theorem 4.1.4 *Let (x^n) be generated by Algorithm 4.1.3. Assume that* $0 <$

$\liminf_{n\to\infty} \alpha_n \leq \limsup_{n\to\infty} \alpha_n < \frac{2}{L}$, $\theta_n \geq 0$, $\eta_n \geq 0$, $\sum_{n=1}^{\infty} \theta_n < +\infty$ *and* $\sum_{n=1}^{\infty} \eta_n < +\infty$. *Then we have*

1. *for each* $x_* \in \Omega \cap \operatorname{argmin}(f+g)$, $\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n} (1 + 2(\theta_j + \eta_j + \theta_j \eta_j))$ *where* $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$.
2. (x^n) *weakly converges to an element of* $\Omega \cap \operatorname{argmin}(f+g)$.

Proof. Let $x_* \in \Omega \cap \operatorname{argmin}(f+g)$. Thus, we obtain

$$\begin{aligned} \|x^{n+1} - x_*\|^2 &= \|P_\Omega(z^n) - x_*\|^2 \\ &\leq \|z^n - x_*\|^2 - \|P_\Omega(z^n) - z^n\|^2. \end{aligned}$$

According to Theorem 4.1.2, we can show that

$$\begin{aligned} & \|x^{n+1} - x_*\|^2 \\ = \;& \|P_\Omega(z^n) - x_*\|^2 \\ \leq \;& \|z^n - x_*\|^2 - \|P_\Omega(z^n) - z^n\|^2 \\ \leq \;& \|y^n - x_*\|^2 - \alpha_n(\frac{2}{L} - \alpha_n)\|\nabla f(y^n) - \nabla f(x_*)\|^2 \\ & -\|y^n - z^n - \alpha_n(\nabla f(y^n) - \nabla f(x_*))\|^2 - \|P_\Omega(z^n) - z^n\|^2 \\ \leq \;& \|v^n - x_*\|^2 + 2\eta_n\|v^n - x_*\|\|v^n - x^{n-1}\| + \eta_n^2\|v^n - x^{n-1}\|^2 \\ & -\alpha_n(\frac{2}{L} - \alpha_n)\|\nabla f(y^n) - \nabla f(x_*)\|^2 \\ & -\|y^n - z^n - \alpha_n(\nabla f(y^n) - \nabla f(x_*))\|^2 - \|P_\Omega(z^n) - z^n\|^2 \\ \leq \;& \|w^n - x_*\|^2 - \alpha_n(\frac{2}{L} - \alpha_n)\|\nabla f(w^n) - \nabla f(x_*)\|^2 \\ & -\|w^n - z^n - \alpha_n(\nabla f(w^n) - \nabla f(x_*))\|^2 + 2\eta_n\|v^n - x_*\|\|v^n - x^{n-1}\| \\ & +\eta_n^2\|v^n - x^{n-1}\|^2 - \alpha_n(\frac{2}{L} - \alpha_n)\|\nabla f(y^n) - \nabla f(x_*)\|^2 \\ & -\|y^n - z^n - \alpha_n(\nabla f(y^n) - \nabla f(x_*))\|^2 - \|P_\Omega(z^n) - z^n\|^2 \\ \leq \;& \|x^n - x_*\|^2 + 2\theta_n\|x^n - x_*\|\|x^n - x^{n-1}\| + \theta_n^2\|x^n - x^{n-1}\|^2 \end{aligned}$$

$$-\alpha_n(\frac{2}{L}-\alpha_n)\|\nabla f(w^n)-\nabla f(x_*)\|^2$$
$$-\|w^n-z^n-\alpha_n(\nabla f(w^n)-\nabla f(x_*))\|^2+2\eta_n\|v^n-x_*\|\|v^n-x^{n-1}\|$$
$$+\eta_n^2\|v^n-x^{n-1}\|^2-\alpha_n(\frac{2}{L}-\alpha_n)\|\nabla f(y^n)-\nabla f(x_*)\|^2$$
$$-\|y^n-z^n-\alpha_n(\nabla f(y^n)-\nabla f(x_*))\|^2-\|P_\Omega(z^n)-z^n\|^2.$$

It can be shown that $\lim_{n\to\infty}\|x^n-x_*\|$ exists and hence

$$\lim_{n\to\infty}\|\nabla f(w^n)-\nabla f(x_*)\|=0 \quad\text{and}\quad \lim_{n\to\infty}\|\nabla f(y^n)-\nabla f(x_*)\|=0.$$

Moreover, we see that

$$\lim_{n\to\infty}\|w^n-z^n-\alpha_n(\nabla f(w^n)-\nabla f(x_*))\|=0$$

and

$$\lim_{n\to\infty}\|y^n-z^n-\alpha_n(\nabla f(y^n)-\nabla f(x_*))\|=0.$$

We can check that $\lim_{n\to\infty}\|w^n-v^n\|=0$ and $\lim_{n\to\infty}\|y^n-z^n\|=0$. Since ∇f is uniform continuous, we obtain $\lim_{n\to\infty}\|\nabla f(w^n)-\nabla f(v^n)\|=0$ and $\lim_{n\to\infty}\|\nabla f(y^n)-\nabla f(z^n)\|=0$. By definition of z^n and (3.3.1), we see that

$$\frac{y^n-\alpha_n\nabla f(y^n)-z^n}{\alpha_n}\in\partial g(z^n).$$

It follows that

$$\frac{y^n-z^n}{\alpha_n}+\nabla f(z^n)-\nabla f(y^n)\in\nabla f(z^n)+\partial g(z^n).$$

We see that

$$\begin{aligned}\left\|\frac{y^n - z^n}{\alpha_n} + \nabla f(z^n) - \nabla f(y^n)\right\| &\leq \frac{1}{\alpha_n}\|y^n - z^n\| + \|\nabla f(z^n) - \nabla f(y^n)\| \\ &\to 0 \text{ as } n \to \infty.\end{aligned}$$

Since (x^n) is bounded, the set of its weak limit points is nonempty. Take any weak limit point $\bar{x}$ of (x^n). So there is a subsequence (x^{n_i}) of (x^n) weakly converging to $\bar{x}$. By Lemma 3.3.3, we have $0 \in \nabla f(\bar{x}) + \partial g(\bar{x})$. Thus $\bar{x} \in \operatorname{argmin}(f+g)$. Moreover, since $\lim_{n\to\infty} \|P_\Omega(z^n) - z^n\| = 0$ and $z^{n_i} \rightharpoonup \bar{x}$ by Lemma 3.2.56 we have $\bar{x} \in \Omega$. Thus $\bar{x} \in \Omega \cap \operatorname{argmin}(f+g)$. By Theorem 3.2.62, we can conclude that Theorem 4.1.4 holds. $\square$

Remark 4.1.5 It is easy to see that the conditions $\sum_{n=1}^{\infty} \theta_n < +\infty$ and $\sum_{n=1}^{\infty} \eta_n < +\infty$ imply conditions

$$\sum_{n=1}^{\infty} \theta_n \|x^n - x^{n-1}\|^2 < +\infty \text{ and } \sum_{n=1}^{\infty} \eta_n \|x^n - x^{n-1}\|^2 < +\infty \tag{4.1.6}$$

when the sequence (x^n) in Algorithm 4.1.1 and Algorithm 4.1.3 is bounded. Moreover, our assumptions can be replaced by (4.1.6) with the same proof line as in Theorem 4.1.2 and 4.1.4.

4.2 Inertial projected forward-backward algorithm for constrained convex minimization problems

In this section, we propose the alternated inertial projection forward-backward algorithm and give the convergence analysis of the proposed algorithm under the assumption that the solution set is nonempty. Let Ω be a nonempty, closed and convex subset of a real Hilbert space H. We assume that $f : H \to \mathbb{R}\cup\{+\infty\}$ and $g : H \to \mathbb{R}\cup\{+\infty\}$ are proper, convex and lower semi-continuous, that f is differentiable on H and ∇f is uniformly continuous on each bounded subset of H. The following is our algorithm.

Algorithm 4.2.1 Alternated inertial projected forward-backward algorithm (AIPFB)

Initialization: Given $\sigma > 0, \delta \in (0, \frac{1}{2}),\ \theta \in (0,1)$ and $\theta_n \geq 0$.

Iterative step: Let $x^0, x^1 \in H$ arbitrarily and calculate x^{n+1} as follows:

Step1. Compute the inertial step:

$$z^n = \begin{cases} x^n + \theta_n(x^n - x^{n-1}), & \text{if } n \text{ is odd} \\ x^n, & \text{if } n \text{ is even.} \end{cases} \tag{4.2.1}$$

Step 2. Calculate:

$$x^{n+1} = P_\Omega(\text{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n)))$$

where $\alpha_n = \sigma\theta^{m_n}$ and m_n is the smallest nonnegative integer such that

$$\begin{aligned} \alpha_n \|\nabla f(\text{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n))) - \nabla f(z^n)\| \\ \leq \delta \|\text{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n)) - z^n\|. \end{aligned} \tag{4.2.2}$$

Set $n := n+1$ and return to **Step1.**

Theorem 4.2.2 *Let (x^n) be generated by Algorithm 4.2.1. Assume that $\theta_n \geq 0$ and $\sum_{n=1}^{\infty} \theta_n < +\infty$ and that $\alpha_n \geq \alpha > 0$ for some α. Then we have*

1. *for each $x_* \in \operatorname{argmin}(f+g)$, $\|x^{2n+2} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1 + 2\theta_{2j+1})$ where $K = \max\{\|x^1 - x_*\|, \|x^3 - x_*\|\}$.*

2. *(x^n) weakly converges to an element of $\operatorname{argmin}(f+g) \cap \Omega$.*

Proof. Let $x_* \in \operatorname{argmin}(f+g) \cap \Omega$ and set

$$y^{2n+1} = \operatorname{prox}_{\alpha_{2n+1}g}(z^{2n+1} - \alpha_{2n+1}\nabla f(z^{2n+1})).$$

Using (3.3.1), we obtain

$$\begin{aligned}
& \frac{z^{2n+1} - y^{2n+1}}{\alpha_{2n+1}} - \nabla f(z^{2n+1}) \\
= \ & \frac{z^{2n+1} - \operatorname{prox}_{\alpha_{2n+1}g}(z^{2n+1} - \alpha_{2n+1}\nabla f(z^{2n+1}))}{\alpha_{2n+1}} - \nabla f(z^{2n+1}) \\
\in \ & \partial g(y^{2n+1}).
\end{aligned}$$

By the convexity of g, we see that

$$g(x_*) - g(y^{2n+1}) \geq \left\langle \frac{z^{2n+1} - y^{2n+1}}{\alpha_{2n+1}} - \nabla f(z^{2n+1}), x_* - y^{2n+1} \right\rangle. \tag{4.2.3}$$

Similarly, by the convexity of f, we also have

$$f(x_*) - f(z^{2n+1}) \geq \langle \nabla f(z^{2n+1}), x_* - z^{2n+1} \rangle. \tag{4.2.4}$$

From (4.2.2), (4.2.3) and (4.2.4), it follows that

$$\begin{aligned}
& (f+g)(x_*) \\
\geq \ & g(y^{2n+1}) + f(z^{2n+1}) + \left\langle \frac{z^{2n+1} - y^{2n+1}}{\alpha_{2n+1}} - \nabla f(z^{2n+1}), x_* - y^{2n+1} \right\rangle \\
& + \langle \nabla f(z^{2n+1}), x_* - z^{2n+1} \rangle \\
= \ & g(y^{2n+1}) + f(z^{2n+1}) + \frac{1}{\alpha_{2n+1}} \langle z^{2n+1} - y^{2n+1}, x_* - y^{2n+1} \rangle \\
& + \langle \nabla f(y^{2n+1}), y^{2n+1} - z^{2n+1} \rangle + \langle \nabla f(z^{2n+1}) - \nabla f(y^{2n+1}), y^{2n+1} - z^{2n+1} \rangle \\
\geq \ & g(y^{2n+1}) + f(z^{2n+1}) + \frac{1}{\alpha_{2n+1}} \langle z^{2n+1} - y^{2n+1}, x_* - y^{2n+1} \rangle \\
& + f(y^{2n+1}) - f(z^{2n+1}) - \|\nabla f(z^{2n+1}) - \nabla f(y^{2n+1})\| \|y^{2n+1} - z^{2n+1}\| \\
\geq \ & (f+g)(y^{2n+1}) + \frac{1}{\alpha_{2n+1}} \langle z^{2n+1} - y^{2n+1}, x_* - y^{2n+1} \rangle - \frac{\delta}{\alpha_{2n+1}} \|y^{2n+1} - z^{2n+1}\|^2.
\end{aligned}$$

Since $\|z^{2n+1}-x_*\|^2-\|z^{2n+1}-y^{2n+1}\|^2-\|y^{2n+1}-x_*\|^2 = 2\langle z^{2n+1}-y^{2n+1}, y^{2n+1}-x_*\rangle$, we get

$$\begin{aligned}\|y^{2n+1}-x_*\|^2 &\leq \|z^{2n+1}-x_*\|^2 - 2\alpha_{2n+1}[(f+g)(y^{2n+1})-(f+g)(x_*)]\\ &\quad -(1-\delta)\|y^{2n+1}-z^{2n+1}\|^2. \end{aligned} \tag{4.2.5}$$

Using (4.2.5) and Lemma 3.2.48, we have

$$\begin{aligned}\|x^{2n+2}-x_*\|^2 &\leq \|y^{2n+1}-x_*\|^2 - \|P_\Omega(y^{2n+1})-y^{2n+1}\|^2\\ &\leq \|z^{2n+1}-x_*\|^2. \end{aligned} \tag{4.2.6}$$

Now, by setting $y^{2n} = \operatorname{prox}_{\alpha_{2n}g}(z^{2n}-\alpha_{2n}\nabla f(z^{2n}))$, we have

$$\begin{aligned}\frac{z^{2n}-y^{2n}}{\alpha_{2n}} - \nabla f(z^{2n}) &= \frac{z^{2n}-\operatorname{prox}_{\alpha_{2n}g}(z^{2n}-\alpha_{2n}\nabla f(z^{2n}))}{\alpha_{2n}} - \nabla f(z^{2n})\\ &\in \partial g(y^{2n}).\end{aligned}$$

Hence

$$g(x_*)-g(y^{2n}) \geq \left\langle \frac{z^{2n}-y^{2n}}{\alpha_{2n}} - \nabla f(z^{2n}), x_*-y^{2n}\right\rangle.$$

Similarly, we can show that

$$\begin{aligned}\|y^{2n}-x_*\|^2 &\leq \|z^{2n}-x_*\|^2 - 2\alpha_{2n}[(f+g)(y^{2n})-(f+g)(x_*)]\\ &\quad -(1-\delta)\|y^{2n}-z^{2n}\|^2. \end{aligned} \tag{4.2.7}$$

Next, we will show that (x^n) is bounded. Indeed, from (4.2.1) and (4.2.6), we obtain

$$\|x^{2n+2}-x_*\| \leq \|z^{2n+1}-x_*\|$$

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

This shows that

$$\|x^{2n+2} - x_*\| \leq (1 + \theta_{2n+1})\|x^{2n+1} - x_*\| + \theta_{2n+1}\|x^{2n} - x_*\|.$$

By Lemma 3.1.15, we have

$$\|x^{2n+2} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1 + 2\theta_{2j+1})$$

where $K = \max\{\|x^1 - x_*\|, \|x^3 - x_*\|\}$. Since $\sum_{n=1}^{\infty}\theta_{2n+1} < +\infty$, it follows that (x^{2n+1}) is bounded.

Next, we will show that (x^{2n}) is bounded. We see that

$$
\begin{aligned}
\|x^{2n+1} - x_*\|^2 &= \|P_\Omega(y^{2n}) - x_*\|^2 \\
&\leq \|y^{2n} - x_*\|^2 - \|P_\Omega(y^{2n}) - y^{2n}\|^2. \qquad (4.2.8)
\end{aligned}
$$

Then, by (4.2.7) and (4.2.8), we get

$$
\begin{aligned}
\|x^{2n+1} - x_*\| &\leq \|y^{2n} - x_*\| \\
&\leq \|z^{2n} - x_*\| \\
&= \|x^{2n} - x_*\|.
\end{aligned}
$$

So, we have $\lim_{n\to\infty}\|x^{2n} - x_*\|$ exists and (x^{2n}) is bounded. Hence (x^n) is bounded. Moreover, from (4.2.7), we obtain $\lim_{n\to\infty}\|x^{2n+1} - x_*\|$ exists by Lemma 3.1.14. We see that

$$
\begin{aligned}
& \|x^{2n+2} - x_*\|^2 \\
\leq\ & \|y^{2n+1} - x_*\|^2 - \|P_\Omega(y^{2n+1}) - y^{2n+1}\|^2 \\
\leq\ & \|z^{2n+1} - x_*\|^2 - 2\alpha_{2n+1}[(f+g)(y^{2n+1}) - (f+g)(x_*)] \\
& -(1-\delta)\|y^{2n+1} - z^{2n+1}\|^2 - \|P_\Omega(y^{2n+1}) - y^{2n+1}\|^2 \\
=\ & \|x^{2n+1} + \theta_{2n+1}(x^{2n+1} - x^{2n}) - x_*\|^2 - 2\alpha_{2n+1}[(f+g)(y^{2n+1}) - (f+g)(x_*)] \\
& -(1-\delta)\|y^{2n+1} - z^{2n+1}\|^2 - \|P_\Omega(y^{2n+1}) - y^{2n+1}\|^2 \\
=\ & \|x^{2n+1} - x_*\|^2 + \theta_{2n+1}^2\|x^{2n+1} - x^{2n}\| + 2\theta_{2n+1}\langle x^{2n+1} - x_*, x^{2n+1} - x^{2n}\rangle \\
& -2\alpha_{2n+1}[(f+g)(y^{2n+1}) - (f+g)(x_*)] \\
& -(1-\delta)\|y^{2n+1} - z^{2n+1}\|^2 - \|P_\Omega(y^{2n+1}) - y^{2n+1}\|^2 \\
\leq\ & \|x^{2n+1} - x_*\|^2 + \theta_{2n+1}^2\|x^{2n+1} - x^{2n}\| + 2\theta_{2n+1}\|x^{2n+1} - x_*\|\|x^{2n+1} - x^{2n}\| \\
& -2\alpha_{2n+1}[(f+g)(y^{2n+1}) - (f+g)(x_*)] \\
& -(1-\delta)\|y^{2n+1} - z^{2n+1}\|^2 - \|P_\Omega(y^{2n+1}) - y^{2n+1}\|^2. \qquad (4.2.9)
\end{aligned}
$$

On the other hand, from (4.2.7), we see that

$$
\begin{aligned}
& \|x^{2n+1} - x_*\|^2 \\
\leq\ & \|y^{2n} - x_*\|^2 - \|P_\Omega(y^{2n}) - y^{2n}\|^2 \\
\leq\ & \|z^{2n} - x_*\|^2 - 2\alpha_{2n}[(f+g)(y^{2n}) - (f+g)(x_*)] - (1-\delta)\|y^{2n} - z^{2n}\|^2 \\
& -\|P_\Omega(y^{2n}) - y^{2n}\|^2 \\
=\ & \|x^{2n} - x_*\|^2 - 2\alpha_{2n}[(f+g)(y^{2n}) - (f+g)(x_*)] - (1-\delta)\|y^{2n} - z^{2n}\|^2 \\
& -\|P_\Omega(y^{2n}) - y^{2n}\|^2. \qquad (4.2.10)
\end{aligned}
$$

Using (4.2.9) and (4.2.10), we obtain

$$\|x^{2n+2} - x_*\|^2 \leq \ldots$$

$$\|x^{2n} - x_*\|^2 - 2\alpha_{2n}[(f+g)(y^{2n}) - (f+g)(x_*)] - (1-\delta)\|y^{2n} - z^{2n}\|^2$$
$$-\|P_\Omega(y^{2n}) - y^{2n}\|^2 + \theta_{2n+1}^2\|x^{2n+1} - x^{2n}\|$$
$$+2\theta_{2n+1}\|x^{2n+1} - x_*\|\|x^{2n+1} - x^{2n}\| - 2\alpha_{2n+1}[(f+g)(y^{2n+1}) - (f+g)(x_*)]$$
$$-(1-\delta)\|y^{2n+1} - z^{2n+1}\|^2 - \|P_\Omega(y^{2n+1}) - y^{2n+1}\|^2.$$

Since $\lim_{n\to\infty}\|x^{2n} - x_*\|$ exists and $\sum_{n=1}^{\infty}\theta_n < +\infty$, we obtain

$$\lim_{n\to\infty}\|y^{2n+1} - z^{2n+1}\| = 0 \text{ and } \lim_{n\to\infty}\|y^{2n} - z^{2n}\| = 0. \tag{4.2.11}$$

Also, we have

$$\lim_{n\to\infty}\|P_\Omega(y^{2n+1}) - y^{2n+1}\| = 0 \text{ and } \lim_{n\to\infty}\|P_\Omega(y^{2n}) - y^{2n}\| = 0. \tag{4.2.12}$$

From (4.2.11) and (4.2.12), it implies that

$$\lim_{n\to\infty}\|y^n - z^n\| = \lim_{n\to\infty}\|\operatorname{prox}_{\alpha_n g}(z^n - \alpha_n\nabla f(z^n)) - z^n\| = 0$$

and

$$\begin{aligned}&\lim_{n\to\infty}\|P_\Omega(y^n) - y^n\|\\ =\ &\lim_{n\to\infty}\|P_\Omega(\operatorname{prox}_{\alpha_n g}(z^n - \alpha_n\nabla f(z^n))) - \operatorname{prox}_{\alpha_n g}(z^n - \alpha_n\nabla f(z^n))\| = 0.\end{aligned}$$

By definition of z^n, we get $\lim_{n\to\infty}\|z^n - x^n\| = 0$. Hence, we obtain

$$\begin{aligned}&\|\operatorname{prox}_{\alpha_n g}(z^n - \alpha_n\nabla f(z^n)) - x^n\|\\ \le\ &\|\operatorname{prox}_{\alpha_n g}(z^n - \alpha_n\nabla f(z^n)) - z^n\| + \|z^n - x^n\| \to 0 \text{ as } n\to\infty.\end{aligned}$$

Since (x^n) is bounded, the set of its weak limit points is nonempty. Take any weak limit point $\bar{x}$ of (x^n). So there is a subsequence (x^{n_i}) of (x^n) weakly converging to

$\bar{x}$. Moreover, (y^{n_i}) also weakly converges to $\bar{x}$. Since (z^{n_i}) is bounded, $\lim_{i\to\infty} \|z^{n_i} - \operatorname{prox}_{\alpha_{n_i} g}(z^{n_i} - \alpha_{n_i}\nabla f(z^{n_i}))\| = 0$ and ∇f is uniformly continuous, we obtain

$$\lim_{i\to\infty} \|\nabla f(z^{n_i}) - \nabla f(\operatorname{prox}_{\alpha_{n_i} g}(z^{n_i} - \alpha_{n_i}\nabla f(z^{n_i})))\| = 0$$

By definition of y^n and (3.3.1), we see that

$$\begin{aligned} &\frac{z^{n_i} - \alpha_{n_i}\nabla f(z^{n_i}) - \operatorname{prox}_{\alpha_{n_i} g}(z^{n_i} - \alpha_{n_i}\nabla f(z^{n_i}))}{\alpha_{n_i}} \\ \in\ & \partial g(\operatorname{prox}_{\alpha_{n_i} g}(z^{n_i} - \alpha_{n_i}\nabla f(z^{n_i}))). \end{aligned}$$

It follows that

$$\begin{aligned} &\frac{z^{n_i} - \operatorname{prox}_{\alpha_{n_i} g}(z^{n_i} - \alpha_{n_i}\nabla f(z^{n_i}))}{\alpha_{n_i}} + \nabla f(\operatorname{prox}_{\alpha_{n_i} g}(z^{n_i} - \alpha_{n_i}\nabla f(z^{n_i}))) - \nabla f(z^{n_i}) \\ \in\ & \nabla f(\operatorname{prox}_{\alpha_{n_i} g}(z^{n_i} - \alpha_{n_i}\nabla f(z^{n_i}))) + \partial g(\operatorname{prox}_{\alpha_{n_i} g}(z^{n_i} - \alpha_{n_i}\nabla f(z^{n_i}))) \\ \subseteq\ & \partial (f+g)(\operatorname{prox}_{\alpha_{n_i} g}(z^{n_i} - \alpha_{n_i}\nabla f(z^{n_i}))). \end{aligned}$$

By passing $i \to \infty$ and using Lemma 3.3.3, we have $0 \in \nabla f(\bar{x}) + \partial g(\bar{x})$. Thus $\bar{x} \in \operatorname{argmin}(f+g)$. Moreover, since $\lim_{n\to\infty} \|P_\Omega(y^n) - y^n\| = 0$ and $y^{n_i} \rightharpoonup \bar{x}$, by Lemma 3.2.56 we have $\bar{x} \in \Omega$. Thus $\bar{x} \in \Omega \cap \operatorname{argmin}(f+g)$. By Theorem 3.2.62, we can show that (x^n) converges weakly to a point in $\operatorname{argmin}(f+g) \cap \Omega$. □

4.3 Inertial proximal gradient method using adaptive stepsize for convex minimization problems

In this section, we propose a new inertial forward-backward algorithm and prove the weak convergence. Assume that $f : H \to \mathbb{R} \cup \{+\infty\}$ and $g : H \to \mathbb{R} \cup \{+\infty\}$ are proper, convex and lower semi-continuous that ∇f is L-Lipschitz continuous on H. The following is our algorithm.

Algorithm 4.3.1 Inertial forward-backward using adaptive stepsize (IF-BAS)

Initialization: Given $\delta \in (0,1)$, $\alpha_n \geq 0$ and $\theta_n \geq 0$.

Iterative step: Let $x^0, x^1 \in H$ be arbitrarily and calculate x^{n+1} as follows:

Step1. Compute the inertial step:

$$z^n = x^n + \theta_n(x^n - x^{n-1}) \tag{4.3.1}$$

Step 2. Calculate:

$$x^{n+1} = \text{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n)), \tag{4.3.2}$$

where

$$\alpha_{n+1} = \begin{cases} \min\{\frac{\delta\|z^n - x^{n+1}\|}{\|\nabla f(z^n) - \nabla f(x^{n+1})\|}, \alpha_n\} & \text{if } \nabla f(z^n) - \nabla f(x^{n+1}) \neq 0 \\ \alpha_n & \text{otherwise.} \end{cases} \tag{4.3.3}$$

Set $n := n+1$ and return to **Step1.**

Theorem 4.3.2 *Let (x^n) be defined by Algorithm 4.3.1. Assume that $\theta_n \geq 0$ and $\sum_{n=1}^{\infty} \theta_n < +\infty$. Then, we have*

1. *for each $x_* \in \operatorname{argmin}(f+g)$, $\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1+2\theta_j)$, where $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$.*

2. *The sequence (x^n) weakly converges to a point in $\operatorname{argmin}(f+g)$.*

Proof. Using (3.3.1) and (4.3.2), we see that

$$\frac{z^n - x^{n+1}}{\alpha_n} - \nabla f(z^n) = \frac{z^n - \text{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n))}{\alpha_n} - \nabla f(z^n) \in \partial g(x^{n+1}).$$

From the convexity of g, we have

$$g(x) - g(x^{n+1}) \geq \left\langle \frac{z^n - x^{n+1}}{\alpha_n} - \nabla f(z^n), x - x^{n+1} \right\rangle \tag{4.3.4}$$

for all $x \in H$. Also the convexity of f gives

$$f(x) - f(y) \geq \langle \nabla f(y), x - y \rangle \tag{4.3.5}$$

for all $x, y \in H$. Combining (4.3.4) and (4.3.5) with $y = z^n$, we obtain

$$\begin{aligned}
& g(x) - g(x^{n+1}) + f(x) - f(z^n) \\
\geq\ & \langle \frac{z^n - x^{n+1}}{\alpha_n} - \nabla f(z^n), x - x^{n+1} \rangle + \langle \nabla f(z^n), x - z^n \rangle \\
=\ & \frac{1}{\alpha_n} \langle z^n - x^{n+1}, x - x^{n+1} \rangle + \langle \nabla f(z^n), x^{n+1} - z^n \rangle \\
=\ & \frac{1}{\alpha_n} \langle z^n - x^{n+1}, x - x^{n+1} \rangle + \langle \nabla f(x^{n+1}), x^{n+1} - z^n \rangle \\
& + \langle \nabla f(z^n) - \nabla f(x^{n+1}), x^{n+1} - z^n \rangle \\
\geq\ & \frac{1}{\alpha_n} \langle z^n - x^{n+1}, x - x^{n+1} \rangle + \langle \nabla f(x^{n+1}), x^{n+1} - z^n \rangle \\
& - \|\nabla f(z^n) - \nabla f(x^{n+1})\| \|x^{n+1} - z^n\| \\
\geq\ & \frac{1}{\alpha_n} \langle z^n - x^{n+1}, x - x^{n+1} \rangle + \langle \nabla f(x^{n+1}), x^{n+1} - z^n \rangle \\
& - \frac{\delta}{\alpha_{n+1}} \|z^n - x^{n+1}\|^2.
\end{aligned}$$

By definition of (α_n), we have

$$\|\nabla f(z^n) - \nabla f(x^{n+1})\| \leq \frac{\delta}{\alpha_{n+1}} \|z^n - x^{n+1}\| \tag{4.3.6}$$

Indeed, if $\nabla f(z^n) = \nabla f(x^{n+1})$, then the inequality (4.3.6) hold. Otherwise, from (4.6.8), we have

$$\alpha_{n+1} = \min\{ \frac{\delta \|z^n - x^{n+1}\|}{\|\nabla f(z^n) - \nabla f(x^{n+1})\|}, \alpha_n \} \leq \frac{\delta \|z^n - x^{n+1}\|}{\|\nabla f(z^n) - \nabla f(x^{n+1})\|}.$$

This implies that

$$\|\nabla f(z^n) - \nabla f(x^{n+1})\| \le \frac{\delta}{\alpha_{n+1}}\|z^n - x^{n+1}\|.$$

Therefore, the inequality follows from (4.3.6). It then follows that

$$\begin{aligned}
&\langle z^n - x^{n+1}, x^{n+1} - x\rangle \\
\ge\ & \alpha_n[f(z^n) + g(x^{n+1}) - (f+g)(x) + \langle \nabla f(x^{n+1}), x^{n+1} - z^n\rangle] \\
&-\frac{\delta\alpha_n}{\alpha_{n+1}}\|z^n - x^{n+1}\|^2.
\end{aligned}$$

Using $2\langle z^n - x^{n+1}, x^{n+1} - x\rangle = \|z^n - x\|^2 - \|z^n - x^{n+1}\|^2 - \|x^{n+1} - x\|^2$, we get

$$\begin{aligned}
\|x^{n+1} - x\|^2 \ \le\ & \|z^n - x\|^2 - 2\alpha_n[(f+g)(x^{n+1}) - (f+g)(x)] \\
&-(1 - \frac{2\delta\alpha_n}{\alpha_{n+1}})\|z^n - x^{n+1}\|^2. \qquad (4.3.7)
\end{aligned}$$

Now, we let $x_* \in \operatorname{argmin}(f+g)$ and we will show that (x^n) is bounded. By (4.3.1) and (4.3.7), we see that

$$\begin{aligned}
\|x^{n+1} - x_*\| &\le \|z^n - x_*\| \\
&= \|x^n + \theta_n(x^n - x^{n-1}) - x_*\| \\
&\le \|x^n - x_*\| + \theta_n(\|x^n - x_*\| + \|x^{n-1} - x_*\|).
\end{aligned}$$

It implies that

$$\|x^{n+1} - x_*\| \le (1+\theta_n)\|x^n - x_*\| + \theta_n\|x^{n-1} - x_*\|.$$

By Lemma 3.1.15, we have

$$\|x^{n+1} - x_*\| \le K \cdot \prod_{j=1}^{n}(1 + 2\theta_j),$$

where $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$. Since $\sum_{n=1}^{\infty} \theta_n < +\infty$, we have (x^n) is bounded. From (4.3.1) and (4.3.7), we see that

$$\begin{aligned}
\|x^{n+1} - x_*\|^2 &\leq \|x^n + \theta_n(x^n - x^{n-1}) - x_*\|^2 - 2\alpha_n[(f+g)(x^{n+1}) - (f+g)(x_*)] \\
&\quad -(1 - \frac{2\delta\alpha_n}{\alpha_{n+1}})\|z^n - x^{n+1}\|^2 \\
&\leq \|x^n - x_*\|^2 + 2\theta_n\|x^n - x_*\|\|x^n - x^{n-1}\| + \theta_n^2\|x^n - x^{n-1}\|^2 \\
&\quad -2\alpha_n[(f+g)(x^{n+1}) - (f+g)(x_*)] - (1 - \frac{2\delta\alpha_n}{\alpha_{n+1}})\|z^n - x^{n+1}\|^2.
\end{aligned}$$

This implies that $\lim_{n\to\infty} \|x^n - x_*\|^2$ exists. Since $\lim_{n\to\infty}(1 - \frac{\delta\alpha_n}{\alpha_{n+1}}) = 1 - \delta > 0$, we have

$$\lim_{n\to\infty} \|z^n - x^{n+1}\| = 0.$$

From definition of z^n, it is easily seen that $\lim_{n\to\infty} \|z^n - x^n\| = 0$ and implies that $\lim_{n\to\infty} \|x^{n+1} - x^n\| = 0$. By the boundedness of (x^n), we know that the set of its weak accumulation points is nonempty. Let x^∞ be a weak accumulation point of (x^n). So there is a subsequence (x^{n_i}) of (x^n) such that (x^n) converges weakly to x^∞. Next, we show that $x^\infty \in \operatorname{argmin}(f+g)$. Let $(v,u) \in \operatorname{gph}(\nabla(f) + \partial(g))$, that is $u - \nabla f(v) \in \partial g(v)$. Since $x^{n_i+1} = \operatorname{prox}_{\alpha_{n_i}g}(I - \alpha_{n_i}\nabla f)z^{n_i}$, we obtain

$$(I - \alpha_{n_i}\nabla f)z^{n_i} \in (I + \alpha_{n_i}\partial g)x^{n_i+1},$$

which yields $\frac{1}{\alpha_{n_i}}(z^{n_i} - x^{n_i+1} - \alpha_{n_i}\nabla f(z^{n_i})) \in \partial g(x^{n_i+1})$. Since ∂g is maximal monotone, we have

$$\langle v - x^{n_i+1}, u - \nabla f(v) - \frac{1}{\alpha_{n_i}}(z^{n_i} - x^{n_i+1} - \alpha_{n_i}\nabla f(z^{n_i}))\rangle \geq 0.$$

This shows that

$$
\begin{aligned}
& \langle v - x^{n_i+1}, u\rangle \\
\geq\ & \langle v - x^{n_i+1}, \nabla f(v) + \frac{1}{\alpha_{n_i}}(z^{n_i} - x^{n_i+1} - \alpha_{n_i}\nabla f(z^{n_i}))\rangle \\
=\ & \langle v - x^{n_i+1}, \nabla f(v) - \nabla f(x^{n_i})\rangle + \langle v - x^{n_i+1}, \frac{1}{\alpha_{n_i}}(z^{n_i} - x^{n_i+1})\rangle \\
=\ & \langle v - x^{n_i+1}, \nabla f(v) - \nabla f(z^{n_i+1})\rangle + \langle v - x^{n_i+1}, \nabla f(x^{n_i+1}) - \nabla f(z^{n_i})\rangle \\
& + \langle v - x^{n_i+1}, \frac{1}{\alpha_{n_i}}(z^{n_i} - x^{n_i+1})\rangle \\
\geq\ & \langle v - x^{n_i+1}, \nabla f(x^{n_i+1}) - \nabla f(z^{n_i})\rangle + \langle v - x^{n_i+1}, \frac{1}{\alpha_{n_i}}(z^{n_i} - x^{n_i+1})\rangle.
\end{aligned}
$$

Since $\lim_{n\to\infty}\|z^n - x^{n+1}\| = 0$ and by assumptions, we have $\lim_{n\to\infty}\|\nabla f(z^n) - \nabla f(x^{n+1})\| = 0$. From Remark 3.1 in [38], we know that (α_n) is bounded from below by $\min\left\{\alpha_0, \frac{\delta}{L}\right\}$ and $\lim_{n\to\infty}\alpha_n = \alpha > 0$. Hence we obtain

$$
\langle v - x^{\infty}, u\rangle = \lim_{i\to\infty}\langle v - x^{n_i+1}, u\rangle \geq 0.
$$

Hence, $0 \in (\nabla f + \partial g)x^{\infty}$, and consequently $x^{\infty} \in \operatorname{argmin}(f+g)$. This gives that (x^n) converges weakly to a point in $\operatorname{argmin}(f+g)$ by applying Theorem 3.2.62. We thus complete the proof. □

4.4 Convergence analysis of a modified forward-backward splitting algorithm for minimization problems

In this section, we assume that $f : H \to \mathbb{R}\cup\{+\infty\}$ and $g : H \to \mathbb{R}\cup\{+\infty\}$ are proper, lower semi-continuous and convex functions, f is uniformly continuous on bounded sets and ∇f is bounded on bounded sets. The following is our algorithm.

Algorithm 4.4.1 New modified forward-backward splitting algorithm (NMFBS)

Initialization: Given $\sigma > 0$, $\theta > 0$, $\delta \in (0,1)$ and $\theta_1 > 0$.

Iterative step: Let $x^0 = x^1 \in H$ and calculate x^{n+1} as follows:

Step 1. Compute the inertial step:

$$z^n = x^n + \theta_n(x^n - x^{n-1}).$$

Step 2. Compute:

$$y^n = \text{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n)) + \alpha_n(\nabla f(z^n) - \nabla f(\text{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n)))),$$

where $\alpha_n = \sigma\theta^{m_n}$ and m_n is the smallest number such that:

$$\begin{aligned} & \alpha_n^2(\|\nabla f(z^n) - \nabla f(\text{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n)))\|^2 \\ & +\|\nabla f(y^n) - \nabla f(\text{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)))\|^2) \\ \leq \; & \delta^2(\|z^n - \text{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n))\|^2 \\ & +\|y^n - \text{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n))\|^2). \end{aligned} \tag{4.4.1}$$

Step 3. Compute the x^{n+1} step:

$$x^{n+1} = \text{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)) + \alpha_n(\nabla f(y^n) - \nabla f(\text{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)))).$$

Set $n := n + 1$ and return to **Step1.**

Following the proof as in [19], we can show the following lemma.

Lemma 4.4.2 *The Linesearch* (4.4.1) *has a finite steps.*

Theorem 4.4.3 *Suppose that $\alpha_n \geq \alpha$ for some $\alpha > 0$, $\theta_n \geq 0$ and $\sum_{n=1}^{\infty} \theta_n < +\infty$. Then, we have for each $x_* \in \text{argmin}(f+g) \cap \Omega$, $\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1 + 2\theta_j)$ where $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$ and (x^n) generated by Algorithm 4.4.1* *converges weakly to an element of* $\operatorname{argmin}(f+g)$.

Proof. Let $x_* \in \operatorname{argmin}(f+g)$ and set $p^n = \operatorname{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n))$. Then we obtain

$$y^n = p^n + \alpha_n(\nabla f(z^n) - \nabla f(p^n)). \tag{4.4.2}$$

Moreover, we have

$$z^n - p^n - \alpha_n \nabla f(z^n) \in \alpha_n \partial g(p^n). \tag{4.4.3}$$

Using (4.4.2), we see that

$$\alpha_n \nabla f(z^n) = y^n - p^n + \alpha_n \nabla f(p^n)). \tag{4.4.4}$$

Combining (4.4.3) and (4.4.4), we have

$$z^n - y^n - \alpha_n \nabla f(p^n)) \in \alpha_n \partial g(p^n). \tag{4.4.5}$$

Now, set $r^n = \operatorname{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n))$. Then, we obtain

$$x^{n+1} = r^n + \alpha_n(\nabla f(y^n) - \nabla f(r^n)). \tag{4.4.6}$$

Also, we have

$$y^n - x^{n+1} - \alpha_n \nabla f(r^n)) \in \alpha_n \partial g(r^n). \tag{4.4.7}$$

Since $x_* \in \operatorname{argmin}(f+g)$, we obtain $-\alpha_n \nabla f(x_*) \in \alpha_n \partial(x_*)$. Thus, by (4.4.5), (4.4.7) and the monotonicity of ∂g, we have

$$\langle z^n - y^n - \alpha_n(\nabla f(p^n) - \nabla f(x_*)), p^n - x_* \rangle \geq 0$$

and

$$\langle y^n - x^{n+1} - \alpha_n(\nabla f(r^n) - \nabla f(x_*)), r^n - x_*\rangle \geq 0.$$

So, we have $\langle z^n - y^n, p^n - x_*\rangle \geq 0$ and $\langle y^n - x^{n+1}, r^n - x_*\rangle \geq 0$ by the monotonicity of ∇f. Thus we have

$$\langle z^n - y^n, p^n - y^n\rangle + \langle z^n - y^n, y^n - x_*\rangle \geq 0 \tag{4.4.8}$$

and

$$\langle y^n - x^{n+1}, r^n - x^{n+1}\rangle + \langle y^n - x^{n+1}, x^{n+1} - x_*\rangle \geq 0. \tag{4.4.9}$$

We note that $\|x \pm y\|^2 = \|x\|^2 \pm 2\langle x, y\rangle + \|y\|^2$ for all $x, y \in H$. Using (4.4.8), we have

$$\begin{aligned} & 2(\langle z^n - y^n, p^n - y^n\rangle + \langle z^n - y^n, y^n - x_*\rangle) \\ = \; & \|z^n - x_*\|^2 - \|y^n - x_*\|^2 + \|y^n - p^n\|^2 - \|z^n - p^n\|^2. \end{aligned} \tag{4.4.10}$$

Using (4.4.9), we have

$$\begin{aligned} & 2(\langle y^n - x^{n+1}, r^n - x^{n+1}\rangle + \langle y^n - x^{n+1}, x^{n+1} - x_*\rangle) \\ = \; & \|y^n - x_*\|^2 - \|x^{n+1} - x_*\|^2 + \|x^{n+1} - r^n\|^2 - \|y^n - r^n\|^2. \end{aligned} \tag{4.4.11}$$

From (4.4.8),(4.4.9), (4.4.10) and (4.4.11), we obtain

$$\begin{aligned} \|x^{n+1} - x_*\|^2 \leq \; & \|z^n - x_*\|^2 + \|y^n - p^n\|^2 - \|z^n - p^n\|^2 + \|x^{n+1} - r^n\|^2 \\ & - \|y^n - r^n\|^2. \end{aligned} \tag{4.4.12}$$

Using (4.4.1), (4.4.2), (4.4.6) and (4.4.12), we obtain

$$\begin{aligned}
& \|x^{n+1} - x_*\|^2 \\
\leq\ & \|z^n - x_*\|^2 + \|p^n + \alpha_n(\nabla f(z^n) - \nabla f(p^n)) - p^n\|^2 - \|z^n - p^n\|^2 \\
& + \|r^n + \alpha_n(\nabla f(y^n) - \nabla f(r^n)) - r^n\|^2 - \|y^n - r^n\|^2 \\
=\ & \|z^n - x_*\|^2 + \alpha_n^2(\|\nabla f(z^n) - \nabla f(p^n)\|^2 + \|\nabla f(y^n) - \nabla f(r^n)\|^2) \\
& - \|z^n - p^n\|^2 - \|y^n - r^n\|^2 \\
\leq\ & \|z^n - x_*\|^2 + \delta^2(\|z^n - p^n\|^2 + \|y^n - r^n\|^2) - \|z^n - p^n\|^2 - \|y^n - r^n\|^2 \\
=\ & \|z^n - x_*\|^2 - (1-\delta^2)\|z^n - p^n\|^2 - (1-\delta^2)\|y^n - r^n\|^2. \qquad (4.4.13)
\end{aligned}$$

Next, we will show that $\lim_{n\to\infty} \|x^n - x_*\|$ exists. From (4.4.13), we see that

$$\begin{aligned}
\|x^{n+1} - x_*\| &\leq \|z^n - x_*\| \\
&= \|x^n - x_*\| + \theta_n\|x^n - x^{n-1}\| \\
&\leq \|x^n - x_*\| + \theta_n(\|x^n - x_*\| + \|x^{n-1} - x_*\|) \\
&= (1+\theta_n)\|x^n - x_*\| + \theta_n\|x^{n-1} - x_*\|. \qquad (4.4.14)
\end{aligned}$$

By Lemma 3.1.15, we have

$$\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1 + 2\theta_j),$$

where $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$. Since $\sum_{n=1}^{\infty} \theta_n < +\infty$, we have (x^n) is bounded. Thus $\sum_{n=1}^{\infty} \theta_n\|x^n - x^{n-1}\| < +\infty$. By Lemma 3.1.14 and (4.4.14), we have $\lim_{n\to\infty} \|x^n - x_*\|$ exists. From (4.4.13), we see that

$$\begin{aligned}
& \|x^{n+1} - x_*\|^2 \\
\leq\ & \|z^n - x_*\|^2 - (1-\delta^2)\|z^n - p^n\|^2 - (1-\delta^2)\|y^n - r^n\|^2
\end{aligned}$$

$$
\begin{aligned}
&= \|x^n + \theta_n(x^n - x^{n-1}) - x_*\|^2 - (1-\delta^2)\|z^n - p^n\|^2 - (1-\delta^2)\|y^n - r^n\|^2 \\
&= \|x^n - x_*\|^2 + 2\theta_n\|x^n - x_*\|\|x^n - x^{n-1}\| + \theta_n^2\|x^n - x^{n-1}\|^2 \\
&\quad -(1-\delta^2)\|z^n - p^n\|^2 - (1-\delta^2)\|y^n - r^n\|^2.
\end{aligned}
$$

Noting $\lim_{n\to\infty} \theta_n\|x^n - x^{n-1}\| = 0$, $\lim_{n\to\infty} \|x^n - x_*\|$ exists and $\delta \in (0,1)$, we have

$$\lim_{n\to\infty} \|z^n - p^n\| = 0 \tag{4.4.15}$$

and

$$\lim_{n\to\infty} \|y^n - r^n\| = 0. \tag{4.4.16}$$

Since ∇f is uniformly continuous on bounded sets, we have

$$\lim_{n\to\infty} \|\nabla f(z^n) - \nabla f(p^n)\| = 0. \tag{4.4.17}$$

and

$$\lim_{n\to\infty} \|\nabla f(y^n) - \nabla f(r^n)\| = 0. \tag{4.4.18}$$

By definition of z^n, it is easy to see that $\lim_{n\to\infty} \|z^n - x^n\| = 0$. Then,

$$
\begin{aligned}
\|p^n - x^n\| &\leq \|z^n - p^n\| + \|z^n - x^n\| \\
&\to 0 \text{ as } n \to \infty. \qquad (4.4.19)
\end{aligned}
$$

From (4.4.15), (4.4.16), (4.4.17) and (4.4.19), we obtain

$$
\begin{aligned}
\|r^n - x^n\| &\leq \|r^n - y^n\| + \|y^n - p^n\| + \|p^n - x^n\| \\
&= \|r^n - y^n\| + \|p^n + \alpha_n(\nabla f(z^n) - \nabla f(p^n)) - p^n\| + \|p^n - x^n\| \\
&= \|r^n - y^n\| + \alpha_n\|\nabla f(z^n) - \nabla f(p^n)\| + \|p^n - x^n\|
\end{aligned}
$$

$$\to \quad 0 \text{ as } n \to \infty.$$

By the boundedness of (x^n), we assume that $\bar{x}$ is a weak limit point of (x^n), *i.e.*, there is a subsequence (x^{n_i}) of (x^n) such that $x^{n_i} \rightharpoonup \bar{x}$. Since $\lim_{i\to\infty} \|r^{n_i} - x^{n_i}\| = 0$, we also obtain $r^{n_i} \rightharpoonup \bar{x}$ as $i \to \infty$. Using (3.3.1), we obtain

$$\begin{aligned} \frac{y^{n_i} - r^{n_i}}{\alpha_{n_i}} - \nabla f(y^{n_i}) &= \frac{y^{n_i} - \operatorname{prox}_{\alpha_{n_i} g}(y^{n_i} - \alpha_{n_i}\nabla f(y^{n_i})) - \alpha_{n_i}\nabla f(y^{n_i})}{\alpha_{n_i}} \\ &\in \partial g(y^{n_i} - \alpha_{n_i}\nabla f(y^{n_i})). \end{aligned}$$

It follows that

$$\begin{aligned} \frac{y^{n_i} - r^{n_i}}{\alpha_{n_i}} + \nabla f(r^{n_i}) - \nabla f(y^{n_i}) &\in \nabla f(r^{n_i}) + \partial g(r^{n_i}) \\ &\subseteq \partial(f+g)(r^{n_i}). \end{aligned}$$

By passing $i \to \infty$ and using (4.4.16) and (4.4.18), we have $0 \in \nabla f(\bar{x}) + \partial g(\bar{x})$ by Lemma 3.3.3. Hence, by Theorem 3.2.62, we can conclude that (x^n) converges weakly to a point in $\operatorname{argmin}(f+g)$. We thus complete the proof. □

4.5 New inertial forward-backward algorithm for convex minimization problems

In this section, we assume that $f : H \to \mathbb{R} \cup \{+\infty\}$ and $g : H \to \mathbb{R} \cup \{+\infty\}$ are proper, lower semi-continuous and convex functions, f is uniformly continuous on bounded sets and ∇f is bounded on bounded sets. The following is our algorithm.

Algorithm 4.5.1 Inertial modified forward-backward splitting algorithm (IMFBS)

Initialization: Given $\sigma, \theta, \mu_1 > 0$, $\delta \in (0, \frac{1}{2})$ and $\rho \in (0, 1)$.

Iterative step: Let $x^0 = x^1 \in H$, calculate x^{n+1} and μ_{n+1} as follows:

Step1. Compute the inertial step:

$$v^n = x^n + \theta_n(x^n - x^{n-1}).$$

Step2. Compute the forward-backward step:

$$y^n = \mathrm{prox}_{\alpha_n g}(v^n - \alpha_n \nabla f(v^n)),$$

where $\alpha_n = \sigma\theta^{m_n}$ and m_n is the smallest nonnegative number such that:

$$\alpha_n \|\nabla f(y^n) - \nabla f(v^n)\| \leq \delta \|y^n - v^n\|. \tag{4.5.1}$$

Step3. Compute the forward-backward step:

$$z^n = \mathrm{prox}_{\mu_n g}(y^n - \mu_n \nabla f(y^n)).$$

Step4. Compute the x^{n+1} step:

$$x^{n+1} = z^n + \mu_n(\nabla f(y^n) - \nabla f(z^n))$$

and update

$$\mu_{n+1} = \begin{cases} \min\left(\frac{\rho\|y^n - z^n\|}{\|\nabla f(y^n) - \nabla f(z^n)\|}, \mu_n\right) & \mathit{if}\ \|\nabla f(y^n) - \nabla f(z^n)\| \neq 0; \\ \mu_n & \mathit{otherwise}. \end{cases} \tag{4.5.2}$$

Set $n := n+1$ and return to **Step1.**

Theorem 4.5.2 *Suppose that* $\alpha_n \geq \alpha$ *for some* $\alpha > 0$, $\theta_n \geq 0$ *and* $\sum_{n=1}^{\infty} \theta_n < +\infty$. *Then, we have for each* $x_* \in \operatorname{argmin}(f+g) \cap \Omega$, $\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1+2\theta_j)$ *where* $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$ *and the sequence* (x^n) *generated by Algorithm*

4.5.1 converges weakly to a minimizer of $f+g$.

Proof. Let $x_* \in \operatorname{argmin}(f+g)$. Following the definition of z^n, we have

$$y^n - z^n - \mu_n \nabla f(y^n) \in \mu_n \partial g(z^n). \tag{4.5.3}$$

By definition of x^{n+1}, we see that

$$\mu_n \nabla f(y^n) = x^{n+1} - z^n + \mu_n \nabla f(z^n). \tag{4.5.4}$$

From (4.5.3) and (4.5.4), we have

$$y^n - x^{n+1} - \mu_n \nabla f(z^n) \in \mu_n \partial g(z^n). \tag{4.5.5}$$

Since $x_* \in \operatorname{argmin}(f+g)$, we get $-\mu_n \nabla f(x_*) \in \mu_n \partial g(x_*)$. Thus, by relation (4.5.5) and the monotonicity of ∂g, we have

$$\langle y^n - x^{n+1} - \mu_n(\nabla f(z^n) - \nabla f(x_*)), z^n - x_* \rangle \geq 0.$$

This together with the monotonicity of ∇f implies that

$$\langle y^n - x^{n+1}, z^n - x_* \rangle \geq 0.$$

Hence, we have

$$\langle y^n - x^{n+1}, z^n - x^{n+1} \rangle + \langle y^n - x^{n+1}, x^{n+1} - x_* \rangle \geq 0. \tag{4.5.6}$$

We know that $\|x \pm y\|^2 = \|x\|^2 \pm 2\langle x, y \rangle + \|y\|^2$. So by (4.5.6), we obtain

$$\frac{1}{2}[\|y^n - x^{n+1}\|^2 + \|x^{n+1} - z^n\|^2 - \|y^n - z^n\|^2]$$

$$+\frac{1}{2}[\|y^n - x_*\|^2 - \|y^n - x^{n+1}\|^2 - \|x^{n+1} - x_*\|^2] \geq 0.$$

It implies that

$$\|x^{n+1} - x_*\|^2 \leq \|y^n - x_*\|^2 + \|x^{n+1} - z^n\|^2 - \|y^n - z^n\|^2. \tag{4.5.7}$$

By definition of x^{n+1} and (4.5.2), we have

$$\begin{aligned} \|x^{n+1} - z^n\|^2 &\leq \|z^n + \mu_n(\nabla f(y^n) - \nabla f(z^n)) - z^n\|^2 \\ &= \mu_n^2 \|\nabla f(y^n) - \nabla f(z^n)\|^2. \end{aligned} \tag{4.5.8}$$

Note that

$$\mu_{n+1} = \min\left(\frac{\rho\|y^n - z^n\|}{\|\nabla f(y^n) - \nabla f(z^n)\|}, \mu_n\right) \leq \frac{\rho\|y^n - z^n\|}{\|\nabla f(y^n) - \nabla f(z^n)\|}.$$

It follows that

$$\|\nabla f(y^n) - \nabla f(z^n)\| \leq \frac{\rho}{\mu_{n+1}}\|y^n - z^n\|. \tag{4.5.9}$$

Combining (4.5.8) and (4.5.9), we have

$$\|x^{n+1} - z^n\|^2 \leq \frac{\rho^2\mu_n^2}{\mu_{n+1}^2}\|y^n - z^n\|^2. \tag{4.5.10}$$

By definition of y^n, we have

$$\frac{v^n - y^n}{\alpha_n} - \nabla f(v^n) \in \partial g(y^n).$$

By the convexity of g, we get

$$g(x_*) - g(y^n) \geq \left\langle \frac{v^n - y^n}{\alpha_n} - \nabla f(v^n), x_* - y^n \right\rangle. \tag{4.5.11}$$

By the convexity of f, we see that

$$f(x_*) - f(v^n) \geq \langle \nabla f(v^n), x_* - v^n \rangle. \tag{4.5.12}$$

Combining (4.5.1), (4.5.11) and (4.5.12), we have

$$\begin{aligned}
& (f+g)(x_*) \\
\geq\ & g(y^n) + f(v^n) + \left\langle \frac{v^n - y^n}{\alpha_n} - \nabla f(v^n), x_* - y^n \right\rangle + \langle \nabla f(v^n), x_* - v^n \rangle \\
=\ & g(y^n) + f(v^n) + \frac{1}{\alpha_n}\langle v^n - y^n, x_* - y^n \rangle + \langle \nabla f(v^n) - \nabla f(y^n), y^n - v^n \rangle \\
& + \langle \nabla f(y^n), y^n - v^n \rangle \\
\geq\ & g(y^n) + f(v^n) + \frac{1}{\alpha_n}\langle v^n - y^n, x_* - y^n \rangle - \|\nabla f(v^n) - \nabla f(y^n)\| \|y^n - v^n\| \\
& + \langle \nabla f(y^n), y^n - v^n \rangle \\
\geq\ & g(y^n) + f(v^n) + \frac{1}{\alpha_n}\langle v^n - y^n, x_* - y^n \rangle - \frac{\delta}{\alpha_n}\|y^n - v^n\|^2 \\
& + \langle \nabla f(y^n), y^n - v^n \rangle.
\end{aligned} \tag{4.5.13}$$

By (4.5.13) and the convexity of f, we get

$$\begin{aligned}
& \frac{1}{\alpha_n}\langle v^n - y^n, y^n - x_* \rangle \\
\geq\ & g(y^n) + f(v^n) - (f+g)(x_*) - \frac{\delta}{\alpha_n}\|y^n - v^n\|^2 + f(y^n) - f(v^n) \\
=\ & (f+g)(y^n) - (f+g)(x_*) - \frac{\delta}{\alpha_n}\|y^n - v^n\|^2.
\end{aligned} \tag{4.5.14}$$

We see that $\|v^n - x_*\|^2 = \|v^n - y^n\|^2 + 2\langle v^n - y^n, y^n - x_* \rangle + \|y^n - x_*\|^2$. By (4.5.14), we have

$$\begin{aligned}
& \|v^n - x_*\|^2 - \|v^n - y^n\|^2 - \|y^n - x_*\|^2 \\
\geq\ & 2\alpha_n[(f+g)(y^n) - (f+g)(x_*)] - 2\delta\|y^n - v^n\|^2.
\end{aligned}$$

It implies that

$$\|y^n - x_*\|^2 \leq \|v^n - x_*\|^2 - (1-2\delta)\|v^n - y^n\|^2 - 2\alpha_n[(f+g)(y^n) - (f+g)(x_*)]. \tag{4.5.15}$$

Hence, from (4.5.7), (4.5.10) and (4.5.15), we obtain

$$\|x^{n+1} - x_*\|^2 \leq \|v^n - x_*\|^2 - (1-2\delta)\|v^n - y^n\|^2 - (1 - \frac{\rho^2\mu_n^2}{\mu_{n+1}^2})\|y^n - z^n\|^2 - 2\alpha_n[(f+g)(y^n) - (f+g)(x_*)]. \tag{4.5.16}$$

Now, we will show that (x^n) is bounded. From (4.5.16) and by definition of (v^n), we have

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

This shows that

$$\|x^{n+1} - x_*\| \leq (1+\theta_n)\|x^n - x_*\| + \theta_n\|x^{n-1} - x_*\|.$$

By Lemma 3.1.15, we conclude that

$$\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1+2\theta_j)$$

where $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$. Since $\sum_{n=1}^{\infty}\theta_n < +\infty$, we have (x^n) is bounded. From (4.5.16), we have

$$\|x^{n+1} - x_*\|^2$$

$$\begin{aligned}
&\leq \|x^n + \theta_n(x^n - x^{n-1}) - x_*\|^2 - (1-2\delta)\|v^n - y^n\|^2 \\
&\quad -2\alpha_n[(f+g)(y^n) - (f+g)(x_*)] - (1 - \frac{\rho^2\mu_n^2}{\mu_{n+1}^2})\|y^n - z^n\|^2 \\
&= \|x^n - x_*\|^2 + 2\theta_n\|x^n - x_*\|\|x^n - x^{n-1}\| + \theta^2\|x^n - x^{n-1}\|^2 \\
&\quad -(1-2\delta)\|v^n - y^n\|^2 - 2\alpha_n[(f+g)(y^n) - (f+g)(x_*)] \\
&\quad -(1 - \frac{\rho^2\mu_n^2}{\mu_{n+1}^2})\|y^n - z^n\|^2. \qquad (4.5.17)
\end{aligned}$$

Since $\lim_{n\to\infty} \theta_n\|x^n - x^{n-1}\| = 0$, $\lim_{n\to\infty} \|x^n - x_*\|$ exists and $1 - 2\delta > 0$, we have

$$\lim_{n\to\infty} \|v^n - y^n\| = 0.$$

By definition of v^n, we have $\lim_{n\to\infty} \|v^n - x^n\| = 0$. Then,

$$\begin{aligned}
\|y^n - x^n\| &\leq \|v^n - y^n\| + \|v^n - x^n\| \\
&\to 0 \text{ as } n \to \infty.
\end{aligned}$$

Since $\lim_{n\to\infty} \left(1 - \frac{\mu_n^2\rho^2}{\mu_{n+1}^2}\right) = 1 - \rho^2 > 0$, we have $\lim_{n\to\infty} \|y^n - z^n\| = 0$. So, we have

$$\begin{aligned}
\|z^n - x^n\| &\leq \|y^n - z^n\| + \|y^n - x^n\| \\
&\to 0 \text{ as } n \to \infty.
\end{aligned}$$

Since the sequence (x^n) is bounded, assume that x^∞ is a weak limit point of (x^n), *i.e.*, there is a subsequence (x^{n_i}) of (x^n) such that $x^{n_i} \rightharpoonup x^\infty$. Since $\lim_{i\to\infty} \|z^{n_i} - x^{n_i}\| = 0$, we also obtain $z^{n_i} \rightharpoonup x^\infty$ as $i \to \infty$. Since (y^{n_i}) is bounded, $\lim_{i\to\infty} \|y^{n_i} - z^{n_i}\| = 0$ and ∇f is uniformly continuous on H, we have

$$\lim_{i\to\infty} \|\nabla f(y^{n_i}) - \nabla f(z^{n_i})\| = 0.$$

From (3.3.1), we obtain

$$\begin{aligned}\frac{y^{n_i}-z^{n_i}}{\alpha_{n_i}}-\nabla f(y^{n_i}) &= \frac{y^{n_i}-\operatorname{prox}_{\alpha_{n_i}g}(y^{n_i}-\alpha_{n_i}\nabla f(y^{n_i}))-\alpha_{n_i}\nabla f(y^{n_i})}{\alpha_{n_i}}\\ &\in \partial g(y^{n_i}-\alpha_{n_i}\nabla f(y^{n_i})).\end{aligned}$$

It follows that

$$\begin{aligned}\frac{y^{n_i}-z^{n_i}}{\alpha_{n_i}}+\nabla f(z^{n_i})-\nabla f(y^{n_i}) &\in \nabla f(z^{n_i})+\partial g(z^{n_i})\\ &\subseteq \partial(f+g)(z^{n_i}).\end{aligned}$$

By passing $i\to\infty$ and using Lemma 3.3.3, we have $0\in\nabla f(x^\infty)+\partial g(x^\infty)$. Thus $x^\infty\in\operatorname{argmin}(f+g)$. Hence, by Theorem 3.2.62, we can conclude that (x^n) converges weakly to a point in $\operatorname{argmin}(f+g)$. We thus complete the proof. □

4.6 Numerical examples and applications

In this section, we first consider a simple linear inverse problem as the following from

$$b = Ax+\varepsilon \tag{4.6.1}$$

where $x\in\mathbb{R}^{n\times 1}$ is the solution of the problem to be approximated, $A\in\mathbb{R}^{m\times n}$ and $\varepsilon\in\mathbb{R}^{m\times 1}$ are known and $\varepsilon\in\mathbb{R}^{m\times 1}$ is an additive noise vector. Such problems (4.6.1) arise in various applications such as the image and signal processing problems, astrophysical problems and data classification problems.

The purpose of the image restoration problem is to minimize the additive noise in which the classical estimator is the least squares (LS) given as follows:

$$\min_x \|Ax-b\|_2^2 \tag{4.6.2}$$

where $\|\cdot\|_2$ is ℓ^2-norm. However, this model still has some ill-conditions in the case that the least square solution has a huge norm which is thus meaningless. In 1977, Tikhonov and Arsenin [41] improved this ill-posed problem by introducing the regularization techniques which are known as the Tikhonov regularization (TR) model and it is of the following form:

$$\min_x(\|Ax-b\|_2^2+\lambda\|Lx\|_2^2) \tag{4.6.3}$$

for some regularization parameter $\lambda > 0$ and Tikhonov matrix L. On the other hand, another successful regularization method for improvement of Tikhonov regularization is known as the least absolute shrinkage and selection operator (LASSO) which was introduced by Tibshirani (1996). The method is to find a solution

$$\min_x(\|Ax-b\|_2^2+\lambda\|x\|_1) \tag{4.6.4}$$

where $\|\cdot\|_1$ is ℓ^1- norm. The LASSO can be applied to regression problems and image restoration problems. For solving 4.6.3 and 4.6.4, we extend them to a general naturally formulation, that is, the problem of finding the minimizer of sum of two functions:

$$\min_x(f(x)+g(x)) \tag{4.6.5}$$

We next discuss some experiment results by using our proposal algorithm to solving the image deblurring problem and image inpainting problem. The numerical experiments are preformed by Matlab 2020b on a 64-bit MacBook Pro Chip Apple M1 and 8 GB of RAM. We use the Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index (SSIM) to measure a quality of restored images as the following:

Peak Signal-to-Noise Ratio (PSNR): PSNR [40] is measured in decibels (dB). The higher the PSNR, the better image has been completed to match the original image.

$$\mathrm{PSNR} = 20\log_{10}\left(\frac{\mathrm{MAX}_f}{\sqrt{\mathrm{MSE}}}\right)$$

and

$$\mathrm{MSE} = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}|x(i,j) - x_*(i,j)|^2$$

where x is the original image, x_* represents completed image through the model, m represents image pixel rows, n represents image pixel columns, i and j represents row and column index respectively. MAX_f is a constant equal to 255.

Structural Similarity Index (SSIM): The Structural Similarity (SSIM) Index [40] depends on computation of terms, namely the luminance, contrast and structural term.

$$\mathrm{SSIM}(x,y) = [l(x,y)] \times [c(x,y)] \times [s(x,y)]$$

where The luminance comparison is defined by

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1},$$

$C_1 = (K_1L)^2$, L is the dynamic range of the pixel values and K_1 is a small constant. The contrast comparison is defined by

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2},$$

$C_2 = (K_2L)^2$, K_2 is a non-negative constant and a small constant. The structural comparison is defined by

$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3},$$

and x_i, x_{*i} are the i pixel value of the estimated image and the reference image. We denote N by the total number of pixel.

- μ_x is the average of x (the original image):

$$\mu_x = \frac{1}{N}\sum_{i=0}^{N} x_i.$$

- μ_{x_*} is the average of x_* (the restored image):

$$\mu_{x_* i} = \frac{1}{N}\sum_{i=0}^{N} x_{*i}.$$

- σ_x is the variance of x:

$$\sigma_x = \left(\frac{1}{N-1}\sum_{i=0}^{N}(x_i - \mu_x)^2\right)^{1/2}.$$

- σ_{x_*} is the variance of x_*:

$$\sigma_{x_*} = \left(\frac{1}{N-1}\sum_{i=0}^{N}(x_{*i} - \mu_x)^2\right)^{1/2}.$$

- σ_{xx_*} is the covariance of x and x_*:

$$\sigma_{xx_*} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)(x_{*i} - \mu_{x_{*i}}).$$

The index simplifies to:

$$\mathrm{SSIM}(x, x_*) = \frac{(2\mu_x\mu_{x_*} + C_1)(2\sigma_{xx_*} + C_2)}{(\mu_x^2 + \mu_{x_*}^2 + C_1)(\sigma_x^2 + \sigma_{x_*}^2 + C_2)}.$$

SSIM ranges from 0 to 1, where 1 means flawless recovery.

4.6.1 Image deblurring

In this experiment, we apply the unconstrained convex minimization problem (2.1.4) to image deblurring problem. We compare efficiency of our algorithms (Algorithm 4.1.1 (DIFB), Algorithm 4.3.1 (IFBAS), Algorithm 4.4.1 (NMFBS) and Algorithm 4.5.1 (IMFBS)) with Algorithm 2.1.6 (RFB), Algorithm 2.1.12 (FBFS), Algorithm 2.1.8 (FISTA-BT), Algorithm 2.1.11 (NAGA), Algorithm 2.1.10 (FBS-CN) and Algorithm 2.1.10 (FISTA-CN).

The image deblurring problem is a basic linear inverse problem (4.6.1):

$$b = Ax + \varepsilon$$

where $x \in \mathbb{R}^n$ is the original image to be estimated, $\varepsilon \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$ are known and ε is an unknown noise. To approximate the original image in (4.6.1), we need to minimize the value of ε by using the LASSO (4.6.4) model:

$$\min_x(\|Ax - b\|_2^2 + \lambda\|x\|_1)$$

where $\lambda > 0$, $\|\cdot\|_2$ is the Euclidean norm and $\|\cdot\|_1$ is the ℓ_1-norm. In this situation, we choose the regularization parameter $\lambda = 5 \times 10^{-7}$. It is noted that problem (4.6.1) can be applied to (4.6.4) by setting

$$f(x) = \|Ax - b\|_2^2 \text{ and } g(x) = \lambda\|x\|_1$$

We choose the Lipschitz constant L of the gradient ∇f which is the maximum value of eigenvalues of the matrix A^TA.

We investigate throughout this experiment under the following setting Table 1. The initial point $x^0 = x^1$ are vectors of ones with the size of original images for all algorithms. Let θ_n be defined by

$$\theta_n = \begin{cases} \dfrac{t_n - 1}{t_{n+1}} \text{ where } t_{n+1} = \dfrac{1+\sqrt{1+4t_n^2}}{2}, & \text{if } 1 \leq n \leq M; \\ \dfrac{1}{n^2}, & \text{otherwise.} \end{cases}$$

Algorithms		Parameters				
Comparision	OUR	t_1	$\alpha_n \in (0, 2/L)$	$\sigma > 0$	$\theta \in (0,1)$	$\delta \in (0, 1/2)$
RFB			$1/\|A\|^2$			
FBFS			$1/\|A\|^2$			
FISTA-BT		1	$1/\|A\|^2$			
NAGA		1	$1/\|A\|^2$			
FBS-CN				0.12	0.51	0.92
FISTA-CN		1		0.12	0.51	0.92
	DIFB	1	$1/\|A\|^2$			
	IFBAS	1		0.12	0.51	0.92
	NMFBS	1		0.12	0.51	0.92
	IMFBS	1		0.12	0.51	0.92
Algorithms		**Parameters**				
Comparision	OUR			$\rho \in (0, 1/2)$	$\alpha_1 \geq 0$	$\mu_1 > 0$
RFB						
FBFS						
FISTA-BT						
NAGA						
FBS-CN						
FISTA-CN						
	DIFB					
	IFBAS				0.21	
	NMFBS			0.42		0.42
	IMFBS					

Table 1: Chosen parameters of each algorithm

The original image size $480 \times 640 \times 3$ and three different types of original image degraded by the blurring matrices are shown in Figures 1-4 and the original image size $640 \times 463 \times 3$ and three different types of original image degraded by the blurring matrices are shown in Figures 14-17. The results of the deblurred images with M iterations for each algorithms are shown in Tables 2.

Algorithms	Fig(A) with BM1-1			Fig(B) with BM2-1		
	PSNR	SSIM	CPU	PSNR	SSIM	CPU
RFB	24.9179	0.7749	65.4483	24.0673	0.6737	72.1916
FBFS	25.0707	0.7781	78.2448	24.1804	0.6785	91.3530
FISTA-BT	34.5671	0.8952	64.8912	33.8992	0.9069	71.4991
NAGA	35.1331	0.9026	79.7754	34.6569	0.9166	92.0715
FBS-CN	25.1850	0.7804	302.9514	24.2644	0.6821	392.5678
FISTA-CN	34.6542	0.8961	300.7232	34.0384	0.9089	395.9044
DIFB	35.8017	0.9105	78.1437	35.2470	0.9220	92.0930
IFBAS	35.6340	0.9087	303.3624	35.1282	0.9219	400.6912
NMFBS	37.2883	0.9283	551.2007	36.2275	0.9385	834.6453
IMFBS	35.6907	0.9090	552.6206	35.2976	0.9237	733.1543
	Fig(A) with BM1-2			Fig(B) with BM2-2		
	PSNR	SSIM	CPU	PSNR	SSIM	CPU
RFB	30.2306	0.8952	66.0884	32.0117	0.9190	79.3981
FBFS	30.3811	0.8974	77.9553	32.1782	0.9213	100.6219
FISTA-BT	39.0064	0.9606	65.2943	43.4527	0.9872	78.1251
NAGA	39.6384	0.9637	79.6475	44.3289	0.9892	102.4385
FBS-CN	30.4931	0.8990	299.3682	32.3014	0.9229	494.0503
FISTA-CN	39.1363	0.9611	301.1905	43.5981	0.9875	490.1089
DIFB	40.1953	0.9663	78.4121	44.9859	0.9903	103.2291
IFBAS	40.0068	0.9657	298.7557	44.7392	0.9901	497.1921
NMFBS	41.6020	0.9729	554.4284	46.6168	0.9930	747.0347
IMFBS	40.2021	0.9663	551.4667	45.0823	0.9907	747.7734
	Fig(A) with BM1-3			Fig(B) with BM2-3		
	PSNR	SSIM	CPU	PSNR	SSIM	CPU
RFB	33.3823	0.9262	65.2973	32.0160	0.9122	73.3129
FBFS	33.6269	0.9282	76.9898	32.2566	0.9149	92.1898
FISTA-BT	44.3344	0.9822	64.7505	44.7931	0.9874	72.0578
NAGA	44.9963	0.9853	78.0458	45.7750	0.9896	93.0511
FBS-CN	33.8078	0.9296	304.1175	32.4343	0.9167	404.1600
FISTA-CN	44.4917	0.9828	307.1597	45.0092	0.9880	400.4293
DIFB	45.6000	0.9872	78.2824	46.4959	0.9912	92.6083
IFBAS	45.4127	0.9861	307.6676	46.3783	0.9910	414.1358
NMFBS	48.0352	0.9921	558.8756	48.4738	0.9944	771.1393
IMFBS	45.8027	0.9877	564.3752	46.6290	0.9915	758.7245

Table 2: The results of deblurred images for each algorithm.

Figure 1: The original image of size $480 \times 640 \times 3$ (Fig(A)).

Figure 2: Out of focus blur (disk) with radius $r = 11$ (BM1-1).

Figure 3: Gaussian blur of the filter size [7 $\times$ 7] with standard deviation $\sigma = 13$ (BM1-2).

Figure 4: Motion blur specified with the motion length of 11 pixels and motion orientation $\theta = 16$ (BM1-3).

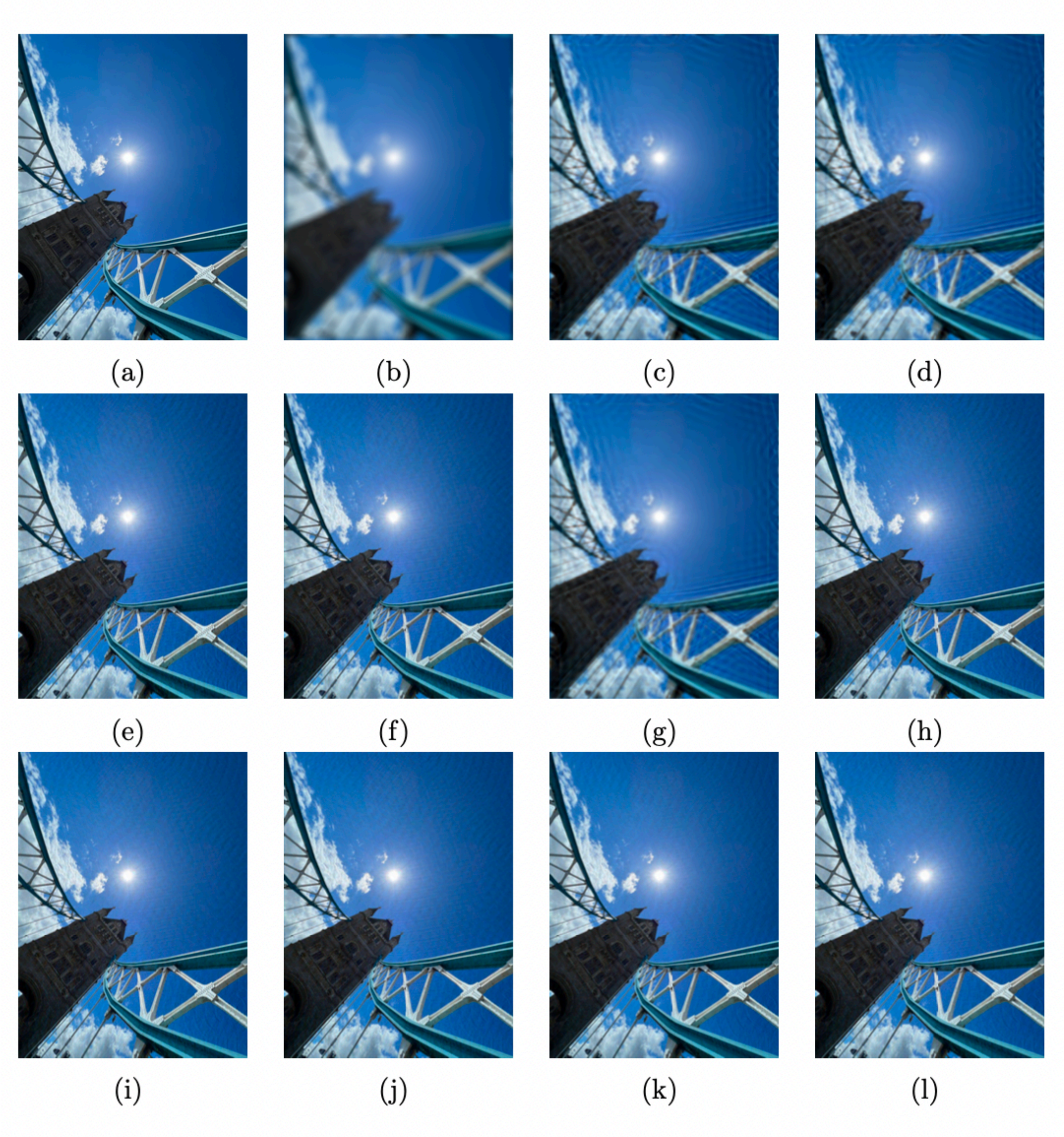
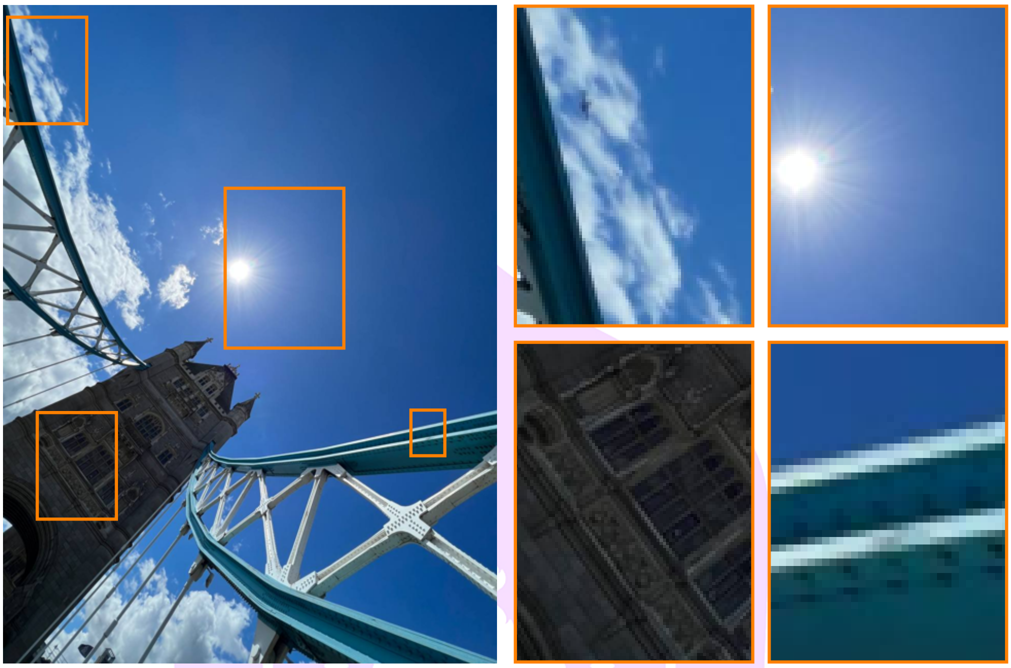


Figure 5: (a) The original image size $480 \times 640 \times 3$ (Fig(A)), (b) the deblurred RGB image by out of focus blur matrices (BM1-1) and the restored images for (c) RFB (PSNR:24.9179, SSIM:0.7749) (d) FBFS (PSNR:25.0707, SSIM:0.7781) (e) FISTA-BT (PSNR:34.5671, SSIM:0.8952) (f) NAGA (PSNR:35.1331, SSIM:0.9026) (g) FBS-CN (PSNR:25.1850, SSIM:0.7804) (h) FISTA-CN (PSNR:34.6542, SSIM:0.8961) (i) DIFB (PSNR:35.8017, SSIM:0.9105) (j) IFBAS (PSNR:35.6340, SSIM:0.9087) (k) NMFBS (PSNR:37.2883, SSIM:0.9283) and (l) IMFBS (PSNR:37.2883, SSIM:0.9283), respectively.



(a) (b)

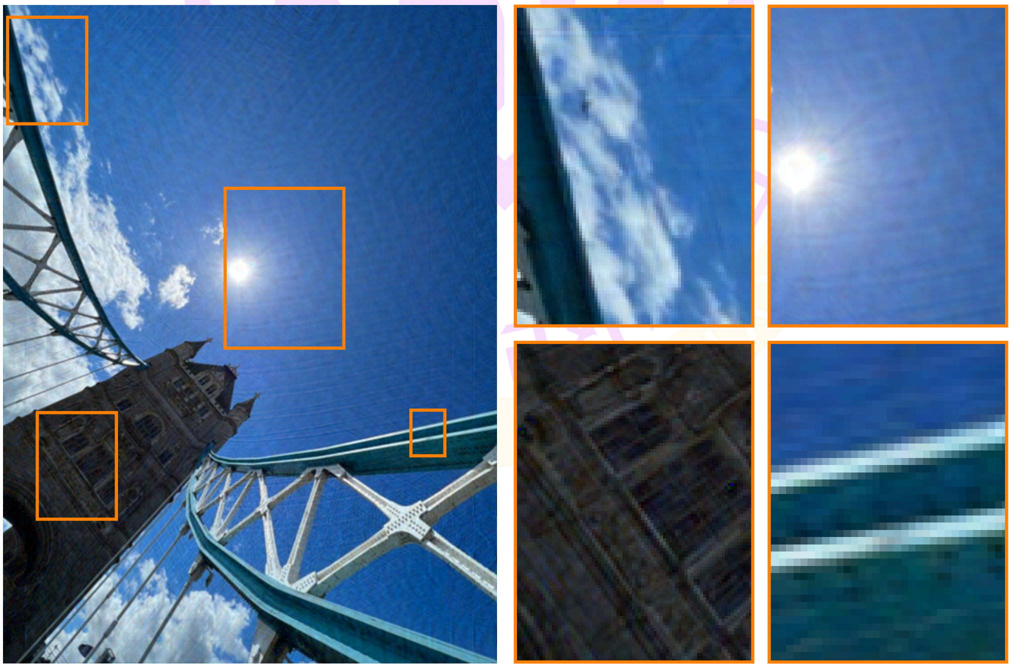
Figure 6: The original image size $480 \times 640 \times 3$ (Fig(A)) and zoom in the figures.

(a) (b)

Figure 7: The deblurred RGB image by out of focus blur matrices (BM1-1) and zoom in the figures.

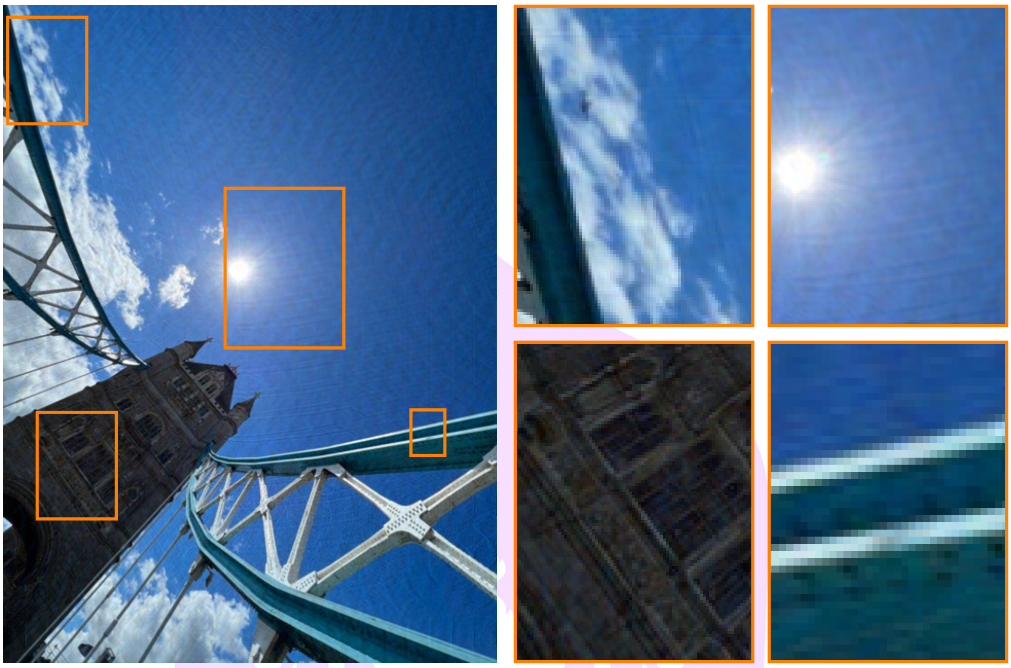
(a) (b)

Figure 8: The restored images for RFB (PSNR:24.9179, SSIM:0.7749) and zoom in the figures.



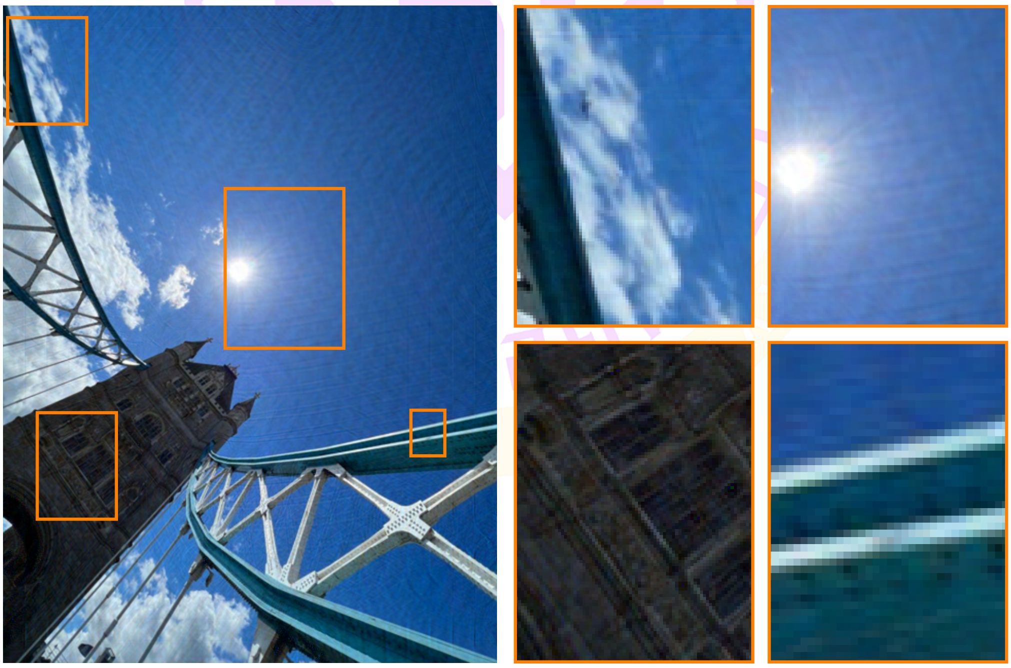
(a) (b)

Figure 9: The restored images for FISTA-BT (PSNR:34.5671, SSIM:0.8952).



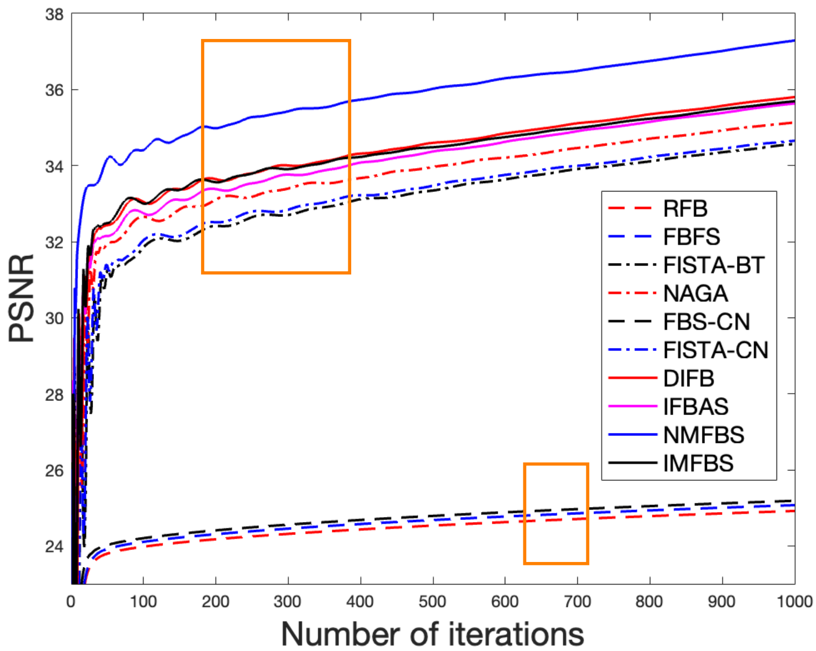
(a) (b)

Figure 10: The restored images for DIFB (PSNR:35.8017, SSIM:0.9105).

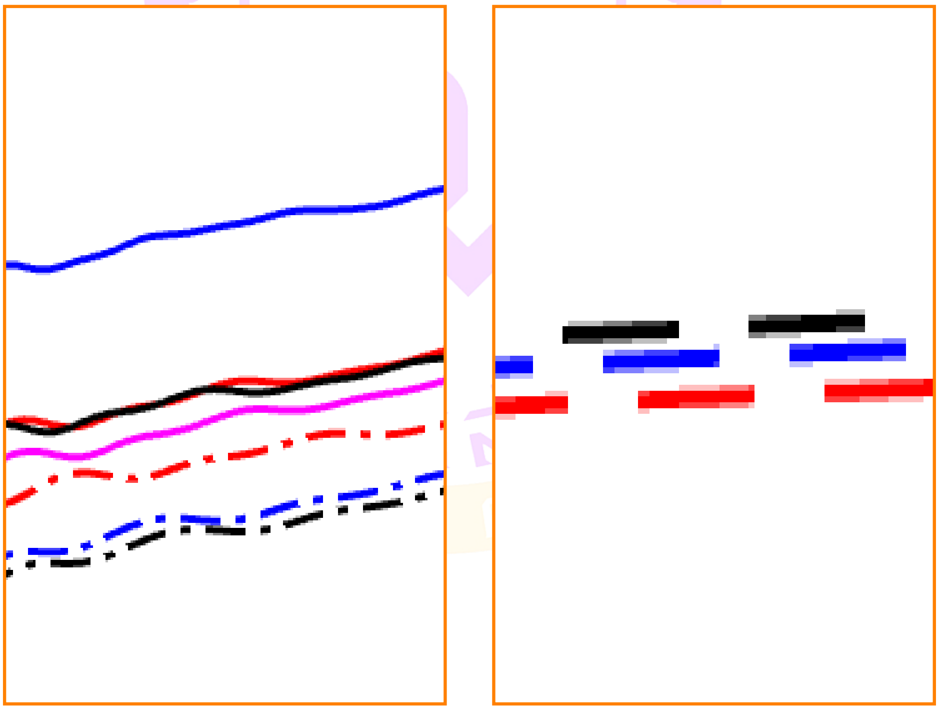


(a) (b)

Figure 11: The restored images for IFBAS (PSNR:35.6340, SSIM:0.9087).

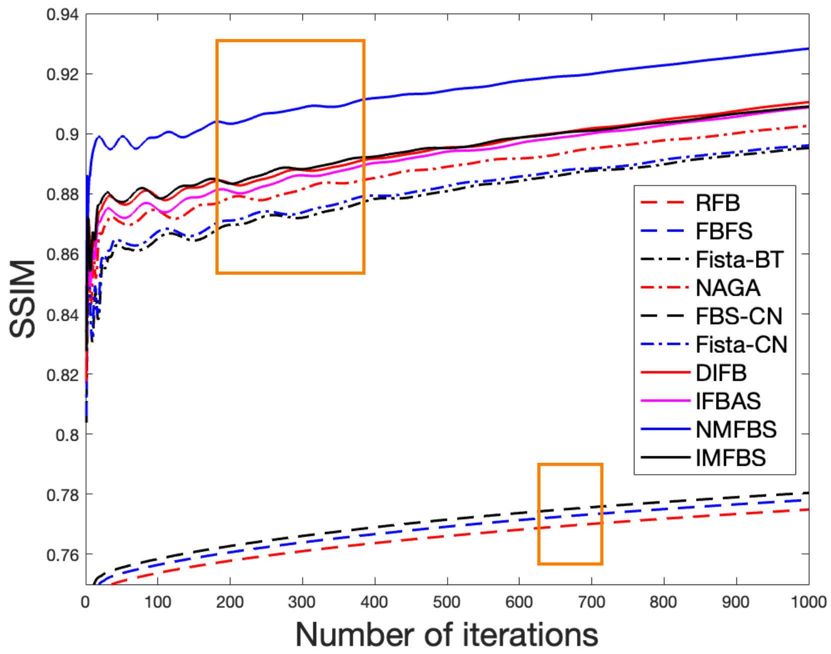


(a)

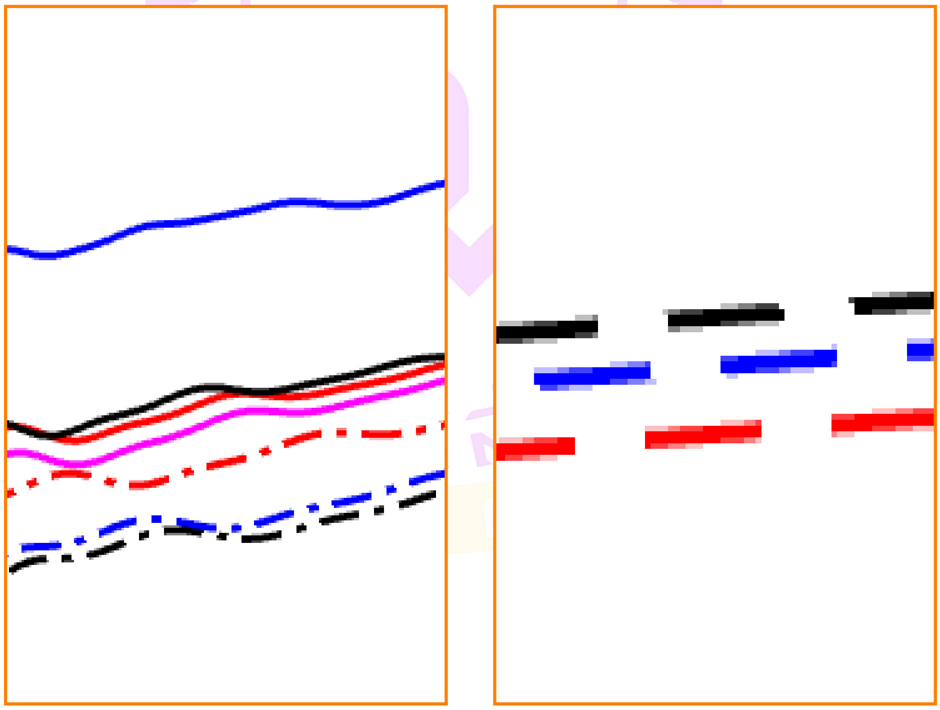


(b)

Figure 12: Graphs of PSNR plotting for out of focus blurs (BM1-1) and methods.



(a)



(b)

Figure 13: Graphs of SSIM plotting for out of focus blurs (BM1-1) and methods.

Figure 14: The original image of size $640 \times 463 \times 3$ (Fig(B)).

Figure 15: Out of focus blur (disk) with radius $r = 6$ (BM2-1).

Figure 16: Gaussian blur of the filter size $[6 \times 6]$ with standard deviation $\sigma = 11$ (BM2-2).

Figure 17: Motion blur specified with the motion length of 13 pixels and motion orientation $\theta = 57$ (BM2-3).

Figure 18: (a) The original image size 640 × 463 × 3 (Fig(B)), (b) the deblurred RGB image by motion blur specified with the motion length of 13 pixels and motion orientation 57 (BM2-3) and the restored images for (c) RFB (PSNR:32.0160, SSIM:0.9122) (d) FBFS (PSNR:32.2566, SSIM:0.9149) (e) FISTA-BT (PSNR:44.7931, SSIM:0.9874) (f) NAGA (PSNR:45.7750, SSIM:0.9896) (g) FBS-CN (PSNR:32.4343, SSIM:0.9167) (h) FISTA-CN (PSNR:45.0092, SSIM: 0.9880) (i) DIFB (PSNR:46.4959, SSIM: 0.9912) (j) IF-BAS (PSNR:46.3783, SSIM:0.9910) (k) NMFBS (PSNR:48.4738, SSIM: 0.9944) and (l) IMFBS (PSNR:46.6290, SSIM:0.9915) respectively.

(a) (b)

Figure 19: The original image size $640 \times 463 \times 3$ (Fig(B)).

(a) (b)

Figure 20: The deblurred RGB image by motion blur specified with the motion length of 13 pixels and motion orientation 57 (BM2-3).

(a) (b)

Figure 21: The restored images for FBFS (PSNR:32.2566, SSIM:0.9149) and zoom in the figures.

(a) (b)

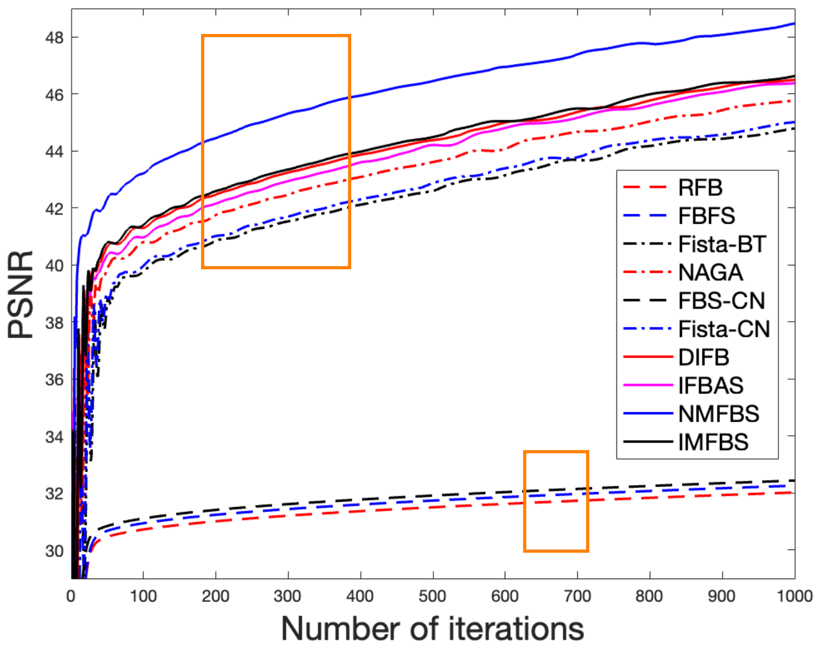
Figure 22: The restored images for FISTA-CN (PSNR:45.0092, SSIM: 0.9880) and zoom in the figures.

(a) (b)

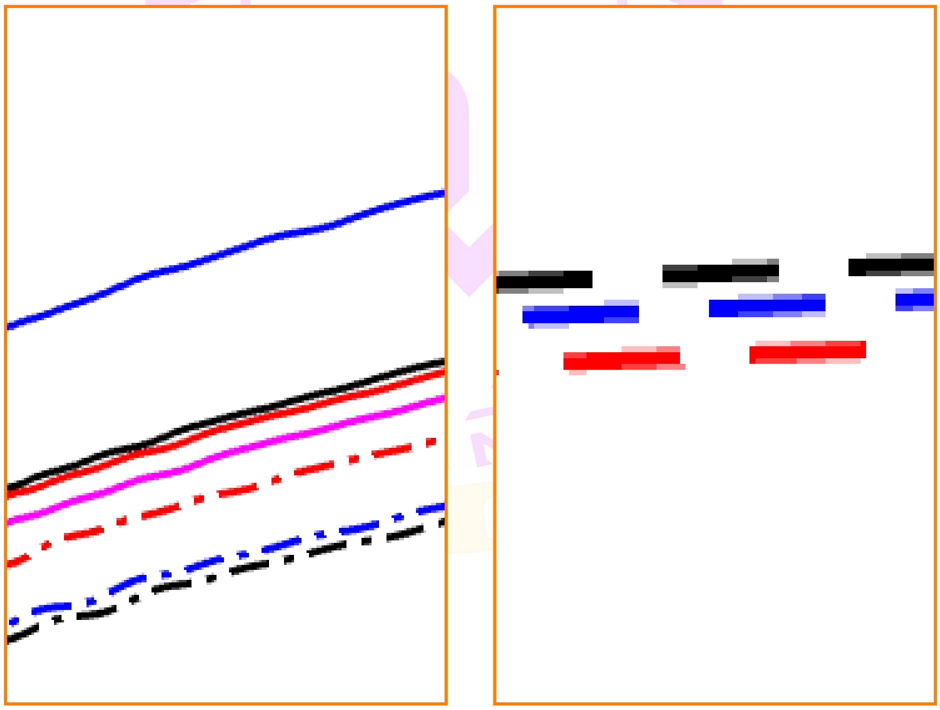
Figure 23: The restored images for NMFBS (PSNR:48.4738, SSIM: 0.9944) and zoom in the figures.

(a) (b)

Figure 24: The restored images for IMFBS (PSNR:46.6290, SSIM:0.9915) and zoom in the figures.

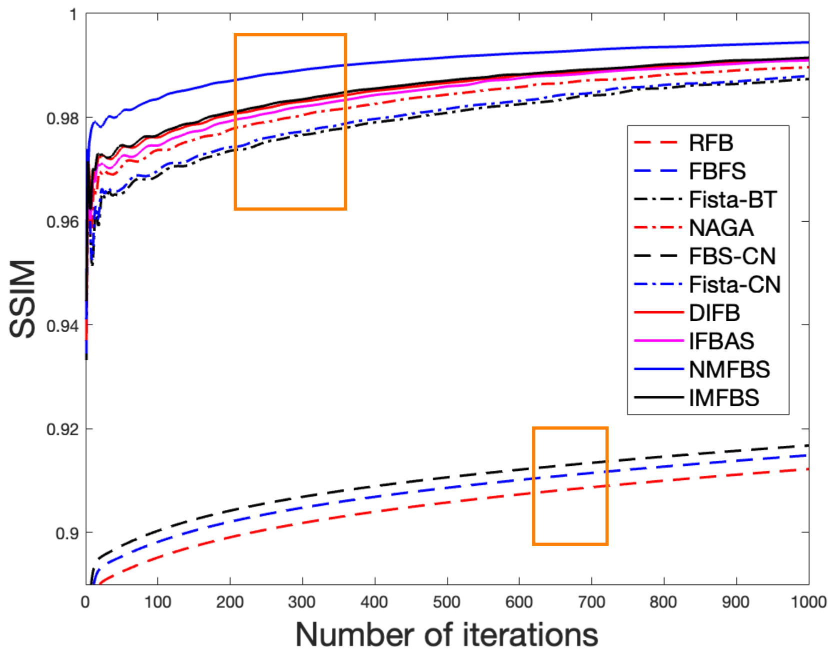


(a)

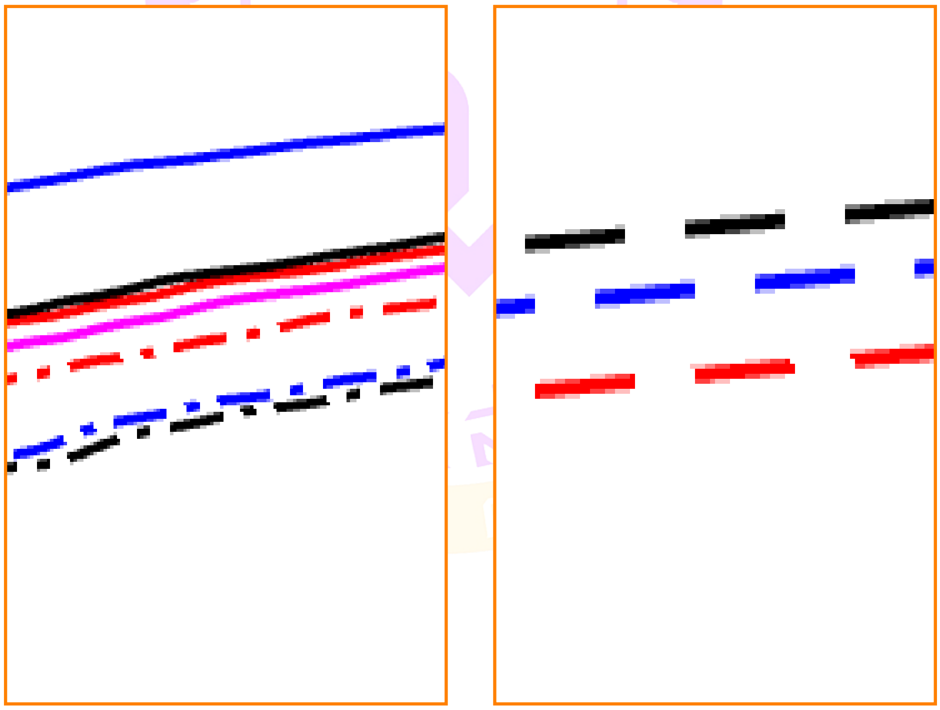


(b)

Figure 25: Graphs of PSNR plotting for motion blurs (BM2-3) and methods.



(a)



(b)

Figure 26: Graphs of SSIM plotting for motion blurs (BM2-3) and methods.

4.6.2 Image inpainting

In this section, we apply the constrained convex minimization problem (2.1.6) to image inpainting problem. We compare efficiency of our algorithms (Algorithm 4.2.1 denoted by AIPFB and Algorithm 4.1.3 denoted by DIPFB) with Algorithm 3.3 in an inertial three-operator splitting algorithm (ITOS) proposed by Cui, et al. [8], Algorithm 2.1.10 (FBS-CN) and Algorithm 2.1.10 (FISTA-CN).

Now, we aim to apply the result for solving an image inpainting problem which is of the following mathematical model:

$$\min_{x\in\mathbb{R}^{m\times n}} \frac{1}{2}\|A(x-x^0)\|_F^2 + \mu\|x\|_* \tag{4.6.6}$$

where $x^0 \in \mathbb{R}^{m\times n}(m<n)$ is a matrix with entries that lie in the interval $[l,u]$, A is a linear map that selects a subset of the entries of an $m\times n$ matrix by setting each unknown entry in the matrix to 0, x is matrix of known entries $A(x^0)$, and $\mu>0$ is a regularization parameter.

In particular, we consider the following image inpainting problem [12, 8]:

$$\min_{x} \frac{1}{2}\|P_\Omega(x) - P_\Omega(x^0)\|_F^2 + \mu\|x\|_* \tag{4.6.7}$$

where $\|\cdot\|_F$ is the Frobenius norm, and $\|\cdot\|_*$ is the nuclear norm. Here we define P_Ω by

$$P_\Omega(x) = \begin{cases} x^{ij}, & (i,j)\in\Omega, \\ 0, & \text{otherwise.} \end{cases} \tag{4.6.8}$$

The nuclear norm has been widely used in image inpainting and matrix completion problem, which is a convex relaxation of low rank constraint. It is obvious that the optimization problem (4.6.6) relates to (2.1.6). In fact, let $f(x) = \frac{1}{2}\|P_\Omega(x) - P_\Omega(x^0)\|_F^2$ and $g(x) = \mu\|x\|_*$. Then $\nabla f(x) = P_\Omega(x) - P_\Omega(x^0)$ with 1-Lipschitz continuous. The proximity operator of $g(x)$ can be computed by the singular value decomposition (SVD) [7].

We investigate throughout this section under the following setting Table 3. The initial point $x^0 = x^1$ are vectors of ones with the size of original images for all algorithms. Let θ_n and α_n be defined by

$$\theta_n = \begin{cases} \dfrac{t_n - 1}{t_{n+1}} \text{ where } t_{n+1} = \dfrac{1 + \sqrt{1 + 4t_n^2}}{2}, & \text{if } 1 \leq n \leq m; \\ \dfrac{1}{n^2}, & \text{otherwise.} \end{cases}$$

Algorithms		Parameters				
Comparision	OUR	$t_1 = 1$	γ $\in (0, 2/L)$	$\sigma > 0$	θ $\in (0,1)$	δ $\in (0, 1/2)$
ITOS		1	$1/\|A\|^2$			
FBS-CN				0.98	0.93	0.45
FISTA-CN		1		0.98	0.93	0.45
	AIPFB	1		0.98	0.93	0.45
	DIPFB	1		0.98	0.93	0.45

Table 3: Chosen parameters of each algorithm.

Algorithms	M=100		M=180		
	Fig(C) with MM1-1		Fig(C) with MM1-2		
	PSNR	SSIM	PSNR	SSIM	CPU
ITOS	23.4189	0.8771	23.4056	0.8769	46.4300
FBS-CN	24.3867	0.8966	24.4614	0.8984	497.6008
FISTA-CN	24.4621	0.8984	24.4621	0.8984	500.2202
AIPFB	24.4563	0.8982	24.4624	0.8984	508.7895
DIPFB	24.4625	0.8984	24.4624	0.8984	84.1010
	Fig(C) with MM1-2		Fig(C) with MM1-2		
	PSNR	SSIM	PSNR	SSIM	CPU
ITOS	22.3288	0.8464	22.3254	0.8464	46.7041
FBS-CN	22.6511	0.8473	23.0530	0.8639	514.9624
FISTA-CN	23.0600	0.8641	23.0600	0.8641	519.2283
AIPFB	23.0220	0.8628	23.0599	0.8641	524.2277
DIPFB	23.0603	0.8641	23.0603	0.8641	84.7715
	Fig(C) with MM1-3		Fig(C) with MM1-3		
	PSNR	SSIM	PSNR	SSIM	CPU
ITOS	19.5402	0.7532	19.8067	0.7653	46.6605
FBS-CN	14.8163	0.4075	20.0113	0.7659	510.8733
FISTA-CN	20.2495	0.7758	20.2491	0.7758	514.8055
AIPFB	19.1909	0.7183	20.2286	0.7750	520.3643
DIPFB	20.2498	0.7758	20.2493	0.7758	85.0540
	M=35		M=50		
	Fig(D) with MM2		Fig(D) with MM2		
	PSNR	SSIM	PSNR	SSIM	CPU
ITOS	25.2340	0.9471	25.5528	0.9472	5.1685
FBS-CN	25.1532	0.9672	27.2289	0.9715	52.3765
FISTA-CN	28.1513	0.9729	28.4510	0.9733	52.9773
AIPFB	26.9692	0.9711	28.1094	0.9729	54.3974
DIPFB	28.4619	0.9733	28.4612	0.9733	10.1617
	M=30		M=50		
	Fig(E) with MM3		Fig(E) with MM3		
	PSNR	SSIM	PSNR	SSIM	CPU
ITOS	21.6963	0.9636	22.4804	0.9671	2.3047
FBS-CN	21.3336	0.9678	23.5173	0.9776	21.6793
FISTA-CN	23.8421	0.9788	25.0734	0.9824	20.9304
AIPFB	22.6265	0.9741	24.3814	0.9805	21.3550
DIPFB	25.1541	0.9826	25.1911	0.9827	3.8593

Table 4: The results of deblurred images for each algorithm.

Figure 27: The original image of size 718 × 717 (Fig(C)).

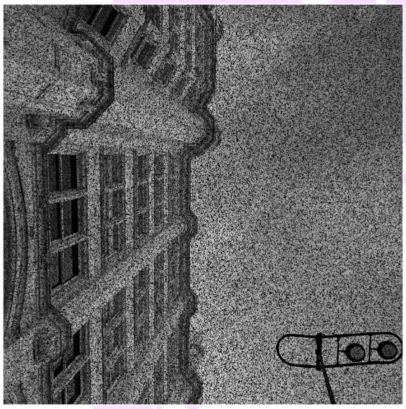


Figure 28: The 20% missing image (MM1-1).

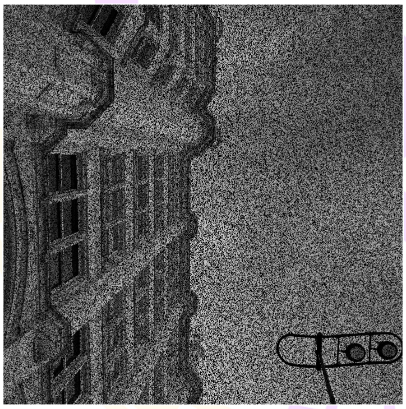


Figure 29: The 40% missing image (MM1-2).

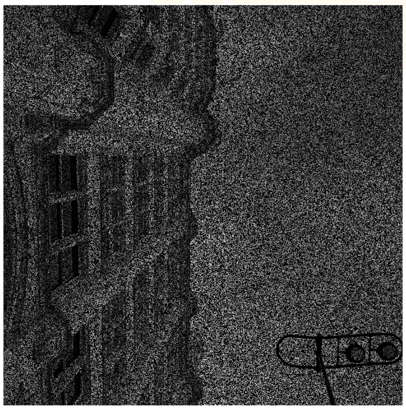
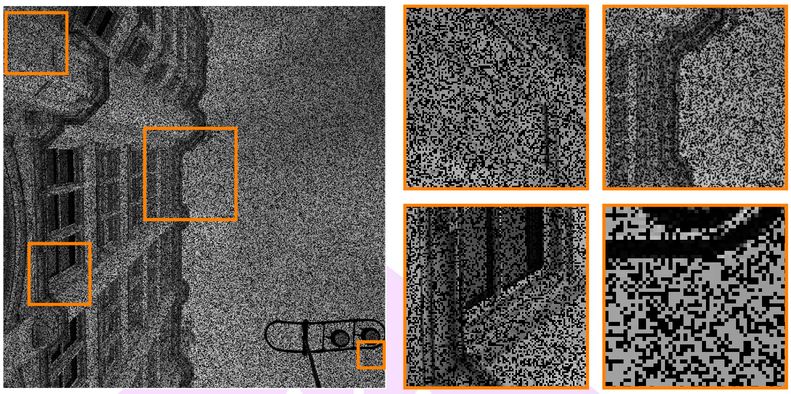


Figure 30: The 60% missing image (MM1-3).



(a) (b)

Figure 31: The 40% missing image (MM1-2).

(a) (b)

Figure 32: The restored image for ITOS (PSNR:22.3254, SSIM:0.8464).

(a) (b)

Figure 33: The restored image for FBS-CN (PSNR:23.0530, SSIM:0.8639).

(a) (b)

Figure 34: The restored image for FISTA-CN (PSNR:23.0600, SSIM:0.8641).

(a) (b)

Figure 35: The restored image for AIPFB (PSNR:23.0599, SSIM:0.8641).

(a) (b)

Figure 36: The restored image for DIPFB (PSNR:23.0603, SSIM:0.8641).

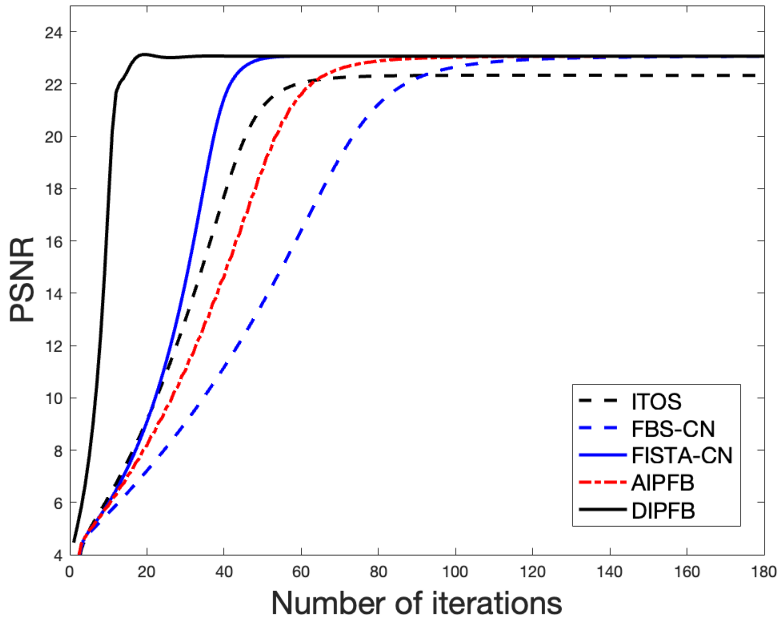


Figure 37: The graphs of PSNR plotting for 40% missing and methods.

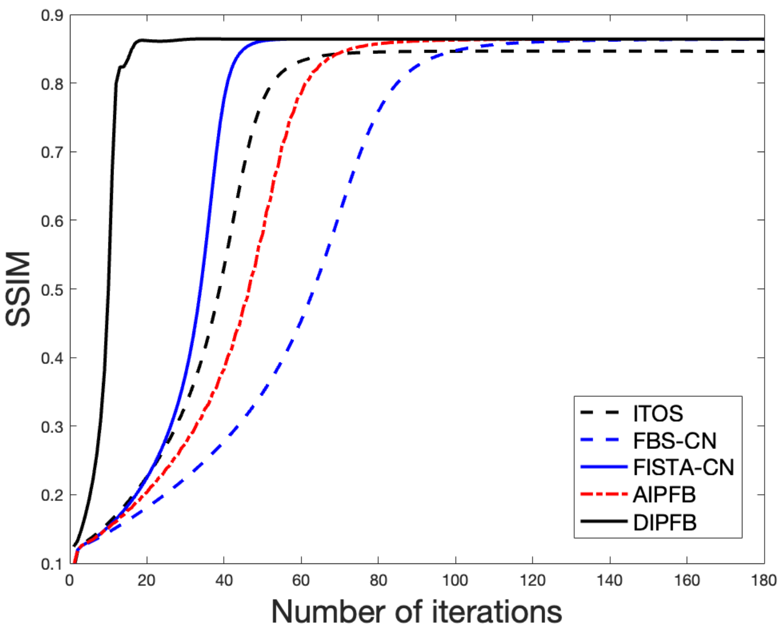
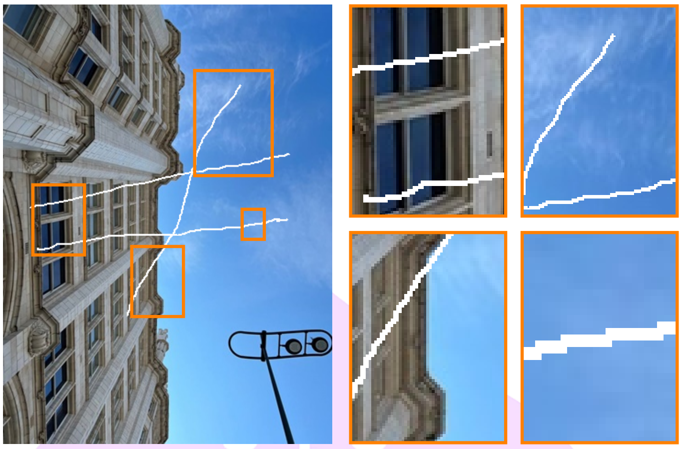


Figure 38: The graphs of SSIM plotting for 40% missing and methods.

Figure 39: The original image of size $359 \times 479 \times 3$ (Fig(D)).

Figure 40: The missing image (MM2).



(a) (b)

Figure 41: The missing image (MM2).

(a) (b)

Figure 42: The restored image for ITOS (PSNR:25.5528, SSIM:0.9472).

(a) (b)

Figure 43: The restored image for FBS-CN (PSNR:27.2289, SSIM:0.9715).

(a) (b)

Figure 44: The restored image for FISTA-CN (PSNR:28.4510, SSIM:0.9733).

(a) (b)

Figure 45: The restored image for AIPFB (PSNR:28.1094, SSIM:0.9729).

(a) (b)

Figure 46: The restored image for AIPFB (PSNR:28.4612, SSIM:0.9733).

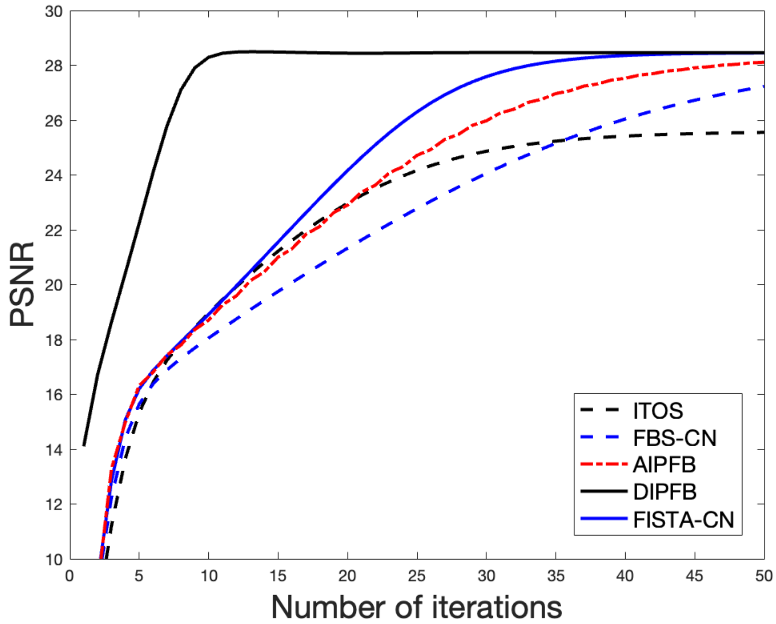


Figure 47: The graphs of PSNR plotting for missing (MM2) and methods.

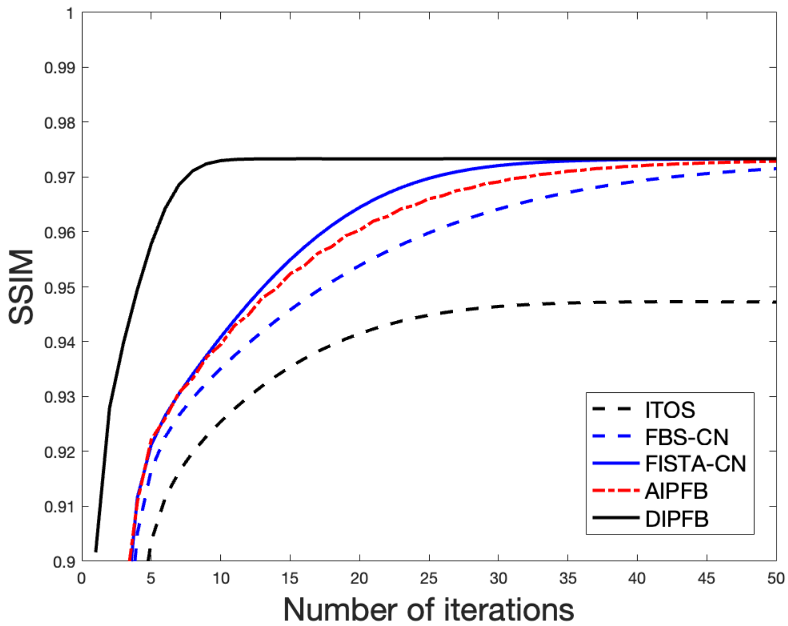


Figure 48: The graphs of SSIM plotting for missing (MM2) and methods.

Figure 49: The original image of size $380 \times 285 \times 3$ (Fig(E)).

Figure 50: The missing image (MM3).

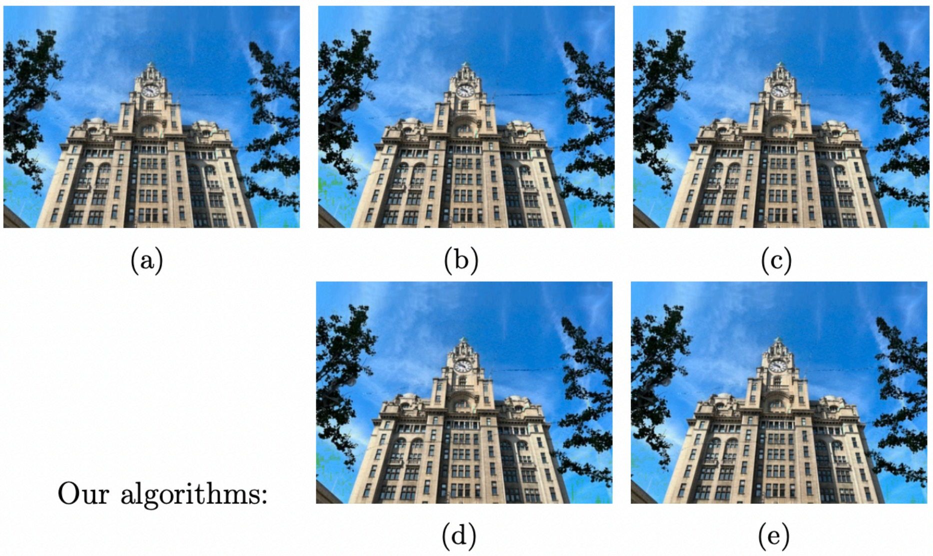


Figure 51: The restored images for (a) ITOS (PSNR:22.4804, SSIM:0.9671) (b) FBS-CN (PSNR:23.5173, SSIM:0.9776) (c) FISTA-CN (PSNR:25.0734, SSIM:0.9824) (d) AIPFB (PSNR:24.3814, SSIM:0.9805) and (e) DIPFB (PSNR:25.1911, SSIM:0.9827), respectively.

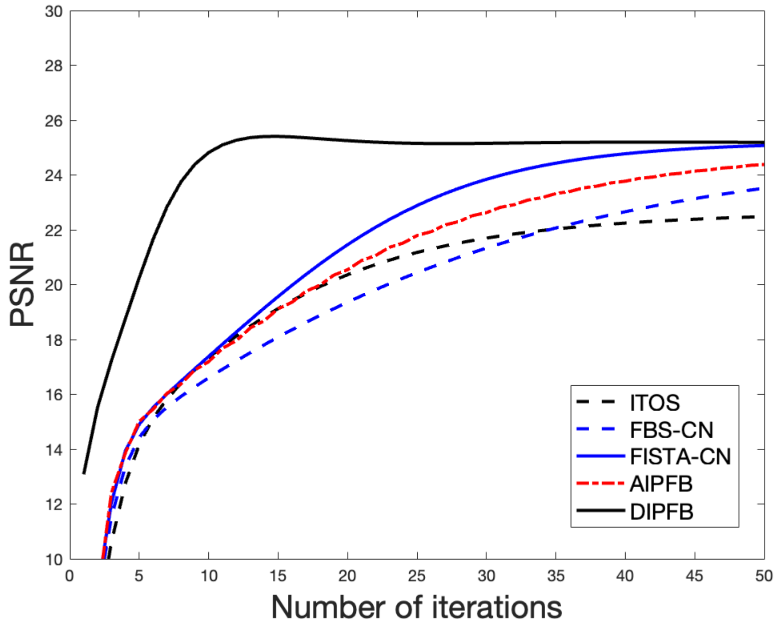


Figure 52: The graphs of PSNR plotting for missing (MM3) and methods.

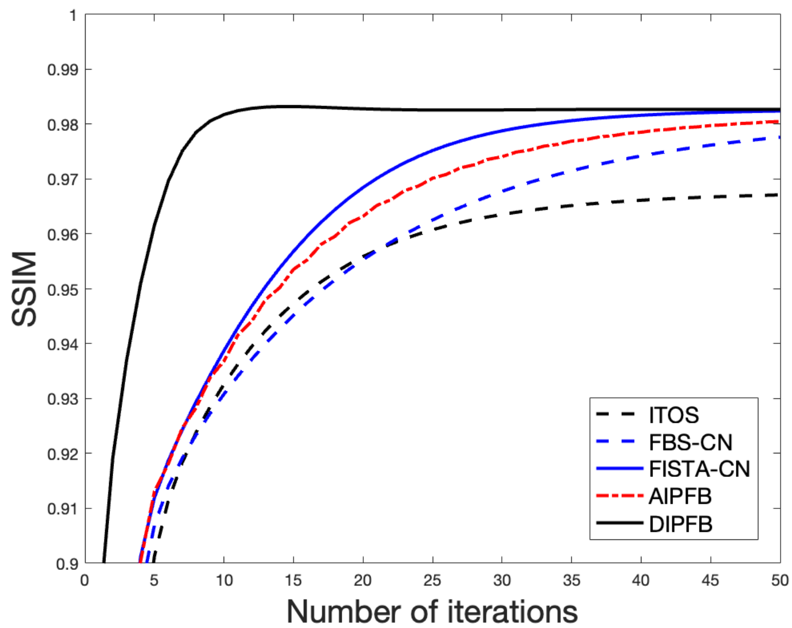


Figure 53: The graphs of SSIM plotting for missing (MM3) and methods.

CHAPTER 5

CONCLUSIONS

In this thesis, we have discussed the convex minimization problem of the sum of two convex function in Hilbert space. First, we introduce a double inertial forward-backward algorithms with Lipschitz continuity condition on gradient of function. We prove that the sequence generated by our proposed algorithms converges weakly to a minimizer of the sum those two convex functions under some mild conditions.

(1) Let H be a Hilbert space and let (x^n) be generated by

Initialization: Given α_n, θ_n and η_n be real positive sequences.

Iterative step: Let $x^0, x^1 \in H$ be arbitrarily and calculate x^{n+1} as follows:

Step1. Compute the inertial step:

$$w^n = x^n + \theta_n(x^n - x^{n-1}).$$

Step 2. Calculate:

$$v^n = \text{prox}_{\alpha_n g}(w^n - \alpha_n \nabla f(w^n))$$

and

$$y^n = v^n + \eta_n(v^n - x^{n-1}).$$

Step 3. Calculate the next iteration via:

$$x^{n+1} = \text{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)).$$

Set $n := n+1$ and return to **Step1.** Assume that $0 < \liminf_{n\to\infty} \alpha_n \leq \limsup_{n\to\infty} \alpha_n < \frac{2}{L}$, $\sum_{n=1}^{\infty} \theta_n < +\infty$ and $\sum_{n=1}^{\infty} \eta_n < +\infty$. Then we have

1. for each $x_* \in \operatorname{argmin}(f+g)$, $\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1 + 2(\theta_j + \eta_j + \theta_j\eta_j))$ where $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$.

2. (x^n) weakly converges to an element of $\operatorname{argmin}(f+g)$.

(2) Let H be a Hilbert space and let (x^n) be generated by

Initialization: Given α_n, θ_n and η_n be real positive sequences.

Iterative step: Let $x^0, x^1 \in H$ be arbitrarily and calculate x^{n+1} as follows:

Step1. Compute the inertial step:

$$w^n = x^n + \theta_n(x^n - x^{n-1}).$$

Step 2. Calculate:

$$v^n = \operatorname{prox}_{\alpha_n g}(w^n - \alpha_n \nabla f(w^n))$$

and

$$y^n = v^n + \eta_n(v^n - x^{n-1}).$$

Step 3. Calculate:

$$z^n = \operatorname{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)).$$

Step 4. Calculate the next iteration via:

$$x^{n+1} = P_\Omega(z^n).$$

Set $n := n+1$ and return to **Step1.** Assume that $0 < \liminf_{n\to\infty} \alpha_n \leq \limsup_{n\to\infty} \alpha_n < \frac{2}{L}$, $\theta_n \geq 0$, $\eta_n \geq 0$, $\sum_{n=1}^{\infty} \theta_n < +\infty$ and $\sum_{n=1}^{\infty} \eta_n < +\infty$. Then we have

1. for each $x_* \in \Omega \cap \operatorname{argmin}(f+g)$, $\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1+2(\theta_j + \eta_j + \theta_j\eta_j))$ where $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$.

2. (x^n) weakly converges to an element of $\Omega \cap \operatorname{argmin}(f+g)$.

Next, to challenge of removing the Lipschitz continuity assumption on the gradient of the function attracts us to study the concept of the linesearch method, we introduce an inertial forward-backward algorithms which the step-size dose not depend on any Lipschitz constant.

(3) Let H be a Hilbert space and let (x^n) be generated by
Initialization: Given $\sigma > 0, \delta \in (0, \frac{1}{2}),\ \theta \in (0,1)$ and $\theta_n \geq 0$.
Iterative step: Let $x^0, x^1 \in H$ be arbitrarily and calculate x^{n+1} as follows:
Step1. Compute the inertial step:

$$z^n = \begin{cases} x^n + \theta_n(x^n - x^{n-1}), & \text{if } n \text{ is odd} \\ x^n, & \text{if } n \text{ is even.} \end{cases}$$

Step 2. Calculate:

$$x^{n+1} = P_\Omega(\operatorname{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n)))$$

where $\alpha_n = \sigma\theta^{m_n}$ and m_n is the smallest nonnegative integer such that

$$\begin{aligned} \alpha_n \|\nabla f(\operatorname{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n))) - \nabla f(z^n)\| \\ \leq \delta \|\operatorname{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n)) - z^n\|. \end{aligned}$$

Set $n := n+1$ and return to **Step1.** Assume that $\theta_n \geq 0$ and $\sum_{n=1}^{\infty} \theta_n < +\infty$ and that $\alpha_n \geq \alpha > 0$ for some α. Then we have

1. for each $x_* \in \operatorname{argmin}(f+g)$, $\|x^{2n+2} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1+2\theta_{2j+1})$ where $K = \max\{\|x^1 - x_*\|, \|x^3 - x_*\|\}$.

2. (x^n) weakly converges to an element of $\operatorname{argmin}(f+g) \cap \Omega$.

(4) Let H be a Hilbert space and let (x^n) be generated by

Initialization: Given $\delta \in (0,1)$, $\alpha_n \geq 0$ and $\theta_n \geq 0$.

Iterative step: Let $x^0, x^1 \in H$ be arbitrarily and calculate x^{n+1} as follows:

Step1. Compute the inertial step:

$$z^n = x^n + \theta_n(x^n - x^{n-1})$$

Step 2. Calculate:

$$x^{n+1} = \operatorname{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n))$$

where

$$\alpha_{n+1} = \begin{cases} \min\{\frac{\delta\|z^n - x^{n+1}\|}{\|\nabla f(z^n) - \nabla f(x^{n+1})\|}, \alpha_n\} & \text{if } \nabla f(z^n) - \nabla f(x^{n+1}) \neq 0 \\ \alpha_n & \text{otherwise.} \end{cases}$$

Set $n := n+1$ and return to **Step1.** Assume that $\theta_n \geq 0$ and $\sum_{n=1}^{\infty} \theta_n < +\infty$. Then, we have

1. for each $x_* \in \operatorname{argmin}(f+g)$, $\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1+2\theta_j)$ where $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$.

2. The sequence (x^n) weakly converges to a point in $\operatorname{argmin}(f+g)$.

(5) Let H be a Hilbert space and let (x^n) be generated by

Initialization: Given $\sigma > 0$, $\theta > 0$, $\delta \in (0,1)$ and $\theta_1 > 0$.

Iterative step: Let $x^0 = x^1 \in H$ and calculate x^{n+1} as follows:

Step 1. Compute the inertial step:

$$z^n = x^n + \theta_n(x^n - x^{n-1}).$$

Step 2. Compute:

$$y^n = \mathrm{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n)) + \alpha_n(\nabla f(z^n) - \nabla f(\mathrm{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n))))$$

where $\alpha_n = \sigma\theta^{m_n}$ and m_n is the smallest number such that:

$$\begin{aligned} & \alpha_n^2(\|\nabla f(z^n) - \nabla f(\mathrm{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n)))\|^2 \\ & +\|\nabla f(y^n) - \nabla f(\mathrm{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)))\|^2) \\ \leq \; & \delta^2(\|z^n - \mathrm{prox}_{\alpha_n g}(z^n - \alpha_n \nabla f(z^n))\|^2 \\ & +\|y^n - \mathrm{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n))\|^2). \end{aligned}$$

Step 3. Compute the x^{n+1} step:

$$x^{n+1} = \mathrm{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)) + \alpha_n(\nabla f(y^n) - \nabla f(\mathrm{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)))).$$

Set $n := n + 1$ and return to **Step1.** Suppose that $\alpha_n \geq \alpha$ for some $\alpha > 0$, $\theta_n \geq 0$ and $\sum_{n=1}^{\infty} \theta_n < +\infty$. Then, we have for each $x_* \in \mathrm{argmin}(f + g) \cap \Omega$, $\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1 + 2\theta_j)$ where $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$ and the sequence (x^n) weakly converges to an element of $\mathrm{argmin}(f + g)$.

(6) Let H be a Hilbert space and let (x^n) be generated by

Initialization: Given $\sigma, \theta, \mu > 0$, $\delta, \rho \in (0, 1)$ and $\theta_1 > 0$.

Iterative step: Let $x^0 = x^1 \in H$, calculate x^{n+1} and μ_{n+1} as follows:

Step1. Compute the inertial step:

$$v^n = x^n + \theta_n(x^n - x^{n-1}).$$

Step2. Compute the forward-backward step:

$$y^n = \text{prox}_{\alpha_n g}(v^n - \alpha_n \nabla f(v^n))$$

where $\alpha_n = \sigma\theta^{m_n}$ and m_n is the smallest nonnegative number such that:

$$\alpha_n \|\nabla f(y^n) - \nabla f(v^n)\| \leq \delta \|y^n - v^n\|.$$

Step3. Compute the forward-backward step:

$$z^n = \text{prox}_{\mu_n g}(y^n - \mu_n \nabla f(y^n)).$$

Step4. Compute the x^{n+1} step:

$$x^{n+1} = z^n + \mu_n(\nabla f(y^n) - \nabla f(z^n))$$

and update

$$\mu_{n+1} = \begin{cases} \min\left(\frac{\rho\|y^n - z^n\|}{\|\nabla f(y^n) - \nabla f(z^n)\|}, \mu_n\right) & if\ \|\nabla f(y^n) - \nabla f(z^n)\| \neq 0; \\ \mu_n & otherwise. \end{cases}$$

Set $n := n + 1$ and return to **Step1.** Suppose that $\alpha_n \geq \alpha$ for some $\alpha > 0$, $\theta_n \geq 0$ and $\sum_{n=1}^{\infty} \theta_n < +\infty$. Then, we have for each $x_* \in \text{argmin}(f + g) \cap \Omega$, $\|x^{n+1} - x_*\| \leq K \cdot \prod_{j=1}^{n}(1 + 2\theta_j)$ where $K = \max\{\|x^1 - x_*\|, \|x^2 - x_*\|\}$ and the sequence (x^n) weakly converges to a point in $\text{argmin}(f + g)$.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Agarwal, R. P., O'Regan, D. and Sahu, D. R. (2009). Fixed Point Theory for Lipschitzian-type Mappings with Applications. New York: **Springer**.

[2] Bauschke, H. H. and Combettes, P. L. (2011). Convex Analysis and Monotone Operator Theory in Hilbert Spaces. New York: **Springer**.

[3] Beck, A., and Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM journal on imaging sciences, 2(1), 183-202.

[4] Bello Cruz, J. Y. and Nghia, T. T. (2016). On the convergence of the forward-backward splitting method with linesearches. **Optimization Methods and Software**, 31(6), 1209 - 1238.

[5] Bertsekas, D. P. (1976). On the Goldstein-Levitin-Polyak gradient projection method. **IEEE Transactions on Automatic Control**, 21(2), 174-184.

[6] Burachik, R. S., and Iusem, A. N. (2008). Enlargements of monotone operators. In Set-Valued Mappings and Enlargements of Monotone Operators (pp. 161-220). **Springer**, Boston, MA.

[7] Cai, J. F., Cands, E. J., and Shen, Z. (2010). A singular value thresholding algorithm for matrix completion. **SIAM Journal on Optimization,** 20(4), 1956-1982.

[8] Cui, F., Tang, Y., & Yang, Y. (2019). An inertial three-operator splitting algorithm with applications to image inpainting. arXiv preprint arXiv:1904.11684.

[9] Cegielski, A. (2012). Iterative methods for fixed point problems in Hilbert spaces (Vol. 2057). **Springer**.

[10] Combettes, P. L., and Wajs, V. R. (2005). Signal recovery by proximal forward-backward splitting. **Multiscale Modeling & Simulation**, 4(4), 1168-1200.

[11] Combettes, P. L., and Pesquet, J. C. (2011). Proximal splitting methods in signal processing. In Fixed-point algorithms for inverse problems in science and engineering (pp. 185-212). **Springer**, New York, USA.

[12] Davis, D., and Yin, W. (2017). A three-operator splitting scheme and its optimization applications. **Set-valued and Variational Analysis,** 25(4), 829-858.

[13] Douglas, J., and Rachford, H. H. (1956). On the numerical solution of heat conduction problems in two and three space variables. **Transactions of the American Mathematical Society**, 82(2), 421-439.

[14] Duan, P., Zhang, Y., and Bu, Q. (2020). New inertial proximal gradient methods for unconstrained convex optimization problems. **Journal of Inequalities and Applications**, 2020(1), 1-18.

[15] Dunn, J. C. (1976). Convexity, monotonicity, and gradient processes in Hilbert space. **Journal of Mathematical Analysis and Applications**, 53(1), 145-158.

[16] Goldstein, A. A. (1964). Convex programming in Hilbert space. **Bulletin of the American Mathematical Society**, 70(5), 709-710.

[17] Hanjing, A., and Suantai, S. (2020). A fast image restoration algorithm based on a fixed point and optimization method. **Mathematics**, 8(3), 378.

[18] Jin, W., Censor, Y., and Jiang, M. (2016). Bounded perturbation resilience of projected scaled gradient methods. **Computational Optimization and Applications**, 63(2), 365-392.

[19] Kankam, K., Pholasa, N., and Cholamjiak, P. (2019). On convergence and complexity of the modified forward-backward method involving new

linesearches for convex minimization. **Mathematical Methods in the Applied Sciences**, 42(5), 1352 - 1362.

[20] Levitin, E. S., and Polyak, B. T. (1966). Constrained minimization methods. **USSR Computational Mathematics and Mathematical Physics**, 6(5), 1-50.

[21] Lions, P. L., and Mercier, B. (1979). Splitting algorithms for the sum of two nonlinear operators. **SIAM Journal on Numerical Analysis**, 16(6), 964-979.

[22] Lorenz, D. A., and Pock, T. (2015). An inertial forward-backward algorithm for monotone inclusions. Journal of Mathematical Imaging and Vision, 51(2), 311-325.

[23] Martinet, B. (1970). Regularisation, d'inéquations variationelles par approximations succesives. **Revue Francaise d'informatique et de Recherche Operationelle**.

[24] Minty, G. J. (1962). Monotone (nonlinear) operators in Hilbert space. **Duke Mathematical Journal**, 29(3), 341-346.

[25] Moreau, J. J. (1965). Proximité et dualit dans un espace hilbertien. **Bulletin de la Société mathématique de France**, 93, 273-299.

[26] Moudafi, A., and Oliny, M. (2003). Convergence of a splitting inertial proximal method for monotone operators. **Journal of Computational and Applied Mathematics**, 155(2), 447-454.

[27] Mu, Z., and Peng, Y. (2015). A note on the inertial proximal point method. Statistics, **Optimization & Information Computing**, 3(3), 241-248.

[28] Nesterov, Y. E. (1983). A method for solving the convex programming problem with convergence rate $O(1/k^2)$. **In Dokl. akad. nauk Sssr** (Vol. 269, pp. 543-547).

[29] Opial, Z. (1967). Weak convergence of the sequence of successive approximations for nonexpansive mappings. **Bulletin of the American Mathematical Society**, 73(4), 591 - 597.

[30] Osilike, M. O., and Aniagbosor, S. C. (2000). Weak and strong convergence theorems for fixed points of asymptotically nonexpensive mappings. **Mathematical and Computer Modelling**, 32(10), 1181 - 1191.

[31] Parikh, N. and Boyd, S. (2014). Proximal algorithms. Foundations and trends in Optimization, 1(3), 127-239.

[32] Polyak, B. T. (1963). Gradient methods for the minimisation of functionals. **USSR Computational Mathematics and Mathematical Physics**, 3(4), 864-878.

[33] Polyak, B. T. (1964). Some methods of speeding up the convergence of iteration methods. **Ussr Computational Mathematics and Mathematical Physics**, 4(5), 1-17.

[34] Rockafellar, R. T. (1976). Monotone operators and the proximal point algorithm. **SIAM Journal on Control and Optimization**, 14(5), 877-898.

[35] Takahashi, W. (2009). Nonlinear and Convex Analysis. Japan: **Yokohama Publishers, Inc.**.

[36] Tan, K. K., and Xu, H. K. (1993). Approximating fixed points of nonexpansive mappings by the Ishikawa iteration process. **Journal of Mathematical Analysis and Applications**, 178, 301 - 301.

[37] Tseng, P. (2000). A modified forward-backward splitting method for maximal monotone mappings. **SIAM Journal on Control and Optimization**, 38(2), 431-446.

[38] Van Hieu, D., and Anh, P. K. (2021). Modified forward-backward splitting method for variational inclusions. **4OR**, 19(1), 127 - 151.

[39] Verma, M., and Shukla, K. K. (2017). A new accelerated proximal gradient technique for regularized multitask learning framework. Pattern Recognition Letters, 95, 98-103.

[40] Hore, A., and Ziou, D. (2010, August). Image quality metrics: PSNR vs. SSIM. In 2010 20th international conference on pattern recognition (pp. 2366-2369). **IEEE**.

[41] Tikhonov, A. N. and Arsenin, V.Y (1977). Solution of Ill-Posed Problems; V.H. Winston: Washington, DC, USA: J. **Wiley**.

BIOGRAPHY

BIOGRAPHY

Name Surname	Kunrada Kankam
Date of Birth	September 11, 1995
Place of Birth	Phrae Province, Thailand
Address	228 Moo 7, Mae Chua Sub - district, Den Chai District, Phrae Province, Thailand 54110

Education Background

2020	M.Sc. (Mathematics), University of Phayao, Phayao, Thailand
2017	B.Sc. (Mathematics), University of Phayao, Phayao, Thailand

Publications

Articles

1. **Kankam, K.**, Pholasa, N., and Cholamjiak, P. (2019). On convergence and complexity of the modified forward-backward method involving new linesearches for convex minimization. Mathematical Methods in the Applied Sciences, 42(5), 1352 - 1362.

2. **Kankam, K.**, Pholasa, N., and Cholamjiak, P. (2019). Hybrid forward-backward algorithms using linesearch rule for minimization problem. Thai Journal of Mathematics, 17(3), 607 - 625.

3. Cholamjiak, P., **Kankam, K.**, Srinet, P., and Pholasa, N. (2019). A double forward-backward algorithm using linesearches for minimization problem. Thai Journal of Mathematics, 18(1), 63-76.

4. Suantai, S., **Kankam, K.**, and Cholamjiak, P. (2020). A novel forward-backward algorithm for solving convex minimization problem in Hilbert spaces. Mathematics, 8(1), 42.

5. **Kankam, K.**, Pheetarakorn, S., and Cholamjiak, W. (2020). The modified SP and Noor iterations with shrinking projection methods for three $G-$nonexpansive mappings in Hilbert spaces with graphs. Thai Journal of Mathematics, 18(3), 1285-1297.

6. **Kankam, K.**, Srinak, P., Cholamjiak, P., and Pholasa, N. (2020). Solving the multiple-set split feasibility problem and the equilibrium problem by a new relaxed CQ algorithm. Rendiconti del Circolo Matematico di Palermo Series 2, 69(3), 1131-1148.

7. Suantai, S., **Kankam, K.**, and Cholamjiak, P. (2021). A projected forward-backward algorithm for constrained minimization with applications to image inpainting. Mathematics, 9(8), 890.

8. Suantai, S., **Kankam, K.**, Cholamjiak, P., and Cholamjiak, W. (2021). A parallel monotone hybrid algorithm for a finite family of $G-$nonexpansive mappings in Hilbert spaces endowed with a graph applicable in signal recovery. Computational and Applied Mathematics, 40(4), 1-17.

9. **Kankam, K.**, Wongjak, W., Kotcharerg, J., and Cholamjiak, W. (2021). Solving common nonmonotone equilibrium problems using a parallel inertial type algorithm with armijo linesearch. Thai Journal of Mathematics, 19(3), 1096-1111.

10. **Kankam, K.**, and Cholamjiak, P. (2021). Strong convergence of the forward-backward splitting algorithms via linesearches in

Hilbert spaces. Applicable Analysis, 1-20.

11. Pholasa, N., **Kankam, K.**, and Cholamjiak, P. (2021). Solving the split feasibility problem and the fixed point problem of left Bregman firmly nonexpansive mappings via the dynamical step sizes in Banach spaces. Vietnam Journal of Mathematics, 49(4), 1011-1026.

12. Suantai, S., Noor, M. A., **Kankam, K.**, and Cholamjiak, P. (2021). Novel forward-backward algorithms for optimization and applications to compressive sensing and image inpainting. Advances in Difference Equations, 2021(1), 1-22.

13. Suantai, S., **Kankam, K.**, Cholamjiak, W., and Yajai, W. (2022). Parallel hybrid algorithms for a finite family of G-nonexpansive mappings and its application in a novel signal recovery. Mathematics, 10(12), 2140.

14. **Kankam, K.**, Cholamjiak, P., and Cholamjiak, W. (2022). A modified parallel monotone hybrid algorithm for a finite family of G-nonexpansive mappings apply to a novel signal recovery. Results in Nonlinear Analysis, 5(3), 393-411.

15. Suantai, S., **Kankam, K.**, Yambangwai, D., and Cholamjiak, W. (2022). A modified inertial parallel viscosity-type algorithm for a finite family of nonexpansive mappings and its applications. Mathematics, 10(23), 4422.

16. **Kankam, K.**, Cholamjiak, W., and Cholamjiak, P. (2022). Convergence analysis of a modified forward-backward splitting algorithm for minimization and application to image recovery. Computational and Mathematical Methods, 2022.

17. **Kankam, K.**, Cholamjiak, W., and Cholamjiak, P. (2023). New inertial forward-backward algorithm for convex minimization with applications. Demonstratio Mathematica, 56(1), 20220188. Demonstratio mathematica, 56(1), 1-13.

18. **Kankam, K.**, and Cholamjiak, P. (2023). Double inertial proximal gradient algorithms for convex optimization problems and applications. Acta Mathematica Scientia, 43, 1462-1476.

Conference presentations

1. **Kankam, K.**, Srinak, P., Pholasa, N., Cholamjiak, P. (May 24 - 25, 2018). Solving the multiple-sets split feasibility problem and the equilibrium problem by a new relaxed CQ algorithm. In The 10^{th} National Science Research Conference, Mahasarakham University, Mahasarakham, Thailand.

2. **Kankam, K.**, Cholamjiak, P. (November 5 - 9, 2018). Solving the multiple-sets split feasibility problem and the equilibrium problem by a new relaxed CQ algorithm. In The 6^{th} Asian Conference on Nonlinear Analysis and Optimization, Okinawa, Japan.

3. **Kankam, K.**, Cholamjiak, P. (September 12 - 14, 2019). Strong convergence of the forward-backward splitting algorithms via linesearches. In International Workshop on Applied Nonlinear Analysis, Bangsaen, Chonburi, Thailand.

4. **Kankam, K.**, Cholamjiak, P. (April 28 - 30, 2021). Convergence theorem for convex minimization problem in Hilbert space. In the 24^{th} International Annual Symposium on Computational Science and Engineering, Kasetsart University (Kamphaeng Saen Campus), Nakhon Pathom, Thailand.

5. **Kankam, K.**, Cholamjiak, P. (August 20 - 21, 2021). A forward-backward algorithm for convex minimization problem with application. In International Workshop on Digital Image Processing and Machine Learning, Thailand.

6. **Kankam, K.**, Cholamjiak, P. (August 29 - September 1, 2022). Convergence theorem for optimization problem and application. In the 11^{th} International Eurasian Conference on Mathematical Sciences and Applications, Yildiz Technical University, Istanbul, Turkey.

7. **Kankam, K.**, Cholamjiak, P. (September 8 - 9, 2022). A forward-backward algorithm for optimization problem in Hilbert space. In The international conference on Digital Image Processing and Machine Learning 2022, University of Phayao, Phayao, Thailand.

8. **Kankam, K.**, Cholamjiak, P. (October 12 - 15, 2022). The convergence of proximal gradient algorithm for minimization problem with application. In International Workshop on Nonlinear Analysis and its Applications 2022, at Sakarya University of Applied Sciences and Sakarya University, Turkey.