J. Nonlinear Var. Anal. 8 (2024), No. 3, pp. 451-471
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.8.2024.3.07

A PARAMETERIZED THREE-OPERATOR SPLITTING ALGORITHM FOR NON-CONVEX MINIMIZATION PROBLEMS WITH APPLICATIONS

LIUYI MIAO[1], YUCHAO TANG[2,*], CHANGLONG WANG[3]

[1]*Department of Mathematics, Nanchang University, Nanchang 330031, China*
[2]*School of Mathematics and Information Science, Guangzhou University, Guangzhou 510006, China*
[3]*Key Laboratory of Electronic Information Countermeasure and Simulation Technology of Ministry of Education, Xidian University, Xi'an 710071, China*

Abstract. In this paper, we propose a parameterized three-operator splitting algorithm to solve non-convex minimization problems with the sum of three non-convex functions, where two of them have Lipschitz continuous gradients. We establish the convergence of the proposed algorithm under the Kurdyka-Łojasiewicz assumption by constructing a suitable energy function with a non-increasing property. As applications, we employ the proposed algorithm to solve low-rank matrix recovery and image inpainting problems. Numerical results demonstrate the efficiency and effectiveness of the proposed algorithm compared to other algorithms.

Keywords. Image inpainting problems; Low rank matrix recovery; Non-convex minimization; Parameterized; Three-operator splitting algorithm.

1. Introduction

In this paper, we consider the following non-convex minimization problem:

$$\min_{x\in\mathscr{H}} F(x)+G(x)+H(x), \tag{1.1}$$

where $\mathscr{H}$ represents a real Hilbert space, $G:\mathscr{H}\to(-\infty,+\infty]$ is a proper lower semi-continuous function, and $F:\mathscr{H}\to R$ and $H:\mathscr{H}\to R$ are Fréchet differentiable with Lipschitz continuous gradient ∇F and ∇G, respectively. This problem covers various applications in sparse signal recovery [1, 2, 3, 4, 5] and low-rank matrix recovery [6, 7, 8, 9, 10]. When $H=0$, problem (1.1) reduces to the following non-convex minimization problem of the sum of two non-convex functions,

$$\min_{x\in\mathscr{H}} F(x)+G(x). \tag{1.2}$$

For problem (1.2), the forward-backward splitting algorithm [11] and the forward-backward-forward splitting algorithm [12] are two classic algorithms. By using new analysis tools and assuming that the objective function satisfies certain regularization conditions and the algorithm parameters meet some requirements, authors explored the theoretical convergence of these two

*Corresponding author.
E-mail address: hhaaoo1331@163.com (Y. Tang).
Received 5 June 2023; Accepted 17 March 2024; Published online 5 April 2024.

algorithms for solving the completely non-convex problem (1.2). Attouch et al. [13] proved that the iterative sequence generated by the forward-backward splitting algorithm converges to a critical point of (1.2) by imposing the Kurdyka-Łojasiewicz inequality [14]. Bot and Csetnek [16] proposed a forward-backward-forward splitting algorithm to solve (1.2), which combined with a Bregman distance. The proposed algorithm generalized the original forward-backward-forward splitting algorithm by Tseng [12] from the convex to the non-convex setting. Bot et al. [16] proposed an iterative algorithm with a combination of the forward-backward splitting algorithm and inertial techniques. They demonstrated that it can converge to a critical point of the objective function under certain assumptions. Themelis et al. [17] constructed a forward-backward splitting algorithm that converges to a stationary point of (1.2), which is based on the forward-backward Envelope.

On the other hand, when F and G are convex functions, and F has no differentiability conditions, the Douglas-Rachford splitting algorithm [18] is a simple and effective method for solving (1.2). However, for the non-convex problem (1.2), it is necessary to add the assumption that G is differentiable. The earliest work on this was proposed by Li and Pong [19]. They first studied the sequence generated by the Douglas-Rachford splitting algorithm to a stationary point of the non-convex minimization problem (1.2). In [20], Li et al. established the convergence of the Peaceman-Rachford splitting algorithm with the additional assumption that F is strongly convex. Further, Li and Wu [21] proposed a generalized splitting algorithm, which extended the Douglas-Rachford splitting algorithm [19] and the Peaceman-Rachford splitting algorithm [20]. Based on the Douglas-Rachford envelope, Themelis and Patrinos [22] developed a unified convergence analysis of the Douglas-Rachford splitting algorithm with over-relaxation parameter ranges in $(0,2]$. Furthermore, Themelis et al. [23] proposed line-search algorithms to improve the Douglas-Rachford splitting algorithm by means of quasi-Newton directions. Recently, Bian and Zhang [24] extended the so-called parameterized Douglas-Rachford splitting algorithm in Wang and Wang [25] for solving the non-convex minimization problem (1.2). They constructed a suitable energy function to obtain the theoretical convergence, which has a close relationship with the one in [19].

To solve (1.1), although we can use minimization algorithms to solve the sum of two non-convex functions, we have to calculate the gradient of $F+H$ or the proximity operator of $F+G$, and the resulting algorithm obviously does not fully utilize the separable structure of problem (1.1). To overcome this drawback, Bian and Zhang [26] employed a there-operator splitting algorithm proposed by Davis and Yin [27] to solve problem (1.1), which is defined by

$$\begin{cases} y^{t+1} = \arg\min\limits_{y\in\mathscr{H}} \{F(y) + \dfrac{1}{2\gamma}\|y - x^t\|^2\}, \\ z^{t+1} \in \arg\min\limits_{z\in\mathscr{H}} \{G(z) + \dfrac{1}{2\gamma}\|z - (2y^{t+1} - \gamma\nabla H(y^{t+1}) - x^t)\|^2\}, \\ x^{t+1} = x^t + (z^{t+1} - y^{t+1}). \end{cases} \tag{1.3}$$

Since functions F, G, and H are non-convex, the convergence of (1.3) can not be derived from the original three-operator splitting algorithm. To tackle this difficulty, Bian and Zhang constructed a non-increasing energy function and established the global convergence of (1.3) under certain conditions on the iterative parameters. Inspired by the parameterized Douglas-Rachford splitting algorithm, Zhang and Chen [28] proposed a parameterized three-operator splitting algorithm for solving problem (1.1) in the convex setting. They proved the convergence of the proposed algorithm under mild assumptions. The numerical results in [28] confirmed that the advantage of the parameterized algorithm over the original non-parameterized algorithm.

The parameterized three-operator splitting algorithm of Zhang and Chen [28] is a generalization of the Davis-Yin three-operator splitting algorithm [27]. It is an interesting topic to investigate whether the parameterized three-operator splitting algorithm can be extended to solve non-convex minimization problems. The difficulty lies in how to construct a suitable energy function with a non-increasing property. Based on the parameterized method, we propose a parameterized three-operator splitting algorithm to solve non-convex minimization problem (1.1). Compared with the three-operator splitting algorithm, our proposed splitting method employs a parameterization technique that provides more flexibility for the iterative parameters, allowing the algorithm to be better adapted to different non-convex problems. We construct a monotone non-increasing energy function with the Kurdyka-Łojasiewicz (KL) property to prove the convergence of the algorithm. Our assumptions are weaker than the original three-operator splitting algorithm in the non-convex case. We do not need to assume that the function is coercive to prove that the sequence is bounded. Finally, we apply the parameterized splitting algorithm to low-rank matrix recovery and low-rank image inpainting problems, and show through numerical experiments that our proposed algorithm outperforms the three-operator splitting algorithm in several aspects.

This paper is organized as follows. In Section 2, we provide basic definitions and lemmas required for our analysis. In Section 3, we present the parameterized three-operator partitioning algorithm and establish its convergence analysis in the non-convex case. In Section 4, we report numerical experiments and results. Finally, we provide some conclusions in Section 5.

2. Preliminaries

In this section, $\mathbb{R}$ is defined as the set of real numbers and $\mathbb{N}$ is defined as the set of non-negative integers. $\mathbb{R}^n$ and $\mathscr{H}$ are n-dimensional Euclidean space and real Hilbert space, respectively. The extended real line as $\bar{\mathbb{R}} = R \cup \{\infty\}$, and a function $f : \mathscr{H} \to \bar{\mathbb{R}}$ is called an extended real-valued function. The inner product is defined as $\langle \cdot, \cdot \rangle$, and $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ is the norm. The symbols $\to$ and $\rightharpoonup$ indicate strong convergence and weak convergence, respectively. We say that the function f is proper if for any $x \in \text{dom} f$, $f(x) > -\infty$, and there exists $x \in \text{dom} f$ such that $f(x) < +\infty$.

For any $a, b, c, d \in \mathscr{H}$, there are the following fundamental equalities:

$$\begin{cases} 2\langle a-b, b-c\rangle = \|a-c\|^2 - \|a-b\|^2 - \|b-c\|^2; & \text{(2.1a)} \\ \|2a-b-c\|^2 - \|a-c-d\|^2 = \|a-c\|^2 - \|b-c\|^2 + 2\|a-b\|^2 + 2\langle d, b-a\rangle. & \text{(2.1b)} \end{cases}$$

Definition 2.1. ([29]) Let f be an extended real-valued function.
(i) If $\text{epi} f = \{(x, \xi) \in \mathscr{H} \times \mathbb{R} | f(x) \leq \xi\}$ is a closed set, then f is said to be a closed function;
(ii) If, for any $x \in \text{dom} f$, $\liminf_{y \to x} f(y) \geq f(x)$, then f is said to be lower semi-continuous.

In fact, let f be an extended real-valued function. Then f is closed function if and only if f is a lower semi-continuous function.

Definition 2.2. ([13]) Let $f : \mathscr{H} \to \bar{\mathbb{R}}$ be a proper lower semi-continuous function.
(i) For a given $x \in \text{dom} f$, the set of all vectors u that satisfy the following conditions is defined

as the *Fréchet* sub-differential of f at point x:

$$\liminf_{y\to x, y\neq x} \frac{f(y)-f(x)-\langle u, y-x\rangle}{\|y-x\|} \geq 0,$$

which is written as $\hat{\partial} f(x)$. When $x \neq 0$, we define $\hat{\partial} f(x)$ as the empty set.
(ii) The limit sub-differential of f at point x is defined as:

$$\partial f(x) = \{u \in \mathbb{R}^n : \exists x^k \to x, f(x^k) \to f(x), u^k \in \hat{\partial} f(x^k) \to u\}. \tag{2.2}$$

Lemma 2.1. ([29]) *Let $f : \mathscr{H} \to \bar{\mathbb{R}}$ be a proper lower semi-continuous function. If f is convex, then ∂f is monotone, i.e., $\langle x-y, u-v\rangle \geq 0$ for all $x, y \in \mathscr{H}, u \in \partial f(x), v \in \partial f(y)$.*

Definition 2.3. ([30]) Let $f : \mathscr{H} \to \bar{\mathbb{R}}$ be a gradient Lipschitz continuous function, that is, there exists a constant $L > 0$ such that $\|\nabla f(y_1) - \nabla f(y_2)\| \leq L\|y_1 - y_2\|$ for all $y_i \in \mathbb{R}^n$, $i = 1, 2$. For any $x, y \in \mathscr{H}$, we have $f(y) \leq f(x) + \langle \nabla f(x), y-x\rangle + \frac{L}{2}\|x-y\|^2$, where L is the Lipschitz constant of ∇f.

Definition 2.4. ([29]) Let $f : \mathscr{H} \to \bar{\mathbb{R}}$ be a proper lower semi-continuous function. The proximal mapping of f with index $\gamma > 0$ is defined as follows:

$$\text{prox}_{\gamma f}(x) = \arg\min_{u\in \text{dom} f} \{f(u) + \frac{1}{2\gamma}\|u-x\|^2\}.$$

Definition 2.5. ([29]) For a non-empty closed set C, its indicator function is defined as follows:

$$I_C(x) = \begin{cases} 0, & x \in C \\ +\infty, & x \notin C \end{cases}$$

Definition 2.6. ([13]) (Φ_η function class) Define Φ_η as a set of concave continuous function $\varphi : [0, \eta) \to R_+$ with the following conditions:
(i) $\varphi(0) = 0$;
(ii) φ is continuously different in $(0, \eta)$, and continuous at 0;
(iii) $\varphi'(s) > 0$ for all $s \in (0, \eta)$.

Definition 2.7. ([13]) (Kurdyka-Łojasiewicz (KL) property) Let $f : \mathscr{H} \to \bar{\mathbb{R}}$ be a proper lower semi-continuous function.
(i) f is said to have the KL property at a given point $\bar{u} \in \text{dom}\partial f \overset{def}{=} \{u|\partial f(u) \neq \emptyset\}$ if there exists $\eta \in (0, +\infty]$, a domain U of $\bar{u}$ and the function $\varphi \in \Phi_\eta$, such that, for all $u \in U \cap [f(\bar{u}) < f < f(\bar{u}) + \eta]$, the following inequality holds: $\varphi'(f(u) - f(\bar{u})) \cdot \text{dist}(0, \partial f(u)) \geq 1$, where $\text{dist}(x, S)$ is the distance from point x to the set S.
(ii) f is said to be a KL function if f satisfies the KL property everywhere on $\text{dom}\partial f$.

3. Main Algorithm and Convergence Analysis

In this section, we present the main algorithm for solving non-convex minimization problem (1.1) and provide a proof of its convergence. The proposed algorithm utilizes the three-operator splitting algorithm and its generalization of parameterized method.

Algorithm 1 Parameterized Three-Operator splitting algorithm.

1: Choose parameters:$\gamma > 0, \alpha \in (\frac{3}{2}, 2]$;
2: Initialization: $x^0 \in \mathbb{R}^n$;
3: General step: for any $t = 0, 1, \cdots$, execute the following:

$$\begin{cases} y^{t+1} = \arg\min_{y \in \mathscr{H}} \{F(y) + \frac{1}{2\gamma}\|y - x^t\|^2\}, & \text{(3.1a)} \\ z^{t+1} \in \arg\min_{z \in \mathscr{H}} \{G(z) + \frac{1}{2\gamma}\|z - (\alpha y^{t+1} - \gamma \nabla H(y^{t+1}) - x^t)\|^2\}, & \text{(3.1b)} \\ x^{t+1} = x^t + (z^{t+1} - y^{t+1}). & \text{(3.1c)} \end{cases}$$

3.1. **Parameterized three-operator splitting algorithm.** First, we present Algorithm 1 which outlines the parameterized three-operator splitting algorithm.

Based on the first-order optimality conditions of ($3.1a$) and ($3.1b$), we obtain:

$$\begin{cases} 0 = \nabla F(y^{t+1}) + \frac{1}{\gamma}(y^{t+1} - x^t), & \text{(3.2a)} \\ 0 \in \partial G(z^{t+1}) + \frac{1}{\gamma}(z^{t+1} - \alpha y^{t+1} + \gamma \nabla H(y^{t+1}) - x^t),. & \text{(3.2b)} \end{cases}$$

We now give some assumptions for the function in problem (1.1).

Assumption 3.1**.**
(A1) The function F has Lipschitz continuous gradient with constant L;
(A2) The function G is a proper lower semi-continuous function with a nonempty mapping $\text{prox}_{\gamma G}(x)$ for any x and $\gamma > 0$;
(A3) The function H has Lipschitz continuous gradient with constant β.

Remark 3.1. (i) When H is zero, Algorithm 1 reduces to the parameterized Douglas-Rachford algorithm [24]. In [24], the assumptions on functions F and G are the same as ours. Moreover, when $\alpha = 2$, Algorithm 1 becomes the standard Douglas-Rachford Splitting algorithm. For more details on this algorithm, we refer to [19].

(ii) There exists a constant l such that $F(\cdot) + \frac{l}{2}\|\cdot\|_2^2$ is a convex function. In particular, l can be taken as L, where L is the gradient Lipschitz constant of the function F.

3.2. **Convergence analysis.** To establish convergence, we use the following energy function associated with Assumption 1 as follows:

$$\begin{aligned} \theta_\gamma(x,y,z) = {} & F(y) + G(z) + H(y) + \frac{1}{2\gamma}\|x - \alpha y + z + \gamma \nabla H(y)\|^2 - \frac{1}{\gamma}\|y - z\|^2 \\ & - \frac{1}{2\gamma}\|x - (\alpha - 1)y + \gamma \nabla H(y)\|^2 + \frac{2 - \alpha}{2\gamma}\|y\|^2. \end{aligned} \tag{3.3}$$

Lemma 3.1. *(Sufficient decrease on the energy function) Suppose that functions F, G, and H satisfy Assumption 3.1. Let $\frac{3}{2} < \alpha \leq 2$, and let $\{(x^t, y^t, z^t)\}$ be a sequence generated by Algorithm 1. Then, for all $t \geq 1$, $\theta_\gamma(x^{t+1}, y^{t+1}, z^{t+1}) - \theta_\gamma(x^t, y^t, z^t) \leq -\Lambda(\gamma)\|y^{t+1} - y^t\|^2$, where*

$$\Lambda(\gamma) = -\frac{2-\alpha}{2\gamma} - \beta + \frac{1}{2}(\frac{1}{\gamma} - l) - \frac{4-\alpha+\beta\gamma}{2\gamma}[(2\gamma l - 1) + (1+\gamma L)^2].$$

Furthermore, if there exist parameters $\gamma > 0$ such that $\Lambda(\gamma) \geq 0$, then the sequence $\theta_\gamma(x^t, y^t, z^t)$ is non-increasing.

Proof. Given that $F(\cdot) + \frac{1}{2\gamma}\|x^t - \cdot\|^2$ is $(\frac{1}{\gamma} - l)$-strongly convex, and y^{t+1} is a minimizer of (3.1a), we have

$$F(y^{t+1}) + \frac{1}{2\gamma}\|y^{t+1} - x^t\|^2 \leq F(y^t) + \frac{1}{2\gamma}\|y^t - x^t\|^2 - \frac{1}{2}(\frac{1}{\gamma} - l)\|y^{t+1} - y^t\|^2. \tag{3.4}$$

Since z^{t+1} minimizes of (3.1b), we obtain

$$G(z^{t+1}) + \frac{1}{2\gamma}\|z^{t+1} - \alpha y^{t+1} + \gamma\nabla H(y^{t+1}) + x^t\|^2 \leq G(z^t) + \frac{1}{2\gamma}\|z^t - \alpha y^{t+1} + \gamma\nabla H(y^{t+1}) + x^t\|^2 \tag{3.5}$$

By adding inequality (3.4) to inequality (3.5), we arrive at

$$\begin{aligned} &F(y^{t+1}) + G(z^{t+1}) + \frac{1}{2\gamma}\|y^{t+1} - x^t\|^2 + \frac{1}{2\gamma}\|z^{t+1} - \alpha y^{t+1} + \gamma\nabla H(y^{t+1}) + x^t\|^2 \\ &\leq F(y^t) + G(z^t) + \frac{1}{2\gamma}\|y^t - x^t\|^2 + \frac{1}{2\gamma}\|z^t - \alpha y^{t+1} + \gamma\nabla H(y^{t+1}) + x^t\|^2 - \frac{1}{2}(\frac{1}{\gamma} - l)\|y^{t+1} - y^t\|^2 \end{aligned} \tag{3.6}$$

Observing the last term on the left-hand side of the inequality above, we have

$$\begin{aligned} &\|\alpha y^{t+1} - z^{t+1} - \gamma\nabla H(y^{t+1}) - x^t\|^2 \\ &= \|\alpha y^{t+1} - x^{t+1} - z^{t+1} - \gamma\nabla H(y^{t+1})\|^2 + \|x^{t+1} - x^t\|^2 \\ &\quad + 2\langle \alpha y^{t+1} - x^{t+1} - z^{t+1} - \gamma\nabla H(y^{t+1}), x^{t+1} - x^t\rangle \\ &= \|\alpha y^{t+1} - x^{t+1} - z^{t+1} - \gamma\nabla H(y^{t+1})\|^2 + \|x^{t+1} - x^t\|^2 + 2\langle y^{t+1} - z^{t+1}, x^{t+1} - x^t\rangle \\ &\quad + 2\langle (\alpha-1)y^{t+1} - x^{t+1} - \gamma\nabla H(y^{t+1}), x^{t+1} - x^t\rangle \\ &= \|\alpha y^{t+1} - x^{t+1} - z^{t+1} - \gamma\nabla H(y^{t+1})\|^2 + \|x^{t+1} - x^t\|^2 - 2\|x^{t+1} - x^t\|^2 \\ &\quad - \|(\alpha-1)y^{t+1} - x^{t+1} - \gamma\nabla H(y^{t+1})\|^2 - \|x^{t+1} - x^t\|^2 + \|(\alpha-1)y^{t+1} - x^t - \gamma\nabla H(y^{t+1})\|^2 \\ &= \|\alpha y^{t+1} - x^{t+1} - z^{t+1} - \gamma\nabla H(y^{t+1})\|^2 - 2\|x^{t+1} - x^t\|^2 - \|(\alpha-1)y^{t+1} - x^{t+1} - \gamma\nabla H(y^{t+1})\|^2 \\ &\quad + \|(\alpha-1)y^{t+1} - x^t - \gamma\nabla H(y^{t+1})\|^2, \end{aligned} \tag{3.7}$$

where the third equality makes use of (3.1c) and (2.1a). Substituting(3.7) into (3.6), we obtain

$$\begin{aligned}
&F(y^{t+1})+G(z^{t+1})+\frac{1}{2\gamma}\|\alpha y^{t+1}-x^{t+1}-z^{t+1}-\gamma\nabla H(y^{t+1})\|^2\\
&\qquad -\frac{1}{2\gamma}\|(\alpha-1)y^{t+1}-x^{t+1}-\gamma\nabla H(y^{t+1})\|^2-\frac{1}{\gamma}\|y^{t+1}-z^{t+1}\|^2\\
&\leq F(y^t)+G(z^t)+\frac{1}{2\gamma}\|y^t-x^t\|^2+\frac{1}{2\gamma}\|y^{t+1}-z^t\|^2+\frac{1}{\gamma}\langle(\alpha-1)y^{t+1}-x^t,y^{t+1}-z^t\rangle\\
&\qquad +\langle\nabla H(y^{t+1}),z^t-y^{t+1}\rangle-\frac{1}{2\gamma}\|y^{t+1}-z^t+z^t-x^t\|^2-\frac{1}{2}(\frac{1}{\gamma}-l)\|y^{t+1}-y^t\|^2\\
&=F(y^t)+G(z^t)+\frac{1}{2\gamma}\|y^t-x^t\|^2+\frac{1}{\gamma}\langle(\alpha-1)y^{t+1}-x^t,y^{t+1}-z^t\rangle+\langle\nabla H(y^{t+1}),z^t-y^{t+1}\rangle\\
&\qquad +\langle\nabla H(y^t),z^t-y^t\rangle-\langle\nabla H(y^t),z^t-y^t\rangle-\frac{1}{2\gamma}\|z^t-x^t\|^2-\frac{1}{\gamma}\langle y^{t+1}-z^t,z^t-x^t\rangle\\
&\qquad -\frac{1}{2}(\frac{1}{\gamma}-l)\|y^{t+1}-y^t\|^2\\
&=F(y^t)+G(z^t)+\frac{1}{2\gamma}\|\alpha y^t-z^t-x^t-\gamma\nabla H(y^t)\|^2-\frac{1}{2\gamma}\|(\alpha-1)y^t-x^t-\gamma\nabla H(y^t)\|^2\\
&\qquad +\frac{1}{\gamma}\langle y^t-z^t,(2-\alpha)y^t\rangle-\frac{1}{\gamma}\|y^t-z^t\|^2+\frac{1}{\gamma}\langle(\alpha-1)y^{t+1}-x^t,y^{t+1}-z^t\rangle\\
&\qquad +\langle\nabla H(y^{t+1}),z^t-y^{t+1}\rangle-\langle\nabla H(y^t),z^t-y^t\rangle-\frac{1}{\gamma}\langle y^{t+1}-z^t,z^t-x^t\rangle-\frac{1}{2}(\frac{1}{\gamma}-l)\|y^{t+1}-y^t\|^2
\end{aligned} \tag{3.8}$$

where the third equality is obtained by utilizing formula (2.1b) with $a=y^t,b=z^t,c=0$, and $d=\gamma\nabla H(y^t)$. Now, we consider the term $\langle\nabla H(y^{t+1}),z^t-y^{t+1}\rangle-\langle\nabla H(y^t),z^t-y^t\rangle$ in (3.8)

$$\begin{aligned}
&\langle\nabla H(y^{t+1}),z^t-y^{t+1}\rangle-\langle\nabla H(y^t),z^t-y^t\rangle\\
&\leq\langle\nabla H(y^{t+1})-\nabla H(y^t),z^t-y^{t+1}\rangle+H(y^{t+1})-H(y^t)+\frac{\beta}{2}\|y^{t+1}-y^t\|^2\\
&\leq H(y^{t+1})-H(y^t)+\beta\|y^{t+1}-y^t\|^2+\frac{\beta}{2}\|y^{t+1}-z^t\|^2,
\end{aligned} \tag{3.9}$$

where the first inequality is derived from the descent lemma, and the last inequality holds due to the Lipschitz continuous of ∇H and the Cauchy-Schwarz inequality.

Next, we analyze the following inner product terms to further understand the behaviour of the algorithm

$$\begin{aligned}
&\frac{1}{\gamma}\langle y^t-z^t,(2-\alpha)y^t\rangle+\frac{1}{\gamma}\langle(\alpha-1)y^{t+1}-x^t,y^{t+1}-z^t\rangle-\frac{1}{\gamma}\langle y^{t+1}-z^t,z^t-x^t\rangle\\
&=\frac{2-\alpha}{2\gamma}\|y^t-z^t\|^2+\frac{2-\alpha}{2\gamma}\|y^t\|^2-\frac{2-\alpha}{2\gamma}\|z^t\|^2+\frac{\alpha-1}{2\gamma}\|y^{t+1}-z^t\|^2+\frac{\alpha-1}{2\gamma}\|y^{t+1}\|^2\\
&\qquad -\frac{\alpha-1}{2\gamma}\|z^t\|^2+\frac{1}{2\gamma}\|y^{t+1}-z^t\|^2+\frac{1}{2\gamma}\|z^t\|^2-\frac{1}{2\gamma}\|y^{t+1}\|^2\\
&\leq\frac{2-\alpha}{\gamma}\|y^t-y^{t+1}\|^2+\frac{4-\alpha}{2\gamma}\|y^{t+1}-z^t\|^2+\frac{2-\alpha}{2\gamma}(\|y^t\|^2-\|y^{t+1}\|^2).
\end{aligned} \tag{3.10}$$

Then, we focus on $\|y^{t+1}-z^t\|^2$. Since $F(\cdot)+\frac{l}{2}\|\cdot\|^2$ is a convex function, we know that $\nabla(F(\cdot)+\frac{l}{2}\|\cdot\|^2)$ is monotone. Hence, we have

$$\langle \nabla(F(y^{t+1})+\frac{l}{2}\|y^{t+1}\|^2)-\nabla(F(y^t)+\frac{l}{2}\|y^t\|^2), y^{t+1}-y^t\rangle \geq 0.$$

By taking advantage of (3.2a), we can simplify the expression and obtain

$$\langle (\frac{1}{\gamma}(x^t-y^{t+1})+ly^{t+1})-(\frac{1}{\gamma}(x^{t-1}-y^t)+ly^t, y^{t+1}-y^t)\rangle \geq 0.$$

Therefore, we obtain

$$\langle y^{t+1}-y^t, x^t-x^{t-1}\rangle \geq (1-\gamma l)\|y^{t+1}-y^t\|^2. \tag{3.11}$$

By the first-order optimality condition of (3.1a) and the Lipschitz continuity of ∇F, we can conclude that

$$\|\nabla F(y^{t+1})-\nabla F(y^t)\|^2 = \|\frac{1}{\gamma}(x^t-y^{t+1})-\frac{1}{\gamma}(x^{t-1}-y^t)\| \leq L\|y^{t+1}-y^t\|.$$

By applying the triangle inequality, we obtain

$$\|x^t-x^{t-1}\|-\|y^{t+1}-y^t\| \leq \|(x^t-y^{t+1})-(x^{t-1}-y^t)\| \leq \gamma L\|y^{t+1}-y^t\|.$$

Thus, we can establish the relationship between $\|y^{t+1}-y^t\|$ and $\|x^t-x^{t-1}\|$ as follows:

$$\|x^t-x^{t-1}\| \leq (1+\gamma L)\|y^{t+1}-y^t\|. \tag{3.12}$$

By combining equations (3.11) and (3.12), we deduce that

$$\begin{aligned}\|y^{t+1}-z^t\|^2 &= \|y^{t+1}-y^t\|^2+\|x^t-x^{t-1}\|^2-2\langle y^{t+1}-y^t, x^t-x^{t-1}\rangle \\ &\leq (2\gamma l-1)\|y^{t+1}-y^t\|^2+\|x^t-x^{t-1}\|^2 \\ &\leq [(2\gamma l-1)+(1+\gamma L)^2]\|y^{t+1}-y^t\|^2.\end{aligned} \tag{3.13}$$

By substituting equations (3.9), (3.10), and (3.13) into equation (3.7), we see that

$$\begin{aligned}
&F(y^{t+1})+G(z^{t+1})+H(y^{t+1})+\frac{1}{2\gamma}\|\alpha y^{t+1}-x^{t+1}-z^{t+1}-\gamma\nabla H(y^{t+1})\|^2 \\
&\quad -\frac{1}{2\gamma}\|(\alpha-1)y^{t+1}-x^{t+1}-\gamma\nabla H(y^{t+1})\|^2-\frac{1}{\gamma}\|y^{t+1}-z^{t+1}\|^2+\frac{2-\alpha}{2\gamma}\|y^{t+1}\|^2 \\
&\leq F(y^t)+G(z^t)+H(y^t)+\frac{1}{2\gamma}\|\alpha y^t-x^t-z^t-\gamma\nabla H(y^t)\|^2 \\
&\quad -\frac{1}{2\gamma}\|(\alpha-1)y^t-x^t-\gamma\nabla H(y^t)\|^2-\frac{1}{\gamma}\|y^t-z^t\|^2+\frac{2-\alpha}{2\gamma}\|y^t\|^2 \\
&\quad +\Big[\frac{2-\alpha}{\gamma}+\beta-\frac{1}{2}(\frac{1}{\gamma}-l)\Big]\|y^{t+1}-y^t\|^2+\frac{4-\alpha+\beta\gamma}{2\gamma}\|y^{t+1}-z^t\|^2 \\
&\leq F(y^t)+G(z^t)+H(y^t)+\frac{1}{2\gamma}\|\alpha y^t-x^t-z^t-\gamma\nabla H(y^t)\|^2 \\
&\quad -\frac{1}{2\gamma}\|(\alpha-1)y^t-x^t-\gamma\nabla H(y^t)\|^2-\frac{1}{\gamma}\|y^t-z^t\|^2+\frac{2-\alpha}{2\gamma}\|y^t\|^2 \\
&\quad +\Big[\frac{2-\alpha}{\gamma}+\beta-\frac{1}{2}(\frac{1}{\gamma}-l)+\frac{4-\alpha+\beta\gamma}{2\gamma}[(2\gamma l-1)+(1+\gamma L)^2]\Big]\|y^{t+1}-y^t\|^2.
\end{aligned}$$

Thus we can confirm that the merit function satisfies the following condition

$$\theta_\gamma(x^{t+1},y^{t+1},z^{t+1}) - \theta_\gamma(x^t,y^t,z^t) \leq -\Lambda(\gamma)\|y^{t+1}-y^t\|^2, \tag{3.14}$$

where $\Lambda(\gamma) = -\frac{2-\alpha}{\gamma} - \beta + \frac{1}{2}(\frac{1}{\gamma} - l) - \frac{4-\alpha+\beta\gamma}{2\gamma}[(2\gamma l-1)+(1+\gamma L)^2]$. If $\frac{3}{2} \leq \alpha \leq 2$, then $\lim_{\gamma\to 0}\gamma\Lambda(\gamma) = \frac{2\alpha-3}{2} \geq 0$ is satisfied, which means that the merit function $\theta_\gamma(x^t,y^t,z^t)$ is non-increasing. □

Based on the proof presented above, we see that the energy function is non-increasing. Under specific assumptions, we are able to prove that the sequence generated by Algorithm 1 is bounded.

Theorem 3.1. *(Boundedness of the sequence). Assuming that Assumption 3.1 is satisfied and the parameter γ in Algorithm 1 is such that $\Lambda(\gamma) > 0$. Suppose that $\frac{3}{2} < \alpha \leq 2$ and the function F, G, and H are both bounded below. Then, the sequence $\{(x^t,y^t,z^t)\}$ generated by Algorithm 1 is bounded.*

Proof. It is obvious that merit function (3.3) is equivalent to the following form:

$$\begin{aligned}\theta_\gamma(x^t,y^t,z^t) &= F(y^t)+G(z^t)+H(y^t)+\frac{1}{2\gamma}\|\alpha y^t - x^t - z^t - \gamma\nabla H(y^t)\|^2 \\ &\quad -\frac{1}{2\gamma}\|(\alpha-1)y^t - x^t - \gamma\nabla H(y^t)\|^2 - \frac{1}{\gamma}\|y^t-z^t\|^2 + \frac{2-\alpha}{2\gamma}\|y^t\|^2 \\ &= F(y^t)+G(z^t)+H(y^t)+\frac{1}{2\gamma}(\|x^t-y^t\|^2 - \|x^t-z^t\|^2) \\ &\quad \langle\nabla H(y^t), z^t-y^t\rangle + \frac{2-\alpha}{2\gamma}\|y^t\|^2 + \frac{2-\alpha}{2\gamma}\langle y^t, z^t-y^t\rangle.\end{aligned} \tag{3.15}$$

Since F and H are proper functions, we can conclude by applying the descent lemma that there exist $\xi^* > -\infty$ and $\zeta^* > -\infty$ such that

$$\begin{aligned}\xi^* &\leq F(x-\frac{1}{L}\nabla F(x)) \leq F(x) + \langle\nabla F(x),(x-\frac{1}{L}\nabla F(x))-x\rangle + \frac{L}{2}\|(x-\frac{1}{L}\nabla F(x))-x\|^2 \\ &= F(x) - \frac{1}{2L}\|\nabla F(x)\|^2.\end{aligned} \tag{3.16}$$

Similarly,

$$\zeta^* \leq H(x-\frac{1}{\beta}\nabla H(x)) \leq H(x) - \frac{1}{2\beta}\|\nabla F(x)\|^2. \tag{3.17}$$

By using the Cauchy-Schwarz inequality and (3.1c), we obtain

$$\begin{aligned}\langle\nabla H(x), z^t-y^t\rangle &\geq -\frac{1}{4\beta}\|\nabla H(x)\|^2 - \beta\|z^t-y^t\|^2 \\ &\geq -\frac{1}{4\beta}\|\nabla H(x)\|^2 - \beta\|x^t-x^{t-1}\|^2 \\ &\geq -\frac{1}{4\beta}\|\nabla H(x)\|^2 - 2\beta\|x^t-y^t\|^2 - 2\beta\|y^t-x^{t-1}\|^2.\end{aligned} \tag{3.18}$$

Based on equations (3.15)∼(3.18) and Theorem 3.1, it can be concluded that

$$\begin{aligned}
&\theta_\gamma(x^1,y^1,z^1)\\
&\geq F(y^t)+G(z^t)+H(y^t)+\frac{1}{2\gamma}(\|x^t-y^t\|^2-\|x^t-z^t\|^2)\\
&\quad+\langle\nabla H(y^t),z^t-y^t\rangle-\frac{2-\alpha}{2\gamma}\|z^t-y^t\|^2+\frac{2-\alpha}{2\gamma}\|z^t\|^2\\
&\geq F(y^t)+G(z^t)+H(y^t)+\frac{1}{2\gamma}(\|x^t-y^t\|^2-\|x^t-z^t\|^2)+\langle\nabla H(y^t),z^t-y^t\rangle\\
&\quad-\frac{2-\alpha}{\gamma}(\|z^t-x^t\|^2+\|x^t-y^t\|^2)+\frac{2-\alpha}{2\gamma}\|z^t\|^2\\
&\geq F(y^t)+G(z^t)+H(y^t)+\frac{2\alpha-3}{2\gamma}\|x^t-y^t\|^2-\frac{5-2\alpha}{2\gamma}\|x^{t-1}-y^t\|^2\\
&\quad-\frac{1}{4\beta}\|\nabla H(y^t)\|^2-2\beta\|x^t-y^t\|^2-2\beta\|y^t-x^{t-1}\|^2+\frac{2-\alpha}{2\gamma}\|z^t\|^2\\
&=\mu F(y^t)+(1-\mu)F(y^t)+G(z^t)+\nu H(y^t)+(1-\nu)H(y^t)-\frac{1-\nu}{2\beta}\|\nabla H(y^t\|^2\\
&\quad+\frac{1-2\nu}{4\beta}\|\nabla H(y^t\|^2+[\frac{1-\mu}{2L}-(\frac{5-2\alpha}{2\gamma}+2\beta)\gamma^2]\|\nabla F(y^t)\|^2-\frac{1-\mu}{2L}\|\nabla F(y^t)\|^2\\
&\quad+(\frac{2\alpha-3}{2\gamma}-2\beta)\|y^t-x^t\|^2+\frac{2-\alpha}{2\gamma}\|z^t\|^2\\
&\geq\mu F(y^t)+(1-\mu)\xi^*+G(z^t)+\nu H(y^t)+(1-\nu)\zeta^*+\frac{1-2\nu}{4\beta}\|\nabla H(y^t)\|^2\\
&\quad+[\frac{1-\mu}{2L}-(\frac{5-2\alpha}{2\gamma}+2\beta)\gamma^2]\|\nabla F(y^t)\|^2+(\frac{2\alpha-3}{2\gamma}-2\beta)\|y^t-x^t\|^2+\frac{2-\alpha}{2\gamma}\|z^t\|^2.
\end{aligned}\tag{3.19}$$

We can choose the parameter $\gamma>0$ to be small enough and select μ and ν from the range of $(0,1)$ such that

$$\frac{1-2\nu}{4\beta},\frac{1-\mu}{2L}-(\frac{5-2\alpha}{2\gamma}+2\beta)\gamma^2,\frac{2\alpha-3}{2\gamma}-2\beta>0.$$

Based on equation (3.19), we can readily determine that $\{z^t\}$, $\{\nabla F(y^t)\}$, and $\{x^t-y^t\}$ are bounded. Additionally, using equations (3.2b) and (3.1c), we can conclude that $\{z^t-y^t\}$ is also bounded, which implies that $\{y^t\}$ is bounded. Finally, since $\{x^t-y^t\}$ is bounded, we can infer that $\{x^t\}$ is also bounded. □

Remark 3.2. It can be observed from the proof of the boundedness of the sequence produced by Algorithm 1 that there is no requirement to assume the coercivity of functions such as F, G, or H. This indicates that our algorithm has less stringent assumptions in comparison to the non-convex three-operator splitting algorithm studied by Bian and Zhang [26].

Next, we prove the convergence of the subsequence generated by Algorithm 1.

Theorem 3.2. *(Subsequential convergence) Assuming that $\frac{3}{2}<\alpha\leq 2$ and that Assumption 3.1 is fulfilled. Let γ be a parameter in Algorithm 1 such that $\Lambda(\gamma)>0$. Moreover, if a cluster of*

the point sequence $\{(x^t, y^t, z^t)\}$ *exist, then the following assertions hold*
(i)

$$\lim_{t\to\infty} \|y^{t+1} - y^t\| = \lim_{t\to\infty} \|x^{t+1} - x^t\| = \lim_{t\to\infty} \|z^{t+1} - y^{t+1}\|; \tag{3.20}$$

(ii) any cluster point (x^*, y^*, z^*) *of sequence* $\{(x^t, y^t, z^t)\}$ *generated by Algorithm 1 satisfies:*

$$0 \in \nabla F(z^*) + \partial G(z^*) + \nabla H(z^*) + \frac{2-\alpha}{\gamma} z^*.$$

Proof. (i) Summing inequality (3.14) from $t = 1$ to $N-1$, we see that

$$\theta_\gamma(x^N, y^N, z^N) - \theta_\gamma(x^1, y^1, z^1) \le -\Lambda(\gamma) \sum_{t=1}^{N} \|y^{t+1} - y^t\|^2$$

Let (x^*, y^*, z^*) be a cluster point of the sequence $\{(x^t, y^t, z^t)\}$ such that there exists a subsequence $(x^{t_j}, y^{t_j}, z^{t_j})$ such that $\lim_{j\to\infty}(x^{t_j}, y^{t_j}, z^{t_j}) = (x^*, y^*, z^*)$. Since θ_γ is lower semi-continuous and F and G are proper functions. Let $N = t_j$. As $j \to \infty$, we obtain

$$-\infty < \theta_\gamma(x^*, y^*, z^*) - \theta_\gamma(x^1, y^1, z^1) \le -\Lambda(\gamma) \sum_{t=1}^{\infty} \|y^{t+1} - y^t\|^2,$$

which implies that $\lim_{t\to\infty} \|y^{t+1} - y^t\|^2 = 0$. Combining (3.1c) and (3.14), we can conclude that $\lim_{t\to\infty} \|z^{t+1} - y^{t+1}\| = \lim_{t\to\infty} \|x^{t+1} - x^t\| = 0$.

(ii) Suppose that $\{(x^{t_j}, y^{t_j}, z^{t_j})\}$ is a convergent subsequence such that

$$\lim_{j\to\infty} (x^{t_j}, y^{t_j}, z^{t_j}) = \lim_{j\to\infty} (x^{t_j-1}, y^{t_j-1}, z^{t_j-1}) = (x^*, y^*, z^*). \tag{3.21}$$

It follows from (3.1b) that

$$G(z^t) + \frac{1}{2\gamma} \|z^t - \alpha y^t - \gamma \nabla H(y^t) - x^t\|^2 \le G(z^*) + \frac{1}{2\gamma} \|z^* - \alpha y^t - \gamma \nabla H(y^t) - x^t\|^2.$$

Replacing z^t with z^{t_j} and using (3.21), we have $\limsup_{j\to\infty} G(z^{t_j}) \le G(z^*)$.

On the other hand, we conclude from the lower semi-continuous of G that $\liminf_{j\to\infty} G(z^{t_j}) \ge G(z^*)$. Hence, $\lim_{j\to\infty} G(z^{t_j}) = G(z^*)$.

Finally, according to (2.2), we obtain

$$0 \in \nabla F(z^*) + \partial G(z^*) + \nabla H(z^*) + \frac{2-\alpha}{\gamma} z^*,$$

which ends the proof. □

To prove the global convergence of the sequence generated by Algorithm 1, we need to first verify the following lemma.

Lemma 3.2. *Let Assumption 3.1 be satisfied, and let H be a twice continuous differentiable function with a bounded Hessian matrix, i.e., there exists a constant M such that $\nabla^2 H \le M$. Let $\{(x^t, y^t, z^t)\}_{t\ge 0}$ be the sequence generated by Algorithm 1. Then, there exists a τ such that, for any $t \ge 1$, $dist(0, \partial\theta_\gamma(x^t, y^t, z^t)) \le \tau \|y^{t+1} - y^t\|$.*

Proof. From (3.1c), it is easy to see

$$\begin{aligned}\nabla_x\theta_\gamma(x^t,z^t,z^t) &= \frac{1}{\gamma}(x^t-\alpha y^t+z^t+\gamma\nabla H(y^t)-\frac{1}{\gamma}(x^t-(\alpha-1)y^t+\gamma\nabla H(y^t)\\ &= \frac{1}{\gamma}(z^t-y^t)=\frac{1}{\gamma}(x^t-x^{t-1}).\end{aligned}\tag{3.22}$$

In view of (3.1c) and (3.2a), as well as the boundedness of $\nabla^2 H(y)$, we obtain

$$\begin{aligned}\partial_y\theta(\gamma)(x^t,z^t,z^t) &= \frac{\alpha-1}{\gamma}(y^t-z^t)+\frac{1}{\gamma}(x^{t-1}-x^t)-\nabla^2H(y^t)(y^t-z^t)\\ &\le \frac{\alpha}{\gamma}(x^{t-1}-x^t)+M(x^t-x^{t-1}).\end{aligned}\tag{3.23}$$

By using (3.2b) and (3.1c), we obtain

$$\begin{aligned}\nabla_z\theta_\gamma(x^t,z^t,z^t) &= \partial G(z^t)+\frac{1}{\gamma}(x^t-\alpha y^t+\gamma\nabla H(y^t))+\frac{2}{\gamma}(y^t-z^t)\\ &\ni -\frac{1}{\gamma}(z^t-\alpha y^t+\gamma\nabla H(y^t)+x^{t-1})+\frac{1}{\gamma}(x^t-\alpha y^t+\gamma\nabla H(y^t))+\frac{2}{\gamma}(y^t-z^t)\\ &= \frac{1}{\gamma}(x^{t-1}-x^t).\end{aligned}\tag{3.24}$$

Finally, it follows from (3.22), (3.23), and (3.24) that there exists a constant τ such that

$$dist(0,\partial\theta_\gamma(x^t,z^t,z^t))\le\tau\|y^{t+1}-y^t\|^2,\ \ \forall t\ge 1.$$

□

Based on the lemmas above, it is not difficult to draw the following conclusions. Since the proof is similar to that in [26], we omit it here.

Theorem 3.3. *(Global convergence) Assume that Assumption 3.1 is satisfied and the parameter γ in Algorithm 1 is such that $\Lambda(\gamma)>0$. Suppose that $\frac{3}{2}<\alpha\le 2$ and the energy function θ_γ is a KL function. If the sequence $\{(x^t,y^t,z^t)\}$ generated by Algorithm 1 has a cluster point, then the entire sequence $\{(x^t,y^t,z^t)\}$ converges to a stationary point of (1.1).*

4. Numerical Experiments

In this section, to verify the effectiveness of our proposed algorithm, we apply the parameterized three-operator splitting algorithm to solve the classical low-rank matrix recovery and low-rank image inpainting problems. We compare the numerical results with the three-operator splitting algorithm in non-convex settings. We programmed all the algorithms in MATLAB R2022A software and performed numerical experiments on a laptop computer equipped with a 2.80 GHz Intel Core processor and 16 GB of RAM.

4.1. **Low-rank matrix recovery.** Low-rank matrix restoration has gained widespread use in image processing applications in recent years, including denoising, deblurring, and more [31, 32, 33, 34]. Let $\text{rank}(X)$ be the rank of the matrix X. The problem can be expressed as

$$\begin{aligned}&\min_{X\in\mathbb{R}^{m\times n}}\ \text{rank}(X),\\ &s.t.\ X_{ij}=M_{ij},\ (i,j)\in\Omega,\end{aligned}$$

where X is an $m\times n$ matrix, M is a noisy and incomplete version of X, and Ω is the set of observed entries. This problem is known to be NP-hard. Researchers have proposed various methods to solve it, such as convex relaxation, nuclear norm minimization, and non-convex optimization. Another way to formulate the problem is to express it as follows:

$$\min_{X\in\mathbb{R}^{m\times n}} \frac{1}{2}\|\mathscr{P}_\Omega(X)-\mathscr{P}_\Omega(M)\|_2^2+\mathscr{I}_{\mathscr{C}(r)}(X)+\frac{\lambda}{2}\|X\|_2^2, \tag{4.1}$$

where λ is the regularization parameter, $\mathscr{P}$ is the orthogonal projector onto the span of matrices vanishing outside Ω, and $\mathscr{I}_{\mathscr{C}(r)}$ is the indicator function of the set of matrices with rank at most r. This formulation is equivalent to (1.1) with $F(X)=\frac{1}{2}\|\mathscr{P}_\Omega(X)-\mathscr{P}_\Omega(M)\|_2^2$, $G(X)=\mathscr{I}_{\mathscr{C}(r)}(X)-\frac{2-\alpha}{2\gamma}\|X\|^2$, and $H(X)=\frac{\lambda}{2}\|X\|_2^2$. Moreover, the conditions of convergence of Algorithm 1 are satisfied as F and H are both gradient Lipschitz continuous with Lipschitz constants of $L=1$ and $\beta=\lambda$, respectively, while G is a proper lower semi-continuous function. As a result, we can easily apply our proposed algorithm to solve (4.1) and obtain the following algorithm:

$$\begin{cases} U^{t+1}=\arg\min_U\{\frac{1}{2}\|\mathscr{P}_\Omega(U)-\mathscr{P}_\Omega(M)\|_2^2+\frac{1}{2\gamma}\|U-X^t\|^2\},\\ V^{t+1}=\arg\min_V\{\mathscr{I}_{\mathscr{C}(r)}(V)-\frac{2-\alpha}{2\gamma}\|V\|^2-\frac{2-\alpha}{2\gamma}\|V\|^2\\ \qquad\qquad +\frac{1}{2\gamma}\|V-(\alpha U^{t+1}-\gamma\nabla H(U^{t+1})-X^t)\|^2\},\\ X^{t+1}=X^t+(V^{t+1}-U^{t+1}). \end{cases}$$

Obviously, when $\alpha=2$, the above algorithm degenerates to the original three-operator splitting algorithm. The solution of the above algorithm sub-problem can be obtained analytically through a closed-form expression.

$$\begin{cases} U^{t+1}=\begin{cases}\frac{1}{1+\gamma}(X^t_{i,j}+\gamma M_{i,j}),\ (i,j)\in\Omega,\\ \qquad\qquad X^t_{i,j},\ (i,j)\in\Omega,\end{cases}\\ V^{t+1}=\mathscr{P}_{\mathscr{C}(r)}(\frac{(\alpha-\gamma\lambda)U^{t+1}-X^t}{\alpha-1},r),\\ X^{t+1}=X^t+(V^{t+1}-U^{t+1}). \end{cases}$$

To ensure that $\Lambda(\gamma)>0$ and (3.20) are satisfied, we set $l=0$ and choose $0<\gamma<\gamma_0=\frac{2\alpha-3}{4}$. In practical computations, the value of γ may become very small. Therefore, we adopt a heuristic used in Li and Pong [19] for both the three-operator splitting algorithm and our algorithm. We initialize $\gamma=k*\gamma_0$ and update it as $\max\{\gamma_0-\varepsilon,k_1*\gamma\}$ whenever $\gamma>\gamma_0$. Here, $k_1<1$ ensures that γ decreases monotonically, and the sequence satisfies either $\|U^t-U^{t-1}\|>1000/t$ or $\|U\|_\infty>1e10$. As the number of iterations increases, γ eventually becomes smaller than γ_0, thus ensuring the algorithm's convergence according to Section 3.2. It is also necessary to choose k_1 for this. We set $k=60$ for both algorithms and consider three different values of $\alpha=1.9,1.8$, and 1.6 for our algorithm.

We address the problem of recovering a rank-r $n\times n$ matrix M by using only a small subset of its entries. Specifically, we randomly generate two $n\times r$ matrices M_L and M_R independently, with i.i.d. Gaussian entries. We then define $M = M_L M_R^T$, where T denotes the matrix transpose. Next, we randomly select a set Ω of m entries uniformly at random from all possible entries, and the sampling ratio is denoted as $p = \frac{m}{n^2}$. Our objective is to recover a matrix of the lowest rank that agrees with M on the entries in Ω. In all experiments, we set the stop criterion to

$$Tol := \frac{\|X^{t+1} - X^t\|_F}{\|X^t\|_F} < 1 \times 10^{-7}$$

and the relative error (Err) is defined as

$$Err = \frac{\|X^{opt} - M\|_F}{\|M\|_F},$$

where $\|\cdot\|_F$ represents the Frobenius norm.

Next, we provide the specific parameter selection used to ensure convergence with the lowest relative error. We set λ to 1.5×10^{-6}. The following are the results of our experiments where we reconstructed matrices of rank 2 or 5 across various sizes ($n = 800, 1000, 2000, 3000$) with sampling ratios of $p = 0.05$ or $p = 0.08$. All results are averaged over five runs.

TABLE 1. Numerical Comparison of $p = 0.05$.

Rank	n	$\alpha = 2$	$\alpha = 1.9$	$\alpha = 1.8$	$\alpha = 1.6$
		Time/Iter./Err	Time/Iter./Err	Time/Iter./Err	Time/Iter./Err
2	800	41/4112/1.7377	35/3257/1.5913	26/2374/1.4454	5/487/1.1528
	1000	66/4148/1.7518	56/3294/1.5991	41/2411/1.4464	9/512/1.1427
	2000	293/2979/1.0201	237/2378/0.9418	173/1747/0.8642	26/299/0.6970
	3000	688/2702/0.8874	558/2161/0.8215	410/1590/0.7538	58/220/0.5784
5	800	66/6168/3.3291	56/4871/3.0506	43/3537/2.7697	9/726/2.1981
	1000	131/7034/3.1843	115/5845/2.9035	97/4620/2.6207	40/2032/2.0470
	2000	341/3552/1.3224	281/2844/1.2173	201/2109/1.1134	48/508/0.9064
	3000	701/2960/1.0373	548/2352/0.9570	392/1714/0.8778	62/259/0.7103

Note that Err is counted in units of 10^{-4}. For example, $Err = 1$ corresponds to $Err = 1 \times 10^{-4}$.

TABLE 2. Numerical Comparison of $p = 0.08$.

Rank	n	$\alpha = 2$	$\alpha = 1.9$	$\alpha = 1.8$	$\alpha = 1.6$
		Time/Iter./Err	Time/Iter./Err	Time/Iter./Err	Time/Iter./Err
2	800	24/2582/0.7799	22/2095/0.7166	17/1574/0.6539	3/ 268/0.4652
	1000	40/2308/0.6887	35/1853/0.6347	25/1360/0.5801	3/207/0.3020
	2000	117/1469/0.5428	83/1057/0.4981	47/577/0.4255	16/154/0.2304
	3000	324/1308/0.4975	231/897/0.4501	103/453/0.3004	45/186/0.2217
5	800	36/3311/1.1901	32/2689/1.0902	24/2043/0.9902	8/633/0.7856
	1000	50/2983/0.9033	43/2446/0.8294	33/1883/0.7555	10/591/0.5992
	2000	208/1967/0.6210	156/1537/0.5735	110/1062/0.5214	23/210/0.2264
	3000	338/1568/0.5510	258/1152/0.5071	148/662/0.4453	78/191/0.2233

Note that Err is counted in units of 10^{-4}. For example, $Err = 1$ corresponds to $Err = 1 \times 10^{-4}$.

From Tables 1 and 2, we see that the performance of Algorithm 1 is influenced by both α and p. Generally, a smaller value of α results in better convergence speed and relative error of Algorithm 1. On the other hand, an increase in p leads to lower convergence speed and relative error of Algorithm 1.

Additionally, we discover that Algorithm 1 requires more time and number of iterations when working with large-scale and high-rank matrices. This is because the complexity of the matrix recovery problem increases with the size and rank of the matrix.

Overall, our experimental data results demonstrate that Algorithm 1 can achieve a lower relative error and higher convergence speed than the original three-operator splitting algorithm. This confirms that our proposed algorithm is not only effective but also stable for low-rank matrix recovery.

4.2. **Low-rank image inpainting.** In this subsection, we demonstrate the effectiveness of our algorithm on images of "building" with a size of 517×493.

First, we perform singular value decomposition (SVD) [10, 35, 36] on the image, resulting in a low-rank image I_r. We then select a set Ω of m entries uniformly at random from all possible entries, with a sampling ratio of $p = \frac{m}{493\times517}$.

The main objective is to use the three-operator splitting algorithm and our proposed algorithm to recover an image of the lowest rank that matches the entries in Ω.

For all experiments, we set the stopping criterion as the relative error between the recovered image and the original low-rank image being less than 10^{-5}, that is,

$$Tol := \frac{\|X^{opt} - I_r\|_F}{\|I_r\|_F} < 1 \times 10^{-5},$$

and the peak-signal-to-noise ratio (PSNR) is used to evaluate the quality of the restored images, defined by

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

and

$$MSE = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}[I_1(i,j) - X^{opt}(i,j)]^2,$$

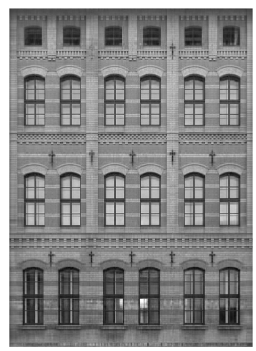
where m and n are the number of rows and columns of the image and I_r is the real image, and X^{opt} is the recovered image.

We carefully selected the optimal parameters for our experiments, taking into account the matrix rank size r and sampling rate p. The following are the results of our experiments, where we reconstructed images with 30, 10, and 5 ranks at varying sampling rates ($p = 0.8, 0.5, 0.3$). All results represent the average of five runs.

TABLE 3. Numerical results of different sampling rates p and the rank r of the image in terms of CPU time, number of iterations (Iter), and PSNR values.

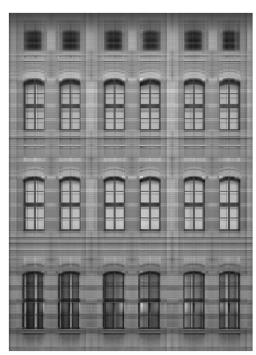
p	r	$\alpha = 2$	$\alpha = 1.6$
		Time/Iter/PSNR	Time/Iter/PSNR
0.8	30	14.7/866/105.86	8.9/519/105.87
	10	3.0/378/106.02	1.8/226/106.07
	5	2.4/288/106.08	1.4/172/106.27
0.5	30	101.4/4579/105.85	47.8/2754/105.85
	10	8.3/1056/105.88	5.2/632/105.89
	5	4.5/668/105.90	2.7/400/106.00
0.3	30	890.5/42743/105.85	546.6/25652/105.85
	10	25.4/2904/105.88	14.8/1740/105.88
	5	9.9/1338/105.91	6.0/802/105.98

(a) original image I



(b) r=30

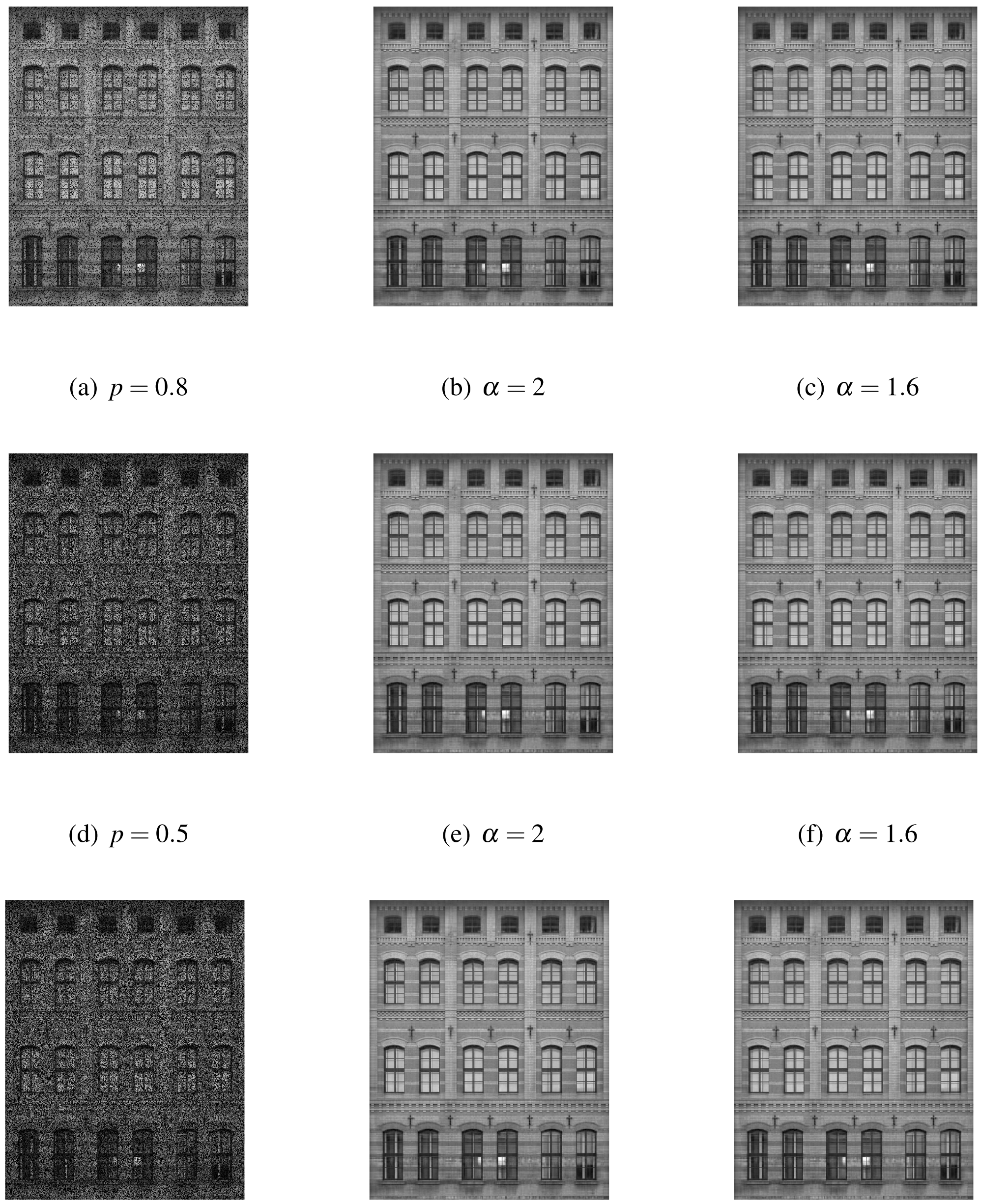
(c) r=10



(d) r=5

FIGURE 4.1. From left to right, these images are the original image and the images of different rank r obtained by the SVD on the original image with $r = 30$, $r = 10$, and $r = 5$, respectively.

Then, we use SVD to obtain an image with $r = 30$ and sample the image with different sampling rates: $p = 0.8$, $p = 0.5$ and $p = 0.3$. The images shown from left to right are the sampled image and the recovered results for $\alpha = 2$ and $\alpha = 1.6$, respectively.



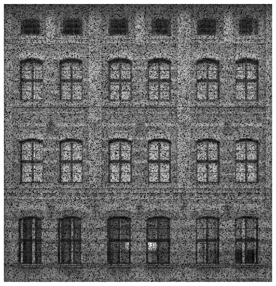
(a) $p = 0.8$ (b) $\alpha = 2$ (c) $\alpha = 1.6$

(d) $p = 0.5$ (e) $\alpha = 2$ (f) $\alpha = 1.6$

(g) $p = 0.3$ (h) $\alpha = 2$ (i) $\alpha = 1.6$

FIGURE 4.2. First column: Low-rank images with randomly sampling and a rank of 30; Second column: Restored images with $\alpha = 2$; Third column: Restored images with $\alpha = 1.6$.

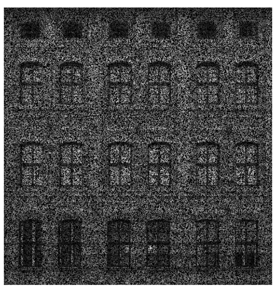
Next, we use SVD to obtain an image with $r = 10$ and sample the image with different sampling rates: $p = 0.8$, $p = 0.5$ and $p = 0.3$. The images shown from left to right are the sampled image and the recovered results for $\alpha = 2$ and $\alpha = 1.6$, respectively.



(a) $p = 0.8$

(b) $\alpha = 2$

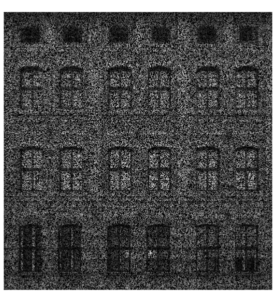
(c) $\alpha = 1.6$



(d) $p = 0.5$

(e) $\alpha = 2$

(f) $\alpha = 1.6$



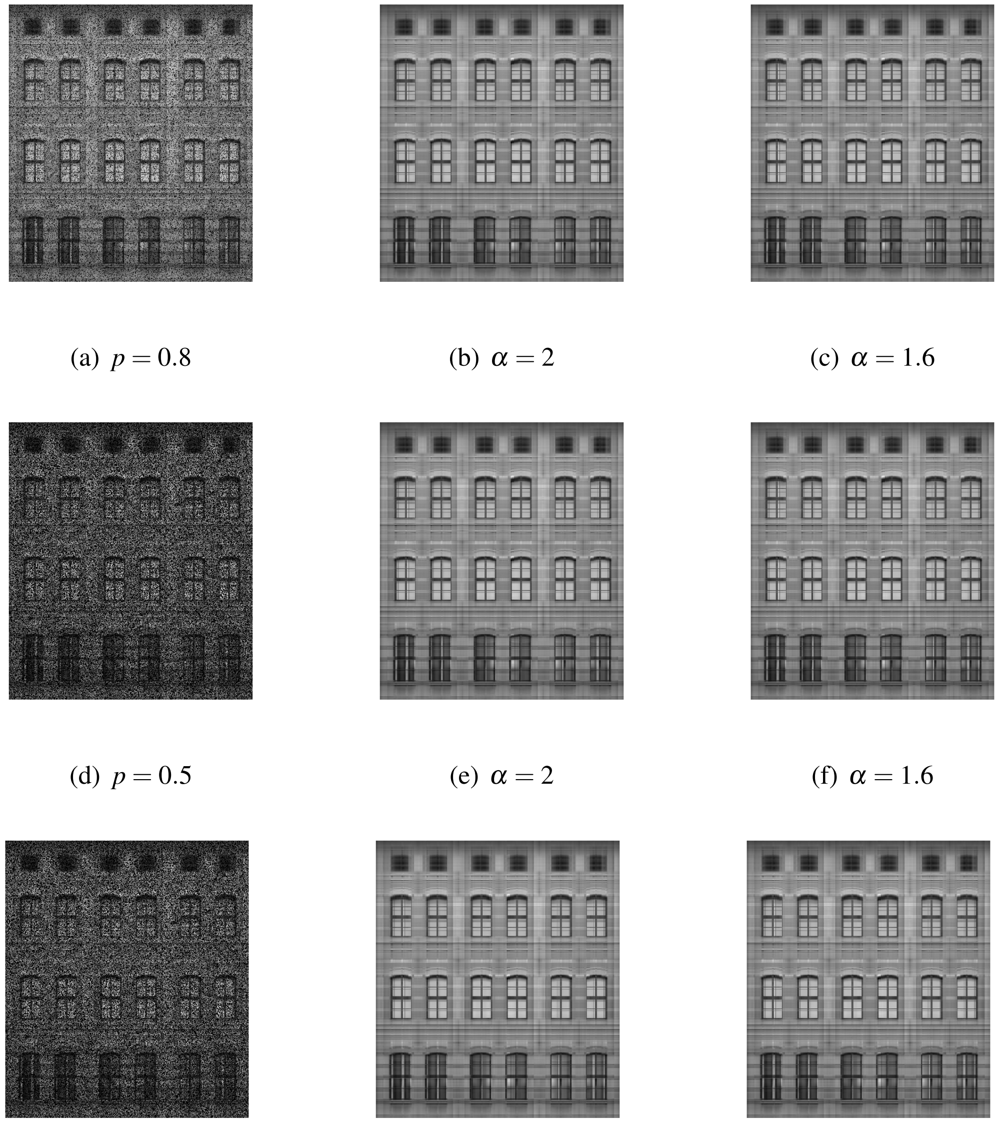
(g) $p = 0.3$

(h) $\alpha = 2$

(i) $\alpha = 1.6$

FIGURE 4.3. First column: Low-rank images with randomly sampling and a rank of 10; Second column: Restored images with $\alpha = 2$; Third column: Restored images with $\alpha = 1.6$.

Finally, we use SVD to obtain an image with $r = 5$ and sample the image with different sampling rates: $p = 0.8$, $p = 0.5$ and $p = 0.3$. The images shown from left to right are the recovery results for $\alpha = 2$ and $\alpha = 1.6$, respectively.



(a) $p = 0.8$ (b) $\alpha = 2$ (c) $\alpha = 1.6$

(d) $p = 0.5$ (e) $\alpha = 2$ (f) $\alpha = 1.6$

(g) $p = 0.3$ (h) $\alpha = 2$ (i) $\alpha = 1.6$

FIGURE 4.4. First column: Low-rank images with randomly sampling and a rank of 5; Second column: Restored images with $\alpha = 2$; Third column: Restored images with $\alpha = 1.6$.

Our experimental results demonstrate that our proposed algorithm performs better than the three-operator splitting algorithm in recovering low-rank images. Specifically, our algorithm requires approximately 40% less time and iteration numbers compared to the three-operator splitting algorithm. Furthermore, when both algorithms meet the same stopping criteria, our algorithm achieves slightly higher PSNR values than the three-operator splitting algorithm. These findings suggest that our algorithm is not only more efficient and accurate in recovering low-rank images, but it also has better performance and practicality.

5. Conclusions

In this paper, we proposed a parameterized three-operator splitting algorithm to solve (1.1). We proved that the sequences generated by the proposed algorithm converge to a critical point of the objective function, assuming that the objective function satisfies the Kurdyka-Łojasiewicz property. As applications, we employed the proposed algorithm to solve the classical problems of low-rank matrix recovery and image inpainting. Moreover, we compared our results with those of the three-operator splitting algorithm and showed that the proposed algorithm exhibits better performance and efficiency.

Acknowledgments

This work was supported by the National Natural Science Foundations of China (12061045 and 62101418), Research project at Guangzhou University (RC2023061), Guangzhou Education Scientific Research Project (202315829) and the Jiangxi Provincial Natural Science Foundation (20224ACB211004).

References

[1] S. Huang, T. Tran, Sparse signal recovery via generalized entropy functions minimization, IEEE Trans. Signal Process. 67 (2017), 1322–1337.

[2] E. Esser, Y. F. Lou, J. Xin, A method for finding structured sparse solutions to nonnegative least squares problems with applications, SIAM J. Imag. Sci. 6 (2013), 2010–2046

[3] P. Yin, Y. Lou, Q. He, J. Xin, Minimization of l_{1-2} for compressed sensing, SIAM J. Sci. Comput. 37 (2015), A536–A563.

[4] Y. Lou, T. Zeng, S. Osher, J. Xin, A weighted difference of anisotropic and isotropic total variation model for image processing, SIAM J. Imag. Sci. 8 (2015), 1798–1823.

[5] Y. F. Lou, P. H. Yin, Q. He, J. Xin. Computing sparse representation in a highly coherent dictionary based on difference of l_1 and l_2, J. Sci. Comput. 64 (2015), 178–196.

[6] H. Zhang, L. Z. Cheng, W. Zhu, A lower bound guaranteeing exact matrix completion via singular value thresholding algorithm, Appl. Comput. Harmon. Anal. 31 (2011), 454–459.

[7] T. Huang, M. Li, X. Qin, W. Zhu, A CNN-based policy for optimizing continuous action control by learning state sequences, Neurocomputing, 468 (2022), 286-295.

[8] R. Cabral, F. De la Torre, J. Costeira, A. Bernardino, Unifying nuclear norm and bilinear factorization approaches for low-rank matrix decomposition, In: IEEE Int. Conf. Comput. Vis. pp. 2488–2495, IEEE, 2013.

[9] R. Meka, P. Jain, I. Dhillon, Guaranteed rank minimization via singular value projection, arXiv:0909.5457, 2009.

[10] J. F. Cai, E. Candès, Z. Shen, A singular value thresholding algorithm for matrix completion, SIAM J. Optim. 20 (2010), 1956–1982.

[11] P. L. Combettes, V. R. Wajs, Signal recovery by proximal forward-backward splitting, Multiscale Model. Simul. 4 (2005), 1168–1200.

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

[13] H. Attouch, J. Bolte, B. F. Svaiter,Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods, Math. Program. Ser. A 137 (2013), 91–129.

[14] H. Attouch, J. Bolte, P. Redont, A. Soubeyran. Proximal alternating minimization and projection methods for nonconvex problems. An approach based on the Kurdyka-Łojasiewicz inequality, Math. Oper. Res. 35 (2013), 438–457.

[15] R. I. Bot, E. R. Csetnek, An inertial tseng's type proximal algorithm for nonsmooth and nonconvex optimization problems, J. Optim. Theory Appl., 171 (2016), 600–616.

[16] R. I. Bot, E. R. Csetnek, S. C. László, An inertial forward–backward algorithm for the minimization of the sum of two nonconvex functions, EURO J. Comput. Optim. 4 (2016), 3–25.

[17] A. Themelis, L. Stella, P. Patrinos, Forward-backward envelope for the sum of two nonconvex functions: Further properties and nonmonotone linesearch algorithms, SIAM J. Optim. 28 (2018), 2274–2303.

[18] P. L. Combettes, J. C. Pesquet, A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery, IEEE J. Sel. Top. Signal Process. 1 (2007), 564–574.

[19] G. Li, T. K. Pong, Douglas-Rachford splitting for nonconvex optimization with application to nonconvex feasibility problems, Math. Program. 159 (2016), 371–401.

[20] G. Li, T. Liu, T. K. Pong. Peaceman-Rachford splitting for a class of nonconvex optimization problems, Comput. Optim. Appl. 68 (2017), 407–436.

[21] M. Li, Z. Wu, Convergence analysis of the generalized splitting methods for a class of nonconvex optimization problems, J. Optim. Theory Appl. 183 (2019), 535–565.

[22] A. Themelis, P. Patrinos. Douglas-Rachford splitting and admm for nonconvex optimization: Tight convergence results, SIAM J. Optim. 30 (2020), 149–181.

[23] A. Themelis, L. Stella, P. Patrinos, Douglas–Rachford splitting and ADMM for nonconvex optimization: accelerated and newton-type linesearch algorithms, Comput. Optim. Appl. 82 (2022), 395–440.

[24] F. Bian, X. Zhang, A parameterized Douglas-Rachford splitting algorithm for non-convex optimization, Appl. Math. Comput. 410 (2021), 126425.

[25] D. Wang, X. Wang, A parameterized Douglas-Rachford algorithm, Comput. Optim. Appl. 73 (2019), 839–869.

[26] F. Bian, X. Zhang, A three-operator splitting algorithm for nonconvex sparsity regularization, SIAM J. Sci. Comput. 43 (2021), A2809–A2839.

[27] D. Davis, W. T. Yin, A three-operator splitting scheme and its optimization applications, Set-Valued Var. Anal. 25 (2017), 829–858.

[28] C. Zhang, J. J. Chen. A parameterized three-operator splitting algorithm and its expansion, J. Nonlinear Var. Anal. 5 (2021), 211–226.

[29] H. H. Bauschke, P. L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer, Cham, 2017.

[30] A. Beck, First-order Methods in Optimization, MOS-SIAM Series on Optimization, SIAM, Philadelphia, 2017.

[31] H. Wang, Y. Cen, Z. He, Z. He, R. Zhao, F. Z. Zhang, Reweighted low-rank matrix analysis with structural smoothness for image denoising, IEEE Trans. Image Process. 27 (2018), 1777–1792.

[32] Y. Yang, J. Zheng, S. Chen, Local low-rank matrix recovery for hyperspectral image denoising with l_0 gradient constraint, Pattern Recognit. Lett. 135 (2020), 167–172.

[33] H. Wang, W. Li, L. Hu, C. Zhang, Q. He, Structural smoothness low-rank matrix recovery via outlier estimation for image denoising, Multimedia Systems 28 (2022), 241–255.

[34] M. Davenport, J. Romberg, An overview of low-rank matrix recovery from incomplete observations, IEEE J. Sel. Topics Signal Process. 10 (2016), 608–622.

[35] Y. Wen, K. Li, H. Chen, Accelerated matrix completion algorithm using continuation strategy and randomized svd, J. Comput. Appl. Math. 429 (2023), 115215.

[36] X. Jin, G. Xie, Q. Wang, G. Zhong, G. Geng, Nonconvex matrix completion with nesterov's acceleration, Big Data Anal. 3 (2018), 1–11.