Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 9578168, 14 pages
https://doi.org/10.1155/2020/9578168

Research Article

Guarantees of Fast Band Restricted Thresholding Algorithm for Low-Rank Matrix Recovery Problem

Fujun Zhao,[1] Jigen Peng ,[2] Kai Sun,[1] and Angang Cui[1]

[1]*School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, 710049, China*
[2]*School of Mathematics and Information Science, Guangzhou University, Guangzhou 510006, China*

Correspondence should be addressed to Jigen Peng; jgpengxjtu@126.com

Received 1 March 2020; Accepted 7 April 2020; Published 30 April 2020

Guest Editor: Junpeng Shi

Affine matrix rank minimization problem is a famous problem with a wide range of application backgrounds. This problem is a combinatorial problem and deemed to be NP-hard. In this paper, we propose a family of fast band restricted thresholding (FBRT) algorithms for low rank matrix recovery from a small number of linear measurements. Characterized via restricted isometry constant, we elaborate the theoretical guarantees in both noise-free and noisy cases. Two thresholding operators are discussed and numerical demonstrations show that FBRT algorithms have better performances than some state-of-the-art methods. Particularly, the running time of FBRT algorithms is much faster than the commonly singular value thresholding algorithms.

1. Introduction

The affine matrix rank minimization (AMRM) problem, which is to recover a low-rank matrix from only a small number of linear measurements, can be described as the following optimization problem:

$$\begin{aligned} \min_{X\in\mathbb{C}^{n_1\times n_2}} \quad & \operatorname{rank}(X) \\ \text{s.t.} \quad & \|\mathscr{A}(X)-b\|_2 \le \eta, \end{aligned} \tag{1}$$

where $\mathscr{A} \in \mathscr{L}(\mathbb{C}^{n_1\times n_2}, \mathbb{C}^d)$ is a given linear map, $b \in \mathbb{C}^d$ is a given vector, and $\eta \ge 0$ is error tolerance. This problem has got much attention in recent years and many applications arising in various areas can be captured by solving model (1), for example, matrix completion [1, 2], background modeling [3], subspace clustering [4], phase retrieval [5, 6], image inpainting [7], and other applications [8, 9]. The lines and the symbols in all figures are too small to see. It is better to enlarge them. Unfortunately, ARMR problem is known to be NP-hard. Therefore, without further assumptions on $\mathscr{A}$ and X, solving this problem would be computationally intractable. To overcome this shortcoming, many researchers have focused on replacing $\operatorname{rank}(X)$ with other penalties $G(X)$, such as Schatten p-norm $(0<p<1)$ [10], Minimax Concave Plus (MCP) [11], Smoothly Clipped Absolute Deviation (SCAD) [12], Logarithm [13], Geman [14], and Laplace [15], and considered the following optimization problem:

$$\begin{aligned} \min_{X\in\mathbb{C}^{n_1\times n_2}} \quad & G(X) \\ \text{s.t.} \quad & \|\mathscr{A}(X)-b\|_2 \le \eta, \end{aligned} \tag{2}$$

or the corresponding unconstrained optimization problem:

$$\min_{X\in\mathbb{C}^{n_1\times n_2}} \frac{1}{2}\|\mathscr{A}(X)-b\|_2^2 + \lambda\cdot G(X), \tag{3}$$

where $\lambda > 0$ is a regularization parameter. Model (3) can be transformed into a fixed point problem described as

$$X = \mathscr{H}\left(X - s\cdot\mathscr{A}^*(AX-b)\right), \tag{4}$$

where $\mathscr{A}^*$ is the dual operator of $\mathscr{A}$, $\mathscr{H}$ is the thresholding operator, and s is a step size parameter. Naturally, we have corresponding iterative thresholding algorithm. Given a fixed penalty $G(\cdot)$, there have been many theoretical guarantees. However, there are still some challenges as follows:

(I) Most of convergence results were developed for model (3) with fixed λ, and it is difficult to choose an appropriate parameter λ

(II) It is hard to design the thresholding operator $\mathscr{H}(\cdot)$ for penalties $G(X)$ except some special penalties, such as Schatten p-norm $(p = 0, 1, 1/2, 2/3)$ [2, 16–18] and SCAD [12]

(III) Every iteration of the iterative thresholding algorithms needs to calculate the SVD, and if the size of the original matrix is large, the iterative thresholding algorithms will be time consuming

Two papers that relate to our work need to be overviewed here. Chartrand [19] firstly studied the way to compute the penalty $G(\cdot)$ via a given thresholding operator. However, his work did not show the stability guarantees of thresholding algorithms. Wei et al. [20] proposed an iterative hard algorithm with subspace projections, which had relatively lower computation complexity. However, his work did not examine the performance in the noisy case. Taking into account these work, we try to design a new algorithm and overcome these shortcomings to a certain extent.

In this paper, we design a family of new fast band restricted thresholding (FBRT) algorithms to solve model (1) and elaborate the theoretical guarantees in both noise-free and noisy cases. In fact, many related works [17, 18, 21, 22] focus on the stability of the optimal solution of the model. For the effect of noise on the algorithm, an empirical explanation is often given through numerical experiments. At the end of this paper, we give some simple examples to roughly verify our theoretical results.

The remainder of this paper is organized in the following. In Section 2, we introduce some basic preliminary results and notations. In Section 3, we describe the fast band restricted thresholding (FBRT) algorithm. In Section 4, we prove the convergence of the FBRT algorithm. In Section 5, we display the numerical simulations, and then conclude this paper in Section 6. All the proofs are presented in Appendix.

2. Preliminaries

In this section, we introduce some notations and useful results to facilitate the following research.

2.1. Notations. We denote $\mathscr{A}$ as a linear operator from $\mathbb{C}^{n_1\times n_2}$ to $\mathbb{C}^d$. For any $X \in \mathbb{C}^{n_1\times n_2}$, $\mathscr{A}(X)$ is $(\langle X, A_i\rangle_F)_{i=1}^d$, where $\langle X, A_i\rangle_F = \mathrm{tr}(XA_i^*)$ is a Frobenius inner product. Thus, we naturally obtain the dual operator $\mathscr{A}^*$ as follows:

$$\begin{aligned}\mathscr{A}^*:\quad \mathbb{C}^d &\longrightarrow \mathbb{C}^{n_1\times n_2}\\ b &\longmapsto \sum_{i=1}^d b_i A_i.\end{aligned} \tag{5}$$

We also denote $\overline{X}$ as the original matrix with $\mathrm{rank}(\overline{X}) = r$ and $e = b - \mathscr{A}(\overline{X})$ as the noise with error tolerance $\eta = \|e\|_2$. For any subspaces $\Omega_1, \Omega_2 \subset \mathbb{C}^{n_1\times n_2}$, we obtain the sum of subspaces $\Omega_1 + \Omega_2$, and we also rewrite $\Omega_1 \cap (\Omega_2)^\perp$ as $\Omega_1 \backslash \Omega_2$.

2.2. Some Useful Results. In this section, we introduce some useful results, which show the relationship between the original low-rank matrix and the solution of model (1). For this purpose, we review the definition of RIC, which plays an important role in the following research.

Definition 1 (see [23]). Given a linear operator $\mathscr{A}: \mathbb{C}^{n_1\times n_2} \longrightarrow \mathbb{C}^d$ and $1 \le r \le \min\{m, n\}$, the restricted isometry constant δ_r is the smallest nonnegative number, such that, for any matrix $X \in \mathbb{C}^{n_1\times n_2}$, with $\mathrm{rank}(X) \le r$, we have the following inequality:

$$(1 - \delta_r)\|X\|_F^2 \le \|\mathscr{A}(X)\|_2^2 \le (1 + \delta_r)\|X\|_F^2, \tag{6}$$

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

For a given linear map $\mathscr{A}$, it is difficult to calculate the restricted isometry constant δ_r, but for Gaussian random linear map, there has been a result that the number of measurements $d \ge c_0 r(n_1 + n_2)\log(n_1 n_2)$ (c_0 only depends on δ) is sufficient to yield an RIC of δ with high probability [23].

In the noisy case, the following theorem claims that the solution of model (1) cannot be far of from the original low-rank matrix, if $\mathscr{A}$ satisfies a certain RIC. The related theorem about compressive sensing is mentioned in [24], see equation (5.24).

Theorem 1. *Let Z, Z^η be two solution of (1) within error tolerance η (i.e., $\|\mathscr{A}(Z) - b\|_2, \|\mathscr{A}(Z^\eta) - b\|_2 \le \eta$). If $\mathrm{rank}(Z) = \mathrm{rank}(Z^\eta) = r$ and $\delta_{2r} < 1$, we can obtain a stability claim of the form*

$$\left\|Z - Z^\eta\right\|_2^2 < \frac{4\eta^2}{1 - \delta_{2r}}. \tag{7}$$

If we take $\eta = 0$ in Theorem 1, we obtain the following uniqueness result, in a noise-free case.

Corollary 1. *Let Z be a r-rank solution of $\mathscr{A}(X) = b$. If $\mathscr{A}$ satisfies*

$$\delta_{2r} < 1, \tag{8}$$

then Z is the unique lowest rank solution.

Taking into account the abovementioned results, it implies that the original low-rank matrix $\overline{X}$ is the solution of model (1), if $\mathscr{A}$ satisfies a certain RIC.

3. Fast Band Restricted Thresholding Algorithm

In this section, we will design a new fast band restricted thresholding (FBRT) algorithm. For this purpose, we need to review some thresholding algorithms for (1) in previous work and compare FBRT algorithm with other algorithms.

Similar to compressive sensing problem, both iterative hard thresholding and iterative soft thresholding, also known as SVP [16] and SVT [2], are simple and efficient algorithms for low-rank matrix recovery. In pursuit of better results, some alternative algorithms, such as HFPA [22], thresholding function for Schatten 2/3-norm [18], SCAD [12], and firm thresholding [25], have been proposed. We present these commonly used thresholding functions in Figure 1, and it is obvious that all of them satisfy the following definition.

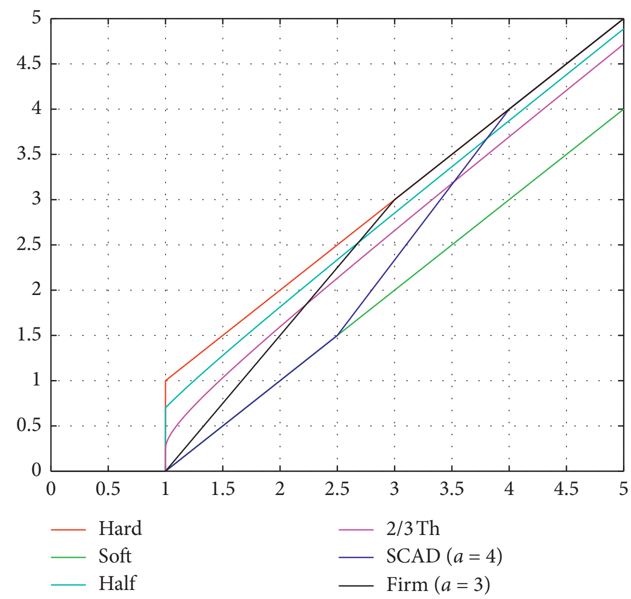


FIGURE 1: Typical thresholding function $h_\tau(\cdot)$ with $\tau = 1$.

Definition 2. A function h_τ of $\mathbb{R}^+$ with parameter $\tau > 0$ is called as band restricted thresholding (BRT) function, if it satisfies

(I) $h_\tau(u) = 0$, with $u \le \tau$

(II) $h_\tau(u) \le h_\tau(v)$, with $\tau < u \le v$

(III) There exists a constant $c \in [0,1]$ (band parameter), such that $u - c\tau \le h_\tau(u) \le u$, with $u \ge \tau$

According to Definition 2, the following Table 1 shows the corresponding band parameter c of these BRT functions.

Given a BRT function h_τ, we can obtain a thresholding operator $\mathscr{H}_\tau(\cdot)$ which is defined as

$$\mathscr{H}_\tau(X) = \sum_{i=1} h_\tau(\sigma_i) u_i v_i^*, \tag{9}$$

where $\sum_{i=1}\sigma_i u_i v_i^*$ is the SVD of X, and we can also have the corresponding thresholding algorithm described as

$$\begin{aligned} W^{(k)} &= X^{(k-1)} + \mathscr{A}^*\left(b - \mathscr{A}\left(X^{(k-1)}\right)\right), \\ X^{(k)} &= \mathscr{H}_\tau\left(W^{(k)}\right), \end{aligned} \tag{10}$$

where τ is the threshold parameter. In fact, every iteration of the abovementioned algorithm needs to compute the SVD of $W^{(k)}$. Thus, if $N = (n_1 + n_2)$ is large, this algorithm will be computationally expensive. To improve the efficiency of algorithm, we propose to orthogonally project $W^{(k)}$ onto a subspace S_k, which was first applied in [20]. S_k, which is the tangent space of the rank r matrix manifold at $X^{(k)}$, is described as

$$S_k := \{U_k P + Q V_k^* : P \in \mathbb{C}^{r\times n_2}, Q \in \mathbb{C}^{n_1 \times r}\}, \tag{11}$$

where $U_k \in \mathbb{C}^{n_1\times r}$ and $V_k \in \mathbb{C}^{r\times n_2}$ are the left and right singular vectors of $X^{(k)}$. Furthermore, for any matrix $Z \in \mathbb{C}^{n_1\times n_2}$, $\mathscr{P}_{S_k}(Z)$ is defined as

TABLE 1: Band parameter c for different thresholding functions.

$h_{\tau,*}$	$h_{\tau,0}$	$h_{\tau,1}$	$h_{\tau,1/2}$	$h_{\tau,2/3}$	$h_{\tau,\mathrm{SCAD}}$	$h_{\tau,\mathrm{firm}}$
c	0	1	1/3	3/4	1	1

$$\begin{aligned} \mathscr{P}_{S_k}(Z) &= U_k U_k^* Z + Z V_k V_k^* - U_k U_k^* Z V_k V_k^* \\ &= (I - U_k U_k^*) Z (I - V_k V_k^*). \end{aligned} \tag{12}$$

Meanwhile, we derive the fast band restricted thresholding (FBRT) algorithm described as in Algorithm 1.

Remark 1. The threshold parameter τ plays an important role which affects the performance of FBRT algorithm. Here, we let $\tau = \sigma_{r+1}^k$, and we will show the theoretical performance of FBRT algorithm in the following discussion.

Remark 2. The stopping criterion is $\|\mathscr{A}(X^{(\bar{k})}) - b\|_2 \le \varepsilon$. Taking into account Theorem 1, $X^{(\bar{k})}$ cannot be far of from $\overline{X}$, if $\mathscr{A}$ satisfies a certain RIC. On the contrary, it is an important situation of classified discussion that $X^{(n)}$ is close to $\overline{X}$, and the detail will be shown in the following section.

It is worth noting that $X^{(k)} + \mathscr{A}^*(b - \mathscr{A}(X^{(k)}))$ is always a full rank matrix and computing the SVD will use $O(N^3)$ $(N = (m+n))$ floating point operations flops. In the mean time, $W^{(k)}$ in every iteration of FBRT algorithm is a $(2r)$-rank matrix, and according to the QR factorization, we can obtain that

$$W^{(k)} = QR = QR_1^* Q_1^*, \tag{13}$$

where $W^{(k)} = QR$ and $R^* = Q_1 R_1$ are the QR factorizations of $W^{(k)}$ and R^*, with $Q \in \mathbb{C}^{n_1\times 2r}$, $R \in \mathbb{C}^{2r\times n_2}$, $R_1 \in \mathbb{C}^{2r\times 2r}$, and $Q_1 \in \mathbb{C}^{n_2\times 2r}$. Thus, the SVD of $W^{(k)}$ can be obtained from the SVD of R_1, and computing the SVD of $W^{(k)}$, will use $O(r^3)$ floating point operations instead of $O(N^3)$ flops [20].

4. Analysis of the FBRT Algorithm

In this section, we will study theoretically the performance of the FBRT algorithm, and the following Theorem 2 and 3 show the theoretical guarantees in both noise-free and noisy cases.

Theorem 2. *Let the sequence $\{X^{(n)}\}$ be defined by the FBRT algorithm with $\mathscr{A}(X) = \mathscr{A}(\overline{X}) = b$ and $\mathrm{rank}(\overline{X}) = r$. If the thresholding operator in the FBRT algorithm is the BRT operator with a band parameter c and the following constant*

$$\rho := \sqrt{4 + 2c^2 r}\left(\delta_{2r} + \delta_{3r} + \frac{\sqrt{4 + 2c^2 r}\,\delta_{2r+1}}{\sigma_{\min}\overline{X}}\|\overline{X}\|_F\right), \tag{14}$$

is less than 1, we have

$$\left\|X^{(n)} - \overline{X}\right\|_F \le \rho^n \left\|X^{(0)} - \overline{X}\right\|_F. \tag{15}$$

Particularly, $\rho < 1$ can be satisfied if

$$\delta_{3r} \le \frac{1}{2\sqrt{4 + 2c^2 r} + (4 + 2c^2 r)\sqrt{r}\,\sigma_{\max}(\overline{X})/\sigma_{\min}(\overline{X})}. \tag{16}$$

Input: $\mathscr{A} \in \mathscr{L}(\mathbb{C}^{n_1 \times n_2}, \mathbb{C}^d)$, $r \in \mathbb{N}^+$, $b \in \mathbb{C}^d$, and error tolerance $\varepsilon \geq 0$
Initialization: $k = 0, X^{(0)} = \mathscr{H}_{\sigma^0_{r+1}}(W^{(0)})$, where $W^{(0)} = \mathscr{A}^*(b)$ and σ^k_{r+1} is the $(r+1)$-th singular value of $W^{(0)}$.
Repeat:
(1) $k \longleftarrow k+1$,
(2) $W^{(k)} = X^{(k-1)} + \mathscr{P}_{S_{k-1}} \mathscr{A}^*(b - \mathscr{A}(X^{(k-1)}))$,
(3) $W^{(k)} = U^{(k)}\Sigma^{(k)}(V^{(k)})^*$, based on the singular value decomposition,
(4) $X^{(k)} = U^{(k)} H_{\sigma^k_{r+1}}(\Sigma^{(k)})(V^{(k)})^*$, where σ^k_{r+1} is the $(r+1)$-th singular value of the matrix $W^{(k)}$.
Until: the matrix $X^{(k)}$, with $\|\mathscr{A}(X^{(k)}) - b\|_2 \leq \varepsilon$.

Algorithm 1: Fast band restricted thresholding (FBRT).

The proof of Theorem 2 is given in Appendix, and we discuss condition of Theorem 2, which shows the performance of the FBRT algorithm in the AMRM problem.

Remark 3. According to Theorem 2, the performance of the FBRT algorithm depends on $\sigma_{\max}(\overline{X})/\sigma_{\min}(\overline{X})$, i.e., the condition number of the original r-rank matrix $\overline{X}$. It plays an important role for the projection operator $\mathscr{P}_{S_n}$.

In the noisy case, the error tolerance η is a significant parameter. According to Theorem 1, we know that there exists a gap between the solution of model (1) and the original low-rank matrix. Thus, we always assume that the model error is small.

Theorem 3. *Assume that $\|\overline{X}\|_F \geq \mu\eta$, where rank$(\overline{X}) = r$ and $\eta = \|e\|_2$ is error tolerance. Let the sequence $\{X^{(n)}\}$ be defined by the FBRT algorithm with $\|\mathscr{A}(X) - b\|_2 \leq \eta$. If the thresholding operator in the FBRT algorithm is the BRT operator with a band parameter c and $\mathscr{A}$ satisfies*

$$\delta_{3r} \leq \frac{\mu}{2\sqrt{4+2c^2 r}(\mu+1)\left(1+\left(\sqrt{4+2c^2 r}/\sigma_{\min}(\overline{X})\right)\left(\|\overline{X}\|_F - \mu\eta\right)\right)}, \tag{17}$$

then the sequence $\{X^{(n)}\}$ must satisfy one of the following result:

(i) There exists a positive integer n, such that

$$\left\|X^{(n)} - \overline{X}\right\|_F \leq c'\eta, \tag{18}$$

where $c' = 4(\mu+1)\sqrt{2+c^2 r}$.

(ii) For any positive integer n, we have

$$\left\|X^{(n)} - \overline{X}\right\|_F \leq \rho^n \left\|X^{(0)} - \overline{X}\right\|_F + \frac{\sqrt{4+2c^2 r}\sqrt{1+\delta_{2r+1}}}{1-\rho}\eta, \tag{19}$$

where $\rho = (\mu/(\mu+1)) < 1$.

The proof of this theorem is also presented in Appendix. First of all, if δ_{3r} satisfies the condition of Theorem 3, $\delta_{3r} \leq \mu/(\mu+1) < 1$. Besides, when the error tolerance η is not too large, the result of Theorem 3 illustrates that there exists a positive integer n, such that $X^{(n)}$ will be close to the original matrix $\overline{X}$. Since $\mathscr{A}$ is a bounded linear operator, it implies that $\mathscr{A}(X^{(n)})$ is close to $\mathscr{A}(\overline{X})$. In the mean time, the fact that $\mathscr{A}(X^{(n)})$ is close to $\mathscr{A}(\overline{X})$ also implies that $X^{(n)}$ is close to the original matrix $\overline{X}$, according to Theorem 1. On the contrary, the parameter μ in this theorem is a key parameter, and performance of algorithm will get better as $\|\overline{X}\|_F - \mu\eta$ is close to 0. Meanwhile, the performance of algorithm also depends on the condition number of the original r-rank matrix $\overline{X}$, which is similar to Theorem 2.

5. Numerical Demonstration

In this section, we present some empirical observations of FBRT algorithms with two thresholding operator:

$$h_{\tau,\mathrm{atan}}(u) = \begin{cases} u - \tau + \dfrac{2\tau}{\pi}\arctan(u-\tau), & \text{if } u \geq \tau, \\ 0, & \text{if } u < \tau, \quad \text{(Atan)}, \end{cases}$$

$$h_{\tau,\mathrm{firm}}(u) = \begin{cases} u, & \text{if } u > a\cdot\tau, \\ \dfrac{a}{a-1}(x-u), & \text{if } \tau \leq u \leq a\cdot\tau, \quad \text{(firm)}, \\ 0, & \text{if } u < \tau, \end{cases} \tag{20}$$

and compare them with some state-of-the-art methods (singular value thresholding (SVT) algorithm [2], singular value pursuit (SVP) algorithm [16], half norm fixed point (Half) algorithm [22], and Riemannian gradient descent (RGrad) algorithm [20]).

Here, we define two quantities to quantify performance of the algorithm:SR $:= d/n_1 n_2$, i.e., the number of measurements divided by the number of entries of the matrix, which denotes the sampling ratio, andOS $:= d/(r(n_1 + n_2 - r))$ is the oversampling ratio, i.e., the ratio between the number of sampled entries and the "true dimensionality" of an $n_1 \times n_2$ matrix of rank r. In fact, if OS < 1, we cannot recover the original low-rank matrix because there are always an infinite number of matrices of rank r with the given entries [26]. We use the relative error (RE) in the Frobenius norm

$$\mathrm{RE} := \frac{\left\|M - X_{\mathrm{opt}}\right\|_F}{\|M\|_F}, \tag{21}$$

to evaluate the closeness of X_{opt} to M, where X_{opt} is the "optimal" solution of the algorithm and M is the original low-rank matrix.

Table 2: Numerical results of two FBRT algorithms: FBRT-Atan and FBRT-firm ($a = 3$) with fixed $m = n = 300$. For any SR, the maximum of 10 relative errors is less than 0.01 when $r \le r_{\min}$ and the minimum of 10 relative errors is greater than or equal to 0.01 when $r > r_{\max}$.

z SR		FBRT-Atan Min RE	FBRT-Atan Max RE		FBRT-firm Min RE	FBRT-firm Max RE
0.1	$r_{\min} = 11$	$1.0260e-5$	0.0023	$r_{\min} = 11$	$2.1594e-5$	0.0068
	$r_{\max} = 13$	0.0079	0.6592	$r_{\max} = 12$	$2.7247e-4$	0.5478
0.2	$r_{\min} = 26$	$6.1794e-6$	$6.5852e-5$	$r_{\min} = 27$	$8.8405e-6$	$2.2534e-4$
	$r_{\max} = 28$	$3.3154e-4$	0.5499	$r_{\max} = 28$	$3.6679e-4$	0.5143
0.3	$r_{\min} = 43$	$7.6326e-6$	$9.6796e-6$	$r_{\min} = 44$	$8.2757e-5$	$3.0298e-4$
	$r_{\max} = 45$	0.0017	0.5188	$r_{\max} = 45$	0.0032	0.4816
0.4	$r_{\min} = 61$	$7.7705e-6$	$1.8222e-5$	$r_{\min} = 61$	$7.8113e-6$	$8.3863e-6$
	$r_{\max} = 62$	$6.8547e-5$	0.4236	$r_{\max} = 63$	0.0011	0.4392
0.5	$r_{\min} = 80$	$7.0491e-6$	$7.3872e-6$	$r_{\min} = 81$	$1.1833e-5$	$2.2862e-5$
	$r_{\max} = 81$	$1.1140e-5$	0.3924	$r_{\max} = 83$	0.0059	0.3871
0.6	$r_{\min} = 102$	$7.4470e-6$	$2.3292e-5$	$r_{\min} = 103$	$1.9080e-5$	$6.4072e-5$
	$r_{\max} = 103$	$2.2913e-5$	0.3510	$r_{\max} = 105$	0.0016	0.3472
0.7	$r_{\min} = 127$	$7.7453e-6$	$5.2701e-5$	$r_{\min} = 128$	$1.8729e-5$	$1.3984e-4$
	$r_{\max} = 128$	$5.1908e-5$	0.2978	$r_{\max} = 129$	$1.5370e-4$	0.2578
0.8	$r_{\min} = 156$	$6.7983e-6$	$7.4203e-6$	$r_{\min} = 158$	$5.5031e-5$	$2.1316e-4$
	$r_{\max} = 158$	$1.3876e-4$	0.2559	$r_{\max} = 159$	$3.2228e-4$	0.2214
0.9	$r_{\min} = 196$	$9.9473e-6$	$4.7502e-4$	$r_{\min} = 197$	$4.6863e-5$	$1.7010e-4$
	$r_{\max} = 197$	$2.5510e-4$	0.1592	$r_{\max} = 198$	0.0011	0.1610

5.1. Empirical Phase Transition. In this section, we test how many measurements are needed to recover a low rank matrix. For the sake of simplicity, we set $n = m$ and generate rank r matrix $X = LR$, where $L \in \mathbb{R}^{n\times r}$ and $R \in \mathbb{R}^{r\times n}$ and the components of L and R is sampled from the standard normal distribution. The stopping criterion is as follows:

$$\frac{\|X^k - X^{k-1}\|_F}{\|X^{k-1}\|_F} \le 10^{-8}. \tag{22}$$

Simulations of FBRT algorithms are repeated for 10 times, and numerical results are reported in Table 2. We consider an algorithm to successfully recover the low rank matrix X if the "optimal" solution of this algorithm with $\mathrm{RE} \le 0.01$. Furthermore, we denote $r_{\min}$ as the largest rank such that the corresponding max $\mathrm{RE} \le 0.01$ and $r_{\max}$ as the smallest rank such that the corresponding min $\mathrm{RE} \ge 0.01$. $\mathrm{OS}_{\max}$ is computed via $r_{\min}$, and $\mathrm{OS}_{\min}$ is computed via $r_{\max}$. Figure 2 shows the empirical of the tested algorithms on the (SR and OS), where we calculate oversampling ratio OS via $r = ((r_{\min} + r_{\max})/2)$.

Table 2 and Figure 2 show that these two FBRT algorithms can affect recover rank r matrices, where OS is close to 1, and there is not much difference between two empirical phase transition curves.

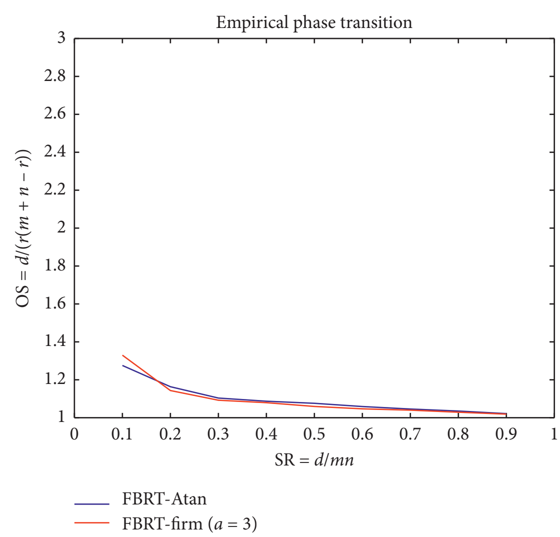


Figure 2: Empirical phase transition curves for FBRT algorithms: FBRT-Atan and FBRT-firm ($a = 3$) with fixed $m = n = 300$

5.2. Comparison with the State-of-the-Art Algorithms in Noise-Free Case. In this section, we consider the noise-free case and test the performances of FBRT algorithms on matrix completion problems. We also generate a low rank matrix in the way of Section 5.1. The stopping criterion is as follows:

$$\frac{\|X^k - X^{k-1}\|_F}{\|X^{k-1}\|_F} \le 10^{-8}. \tag{23}$$

Simulations are repeated for 10 times, and numerical results are reported in Tables 3 and 4 and Figure 3.

In Table 3, we show the comparison experiments with SVT, SVP, and Half. We conduct the tests for $n \in \{400, 1000\}$, $r = 50$, and $\mathrm{OS} = 2$ and record the average, maximum, and minimum values of relative errors and running time, respectively. Based on Table 3, we can obverse that the relative errors of FBRT algorithms are smaller, and the running time is also faster. In the meantime, the running time of FBRT algorithms increases slowly as n increases. We explore the trends, displayed in Figure 3, between the relative error and the running time of the first time algorithm operation results. In Table 4, we show the comparison experiments with RGrad. We conduct the tests in two cases: $n = 100, r = 19, \mathrm{SR} = 0.4$ and $n = 1000, r = 50, \mathrm{OS} = 2$. Based on Table 4, we see that the running time of different

Table 3: Numerical results of matrix recovery algorithms: SVT, SVP, Half, FBRT-Atan, and FBRT-firm ($a = 3$) with different n but fixed rank = 300 and OS = 2.

		SVT	SVP	Half	FBRT-Atan	FBRT-firm
$n = 400$ $r = 50$ OS = 2	taver	12.45	6.88	12.42	5.00	5.04
	REaver	0.20	$1.76e-7$	$1.61e-4$	$1.76e-7$	$1.77e-7$
	$t$max	13.47	7.07	13.32	5.15	5.27
	$t$min	11.61	6.72	7.21	4.88	4.94
	REmax	0.21	$1.92e-7$	$1.65e-4$	$1.85e-7$	$1.87e-7$
	REmin	0.19	$1.70e-7$	$1.57e-4$	$1.69e-7$	$1.68e-7$
$n = 1000$ $r = 50$ OS = 2	taver	131.52	172.59	420.00	92.75	93.00
	REaver	0.44	$5.37e-7$	$1.20e-4$	$5.54e-7$	$5.48e-7$
	$t$max	139.66	174.62	443.33	93.81	94.78
	$t$min	122.04	170.64	379.53	91.22	91.62
	REmax	0.45	$5.48e-7$	$1.21e-4$	$5.92e-7$	$5.73e-7$
	REmin	0.41	$5.33e-7$	$1.18e-4$	$5.32e-7$	$5.36e-7$

Table 4: Numerical results of matrix recovery algorithms: RGrad, FBRT-Atan, and FBRT-firm ($a = 3$) with different n, rank and OS = 2.l

		RGrad	FBRT-Atan	FBRT-firm
$n = 100$ $r = 19$ SR = 0.4	Aver. RE	0.0354	5.5880e-6	4.7998e-6
	Aver. Time	8.1457	9.3389	8.5035
	Max RE	0.3541	$8.2501e-6$	$6.1506e-6$
	Max time	13.1745	11.6156	10.0949
	Min RE	$4.1929e-6$	$4.3110e-6$	$3.9417e-6$
	Min time	0.5778	7.4130	6.9589
$n = 1000$ $r = 50$ OS = 2	Aver. RE	$5.6007e-7$	$5.4427e-7$	$5.4482e-7$
	Aver. Time	99.6237	95.4349	97.2309
	Max RE	$6.3984e-7$	$5.5862e-7$	$5.6776e-7$
	Max time	103.0736	98.7222	100.1280
	Min RE	$5.3833e-7$	$5.2797e-7$	$5.2966e-7$
	Min time	98.2743	93.0526	94.3025

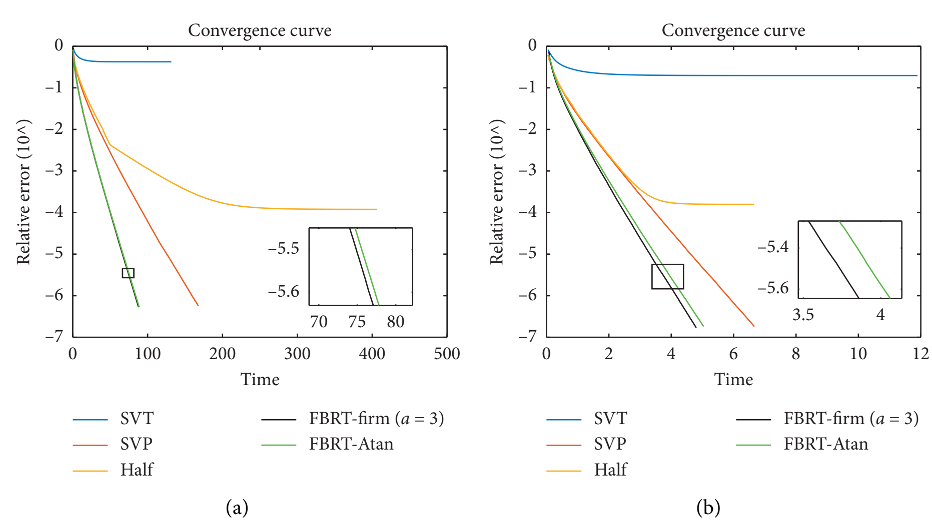
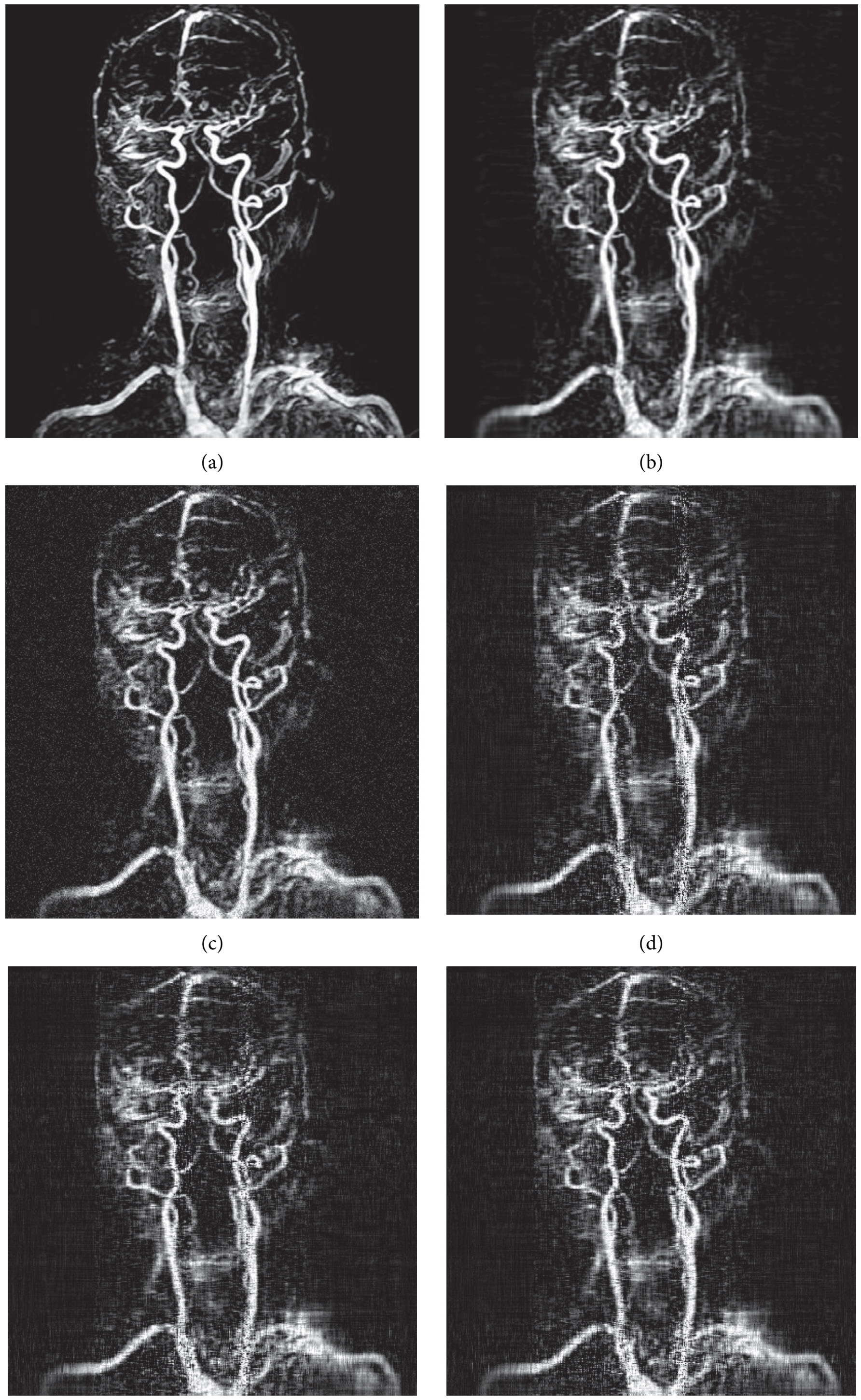


Figure 3: Convergence curves for matrix recovery algorithms: SVT, SVP, Half, FBRT-Atan and FBRT-firm ($a = 3$) with different $n = 1000$, 400 but fixed rank = 50 and OS = 2. (a) $n = 1000$. (b) $n = 400$.

algorithms is about the same, but the relative errors of FBRT algorithms are smaller than the relative error of RGrad. Therefore, we can find that FBRT algorithms perform better than others in noise-free case.

5.3. Comparison with the State-of-the-Art Algorithms in Noisy Case. In this section, we consider the noisy case and compare the performances of FBRT algorithms with RGard algorithm on the image inpainting problems. Here,



(a) (b) (c) (d) (e) (f)

Figure 4: (a) Original 419×400 IVI image with full rank; (b) approximated image of (a) with rank 35; c noised image of (b) with $\sigma^2 = 0.01$; (d, e, and f) the recovering images via RGrad, FBRT-Atan, and FBRT-firm, respectively, SR = 0.4 (noisy case). (a) Original IVI. (b) Low-rank IVI. (c) IVI with noise. (d) RGrad0.4. (e) FBRT-atan0.4. (f) FBRT-_rm0.4.

we test these algorithms on the grayscale image: 419×400 intracranial venous image (IVI), and we obtain approximated low-rank image truncated from IVI with rank $r = 35$. We generate the noised image with normal distribution by

$$\text{imnoise}\left(\text{image}, \text{“gaussian”}, 0, \sigma^2\right), \tag{24}$$

where σ^2 is the variance of normal distribution, and generate the sample image from noised image with SR = 0.4. We consider the variance $\sigma^2 \in \{0.01, 0.001\}$ and iterations are

Table 5: Numerical results of four algorithms on image inpainting (approximated low-rank IVI) with different ranks (noisy case).

	RE		
	RGrad	FBRT-Atan	FBRT-firm
$\sigma^2 = 0.01$	0.4577 (156)	0.4108 (114)	0.4017 (121)
$\sigma^2 = 0.001$	0.1262 (524)	0.1251 (514)	0.1250 (578)

150 and 550, respectively. The original image, its approximated low-rank image, its noised image, and its recovering images of different algorithms and displayed in Figure 4, respectively. Figure 4(a) is original IVI with full rank. Figure 4(b) is a approximated low-rank image truncated from Figure 4(a) with rank $r = 35$. Figure 4(c) is a noised image of Figure 4(b) with $\sigma^2 = 0.01$. Figures 4(d)–4(f) are the recovering image via different algorithms, respectively. Numerical results for image inpainting are reported in Table 5. Comparing these comparison experiments, we find that FBRT algorithms perform much better that RGard algorithm in image inpainting noisy case.

6. Conclusions

In this paper, we proposed the FBRT algorithm and developed the theoretical guarantees in both noise-free and noisy case. The numerical demonstration showed that this algorithm is effective. However, it is important to estimate error tolerance η, which has a great influence on FBRT algorithm. In the low SNR cases, more priori assumptions of the original low rank matrix are necessary, and we will study it in our future work. On the contrary, the condition number of the original matrix is also a significant parameter. As we know, phase retrieval [5, 6] and blind deconvolution [27, 28] problems can be transformed into a rank one matrix recovery problem. Since the condition number of a rank one matrix is always equal to one, it may be worth to study phase retrieval and blind deconvolution problems via the FBRT algorithm.

Appendix

Proof of Theorems 2 and 3

Now, let us turn to the proof of Theorems 2 and 3. Before Theorems 2 and 3, we need to denote some notations and introduce some lemmas.

We denote $\overline{I}$, I^n, and I_+^n as subspaces of $\mathbb{C}^{n_1\times n_2}$, which are described, respectively, as

$$\begin{aligned}\overline{I} &:= \left\{U\overline{V}^*: U \in \mathbb{C}^{n_1\times r}\right\},\\ I^n &:= \left\{U(V_n)^*: U \in \mathbb{C}^{n_1\times r}\right\},\\ I_+^n &:= \left\{U(V_{n+})^*: U \in \mathbb{C}^{n_1\times(r+1)}\right\},\end{aligned} \tag{A.1}$$

where $\overline{V} \in \mathbb{C}^{r\times n_2}$ is the right singular vectors of $\overline{X}$, $V_n \in \mathbb{C}^{r\times n_2}$ is the right singular vectors of $X^{(n)}$, and $V_{n+} \in \mathbb{C}^{(r+1)\times n_2}$ is the right singular vectors corresponding to the first $r+1$ singular values of $W^{(n)}$. According to FBRT algorithm, it holds that $I^n \subset I_+^n$. For any matrix $M \in \mathbb{C}^{n_1\times n_2}$ and matrix subspaces $I, J \subset \mathbb{C}^{n_1\times n_2}$, M_I represents $\mathscr{P}_I(M)$, where $\mathscr{P}_I(\cdot)$ is an orthogonal projection onto I, and we also denote $I + J$ as the sum of two subspaces and rewrite $I \cap (J)^\perp$ as $I\backslash J$.

Lemma A.1. *Let σ_{r+1}^n be the $(r+1)$th singular value of the matrix $W^{(n)}$. Then,*

$$\sigma_{r+1}^n \le \left\|\left(W^{(n)} - \overline{X}\right)_{I_+^n}\right\|_{2\longrightarrow 2} \le \left\|\left(W^{(n)} - \overline{X}\right)_{I_+^n}\right\|_F. \tag{A.2}$$

Proof. Because of the definitions of σ_{r+1}^n and I_+^n, it holds that

$$\sigma_{r+1}^n = \min_{\dim(M)\le n_2 - r}\ \max_{x\in M\subset\mathbb{C}^{n_2},\|x\|_2=1} \left\|W_{I_+^n}^{(n)}x\right\|_2, \tag{A.3}$$

where the second equality is based on the definition of I_+^n. In the mean time, we project $\overline{X}$ onto a subspace I_+^n and obtain $\overline{X}_{I_+^n}$ described as

$$\overline{X}_{I_+^n} = \overline{X}V_{n+}(V_{n+})^*. \tag{A.4}$$

Thus, $\operatorname{rank}(\overline{X}_{I_+^n}) \le \operatorname{rank}(\overline{X}) = r$, and we obtain a subspace $\ker(\overline{X}_{I_+^n})$ with $\dim(\ker(\overline{X}_{I_+^n})) \ge n_2 - r$. Furthermore, it holds that

$$\begin{aligned}\sigma_{r+1}^n &= \min_{\dim(M)\le n_2 - r}\ \max_{x\in M\subset\mathbb{C}^{n_2},\|x\|_2=1} \left\|W_{I_+^n}^{(n)}x\right\|_2\\ &\le \min_{\dim(M)\le n_2 - r}\ \max_{x\in M\subset\ker\left(\overline{X}_{I_+^n}\right),\|x\|_2=1} \left\|W_{I_+^n}^{(n)}x\right\|_2\\ &\le \max_{x\in\ker\left(\overline{X}_{I_+^n}\right),\|x\|_2=1} \left\|W_{I_+^n}^{(n)}x\right\|_2\\ &= \max_{x\in\ker\left(\overline{X}_{I_+^n}\right),\|x\|_2=1} \left\|\left(W^{(n)} - \overline{X}\right)_{I_+^n}x\right\|_2\\ &\le \left\|\left(W^{(n)} - \overline{X}\right)_{I_+^n}\right\|_{2\longrightarrow 2} \le \left\|\left(W^{(n)} - \overline{X}\right)_{I_+^n}\right\|_F.\end{aligned} \tag{A.5}$$

□

Lemma A.2. *Let $e \in \mathbb{C}^d$ and S be a linear subspace of $\mathbb{C}^{n_1\times n_2}$. If $r = \max\{\operatorname{rank}(M): M \in S\}$, we have*

$$\left\|(\mathscr{A}^*(e))_S\right\|_F \le \sqrt{1+\delta_r}\|e\|_2. \tag{A.6}$$

Proof. In fact, it holds that

$$\begin{aligned}\left\|(\mathscr{A}^*(e))_S\right\|_F^2 &= \langle \mathscr{A}^*(e), (\mathscr{A}^*(e))_S\rangle_F = \langle e, \mathscr{A}((\mathscr{A}^*(e))_S)\rangle\\ &\le \|e\|_2\left\|\mathscr{A}((\mathscr{A}^*(e))_S)\right\|_2 \le \|e\|_2\sqrt{1+\delta_r}\left\|(\mathscr{A}^*(e))_S\right\|_F,\end{aligned} \tag{A.7}$$

and divide through by $\|(\mathscr{A}^*(e))_S\|_F$ to complete the proof. □

Lemma A.3. *Let $M_1, M_2 \in \mathbb{C}^{n_1\times n_2}$ and I be a linear subspace of $\mathbb{C}^{n_1\times n_2}$. If $r_1 = \max\{\operatorname{rank}(M): M = k_1M_1 + k_2M_2, k_1, k_2 \in \mathbb{C}\}$ and $r_2 = \max\left\{\operatorname{rank}(M): M \in \overline{\operatorname{span}\{I, M_1\}}\right\}$, we have*

$$\begin{aligned}\left|\langle M_1, (Id - \mathscr{A}^*\mathscr{A})(M_2)\rangle_F\right| &\le \delta_r\|M_1\|_F\|M_2\|_F,\\ \left\|((Id - \mathscr{A}^*\mathscr{A})(M_1))_I\right\|_F &\le \delta_r\|M_1\|_F.\end{aligned} \tag{A.8}$$

Proof. To the first inequality, let $S = \overline{\text{span}\{M_1, M_2\}}$; thus, $(M_1)_S = M_1$, $(M_2)_S = M_2$ and $r_1 = \max\{\text{rank}(M)\colon M \in S\}$. Therefore, we have

$$\begin{aligned}\left|\langle M_1, (Id - \mathscr{A}^*\mathscr{A})(M_2)\rangle_F\right| &= \left|\langle M_1, M_2\rangle_F - \langle \mathscr{A}(M_1), \mathscr{A}(M_2)\rangle_F\right| \\ &= \left|\langle (M_1)_S, (M_2)_S\rangle_F - \langle \mathscr{A}(M_1)_S, \mathscr{A}(M_2)_S\rangle_F\right| \\ &= \left|\langle (M_1)_S, (Id - \mathscr{A}^*\mathscr{A})(M_2)_S\rangle\right| \le \|M_1\|_F \|(Id - \mathscr{A}^*\mathscr{A})(M_2)_S\|_F \\ &\le \delta_{r_1}\|M_1\|_F\|M_2\|_F.\end{aligned} \tag{A.9}$$

To the second inequality, we have

$$\begin{aligned}&\|((Id - \mathscr{A}^*\mathscr{A})(M_1))_I\|_F^2 \\ &\quad = \langle ((Id - \mathscr{A}^*\mathscr{A})(M_1))_I, (Id - \mathscr{A}^*\mathscr{A})(M_1)\rangle_F \\ &\quad \le \delta_{r_2}\|((Id - \mathscr{A}^*\mathscr{A})(M_2))_I\|_F\|M_1\|_F,\end{aligned} \tag{A.10}$$

and divide through by $\|((Id - \mathscr{A}^*\mathscr{A})(M_2))_I\|_F$ to complete the proof. □

Lemma A.4 *(see [20]). Let $X^{(l)} = U_l\Sigma_l V_l^*$ be a rank r matrix and S_l be the tangent space of the rank r matrix manifold at $X^{(l)}$. Let X be another rank r matrix. Then,*

$$\left\|\left(I - \mathscr{P}_{S_l}\right)(X)\right\|_F \le \frac{1}{\sigma_{\min}(X)}\left\|X^{(l)} - X\right\|_F^2. \tag{A.11}$$

Lemma A.5 *(see [20]). Let $X^{(l)} = U_l\Sigma_l V_l^*$ be a rank r matrix with the tangent space S_l. Let X be another rank r matrix. Then, the Frobenius norm of $\mathscr{P}_{S_l}\mathscr{A}^*\mathscr{A}(I - \mathscr{P}_{S_l})(X)$ can be bounded as*

$$\left\|\mathscr{P}_{S_l}\mathscr{A}^*\mathscr{A}\left(I - \mathscr{P}_{S_l}\right)(X)\right\|_F \le \delta_{3r}\left\|\left(I - \mathscr{P}_{S_l}\right)(X)\right\|_F. \tag{A.12}$$

Lemma A.4 and A.5 can be found in [20] and the detail proofs are omitted here. Based on lemmas above, we start to prove Theorems 2 and 3.

Proof of Theorem 2. To prove inequality (15), we just need to prove

$$\left\|X^{(n+1)} - \overline{X}\right\|_F \le \rho\left\|X^{(n)} - \overline{X}\right\|_F. \tag{A.13}$$

First of all, we have the following inequality:

$$\begin{aligned}&\left\|X^{(n)} - \overline{X}\right\|_F^2 \\ &\quad = \left\|\left(X^{(n)}\right)_{I^n} - (\overline{X})_{I^n}\right\|_F^2 + \left\|(\overline{X})_{\overline{I}\setminus I^n}\right\|_F^2 \\ &\quad \le 2\Big(\left\|\left(W^{(n)}\right)_{I^n} - (\overline{X})_{I^n}\right\|_F^2 + \left\|\left(W^{(n)}\right)_{I^n} - \left(X^{(n)}\right)_{I^n}\right\|_F^2 \\ &\qquad + \left\|\left(W^{(n)}\right)_{\overline{I}\setminus I^n} - (\overline{X})_{\overline{I}\setminus I^n}\right\|_F^2 + \left\|\left(W^{(n)}\right)_{\overline{I}\setminus I^n}\right\|_F^2\Big) \\ &\quad = 2\Big(\left\|\left(W^{(n)}\right)_{I^n + \overline{I}} - (\overline{X})_{I^n + \overline{I}}\right\|_F^2 + \left\|\left(W^{(n)}\right)_{I^n} - \left(X^{(n)}\right)_{I^n}\right\|_F^2 + \left\|\left(W^{(n)}\right)_{\overline{I}\setminus I^n}\right\|_F^2\Big) \\ &\quad \triangleq J_1 + J_2 + J_3,\end{aligned} \tag{A.14}$$

since the space $(\mathrm{C}^{n_1\times n_2}, \|\cdot\|_F)$ is an Euclidean space and $(X^{(n)})_{\overline{I}\setminus I^n} = 0$. In the following, we will bound J_1, J_2, and J_3 one by one.

According to the property of the orthogonal projection, we obtain $J_1 \le 2\|(W^{(n)} - \overline{X})_{I_+^n + \overline{I}}\|_F^2$.

Besides, bound of J_3 begins with the following equations:

$$\begin{aligned}\left\|W^{(n)}\right\|_F^2 &= \left\|W^{(n)} - \left(W^{(n)}\right)_{I^n}\right\|_F^2 + \left\|\left(W^{(n)}\right)_{I^n}\right\|_F^2 \\ &= \left\|W^{(n)} - \left(W^{(n)}\right)_{I^n}\right\|_F^2 + \left\|\left(W^{(n)}\right)_{I^n \cap \overline{I}}\right\|_F^2 + \left\|\left(W^{(n)}\right)_{I^n \setminus \overline{I}}\right\|_F^2 \\ &= \left\|W^{(n)} - \left(W^{(n)}\right)_{\overline{I}}\right\|_F^2 + \left\|\left(W^{(n)}\right)_{\overline{I}}\right\|_F^2 \\ &= \left\|W^{(n)} - \left(W^{(n)}\right)_{\overline{I}}\right\|_F^2 + \left\|\left(W^{(n)}\right)_{\overline{I}\cap I^n}\right\|_F^2 + \left\|\left(W^{(n)}\right)_{\overline{I}\setminus I^n}\right\|_F^2.\end{aligned} \tag{A.15}$$

Since the definition of I^n, $(W^{(n)})_{I^n}$ is the best rank r approximation of $W^{(n)}$. Thus, it implies that

$$J_3 \le 2\left\|\left(W^{(n)}\right)_{I^n\setminus\bar{I}}\right\|_F^2 = 2\left\|\left(W^{(n)} - \bar{X}\right)_{I^n\setminus\bar{I}}\right\|_F^2 \le 2\left\|\left(W^{(n)} - \bar{X}\right)_{I_+^n+\bar{I}}\right\|_F^2, \quad (A.16)$$

where $\bar{X}_{I^n\setminus\bar{I}} = 0$.

In addition, the thresholding algorithm is a kind of algorithm for singular value of the matrix. It is easy to obtain that

$$J_2 = 2\left\|\Sigma^{(n)} - H_{\sigma_{r+1}^n}\left(\Sigma^{(n)}\right)\right\|_F^2 = 2\sum_{i=1}^{r}\left(\sigma_i^n - h_{\sigma_{r+1}^n}\left(\sigma_i^n\right)\right)^2. \quad (A.17)$$

Because the thresholding function is a BRT function with a band parameter c, it holds that

$$J_2 \le 2c^2 r\left(\sigma_{r+1}^n\right)^2 \le 2c^2 r\left\|\left(W^{(n)}\right)_{I_+^n} - (\bar{X})_{I_+^n}\right\|_F^2 \le 2c^2 r\left\|\left(W^{(n)} - \bar{X}\right)_{I_+^n+\bar{I}}\right\|_F^2, \quad (A.18)$$

where the second inequality follows from Lemma A.1. Therefore, we have

$$\left\|X^{(n)} - \bar{X}\right\|_F^2 \le J_1 + J_2 + J_3 \le \left(4 + 2c^2 r\right)\left\|\left(W^{(n)} - \bar{X}\right)_{I_+^n+\bar{I}}\right\|_F^2. \quad (A.19)$$

In the meantime, we have $W^{(n+1)} = X^{(n)} + \mathscr{P}_{S_n}\mathscr{A}^*(b - \mathscr{A}(X^{(n)}))$. Thus, it implies that

$$\begin{aligned}\left\|W^{(n+1)} - \bar{X}\right\|_F &= \left\|X^{(n)} - \bar{X} + \mathscr{P}_{S_n}\mathscr{A}^*\left(b - \mathscr{A}\left(X^{(n)}\right)\right)\right\|_F \\ &= \left\|X^{(n)} - \bar{X} + \mathscr{P}_{S_n}\mathscr{A}^*\mathscr{A}\left(\bar{X} - X^{(n)}\right)\right\|_F \\ &\le \left\|X^{(n)} - \bar{X} - \mathscr{P}_{S_n}\mathscr{A}^*\mathscr{A}\mathscr{P}_{S_n}\left(X^{(n)} - \bar{X}\right)\right\|_F \\ &\quad + \left\|\mathscr{P}_{S_n}\mathscr{A}^*\mathscr{A}\left(I - \mathscr{P}_{S_n}\right)\left(X^{(n)} - \bar{X}\right)\right\|_F \\ &\le \left\|\left(\mathscr{P}_{S_n} - \mathscr{P}_{S_n}\mathscr{A}^*\mathscr{A}\mathscr{P}_{S_n}\right)\left(X^{(n)} - \bar{X}\right)\right\|_F \\ &\quad + \left\|\left(I - \mathscr{P}_{S_n}\right)\left(X^{(n)} - \bar{X}\right)\right\|_F \\ &\quad + \left\|\mathscr{P}_{S_n}\mathscr{A}^*\mathscr{A}\left(I - \mathscr{P}_{S_n}\right)\left(X^{(n)} - \bar{X}\right)\right\|_F \\ &= \left\|\left(\mathscr{P}_{S_n} - \mathscr{P}_{S_n}\mathscr{A}^*\mathscr{A}\mathscr{P}_{S_n}\right)\left(X^{(n)} - \bar{X}\right)\right\|_F \\ &\quad + \left\|\left(I - \mathscr{P}_{S_n}\right)\bar{X}\right\|_F + \left\|\mathscr{P}_{S_n}\mathscr{A}^*\mathscr{A}\left(I - \mathscr{P}_{S_n}\right)\bar{X}\right\|_F \\ &\triangleq J_1' + J_2' + J_3',\end{aligned} \quad (A.20)$$

where the second equality follows from the fact $(I - \mathscr{P}_{S_n})(X^{(n)}) = 0$.

Bound of J_1': it is a fact that RIC can be represented as

$$\begin{aligned}\delta_{2r} &= \sup_{0<\operatorname{rank}(X)\le 2r} \frac{\|\mathscr{A}(X)\|_2^2 - \|X\|_F^2}{\|X\|_F^2} \\ &\ge \sup_{0<\operatorname{rank}(X)} \frac{\left\|\left(\mathscr{A}\mathscr{P}_{S_n}\right)(X)\right\|_2^2 - \left\|\mathscr{P}_{S_n}X\right\|_F^2}{\left\|\mathscr{P}_{S_n}X\right\|_F^2} \\ &\ge \sup_{0<\operatorname{rank}(X)} \frac{\left\langle\left(\mathscr{P}_{S_n} - \mathscr{P}_{S_n}^*\mathscr{A}^*\mathscr{A}\mathscr{P}_{S_n}\right)(X), X\right\rangle}{\|X\|_F^2} \\ &= \sup_{0<\operatorname{rank}(X)} \frac{\left\|\left(\mathscr{P}_{S_n} - \mathscr{P}_{S_n}^*\mathscr{A}^*\mathscr{A}\mathscr{P}_{S_n}\right)(X)\right\|_F}{\|X\|_F} \\ &\ge \frac{\left\|\left(\mathscr{P}_{S_n} - \mathscr{P}_{S_n}^*\mathscr{A}^*\mathscr{A}\mathscr{P}_{S_n}\right)\left(X^{(n)} - \bar{X}\right)\right\|_F}{\left\|X^{(n)} - \bar{X}\right\|_F},\end{aligned} \quad (A.21)$$

where the second equal sign is based on the fact that $\mathscr{P}_{S_n} - \mathscr{P}_{S_n}^*\mathscr{A}^*\mathscr{A}\mathscr{P}_{S_n}$ is a self-adjoint operator. Thus, the first term J_1' can be bounded as

$$J_1' \le \delta_{2r}\left\|X^{(n)} - \bar{X}\right\|_F. \quad (A.22)$$

Bound of J_2': according to Lemma A.4, the second term J_2' can be bounded as

$$J_2' \le \frac{1}{\sigma_{\min}(\bar{X})}\left\|X^{(n)} - \bar{X}\right\|_F^2. \quad (A.23)$$

Bound of J_3': according to Lemma A.5, the third term J_3' can be bounded as

$$\begin{aligned}J_3' &\le \delta_{3r}\left\|\left(I - \mathscr{P}_{S_n}\right)\bar{X}\right\|_F = \delta_{3r}\left\|\left(I - \mathscr{P}_{S_n}\right)\left(X^{(n)} - \bar{X}\right)\right\|_F \\ &\le \delta_{3r}\left\|X^{(n)} - \bar{X}\right\|_F,\end{aligned} \quad (A.24)$$

where the equality follows from the fact $(I - \mathscr{P}_{S_n})(\bar{X}) = 0$.

Therefore, we have

$$\left\|W^{(n+1)} - \bar{X}\right\|_F \le \left(\delta_{2r} + \delta_{3r} + \frac{\left\|X^{(n)} - \bar{X}\right\|_F}{\sigma_{\min}(\bar{X})}\right)\left\|X^{(n)} - \bar{X}\right\|_F. \quad (A.25)$$

Combining (A.19) and (A.25), it holds that

$$\begin{aligned}&\left\|X^{(n+1)} - \bar{X}\right\|_F \\ &\quad \le \sqrt{4 + 2c^2 r}\left(\delta_{2r} + \delta_{3r} + \frac{\left\|X^{(n)} - \bar{X}\right\|_F}{\sigma_{\min}(\bar{X})}\right)\left\|X^{(n)} - \bar{X}\right\|_F.\end{aligned} \quad (A.26)$$

Considering $X^{(0)} = H_{\sigma_{r+1}^0}(W^{(0)}) = H_{\sigma_{r+1}^0}(\mathscr{A}^*(b))$ and (A.19), it implies that

$$\begin{aligned}\left\|X^{(0)} - \bar{X}\right\|_F &\le \sqrt{4 + 2c^2 r}\left\|\left(W^{(0)} - \bar{X}\right)_{I_+^0+\bar{I}}\right\|_F \\ &= \sqrt{4 + 2c^2 r}\left\|\left(\left(\mathscr{A}^*\mathscr{A} - I\right)(\bar{X})\right)_{I_+^0+\bar{I}}\right\|_F \\ &\le \sqrt{4 + 2c^r}\,\delta_{2r+1}\|\bar{X}\|_F,\end{aligned} \quad (A.27)$$

where the second inequality follows from Lemma A.3.

Define

$$\rho := \sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{\sqrt{4+2c^2r}\,\delta_{2r+1}}{\sigma_{\min}(\overline{X})}\right). \quad \text{(A.28)}$$

If $\rho < 1$, we use mathematical induction to prove (A.13). When $n = 0$, we obtain

$$\begin{aligned}\left\|X^{(1)}-\overline{X}\right\|_F &\le \sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{\|X^{(0)}-\overline{X}\|_F}{\sigma_{\min}(\overline{X})}\right)\left\|X^{(0)}-\overline{X}\right\|_F\\ &\le \sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{\sqrt{4+2c^2r}\,\delta_{2r+1}\|\overline{X}\|_F}{\sigma_{\min}(\overline{X})}\right)\left\|X^{(0)}-\overline{X}\right\|_F\\ &= \rho\left\|X^{(0)}-\overline{X}\right\|_F,\end{aligned} \quad \text{(A.29)}$$

and inequality (A.13) holds true.

If inequality (A.13) holds true for $n \le k-1$ (where k is a positive integer), then when $n = k$, we obtain

$$\begin{aligned}\left\|X^{(k+1)}-\overline{X}\right\|_F &\le \sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{\|X^{(k)}-\overline{X}\|_F}{\sigma_{\min}(\overline{X})}\right)\left\|X^{(k)}-\overline{X}\right\|_F\\ &\le \sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{\rho\|X^{(k)}-\overline{X}\|_F}{\sigma_{\min}(\overline{X})}\right)\left\|X^{(k)}-\overline{X}\right\|_F \le \cdots\\ &\le \sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{\rho^{k+1}\|X^{(0)}-\overline{X}\|_F}{\sigma_{\min}(\overline{X})}\right)\left\|X^{(k)}-\overline{X}\right\|_F\\ &\le \sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{\|X^{(0)}-\overline{X}\|_F}{\sigma_{\min}(\overline{X})}\right)\left\|X^{(k)}-\overline{X}\right\|_F\\ &\le \sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{\sqrt{4+2c^2r}\,\delta_{2r+1}\|\overline{X}\|_F}{\sigma_{\min}(\overline{X})}\right)\left\|X^{(k)}-\overline{X}\right\|_F\\ &= \rho\left\|X^{(k)}-\overline{X}\right\|_F,\end{aligned} \quad \text{(A.30)}$$

and inequality (A.13) holds true.

Furthermore, if

$$\delta_{3r} \le \frac{1}{2\sqrt{4+2c^2r}+(4+2c^2r)\sqrt{r}\,\sigma_{\max}(\overline{X})/\sigma_{\min}(\overline{X})}, \quad \text{(A.31)}$$

it implies that

$$\rho \le \sqrt{4+2c^2r}\,\delta_{3r}\left(2+\frac{\sqrt{4+2c^2r}\,\sqrt{r}\,\sigma_{\max}(\overline{X})}{\sigma_{\min}(\overline{X})}\right) < 1, \quad \text{(A.32)}$$

where $\delta_{2r} \le \delta_{2r+1} \le \delta_{3r}$ and $\|\overline{X}\|_F \le \sqrt{r}\,\sigma_{\max}(\overline{X})$. □

Proof of Theorem 3. If there exists a positive integer n such that

$$\left\|X^{(n)}-\overline{X}\right\|_F \le 4(\mu+1)\sqrt{2+c^2r}\,\eta, \quad \text{(A.33)}$$

the result of Theorem 3 holds true. If

$$\left\|X^{(n)}-\overline{X}\right\|_F > 4(\mu+1)\sqrt{2+c^2r}\,\eta, \quad \text{(A.34)}$$

for any positive integer n, it implies that

$$\begin{aligned}&2\left(\left\|X^{(n)}-\overline{X}\right\|_F-(\mu+1)\sqrt{4+2c^2r}\,\sqrt{1+\delta_{2r+1}}\,\eta\right)\\ &\quad\ge \left\|X^{(n)}-\overline{X}\right\|_F,\end{aligned} \quad \text{(A.35)}$$

since $\delta_{2r+1} \le \delta_{3r} < 1$.

To prove inequality (19), we firstly need to prove

$$\left\|X^{(n+1)}-\overline{X}\right\|_F \le \rho\left\|X^{(n)}-\overline{X}\right\|_F+\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\,\eta, \quad \text{(A.36)}$$

where $\rho = (\mu/(\mu+1)) < 1$. Similar to the discussion of Theorem 2, inequality (A.19) holds. Meanwhile, we also have

$$\begin{aligned}\left\|W^{(n+1)}-\overline{X}\right\|_F &= \left\|X^{(n)}-\overline{X}+\mathscr{P}_{S_n}\mathscr{A}^*\left(b-\mathscr{A}\left(X^{(n)}\right)\right)\right\|_F \\ &= \left\|X^{(n)}-\overline{X}+\mathscr{P}_{S_n}\mathscr{A}^*\left(\mathscr{A}(\overline{X})+e-\mathscr{A}\left(X^{(n)}\right)\right)\right\|_F \\ &\le \left\|X^{(n)}-\overline{X}-\mathscr{P}_{S_n}\mathscr{A}^*\mathscr{A}\left(X^{(n)}-\overline{X}\right)\right\|_F + \left\|\mathscr{P}_{S_n}\mathscr{A}^*(e)\right\|_F \\ &\le \left\|\left(\mathscr{P}_{S_n}-\mathscr{P}_{S_n}\mathscr{A}^*\mathscr{A}\mathscr{P}_{S_n}\right)\left(X^{(n)}-\overline{X}\right)\right\|_F + \left\|\left(I-\mathscr{P}_{S_n}\right)\left(X^{(n)}-\overline{X}\right)\right\|_F \\ &\quad + \left\|\mathscr{P}_{S_n}\mathscr{A}^*\mathscr{A}\left(I-\mathscr{P}_{S_n}\right)\left(X^{(n)}-\overline{X}\right)\right\|_F + \left\|\mathscr{P}_{S_n}\mathscr{A}^*(e)\right\|_F \\ &:= J_1'+J_2'+J_3'+J_4'.\end{aligned} \tag{A.37}$$

In the meantime, similar to the proof of Theorem 2, we can also obtain bounds of J_1', J_2', and J_3', and according to Lemma A.2, bound of J_4' can be described as

$$\left\|\mathscr{P}_{S_n}\mathscr{A}^*(e)\right\|_F \le \sqrt{1+\delta_{2r}}\|e\|_2. \tag{A.38}$$

Thus, we obtain

$$\begin{aligned}\left\|W^{(n+1)}-\overline{X}\right\|_F &\le \left(\delta_{2r}+\delta_{3r}+\frac{\left\|X^{(n)}-\overline{X}\right\|_F}{\sigma_{\min}(\overline{X})}\right)\left\|X^{(n)}-\overline{X}\right\|_F \\ &\quad + \sqrt{1+\delta_{2r}}\|e\|_2.\end{aligned} \tag{A.39}$$

Combining (A.19) and (A.39), it holds that

$$\begin{aligned}\left\|X^{(n+1)}-\overline{X}\right\|_F &\le \sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{\left\|X^{(n)}-\overline{X}\right\|_F}{\sigma_{\min}(\overline{X})}\right)\left\|X^{(n)}-\overline{X}\right\|_F \\ &\quad + \sqrt{4+2c^2r}\sqrt{1+\delta_{2r}}\|e\|_2 \\ &\le \sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{\left\|X^{(n)}-\overline{X}\right\|_F}{\sigma_{\min}(\overline{X})}\right)\left\|X^{(n)}-\overline{X}\right\|_F \\ &\quad + \sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\|e\|_2.\end{aligned} \tag{A.40}$$

Because we have (A.35), it implies that

$$\begin{aligned}\left\|X^{(n+1)}-\overline{X}\right\|_F &\le \sqrt{4+2c^2r}\Bigg(\delta_{2r}+\delta_{3r}+\frac{2}{\sigma_{\min}(\overline{X})}\Bigg(\left\|X^{(n)}-\overline{X}\right\|_F \\ &\quad -(\mu+1)\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\,\eta\Bigg)\Bigg)\left\|X^{(n)}-\overline{X}\right\|_F \\ &\quad + \sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\|e\|_2.\end{aligned} \tag{A.41}$$

Considering $X^{(0)} = H_{\sigma^0_{r+1}}(W^{(0)}) = H_{\sigma^0_{r+1}}(\mathscr{A}^*(b))$ and (A.19), it implies that

$$\begin{aligned}\left\|X^{(0)}-\overline{X}\right\|_F &\le \sqrt{4+2c^2r}\left\|\left(W^{(0)}-\overline{X}\right)\right\|_F \\ &= \sqrt{4+2c^2r}\left\|\left((\mathscr{A}^*\mathscr{A}-I)(\overline{X})+\mathscr{A}^*(e)\right)_{I^0_+ + \overline{I}}\right\|_F \\ &\le \sqrt{4+2c^2r}\left\|\left((\mathscr{A}^*\mathscr{A}-I)(\overline{X})\right)_{I^0_+ + \overline{I}}\right\|_F \\ &\quad + \sqrt{4+2c^2r}\left\|\left(\mathscr{A}^*(e)\right)_{I^0_+ + \overline{I}}\right\|_F \\ &\le \sqrt{4+2c^2r}\,\delta_{2r+1}\|\overline{X}\|_F + \sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\|e\|_2,\end{aligned} \tag{A.42}$$

where the third inequality follows from Lemmas A.2 and A.3. To prove inequality (A.36), we just need to prove the following inequality:

$$\begin{aligned}&\sqrt{4+2c^2r}\Bigg(\delta_{2r}+\delta_{3r}+\frac{2}{\sigma_{\min}(\overline{X})} \\ &\quad \cdot \left(\left\|X^{(n)}-\overline{X}\right\|_F-(\mu+1)\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\,\eta\right)\Bigg) \le \rho.\end{aligned} \tag{A.43}$$

Based on the fact that

$$\delta_{3r} \le \frac{\mu}{2\sqrt{4+2c^2r}(\mu+1)\left(1+\left(\sqrt{4+2c^2r}/\sigma_{\min}(\overline{X})\right)\left(\|\overline{X}\|_F-\mu\eta\right)\right)}, \tag{A.44}$$

now we use mathematical induction to prove (A.43).

When $n=0$, we obtain

$$\begin{aligned}
&\sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{2}{\sigma_{\min}(\overline{X})}\left(\left\|X^{(0)}-\overline{X}\right\|_F-(\mu+1)\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\,\eta\right)\right)\\
&\le\sqrt{4+2c^2r}\left(2\delta_{3r}+\frac{2}{\sigma_{\min}(\overline{X})}\left(\left\|X^{(0)}-\overline{X}\right\|_F-(\mu+1)\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\,\eta\right)\right)\\
&\le\sqrt{4+2c^2r}\left(2\delta_{3r}+\frac{2}{\sigma_{\min}(\overline{X})}\left(\sqrt{4+2c^2r}\,\delta_{2r+1}\|\overline{X}\|_F+\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\,\eta-(\mu+1)\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\,\eta\right)\right)\\
&\le\sqrt{4+2c^2r}\left(2\delta_{3r}+\frac{2\sqrt{4+2c^2r}}{\sigma_{\min}(\overline{X})}\left(\delta_{3r}\|\overline{X}\|_F-\mu\sqrt{1+\delta_{2r+1}}\,\eta\right)\right)\\
&\le\sqrt{4+2c^2r}\left(2\delta_{3r}+\frac{2\sqrt{4+2c^2r}}{\sigma_{\min}(\overline{X})}\left(\delta_{3r}\|\overline{X}\|_F-\mu\delta_{3r}\eta\right)\right)\\
&=\delta_{3r}\sqrt{4+2c^2r}\left(2+\frac{2\sqrt{4+2c^2r}}{\sigma_{\min}(\overline{X})}\left(\|\overline{X}\|_F-\mu\eta\right)\right)\\
&\le\rho,
\end{aligned} \tag{A.45}$$

where the fourth inequality sign is based on $\sqrt{1+\delta_{2r+1}}\ge 1\ge\delta_{3r}$. Thus, inequality (A.43) holds true.

If inequality (A.43) holds true for $n\le k-1$ (where k is a positive integer), it implies that

$$\left\|X^{(n+1)}-\overline{X}\right\|_F\le\rho\left\|X^{(n)}-\overline{X}\right\|_F+\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\eta. \tag{A.46}$$

Then, when $n=k$, we can obtain

$$\begin{aligned}
&\sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{2}{\sigma_{\min}(\overline{X})}\left(\left\|X^{(k)}-\overline{X}\right\|_F-(\mu+1)\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\,\eta\right)\right)\\
&\le\sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{2}{\sigma_{\min}(\overline{X})}\left(\rho\left\|X^{(k-1)}-\overline{X}\right\|_F-\mu\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\,\eta\right)\right)\\
&=\sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{2}{\sigma_{\min}(\overline{X})}\rho\left(\left\|X^{(k-1)}-\overline{X}\right\|_F-(\mu+1)\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\,\eta\right)\right)\\
&\le\sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{2}{\sigma_{\min}(\overline{X})}\left(\left\|X^{(k-1)}-\overline{X}\right\|_F-(\mu+1)\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\,\eta\right)\right)\\
&\le\cdots\le\sqrt{4+2c^2r}\left(\delta_{2r}+\delta_{3r}+\frac{2}{\sigma_{\min}(\overline{X})}\left(\left\|X^{(0)}-\overline{X}\right\|_F-(\mu+1)\sqrt{4+2c^2r}\sqrt{1+\delta_{2r+1}}\,\eta\right)\right)\le\rho,
\end{aligned} \tag{A.47}$$

and inequality (A.43) holds true. □

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request..

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundations of China (11771347, 91730306, 41390454, and 11271297).

References

[1] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[2] J. F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *Siam Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2008.

[3] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM*, vol. 58, no. 3, 2000.

[4] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 171–184, 2013.

[5] E. J. Candès, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM Review*, vol. 57, no. 2, pp. 225–251, 2015.

[6] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via wirtinger flow: theory and algorithms," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1985–2007, 2015.

[7] Y. Yu, J. Peng, and S. Yue, "A new nonconvex approach to low-rank matrix completion with application to image inpainting," *Multidimensional Systems and Signal Processing*, vol. 30, no. 1, pp. 145–174, 2019.

[8] Y. Gao, X. Han, and M. Ma, "Recovery of low-rank matrices based on the rank null space properties," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 15, no. 4, p. 1750032, 2017.

[9] F. Zhao, J. Peng, and A. Cui, "Design strategy of thresholding operator for low-rank matrix recovery problem," *Signal Processing*, vol. 171, Article ID 107510, 2020.

[10] l. E. Frank and J. H. Friedman, "A statistical view of some chemometrics regression tools," *Technometrics*, vol. 35, no. 2, pp. 109–135, 1993.

[11] C.-H. Zhang, "Nearly unbiased variable selection under minimax concave penalty," *The Annals of Statistics*, vol. 38, no. 2, pp. 894–942, 2010.

[12] J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties," *Journal of the American Statistical Association*, vol. 96, no. 456, pp. 1348–1360, 2001.

[13] J. H. Friedman, "Fast sparse regression and classification," *International Journal of Forecasting*, vol. 28, no. 3, pp. 722–738, 2012.

[14] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Transactions on Image Processing*, vol. 4, no. 7, pp. 932–946, 1995.

[15] J. Trzasko and A. Manduca, "Highly undersampled magnetic resonance image reconstruction via homotopic l_0-minimization," *IEEE Transactions on Medical Imaging*, vol. 28, no. 1, pp. 106–121, 2008.

[16] L. Cui, L. Liu, and X. Z. Huang, "A sufficient condition for low-rank recovery via iterative hard thresholding pursuit," *Applied Mechanics and Materials*, vol. 644–650, pp. 2378–2381, 2014.

[17] Z. Xu, X. Chang, F. Xu, and H. Zhang, "$L_{1/2}$ regularization: a thresholding representation theory and a fast solver," *IEEE Transactions on Neural Networks & Learning Systems*, vol. 23, no. 7, pp. 1013–1027, 2012.

[18] W. Cao, J. Sun, and Z. Xu, "Fast image deconvolution using closed-form thresholding formulas of regularization," *Journal of Visual Communication and Image Representation*, vol. 24, no. 1, pp. 31–41, 2013.

[19] R. Chartrand, "Shrinkage mappings and their induced penalty functions," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1026–1029, Florence, Italy, May 2014.

[20] K. Wei, J.-F. Cai, T. F. Chan, and S. Leung, "Guarantees of riemannian optimization for low rank matrix recovery," *SIAM Journal on Matrix Analysis and Applications*, vol. 37, no. 3, pp. 1198–1222, 2016.

[21] A. Cui, J. Peng, H. Li, C. Zhang, and Y. Yu, "Affine matrix rank minimization problem via non-convex fraction function penalty," *Journal of Computational and Applied Mathematics*, vol. 336, pp. 353–374, 2018.

[22] D. Peng, N. Xiu, and J. Yu, "$s_{0.5}$ regularization methods and fixed point algorithms for affine rank minimization problems," *Computational Optimization and Applications*, vol. 67, no. 3, pp. 543–569, 2013.

[23] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Review*, vol. 52, no. 3, pp. 471–501, 2010.

[24] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer Science & Business Media, Berlin, Germany, 2010.

[25] J. Woodworth and R. Chartrand, "Compressed sensing recovery via nonconvex shrinkage penalties," *Inverse Problems*, vol. 32, no. 7, Article ID 075004, 2016.

[26] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and bregman iterative methods for matrix rank minimization," *Mathematical Programming*, vol. 128, no. 1-2, pp. 321–353, 2009.

[27] A. Ahmed, B. Recht, and J. Romberg, "Blind deconvolution using convex programming," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1711–1732, 2014.

[28] X. Li, S. Ling, T. Strohmer, and K. Wei, "Rapid, robust, and reliable blind deconvolution via nonconvex optimization," *Applied and Computational Harmonic Analysis*, vol. 47, no. 3, pp. 893–934, 2019.