SPLITTING METHOD FOR SOLVING CONVEX MINIMIZATION PROBLEMS AND APPLICATIONS

SUPARAT KESORNPROM

A Thesis Submitted to University of Phayao
in Partial Fulfillment of the Requirements
for the Doctor of Philosophy in Mathematics
October 2022

Thesis

Title

Splitting Method For Solving Convex Minimization Problems And Applications

Submitted by Suparat Kesornprom

Approved in partial fulfillment of the requirements for the

Doctor of Philosophy in Mathematics

University of Phayao

....Chairman

(Professor Dr.Suthep Suantai)

....Committee

(Professor Dr. Poom Kumam)

....Committee

(Associate Professor Dr.Prasit Cholamjiak)

....Committee

(Associate Professor Dr.Tanakit Thianwan)

....Committee

(Associate Professor Dr.Watcharaporn Cholamjiak)

....Committee

(Associate Professor Dr.Damrongsak Yambangwai)

Approved by

....

(Associate Professor Dr.Chayan Boonyarak)

Dean of School of Science

October 2022

ACKNOWLEDGEMENT

I would like to express my sincere thanks to my thesis advisors, Associate Professor Dr. Prasit Cholamjiak and Associate Professor Dr. Watcharaporn Cholamjiak for the continuous support of my Ph.D. study and related research, for their motivation and immense knowledge. I could not have imagined having a better advisors and mentors for my Ph.D. study.

I would like to take this opportunity to express my thankfulness to Assistant professor Dr. Yann Savoye for his kindness, helpful guidance and academic support when I was doing research at Liverpool John Moores University, Liverpool, UK.

My sincere gratitude also goes to my thesis committees, Professor Dr. Suthep Suantai, Professor Dr. Poom Kumam, Associate Professor Dr.Tanakit Thianwan, Associate Professor Dr.Damrongsak Yambangwai. It is my pleasure to express my thanks to all of them for their generous.

I appreciate mentioning that my graduate study was financially supported by International Research Network in Digital Image Processing and Machine Learning (IRN).

Finally, I most gratefully acknowledge my family and friends for all their support throughout the period of this thesis.

Suparat Kesornprom

เรื่อง: วิธีการแยกสำหรับการแก้ปัญหาค่าต่ำสุดเชิงคอนเวกซ์และการประยุกต์
ผู้วิจัย: ศุภารัตน์ เกษรพรม, วิทยานิพนธ์: ปร.ด. (คณิตศาสตร์), มหาวิทยาลัยพะเยา, 2565
ประธานที่ปรึกษา: รองศาสตราจารย์ ดร.ประสิทธิ์ ช่อลำเจียก
กรรมการที่ปรึกษา: รองศาสตราจารย์ ดร.ธนกฤต เทียนหวาน, รองศาสตราจารย์ ดร.วัชรภรณ์ ช่อลำเจียก
คำสำคัญ: ปัญหาค่าต่ำสุดเชิงคอนเวกซ์, วิธีการข้างหน้า–ข้างหลัง, กฎไลน์เสิร์ช, ขนาดขั้นแบบปรับตัว, วิธีการแบบเฉื่อย, การลู่เข้าแบบอ่อน, การกู้คืนภาพ, การซ่อมแซมภาพ, การจำแนกข้อมูล

บทคัดย่อ

ทฤษฎีการหาค่าเหมาะที่สุดมีความสำคัญอย่างมากในเชิงวิทยาศาสตร์ประยุกต์ เช่น กลศาสตร์ วิศวกรรมศาสตร์ เศรษฐศาสตร์ คอมพิวเตอร์ การจัดการข้อมูล และการแพทย์ ในปัจจุบันมีนักวิจัยมากมาย ได้ให้ความสนใจในการพัฒนาขั้นตอนวิธีให้มีประสิทธิภาพในการใช้งานจริง ปัญหาหลักในทฤษฎีการหาค่าเหมาะสมที่สุดคือปัญหาค่าต่ำสุดเชิงคอนเวกซ์ ซึ่งสามารถเป็นแบบจำลองเดียวกันกับปัญหาที่มีอยู่จริงมากมาย เช่น การประมวลผลสัญญาณ การประมวลผลภาพ การจัดจำแนกข้อมูล และการประยุกต์ในสาขาอื่นๆ

ในวิทยานิพนธ์นี้ถูกแบ่งออกเป็นสองส่วน ส่วนแรกของวิทยานิพนธ์คือ การปรับปรุงของขั้นตอนวิธีแบบแยกที่เรียกว่า ขั้นตอนวิธีข้างหน้า ข้างหลัง ซึ่งขั้นตอนวิธีนี้ถูกออกแบบโดยใช้เทคนิคแบบเฉื่อย และใช้ขนาดขั้นแบบใหม่ของไลน์เสิร์ช สำหรับแก้ปัญหาค่าต่ำสุดเชิงคอนเวกซ์ ในกรอบของปริภูมิฮิลเบิร์ต และยังได้ทฤษฎีการลู่เข้าของขั้นตอนวิธีที่นำเสนอซึ่งถูกพิสูจน์ภายใต้เงื่อนไขที่เหมาะสม ส่วนที่สอง คือการประยุกต์ของขั้นตอนวิธีที่นำเสนอไปยังการกู้คืนภาพ การซ่อมแซมภาพ และการจัดจำแนกข้อมูล ซึ่งการประยุกต์นี้ให้การสนับสนุนประสิทธิภาพของขั้นตอนวิธีที่พัฒนาขึ้นมา

Title: SPLITTING METHOD FOR SOLVING CONVEX MINIMIZATION PROBLEMS AND APPLICATIONS

Author: Suparat Kesornprom, Thesis: Ph.D. (Mathematics), University of Phayao, 2022

Advisor: Associate Professor Dr.Prasit Cholamjiak

Co-advisor: Associate Professor Dr.Tanakit Thianwan, Associate Professor Dr.Watcharaporn Cholamjiak

Keywords: convex minimization problem, forward-backward method, linesearch rule, adaptive stepsize, inertial method, weak convergence, image deblurring, image inpanting, data classification

ABSTRACT

Optimization theory is important in applied sciences such as mechanics, economics, computing, data management, and medicine. In the modern world, many researchers are interested in developing algorithms to solve the optimization problem by focusing on developing methods that have efficiency in practical applications. In optimization theory, a significant problem is the convex minimization problem, which many researchers are interested in studying and developing algorithms to solve this problem. It can be a unified model for many practical problems such as signal processing, image processing, data classification, and many other applied fields.

In this dissertation has separated by two parts. The dissertation's first purpose is to modify splitting methods, called the forward-backward method. The splitting methods are designed using inertial technique and new linesearch in stepsize. Moreover, the convergence of the proposed methods is proved under suitable conditions for solving the convex minimization problem in the framework of Hilbert spaces. The second purpose is the applications of the proposed methods to image deblurring, image inpainting, and date classification. The applications are given to support the efficiency of the proposed methods.

LIST OF CONTENTS

Chapter **Page**

LIST OF TABLES

Tables **Page**

LIST OF FIGURES

Figures **Page**

LIST OF FIGURES (CONT.)

Figures **Page**

CHAPTER I

INTRODUCTION

Optimization has been expanding in all directions at an astonishing rate during the last few decades. New algorithmic and theoretical techniques have been developed, the diffusion into other disciplines has proceeded at a rapid pace, and our knowledge of all aspects of the field has grown even more profound. At the same time, one of the most striking trends in optimization is the constantly increasing emphasis on the interdisciplinary nature of the field. Optimization has been a basic tool in all areas of applied mathematics, engineering, medicine, economics and other sciences.

An optimization problem consists in maximizing or minimizing some function relative to some set, representing a range of choices available in a certain situation. The function allows comparison of the different choices for determining which might be best. In this work, to focus in convex minimization problems from optimization problem. There are several methods for solving the convex minimization problems such as gradient methods, proximal point algorithm, forward-backward algorithm, Douglas-Rachford algorithm, Peaceman-Rachford algorithm, extragradient method and Tseng's splitting algorithm and so on.

Actually, Dunn [23] (1976), Bertsekas and Tsitsiklis [7] (1997) proposed the gradient method. Next, Martinet [34, 35] and Rockafellar [42, 43] proposed the proximal point algorithm for minimizing a nondifferentiable function. In particular, it will be shown that a number of apparently unrelated, well-known algorithms (e.g., iterative thresholding, projected Landweber, projected gradient, alternating projections, alternating-direction method of multipliers, alternating split Bregman) are special instances of proximal algorithm. In 1979, Lions and Mercier [30] and Passty [40] proposed the forward-backward splitting method which requires one of the function be differentiable, with a Lipschitz continuous gradient. In 1956, Douglas and Rachford proposed the Douglas-Rachford algorithm for solving matrix equations. It can be viewed as more general in scope than the forward-backward algorithm in that it does not require that any of the functions have a Lipschitz continuous gradient. However, this observation must be weighed against the fact that it may be more demanding numerically as it requires the implementation of two proximal steps at each iteration, whereas only one is needed in the forward-backward algorithm. In some problems, both may be easily implementable and it is not clear a priori which algorithm may be more efficient. Applications of the Douglas-Rachford algorithm to signal and image processing can be found in [22]. The Peaceman-Rachford splitting algorithm [18, 24, 30] its convergence requires additional assumptions for the some limiting case Douglas-Rachford algorithm. Tseng [51] (2000) introduced the following modified forward-backward splitting method, also known as Tseng's splitting method. In 2009, Tseng and Yun [52] is the first that considers a variable metric forward-backward algorithm in finite dimensional spaces, where the smooth part is only continuously differentiable (possibly nonconvex). They proposes a general Armjio-type line search rule and prove that cluster points of the generated sequence are stationary points. In 2017, Salzo [44] modified the forward-backward splitting algorithm in infinite dimensional Hilbert spaces without the assumption of the Lipschitz continuity of the gradient and using different types of line search procedures. In 1976, Korpelevich [28] proposed the extragradient method has become a classical method for solving variational inequality problems. For optimization problems, this method generates a sequence of estimates based on two projected gradient steps at each iteration.

Motivated and inspired by the work mentioned above, in this research, we aim to design new algorithms for solving the convex minimization problems in Hilbert spaces. We prove the convergence theorems under some suitable conditions. Finally, we give some applications of minimization problem including its numerical experiments and apply the proposed algorithm to image deblurring, image inpainting and data classification.

CHAPTER II

REVIEW OF RELATED LITERATURE AND RESEARCH

Let H be a real Hilbert space. In this research, the objective of our investigation is to solve the following convex minimization problem:

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

where $g : H \to (-\infty, +\infty]$ is proper, lower semicontinuous and convex, and $f : H \to \mathbb{R}$ is convex and differentiable with the Lipschitz continuous gradient denoted by ∇f. By Fermat's rule, the minimization problem (2.1.1) is to find $x^* \in H$ which satisfies the first order optimality condition:

$$0 \in (\partial g + \nabla f)(x^*),$$

where ∂g denotes the subdifferential of g. It is known that the convex minimization problem has been applied successfully in many real-world problems, such as for signal processing, image reconstruction and many more. We consult the readers to [3, 10, 11, 14, 15, 20, 31, 54, 55, 56].

In what follows, we let Ω represent the solution set of problem (2.1.1). In order to solve (2.1.1), several algorithms based on the proximal operator have been proposed by using the following fixed point equation:

$$x = \text{prox}_{\lambda g}(x - \lambda \nabla f(x)), \ \forall x \in H, \ \lambda > 0,$$

where prox_g is the proximal operator of g given by $\text{prox}_g = (I + \partial g)^{-1}$, where I denotes the identity operator on H.

Formally, for $x_0 \in H$ and the stepsize $\lambda \in (0, 2/L)$, L is the Lipschitz constant of ∇f, the forward-backward algorithm (FB) reads:

$$x_{n+1} = \text{prox}_{\lambda g}(x_n - \lambda \nabla f(x_n)).$$

It was proved that the generated iterates weakly converge to a minimizer of $f + g$. Over the years, various schemes have been proposed to improve the forward-backward algorithm. See also [5, 16, 29, 45, 47].

Alvarez and Attouch [2] introduced the following inertial proximal algorithm. Let $x_0 = x_1$ is chosen arbitrarily,

$$x_{n+1} = \text{prox}_{\lambda_n g}(x_n + \theta_n(x_n - x_{n-1})),$$

where $\{\lambda_n\}$ is a sequence in $(0, \infty)$ and $\{\theta_n\}$ is a sequence in $[0, \infty)$. Polyak [41] proposed an inertial extrapolation as an acceleration process to solve a convex minimization problem.

In 2005, Combettes and Wajs [17] introduced the following relaxed forward-backward method **(RFB)**. Let $\varepsilon \in (0, \min\{1, \frac{1}{L}\})$ and $x_0 \in H$, define

$$\begin{aligned} y_n &= x_n - \lambda_n \nabla f(x_n) \\ x_{n+1} &= x_n + \alpha_n(\text{prox}_{\lambda_n g} y_n - x_n), \end{aligned}$$

where $\lambda_n \in [\varepsilon, \frac{2}{L} - \varepsilon]$, $\alpha_n \in [\varepsilon, 1]$ and L is the Lipschitz constant of the gradient of ∇f.

In the spirit of Nesterov [37], Cruz and Nghia [6] proposed a fast multistep forward-backward method **(FMFB)** with linesearch. Let $x_0 = x_1 \in \text{dom} g$, $t_0 =$

1, $\sigma > 0$, $\delta \in (0, \frac{1}{2})$, $\gamma \in (0,1)$:

$$\begin{aligned} t_{n+1} &= \frac{1+\sqrt{1+4t_{n-1}^2}}{2} \\ \theta_n &= \frac{t_{n-1}-1}{t_n} \\ y_n &= x_n + \theta_n(x_n - x_{n-1}) \\ z_n &= P_{\text{dom}g}(y_n) \\ x_{n+1} &= prox_{\lambda_n g}(z_n - \lambda_n \nabla f(z_n)), \end{aligned}$$

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

$$\lambda_n \|\nabla f(x_{n+1}) - \nabla f(z_n)\| \leq \delta \|x_{n+1} - z_n\|.$$

Recently, Suantai et al. [46] designed the following new linesearch rule.

Linesearch 2.1.1 *Let $x \in H$, $\sigma > 0$, $\rho \in (0,1)$ and $\delta > 0$.*

Input $\lambda = \sigma$.

While

$\frac{\lambda}{2}(\|\nabla f(prox^2_{\lambda g}(x - \lambda\nabla f(x))) - \nabla f(x)\| + \|\nabla f(prox_{\lambda g}(x - \lambda\nabla f(x))) - \nabla f(x)\|)$

$> \delta(\|prox^2_{\lambda g}(x - \lambda\nabla f(x)) - prox_{\lambda g}(x - \lambda\nabla f(x))\| + \|prox_{\lambda g}(x - \lambda\nabla f(x)) - x\|)$,

do $\lambda = \rho\lambda$

End While

Out put λ.

Here $prox^2_{\lambda g}(x - \lambda\nabla f(x)) = prox_{\lambda g}(prox_{\lambda g}(x - \lambda\nabla f(x)))$.

In 2003, Moudafi and Oliny [36] introduced the following inertial forward-backward method (IFB).

Method IFB: The inertial forward-backward method.

Let $x_0, x_1 \in H$, $\lambda_n \in (0, 2/L)$ for all $n \geq 1$. Define

$$\begin{aligned} y_n &= x_n + \theta_n(x_n - x_{n-1}) \\ x_{n+1} &= \text{prox}_{\lambda_n g}(y_n - \lambda_n \nabla f(x_n)), \end{aligned}$$

where θ_n is the inertial parameter which controls the momentum $x_n - x_{n-1}$.

In 2009, the FISTA scheme by Beck and Teboulle [5] is the most-known one, which achieves $O(1/n^2)$ convergence rate.

Method FISTA Let $x_0 = x_1 \in H$, $t_0 = 1$ and $\lambda_n = 1/L$ for all $n \geq 1$. Calculate

$$\begin{aligned} t_n &= \frac{1 + \sqrt{1 + 4t_{n-1}^2}}{2}, \ \theta_n = \frac{t_{n-1} - 1}{t_n} \\ y_n &= x_n + \theta_n(x_n - x_{n-1}) \\ x_{n+1} &= \text{prox}_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n)). \end{aligned}$$

In 2017, Verma and Shukla [53] proposed the new accelerated proximal gradient algorithm (NAGA) as follows:

Method NAGA The new accelerated proximal gradient algorithm.
Let $x_0, x_1 \in H$, $\alpha_n \in (0, 1)$ and $\lambda_n \in (0, 2/L)$ for all $n \geq 1$. Define

$$\begin{aligned} y_n &= x_n + \theta_n(x_n - x_{n-1}) \\ z_n &= (1 - \alpha_n)y_n + \alpha_n \text{prox}_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n)) \\ x_{n+1} &= \text{prox}_{\lambda_n g}(z_n - \lambda_n \nabla f(z_n)). \end{aligned}$$

In 2000, Tseng [51] proposed a modified forward-backward algorithm **(MFB)** via the stepsize with linesearch technique as follows. Given $\sigma > 0$,

$\rho \in (0,1)$, $\delta \in (0,1)$ and $x_1 \in H$. Compute

$$y_n = \text{prox}_{\lambda_n g}(x_n - \lambda_n \nabla f(x_n)),$$
$$x_{n+1} = \text{prox}_{\lambda_n g}(y_n - \lambda_n(\nabla f(y_n) - \nabla f(x_n))),\ n \geq 1$$

where λ_n is the largest $\lambda \in \{\sigma, \sigma\rho, \sigma\rho^2, ...\}$ satisfying $\lambda\|\nabla f(y_n) - \nabla f(x_n)\| \leq \|y_n - x_n\|$.

In 2020, Padcharoen et al. [39] proposed the modified forward-backward splitting method based on inertial Tseng method **(IMFB)**. Given $\{\lambda_n\} \subset (0, \frac{1}{L})$, $\{\alpha_n\} \subset [0, \alpha] \subset [0, 1)$. Let $x_0, x_1 \in H$ and compute

$$w_n = x_n + \theta_n(x_n - x_{n-1}),$$
$$y_n = \text{prox}_{\lambda_n g}(w_n - \lambda_n \nabla f(w_n)),$$
$$x_{n+1} = y_n - \lambda_n(\nabla f(y_n) - \nabla f(w_n)),\ n \geq 1.$$

They established weak convergence of the proposed method.

In 2015, Shehu et al. [45] introduced the modified split proximal method **(MSP)**. Let $r : H \to H$ be a contraction mapping with a constant $\alpha \in (0, 1)$. Set $\varphi(x) = \sqrt{\|\nabla h(x)\|^2 + \|\nabla \ell(x)\|^2}$ with $h(x) = \frac{1}{2}\|(I - \text{prox}_{\lambda g})Ax\|^2$, $\ell(x) = \frac{1}{2}\|(I - \text{prox}_{\lambda\mu_n f})x\|^2$. Given an initial point $x_1 \in H$ and construct

$$y_n = x_n - \mu_n A^*(I - \text{prox}_{\lambda g})Ax_n,$$
$$x_{n+1} = \alpha_n r(x_n) + (1 - \alpha_n)\text{prox}_{\lambda\mu_n f} y_n,\ n \geq 1,$$

where the stepsize $\mu_n = \psi_n \frac{h(x_n)+\ell(x_n)}{\varphi^2(x_n)}$ with $0 < \psi_n < 4$. They proved strong convergence theorem for proximal split feasibility problems.

Very recently, Malitsky and Tam [33] introduced the forward-reflected-backward algorithm **(FRB)**. Given $\lambda_0 > 0$, $\delta \in (0,1)$, $\gamma \in \{1, \beta^{-1}\}$ and $\beta \in (0,1)$. Compute

$$x_{n+1} = \text{prox}_{\lambda_n g}(x_n - \lambda_n \nabla f(x_n) - \lambda_{n-1}(\nabla f(x_n) - \nabla f(x_{n-1}))), \ n \geq 1$$

where the stepsize $\lambda_n = \gamma \lambda_{n-1} \beta^i$ with i being the smallest nonnegative integer satisfying $\lambda_n \|\nabla f(x_{n+1}) - \nabla f(x_n)\| \leq \frac{\delta}{2} \|x_{n+1} - x_n\|$.

Very recently, Hieu et al. [26] proposed the modified forward-reflected-backward method **(MFRB)** with adaptive stepsize. Given $x_0, x_1 \in H$, $\lambda_0, \lambda_1 > 0$, $\mu \in (0, \frac{1}{2})$:

$$\begin{aligned} x_{n+1} &= \text{prox}_{\lambda_n g}(x_n - \lambda_n \nabla f(x_n) - \lambda_{n-1}(\nabla f(x_n) - \nabla f(x_{n-1}))), \\ \lambda_{n+1} &= \min\{\lambda_n, \frac{\mu \|x_{n+1} - x_n\|}{\|\nabla f(x_{n+1}) - \nabla f(x_n)\|}\}, \ n \geq 1. \end{aligned}$$

This stepsize allows the proposed method without knowing the Lipschitz constant to solve the problem.

In our research, we design a new forward-backward methods with the inertial term. We use the adaptive stepsize and new linesearch stepsize in our methods for solving the convex minimization problem. Moreover, we provide weak convergence theorems for the proposed methods. Finally, we present numerical experiments to illustrate and application to image deblurring, image inpanting and data classification. Some comparisons to other methods are also given to show the efficiency of our methods.

CHAPTER III

PRELIMINARIES

In this section, we provide some basic concepts, definitions and lemmas which will be used in the sequel.

3.1 Fundamentals

Definition 3.1.1 [48]**(Fixed point)** Let X be a nonempty set and $T : X \to X$. We say that $x \in X$ is a fixed point of T if

$$T(x) = x$$

and denote by $Fix(T)$ the set of all fixed points of T.

Definition 3.1.2 [4]**(Convex set)** Let C be a subset of a linear space X. Then C is said to be *convex* if $(1-\lambda)x + \lambda y \in C$ for all $x, y \in C$ and all scalar $\lambda \in [0, 1]$.

Proposition 3.1.3 *[4] Let C be a subset of a linear space X. Then C is convex if and only if $\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_k x_k \in C$ for any finite set $\{x_1, x_2, ..., x_k\} \subseteq C$ and scalars $\lambda_i \geq 0$ with $\lambda_1 + \lambda_2 + ... + \lambda_k = 1$.*

Definition 3.1.4 [4]**(Convex function)** Let X be a linear space and $f : X \to (-\infty, \infty]$ be a function. Then f is said to be *convex* if $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ for all $x, y \in X$ and $\lambda \in [0, 1]$.

Definition 3.1.5 [4]**(Proper function)** Let function $f : X \to (-\infty, \infty]$. Then f is said to be *proper* if there exists $x \in X$ with $f(x) < \infty$.

Example 3.1.6 1. $f(x) = |x|^p$ where $p \geq 1$ is a convex function in $\mathbb{R}$.

2. $f(x) = x^3 - x^2$ is a convex function in $[\frac{1}{3}, \infty)$.

3. $f(x) = x \log x$ is a convex function in $\mathbb{R}^+$.

Definition 3.1.7 [1]**(Normed space)** Let X be a norm linear space over field $\mathbb{K}$ ($\mathbb{R}$ 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:

1. $\|x\| \geq 0$, and $\|x\| = 0 \Leftrightarrow x = 0$;
2. $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in X$ and $\alpha \in \mathbb{K}$;
3. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

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

Example 3.1.8 $\mathbb{R}^k$ is a normed space with the following norms:

$$\begin{aligned} \|x\|_1 &= \sum_{i=1}^{k} |x_i| \text{ for all } x = (x_1, x_2, .., x_k) \in \mathbb{R}^k; \\ \|x\|_p &= \Big(\sum_{i=1}^{k} |x_i|^p\Big)^{1/p} \text{ for all } x = (x_1, x_2, .., x_k) \in \mathbb{R}^k \text{ and } p \in (1, \infty); \\ \|x\|_\infty &= \max_{1 \leq i \leq k} |x_i| \text{ for all } x = (x_1, x_2, .., x_k) \in \mathbb{R}^k. \end{aligned}$$

Example 3.1.9 Let $X = l_1$, the linear space whose elements consist of all absolutely convergent sequences $(x_1, x_2, ..., x_i, ...)$ of scalars ($\mathbb{R}$ *or* $\mathbb{C}$),

$$l_1 = \{x : x = (x_1, x_2, ..., x_i, ...) \text{ and } \sum_{i=1}^{\infty} |x_i| < \infty\}.$$

Then l_1 is a normed space with the norm defined by $\|x\|_1 = \sum_{i=1}^{\infty} |x_i|$.

Example 3.1.10 Let $X = l_p$ $(1 < p < \infty)$ be the linear space whose elements consist of all p-summable sequences $(x_1, x_2, ..., x_i, ...)$ of scalars ($\mathbb{R}$ *or* $\mathbb{C}$),

$$l_p = \{x : x = (x_1, x_2, ..., x_i, ...) \text{ and } \sum_{i=1}^{\infty} |x_i|^p < \infty\}.$$

Then l_p is a normed space with the norm defined by $\|x\|_p = (\sum_{i=1}^{\infty} |x_i|^p)^{1/p}$.

Example 3.1.11 Let $X = l_\infty$, the linear space whose elements consist of all bounded sequences $(x_1, x_2, ..., x_i, ...)$ of scalars ($\mathbb{R}$ *or* $\mathbb{C}$),

$$l_\infty = \{x : x = (x_1, x_2, ..., x_i, ...) \text{ and } \{x_i\}_{i=1}^{\infty} \text{ is bounded}\}.$$

Then l_∞ is a normed space with the norm defined by $\|x\|_\infty = \sup_{i\in\mathbb{N}} |x_i|$.

Definition 3.1.12 [1]**(Completeness)** The space X is said to be *complete* if every Cauchy sequence in X converges.

Example 3.1.13 The Euclidean space $\mathbb{R}^k$ is complete with

$$d(x, y) = \sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2 + (\xi_3 - \eta_3)^2 + ... + (\xi_k - \eta_k)^2}$$

where $x = (\xi_1, \xi_2, \xi_3, ..., \xi_k), y = (\eta_1, \eta_2, \eta_3, ..., \eta_k) \in \mathbb{R}^k$.

Example 3.1.14 The sequence space l_∞ is complete.

Example 3.1.15 The sequence space l_p is complete.

Definition 3.1.16 [1]**(Inner product space)** An inner product space is a vector space X with an inner product defined on X. Here, an inner product on X is a mapping of $X \times X$ into the scalar field $\mathbb{K}$ of X; that is, with every pair of vectors x and y there is associated a scalar which is written by $\langle x, y\rangle$ and called the *inner product* of x and y, such that for all vectors x, y, z and scalars α we have

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

Example 3.1.17 The function $\langle \cdot, \cdot \rangle : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$ defined by

$$\langle x, y \rangle = \sum_{i=1}^{k} x_i y_i \text{ for all } x = (x_1, x_2, ..., x_k),\ y = (y_1, y_2, ..., y_k) \in \mathbb{R}^k$$

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

Example 3.1.18 Let $\mathbb{C}^k$ be the set of k-tuples of complex numbers. Then the function $\langle \cdot, \cdot \rangle : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$ defined by

$$\langle x, y \rangle = \sum_{i=1}^{k} x_i \overline{y_i} \text{ for all } x = (x_1, x_2, ..., x_k),\ y = (y_1, y_2, ..., y_k) \in \mathbb{C}^k$$

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

Example 3.1.19 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 $\langle \cdot, \cdot \rangle : l_2 \times l_2 \to \mathbb{C}$ defined by

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i \overline{y_i} \text{ for all } x = \{x_i\}_{i=1}^{\infty}, y = \{y_i\}_{i=1}^{\infty} \in l_2$$

is an inner product on l_2.

Definition 3.1.20 [4]**(Hilbert space)** An inner product space which is complete with respect to the induced norm is called a *Hilbert space*.

Example 3.1.21 The Euclidean space $\mathbb{R}^k$ is a Hilbert space with inner product defined by

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

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

Example 3.1.22 The space l_2 is a Hilbert space with inner product defined by

$$\langle x, y\rangle = \sum_{j=1}^{\infty} x_j \overline{y_j},$$

where $x, y \in l_2$.

Proposition 3.1.23 *[13]***(The Cauchy-Schwarz inequality)** *Let X be an inner product space. Then the following holds:*

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

i.e.,

$$|\langle x, y\rangle| \leq \|x\|\|y\| \textit{ for all } x, y \in X.$$

Definition 3.1.24 [1]**(Bounded linear operator)** Let X and Y be normed spaces and $T : X \to Y$ be a linear operator. The operator T is said to be *bounded* if there is a real number $c > 0$ such that for all $x \in X$,

$$\|Tx\| \leq c\|x\|.$$

Definition 3.1.25 [4]**(Level set of convex function)** Let $f : H \to \mathbb{R}$ be a convex function with the domain H. Then, for any $\lambda \in \mathbb{R}$, the set

$$V_\lambda = \{x \in H | f(x) \leq \lambda\}$$

Definition 3.1.26 [4] A sequence $\{x_n\}$ in a Hilbert space H is said to converge weakly to a point x in H if

$$\langle x_n, y\rangle \to \langle x, y\rangle$$

for all $y \in H$ and denote that $x_n \rightharpoonup x$.

Definition 3.1.27 [1]**(Contraction mapping)** Let H be a real Hilbert space and C be a nonempty subset of H. Then a map $F : C \to C$ is said to be *contraction* if there exists $k \in [0, 1)$ such that

$$\|F(x) - F(y)\| \le k\|x - y\|,$$

for all $x, y \in C$.

Definition 3.1.28 [4]**(Nonexpansive mapping)** Let H be a real Hilbert space and C be a nonempty subset of H. A mapping $T : C \to C$ is said to be *nonexpansive* if

$$\|Tx - Ty\| \le \|x - y\|, \forall x, y \in C.$$

A mapping $T : C \to C$ is said to be *firmly nonexpansive* if, for all $x, y \in C$,

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

The operator $I - P_C$ is also *firmly nonexpansive*, where I denotes the identity operator, *i.e.*, for any $x, y \in H$,

$$\langle (I - P_C)x - (I - P_C)y, x - y\rangle \ge \|(I - P_C)x - (I - P_C)y\|^2.$$

In a real Hilbert space, we know that for any point $x \in H$, there exists a unique point $P_C x \in C$ such that

$$\|x - P_C x\| \le \|x - y\|, \forall y \in C.$$

Here P_C is called the metric projection of H onto C. We know that P_C is a nonexpansive mapping of H onto C. It is also known that P_C satisfies

$$\langle x - y, P_C x - P_C y\rangle \ge \|P_C x - P_C y\|^2,$$

for all $x, y \in H$. Furthermore, $P_C x$ is characterized by the property

$$\langle x - P_C x, P_C x - y \rangle \geq 0,$$

for all $y \in C$. Moreover, we know that

$$\|y - P_C x\|^2 + \|x - P_C x\|^2 \leq \|x - y\|^2, \ \forall x, y \in H.$$

We also know that all Hilbert space has the Kadec-Klee property, that is, (x^k) converges weakly to x and $\|x_n\| \to \|x\|$ imply x_n converges strongly to x.

We know the following equality:

$$\begin{aligned} \|\beta x + (1-\beta) y\|^2 &= \beta \|x\|^2 + (1-\beta)\|y\|^2 \\ &\quad -\beta(1-\beta)\|x - y\|^2, \ \forall x, y \in H, \end{aligned} \tag{3.1.1}$$

where $\beta \in (0, 1)$.

Definition 3.1.29 [4] Let H be a real Hilbert space and let $f : H \to R$, function f is said to be lower semi-continuous at x if $x_n \to x$, then

$$f(x) \leq \liminf_{n \to \infty} f(x_n).$$

Definition 3.1.30 [4] Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a closed proper convex function, The proximal operator $\text{prox}_f : \mathbb{R}^n \to \mathbb{R}^n$ of f is defined by

$$\text{prox}_f(v) = \arg\min_x (f(x) + (1/2)\|x - v\|_2^2),$$

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

$$\text{prox}_{\alpha f}(v) = \arg\min_x (f(x) + (1/2\alpha)\|x - v\|_2^2),$$

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

Let $g : H \to (-\infty, +\infty]$ be a proper, lower semicontinuous and convex function. We denote the domain of g by $\text{dom}g = \{x \in H | g(x) < +\infty\}$. For any $x \in \text{dom}g$, the subdifferential of g at x is defined by

$$\partial g(x) = \{v \in H | \langle v, y - x \rangle \leq g(y) - g(x),\ y \in H\}.$$

Recall that the proximal operator $\text{prox}_g : \text{dom}(g) \to H$ is defined by $\text{prox}_g(x) = (I + \partial g)^{-1}(z),\ z \in H$. It is well-known that the proximal operator is single-valued. Furthermore, we have

$$\frac{z - \text{prox}_{\lambda g}(z)}{\lambda} \in \partial g(\text{prox}_{\lambda g}(z)) \qquad \text{for all } z \in H,\ \lambda > 0. \tag{3.1.2}$$

A differentiable function f is convex if and only if there holds the inequality

$$f(z) \geq f(x) + \langle \nabla f(x), z - x \rangle,\ \forall z \in H. \tag{3.1.3}$$

Definition 3.1.31 [4] Let S be a nonempty subset of H. A sequence $\{x_n\}$ in H is said to be quasi-Fejér convergent to S if and only if for all $x \in S$ there exists a positive sequence $\{\varepsilon_n\}$ such that $\sum_{n=0}^{\infty} \varepsilon_n < +\infty$ and $\|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 + \varepsilon_n$ for all $n \geq 1$. When $\{\varepsilon_n\}$ is a null sequence, we say that $\{x_n\}$ is Fejér convergent to S.

3.2 Lemmas

Lemma 3.2.1 *[4] Let Ω be a nonempty closed convex subset of a real Hilbert space H. Then, for any $x \in H$, the following assertions hold:*

(1) $\langle x - P_\Omega x, z - P_\Omega x \rangle \leq 0$ *for all* $z \in \Omega$*;*

(2) $\|P_\Omega x - P_\Omega y\|^2 \leq \langle P_\Omega x - P_\Omega y, x - y \rangle$ *for all* $x, y \in H$*;*

(3) $\|P_\Omega x - z\|^2 \leq \|x - z\|^2 - \|P_\Omega x - x\|^2$ *for all* $z \in \Omega$*.*

Lemma 3.2.2 *[9] The subdifferential operator* ∂g *is maximal monotone. Moreover, the graph of* ∂g*,* $\mathrm{Gph}(\partial g) = \{(x, v) \in H \times H : v \in \partial g(x)\}$ *is demiclosed, i.e., if the sequence* $\{(x_n, v_n)\} \subset \mathrm{Gph}(\partial g)$ *satisfies that* $\{x_n\}$ *converges weakly to* x *and* $\{v_n\}$ *converges strongly to* v*, then* $(x, v) \in \mathrm{Gph}(\partial g)$*.*

Lemma 3.2.3 *[38] Let* $\{a_n\}$*,* $\{b_n\}$ *and* $\{c_n\}$ *be real positive sequences such that*

$$a_{n+1} \leq (1 + c_n)a_n + b_n, \;\; n \geq 1.$$

If $\Sigma_{n=1}^{\infty} c_n < +\infty$ *and* $\Sigma_{n=1}^{\infty} b_n < +\infty$*, then* $\lim\limits_{n\to+\infty} a_n$ *exists.*

Lemma 3.2.4 *[25] Let* $\{a_n\}$ *and* $\{\theta_n\}$ *be real positive sequences such that*

$$a_{n+1} \leq (1 + \theta_n)a_n + \theta_n a_{n-1}, \;\; n \geq 1.$$

Then, $a_{n+1} \leq K \cdot \prod_{i=1}^{n}(1 + 2\theta_i)$ *where* $K = \max\{a_1, a_2\}$*. Moreover, if* $\sum_{n=1}^{\infty} \theta_n < +\infty$*, then* $\{a_n\}$ *is bounded.*

Lemma 3.2.5 *[4, 27] If* $\{x_n\}$ *is quasi-Fejér convergent to* S*, then we have:*

(i) $\{x_n\}$ *is bounded.*

(ii) If all weak accumulation points of $\{x_n\}$ *is in* S*, then* $\{x_n\}$ *weakly converges to a point in* S*.*

CHAPTER IV

SPLITTING METHODS

4.1 Inertial modified forward-backward method

In this section, we assume that the following conditions are satisfied for our convergence analysis:

(A1) The solution set of the convex minimization problem (2.1.1) is nonempty, *i.e.*, $\Omega = \text{argmin}(f+g) \neq \emptyset$.

(A2) $f, g : H \to (-\infty, +\infty]$ are two proper, lower semicontinuous and convex functions.

(A3) f is differentiable on H and ∇f is Lipschitz continuous on H with the Lipschitz constant $L > 0$.

We next introduce a new inertial forward-backward method for solving (2.1.1).

Algorithm 4.1.1 Inertial modified forward-backward method (IMFB)
Initialization: *Let* $x_0 = x_1 \in H$, $\lambda_1 > 0$, $\theta_1 > 0$ *and* $\delta \in (0,1)$.
Iterative step: *For* $n \geq 1$, *calculate* x_{n+1} *as follows:*
Step 1. *Compute the inertial step:*

$$w_n = x_n + \theta_n(x_n - x_{n-1}). \tag{4.1.1}$$

Step 2. *Compute the forward-backward step:*

$$y_n = prox_{\lambda_n g}(w_n - \lambda_n \nabla f(w_n)). \tag{4.1.2}$$

Step 3. *Compute the x_{n+1} step:*

$$x_{n+1} = y_n - \lambda_n(\nabla f(y_n) - \nabla f(w_n)) \tag{4.1.3}$$

where

$$\lambda_{n+1} = \begin{cases} \min\{\frac{\delta\|w_n - y_n\|}{\|\nabla f(w_n) - \nabla f(y_n)\|}, \lambda_n\} & if\ \|\nabla f(w_n) - \nabla f(y_n)\| \neq 0; \\ \lambda_n & otherwise. \end{cases}$$

Set $n = n+1$ and return to **Step 1**.

Remark 4.1.2 It is easy to see that the sequence $\{\lambda_n\}$ is non-increasing. From the Lipschitz continuity of ∇f, there exists $L > 0$ such that $\|\nabla f(w_n) - \nabla f(y_n)\| \leq L\|w_n - y_n\|$. Hence,

$$\lambda_{n+1} = \min\left\{\frac{\delta\|w_n - y_n\|}{\|\nabla f(w_n) - \nabla f(y_n)\|}, \lambda_n\right\} \geq \min\{\frac{\delta}{L}, \lambda_n\}.$$

By the definition of $\{\lambda_n\}$, it implies that the sequence $\{\lambda_n\}$ is bounded from below by $\min\{\lambda_0, \frac{\delta}{L}\}$. So, we obtain $\lim_{n\to\infty} \lambda_n = \lambda > 0$.

Lemma 4.1.3 *Let $\{x_n\}$ be generated by Algorithm 4.1.1. Then*

$$\|x_{n+1} - x^*\|^2 \leq \|w_n - x^*\|^2 - (1 - \frac{\delta^2\lambda_n^2}{\lambda_{n+1}^2})\|y_n - w_n\|^2, \forall x^* \in \Omega.$$

Proof. Let $x^* \in \Omega$. Then

$$\begin{aligned} \|x_{n+1} - x^*\|^2 &= \|y_n - \lambda_n(\nabla f(y_n) - \nabla f(w_n)) - x^*\|^2 \\ &= \|y_n - x^*\|^2 + \lambda_n^2\|\nabla f(y_n) - \nabla f(w_n)\|^2 \\ &\quad -2\lambda_n\langle y_n - x^*, \nabla f(y_n) - \nabla f(w_n)\rangle \\ &= \|y_n - w_n + w_n - x^*\|^2 + \lambda_n^2\|\nabla f(y_n) - \nabla f(w_n)\|^2 \\ &\quad -2\lambda_n\langle y_n - x^*, \nabla f(y_n) - \nabla f(w_n)\rangle \end{aligned}$$

$$
\begin{aligned}
&= \|w_n - x^*\|^2 + \|y_n - w_n\|^2 + 2\langle w_n - x^*, y_n - w_n\rangle \\
&\quad -2\lambda_n\langle y_n - x^*, \nabla f(y_n) - \nabla f(w_n)\rangle + \lambda_n^2\|\nabla f(y_n) - \nabla f(w_n)\|^2 \\
&= \|w_n - x^*\|^2 + \|y_n - w_n\|^2 + 2\langle w_n - y_n + y_n - x^*, y_n - w_n\rangle \\
&\quad -2\lambda_n\langle y_n - x^*, \nabla f(y_n) - \nabla f(w_n)\rangle + \lambda_n^2\|\nabla f(y_n) - \nabla f(w_n)\|^2 \\
&= \|w_n - x^*\|^2 + \|y_n - w_n\|^2 - 2\langle y_n - w_n, y_n - w_n\rangle \\
&\quad +2\langle y_n - x^*, y_n - w_n\rangle - 2\lambda_n\langle y_n - x^*, \nabla f(y_n) - \nabla f(w_n)\rangle \\
&\quad +\lambda_n^2\|\nabla f(y_n) - \nabla f(w_n)\|^2 \\
&= \|w_n - x^*\|^2 + \|y_n - w_n\|^2 - 2\|y_n - w_n\|^2 + 2\langle y_n - x^*, y_n - w_n\rangle \\
&\quad -2\langle y_n - x^*, \lambda_n(\nabla f(y_n) - \nabla f(w_n))\rangle + \lambda_n^2\|\nabla f(y_n) - \nabla f(w_n)\|^2 \\
&= \|w_n - x^*\|^2 - \|y_n - w_n\|^2 \\
&\quad -2\langle y_n - x^*, w_n - y_n + \lambda_n(\nabla f(y_n) - \nabla f(w_n))\rangle \\
&\quad +\lambda_n^2\|\nabla f(y_n) - \nabla f(w_n)\|^2. \qquad (4.1.4)
\end{aligned}
$$

Note that

$$\lambda_{n+1} = \min\left\{\frac{\delta\|w_n - y_n\|}{\|\nabla f(w_n) - \nabla f(y_n)\|}, \lambda_n\right\} \le \frac{\delta\|w_n - y_n\|}{\|\nabla f(w_n) - \nabla f(y_n)\|}.$$

It follows that

$$\|\nabla f(w_n) - \nabla f(y_n)\| \le \frac{\delta}{\lambda_{n+1}}\|w_n - y_n\|. \qquad (4.1.5)$$

Combining (4.1.4) and (4.1.5), we obtain

$$
\begin{aligned}
\|x_{n+1} - x^*\|^2 &\le \|w_n - x^*\|^2 - \|y_n - w_n\|^2 + \frac{\delta^2\lambda_n^2}{\lambda_{n+1}^2}\|y_n - w_n\|^2 \\
&\quad -2\langle y_n - x^*, w_n - y_n + \lambda_n(\nabla f(y_n) - \nabla f(w_n))\rangle \\
&= \|w_n - x^*\|^2 - (1 - \frac{\delta^2\lambda_n^2}{\lambda_{n+1}^2})\|y_n - w_n\|^2
\end{aligned}
$$

$$-2\langle y_n - x^*, w_n - y_n + \lambda_n(\nabla f(y_n) - \nabla f(w_n))\rangle. \quad (4.1.6)$$

From (4.1.2), we see that $w_n - \lambda_n \nabla f(w_n) \in (I + \lambda_n \partial g)y_n$. Since ∂g is maximal monotone, then there is $u_n \in \partial g(y_n)$ such that

$$w_n - \lambda_n \nabla f(w_n) = y_n + \lambda_n u_n.$$

This shows that

$$u_n = \frac{1}{\lambda_n}(w_n - \lambda_n \nabla f(w_n) - y_n). \quad (4.1.7)$$

Since $0 \in (\nabla f + \partial g)(x^*)$ and $\nabla f(y_n) + u_n \in (\nabla f + \partial g)y_n$, we get

$$\langle \nabla f(y_n) + u_n, y_n - x^* \rangle \geq 0. \quad (4.1.8)$$

Substituting (4.1.7) into (4.1.8), we have

$$\frac{1}{\lambda_n}\langle w_n - \lambda_n \nabla f(w_n) - y_n + \lambda_n \nabla f(y_n), y_n - x^* \rangle \geq 0.$$

This implies that $\langle w_n - \lambda_n \nabla f(w_n) - y_n + \lambda_n \nabla f(y_n), y_n - x^* \rangle \geq 0$. Using (4.1.6), we derive

$$\|x_{n+1} - x^*\|^2 \leq \|w_n - x^*\|^2 - (1 - \frac{\delta^2 \lambda_n^2}{\lambda_{n+1}^2})\|y_n - w_n\|^2. \quad (4.1.9)$$

□

Lemma 4.1.4 *Let* $\{x_n\}$ *be generated by Algorithm 4.1.1. If* $\sum_{n=1}^{\infty} \theta_n < \infty$*, then* $\lim_{n\to\infty} \|x_n - x^*\|$ *exists for all* $x^* \in \Omega$.

Proof. Let $x^* \in \Omega$. From Lemma 4.1.3, we see that

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

So, we have

$$\begin{aligned}
\|x_{n+1} - x^*\| &\leq \|w_n - x^*\| \\
&= \|x_n + \theta_n(x_n - x_{n-1}) - x^*\| \\
&\leq \|x_n - x^*\| + \theta_n\|x_n - x_{n-1}\| \\
&\leq \|x_n - x^*\| + \theta_n(\|x_n - x^*\| + \|x_{n-1} - x^*\|). \qquad (4.1.10)
\end{aligned}$$

Hence

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

By Lemma 3.2.4, we conclude that

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

where $K = \max\{\|x_1 - x^*\|, \|x_2 - x^*\|\}$. Since $\sum_{n=1}^{\infty} \theta_n < +\infty$, by Lemma 3.2.4, we have $\{x_n - x^*\}$ is bounded. Hence $\sum_{n=1}^{\infty} \theta_n\|x_n - x_{n-1}\| < +\infty$. By Lemma 3.2.3 and (4.1.10), we have $\lim_{n\to\infty} \|x_n - x^*\|$ exists. □

Lemma 4.1.5 *Let $\{x_n\}$ be generated by Algorithm 4.1.1. If $\sum_{n=1}^{\infty} \theta_n < \infty$, then*

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

Proof. We see that

$$
\begin{aligned}
& \|w_n - x^*\|^2 \\
= \;& \|x_n + \theta_n(x_n - x_{n-1}) - x^*\|^2 \\
= \;& \|x_n - x^*\|^2 + 2\theta_n\langle x_n - x^*, x_n - x_{n-1}\rangle + \theta_n^2\|x_n - x_{n-1}\|^2 \\
\leq \;& \|x_n - x^*\|^2 + 2\theta_n\|x_n - x^*\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2. \quad (4.1.11)
\end{aligned}
$$

From (4.1.9) and (4.1.11), we have

$$
\begin{aligned}
\|x_{n+1} - x^*\|^2 \;\leq\; & \|x_n - x^*\|^2 + 2\theta_n\|x_n - x^*\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2 \\
& -(1 - \frac{\delta^2\lambda_n^2}{\lambda_{n+1}^2})\|w_n - y_n\|^2. \quad (4.1.12)
\end{aligned}
$$

Note that $\theta_n\|x_n - x_{n-1}\| \to 0$ and $\lim_{n\to\infty}\|x_n - x^*\|$ exists by Lemma 4.1.4. From (4.1.1) and (4.1.12), we have $\|w_n - x_n\| \to 0$ and $\|w_n - y_n\| \to 0$, respectively. It is easy to see that $\|x_n - y_n\| \to 0$. Since ∇f is uniformly continuous, we obtain

$$
\lim_{n\to\infty}\|\nabla f(w_n) - \nabla f(y_n)\| = 0. \quad (4.1.13)
$$

From (4.1.3) and (4.1.13), we get

$$
\begin{aligned}
\|x_{n+1} - y_n\| &= \lambda_n\|\nabla f(y_n) - \nabla f(w_n)\| \\
&\to 0 \text{ (as } n \to \infty).
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
\|x_{n+1} - x_n\| &\leq \|x_{n+1} - y_n\| + \|y_n - x_n\| \\
&\to 0 \text{ (as } n \to \infty).
\end{aligned}
$$

□

Theorem 4.1.6 *Let $\{x_n\}$ be generated by Algorithm 4.1.1. If $\sum_{n=1}^{\infty} \theta_n < \infty$, then $\{x_n\}$ weakly converges to a point in Ω.*

Proof. Since $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \rightharpoonup \bar{x} \in H$. From Lemma 4.1.5, we obtain $x_{n_k+1} \rightharpoonup \bar{x}$. We note that

$$y_{n_k} = prox_{\lambda_{n_k} g}(w_{n_k} - \lambda_{n_k} \nabla f(w_{n_k})).$$

From (3.1.2), we obtain

$$\frac{w_{n_k} - \lambda_{n_k} \nabla f(w_{n_k}) - y_{n_k}}{\lambda_{n_k}} \in \partial g(y_{n_k}).$$

Hence

$$\frac{w_{n_k} - y_{n_k}}{\lambda_{n_k}} - \nabla f(w_{n_k}) + \nabla f(y_{n_k}) \in \partial g(y_{n_k}) + \nabla f(y_{n_k}). \tag{4.1.14}$$

Since $\|x_n - y_n\| \to 0$, we also have $y_{n_k} \rightharpoonup \bar{x}$. Letting $k \to \infty$ in (4.1.14) and using (4.1.13), by Lemma 3.2.2 and Remark 4.1.2, we get

$$0 \in (\nabla f + \partial g)(\bar{x}).$$

So $\bar{x} \in \Omega$. From (4.1.12) we see that $\{x_n\}$ is a quasi-Fejer sequence. Hence, by Lemma 3.2.5, we conclude that $\{x_n\}$ weakly converges to a point in Ω. This completes the proof. □

4.2 Inertial modified relaxed forward-backward-forward method

In this section, we introduce an algorithm using the inertial extrapolation and the adaptive stepsize. We assume that the following conditions are satisfied for our convergence analysis:

(A1) The solution set of the convex minimization problem (2.1.1) is nonempty, *i.e.*, $S = \text{argmin}(f+g) \cap \Omega \neq \emptyset$, where Ω is a subset of H.

(A2) $f, g : H \to (-\infty, +\infty]$ are two proper, lower semicontinuous and convex functions.

(A3) f is differentiable on H and ∇f is Lipschitz continuous on H with the Lipschitz constant $L > 0$.

We next introduce a relaxed inertial forward-backward-forward method for solving (2.1.1).

Algorithm 4.2.1 Inertial modified relaxed forward-backward-forward method (IRFBF)

Initialization: *Let* $x_0 = x_1 \in H$, $\lambda_1 > 0$, $\rho_1 \in (0,1)$, $\mu \in (0,1)$ *and* $\theta_1 \geq 0$.

Iterative step: *Let* Ω *be a nonempty closed convex subset of* H. *Given* $n \geq 1$, *calculate* x_{n+1} *as follows:*

Step 1. *Compute the inertial step:*

$$w_n = x_n + \theta_n(x_n - x_{n-1}).$$

Step 2. *Compute the forward-backward-forward step:*

$$\begin{aligned} y_n &= prox_{\lambda_n g}(w_n - \lambda_n \nabla f(w_n)), \\ z_n &= (1-\rho_n)w_n + \rho_n(y_n - \lambda_n(\nabla f(y_n) - \nabla f(w_n))), \end{aligned}$$

Step 3. *Compute the projection step:*

$$x_{n+1} = P_\Omega(z_n).$$

Step 4. *Compute the stepsize step:*

$$\lambda_{n+1} = \begin{cases} \min\{\lambda_n, \frac{\mu\|y_n - w_n\|}{\|\nabla f(y_n) - \nabla f(w_n)\|}\} & if\ \|\nabla f(y_n) - \nabla f(w_n)\| \neq 0; \\ \lambda_n & otherwise. \end{cases} \tag{4.2.1}$$

Set $n = n + 1$ *and return to* **Step 1**.

Using the proof line as in [8], we obtain the following lemma.

Lemma 4.2.2 *Let* $\mu \in (0, 1)$ *and* $\lambda_1 > 0$. *The sequence* $\{\lambda_n\}$ *generated by (4.2.1) is nonincreasing and*

$$\lim_{n\to\infty} \lambda_n = \lambda \geq \min\{\lambda_1, \frac{\mu}{L}\}.$$

Hence,

$$\|\nabla f(y_n) - \nabla f(w_n)\| \leq \frac{\mu}{\lambda_{n+1}}\|y_n - w_n\|. \tag{4.2.2}$$

Theorem 4.2.3 *Let* $\{x_n\}$ *be generated by Algorithm 4.2.1. If* $\sum_{n=1}^{\infty} \theta_n < +\infty$ *and* $0 < \liminf_{n\to\infty} \rho_n \leq \limsup_{n\to\infty} \rho_n < 1$, *then the sequence* $\{x_n\}$ *weakly converges to an element of* S.

Proof. Let $x^* \in S$. Then, by Lemma 3.2.1(3), we have

$$\begin{aligned} \|x_{n+1} - x^*\|^2 &= \|P_\Omega(z_n) - x^*\|^2 \\ &\leq \|z_n - x^*\|^2 - \|P_\Omega(z_n) - z_n\|^2. \end{aligned} \tag{4.2.3}$$

From (3.1.1) and by setting $v_n = y_n - \lambda_n(\nabla f(y_n) - \nabla f(w_n))$, we obtain

$$\begin{aligned} \|z_n - x^*\|^2 &= \|(1 - \rho_n)w_n + \rho_n v_n - x^*\|^2 \\ &= \|(1 - \rho_n)(w_n - x^*) + \rho_n(v_n - x^*)\|^2 \end{aligned}$$

$$= (1-\rho_n)\|w_n - x^*\|^2 + \rho_n\|v_n - x^*\|^2 - \rho_n(1-\rho_n)\|v_n - w_n\|^2. \quad (4.2.4)$$

From definition of v_n and (4.2.2), we have

$$\begin{aligned}
\|w_n - x^*\|^2 &= \|w_n - y_n + y_n - v_n + v_n - x^*\|^2 \\
&= \|w_n - y_n\|^2 + \|y_n - v_n\|^2 + \|v_n - x^*\|^2 + 2\langle w_n - y_n, y_n - v_n\rangle \\
&\quad +2\langle y_n - v_n, v_n - x^*\rangle + 2\langle v_n - x^*, w_n - y_n\rangle \\
&= \|w_n - y_n\|^2 + \|y_n - v_n\|^2 + \|v_n - x^*\|^2 + 2\langle w_n - y_n, y_n - x^*\rangle \\
&\quad +2\langle y_n - v_n, v_n - x^*\rangle \\
&= \|w_n - y_n\|^2 + \|y_n - v_n\|^2 + \|v_n - x^*\|^2 + 2\langle w_n - y_n, y_n - x^*\rangle \\
&\quad +2\langle y_n - v_n, v_n - y_n + y_n - x^*\rangle \\
&= \|w_n - y_n\|^2 + \|y_n - v_n\|^2 + \|v_n - x^*\|^2 + 2\langle w_n - y_n, y_n - x^*\rangle \\
&\quad -2\langle y_n - v_n, y_n - v_n\rangle + 2\langle y_n - v_n, y_n - x^*\rangle \\
&= \|w_n - y_n\|^2 - \|y_n - v_n\|^2 + \|v_n - x^*\|^2 + 2\langle w_n - v_n, y_n - x^*\rangle \\
&= \|w_n - y_n\|^2 - \lambda_n^2\|\nabla f(y_n) - \nabla f(w_n)\|^2 + \|v_n - x^*\|^2 \\
&\quad +2\langle w_n - v_n, y_n - x^*\rangle \\
&\geq \|w_n - y_n\|^2 - \frac{\lambda_n^2\mu^2}{\lambda_{n+1}^2}\|y_n - w_n\|^2 + \|v_n - x^*\|^2 \\
&\quad +2\langle w_n - v_n, y_n - x^*\rangle.
\end{aligned}$$

It follows that

$$\|v_n - x^*\|^2 \leq \|w_n - x^*\|^2 - \|w_n - y_n\|^2 + \frac{\lambda_n^2\mu^2}{\lambda_{n+1}^2}\|y_n - w_n\|^2 - 2\langle w_n - v_n, y_n - x^*\rangle. \quad (4.2.5)$$

Since

$$(I-\lambda_n\nabla f)(w_n)\in(I+\lambda_n\partial g)(y_n),$$

we have

$$\begin{aligned}w_n &\in y_n+\lambda_n\partial g(y_n)+\lambda_n\nabla f(w_n)\\ &= y_n-\lambda_n(\nabla f(y_n)-\nabla f(w_n))+\lambda_n(\partial g+\nabla f)(y_n)\\ &= v_n+\lambda_n(\partial g+\nabla f)(y_n).\end{aligned}$$

Hence,

$$\frac{1}{\lambda_n}(w_n-v_n)\in(\partial g+\nabla f)(y_n).$$

This, together with $0\in(\partial g+\nabla f)(x^*)$ and the monotonicity of $\partial g+\nabla f$, implies

$$\langle w_n-v_n,y_n-x^*\rangle\geq 0. \tag{4.2.6}$$

From (4.2.5) and (4.2.6), we have

$$\begin{aligned}\|v_n-x^*\|^2 &\leq \|w_n-x^*\|^2-\|w_n-y_n\|^2+\frac{\lambda_n^2\mu^2}{\lambda_{n+1}^2}\|w_n-y_n\|^2\\ &= \|w_n-x^*\|^2-\left(1-\frac{\lambda_n^2\mu}{\lambda_{n+1}^2}\right)\|w_n-y_n\|^2. \qquad (4.2.7)\end{aligned}$$

From (4.2.4) and (4.2.7), we have

$$\begin{aligned}\|z_n-x^*\|^2 &\leq (1-\rho_n)\|w_n-x^*\|^2+\rho_n\|w_n-x^*\|^2-\rho_n\left(1-\frac{\lambda_n^2\mu^2}{\lambda_{n+1}^2}\right)\|w_n-y_n\|^2\\ &\quad -\rho_n(1-\rho_n)\|v_n-w_n\|^2\\ &= \|w_n-x^*\|^2-\left(1-\frac{\lambda_n^2\mu}{\lambda_{n+1}^2}\right)\|w_n-y_n\|^2\\ &\quad -\rho_n(1-\rho_n)\|v_n-w_n\|^2. \qquad (4.2.8)\end{aligned}$$

This shows that

$$\|z_n - x^*\| \leq \|w_n - x^*\|. \tag{4.2.9}$$

From (4.2.3) and (4.2.9), we also have

$$\|x_{n+1} - x^*\| \leq \|z_n - x^*\| \leq \|w_n - x^*\|.$$

Hence,

$$\begin{aligned} \|x_{n+1} - x^*\| &\leq \|w_n - x^*\| \\ &= \|x_n + \theta_n(x_n - x_{n-1}) - x^*\| \\ &\leq \|x_n - x^*\| + \theta_n\|x_n - x_{n-1}\| \\ &= \|x_n - x^*\| + \theta_n\|x_n - x^* + x^* - x_{n-1}\| \\ &\leq \|x_n - x^*\| + \theta_n(\|x_n - x^*\| + \|x_{n-1} - x^*\|). \end{aligned} \tag{4.2.10}$$

Therefore

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

By Lemma 3.2.4, we conclude that

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

where $K = \max\{\|x_1 - x^*\|, \|x_2 - x^*\|\}$. Since $\sum_{n=1}^{\infty} \theta_n < +\infty$ and $\{x_n\}$ is bounded, we have $\sum_{n=1}^{\infty} \theta_n\|x_n - x_{n-1}\| < +\infty$. By Lemma 3.2.3 and (4.2.10), we have $\lim_{n\to\infty} \|x_n - x^*\|$ exists.

Next, we consider

$$\begin{aligned}
&\|w_n - x^*\|^2 \\
= \;& \|x_n + \theta_n(x_n - x_{n-1}) - x^*\|^2 \\
= \;& \|x_n - x^*\|^2 + 2\theta_n\langle x_n - x^*, x_n - x_{n-1}\rangle + \theta_n^2\|x_n - x_{n-1}\|^2 \\
\leq \;& \|x_n - x^*\|^2 + 2\theta_n\|x_n - x^*\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2. \quad (4.2.11)
\end{aligned}$$

From (4.2.3), (4.2.8) and (4.2.11), we get

$$\begin{aligned}
\|x_{n+1} - x^*\|^2 \leq \;& \|x_n - x^*\|^2 + 2\theta_n\|x_n - x^*\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2 \\
& -\rho_n\left(1 - \frac{\lambda_n^2\mu^2}{\lambda_{n+1}^2}\right)\|w_n - y_n\|^2 - \rho_n(1-\rho_n)\|v_n - w_n\|^2 \\
& -\|x_{n+1} - z_n\|^2. \qquad (4.2.12)
\end{aligned}$$

From (4.2.12), we have $\|w_n - y_n\| \to 0$, $\|v_n - w_n\| \to 0$ and $\|x_{n+1} - z_n\| \to 0$. From definition of w_n, we see that $\|x_n - w_n\| \to 0$. Since ∇f is uniformly continuous, we obtain

$$\lim_{n\to\infty} \|\nabla f(w_n) - \nabla f(y_n)\| = 0. \qquad (4.2.13)$$

From definition of v_n and (4.2.13), we have

$$\begin{aligned}
\|y_n - v_n\| &= \lambda_n\|\nabla f(y_n) - \nabla f(w_n)\| \\
&\to 0. \qquad (4.2.14)
\end{aligned}$$

On the other hand, we see that

$$\begin{aligned}
\|z_n - y_n\|^2 &= \|(1-\rho_n)w_n + \rho_n v_n - y_n\|^2 \\
&= (1-\rho_n)\|w_n - y_n\|^2 + \rho_n\|v_n - y_n\|^2 \\
&\quad -\rho_n(1-\rho_n)\|w_n - v_n\|^2.
\end{aligned}$$

Since $\|w_n - y_n\| \to 0$, $\|v_n - w_n\| \to 0$ and (4.2.14), it follows that

$$\|z_n - y_n\| \to 0.$$

Also, we have

$$\begin{aligned} \|x_n - y_n\| &\leq \|x_n - w_n\| + \|w_n - y_n\| \\ &\to 0 \end{aligned}$$

and

$$\begin{aligned} \|z_n - x_n\| &\leq \|z_n - y_n\| + \|y_n - x_n\| \\ &\to 0. \end{aligned} \tag{4.2.15}$$

Hence, we obtain

$$\begin{aligned} \|x_{n+1} - x_n\| &\leq \|x_{n+1} - z_n\| + \|z_n - x_n\| \\ &\to 0. \end{aligned} \tag{4.2.16}$$

Since the sequence $\{x_n\}$ is bounded, the set of its weak accumulation points is nonempty. Take any weak accumulation point $\bar{x}$ of the sequence y_n. Since $\|x_n - y_n\| \to 0$, there is a subsequence $\{y_{n_k}\}$ of $\{y_n\}$ weakly converging to $\bar{x}$.

From

$$y_{n_k} = prox_{\lambda_{n_k} g}(w_{n_k} - \lambda_{n_k} \nabla f(w_{n_k})),$$

it follows that

$$\frac{w_{n_k} - \lambda_{n_k} \nabla f(w_{n_k}) - y_{n_k}}{\lambda_{n_k}} \in \partial g(y_{n_k}).$$

This implies that

$$\frac{w_{n_k} - y_{n_k}}{\lambda_{n_k}} - \nabla f(w_{n_k}) + \nabla f(y_{n_k}) \in \partial g(y_{n_k}) + \nabla f(y_{n_k}). \tag{4.2.17}$$

Letting $k \to \infty$ in (4.2.17), we obtain by Lemma 3.2.5,

$$0 \in (\partial g + \nabla f)(\bar{x}).$$

Thus $\bar{x} \in \text{argmin}(f + g)$.

Next, we will show that $\bar{x} \in \Omega$. Since P_Ω is nonexpansive, by (4.2.15) and (4.2.16), we have

$$\begin{aligned} \|P_\Omega(x_n) - x_n\| &\leq \|P_\Omega(x_n) - P_\Omega(z_n)\| + \|P_\Omega(z_n) - x_n\| \\ &\leq \|x_n - z_n\| + \|x_{n+1} - x_n\| \\ &\to 0. \end{aligned}$$

Hence, by the demiclosedness of P_Ω, we obtain $\bar{x} \in \Omega$. By Lemma 3.2.5, we conclude that the sequence $\{x_n\}$ weakly converges to a point in S. This completes the proof. □

4.3 New proximal gradient method

In this section, we introduce our algorithm for solving the convex minimization problem (2.1.1). Following Cruz and Nghia [6], we assume that

(I) $f, g : H \to (-\infty, +\infty]$ are two proper, lower semicontinuous and convex functions with $\text{dom}g \subseteq \text{dom}f$.

(II) The function f is Fréchet differentiable on an open set containing $\text{dom}g$. The gradient ∇f is uniformly continuous on any bounded subset of $\text{dom}g$

and maps any bounded subset of dom$g$ to a bounded set in H.

Algorithm 4.3.1 Initialization: *Let* $x_0 = x_1 \in H$, $\sigma > 0$, $\rho \in (0,1)$, $\delta \in (0,1)$, $\theta_1 > 0$

Iterative step: *For* $n \geq 1$, *calculate* x_{n+1} *as follows:*

Step 1. *Compute the inertial step:*

$$w_n = x_n + \theta_n(x_n - x_{n-1}). \tag{4.3.1}$$

Step 2. *Compute the forward-backward step:*

$$y_n = prox_{\lambda_n g}(w_n - \lambda_n \nabla f(w_n)),$$

where $\lambda_n = \sigma\rho^{m_n}$ *and* m_n *is the smallest nonnegative integer such that*

$$\begin{aligned}&\lambda_n(\|\nabla f(prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n))) - \nabla f(y_n)\| + \|\nabla f(w_n) - \nabla f(y_n)\|)\\ &\leq \frac{\delta}{2}(\|prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n)) - y_n\| + \|w_n - y_n\|).\end{aligned}$$

Step 3. *Compute the* x_{n+1} *step:*

$$x_{n+1} = prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n)).$$

Set $n = n + 1$ *and return to Step 1.*

Theorem 4.3.2 *Let* $\{x_n\}$ *be generated by Algorithm 4.3.1. If* $\sum_{n=1}^{\infty} \theta_n < +\infty$ *and* $\lambda_n \geq \lambda$ *for some* $\lambda > 0$, *then the sequence* $\{x_n\}$ *weakly converges to an element of* Ω.

Proof. By definition of proximal operator, we have

$$\frac{w_n - y_n}{\lambda_n} - \nabla f(w_n) = \frac{w_n - prox_{\lambda_n g}(w_n - \lambda_n \nabla f(w_n))}{\lambda_n} - \nabla f(w_n) \in \partial g(y_n).$$

It follows from the convexity of g that

$$g(x) - g(y_n) \geq \big\langle \frac{w_n - y_n}{\lambda_n} - \nabla f(w_n), x - y_n \big\rangle, \forall x \in \text{dom} g. \tag{4.3.2}$$

Also, we have

$$\frac{y_n - x_{n+1}}{\lambda_n} - \nabla f(y_n) = \frac{y_n - prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n))}{\lambda_n} - \nabla f(y_n) \in \partial g(x_{n+1}).$$

By the convexity of g, we also have

$$g(x) - g(x_{n+1}) \geq \big\langle \frac{y_n - x_{n+1}}{\lambda_n} - \nabla f(y_n), x - x_{n+1} \big\rangle, \forall x \in \text{dom} g. \tag{4.3.3}$$

On the other hand, for any $x \in \text{dom} g \subseteq \text{dom} f$, we have

$$f(x) - f(w_n) \geq \langle \nabla f(w_n), x - w_n \rangle \tag{4.3.4}$$

and

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

Combining (4.3.2), (4.3.3), (4.3.4) and (4.3.5), we obtain

$$\begin{aligned}
& g(x) - g(x_{n+1}) + g(x) - g(y_n) + f(x) - f(y_n) + f(x) - f(w_n) \\
\geq \; & \big\langle \frac{y_n - x_{n+1}}{\lambda_n} - \nabla f(y_n), x - x_{n+1} \big\rangle + \big\langle \frac{w_n - y_n}{\lambda_n} - \nabla f(w_n), x - y_n \big\rangle \\
& + \langle \nabla f(y_n), x - y_n \rangle + \langle \nabla f(w_n), x - w_n \rangle \\
= \; & \frac{1}{\lambda_n} \langle y_n - x_{n+1}, x - x_{n+1} \rangle + \langle \nabla f(y_n), x_{n+1} - x \rangle + \frac{1}{\lambda_n} \langle w_n - y_n, x - y_n \rangle \\
& + \langle \nabla f(w_n), y_n - x \rangle + \langle \nabla f(y_n), x - y_n \rangle + \langle \nabla f(w_n), x - w_n \rangle \\
= \; & \frac{1}{\lambda_n} \langle y_n - x_{n+1}, x - x_{n+1} \rangle + \frac{1}{\lambda_n} \langle w_n - y_n, x - y_n \rangle + \langle \nabla f(y_n), x_{n+1} - y_n \rangle \\
& + \langle \nabla f(w_n), y_n - w_n \rangle
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\lambda_n}\langle y_n - x_{n+1}, x - x_{n+1}\rangle + \frac{1}{\lambda_n}\langle w_n - y_n, x - y_n\rangle \\
&\quad + \langle \nabla f(y_n) - \nabla f(x_{n+1}) + \nabla f(x_{n+1}), x_{n+1} - y_n\rangle \\
&\quad + \langle \nabla f(w_n) - \nabla f(y_n) + \nabla f(y_n), y_n - w_n\rangle \\
&= \frac{1}{\lambda_n}\langle y_n - x_{n+1}, x - x_{n+1}\rangle + \frac{1}{\lambda_n}\langle w_n - y_n, x - y_n\rangle \\
&\quad + \langle \nabla f(y_n) - \nabla f(x_{n+1}), x_{n+1} - y_n\rangle + \langle \nabla f(x_{n+1}), x_{n+1} - y_n\rangle \\
&\quad + \langle \nabla f(w_n) - \nabla f(y_n), y_n - w_n\rangle + \langle \nabla f(y_n), y_n - w_n\rangle \\
&\geq \frac{1}{\lambda_n}[\langle y_n - x_{n+1}, x - x_{n+1}\rangle + \langle w_n - y_n, x - y_n\rangle] \\
&\quad - \|\nabla f(y_n) - \nabla f(x_{n+1})\|\|x_{n+1} - y_n\| - \|\nabla f(w_n) - \nabla f(y_n)\|\|y_n - w_n\| \\
&\quad + f(x_{n+1}) - f(y_n) + f(y_n) - f(w_n) \\
&\geq \frac{1}{\lambda_n}[\langle y_n - x_{n+1}, x - x_{n+1}\rangle + \langle w_n - y_n, x - y_n\rangle] \\
&\quad - \|\nabla f(y_n) - \nabla f(x_{n+1})\|(\|y_n - w_n\| + \|x_{n+1} - y_n\|) \\
&\quad - \|\nabla f(w_n) - \nabla f(y_n)\|(\|y_n - w_n\| + \|x_{n+1} - y_n\|) + f(x_{n+1}) - f(w_n) \\
&= \frac{1}{\lambda_n}[\langle y_n - x_{n+1}, x - x_{n+1}\rangle + \langle w_n - y_n, x - y_n\rangle] \\
&\quad - (\|\nabla f(y_n) - \nabla f(x_{n+1})\| + \|\nabla f(w_n) - \nabla f(y_n)\|)(\|y_n - w_n\| + \|x_{n+1} - y_n\|) \\
&\quad + f(x_{n+1}) - f(w_n).
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\frac{1}{\lambda_n}[\langle y_n - x_{n+1}, x_{n+1} - x\rangle + \langle w_n - y_n, y_n - x\rangle] \\
\geq\ & g(x_{n+1}) - g(x) + g(y_n) - g(x) + f(y_n) - f(x) + f(w_n) - f(x) \\
&+ f(x_{n+1}) - f(w_n) - (\|\nabla f(y_n) - \nabla f(x_{n+1})\| + \|\nabla f(w_n) - \nabla f(y_n)\|) \\
&\times(\|y_n - w_n\| + \|x_{n+1} - y_n\|).
\end{aligned} \tag{4.3.6}
$$

From (4.3.6), it follows that

$$\langle y_n - x_{n+1}, x_{n+1} - x\rangle + \langle w_n - y_n, y_n - x\rangle$$

$$
\begin{aligned}
&\geq \lambda_n[(f+g)(x_{n+1})-(f+g)(x)+(f+g)(y_n)-(f+g)(x)] \\
&\quad -\lambda_n(\|\nabla f(y_n)-\nabla f(x_{n+1})\|+\|\nabla f(w_n)-\nabla f(y_n)\|) \\
&\quad \times(\|y_n-w_n\|+\|x_{n+1}-y_n\|) \\
&\geq \lambda_n[(f+g)(x_{n+1})-(f+g)(x)+(f+g)(y_n)-(f+g)(x)] \\
&\quad -\frac{\delta}{2}(\|x_{n+1}-y_n\|+\|w_n-y_n\|)(\|y_n-w_n\|+\|x_{n+1}-y_n\|) \\
&\geq \lambda_n[(f+g)(x_{n+1})-(f+g)(x)+(f+g)(y_n)-(f+g)(x)] \\
&\quad -\frac{\delta}{2}(\|x_{n+1}-y_n\|^2+\|w_n-y_n\|^2). \qquad (4.3.7)
\end{aligned}
$$

We know that

$$
2\langle y_n-x_{n+1},x_{n+1}-x\rangle=\|y_n-x\|^2-\|y_n-x_{n+1}\|^2-\|x_{n+1}-x\|^2 \qquad (4.3.8)
$$

and

$$
2\langle w_n-y_n,y_n-x\rangle=\|w_n-x\|^2-\|w_n-y_n\|^2-\|y_n-x\|^2. \qquad (4.3.9)
$$

From (4.3.7), (4.3.8) and (4.3.9), we get

$$
\begin{aligned}
&-\frac{1}{2}\|y_n-x_{n+1}\|^2-\frac{1}{2}\|x_{n+1}-x\|^2+\frac{1}{2}\|w_n-x\|^2-\frac{1}{2}\|w_n-y_n\|^2 \\
&\geq \lambda_n[(f+g)(x_{n+1})-(f+g)(x)+(f+g)(y_n)-(f+g)(x)] \\
&\quad -\frac{\delta}{2}(\|x_{n+1}-y_n\|^2+\|w_n-y_n\|^2).
\end{aligned}
$$

This shows that

$$
\begin{aligned}
&\|w_n-x\|^2-\|x_{n+1}-x\|^2 \\
&\geq 2\lambda_n[(f+g)(x_{n+1})-(f+g)(x)+(f+g)(y_n)-(f+g)(x)] \\
&\quad +(1-\delta)\|x_{n+1}-y_n\|^2+(1-\delta)\|w_n-y_n\|^2.
\end{aligned}
$$

Setting $x = z$, we obtain

$$\begin{aligned}
\|x_{n+1} - z\|^2 &\leq \|w_n - z\|^2 - 2\lambda_n[(f+g)(x_{n+1}) - (f+g)(z) \\
&\quad +(f+g)(y_n) - (f+g)(z)] - (1-\delta)\|x_{n+1} - y_n\|^2 \\
&\quad -(1-\delta)\|w_n - y_n\|^2 \\
&\leq \|w_n - z\|^2. \qquad (4.3.10)
\end{aligned}$$

Hence

$$\begin{aligned}
\|x_{n+1} - z\| &\leq \|w_n - z\| \\
&= \|x_n + \theta_n(x_n - x_{n-1}) - z\| \\
&\leq \|x_n - z\| + \theta_n\|x_n - x_{n-1}\| \\
&\leq \|x_n - z\| + \theta_n(\|x_n - z\| + \|x_{n-1} - z\|).
\end{aligned}$$

This shows that $\|x_{n+1} - z\| \leq (1+\theta_n)\|x_n - z\| + \theta_n\|x_{n-1} - z\|$. By Lemma 3.2.4, we get

$$\|x_{n+1} - z\| \leq K \cdot \prod_{i=1}^{n}(1 + 2\theta_i),$$

where $K = \max\{\|x_1 - z\|, \|x_2 - z\|\}$. Since $\sum_{n=1}^{\infty} \theta_n < +\infty$ and $\{x_n\}$ is bounded, we have $\sum_{n=1}^{\infty} \theta_n\|x_n - x_{n-1}\| < +\infty$. By Lemma 3.2.3, we also have $\lim_{n\to\infty} \|x_n - z\|$ exists.

Consider

$$\begin{aligned}
\|w_n - z\|^2 &= \|x_n + \theta_n(x_n - x_{n-1}) - z\|^2 \\
&= \|x_n - z\|^2 + 2\theta_n\langle x_n - z, x_n - x_{n-1}\rangle + \theta_n^2\|x_n - x_{n-1}\|^2 \\
&\leq \|x_n - z\|^2 + 2\theta_n\|x_n - z\|\|x_n - x_{n-1}\| \\
&\quad +\theta_n^2\|x_n - x_{n-1}\|^2. \qquad (4.3.11)
\end{aligned}$$

From (4.3.10) and (4.3.11), we obtain

$$\begin{aligned}\|x_{n+1}-z\|^2 &\le \|x_n-z\|^2+2\theta_n\|x_n-z\|\|x_n-x_{n-1}\|+\theta_n^2\|x_n-x_{n-1}\|^2\\ &\quad -2\lambda_n[(f+g)(x_{n+1})-(f+g)(z)+(f+g)(y_n)-(f+g)(z)]\\ &\quad -(1-\delta)\|x_{n+1}-y_n\|^2-(1-\delta)\|w_n-y_n\|^2\\ &\le \|x_n-z\|^2+2\theta_n\|x_n-z\|\|x_n-x_{n-1}\|+\theta_n^2\|x_n-x_{n-1}\|^2\\ &\quad -(1-\delta)\|x_{n+1}-y_n\|^2-(1-\delta)\|w_n-y_n\|^2. \end{aligned} \tag{4.3.12}$$

Since $\sum_{n=1}^{\infty}\theta_n\|x_n-x_{n-1}\|<+\infty$ and $\lim_{n\to\infty}\|x_n-z\|$ exists. Hence, by (4.3.12), $\|y_n-w_n\|\to 0$ and $\|x_{n+1}-y_n\|\to 0$. From (4.3.1), it is easy to see that $\|x_n-w_n\|\to 0$.

Moreover, we have

$$\begin{aligned}\|x_n-y_n\| &\le \|x_n-w_n\|+\|w_n-y_n\|\\ &\to 0 \text{ (as } n\to\infty). \end{aligned} \tag{4.3.13}$$

From (4.3.13), we obtain

$$\begin{aligned}\|x_n-x_{n+1}\| &\le \|x_n-y_n\|+\|y_n-x_{n+1}\|\\ &\to 0 \text{ (as } n\to\infty). \end{aligned}$$

Since $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k}\rightharpoonup z^*\in H$. Moreover, we obtain $x_{n_k+1}\rightharpoonup z^*$. Since $\{x_{n_k}\}$ is bounded and $\|x_{n_k+1}-y_{n_k}\|\to 0$, by assumption (II), we obtain

$$\lim_{k\to\infty}\|\nabla f(x_{n_k+1})-\nabla f(y_{n_k})\|=0. \tag{4.3.14}$$

Note that

$$x_{n_k+1} = \text{prox}_{\lambda_{n_k} g}(y_{n_k} - \lambda_{n_k} \nabla f(y_{n_k})).$$

It follows from (3.1.2) that

$$\frac{y_{n_k} - \lambda_{n_k} \nabla f(y_{n_k}) - x_{n_k+1}}{\lambda_{n_k}} \in \partial g(x_{n_k+1}).$$

Hence

$$\frac{y_{n_k} - x_{n_k+1}}{\lambda_{n_k}} + \nabla f(x_{n_k+1}) - \nabla f(y_{n_k}) \in \nabla f(x_{n_k+1}) + \partial g(x_{n_k+1}).$$

Letting $k \to \infty$ in (4.3.15), from (4.3.14) and Lemma 3.2.2, we get

$$0 \in \partial(f+g)(z^*).$$

Hence $z^* \in \Omega$. By Lemma 3.2.5, we conclude that $\{x_n\}$ weakly converges to a point in Ω. This completes the proof. □

Remark 4.3.3 The condition that $\{\lambda_n\}$ is bounded away from 0 can be removed when ∇f is Lipschitz continuous. Indeed, if ∇f is L-Lipschitz continuous on H, then $\lambda_n \leq \sigma$, $n \geq 1$. If $\lambda_n < \sigma$, then $\lambda_n = \sigma \rho^{m_n}$ where m_n is the smallest positive integer such that

$$\begin{aligned} &\lambda_n(\|\nabla f(prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n))) - \nabla f(y_n)\| + \|\nabla f(w_n) - \nabla f(y_n)\|) \\ &\leq \frac{\delta}{2}(\|prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n)) - y_n\| + \|w_n - y_n\|). \end{aligned} \tag{4.3.15}$$

Set $\widehat{\lambda_n} = \frac{\lambda_n}{\rho}$. By the Lipschitz continuity of ∇f and (4.3.15), we have

$$\begin{aligned} &\widehat{\lambda_n} L(\|prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n)) - y_n\| + \|w_n - y_n\|) \\ \geq\ &\widehat{\lambda_n}(\|\nabla f(prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n))) - \nabla f(y_n)\| + \|\nabla f(w_n) - \nabla f(y_n)\|) \end{aligned}$$

$$> \frac{\delta}{2}(\|prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n)) - y_n\| + \|w_n - y_n\|).$$

So $\widehat{\lambda_n} L > \frac{\delta}{2}$. It follows that $\lambda_n > \frac{\delta\rho}{2L}$. Hence $\lambda_n \geq \min\{\sigma, \frac{\delta\rho}{2L}\}$ for all $n \geq 1$.

4.4 Double proximal gradient method with new linesearch

In this section, we assume that the following conditions are satisfied for our convergence analysis:

(A1) The solution set of the convex minimization problem (2.1.1) is nonempty, *i.e.*, $\Omega = \operatorname{argmin}(f + g) \neq \emptyset$.

(A2) $f, g : H \to (-\infty, +\infty]$ are two proper, lower semicontinuous and convex functions.

(A3) The gradient ∇f is uniformly continuous on bounded subset of H and maps any bounded subset of H.

We next introduce an inertial double forward-backward method for solving (2.1.1).

Algorithm 4.4.1 An inertial double proximal forward-backward method (IDFB)

Initialization: *Let* $x_0 = x_1 \in H$, $\theta_1 > 0$, $\gamma > 0$, $\ell \in (0, 1)$ *and* $0 < \mu < 1$.

Iterative step: *For* $n \geq 1$, *calculate* x_{n+1} *as follows:*

Step 1. *Compute the inertial step:*

$$w_n = x_n + \theta_n(x_n - x_{n-1}).$$

Step 2. *Compute the forward-backward step:*

$$y_n = prox_{\lambda_n g}(w_n - \lambda_n \nabla f(w_n)).$$

Step 3. *Compute the x_{n+1} step:*

$$x_{n+1} = prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n)),$$

where the linesearch $\lambda_n = \gamma \ell^{m_n}$ is the smallest nonegative integer such that

$$\begin{aligned} & \lambda_n(\langle \nabla f(x_{n+1}) - \nabla f(y_n), x_{n+1} - y_n\rangle + \langle \nabla f(y_n) - \nabla f(w_n), y_n - w_n\rangle) \\ \leq \ & \frac{\mu^2+1}{4}\|x_{n+1} - y_n\|^2 + \frac{\mu}{\mu+1}\|y_n - w_n\|^2. \qquad (4.4.1) \end{aligned}$$

Set $n = n+1$ and return to **Step 1**.

Lemma 4.4.2 *Let $x \in H$, $\gamma > 0$, $\ell \in (0,1)$ and $0 < \mu < 1$. For $i = 1, 2, 3, ...,$ set*

$$\begin{aligned} H(x,i) &= \text{prox}_{\gamma \ell^i g}(x - \gamma \ell^i \nabla f(x)) \\ P(x,i) &= \text{prox}_{\gamma \ell^i g}(H(x,i) - \gamma \ell^i \nabla f(H(x,i))). \end{aligned}$$

If

$$\begin{aligned} & \gamma \ell^i(\langle \nabla f(P(x,i)) - \nabla f(H(x,i)), P(x,i) - H(x,i)\rangle \\ & + \langle \nabla f(H(x,i)) - \nabla f(x), H(x,i) - x\rangle) \\ \leq \ & \frac{\mu^2+1}{4}\|P(x,i) - H(x,i)\|^2 + \frac{\mu}{\mu+1}\|H(x,i) - x\|^2, \end{aligned}$$

then $\lambda = \gamma \ell^i$.

Else $i = i+1$. The linesearch (4.4.1) stops after finitely many steps.

Proof. If $x \in \Omega$, then $x = \text{prox}_{\gamma g}(x - \gamma \nabla f(x)) = H(x,0)$. It follows that $H(x,0) = x$ and the linesearch stops with zero step, hence $\lambda = \gamma$.

If $x \notin \Omega$, then

$$\gamma \ell^i(\langle \nabla f(P(x,i)) - \nabla f(H(x,i)), P(x,i) - H(x,i)\rangle$$

$$
\begin{aligned}
&+\langle \nabla f(H(x,i)) - \nabla f(x), H(x,i) - x\rangle) \\
> \ & \frac{\mu^2+1}{4}\|P(x,i) - H(x,i)\|^2 + \frac{\mu}{\mu+1}\|H(x,i) - x\|^2,
\end{aligned}
$$

it follows that

$$
\begin{aligned}
&\gamma\ell^i(\|\nabla f(P(x,i)) - \nabla f(H(x,i))\|\|P(x,i) - H(x,i)\| \\
&+\|\nabla f(H(x,i)) - \nabla f(x)\|\|H(x,i) - x\|) \\
> \ & \frac{\mu^2+1}{4}\|P(x,i) - H(x,i)\|^2 + \frac{\mu}{\mu+1}\|H(x,i) - x\|^2. \qquad (4.4.2)
\end{aligned}
$$

So we have as $i \to \infty$, $\|P(x,i) - H(x,i)\| \to 0$ and $\|H(x,i) - x\| \to 0$. Since ∇f is uniformly continuous, we get $\|\nabla f(P(x,i)) - \nabla f(H(x,i))\| \to 0$ and $\|\nabla f(H(x,i)) - \nabla f(x)\| \to 0$ as $i \to \infty$. By (4.4.2), we have

$$
\frac{\|H(x,i) - x\|}{\gamma\ell^i} \to 0 \text{ as } i \to \infty.
$$

We see that

$$
\frac{x - \gamma\ell^i\nabla f(x) - H(x,i)}{\gamma\ell^i} \in \partial g(H(x,i)).
$$

Hence,

$$
\frac{x - H(x,i)}{\gamma\ell^i} \in \partial g(H(x,i)) + \nabla f(x).
$$

By Lemma 3.2.2, we have $0 \in \partial g(x) + \nabla f(x)$. Thus, $x \in \Omega$ which is a contradiction. This completes the proof. □

Theorem 4.4.3 *Let $\{x_n\}$ be generated by Algorithm 4.4.1. If $\sum_{n=1}^{\infty}\theta_n < \infty$ and $\lambda_n \geq \lambda$ for some $\lambda > 0$, then $\{x_n\}$ weakly converges to point in Ω.*

Proof. By definition of proximal operator, we have

$$\begin{aligned} \frac{w_n - y_n}{\lambda_n} - \nabla f(w_n) &= \frac{w_n - prox_{\lambda_n g}(w_n - \lambda_n \nabla f(w_n))}{\lambda_n} - \nabla f(w_n) \\ &\in \partial g(y_n). \end{aligned}$$

By property of the convexity of g, we obtain

$$g(x) - g(y_n) \geq \langle \frac{w_n - y_n}{\lambda_n} - \nabla f(w_n), x - y_n \rangle, \ \forall x \in H. \tag{4.4.3}$$

Also, we have

$$\begin{aligned} \frac{y_n - x_{n+1}}{\lambda_n} - \nabla f(y_n) &= \frac{y_n - prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n))}{\lambda_n} - \nabla f(y_n) \\ &\in \partial g(x_{n+1}), \end{aligned}$$

by the convexity of g, we also have

$$g(x) - g(x_{n+1}) \geq \langle \frac{y_n - x_{n+1}}{\lambda_n} - \nabla f(y_n), x - x_{n+1} \rangle, \ \forall x \in H. \tag{4.4.4}$$

For any $x \in H$, we also have

$$f(x) - f(w_n) \geq \langle \nabla f(w_n), x - w_n \rangle \tag{4.4.5}$$

and

$$f(x) - f(y_n) \geq \langle \nabla f(y_n), x - y_n \rangle. \tag{4.4.6}$$

Using (3.1.3) and (4.4.1) and combining (4.4.3), (4.4.4), (4.4.5) and (4.4.6), we obtain

$$g(x) - g(x_{n+1}) + g(x) - g(y_n) + f(x) - f(w_n) + f(x) - f(y_n)$$

$$
\begin{aligned}
&\geq \langle \frac{y_n - x_{n+1}}{\lambda_n} - \nabla f(y_n), x - x_{n+1}\rangle + \langle \frac{w_n - y_n}{\lambda_n} - \nabla f(w_n), x - y_n\rangle \\
&\quad +\langle \nabla f(w_n), x - w_n\rangle + \langle \nabla f(y_n), x - y_n\rangle \\
&= \frac{1}{\lambda_n}\langle y_n - x_{n+1}, x - x_{n+1}\rangle + \langle \nabla f(y_n), x_{n+1} - x\rangle + \frac{1}{\lambda_n}\langle w_n - y_n, x - y_n\rangle \\
&\quad +\langle \nabla f(w_n), y_n - x\rangle + \langle \nabla f(w_n), x - w_n\rangle + \langle \nabla f(y_n), x - y_n\rangle \\
&= \frac{1}{\lambda_n}\langle y_n - x_{n+1}, x - x_{n+1}\rangle + \frac{1}{\lambda_n}\langle w_n - y_n, x - y_n\rangle + \langle \nabla f(y_n), x_{n+1} - y_n\rangle \\
&\quad +\langle \nabla f(w_n), y_n - w_n\rangle \\
&= \frac{1}{\lambda_n}\langle y_n - x_{n+1}, x - x_{n+1}\rangle + \frac{1}{\lambda_n}\langle w_n - y_n, x - y_n\rangle \\
&\quad +\langle \nabla f(y_n) - \nabla f(x_{n+1}) + \nabla f(x_{n+1}), x_{n+1} - y_n\rangle \\
&\quad +\langle \nabla f(w_n) - \nabla f(y_n) + \nabla f(y_n), y_n - w_n\rangle \\
&= \frac{1}{\lambda_n}\langle y_n - x_{n+1}, x - x_{n+1}\rangle + \frac{1}{\lambda_n}\langle w_n - y_n, x - y_n\rangle \\
&\quad +\langle \nabla f(y_n) - \nabla f(x_{n+1}), x_{n+1} - y_n\rangle + \langle \nabla f(x_{n+1}), x_{n+1} - y_n\rangle \\
&\quad +\langle \nabla f(w_n) - \nabla f(y_n), y_n - w_n\rangle + \langle \nabla f(y_n), y_n - w_n\rangle \\
&= \frac{1}{\lambda_n}[\langle y_n - x_{n+1}, x - x_{n+1}\rangle + \langle w_n - y_n, x - y_n\rangle] \\
&\quad -\langle \nabla f(x_{n+1}) - \nabla f(y_n), x_{n+1} - y_n\rangle - \langle \nabla f(y_n) - \nabla f(w_n), y_n - w_n\rangle \\
&\quad +\langle \nabla f(x_{n+1}), x_{n+1} - y_n\rangle + \langle \nabla f(y_n), y_n - w_n\rangle \\
&= \frac{1}{\lambda_n}[\langle y_n - x_{n+1}, x - x_{n+1}\rangle + \langle w_n - y_n, x - y_n\rangle] \\
&\quad -[\langle \nabla f(x_{n+1}) - \nabla f(y_n), x_{n+1} - y_n\rangle + \langle \nabla f(y_n) - \nabla f(w_n), y_n - w_n\rangle] \\
&\quad +\langle \nabla f(x_{n+1}), x_{n+1} - y_n\rangle + \langle \nabla f(y_n), y_n - w_n\rangle \\
&\geq \frac{1}{\lambda_n}[\langle y_n - x_{n+1}, x - x_{n+1}\rangle + \langle w_n - y_n, x - y_n\rangle] \\
&\quad -[\frac{\mu^2+1}{4\lambda_n}\|x_{n+1} - y_n\|^2 + \frac{\mu}{(\mu+1)\lambda_n}\|y_n - w_n\|^2] \\
&\quad +f(x_{n+1}) - f(y_n) + f(y_n) - f(w_n).
\end{aligned}
$$

It follows that

$$
2\langle y_n - x_{n+1}, x_{n+1} - x\rangle + 2\langle w_n - y_n, y_n - x\rangle
$$

$$\begin{aligned}
&\geq 2\lambda_n[g(x_{n+1}) - g(x) + g(y_n) - g(x) - f(x) + f(y_n) - f(x) + f(x_{n+1})] \\
&\quad -2\lambda_n[\frac{\mu^2+1}{4\lambda_n}\|x_{n+1} - y_n\|^2 + \frac{\mu}{(\mu+1)\lambda_n}\|y_n - w_n\|^2] \\
&= 2\lambda_n[(f+g)(x_{n+1}) - (f+g)(x) + (f+g)(y_n) - (f+g)(x)] \\
&\quad -[\frac{\mu^2+1}{2}\|x_{n+1} - y_n\|^2 + \frac{2\mu}{\mu+1}\|y_n - w_n\|^2]. \qquad (4.4.7)
\end{aligned}$$

We have

$$2\langle y_n - x_{n+1}, x_{n+1} - x\rangle = \|y_n - x\|^2 - \|y_n - x_{n+1}\|^2 - \|x_{n+1} - x\|^2, \qquad (4.4.8)$$

and

$$2\langle w_n - y_n, y_n - x\rangle = \|w_n - x\|^2 - \|w_n - y_n\|^2 - \|y_n - x\|^2. \qquad (4.4.9)$$

Substituting (4.4.8) and (4.4.9) into (4.4.7), we obtain

$$\begin{aligned}
&-\|y_n - x_{n+1}\|^2 - \|x_{n+1} - x\|^2 + \|w_n - x\|^2 - \|w_n - y_n\|^2 \\
\geq\ & 2\lambda_n[(f+g)(x_{n+1}) - (f+g)(x) + (f+g)(y_n) - (f+g)(x)] \\
&-[\frac{\mu^2+1}{2}\|x_{n+1} - y_n\|^2 + \frac{2\mu}{\mu+1}\|y_n - w_n\|^2].
\end{aligned}$$

It follows that

$$\begin{aligned}
&\|x_{n+1} - x\|^2 \\
\leq\ & \|w_n - x\|^2 - \|y_n - x_{n+1}\|^2 - \|w_n - y_n\|^2 \\
&-2\lambda_n[(f+g)(x_{n+1}) - (f+g)(x) + (f+g)(y_n) - (f+g)(x)] \\
&+\frac{\mu^2+1}{2}\|x_{n+1} - y_n\|^2 + \frac{2\mu}{\mu+1}\|y_n - w_n\|^2 \\
=\ & \|w_n - x\|^2 - (1 - \frac{\mu^2+1}{2})\|y_n - x_{n+1}\|^2 - (1 - \frac{2\mu}{\mu+1})\|w_n - y_n\|^2 \\
&-2\lambda_n[(f+g)(x_{n+1}) - (f+g)(x) + (f+g)(y_n) - (f+g)(x)].
\end{aligned}$$

Setting $x = z \in \Omega$ and using $0 < \mu < 1$, we have

$$\begin{aligned}\|x_{n+1} - z\|^2 &= \|w_n - z\|^2 - (1 - \frac{\mu^2+1}{2})\|y_n - x_{n+1}\|^2 - (1 - \frac{2\mu}{\mu+1})\|w_n - y_n\|^2 \\ &\quad -2\lambda_n[(f+g)(x_{n+1}) - (f+g)(z) + (f+g)(y_n) - (f+g)(z)] \\ &\leq \|w_n - z\|^2. \end{aligned} \tag{4.4.10}$$

So,

$$\begin{aligned}\|x_{n+1} - z\| &\leq \|w_n - z\| \\ &= \|x_n + \theta_n(x_n - x_{n-1}) - z\| \\ &\leq \|x_n - z\| + \theta\|x_n - x_{n-1}\| \\ &\leq \|x_n - z\| + \theta_n(\|x_n - z\| + \|x_{n-1} - z\|), \end{aligned} \tag{4.4.11}$$

which gives $\|x_{n+1} - z\| \leq (1+\theta_n)\|x_n - z\| + \theta_n\|x_{n-1} - z\|$. By Lemma 3.2.4, we obtain

$$\|x_{n+1} - z\| \leq K \cdot \prod_{i=1}^{n}(1 + 2\theta_i),$$

where $K = \max\{\|x_1 - z\|, \|x_2 - z\|\}$. By Lemma 3.2.4 and $\sum_{n=1}^{\infty}\theta_n < +\infty$, we obtain $\{x_n\}$ is bounded. So

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

By Lemma 3.2.3 and (4.4.11), we obtain $\lim_{n\to\infty}\|x_n - z\|$ exists.

Consider,

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

$$\leq \|x_n - z\|^2 + 2\theta_n\|x_n - z\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2. \quad (4.4.12)$$

From (4.4.10) and (4.4.12), we have

$$\begin{aligned}
\|x_{n+1} - z\|^2 &\leq \|x_n - z\|^2 + 2\theta_n\|x_n - z\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2 \\
&\quad -(1 - \frac{\mu^2+1}{2})\|y_n - x_{n+1}\|^2 - (1 - \frac{2\mu}{\mu+1})\|w_n - y_n\|^2 \\
&\quad -2\lambda_n[(f+g)(x_{n+1}) - (f+g)(z) + (f+g)(y_n) - (f+g)(z)] \\
&\leq \|x_n - z\|^2 + 2\theta_n\|x_n - z\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2 \\
&\quad -(1 - \frac{\mu^2+1}{2})\|y_n - x_{n+1}\|^2 - (1 - \frac{2\mu}{\mu+1})\|w_n - y_n\|^2. \quad (4.4.13)
\end{aligned}$$

Since $\sum_{n=1}^{\infty} \theta_n\|x_n - x_{n-1}\| < +\infty$ and $\lim_{n\to\infty}\|x_n - z\|$ exists, from (4.4.13), we obtain

$$\lim_{n\to\infty}\|y_n - x_{n+1}\| = 0 \quad (4.4.14)$$

and

$$\lim_{n\to\infty}\|w_n - y_n\| = 0. \quad (4.4.15)$$

From definition of w_n, we have

$$\lim_{n\to\infty}\|x_n - w_n\| = 0. \quad (4.4.16)$$

From (4.4.15) and (4.4.16), we see that

$$\begin{aligned}
\lim_{n\to\infty}\|x_n - y_n\| &\leq \lim_{n\to\infty}\|x_n - w_n\| + \lim_{n\to\infty}\|w_n - y_n\| \\
&= 0. \quad (4.4.17)
\end{aligned}$$

From (4.4.14) and (4.4.17), we obtain

$$\begin{aligned}\lim_{n\to\infty}\|x_{n+1}-x_n\| &\leq \lim_{n\to\infty}\|x_{n+1}-y_n\| + \lim_{n\to\infty}\|y_n - x_n\| \\ &= 0.\end{aligned}$$

Since $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \rightharpoonup \bar{z} \in H$. Moreover, we obtain $x_{n_k+1} \rightharpoonup \bar{z}$. Since $\{x_{n_k}\}$ is bounded, $\lim_{k\to\infty}\|x_{n_k+1} - y_{n_k}\| = 0$ and ∇f is uniformly continuous, we have

$$\lim_{k\to\infty}\|\nabla f(x_{n_k+1}) - \nabla f(y_{n_k})\| = 0. \tag{4.4.18}$$

From

$$x_{n_k+1} = prox_{\lambda_{n_k} g}(y_{n_k} - \lambda_{n_k}\nabla f(y_{n_k})),$$

it follows that

$$\frac{y_{n_k} - \lambda_{n_k}\nabla f(y_{n_k}) - x_{n_k+1}}{\lambda_{n_k}} \in \partial g(x_{n_k+1}).$$

Hence,

$$\frac{y_{n_k} - x_{n_k+1}}{\lambda_{n_k}} + \nabla f(x_{n_k+1}) - \nabla f(y_{n_k}) \in \nabla f(x_{n_k+1}) + \partial g(x_{n_k+1}). \tag{4.4.19}$$

Using (4.4.18), letting $k \to \infty$ in (4.4.19) and applying Lemma 3.2.2, we get

$$0 \in (\nabla f + \partial g)(\bar{z}).$$

Hence $\bar{z} \in \Omega$. From (4.4.13) and Definition 3.1.31, we have $\{x_n\}$ is a quasi- Fej r sequence. By Lemma 3.2.5, we conclude that $\{x_n\}$ weakly converges to a point in Ω. This completes the proof. □

CHAPTER V

APPLICATION

5.1 Application to image deblurring

This section is devoted to the presentation of numerical experiments which illustrate the application to image deblurring of an inertial modified forward-backward method (IMFB) and new proximal gradient method (Agorithm 4.3.1) in Section 4.1 and 4.3.

Let $A \in \mathbb{R}^{N\times M}$ be the blurring matrix, $x \in \mathbb{R}^N$ the original image and $b \in \mathbb{R}^M$ the degraded image. Adding the random noise $v \in \mathbb{R}^M$, we have the following image recovery problem

$$Ax = b + v.$$

For solving this problem, we make use of the model of Tibshiranit [49] which is known as LASSO problem

$$\min_{x\in\mathbb{R}^N} \left\{ \frac{1}{2}\|b - Ax\|_2^2 + \tau\|x\|_1 \right\}, \tag{5.1.1}$$

where τ is a positive parameter, $\|\cdot\|_1$ is the ℓ_1-norm, and $\|\cdot\|_2$ is the Euclidean norm.

In this connection, problem (5.1.1) can be seen as (2.1.1) by setting

$$f(x) = \frac{1}{2}\|b - Ax\|_2^2 \text{ and } g(x) = \tau\|x\|_1.$$

Peak-signal-to-noise ratio (PSNR) in decibel (dB) [50] is defined by

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right)$$

where MSE= $\|x_n - x\|^2$. Structural similarity index measure (SSIM) [54] is used for measuring the similarity between two images. It is noted that, a higher PSNR generally indicates that the reconstruction is of higher quality. The resultant SSIM index is a decimal value between 0 and 1, and value 1 is indicates perfect structural similarity. We use a Fast Fourier Transform (FFT) for converting it to the frequency domain.

The numerical experiments have been carried out in Matlab environment (version R2020b) on MacBook Pro M1 with ram 8 GB.

We will present restoration of images corrupted by the following blur types:

- Motion blur with motion length of 45 pixels and motion orientation 180°.
- Gaussian blur of filter size 5×5 with standard deviation 5.
- Out of focus blur with radius 7.

Example 5.1.1 Set a regularization $\tau = 10^{-5}$ and $x_0 = x_1 = (1, 1, 1, ..., 1) \in \mathbb{R}^N$.

For IFB, choose $\theta_n = \begin{cases} \frac{1}{n^2\|x_n - x_{n-1}\|_2^2} & \text{if } x_n \neq x_{n-1} \\ 0 & \text{otherwise.} \end{cases}$

For Algorithm 4.3.1, we set $t_0 = 1$, $t_n = \frac{1+\sqrt{1+4t_{n-1}^2}}{2}$ and

$$\theta_n = \begin{cases} \frac{t_{n-1}-1}{t_n}, & \text{if } 1 \leq n \leq 1000 \\ 0 & \text{otherwise.} \end{cases}$$

In the implementation, we choose the following parameters:

Table 1: The parameters for each methods in section 4.3

Methods	$\lambda_n = \frac{n}{L(n+1)}$	$\alpha_n = 0.5$	$\mu = 0.4$	$\rho = 0.5$	$\sigma = 2$	$\delta = 0.9$
IFB	✓	-	-	-	-	-
FISTA	✓	-	-	-	-	-
NAGA	✓	✓	-	-	-	-
FMFB	-	-	✓	✓	✓	-
Algorithm 4.3.1	-	-	-	✓	✓	✓

Here L was computed by the maximum eigenvalues of the matrix $A^T A$. The quality improvements of the reconstructed grey image sizes 320×266 and RGB image sizes 1000×744.

Numerical results for recovering the degraded grey and RGB images by using the proposed method with 1,000 iterations are reported in Table 2.

Table 2: The comparison of PSNR and SSIM of the restored images for grey images

Methods	Motion blur		Gaussian blur		Out of focus	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
IFB	26.5404	0.6317	48.9764	0.9945	34.6308	0.9322
FISTA	56.2778	0.9984	49.4321	0.9959	44.6754	0.9856
NAGA	59.2419	0.9992	50.3907	0.9966	45.2599	0.9871
FMFB	57.5154	0.9987	50.4814	0.9966	45.2917	0.9872
Algorithm 4.3.1	62.7602	0.9996	52.4786	0.9978	45.6101	0.9880

Table 3: The comparison of PSNR and SSIM of the restored images for RGB images

Methods	Motion blur		Gaussian blur		Out of focus	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
IFB	28.2037	0.8686	35.5792	0.9745	28.5151	0.8650
FISTA	39.6261	0.9871	45.1640	0.9931	38.4775	0.9823
NAGM	41.6376	0.9917	42.9277	0.9941	39.3708	0.9849
FMFB	41.6352	0.9918	42.9499	0.9941	39.4220	0.9850
Algorithm 4.3.1	45.6777	0.9966	44.5793	0.9957	40.9983	0.9889

From Tables 2 and 3, we observe that Algorithm 4.3.1 provides a higher PSNR and SSIM than those of IFB, FISTA, NAGA and FMFB methods. This shows that the proposed method has a better convergence behaviour than others.

We next show the restored images of Grey and RGB images for motion blur.

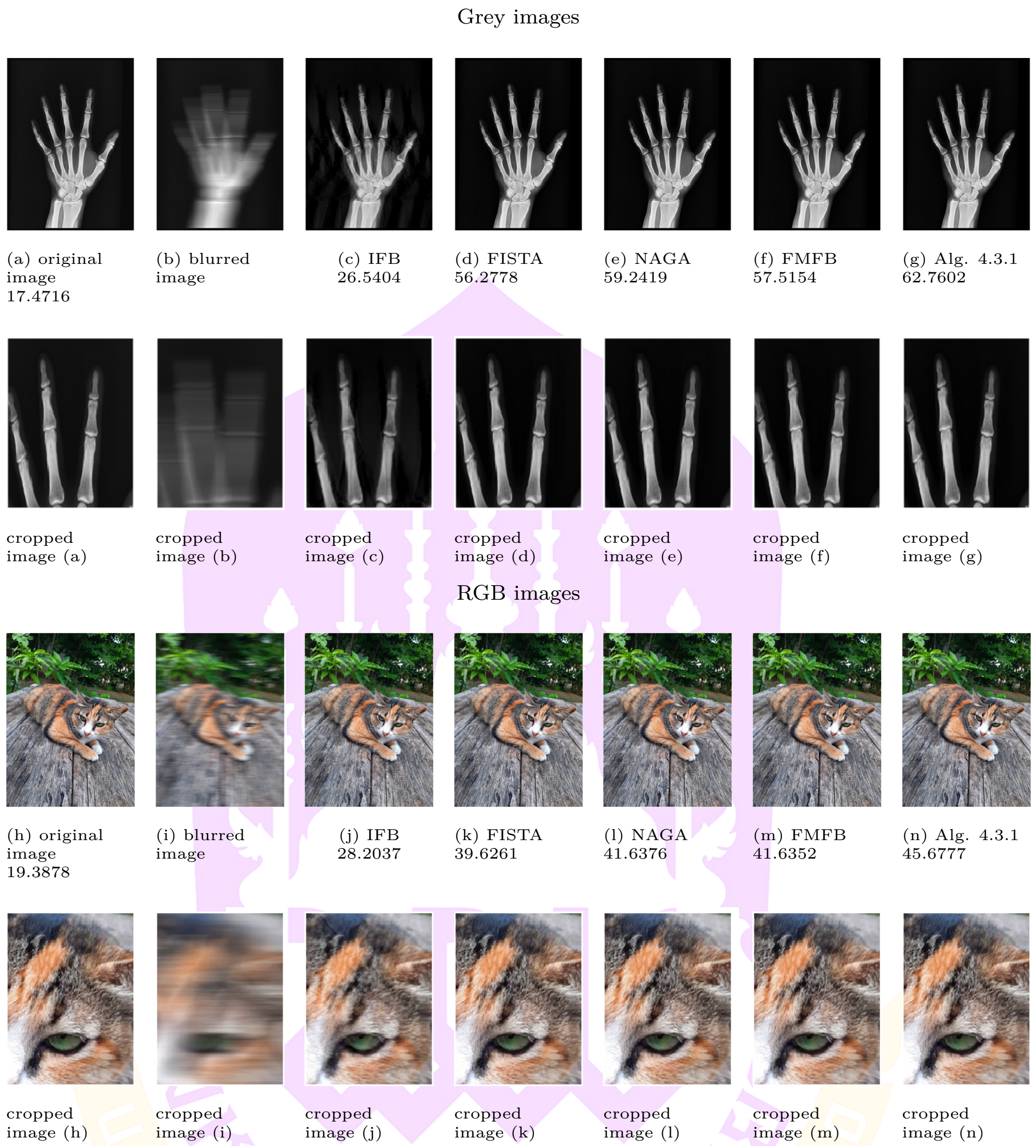


Figure 1: The original images, images degraded by motion blur and restored images for grey images and RGB images.

We next show the restored images of Grey and RGB images for Gaussian blur.

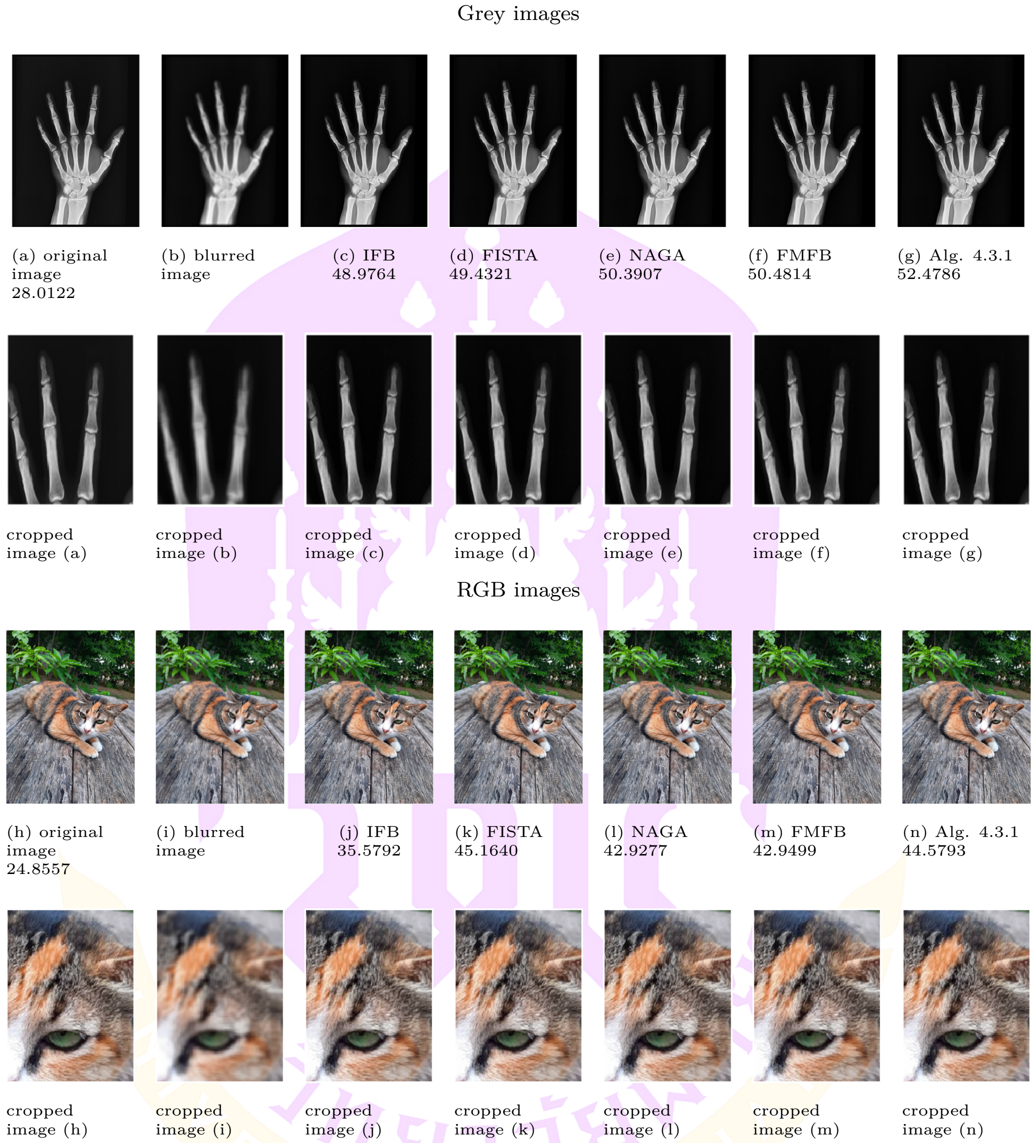
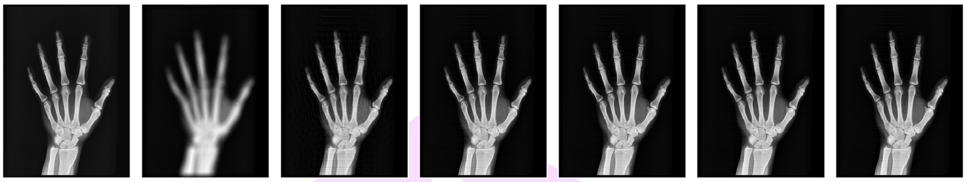


Figure 2: The original images, images degraded by Gaussian blur and restored images for grey images and RGB images.

We next show the restored images of Grey and RGB images for out of focus.

Grey images



(a) original image 23.3081 | (b) blurred image | (c) IFB 34.6308 | (d) FISTA 44.6754 | (e) NAGA 45.2599 | (f) FMFB 45.2917 | (g) Alg. 4.3.1 45.6101

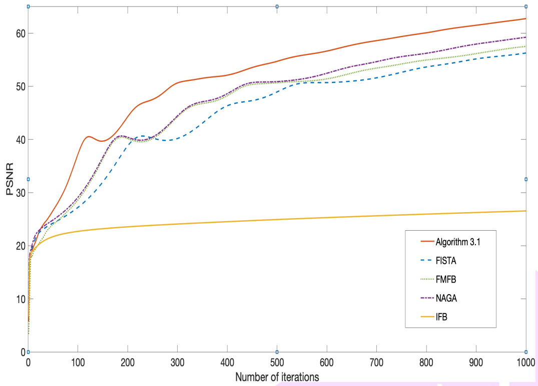
cropped image (a) | cropped image (b) | cropped image (c) | cropped image (d) | cropped image (e) | cropped image (f) | cropped image (g)

RGB images

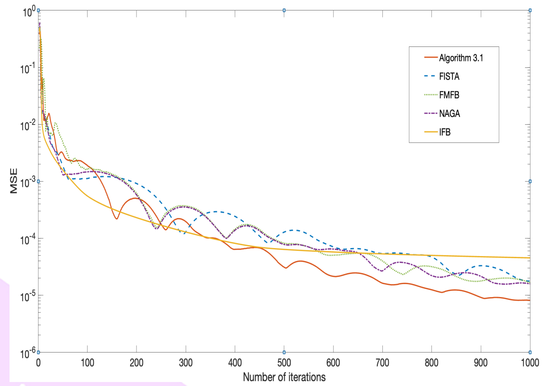
(h) original image 21.6416 | (i) blurred image | (j) IFB 28.5151 | (k) FISTA 38.4775 | (l) NAGA 39.3708 | (m) FMFB 39.4220 | (n) Alg. 4.3.1 40.9983

cropped image (h) | cropped image (i) | cropped image (j) | cropped image (k) | cropped image (l) | cropped image (m) | cropped image (n)

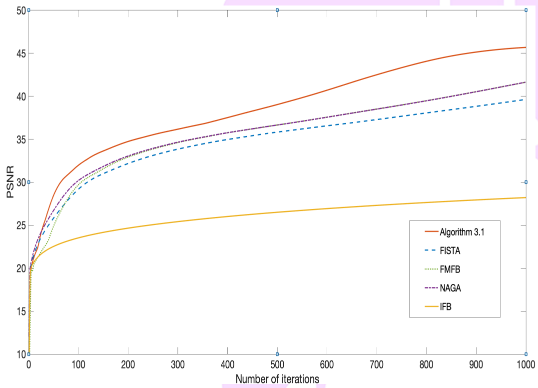
Figure 3: The original images, images degraded for out of focus and restored images for grey images and RGB images.



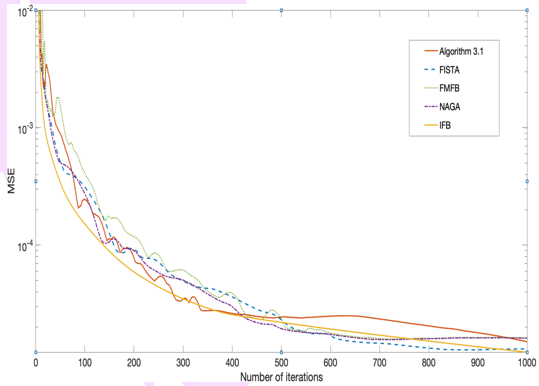
(a) PSNR plotting of grey images



(b) MSE plotting of grey images

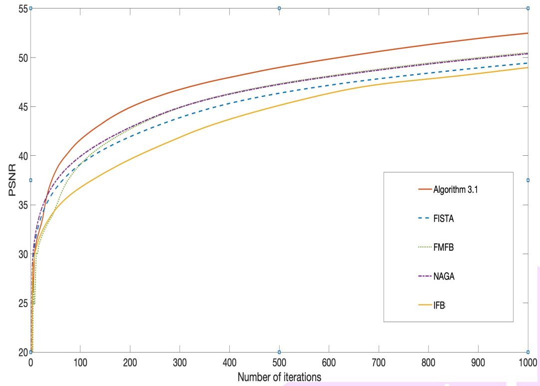


(c) PSNR plotting of RGB images

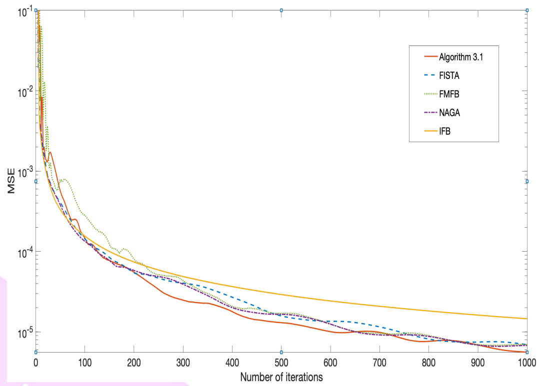


(d) MSE plotting of RGB images

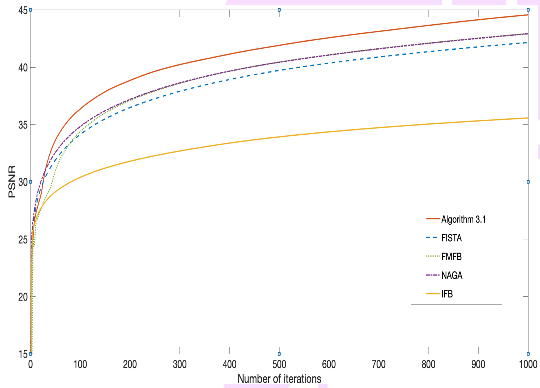
Figure 4: The PSNR plotting and MSE plotting of comparison of IFB, FISTA, NAGA, FMFB and Algorithm 4.3.1 for motion blur.



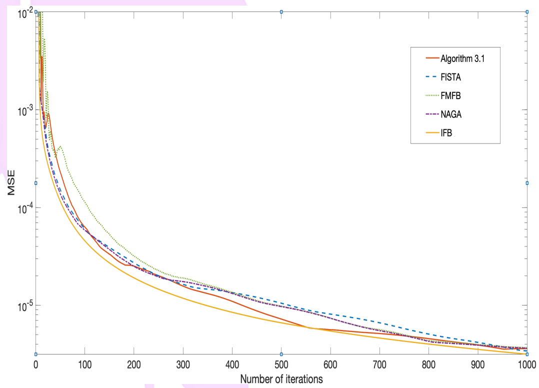
(e) PSNR plotting of grey images



(f) MSE plotting of grey images

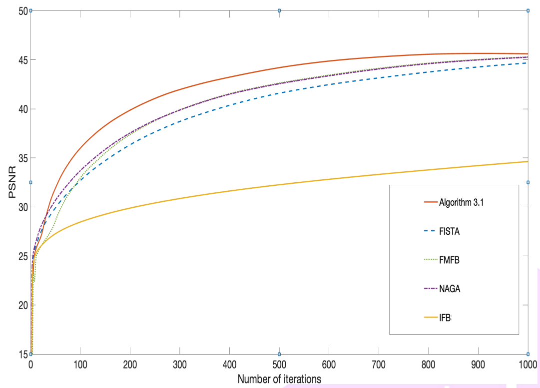


(g) PSNR plotting of RGB images

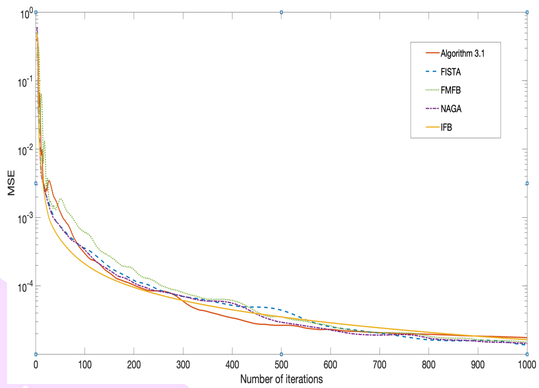


(h) MSE plotting of RGB images

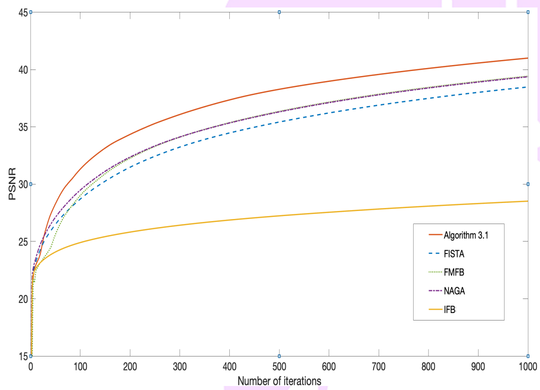
Figure 5: The PSNR plotting and MSE plotting of comparison of IFB, FISTA, NAGA, FMFB and Algorithm 4.3.1 for gaussian blur.



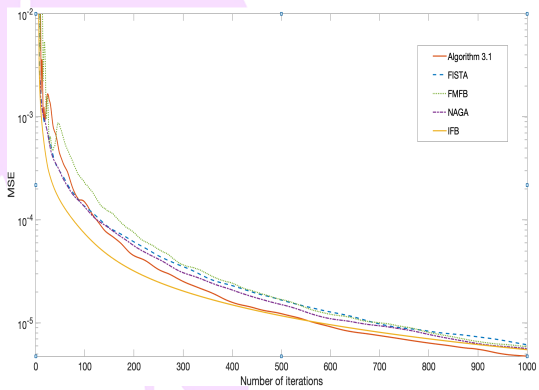
(i) PSNR plotting of grey images



(j) MSE plotting of grey images



(k) PSNR plotting of RGB images



(l) MSE plotting of RGB images

Figure 6: The PSNR plotting and MSE plotting of comparison of IFB, FISTA, NAGA, FMFB and Algorithm 4.3.1 for out of focus.

Example 5.1.2 We compare our algorithm (IMFB) in section 4.1 with FISTA, MFB, FRB, MFRB, IMFB, MSP and FMFB.

In method IMFB, we set $t_0 = 1$, $t_n = \frac{1+\sqrt{1+4t_{n-1}^2}}{2}$ and

$$\theta_n = \begin{cases} \frac{t_{n-1}-1}{t_n}, & \text{if } 1 \leq n \leq 1000 \\ 0 & \text{otherwise.} \end{cases}$$

The regularization parameters are chosen by $\tau = 10^{-5}$ and $x_0 = x_1 = (1, 1, 1, ..., 1) \in \mathbb{R}^N$. We set the following parameters:

Table 4: The parameters for each methods in section 4.1

Parameters	FISTA	MFB	FRB	MFRB	IMFB	MSP	FMFB	IMFB
$\lambda_0 = 0.1$	$\lambda_n = 1/\|A\|^2$	-	✓	✓	$\lambda_n = 1/\|A\|^2$	-	-	-
$\lambda_1 = 0.5$		-	-	✓		-	-	✓
$\sigma = 0.1$		✓	-	-		-	✓	-
$\rho = 0.8$		✓	-	-		-	✓	-
$\beta = 0.5$		-	✓	-		-	-	-
$\gamma = 1/\beta$		-	✓	-		-	-	-
$\delta = 0.5$		✓	✓	-		-	-	✓
$\mu = 0.2$		-	-	✓		-	✓	-
$\alpha_n = \frac{1}{n+1}$		-	-	-		✓	-	-
$\psi = 2$		-	-	-		✓	-	-
$r(x_n) = \frac{1}{2}x_n$		-	-	-		✓	-	-

For the experiments, we use the sizes 251×189 for RGB images. We add Poisson noise. For the results recovering the degraded RGB images, we limit the iterations to 1,000. We report the numerical results as follows:

Table 5: The comparison of PSNR, SSIM and CPU time in seconds for each methods of the restored images

Methods	Motion blur			Gaussian blur			Out of focus		
	PSNR	SSIM	CPU	PSNR	SSIM	CPU	PSNR	SSIM	CPU
FISTA	25.1122	0.7694	48.0184	34.3744	0.9320	47.2637	30.9043	0.8672	47.2831
MFB	24.8516	0.7640	71.2241	34.9546	0.9405	71.4400	28.3107	0.8152	70.4725
FRB	27.5733	0.8536	113.2112	38.6119	0.9703	112.5153	31.8188	0.8886	111.3155
MFRB	25.5158	0.7893	77.6740	36.0870	0.9515	70.3914	29.3660	0.8412	69.7562
IMFB	33.6105	0.9453	42.8921	41.1854	0.9818	42.7064	34.8978	0.9293	42.9258
MSP	34.8218	0.9560	92.9287	38.0609	0.9675	93.2503	32.0816	0.8941	93.2242
FMFB	40.8550	0.9785	64.8279	43.9280	0.9888	64.9999	38.0780	0.9544	64.8632
IMFB	46.7885	0.9920	75.7321	47.3368	0.9939	75.5435	41.0665	0.9743	74.7891

In Table 5, we see that IMFB has a higher PSNR than FISTA, MFB, FRB, MFRB, IMFB, MSP, FMFB for the same number of iterations. Moreover, SSIM of IMFB is closer to 1 than other methods. This shows that our algorithm has a better convergence than other methods for this example. However, we observe that IMFB has a less CPU time than other methods.

We next show the different types of blurred RGB images with the PSNR.

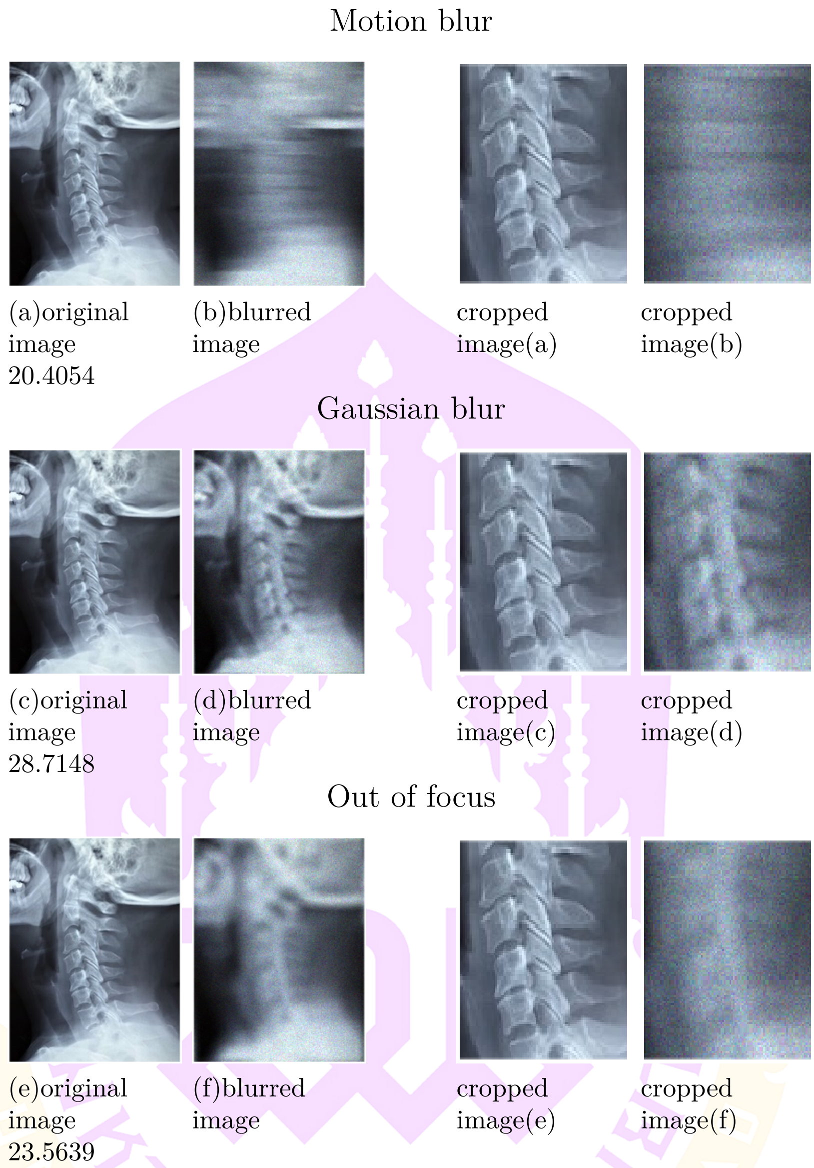


Figure 7: The original images and images degraded for each blurred RGB images with noise.

We next show the restored images of RGB images for Motion blur with the PSNR.

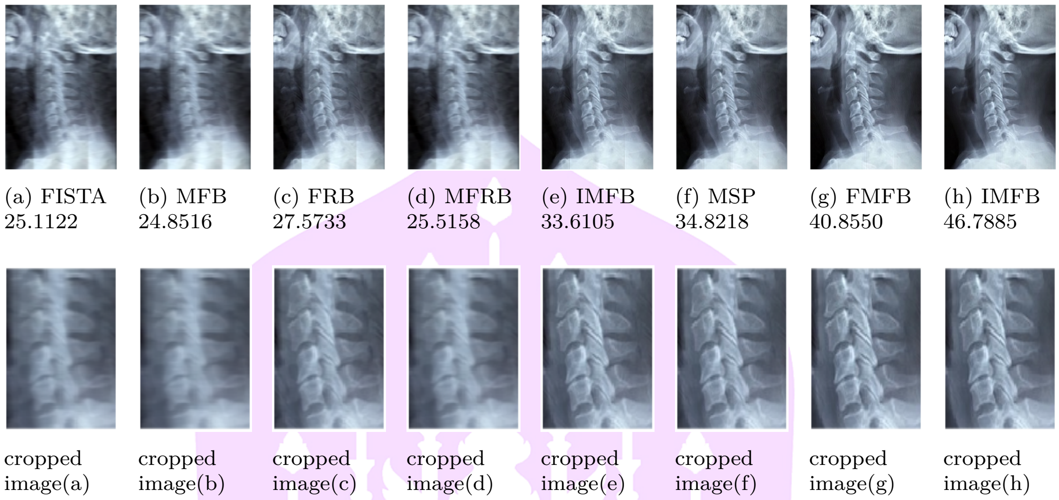


Figure 8: Recovered images via the different methods for degraded images by Motion blur

We next show the restored images of RGB images for Gaussian blur with the PSNR.

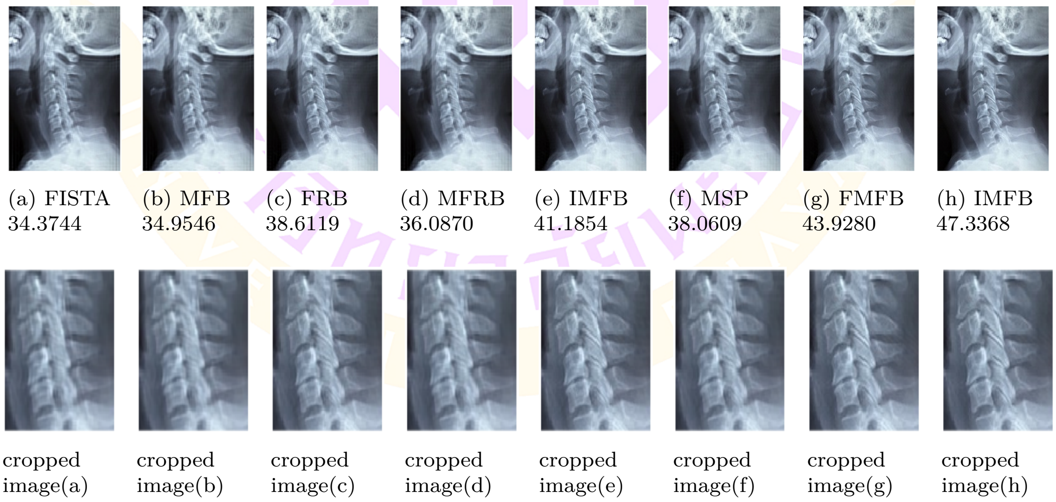
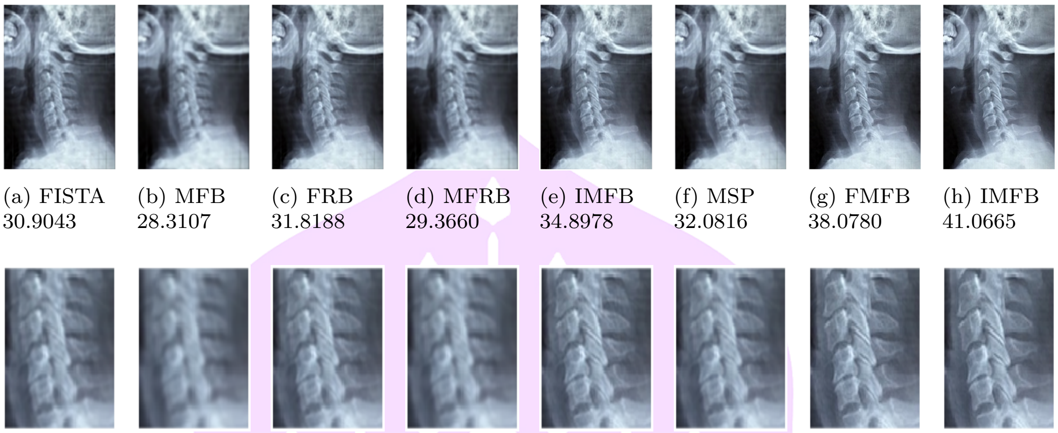


Figure 9: Recovered images via the different methods for degraded images by Gaussian blur

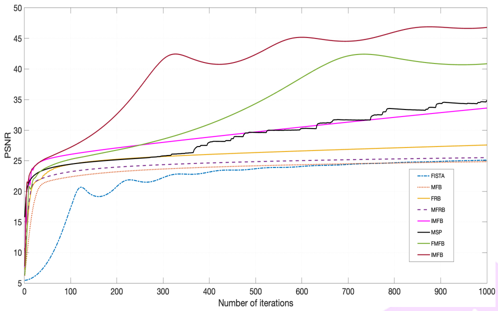
We next show the restored images of RGB images for out of focus with the PSNR.



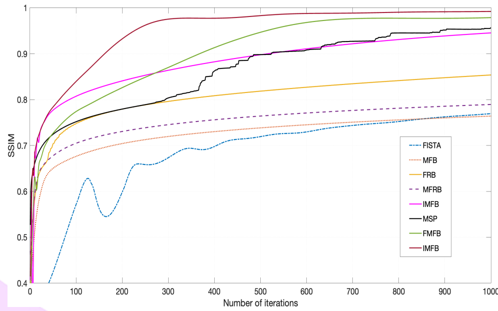
(a) FISTA 30.9043 | (b) MFB 28.3107 | (c) FRB 31.8188 | (d) MFRB 29.3660 | (e) IMFB 34.8978 | (f) MSP 32.0816 | (g) FMFB 38.0780 | (h) IMFB 41.0665

cropped image(a) | cropped image(b) | cropped image(c) | cropped image(d) | cropped image(e) | cropped image(f) | cropped image(g) | cropped image(h)

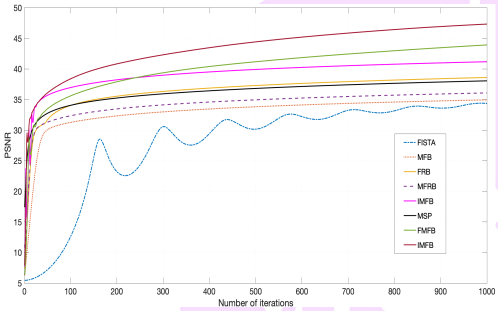
Figure 10: Recovered images via the different methods for degraded images by out of focus



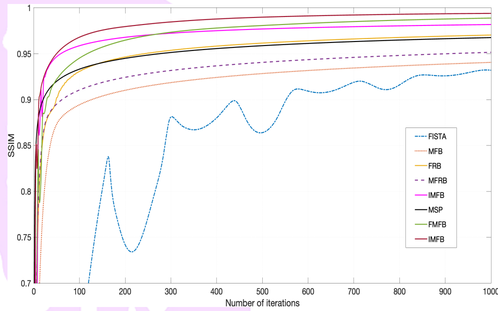
(a) PSNR plotting of Motion blur



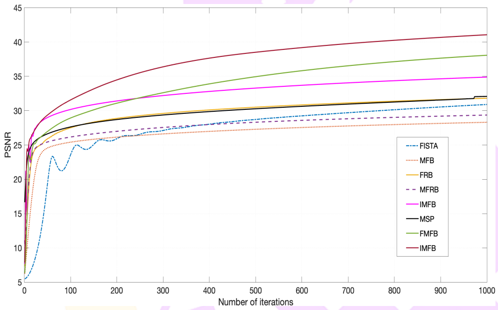
(b) SSIM plotting of Motion blur



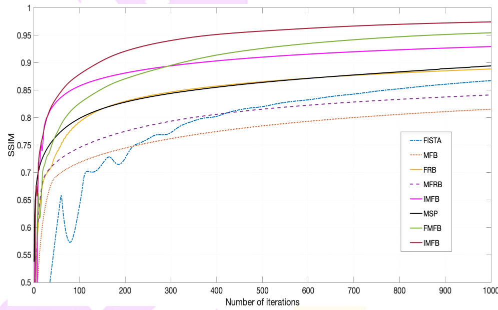
(c) PSNR plotting of Gaussian blur



(d) SSIM plotting of Gaussian blur



(c) PSNR plotting of Out of focus



(d) SSIM plotting of Out of focus

Figure 11: Graphs of PSNR and SSIM values for each blurred image and restored images for each methods

The results of the numerical experiments are summarized in Table 5. In Figures 2, 3, 4 and 11, we report all results that include the recovered images via each algorithms. It is shown that IMFB outperforms FISTA, MFB, FRB, MFRB, IMFB, MSP and FMFB in terms of PSNR and SSIM.

5.2 Application to image inpainting

In this section, we provide numerical experiments that support an inertial modified relaxed forward-backward-forward method (IRFBF) in section 4.2. We aim to apply our result to solve an image inpainting problem which is the following minimization:

$$\min_{x\in\mathbb{R}^{M\times N}} \frac{1}{2}\|A(x-x_0)\|_F^2 + \tau\|x\|_*$$

where $x_0 \in \mathbb{R}^{M\times N}(M < N)$, 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 $\tau > 0$ is regularization parameter.

In particular, we aim to solve the following image inpainting problem [19, 21]:

$$\min_{x\in\mathbb{R}^{M\times N}} \frac{1}{2}\|P_\Omega(x) - P_\Omega(x_0)\|_F^2 + \tau\|x\|_* \tag{5.2.1}$$

where $\|\cdot\|_F$ is the Frobenius matrix norm, and $\|\cdot\|_*$ is the nuclear matrix norm. Define P_Ω as follows:

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

The image inpainting problem (5.2.1) is problem (2.1.1), when $f(x) = \frac{1}{2}\|P_\Omega(x) - P_\Omega(x_0)\|_F^2$ and $g(x) = \tau\|x\|_*$. We know that $\nabla f(x) = P_\Omega(x) - P_\Omega(x_0)$ is 1-Lipschitz continuous and $prox_g$ is obtained by the singular value decomposition (SVD) [12].

To measure the quality of images, we use the signal-to-noise ratio (SNR) and the structural similarity index (SSIM) [54] which are defined by:

$$\text{SNR} = 20 \log \frac{\|x\|_F}{\|x - x_r\|_F}$$

and

$$\text{SSIM} = \frac{(2a_x a_{x_r} + c_1)(2\sigma_{xx_r} + c_2)}{(a_x^2 + a_{x^r}^2 + c_1)(\sigma_x^2 + \sigma_{x_r}^2 + c_2)}$$

where x is the original image, x_r is the restored image, a_x and a_{x_r} are the mean values of the original image a and restored image x_r, respectively, σ_x^2 and $\sigma_{x_r}^2$ are the variances, $\sigma_{xx_r}^2$ is the covariance of two images, $c_1 = (0.01L)^2$ and $c_2 = (0.03L)^2$, and L is the dynamic range of pixel values. SSIM ranges from 0 to 1, and 1 means perfect recovery.

Next, we present the performance of **IRFBF** in section 4.2 and its comparison to the projected version of **RFB** and **FMFB**. In all tests, the starting point $x_0 = x_1 = (0, 0, 0, ..., 0) \in \mathbb{R}^N$. Set

$\lambda_n = 1/\|A\|^2, \alpha_n = 0.09$ for **RFB**;

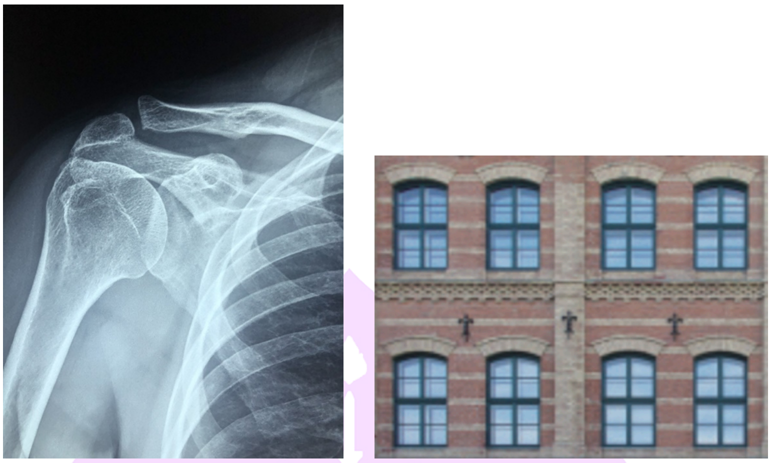
$\sigma = 0.1, \delta = 0.2, \gamma = 0.5$ for **FMFB**;

$\lambda_1 = 0.2, \mu = 0.2, \rho = 2$ for **IRFBF**.

We test two images. The first one, we use x-ray image with size 700×525 (see Figure 12(a)). For the second, we use windows image with size 466×572 (see Figure 12(b)).

The maximum number of iterations was set to be 300th. All codes are written in Matlab (version R2020b) on MacBook Pro M1 with ram 8 GB. Then, we have the numerical results as follows:

From Table 6, we can see that our algorithm **(IRFBF)** is effective and



(a)x-ray image (b)windows image

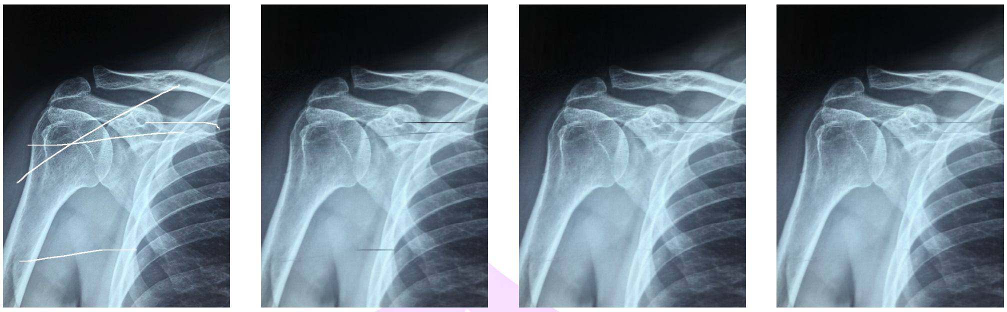
Figure 12: The original images for image inpanting

Table 6: The SNR and SSIM for each methods in section 4.2

Methods	x-ray		windows	
	SNR	SSIM	SNR	SSIM
RFB	19.9964	0.9614	12.5741	0.9365
FMFB	20.4239	0.9649	12.6261	0.9388
IRFBF	22.0834	0.9653	13.6382	0.9391

has higher SNR and SSIM than **RFB** and **FMFB** for both images. This means that our proposed algorithm is better than other methods.

Next, we show the figures of inpainting for each methods.

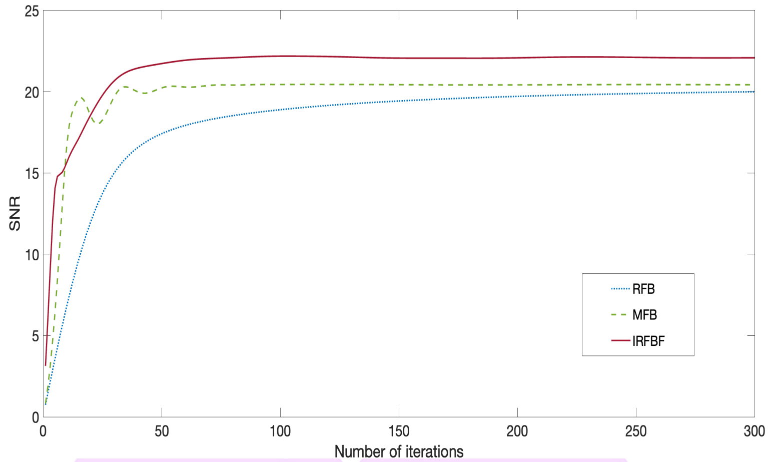


(a)painted image | (b)restored image by RFB (SNR=19.9964) | (c)restored image by FMFB (SNR=20.4239) | (d)restored image by IRFBF (SNR=22.0834)

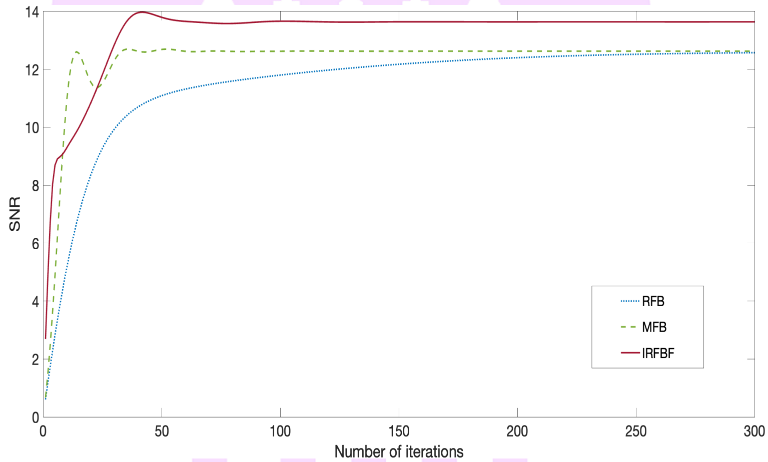
Figure 13: The painted x-ray image and restored images by RFB, FMFB and IRFBF

(a)painted image | (b)restored image by RFB (SNR=12.5741) | (c)restored image by FMFB (SNR=12.6261) | (d)restored image by IRFBF (SNR=13.6382)

Figure 14: The painted windows image and restored images by RFB, FMFB and IRFBF



(a) SNR plotting of x-ray image



(b) SNR plotting of windows image

Figure 15: Graphs of SNR values of x-ray and windows images for each methods

We see that the proposed method does not require the computation of Lipschitz constant of the gradient of functions. Moreover, the linesearch of iterations is not necessary in algorithms.

5.3 Application to data classification

In this section, we apply an inertial double proximal forward-backward method (IDFB) in section 4.4 to data classification problems. The data classification problems based on a learning technique called extreme learning machine

(ELM).

Let $\{(x_n, y_n) : x_n \in \mathbb{R}^N, y_n \in \mathbb{R}^M, n = 1, 2, 3, ..., K\}$ be a training set of K distinct samples, x_n is an input training data and y_n is a training target. For the output of ELM with single hidden layer at the i-th hidden node is

$$h_i(x) = f(a_i \cdot x + b_i),$$

where f is an activation function, a_i is the weight at the i-th hidden node and b_i is the bias at the i-th hidden node.

The output function with L hidden nodes is the single-hidden layer feed forward neural networks (SLFNs)

$$O_n = \sum_{i=1}^{L} \beta_i h_i(x_n),$$

where β_i is the optimal output weight at the i-th hidden node. The hidden layer output matrix A is defined by

$$A = \begin{bmatrix} f(a_1 \cdot x_1 + b_1) & \cdots & f(a_L \cdot x_1 + b_L) \\ \vdots & \ddots & \vdots \\ f(a_1 \cdot x_K + b_1) & \cdots & f(a_L \cdot x_K + b_L) \end{bmatrix}$$

The main aim of ELM is to calculate an optimal weight $\beta = [\beta_1, ..., \beta_L]^T$ such that $A\beta = B$, where $B = [t_1, ..., t_K]^T$ is the training target data. We find the solution β via convex minimization problem. Next, we present the least absolute shrinkage and selection operator (LASSO) [49] to find the parameter β. It can be formulated as follows:

$$\min_{\beta \in \mathbb{R}^L} \{\|A\beta - B\|_2^2 + \tau\|\beta\|_1\}, \tag{5.3.1}$$

where τ is a regularization parameter. We see that if $f(\beta) = \|A\beta - B\|_2^2$ and $g(\beta) = \tau\|\beta\|_1$, then the problem (5.3.1) is reduced to the problem (2.1.1).

In experiments, we use a cervical cancer behaviour risk data set from UCI Machine Learning Repository [32] for training processing. This data set contains 72 samples which has 19 attributes. We classify two classes of data.

We use the sigmoid as the activation function and the hidden nodes $L = 300$. For efficiency of algorithms, we measure by the accuracy of the output data as follows:

$$\text{accuracy} = \frac{\text{correctly predicted data}}{\text{all data}} \times 100.$$

For the loss of an example, it is computed by the binary cross entropy loss function:

$$\text{Loss} = -\frac{1}{\text{output size}} \sum_{i=1}^{\text{output size}} y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i),$$

where $\hat{y}_i$ is the i-th scalar value in the model output, y_i is the corresponding target value, and output size is the number of scalar values in the model output.

We set parameters for each methods as follows:

Table 7: Chosen control parameters of each methods in section 4.4

Methods	γ	ℓ	δ	μ
FISTA		$L = 1/\|A\|$		
FMFB	2	0.5	0.1	-
IDFB	2	0.5	-	0.8

In our method (IDFB), we set $t_0 = 1$, $t_n = \frac{1+\sqrt{1+4t_{n-1}^2}}{2}$ and

$$\theta_n = \begin{cases} \frac{t_{n-1}-1}{t_n} & \text{if } n \leq 1000, \\ 0 & \text{otherwise.} \end{cases}$$

The regularization parameter is $\tau = 10^{-5}$. The stopping criteria is the binary cross entropy (Loss=0.119). We obtain the results in Table 8.

Table 8: The performance of each methods with the stopping criteria as training accuracy > 90 and testing accuracy > 90

Methods	Iter	Training time	Acc(%)
FISTA	49	0.0491	90.91
FMFB	32	0.9747	90.91
IDFB	25	0.8014	90.91

From Table 8, we see that IDFB has the number of iterations less than FISTA and FMFB at the testing accuracy 90.91. It shows that IDFB has a better efficiency than other methods.

Next, we show graphs of the accuracy and loss of training data and testing data for overfitting of IDFB.

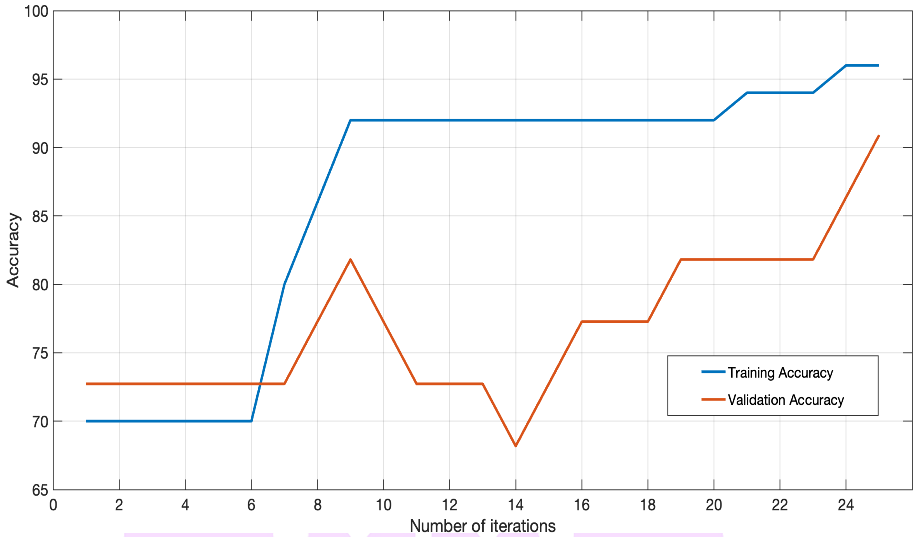


Figure 16: Graph accuracy of IDFB

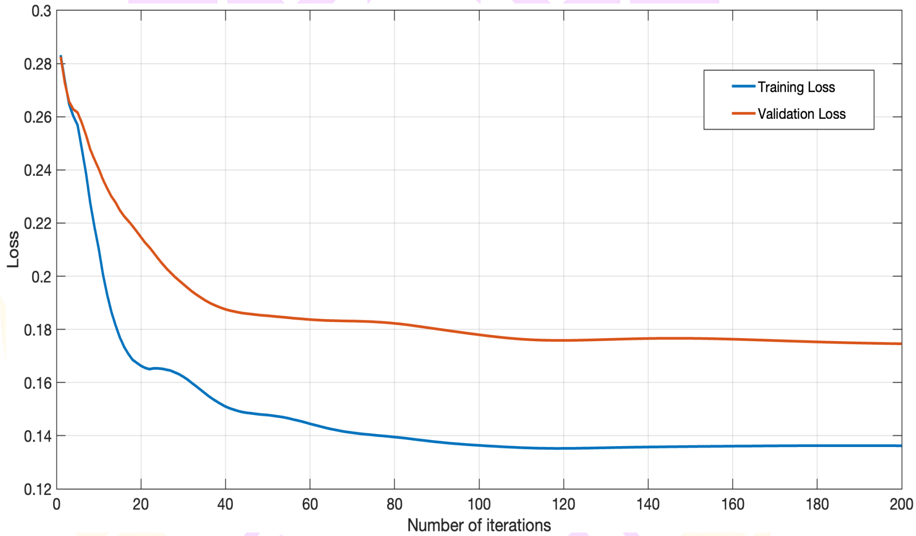


Figure 17: Graph loss of IDFB

From Figure 16, we see that training accuracy and validation accuracy have a high gap. It shows that a few training data set are not good enough to train model. Also, Figure 17 has a gap between training loss and testing loss. However, graphs of accuracy and loss values tends in the same way which show that our method (IDFB) can still classify data set even if there are a few data set.

CHAPTER VI

CONCLUSIONS

The following results are all main theorems of this thesis:

Algorithm 6.1.1 Inertial modified forward-backward method (IMFB)

Initialization: *Let* $x_0 = x_1 \in H$, $\lambda_1 > 0$, $\theta_1 > 0$ *and* $\delta \in (0, 1)$.

Iterative step: *For* $n \geq 1$, *calculate* x_{n+1} *as follows:*

Step 1. *Compute the inertial step:*

$$w_n = x_n + \theta_n(x_n - x_{n-1}).$$

Step 2. *Compute the forward-backward step:*

$$y_n = prox_{\lambda_n g}(w_n - \lambda_n \nabla f(w_n)).$$

Step 3. *Compute the* x_{n+1} *step:*

$$x_{n+1} = y_n - \lambda_n(\nabla f(y_n) - \nabla f(w_n))$$

where

$$\lambda_{n+1} = \begin{cases} \min\{\frac{\delta\|w_n - y_n\|}{\|\nabla f(w_n) - \nabla f(y_n)\|}, \lambda_n\} & if \ \|\nabla f(w_n) - \nabla f(y_n)\| \neq 0; \\ \lambda_n & otherwise. \end{cases}$$

Set $n = n + 1$ *and return to* **Step 1**.

Lemma 6.1.2 *Let* $\{x_n\}$ *be generated by Algorithm 6.1.1. Then*

$$\|x_{n+1} - x^*\|^2 \leq \|w_n - x^*\|^2 - (1 - \frac{\delta^2\lambda_n^2}{\lambda_{n+1}^2})\|y_n - w_n\|^2, \forall x^* \in \Omega.$$

Lemma 6.1.3 *Let* $\{x_n\}$ *be generated by Algorithm 6.1.1. If* $\sum_{n=1}^{\infty}\theta_n < \infty$*, then* $\lim_{n\to\infty}\|x_n - x^*\|$ *exists for all* $x^* \in \Omega$*.*

Lemma 6.1.4 *Let* $\{x_n\}$ *be generated by Algorithm 6.1.1. If* $\sum_{n=1}^{\infty}\theta_n < \infty$*, then*

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

Theorem 6.1.5 *Let* $\{x_n\}$ *be generated by Algorithm 6.1.1. If* $\sum_{n=1}^{\infty}\theta_n < \infty$*, then* $\{x_n\}$ *weakly converges to a point in* Ω*.*

Algorithm 6.1.6 Inertial modified relaxed forward-backward-forward method (IRFBF)

Initialization: *Let* $x_0 = x_1 \in H$*,* $\lambda_1 > 0$*,* $\rho_1 \in (0,1)$*,* $\mu \in (0,1)$ *and* $\theta_1 \geq 0$*.*

Iterative step: *Let* Ω *be a nonempty closed convex subset of* H*. Given* $n \geq 1$*, calculate* x_{n+1} *as follows:*

Step 1. *Compute the inertial step:*

$$w_n = x_n + \theta_n(x_n - x_{n-1}).$$

Step 2. *Compute the forward-backward-forward step:*

$$\begin{aligned} y_n &= prox_{\lambda_n g}(w_n - \lambda_n \nabla f(w_n)), \\ z_n &= (1-\rho_n)w_n + \rho_n(y_n - \lambda_n(\nabla f(y_n) - \nabla f(w_n))), \end{aligned}$$

Step 3. *Compute the projection step:*

$$x_{n+1} = P_{\Omega}(z_n).$$

Step 4. *Compute the stepsize step:*

$$\lambda_{n+1} = \begin{cases} \min\{\lambda_n, \frac{\mu\|y_n - w_n\|}{\|\nabla f(y_n) - \nabla f(w_n)\|}\} & \text{if } \|\nabla f(y_n) - \nabla f(w_n)\| \neq 0; \\ \lambda_n & \text{otherwise.} \end{cases} \tag{6.1.2}$$

Set $n = n + 1$ *and return to* **Step 1**.

Lemma 6.1.7 *Let* $\mu \in (0, 1)$ *and* $\lambda_1 > 0$. *The sequence* $\{\lambda_n\}$ *generated by (6.1.2) is nonincreasing and*

$$\lim_{n\to\infty} \lambda_n = \lambda \geq \min\{\lambda_1, \frac{\mu}{L}\}.$$

Hence,

$$\|\nabla f(y_n) - \nabla f(w_n)\| \leq \frac{\mu}{\lambda_{n+1}}\|y_n - w_n\|.$$

Theorem 6.1.8 *Let* $\{x_n\}$ *be generated by Algorithm 6.1.6. If* $\sum_{n=1}^{\infty} \theta_n < +\infty$ *and* $0 < \liminf_{n\to\infty} \rho_n \leq \limsup_{n\to\infty} \rho_n < 1$, *then the sequence* $\{x_n\}$ *weakly converges to an element of* S.

Algorithm 6.1.9 Initialization: *Let* $x_0 = x_1 \in H$, $\sigma > 0$, $\rho \in (0, 1)$, $\delta \in (0, 1)$, $\theta_1 > 0$

Iterative step: *For* $n \geq 1$, *calculate* x_{n+1} *as follows:*

Step 1. *Compute the inertial step:*

$$w_n = x_n + \theta_n(x_n - x_{n-1}).$$

Step 2. *Compute the forward-backward step:*

$$y_n = prox_{\lambda_n g}(w_n - \lambda_n \nabla f(w_n)),$$

where $\lambda_n = \sigma\rho^{m_n}$ *and* m_n *is the smallest nonnegative integer such that*

$$\lambda_n(\|\nabla f(prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n))) - \nabla f(y_n)\| + \|\nabla f(w_n) - \nabla f(y_n)\|)$$
$$\leq \frac{\delta}{2}(\|prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n)) - y_n\| + \|w_n - y_n\|).$$

Step 3. *Compute the* x_{n+1} *step:*

$$x_{n+1} = prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n)).$$

Set $n = n+1$ *and return to Step 1.*

Theorem 6.1.10 *Let* $\{x_n\}$ *be generated by Algorithm 6.1.9. If* $\sum_{n=1}^{\infty} \theta_n < +\infty$ *and* $\lambda_n \geq \lambda$ *for some* $\lambda > 0$*, then the sequence* $\{x_n\}$ *weakly converges to an element of* Ω.

Algorithm 6.1.11 An inertial double proximal forward-backward method (IDFB)

Initialization: *Let* $x_0 = x_1 \in H$, $\theta_1 > 0$, $\gamma > 0$, $\ell \in (0,1)$ *and* $0 < \mu < 1$.

Iterative step: *For* $n \geq 1$*, calculate* x_{n+1} *as follows:*

Step 1. *Compute the inertial step:*

$$w_n = x_n + \theta_n(x_n - x_{n-1}).$$

Step 2. *Compute the forward-backward step:*

$$y_n = prox_{\lambda_n g}(w_n - \lambda_n \nabla f(w_n)).$$

Step 3. *Compute the* x_{n+1} *step:*

$$x_{n+1} = prox_{\lambda_n g}(y_n - \lambda_n \nabla f(y_n))$$

where the linesearch $\lambda_n = \gamma\ell^{m_n}$ *is the smallest nonegative integer such that*

$$\begin{aligned} & \lambda_n(\langle \nabla f(x_{n+1}) - \nabla f(y_n), x_{n+1} - y_n\rangle + \langle \nabla f(y_n) - \nabla f(w_n), y_n - w_n\rangle) \\ \leq \; & \frac{\mu^2+1}{4}\|x_{n+1} - y_n\|^2 + \frac{\mu}{\mu+1}\|y_n - w_n\|^2 \end{aligned} \tag{6.1.3}$$

Set $n = n+1$ *and return to* **Step 1**.

Lemma 6.1.12 *Let* $x \in H$, $\gamma > 0$, $\ell \in (0,1)$ *and* $0 < \mu < 1$. *For* $i = 1, 2, 3, ...,$ *set*

$$\begin{aligned} H(x,i) &= \text{prox}_{\gamma\ell^i g}(x - \gamma\ell^i \nabla f(x)) \\ P(x,i) &= \text{prox}_{\gamma\ell^i g}(H(x,i) - \gamma\ell^i \nabla f(H(x,i))). \end{aligned}$$

If

$$\begin{aligned} & \gamma\ell^i(\langle \nabla f(P(x,i)) - \nabla f(H(x,i)), P(x,i) - H(x,i)\rangle \\ & + \langle \nabla f(H(x,i)) - \nabla f(x), H(x,i) - x\rangle) \\ \leq \; & \frac{\mu^2+1}{4}\|P(x,i) - H(x,i)\|^2 + \frac{\mu}{\mu+1}\|H(x,i) - x\|^2, \end{aligned}$$

then $\lambda = \gamma\ell^i$.

Else $i = i+1$. *The linesearch (6.1.3) stops after finitely many steps.*

Theorem 6.1.13 *Let* $\{x_n\}$ *be generated by Algorithm 6.1.11. If* $\sum_{n=1}^{\infty}\theta_n < \infty$ *and* $\lambda_n \geq \lambda$ *for some* $\lambda > 0$, *then* $\{x_n\}$ *weakly converges to point in* Ω.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Agarwal, R. P., O'Regan,D., Sahu, D.R. (2009). Fixed point theory for Lipschitzian-type mappings with applications. **New York: Springer.**

[2] Alvarez, F. and Attouch, H. (2001). An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. **Set-Valued Analysis**, 9(1), 3-11.

[3] Ansari, Q. H. and Rehan, A. (2014). Split feasibility and fixed point problems. **Nonlinear analysis**, 281-322.

[4] Bauschke, H. H. and Combettes, P. L. (2011). **Convex analysis and monotone operator theory in Hilbert spaces** (Vol. 408). New York: Springer.

[5] 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.

[6] 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.

[7] Bertsekas, D. and Tsitsiklis, J. (2015). Parallel and distributed computation: numerical methods. **Athena Scientific**.

[8] Bot, R. I., Sedlmayer, M., and Vuong, P. T. (2020). A relaxed inertial forward-backward-forward algorithm for solving monotone inclusions with application to GANs. **arXiv preprint arXiv:2003.07886**.

[9] 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.

[10] Byrne, C. (2002). Iterative oblique projection onto convex sets and the split feasibility problem. **Inverse problems**, 18(2), 441.

[11] Byrne, C. (2003). A unified treatment of some iterative algorithms in signal processing and image reconstruction. **Inverse problems**, 20(1), 103.

[12] Cai, J. F., Cand s, E. J. and Shen, Z. (2010). A singular value thresholding algorithm for matrix completion. **SIAM Journal on optimization**, 20(4), 1956-1982.

[13] Cegielski, A. (2012). Iterative Methods for Fixed Point Problems in Hilbert Spaces Preface. **Iterative methods for fixed point problems in hilbert spaces**, 2057.

[14] Ceng, L. C., Ansari, Q. H. and Yao, J. C. (2012). An extragradient method for solving split feasibility and fixed point problems. **Computers and Mathematics with Applications**, 64(4), 633-642.

[15] Censor, Y. and Elfving, T. (1994). A multiprojection algorithm using Bregman projections in a product space. **Numerical Algorithms**, 8(2), 221-239.

[16] Cholamjiak, W., Cholamjiak, P. and Suantai, S. (2018). An inertial forward-backward splitting method for solving inclusion problems in Hilbert spaces. **Journal of Fixed Point Theory and Applications**, 20(1), 1-17.

[17] Combettes, P. L. and Wajs, V. R. (2005). Signal recovery by proximal forward-backward splitting. **Multiscale modeling and simulation**, 4(4), 1168-1200.

[18] Combettes, P.L. (2004). Solving monotone inclusions via compositions of nonexpansive averaged operators. **Optimization**, 53, 475-504.

[19] Cui, F., Tang, Y. and Yang, Y. (2019). An inertial three-operator splitting algorithm with applications to image inpainting. **arXiv preprint arXiv:1904.11684**.

[20] Dang, Y. and Gao, Y. (2010). The strong convergence of a KM-CQ-like algorithm for a split feasibility problem. **Inverse Problems**, 27(1), 015007.

[21] Davis, D. and Yin, W. (2017). A three-operator splitting scheme and its optimization applications. **Set-valued and variational analysis**, 25(4), 829-858.

[22] Douglas, J. and Rachford, H.H. (1956). On the numerical solution of heat conduction problems in two or three space variables. **Transactions of the American Mathematical Society**, 82, 421-439.

[23] Dunn, J.C. (1976). Convexity, monotonicity, and gradient processes in Hilbert space. **Journal of Mathematical Analysis and Applications**, 53, 145-158.

[24] Eckstein, J. and Bertsekas, D.P. (1992). On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. **Mathematical Programming**, 55, 293-318.

[25] Hanjing, A. and Suantai, S. (2020). A fast image restoration algorithm based on a fixed point and optimization method. **Mathematics**, 8(3), 378.

[26] Van Hieu, D., Anh, P. K. and Muu, L. D. (2021). Modified forward?backward splitting method for variational inclusions. **4OR**, 19(1), 127-151.

[27] Iusem, A. N., Svaiter, B. F. and Teboulle, M. (1994). Entropy-like proximal methods in convex programming. **Mathematics of Operations Research**, 19(4), 790-814.

[28] Korpelevich, G.M. (1976). The extragradient method for finding saddle points and other problems. **Matecon**, 12, 747-756.

[29] Liang, J., Luo, T. and Schonlieb, C. B. (2022). Improving Fast Iterative Shrinkage-Thresholding Algorithm: Faster, Smarter, and Greedier. **SIAM Journal on Scientific Computing**, 44(3), A1069-A1091.

[30] Lions, P.L. and Mercier, B. (1979). Splitting algorithms for the sum of two nonlinear operators. **SIAM Journal on Numerical Analysis**, 16, 964-979.

[31] L pez, G., Mart n-M rquez, V., Wang, F. and Xu, H. K. (2012). Solving the split feasibility problem without prior knowledge of matrix norms. **Inverse Problems**, 28(8), 085004.

[32] Machmud, R. and Wijaya, A. (2016). Behavior determinant based cervical cancer early detection with machine learning algorithm. **Advanced Science Letters**, 22(10), 3120-3123.

[33] Malitsky, Y. and Tam, M. K. (2020). A forward-backward splitting method for monotone inclusions without cocoercivity. **SIAM Journal on Optimization**, 30(2), 1451-1472.

[34] Martinet, B. (1970). Regularisation d'inéquations variationelles par approximations successives, Rev. Française Informat. **Recherche Opérationnelle**, 4, 154-158.

[35] Martinet, B. (1972). Determination approchée d'un point fixe d'une application pseudocontractante. **Comptes Rendus de l'Acadmie des Sciences**, 274, 163-165.

[36] 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.

[37] 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).

[38] Osilike, M. O., Aniagbosor, S. C. and Akuchu, B. G. (2002). Fixed points of asymptotically demicontractive mappings in arbitrary Banach spaces, Panamerican Math. **J**, 12(2), 77-88.

[39] Padcharoen, A., Kitkuan, D., Kumam, W. and Kumam, P. (2021). Tseng methods with inertial for solving inclusion problems and application to image deblurring and image recovery problems. **Computational and Mathematical Methods**, 3(3), e1088.

[40] Passty, G. B. (1979). Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. **Journal of Mathematical Analysis and Applications**, 72, 383-390.

[41] Polyak, B. T. (1964). Some methods of speeding up the convergence of iteration methods. **Ussr computational mathematics and mathematical physics**, 4(5), 1-17.

[42] Rockafellar, R. T. (1976). Monotone operators and the proximal point algorithm. **SIAM journal on control and optimization**, 14(5), 877-898.

[43] Rockafellar, R. T. (1976). Augmented Lagrangians and applications of the proximal point algorithm in convex programming. **Mathematics of operations research**, 1(2), 97-116.

[44] Salzo, S. (2017). The variable metric forward-backward splitting algorithm under mild differentiability assumptions. **SIAM Journal on Optimization**, 27(4), 2153-2181.

[45] Shehu, Y. and Cholamjiak, P. (2019). Iterative method with inertial for variational inequalities in Hilbert spaces. **Calcolo**, 56(1), 1-21.

[46] Suantai, S., Jailoka, P. and Hanjing, A. (2021). An accelerated viscosity forward-backward splitting algorithm with the linesearch process for convex minimization problems. **Journal of Inequalities and Applications**, 2021(1), 1-19.

[47] Suantai, S., Pholasa, N. and Cholamjiak, P. (2018). The modified inertial relaxed CQ algorithm for solving the split feasibility problems. **Journal of Industrial and Management Optimization**, 14(4), 1595.

[48] Takahashi, W. (2000). Nonlinear functional analysis. **Fixed Point Theory and its Applications.**

[49] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. **Journal of the Royal Statistical Society: Series B (Methodological)**, 58(1), 267-288.

[50] Thung, K.H., Raveendran, P. (2009, December). A survey of image quality measures. **In 2009 International Conference for Technical Postgraduates (TECHPOS)** (pp. 1-4). IEEE.

[51] Tseng, P. (2000). A modified forward-backward splitting method for maximal monotone mappings. **SIAM Journal on Control and Optimization**, 38(2), 431-446.

[52] Tseng, P. and Yun, S. (2009). A coordinate gradient descent method for nonsmooth separable minimization. **Mathematical Programming**, 117(1), 387-423.

[53] Verma, M. and Shukla, K. K. (2017). A new accelerated proximal gradient technique for regularized multitask learning framework. **Pattern Recognition Letters**, 95, 98-103.

[54] Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P. (2004). Image quality assessment: from error visibility to structural similarity. **IEEE Transactions on Image Processing**, 13(4), 600-612.

[55] Xu, H. K. (2010). Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces. **Inverse problems**, 26(10), 105018.

[56] Yang, Q. (2004). The relaxed CQ algorithm solving the split feasibility problem. **Inverse problems**, 20(4), 1261.

BIOGRAPHY

BIOGRAPHY

Name Surname Suparat Kesornprom

Date of Birth December 16, 1993

Place of Birth Chiang Rai Province, Thailand

Address 93 Moo 3, Sanmaka Sub - district,
Pa Daet District, Chiang Rai Province,
Thailand 57190

Work Experiences

2016 The student teacher at Padadwittayakom School,
Chiangrai, Thailand

Education Background

2022 Ph.D. (Mathematics), University of Phayao,
Phayao, Thailand

2019 M.Sc. (Mathematics), University of Phayao,
Phayao, Thailand

2017 B.Sc. (Mathematics), University of Phayao,
Phayao, Thailand

2017 B.Ed. (Education), University of Phayao,
Phayao, Thailand

Publications

Articles

1. **Kesornprom, S.** and Cholamjiak, P. (2022). A modified inertial proximal gradient method for minimization problems and applications. AIMS Mathematics, 7(5), 8147-8161.

2. **Kesornprom, S.** and Cholamjiak, P.(2022). A double proximal gradient method with new linesearch for solving convex minimization problem with application to data classification. Re-

sults in Nonlinear Analysis. (Accepted)

3. **Kesornprom, S.** and Cholamjiak, P.(2022). A new relaxed inertial forward-backward-forward method for solving the convex minimization problem with application to image inpainting. Applied Set-Valued Analysis and Optimization.(Submitted)

4. **Kesornprom, S.**, Park, C. and Cholamjiak, P.(2022). New proximal type algorithms for convex minimization and its application to image deblurring. Computational and Applied Mathematics. 41, 333.

Others

1. Suantai, S., **Kesornprom, S.**, Cholamjiak, W. and Cholamjiak, P. (2022). Modified Projection Method with Inertial Technique and Hybrid Stepsize for the Split Feasibility Problem. Mathematics, 10(6), 933.

2. Suantai, S., Panyanak, B., **Kesornprom, S.** and Cholamjiak, P. (2022). Inertial projection and contraction methods for split feasibility problem applied to compressed sensing and image restoration. Optimization Letters, 16(6), 1725-1744.

3. Suantai, S., **Kesornprom, S.**, Pholasa, N., Cho, Y. J. and Cholamjiak, P. (2021). A relaxed projection method using a new linesearch for the split feasibility problem. AIMS Mathematics, 6(3), 2690-2703.

4. **Kesornprom, S.**, Pholasa, N. and Cholamjiak, P. (2021). Efficient projective methods for the split feasibility problem and its applications to compressed sensing and image debluring. Filomat, 35(10), 3241-3266.

5. **Kesornprom, S.**, Cholamjiak, P. and Cholamjiak, W. (2021). Strong Convergence of a Parallel Extragradient-Like Algorithm Involving Pseudo-Monotone Mappings for Solving Common Variational Inequality Problems. Thai Journal of Mathematics, 19(3), 854-864.

6. Saelii, T., **Kesornprom, S.** and Cholamjiak, P. (2020). A novel relaxed projective method for split feasibility problems. Thai Journal of Mathematics, 18(3), 1359-1373.

7. **Kesornprom, S.** and Cholamjiak, P. (2020). A new iterative scheme using inertial technique for the split feasibility problem with application to compressed sensing. Thai Journal of Mathematics, 18(1), 315-332.

8. **Kesornprom, S.**, Pholasa, N. and Cholamjiak, P. (2020). On the convergence analysis of the gradient-CQ algorithms for the split feasibility problem. Numerical Algorithms, 84(3), 997-1017.

9. **Kesornprom, S.** and Cholamjiak, P. (2019). Proximal type algorithms involving linesearch and inertial technique for split variational inclusion problem in hilbert spaces with applications. Optimization, 68(12), 2369-2395.

10. Suantai, S., **Kesornprom, S.** and Cholamjiak, P. (2019). A new hybrid CQ algorithm for the split feasibility problem in Hilbert spaces and its applications to compressed sensing. Mathematics, 7(9), 789.

11. Suantai, S., **Kesornprom, S.** and Cholamjiak, P. (2019). Modified

proximal algorithms for finding solutions of the split variational inclusions. Mathematics, 7(8), 708.

12. Cholamjiak, P., **Kesornprom, S.** and Pholasa, N. (2019). Weak and strong convergence theorems for the inclusion problem and the fixed-point problem of nonexpansive mappings. Mathematics, 7(2), 167.

13. Cholamjiak, W., Suantai, S., Suparatulatorn, R., **Kesornprom, S.** and Cholamjiak, P. (2019). Viscosity approximation methods for fixed point problems in Hilbert spaces endowed with graphs. J. Appl. Numer. Optim, 1(1), 25-38.

14. **Kesornprom, S.**, Pholasa, N. and Cholamjiak, P. (2019). A modified CQ algorithm for solving the multiple-sets split feasibility problem and the fixed point problem for nonexpansive mappings. Thai Journal of Mathematics, 17(2), 475-493.

Conference presentations

1. **Kesornprom, S.** (August 29-September 1, 2022). A modified inertial proximal gradient method for minimization problems and applications. In 11th International Eurasian Conference on Mathematical Sciences and Applications (IECMSA), Istanbul, Turkey.

2. **Kesornprom, S.** (September 8-9, 2022). Note on the modified inertial forward-backward method for minimization problems and image deblurring. In International conference on Digital Image Processing and Machine Learning (ICDIPML), University of Phayao, Phayao, Thailand.

3. **Kesornprom, S.** (August 20-22, 2021). Splitting methods involving inertial technique for split feasibility problem and its applications. In International Workshop on Digital Image Processing and Machine Learning (IWDM), Online.

4. **Kesornprom, S.** (April 28-30, 2021). Efficient projective methods for the split feasibility problem and its applications to compressed sensing and image debluring. In The 24th international annual symposium on computational science and engineering (ANSCSE24), (Online).

5. **Kesornprom, S.** (September 12-14, 2019). Proximal type algorithms involving linesearch and inertial technique for split variational inclusion problem in Hilbert spaces with applications. In International Workshop on Applied Nonlinear Analysis (IWANA), Bangsaen, Chonburi, Thailand.

6. **Kesornprom, S.** (May 15-17, 2019). The gradient-based algorithms for solving the split feasibility problems. In The 24th Annual Meeting in Mathematics (AMM), Burapha University, Chonburi, Thailand.