The perturbation analysis of low-rank matrix stable recovery

Jianwen Huang[1], Jianjun Wang[1,2]*, Feng Zhang[1], Hailin Wang[1]

[1] *School of Mathematic & Statistics, Southwest University, Chongqing, 400715, China*

[2] *Research Center for Artificial Intelligence & Education Big Data, Southwest University, Chongqing, 400715, China*

Abstract. In this paper, we bring forward a completely perturbed nuclear norm minimization method to tackle a formulation of completely perturbed low-rank matrices recovery. In view of the matrix version of the restricted isometry property (RIP) and the Frobenius-robust rank null space property (FRNSP), this paper extends the investigation to a completely perturbed model taking into consideration not only noise but also perturbation, derives sufficient conditions guaranteeing that low-rank matrices can be robustly and stably reconstructed under the completely perturbed scenario, as well as finally presents an upper bound estimation of recovery error. The upper bound estimation can be described by two terms, one concerning the total noise, and another regarding the best r-approximation error. Specially, we not only improve the condition corresponding with RIP, but also ameliorate the upper bound estimation in case the results reduce to the general case. Furthermore, in the case of $\mathcal{E}=0$, the obtaining conditions are optimal.

Key words. Compressed sensing; low-rank matrices; perturbation of measurement operator; constrained nuclear norm minimization.

1 Introduction

Low-rank matrix recovery (LMR) is a rapidly developing research field which has a variety of applications including quantum state tomograph [1], machine learning [2] [3], system identification [4], and computer vision [5]. From the mathematical point of view, we can depict it as

$$y=\mathcal{A}(X), \tag{1.1}$$

where $y\in\mathbb{R}^m$ is a given vector, $X\in\mathbb{R}^{n_1\times n_2}$ is the desired matrix being low-rank or approximately low-rank, and $\mathcal{A}:\mathbb{R}^{n_1\times n_2}\to\mathbb{R}^m$ is a known measurement operator which is given by

$$\mathcal{A}(X)=\left[\mathrm{tr}(X^\top A^{(1)}),\mathrm{tr}(X^\top A^{(2)}),\cdots,\mathrm{tr}(X^\top A^{(m)})\right]^\top. \tag{1.2}$$

*Corresponding author, E-mail: wjjmath@gmail.com, wjj@swu.edu.cn(J.J. Wang)

Here, $\mathrm{tr}(\cdot)$ is the trace function, and $X^\top$ is the transpose of X and $A^{(1)}, A^{(2)}, \cdots, A^{(m)}$ are called measurement matrices. The objective of LMR is to reconstruct the unknown low-rank matrix X basing on the linear measurement y and the measurement operator $\mathcal{A}$.

In practice, the measurement y is generally contaminated by noise z, and think about the noisy LMR model as follows:

$$\hat{y} = \mathcal{A}(X) + z, \tag{1.3}$$

where $\hat{y}$ is the observed measurement which is corrupted by the noise vector z, and z is the additive noise, independent of X. However, more LMR models can be encountered in which not only the measurement y is perturbed by z, but also the measurement operator $\mathcal{A}$ is obstructed by $\mathcal{E}$, namely, the measurement operator $\mathcal{A}$ is substituted with $\hat{\mathcal{A}} = \mathcal{A} + \mathcal{E}$ in (1.3) brought in a multiplicative noise $\mathcal{E}(X)$ relating with X. The totally perturbed problems usually arise in a great deal of applications involving remote sensing [6], source separation [7], telecommunications [8] and so forth. To look for the optimal solution from such completely perturbed problem, a general method is to solve the constrained nuclear norm minimization (for short as NNM) as follows:

$$\min_{\tilde{Z}\in\mathbb{R}^{n_1\times n_2}} \|\tilde{Z}\|_* \text{ s.t. } \|\hat{\mathcal{A}}(\tilde{Z}) - \hat{y}\|_2 \le \epsilon'_{\mathcal{A},r,y}, \tag{1.4}$$

where $\epsilon'_{\mathcal{A},r,y} \ge 0$ is a total noise, and $\|\tilde{Z}\|_*$ is the nuclear norm of the matrix X, viz, the sum of its singular values. When $n_1 = n_2$ and the matrix $X = \mathrm{diag}(x)$ ($x \in \mathbb{R}^{n_1}$) is diagonal, the model (1.4) degenerates to the completely perturbed sparse signal recovery problem

$$\min_{\tilde{z}\in\mathbb{R}^{n_1}} \|\tilde{z}\|_1 \text{ s.t. } \|\hat{A}\tilde{z} - \hat{y}\|_2 \le \epsilon'_{A,r,y}, \tag{1.5}$$

where $\|\tilde{z}\|_1$ is the l_1-norm of the vector $\hat{z}$, viz, the sum of the absolute value of its elements, and $\hat{A} = A + E \in \mathbb{R}^{m\times n_1}$ is a sensing matrix with E denoting perturbations to the matrix A. In [9], Herman and Strohmer gave a sufficient condition to ensure the stable and robust reconstruction of sparse signals, and at the same moment they provided an upper bound estimation of error. Later, Zhang et al. [10] generalized Herman and Strohmer' result to the block-sparse setting. In particular, they not only enhanced the condition associating with block-restricted isometry property, but also bettered the upper bound estimation of error.

In [11], Candès and Plan first proposed the concept of restricted isometry constant (for short as RIC) of a measurement operator, which is given by as follows:

Definition 1.1. *(RIP for operator [11]) We say that a measurement operator $\mathcal{A}$ satisfies the restricted isometry property with constant δ_r if δ_r is the smallest value $\delta \in (0,1)$ such that*

$$(1-\delta)\|X\|_F^2 \le \|\mathcal{A}(X)\|_2^2 \le (1+\delta)\|X\|_F^2 \tag{1.6}$$

holds for all matrices X of rank at most r (wrote as r-rank), where $\|X\|_F := \sqrt{\langle X, X\rangle} = \sqrt{tr(X^\top X)}$ is the Frobenius norm of the matrix X.

The matrix version of the restricted isometry property (abbreviated as RIP) is an important tool in theoretical analysis of LMR. There exist a few of sufficient conditions based on RIP for exact LMR (i.e., in (1.4),

$z = 0$ and $\mathcal{E} = 0$) or noisy/partially perturbed LMR (i.e., in (1.4), $\mathcal{E} = 0$). These comprise $\delta_{4r} < 0.558$, and $\delta_{3r} < 0.4721$ [12], $\delta_{2k} < 0.4931$ [13], $\delta_k < 1/3$ [14], $\delta_k + \theta\delta_{2r+k} < \theta - 1$ with $\theta > 1$ and $k = \lceil 2r\theta^{2/(2-p)} \rceil$ $(0 < p \leq 1)$ [15], $\sqrt{2}\delta_{\max\{r+\lceil 3k/2 \rceil, 2k\}}(k/(2r))^{1/p-1/2}\delta_{2r+k} < (k/(2r))^{1/p-1/2}$ [16], $\delta_{4r} \in [\sqrt{3}/2, 1)$ [17], and $\delta_{tr} < \sqrt{(t-1)/t}$ for fixed $t > 1$ consisting of constrained NNM [18] and regularized NNM [19].

Moreover, another critical tool for the analysis of low-rank recovery is the Frobenius-robust rank null space property (FRNSP) of the measurement operator $\mathcal{A}$, which was introduced by Kabanava et al. [21], extending the sparse vector recovery to the situation of low-rank matrices. The concept states as follows:

Definition 1.2. *(FRNSP for operator [21]) The measurement operator $\mathcal{A}$ is said to satisfy the Frobenius-robust rank null space property of order r with constants $0 < \rho < 1$ and $\tau > 0$ if for all $X \in \mathbb{R}^{n_1 \times n_2}$, the singular values of X fulfills*

$$\|X_{[r]^c}\|_F \leq \frac{\rho}{\sqrt{r}}\|X_{[r]}\|_* + \tau\|\mathcal{A}(X)\|_2. \tag{1.7}$$

However, the aforementioned works are taken into account only in unperturbed scenario (i.e., $\mathcal{E} = 0$), viz, the measurement operator $\mathcal{A}$ is not perturbed by $\mathcal{E}$. From the viewpoint of application, it is more practical to investigate the low-rank matrix recovery in the completely perturbed scenario. Specifically, we want to realize what a requirement to assure robust reconstruction of low-rank matrices, as well as what is the estimation of reconstruction error. In this paper, we make an unflagging effort to extend the constrained nuclear norm minimization to make sure the capability of the completely perturbed model. Based on RIP condition for $\hat{\mathcal{A}}$, the present paper will show the performance of low-rank matrices recovery via the constrained nuclear norm minimization in the completely perturbed model. For better understanding, the main contributions of this work are as follows. First of all, we establish a sufficient condition to ensure the robust and stable reconstruction of low-rank matrices via the completely perturbed nuclear norm minimization. The derived condition extends and improves the previous works concerning complete perturbation, which is proved sharp by [13]. Secondly, the approximation accuracy between the solution and the original matrix is described by a best r-rank approximation error and a total noise, which gives a theoretical support to refine the recovery precision. The result reveals that stableness and robustness regarding the reconstruction of low-rank matrices in the presence of total noise. Finally, discuss another sufficient condition ensuring reconstruction via certain properties of null space of the measurement operator.

The reminder of the paper is constructed as follows. In Section 2, we provide some notations and our main results. We present some necessary lemmas and the proof of main results in Section 3. Finally, conclusion are given in 4.

2 Notations and the main results

Before introducing our main results, we present some symbols similar to [9], which quantify the perturbations $\mathcal{E}$ and z with different upper bounds as follows:

$$\frac{\|\mathcal{E}\|_{op}}{\|\mathcal{A}\|_{op}} \leq \epsilon_{\mathcal{A}}, \ \frac{\|\mathcal{E}\|_{op}^{(r)}}{\|\mathcal{A}\|_{op}^{(r)}} \leq \epsilon_{\mathcal{A}}^{(r)}, \ \frac{\|z\|_2}{\|y\|_2} \leq \epsilon_y, \tag{2.8}$$

where $\|\mathcal{A}\|_{op}$ is the operator norm of the measurement operator $\mathcal{A}$, i.e., $\|\mathcal{A}\|_{op} = \sup\{\|\mathcal{A}(X)\|_2/\|X\|_F : X \in \mathbb{R}^{n_1 \times n_2}$ with $X \neq 0\}$, and $\|\mathcal{A}\|_{op}^{(r)}$ is the operator norm of $\mathcal{A}$ whose original image set is defined in matrix space composed of r-rank nonzero matrices, i.e., $\|\mathcal{A}\|_{op}^{(r)} = \sup\{\|\mathcal{A}(X)\|_2/\|X\|_F : X \in \mathbb{R}^{n_1 \times n_2}$ with $X \neq 0$ and being r-rank$\}$, and meanwhile we represent

$$t_r = \frac{\|X_{[r]^c}\|_F}{\|X_{[r]}\|_F}, \ s_r = \frac{\|X_{[r]^c}\|_*}{\sqrt{r}\|X_{[r]}\|_F}, \ \kappa_{\mathcal{A}}^{(r)} = \frac{\sqrt{1+\delta_r}}{\sqrt{1-\delta_r}}, \ \alpha_{\mathcal{A}} = \frac{\|\mathcal{A}\|_{op}}{\sqrt{1-\delta_r}}, \tag{2.9}$$

where $X_{[r]^c} = X - X_{[r]}$ with $X_{[r]}$ being the best r-rank approximation of the matrix X, and its singular values are made up of r-largest singular values of the matrix X.

We additionally suppose that $n_1 \leq n_2$ and we denote the singular value decomposition (SVD) of $X \in \mathbb{R}^{n_1 \times n_2}$ by $X = \sum_{i=1}^{n_1} \sigma_i(X) u_i(X)(v_i(X))^\top$, where $u_i(X)$ and $v_i(X)$ are respectively the left and right singular value vectors of X. Without loss of generality, we assume that $\sigma_1(X) \geq \sigma_2(X) \geq \cdots \geq \sigma_{n_1}(X)$. For any positive integer s, we represent $[s] = \{1, 2, \cdots, s\}$, and B^c stands for the complement set of B in $[n_1]$, i.e., $B^c = [n_1] \backslash B$ for any $B \subset [n_1]$. For an index set $B \subset [n_1]$ and a vector $x \in \mathbb{R}^{n_1}$, we represent x_B to be the vector that is equal to x on B, and zero otherwise, and similarly $X_B = \sum_{i \in B} \sigma_i(X) u_i(X)(v_i(X))^\top$ and $X_{[s]} = \sum_{i=1}^{s} \sigma_i(X) u_i(X)(v_i(X))^\top$. Using the notations and symbols above, we provide the main results for reconstruction of low-rank matrices via the completely perturbed nuclear norm minimization as follows:

Theorem 2.1. *For given relative perturbations $\epsilon_{\mathcal{A}}$, $\epsilon_{\mathcal{A}}^{(r)}$, $\epsilon_{\mathcal{A}}^{(2r)}$, and ϵ_y in (2.8), suppose that the RIC of the measurement operator $\mathcal{A}$ satisfies*

$$\delta_{2r} < \frac{\frac{\sqrt{2}}{2} + 1}{(1 + \epsilon_{\mathcal{A}}^{(2r)})^2} - 1, \tag{2.10}$$

a general r-rank matrix X fulfills

$$t_r + s_r < \frac{1}{\kappa_{\mathcal{A}}^{(r)}}, \tag{2.11}$$

and the total noise is

$$\epsilon'_{\mathcal{A},r,y} = \left[\frac{\epsilon_{\mathcal{A}}^{(r)} \kappa_{\mathcal{A}}^{(r)} + \epsilon_{\mathcal{A}} \alpha_{\mathcal{A}} t_r}{1 - \kappa_{\mathcal{A}}^{(r)}(t_r + s_r)} + \epsilon_y \right] \|y\|_2. \tag{2.12}$$

Then the solution X^* of the completely perturbed nuclear norm minimization (1.4) obeys

$$\|X - X^*\|_F \leq C\epsilon'_{\mathcal{A},r,y} + D\frac{\|X_{[s]^c}\|_*}{\sqrt{r}}, \tag{2.13}$$

where

$$C = \frac{2\sqrt{1+\delta_{2r}}(1 + \epsilon_{\mathcal{A}}^{(2r)})}{\frac{\sqrt{2}}{2} + 1 - (1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2},$$

$$D=\frac{\frac{\sqrt{2}}{2}[(1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2-1]+\sqrt{2[\frac{\sqrt{2}}{2}-(1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2+1][(1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2-1]+\frac{1}{2}}}{\frac{\sqrt{2}}{2}+1-(1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2}.$$

Remark 2.2. *Theorem 2.1 presents a sufficient condition for the reconstruction of low-rank matrices via constrained nuclear norm minimization. The gained condition mainly involves two aspects: on the one hard, as far as RIC is concerned, (2.10) characterizes the reconstruction condition for low-rank matrices, and simultaneously improves the results in [9] [10] because* $n_1 = n_2$ *and the matrix* $X = diag(x)$ $(x \in \mathbb{R}^{n_1})$ *is diagonal, our results reduce to the conventional compressed sensing. Specifically, the upper bound of RIC* $\delta_{2r} < \frac{\frac{\sqrt{2}}{2}+1}{(1+\epsilon_{\mathcal{A}}^{(2r)})^2}-1$ *is weaker than* $\delta_{2r} < \frac{\sqrt{2}}{(1+\epsilon_{\mathcal{A}}^{(2r)})^2}-1$ *in [9] and* $\delta_{2r} < \frac{1.5}{(1+\epsilon_{\mathcal{A}}^{(2r)})^2}-1$ *in [10]. Factually, in the case of* $\mathcal{E}=0$ *(i.e.,* $\epsilon_{\mathcal{A}}^{(2r)}=0$*), which means that the measurement operator* $\mathcal{A}$ *is not perturbed by* $\mathcal{E}$*, it has been proved [13] that the bound of RIC* $\delta_{2r} < \sqrt{2}/2$ *is sharp, namely, it is not possible to improve that bound. On the other hard, with respect to the restriction to matrices to be recovered, one can easily check that (2.11) is able to be satisfied when the unknown matrices are low-rank.*

Remark 2.3. *The inequality (2.13) shows that the upper bound estimation of recovery error is controlled by the total noise* $\epsilon'_{\mathcal{A},r,y}$ *and the best* r*-rank approximation error. When* $n_1 = n_2$ *and the matrix* $X = diag(x)$ $(x \in \mathbb{R}^{n_1})$ *is diagonal, our results not only incorporate portion of results in [9] [10] [13], but also acquire a tighter upper bound estimation of reconstruction error than that in [9] [10]. In particular, the error bound noise constant* C *and the error bound compressibility constant* D *are less than* C_1 *and* C_0 *in [9],* D_1 *and* D_2 *in [10], respectively, i.e.,*

$$C < C_1 = \frac{4\sqrt{1+\delta_{2r}}(1+\epsilon_{\mathcal{A}}^{(2r)})}{1-(\sqrt{2}+1)[(1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2-1]},$$
$$D < C_0 = \frac{2\{1+(\sqrt{2}-1)[(1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2-1]\}}{1-(\sqrt{2}+1)[(1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2-1]}$$

and

$$C < D_1 = \frac{2\sqrt{2}\sqrt{1+\delta_{2r}}(1+\epsilon_{\mathcal{A}}^{(2r)})}{1-2[(1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2-1]},$$
$$D < D_2 = \frac{2\{1-(2-\sqrt{2})[(1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2-1]\}}{1-2[(1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2-1]}.$$

Remark 2.4. *When no noise and perturbation are introduced, our completely perturbed nuclear norm model will return to the general low-rank matrices recovery, and will bring about exact reconstruction when matrices are* r*-rank. In this case, our results contain part of the Theorem* 1.4 *in [13].*

In the following, we extend the result of Theorem 2.1 to the general case, whose proof is similar to that of Theorem 2.1, stated below.

Corollary 2.5. *For given relative perturbations* $\epsilon_{\mathcal{A}}$, $\epsilon_{\mathcal{A}}^{(r)}$, $\epsilon_{\mathcal{A}}^{(2r)}$, *and* ϵ_y *in (2.8), we assume that the RIC of the measurement operator* $\mathcal{A}$ *satisfies*

$$\delta_{tr} < \frac{\sqrt{\frac{t-1}{t}}+1}{(1+\epsilon_{\mathcal{A}}^{(tr)})^2}-1, \tag{2.14}$$

for $t \geq 1$ and (2.11) and (2.12) hold. Then the solution X^* of the completely perturbed nuclear norm minimization (1.4) satisfies

$$\|X - X^*\|_F \leq C'\epsilon'_{\mathcal{A},r,y} + D'\frac{\|X_{[s]^c}\|_*}{\sqrt{r}}, \tag{2.15}$$

where

$$C' = \frac{2\sqrt{2}\sqrt{t(t-1)(1+\delta_{tr})}(1+\epsilon_{\mathcal{A}}^{(tr)})}{t[\sqrt{\frac{t-1}{t}}+1-(1+\delta_{tr})(1+\epsilon_{\mathcal{A}}^{(tr)})^2]},$$

$$D' = \left\{t\left[\sqrt{\frac{t-1}{t}}+1-(1+\delta_{tr})(1+\epsilon_{\mathcal{A}}^{(tr)})^2\right]\right\}^{-1} \times \left\{4\sqrt{2}[(1+\delta_{tr})(1+\epsilon_{\mathcal{A}}^{(tr)})^2-1] + \sqrt{2}t\left(\sqrt{\frac{t-1}{t}}-\delta_{tr}\right)\right.$$
$$\left. + 4\sqrt{t\left[\sqrt{\frac{t-1}{t}}-(1+\delta_{tr})(1+\epsilon_{\mathcal{A}}^{(tr)})^2+1\right][(1+\delta_{tr})(1+\epsilon_{\mathcal{A}}^{(tr)})^2-1]}\right\}.$$

Remark 2.6. *When $t = 2$, the results of Corollary 2.5 coincide with Theorem 2.1.*

Remark 2.7. *When $\mathcal{E} = 0$, i.e., $\epsilon_{\mathcal{A}} = 0$, the results of Corollary 2.5 not only are the same as part of Proposition* 3.1 *[18], but also we gain a tighter upper bound estimation of error than that in [18]. Particularly, the error bound compressibility constant D' are less than D'', i.e.,*

$$D' = \frac{\sqrt{2}\delta_{tr} + \sqrt{t(\sqrt{(t-1)/t}-\delta_{tr})\delta_{tr}}}{t(\sqrt{(t-1)/t}-\delta_{tr})} + \frac{\sqrt{2}}{2}$$
$$< \frac{\sqrt{2}\delta_{tr} + \sqrt{t(\sqrt{(t-1)/t}-\delta_{tr})\delta_{tr}}}{t(\sqrt{(t-1)/t}-\delta_{tr})} + 1 = D''.$$

Besides, in the case of $\mathcal{E} = 0$, it has been showed [18] that the condition (2.14) for $t \geq 4/3$ is sharp.

Theorem 2.8. *Let $\epsilon_{\mathcal{A}}$, $\epsilon_{\mathcal{A}}^{(r)}$, $\epsilon_{\mathcal{A}}^{(2r)}$, and ϵ_y in (2.8), assume that the measurement operator $\mathcal{A}$ obeys the Frobenius-robust space property of order r with constants $\sqrt{r}\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op} < \rho < 1$ and $\tau > 0$, and (2.11) and (2.12) hold. Then the minimizer X^* of the model (1.4) fullfils*

$$\|X - X^*\|_F \leq \hat{C}\epsilon'_{\mathcal{A},r,y} + \hat{D}\frac{\|X_{[s]^c}\|_*}{\sqrt{r}}, \tag{2.16}$$

where

$$\hat{C} = \frac{2\tau}{1+\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}\left[3+\rho-\frac{\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\rho+\sqrt{r})}{1+\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}\right]\cdot\left[1-\rho+\frac{\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\rho+\sqrt{r})}{1+\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}\right]^{-1},$$
$$\hat{D} = 2\left[1+\rho-\frac{\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\rho+\sqrt{r})}{1+\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}\right]^2\cdot\left[1-\rho+\frac{\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\rho+\sqrt{r})}{1+\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}\right]^{-1}.$$

Remark 2.9. *The theorem demonstrates the stability and robustness of (1.4). The inequality (2.16) indicates the estimation of the reconstruction error is bounded by two terms, one about total noise, and another regarding best r-rank approximation error.*

Remark 2.10. *In the case of $\mathcal{E} = 0$, i.e., $\epsilon_{\mathcal{A}} = 0$, we obtain the same results as Theorem* 3.1 *in [21].*

3 The proof of main results

In this section, we will prove our main results. To establish the results, the following lemmas will be useful. First, we provide the lemma below comprising an inequality concerning δ_r and $\hat{\delta}_r$.

Lemma 3.1. *(RIP for perturbation of measurement operator $\hat{\mathcal{A}}$) Suppose that the RIC of order r for the measurement operator $\mathcal{A}$ is δ_r, relative perturbation upper bound corresponding with the measurement operator $\mathcal{E}$ is $\epsilon_{\mathcal{A}}^{(r)}$ and fix the constant $\hat{\delta}_{r,\max} = (1+\delta_r)(1+\epsilon_{\mathcal{A}}^{(r)})^2 - 1$. Then the RIC $\hat{\delta}_r \leq \hat{\delta}_{r,\max}$ for $\hat{\mathcal{A}} = \mathcal{A} + \mathcal{E}$ is the smallest nonnegative constant such that*

$$(1-\hat{\delta}_r)\|X\|_F^2 \leq \|\hat{\mathcal{A}}(X)\|_2^2 \leq (1+\hat{\delta}_r)\|X\|_F^2 \tag{3.17}$$

holds for all r-rank matrices $X \in \mathbb{R}^{n_1 \times n_2}$.

Proof of the lemma 3.1. The proof of lemma is inspired by [9]. First of all, define l_r and u_r as the smallest nonnegative numbers such that

$$(1-l_r)\|X\|_F^2 \leq \|\hat{\mathcal{A}}(X)\|_2^2 \leq (1+u_r)\|X\|_F^2 \tag{3.18}$$

holds for all matrices X rank at most r. By applying the triangle inequality, the concept of RIC and (2.8), we get

$$\begin{aligned}
\|\hat{\mathcal{A}}(X)\|_2^2 &\leq (\|\mathcal{A}(X)\|_2 + \|\mathcal{E}(X)\|_2)^2 \\
&\leq (\sqrt{1+\delta_r} + \|\mathcal{E}\|_{op}^{(r)})^2\|X\|_F^2 \\
&\leq (\sqrt{1+\delta_r} + \epsilon_{\mathcal{A}}^{(r)}\|\mathcal{A}\|_{op}^{(r)})^2\|X\|_F^2 \\
&\overset{(a)}{\leq} (\sqrt{1+\delta_r} + \epsilon_{\mathcal{A}}^{(r)}\sqrt{1+\delta_r})^2\|X\|_F^2 \\
&= (1+\delta_r)(1+\epsilon_{\mathcal{A}}^{(r)})^2\|X\|_F^2,
\end{aligned} \tag{3.19}$$

where (a) follows from the fact that $\|\mathcal{A}\|_{op}^{(r)} \leq \sqrt{1+\delta_r}$. Due to the notion of u_r, it implies that

$$1 + u_r \leq (1+\delta_r)(1+\epsilon_{\mathcal{A}}^{(r)})^2. \tag{3.20}$$

By employing the inequality above, we get an minimal upper bound

$$u_r = (1+\delta_r)(1+\epsilon_{\mathcal{A}}^{(r)})^2 - 1, \tag{3.21}$$

fulfilling the concept of u_r. Likewise, combining with the reverse triangular inequality, the concept of RIC and (2.8), we obtain

$$l_r = 1 - (1-\delta_r)(1-\epsilon_{\mathcal{A}}^{(r)})^2, \tag{3.22}$$

meeting the notion of l_r. Observe that $1 - u_r \leq 1 - l_r$ and $1 + l_r \leq 1 + u_r$. Accordingly, for given δ_r and $\epsilon_{\mathcal{A}}^{(r)}$, we select $\hat{\delta}_{r,\max} = u_r$ as the smallest nonnegative constant making (3.18) symmetric. Obviously, the real RIC $\hat{\delta}_r$ for $\hat{\mathcal{A}}$ satisfies $\hat{\delta}_r \leq \hat{\delta}_{r,\max}$. The proof is complete. □

Lemma 3.2. *([20]) Suppose that the measurement operator $\mathcal{A}$ fulfills the upper bound of the RIP in (1.6). Then, for every matrix X, we obtain*

$$\|\mathcal{A}(X)\|_2 \leq \sqrt{1+\delta_r}\left(\|X\|_F + \frac{\|X\|_*}{\sqrt{r}}\right). \tag{3.23}$$

Now, we present a sufficient condition for the lower bound of the image of an arbitrary matrix related to the best r-rank approximation and truncation of X and the RIC of $\mathcal{A}$.

Lemma 3.3. *Suppose that condition (2.11) in Theorem 2.1. Then, for general matrix X, its image under $\mathcal{A}$ can be controlled by the following great than 0 quantity*

$$\|\mathcal{A}(X)\|_2 \geq \sqrt{1-\delta_r}\left[\|X_{[r]}\|_F - \kappa_{\mathcal{A}}^{(r)}\left(\|X_{[r]^c}\|_F + \frac{\|X_{[r]^c}\|_*}{\sqrt{r}}\right)\right].$$

Proof of the lemma 3.3. By applying Lemma 3.2 and (2.11), we get

$$\begin{aligned}
\|\mathcal{A}(X)\|_2 &\geq \|\mathcal{A}(X_{[r]})\|_2 - \|\mathcal{A}(X_{[r]^c})\|_2 \\
&\geq \sqrt{1-\delta_r}\|X_{[r]}\|_F - \sqrt{1+\delta_r}\left(\|X_{[r]^c}\|_F + \frac{\|X_{[r]^c}\|_*}{\sqrt{r}}\right) \\
&= \sqrt{1-\delta_r}\left[1 - \frac{\sqrt{1+\delta_r}}{\sqrt{1-\delta_r}}\left(\frac{\|X_{[r]^c}\|_F}{\|X_{[r]}\|_F} + \frac{\|X_{[r]^c}\|_*}{\sqrt{r}\|X_{[r]}\|_F}\right)\right]\|X_{[r]}\|_F \\
&= \sqrt{1-\delta_r}[1 - \kappa_{\mathcal{A}}^{(r)}(t_r + s_r)]\|X_{[r]}\|_F \\
&> 0.
\end{aligned}$$

□

The following lemma gives an upper bound of the size of the total perturbation, which is induced by $\mathcal{E}$ and z.

Lemma 3.4. *Suppose that (2.11) in Theorem 2.1 and denote*

$$\epsilon'_{\mathcal{A},r,y} = \left[\frac{\epsilon_{\mathcal{A}}^{(r)}\kappa_{\mathcal{A}}^{(r)} + \epsilon_{\mathcal{A}}\alpha_{\mathcal{A}}t_r}{1 - \kappa_{\mathcal{A}}^{(r)}(t_r + s_r)} + \epsilon_y\right]\|y\|_2,$$

where $\epsilon_{\mathcal{A}}, \epsilon_{\mathcal{A}}^{(r)}, \epsilon_y$ are determined by (2.8), and $\kappa_{\mathcal{A}}^{(r)}, \alpha_{\mathcal{A}}, t_r, s_r$ in (2.9). Then the total perturbation satisfies

$$\|\mathcal{E}(X)\|_2 + \|z\|_2 \leq \epsilon'_{\mathcal{A},r,y}. \tag{3.24}$$

Proof of the lemma 3.4. By making use of Lemma 3.3, we get

$$\begin{aligned}
\frac{\|\mathcal{E}(X)\|_2}{\|\mathcal{A}(X)\|_2} &\leq \frac{\|\mathcal{E}(X_{[r]})\|_2 + \|\mathcal{E}(X_{[r]^c})\|_2}{\|\mathcal{A}(X)\|_2} \\
&\leq \frac{\|\mathcal{E}\|_{op}^{(r)}\|X_{[r]}\|_F + \|\mathcal{E}\|_{op}\|X_{[r]^c}\|_F}{\sqrt{1-\delta_r}\left[\|X_{[r]}\|_F - \kappa_{\mathcal{A}}^{(r)}\left(\|X_{[r]^c}\|_F + \frac{\|X_{[r]^c}\|_*}{\sqrt{r}}\right)\right]} \\
&\leq \frac{(\epsilon_{\mathcal{A}}^{(r)}\|\mathcal{A}\|_{op}^{(r)} + \epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}t_r)}{\sqrt{1-\delta_r}[1 - \kappa_{\mathcal{A}}^{(r)}(t_r + s_r)]} \\
&\leq \frac{\epsilon_{\mathcal{A}}^{(r)}\kappa_{\mathcal{A}}^{(r)} + \epsilon_{\mathcal{A}}\alpha_{\mathcal{A}}t_r}{1 - \kappa_{\mathcal{A}}^{(r)}(t_r + s_r)}.
\end{aligned} \tag{3.25}$$

Combining with (3.25) and the inequality $\|z\|_2 \leq \epsilon_y\|y\|_2$, the desired result follows. □

Cai and Zhang [18] developed a novel technique which plays a crucial role in the proof of main results. This result shows that any point in a polytope can be expressed as a convex combination of sparse vectors.

Lemma 3.5. *([18]) For a positive number α and a positive integer k, the polytope $T(\alpha, k) \subset \mathbb{R}^n$ is defined by*

$$T(\alpha, k) = \{v \in \mathbb{R}^n : \ \|v\|_\infty \leq \alpha, \|v\|_1 \leq k\alpha\},$$

where $\|v\|_\infty = \max_{1 \leq i \leq n} |v_i|$. For arbitrary $v \in \mathbb{R}^n$, the set $U(\alpha, k, v) \in \mathbb{R}^n$ is defined by

$$U(\alpha, k, v) = \{u \in \mathbb{R}^n : \ supp(u) \subset supp(v), \|u\|_0 \leq k, \|u\|_1 = \|v\|_1, \|u\|_\infty \leq \alpha\},$$

where $supp(u)$ is the support of u, i.e., $supp(u) = \{i : \ u_i \neq 0\}$, and $\|u\|_0 = |supp(u)|$. Then $v \in T(\alpha, k)$ iff v is in the convex hull of $U(\alpha, k, v)$. In particular, any $v \in T(\alpha, k)$ can be represented as

$$v = \sum_{i=1}^{L} \gamma_i u_i,$$

where $u_i \in U(\alpha, k, v)$ and $0 \leq \gamma_i \leq 1$, $\sum_{i=1}^{L} \gamma_i = 1$.

Lemma 3.6. *([16]) (a) Let $M, N \in \mathbb{R}^{n_1 \times n_2}$ be matrices satisfying $M^\top N = 0$ and $MN^\top = 0$. Then*

$$\|M + N\|_F^2 = \|M\|_F^2 + \|N\|_F^2. \tag{3.26}$$

(b) Let $M, N \in \mathbb{R}^{n_1 \times n_2}$ be matrices with the row and column spaces of M and N being orthogonal. Then (3.26) holds.

Lemma 3.7. *For given $\epsilon_{\mathcal{A}}$ as (2.8), let the measurement operator $\mathcal{A}$ satisfy the Frobenius-robust rank null space property with constants $\sqrt{r}\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op} < \rho < 1$ and $\tau > 0$ if for all $X \in \mathbb{R}^{n_1 \times n_2}$, the singular values of X meet (1.7) and fix constants $\hat{\tau} = \tau/(1 + \tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op})$ and $\hat{\rho} = \rho - \frac{\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\rho+\sqrt{r})}{1+\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}$. Then the measurement operator $\hat{\mathcal{A}}$ obeys the Frobenius-robust rank null space property of order r with constants $0 < \hat{\rho} < 1$ and $\hat{\tau} > 0$.*

Remark 3.8. *When $\mathcal{E} = 0$, i.e., $\epsilon_{\mathcal{A}} = 0$, signifying that there exists no perturbation in the measurement operator $\mathcal{A}$, then $\hat{\tau} = \tau$ and $\hat{\rho} = \rho$. In this case, the Definition 3.1 [21] is incorporated in our lemma, namely, we derive the same result as Definition 1.2.*

Proof of the lemma 3.7. By applying the triangle inequality and (2.8), we get

$$\begin{aligned} \|\hat{\mathcal{A}}(X)\|_2 &\leq \|\mathcal{A}(X)\|_2 + \|\mathcal{E}(X)\|_2 \\ &\leq \|\mathcal{A}(X)\|_2 + \|\mathcal{E}\|_{op}\|X\|_F \\ &\leq \|\mathcal{A}(X)\|_2 + \epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}\|X\|_F, \end{aligned} \tag{3.27}$$

Applying the Frobenius-robust rank null space property of $\hat{\mathcal{A}}$ and noting the fact that $\|X\|_F \leq \|X\|_*$ for any X and combining with (3.27),

$$\|X_{[r]}\|_F \leq \frac{\hat{\rho}}{\sqrt{r}}\|X\|_* + \hat{\tau}\|\mathcal{A}(X)\|_2 + \hat{\tau}\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}\|X\|_F$$

$$\leq \frac{\hat{\rho}}{\sqrt{r}}\|X\|_* + \hat{\tau}\|\mathcal{A}(X)\|_2 + \hat{\tau}\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\|X_{[r]}\|_F + \|X_{[r]^c}\|_F)$$
$$\leq \frac{\hat{\rho}}{\sqrt{r}}\|X\|_* + \hat{\tau}\|\mathcal{A}(X)\|_2 + \hat{\tau}\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\|X_{[r]}\|_F + \|X_{[r]^c}\|_*). \tag{3.28}$$

Observing the condition of $\hat{\tau} = \tau/(1 + \tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op})$ with $\tau > 0$, by rearranging the terms in (3.28), we get

$$\|X_{[r]}\|_F \leq \frac{1}{1 - \hat{\tau}\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}\left(\frac{\hat{\rho}}{\sqrt{r}} + \hat{\tau}\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}\right)\|X_{[r]^c}\|_* + \frac{\hat{\tau}}{1 - \hat{\tau}\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}\|\mathcal{A}(X)\|_2. \tag{3.29}$$

Hence, $\hat{\tau} = \tau/(1 + \tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op})$ and $\hat{\rho} = \rho - \frac{\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\rho+\sqrt{r})}{1+\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}$. For given $\epsilon_{\mathcal{A}}$ as (2.8), $\|\mathcal{A}\|_{op}$ and $\tau > 0$, to make $\hat{\rho} \in (0, 1)$, we solve the inequality below

$$0 < \rho - \frac{\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\rho + \sqrt{r})}{1 + \tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}} < 1,$$

which implies

$$\sqrt{r}\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op} < \rho < 1 + (\sqrt{r} + 1)\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}.$$

Combining with the inequality above and $\rho \in (0, 1)$, the lemma is proved. □

The following lemma presents a matrix version of Stechkin's bound which extends the result of sparse vectors [22] to that case of low-rank matrices.

Lemma 3.9. *([21]) Let $X \in \mathbb{R}^{n_1 \times n_2}$ and $r \leq n_1$. Then, for $p > 0$,*

$$\|X_{[r]^c}\|_p \leq \frac{\|X\|_*}{r^{1-1/p}}.$$

Lemma 3.10. *([23]) For any $X, Y \in \mathbb{R}^{n_1 \times n_2}$, we have*

$$\sum_{i=1}^{n_1} |\sigma_i(X) - \sigma_i(Y)|_* \leq \|X - Y\|_*.$$

With preparations above, we now prove our main results.

Proof of the theorem 2.1. Firstly, we reveal the estimation (2.13). Let $X^* = X - Z$, where X is the matrix that we wish to recover, and X^* is the minimizer of (1.4). We exploit the following inequality which has been showed by Mohan and Fazel (see (7) in [12]).

$$\|Z_{[r]^c}\|_* \leq \|Z_{[r]}\|_* + 2\|X_{[r]^c}\|_*. \tag{3.30}$$

Denote $\beta = (\|Z_{[r]}\|_* + 2\|X_{[r]^c}\|_*)/r$, and

$$D_1 = \{i \in [r]^c : \ \sigma_i(Z) > \beta\},$$
$$D_2 = \{i \in [r]^c : \ \sigma_i(Z) \leq \beta\}.$$

Hence, $Z_{[r]^c}$ is partitioned into two parts, i.e.,

$$Z_{[r]^c} = Z_{D_1} + Z_{D_2}.$$

Due to the definitions above and (3.30), we get

$$\|Z_{D_1}\|_* \leq \|Z_{[r]^c}\|_* \leq \beta r.$$

Set $\text{rank}(Z_{D_1}) = |D_1| = n$. In view of the definition of D_1, it leads to

$$\beta r \geq \|Z_{D_1}\|_* = \sum_{i\in D_1} \sigma_i \geq \sum_{i\in D_1} \beta = n\beta.$$

That is, $n \leq r$. Furthermore, we get $\text{rank}(Z_{[r]} + Z_{D_1}) = r + n \leq 2r$,

$$\begin{aligned} \|\sigma_{D_2}(Z)\|_1 &= \|Z_{D_2}\|_* = \|Z_{[r]^c}\|_* - \|Z_{D_1}\|_* \\ &\leq \beta r - n\beta = (r-n)\beta, \\ \|\sigma_{D_2}(Z)\|_\infty &= \|Z_{D_2}\|_{op} \leq \beta. \end{aligned} \tag{3.31}$$

From the triangular inequality and $\|\hat{\mathcal{A}}(X) - \hat{y}\|_2 \leq \epsilon'_{\mathcal{A},r,y}$, it implies that

$$\|\hat{\mathcal{A}}(Z)\|_2 \leq \|\hat{\mathcal{A}}(X) - \hat{y}\|_2 + \|\hat{\mathcal{A}}(X^*) - \hat{y}\|_2 \leq 2\epsilon'_{\mathcal{A},r,y}, \tag{3.32}$$

which deduces

$$\begin{aligned} \left\langle \hat{\mathcal{A}}(Z_{[r]} + Z_{D_1}), \hat{\mathcal{A}}(Z) \right\rangle &\overset{(a)}{\leq} \|\hat{\mathcal{A}}(Z_{[r]} + Z_{D_1})\|_2 \|\hat{\mathcal{A}}(Z)\|_2 \\ &\overset{(b)}{\leq} 2\epsilon'_{\mathcal{A},r,y}\sqrt{1+\hat{\delta}_{2r}}\|Z_{[r]} + Z_{D_1}\|_F, \end{aligned} \tag{3.33}$$

where (a) follows from Hölder's inequality, and (b) is due to the definition of RIC for $\hat{\mathcal{A}}$ and (3.32). By employing Lemma 3.5 and (3.31), we can represent $\sigma_{D_2}(Z)$ as $\sigma_{D_2}(Z) = \sum_{i=1}^{L} \gamma_i w_i$, where $w_i \in U(\beta, r-n, \sigma_{D_2}(Z))$. By further defining

$$\Phi_i = \sum_{j=1}^{n_1} w_i[j] u_j(Z)(v_j(Z))^\top, \ i = 1, 2, \cdots, L,$$

where $w_i[j]$ denotes the jth element of w_i, we can express Z_{D_2} as

$$Z_{D_2} = \sum_{i=1}^{L} \gamma_i \Phi_i, \tag{3.34}$$

and one can easily see that Φ_i is $(r-n)$-rank and $\|\Phi_i\|_{op} = \|w_i\|_\infty \leq \beta$. Therefore,

$$\|\Phi_i\|_F \leq \sqrt{rank(\Phi_i)}\|\Phi_i\|_{op} \leq \sqrt{r-n}\|\Phi_i\|_{op} \leq \sqrt{r}\beta. \tag{3.35}$$

Define $\Psi_i = Z_{[r]} + Z_{D_1} + \lambda\Phi_i$, where $0 \leq \lambda \leq 1$. It is easy to see that

$$\begin{aligned} \sum_{j=1}^{L} \gamma_j \Psi_j - \frac{1}{2}\Psi_i &= Z_{[r]} + Z_{D_1} + \lambda Z_{D_2} - \frac{1}{2}\Psi_i \\ &= \left(\frac{1}{2} - \lambda\right)(Z_{[r]} + Z_{D_1}) - \frac{1}{2}\lambda\Phi_i + \lambda Z. \end{aligned} \tag{3.36}$$

Since $Z_{[r]}, Z_{D_1}, \Phi_i$ are r-rank, n-rank, $(r-n)$-rank, respectively, Ψ_i and $\sum_{j=1}^{L}\gamma_j\Psi_j - \frac{1}{2}\Psi_i - \lambda Z = (\frac{1}{2}-\lambda)(Z_{[r]} + Z_{D_1}) - \frac{1}{2}\lambda\Phi_i$ are all $2r$-rank matrices. Put

$$\Delta = \|Z_{[r]} + Z_{D_1}\|_F,\ Q = \frac{2\|X_{[r]^c}\|_*}{\sqrt{r}}.$$

By utilizing Hölder's inequality, we get

$$\begin{aligned}
\|\Phi_i\|_F &\leq \sqrt{r}\beta = \sqrt{r}\frac{\|Z_{[r]}\|_* + 2\|X_{[r]^c}\|_*}{r} \\
&\leq \|Z_{[r]}\|_F + \frac{2\|X_{[r]^c}\|_*}{\sqrt{r}} \\
&\leq \|Z_{[r]} + Z_{D_1}\|_F + \frac{2\|X_{[r]^c}\|_*}{\sqrt{r}} \\
&= \Delta + Q. \qquad (3.37)
\end{aligned}$$

We make use of the identity below (see (25) in [18]).

$$\sum_{i=1}^{L}\gamma_i\left\|\hat{\mathcal{A}}\left(\sum_{j=1}^{L}\gamma_j\Psi_j - \frac{1}{2}\Psi_i\right)\right\|_2^2 = \sum_{i=1}^{L}\frac{\gamma_i}{4}\|\hat{\mathcal{A}}(\Psi_i)\|_2^2. \qquad (3.38)$$

A combination of (3.33), (3.36), the definition of RIP for $\hat{\mathcal{A}}$ and Lemma 3.6, we can reckon the left side of (3.38)

$$\begin{aligned}
&\sum_{i=1}^{L}\gamma_i\|\hat{\mathcal{A}}(\sum_{j=1}^{L}\gamma_j\Psi_j - \frac{1}{2}\Psi_i)\|_2^2 \\
&= \sum_{i=1}^{L}\gamma_i\|\hat{\mathcal{A}}((\frac{1}{2}-\lambda)(Z_{[r]} + Z_{D_1}) - \frac{1}{2}\lambda\Phi_i + \lambda Z)\|_2^2 \\
&= \sum_{i=1}^{L}\gamma_i\|\hat{\mathcal{A}}((\frac{1}{2}-\lambda)(Z_{[r]} + Z_{D_1}) - \frac{1}{2}\lambda\Phi_i)\|_2^2 \\
&+ \lambda(1-\lambda)\left\langle\hat{\mathcal{A}}(Z_{[r]} + Z_{D_1}), \hat{\mathcal{A}}(Z)\right\rangle \\
&\leq (1+\hat{\delta}_{2r})\sum_{i=1}^{L}\gamma_i\|(\frac{1}{2}-\lambda)(Z_{[r]} + Z_{D_1}) - \frac{1}{2}\lambda\Phi_i\|_F^2 \\
&+ \lambda(1-\lambda)\sqrt{1+\hat{\delta}_{2r}}\|Z_{[r]} + Z_{D_1})\|_F \cdot (2\epsilon'_{\mathcal{A},r,y}) \\
&= (1+\hat{\delta}_{2r})\sum_{i=1}^{L}\gamma_i\left[(\frac{1}{2}-\lambda)^2\|Z_{[r]} + Z_{D_1}\|_F^2 + \frac{1}{4}\lambda^2\|\Phi_i\|_F^2\right] \\
&+ \lambda(1-\lambda)\sqrt{1+\hat{\delta}_{2r}}\|Z_{[r]} + Z_{D_1}\|_F \cdot (2\epsilon'_{\mathcal{A},r,y}). \qquad (3.39)
\end{aligned}$$

On the other side, due to the definition of Ψ_i, we derive

$$\begin{aligned}
\sum_{i=1}^{L}\frac{\gamma_i}{4}\|\hat{\mathcal{A}}(\Psi_i)\|_2^2 &= \sum_{i=1}^{L}\frac{\gamma_i}{4}\|\hat{\mathcal{A}}(Z_{[r]} + Z_{D_1} + \lambda\Phi_i)\|_2^2 \\
&\overset{(a)}{\leq} \sum_{i=1}^{L}\frac{\gamma_i}{4}(1-\hat{\delta}_{2r})\|Z_{[r]} + Z_{D_1} + \lambda\Phi_i\|_F^2 \\
&\overset{(b)}{=} \sum_{i=1}^{L}(1-\hat{\delta}_{2r})\frac{\gamma_i}{4}(\|Z_{[r]} + Z_{D_1}\|_F^2 + \lambda^2\|\Phi_i\|_F^2), \qquad (3.40)
\end{aligned}$$

where (a) is due to the definition of RIP, and (b) follows from Lemma 3.6. Combining with (3.37), (3.39) and (3.40), we get

$$
\begin{aligned}
0 &= \sum_{i=1}^{L} \gamma_i \left\| \hat{\mathcal{A}} \left(\sum_{j=1}^{L} \gamma_j \Psi_j - \frac{1}{2}\Psi_i \right) \right\|_2^2 - \sum_{i=1}^{L} \frac{\gamma_i}{4} \|\hat{\mathcal{A}}(\Psi_i)\|_2^2 \\
&\leq (1+\hat{\delta}_{2r}) \sum_{i=1}^{L} \gamma_i \left[(\frac{1}{2}-\lambda)^2 \|Z_{[r]} + Z_{D_1}\|_F^2 + \frac{1}{4}\lambda^2 \|\Phi_i\|_F^2 \right] \\
&+ \lambda(1-\lambda)\sqrt{1+\hat{\delta}_{2r}} \|Z_{[r]} + Z_{D_1}\|_F \cdot (2\epsilon'_{\mathcal{A},r,y}) \\
&- \sum_{i=1}^{L} (1-\hat{\delta}_{2r}) \frac{\gamma_i}{4} (\|Z_{[r]} + Z_{D_1}\|_F^2 + \lambda^2 \|\Phi_i\|_F^2) \\
&= \sum_{i=1}^{L} \gamma_i \left\{ \left[(1+\hat{\delta}_{2r})(\frac{1}{2}-\lambda)^2 - \frac{1}{4}(1-\hat{\delta}_{2r}) \right] \|Z_{[r]} + Z_{D_1}\|_F^2 + \frac{1}{2}\lambda^2 \hat{\delta}_{2r} \|\Phi_i\|_F^2 \right\} \\
&+ \lambda(1-\lambda)\sqrt{1+\hat{\delta}_{2r}} \|Z_{[r]} + Z_{D_1}\|_F \cdot (2\epsilon'_{\mathcal{A},r,y}) \\
&\leq \left[\lambda(\lambda-1) + (\frac{1}{2} - \lambda + \frac{3}{2}\lambda^2)\hat{\delta}_{2r} \right] \Delta^2 \\
&+ \left[\lambda(1-\lambda)\sqrt{1+\hat{\delta}_{2r}} \cdot (2\epsilon'_{\mathcal{A},r,y}) + \hat{\delta}_{2r}\lambda^2 Q \right] \Delta + \frac{1}{2}\hat{\delta}_{2r}\lambda^2 Q^2.
\end{aligned}
$$

Select $\lambda = \sqrt{2} - 1$, we get

$$
-2\lambda^2 \left(\frac{\sqrt{2}}{2} - \hat{\delta}_{2r} \right) \Delta^2 + \left[\lambda^2 \sqrt{2(1+\hat{\delta}_{2r})} \cdot (2\epsilon'_{\mathcal{A},r,y}) + \hat{\delta}_{2r}\lambda^2 Q \right] \Delta + \frac{1}{2}\hat{\delta}_{2r}\lambda^2 Q^2 \geq 0.
$$

That is,

$$
\lambda^2 \left\{ -2 \left(\frac{\sqrt{2}}{2} - \hat{\delta}_{2r} \right) \Delta^2 + \left[\sqrt{2(1+\hat{\delta}_{2r})} \cdot (2\epsilon'_{\mathcal{A},r,y}) + \hat{\delta}_{2r} Q \right] \Delta + \frac{1}{2}\hat{\delta}_{2r} Q^2 \right\} \geq 0,
$$

which is a second-order inequality for Δ. Consequently, we get

$$
\begin{aligned}
\Delta &\leq \frac{1}{4(\frac{\sqrt{2}}{2} - \hat{\delta}_{2r})} \left\{ \sqrt{2(1+\hat{\delta}_{2r})} \cdot (2\epsilon'_{\mathcal{A},r,y}) + \hat{\delta}_{2r} Q \right. \\
&\left. + \left\{ \left[\sqrt{2(1+\hat{\delta}_{2r})} \cdot (2\epsilon'_{\mathcal{A},r,y}) + \hat{\delta}_{2r} Q \right]^2 + 4(\frac{\sqrt{2}}{2} - \hat{\delta}_{2r})\hat{\delta}_{2r} Q^2 \right\}^{1/2} \right\} \\
&\leq \frac{\sqrt{2(1+\hat{\delta}_{2r})}}{\frac{\sqrt{2}}{2} - \hat{\delta}_{2r}} \epsilon'_{\mathcal{A},r,y} + \frac{\hat{\delta}_{2r} + \sqrt{(\frac{\sqrt{2}}{2} - \hat{\delta}_{2r})\hat{\delta}_{2r}}}{2(\frac{\sqrt{2}}{2} - \hat{\delta}_{2r})} Q.
\end{aligned}
$$

By taking advantage of (3.30) and Hölder's inequality, one can easily verify that

$$
\begin{aligned}
\|Z_{[r]^c}\|_F^2 &\leq \|Z_{[r]^c}\|_{op} \cdot \|Z_{[r]^c}\|_* \\
&\leq \frac{\|Z_{[r]}\|_*}{r} (\|Z_{[r]}\|_* + 2\|X_{[r]^c}\|_*) \\
&\leq \|Z_{[r]}\|_F^2 + \frac{2\|Z_{[r]}\|_F \|X_{[r]^c}\|_*}{\sqrt{r}}.
\end{aligned}
$$

Finally, by the aforementioned inequalities and Lemma 3.1, we obtain

$$
\begin{aligned}
\|Z_{[r]}\|_F &\le \sqrt{\|Z_{[r]}\|_F^2 + \|Z_{[r]^c}\|_F^2} \\
&\le \sqrt{2\|Z_{[r]}\|_F^2 + \frac{2\|Z_{[r]}\|_F\|X_{[r]^c}\|_*}{\sqrt{r}}} \\
&\le \sqrt{2}\|Z_{[r]}\|_F + \frac{\sqrt{2}}{2}\frac{\|X_{[r]^c}\|_*}{\sqrt{r}} \\
&\le \sqrt{2}\Delta + \frac{\sqrt{2}}{2}\frac{\|X_{[r]^c}\|_*}{\sqrt{r}} \\
&\le \frac{2\sqrt{1+\hat{\delta}_{2r}}}{\frac{\sqrt{2}}{2} - \hat{\delta}_{2r}}\epsilon'_{\mathcal{A},r,y} \\
&+ \frac{\frac{\sqrt{2}}{2}\hat{\delta}_{2r} + \sqrt{2(\frac{\sqrt{2}}{2} - \hat{\delta}_{2r})\hat{\delta}_{2r}} + \frac{1}{2}}{\frac{\sqrt{2}}{2} - \hat{\delta}_{2r}}\frac{\|X_{[r]^c}\|_*}{\sqrt{r}} \\
&\le \frac{2\sqrt{1+\delta_{2r}}(1+\epsilon_{\mathcal{A}}^{(2r)})}{\frac{\sqrt{2}}{2} + 1 - (1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2}\epsilon'_{\mathcal{A},r,y} \\
&+ \left[\frac{\sqrt{2}}{2} + 1 - (1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2\right]^{-1}\left\{\frac{\sqrt{2}}{2}[(1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2 - 1]\right. \\
&+ \left.\sqrt{2[\frac{\sqrt{2}}{2} - (1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2 + 1][(1+\delta_{2r})(1+\epsilon_{\mathcal{A}}^{(2r)})^2 - 1]} + \frac{1}{2}\right\}\frac{\|X_{[r]^c}\|_*}{\sqrt{r}}.
\end{aligned}
$$

□

Proof of the theorem 2.8. By using Lemma 3.10, we get

$$
\begin{aligned}
\|X^*\|_* = \|X - (X - X^*)\|_* &\ge \sum_{i=1}^{n_1}|\sigma_i(X) - \sigma_i(X - X^*)| \\
&= \sum_{i=1}^{r}|\sigma_i(X) - \sigma_i(X - X^*)| + \sum_{i=r+1}^{n_1}|\sigma_i(X) - \sigma_i(X - X^*)| \\
&\ge \sum_{i=1}^{r}(\sigma_i(X) - \sigma_i(X - X^*)) + \sum_{i=r+1}^{n_1}(\sigma_i(X - X^*) - \sigma_i(X)).
\end{aligned}
$$

Therefore, by Hölder's inequality and the minimality of X^*,

$$
\begin{aligned}
\|(X - X^*)_{[r]^c}\|_* &\le \|X^*\|_* - \|X_{[r]}\|_* + \|(X - X^*)_{[r]}\|_* + \|X_{[r]^c}\|_* \\
&\le \|X^*\|_* - \|X\|_* + \sqrt{r}\|(X - X^*)_{[r]}\|_F + 2\|X_{[r]^c}\|_* \\
&\le \sqrt{r}\|(X - X^*)_{[r]}\|_F + 2\|X_{[r]^c}\|_*. \qquad (3.41)
\end{aligned}
$$

Making use of the Frobenius-robust rank null space property of $\hat{\mathcal{A}}$ and (3.32), we get

$$
\|(X - X^*)_{[r]^c}\|_* \le \hat{\rho}\|(X - X^*)_{[r]^c}\|_* + \sqrt{r}\hat{\tau}\cdot(2\epsilon'_{\mathcal{A},r,y}) + 2\|X_{[r]^c}\|_*. \qquad (3.42)
$$

By arranging the terms in (3.42), we get

$$
\|(X - X^*)_{[r]^c}\|_* \le \frac{1}{1-\hat{\rho}}(\sqrt{r}\hat{\tau}\cdot(2\epsilon'_{\mathcal{A},r,y}) + 2\|X_{[r]^c}\|_*). \qquad (3.43)
$$

Combining with Hölder's inequality, the Frobenius-robust rank null space property of $\hat{\mathcal{A}}$ and (3.43), we get

$$\begin{aligned}\|X-X^*\|_* &= \|(X-X^*)_{[r]}\|_* + \|(X-X^*)_{[r]^c}\|_* \\ &\le \sqrt{r}\|(X-X^*)_{[r]}\|_F + \|(X-X^*)_{[r]^c}\|_* \\ &\le (1+\hat{\rho})\|(X-X^*)_{[r]^c}\|_* + \hat{\tau}\sqrt{r}\cdot(2\epsilon'_{\mathcal{A},r,y}) \\ &\le \frac{2(1+\hat{\rho})}{1-\hat{\rho}}\|X_{[r]^c}\|_* + \frac{2\hat{\tau}\sqrt{r}}{1-\hat{\rho}}\cdot(2\epsilon'_{\mathcal{A},r,y}). \end{aligned} \tag{3.44}$$

By the Frobenius-robust rank null space property of $\hat{\mathcal{A}}$, Lemmas 3.7 and 3.9 and (3.44), we get

$$\begin{aligned}\|X-X^*\|_F &\le \|(X-X^*)_{[r]}\|_F + \|(X-X^*)_{[r]^c}\|_F \\ &\le \frac{\hat{\rho}}{\sqrt{r}}\|(X-X^*)_{[r]^c}\|_* + \hat{\tau}\|\mathcal{A}(X-X^*)\|_2 + \frac{\|X-X^*\|_*}{\sqrt{r}} \\ &\le \frac{1+\hat{\rho}}{\sqrt{r}}\|X-X^*\|_* + \hat{\tau}\cdot(2\epsilon'_{\mathcal{A},r,y}) \\ &\le \frac{2(1+\hat{\rho})^2}{1-\hat{\rho}}\frac{\|X_{[r]^c}\|_*}{\sqrt{r}} + \frac{2(3+\hat{\rho})\hat{\tau}}{1-\hat{\rho}}\epsilon'_{\mathcal{A},r,y} \\ &= \frac{2\tau}{1+\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}\left[3+\rho-\frac{\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\rho+\sqrt{r})}{1+\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}\right]\cdot\left[1-\rho+\frac{\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\rho+\sqrt{r})}{1+\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}\right]^{-1}\epsilon'_{\mathcal{A},r,y} \\ &+ 2\left[1+\rho-\frac{\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\rho+\sqrt{r})}{1+\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}\right]^2\cdot\left[1-\rho+\frac{\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}(\rho+\sqrt{r})}{1+\tau\epsilon_{\mathcal{A}}\|\mathcal{A}\|_{op}}\right]^{-1}\frac{\|X_{[r]^c}\|_*}{\sqrt{r}}.\end{aligned}$$

We complete the proof. □

4 Conclusion

This work primarily considers completely perturbed issue exploiting the constrained nuclear norm minimization for the low-rank matrices recovery. We establish two main results, which present sufficient conditions and the associating upper bound estimations of recovery error. The gained results give a robust and stable assurance for reconstructing low-rank matrices in the presence of total noise. The actual meanings of gained results include two aspects: firstly, it can instruct the choice of the measurement operators for the low-rank matrices reconstruction, viz, the recovery ability can be better facilitated by an operator with smaller RIC than a bigger one; secondly, it can also offer a theoretical support for bounding error.

5 Acknowledgements

This work was supported by Natural Science Foundation of China (No. 61673015, 61273020), Fundamental Research Funds for the Central Universities (No. XDJK2015A007, XDJK2018C076, SWU1809002), Youth Science and technology talent development project (No.Qian jiao he KY zi [2018]313), Science, Science and technology Foundation of Guizhou province (Qian ke he Ji Chu [2016]1161), Guizhou province natural science foundation in China (Qian Jiao He KY [2016]255).

References

[1] Recht B, Fazel M, Parrilo P A. Guaranteed minimum rank solutions of linear matrix equations via nuclear norm minimization. SIAM Rev, 2010, 52: 471-501

[2] Chang, X., Zhong, Y., Wang, Y., & Lin, S. (2018). Unified Low-Rank Matrix Estimate via Penalized Matrix Least Squares Approximation. IEEE transactions on neural networks and learning systems, (99), 1-12.

[3] Lin J, Li S. Convergence of projected Landweber iteration for matrix rank minimization. Appl Comput Harmon Anal, 2014, 36: 316-325

[4] R. Mazumder, T. Hastie, and R. Tibshirani, Spectral regularization algorithms for learning large incomplete matrices, J. Mach. Learn. Res., vol. 11, pp. 2287-2322, 2010.

[5] E.J. Candès, X.D. Li, Y. Ma, and J. Wright, Robust principal component analysis? J. ACM, vol, 58, no. 3, pp. 11-1: 11-37, 2011.

[6] Y. Wang, L. Lin, Q. Zhao, T. Yue, D. Meng, and Y. Leung, Compressive sensing of hyperspectral images via joint tensor tucker decomposition and weighted total variation regularization, IEEE Geosci. Remote Sens. Lett., vol. 14, no. 12, pp. 2457-2461, Dec. 2017.

[7] T. Blumensath and M. Davies, Compressed sensing and source separation, Int. Conf. Independent Component Anal. Signal Separation 4666 (2007) 341-348.

[8] Herman MA, Strohmer T. High-resolution radar via compressed sensing. IEEE Trans Signal Process 2009; 57(6): 2275-84.

[9] M. A. Herman and T. Strohmer, General deviants: An analysis of perturbations in compressed sensing, IEEE J. Sel. Top. Signal Process. 4(2) (2010) 342-349.

[10] Zhang, J., Wang, J., & Wang, W. (2016). A perturbation analysis of block-sparse compressed sensing via mixed ℓ_2/ℓ_1 minimization. International Journal of Wavelets, Multiresolution and Information Processing, 14(04), 1650026.

[11] E.J. Candès and Y. Plan, Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements, IEEE Trans. Inf. Theory, vol. 57, no. 4, pp. 2342-2359, 2011.

[12] K. Mohan and M. Fazel, New restricted isometry results for noisy low-rank recovery, in Proc. ISIT, pp. 1573-1577, Jun. 2010.

[13] H. Wang and S. Li, The bounds of restricted isometry constants for low rank matrices recovery, Sci. China, Math., vol. 56, no. 6, pp. 1117-1127, 2012.

[14] T.T Cai and A.R. Zhang, Sharp RIP bound for sparse signal and low-rank matrix recovery, Appl. Comput. Harmon. Anal., vol. 35, pp. 74-93, Jan. 2013.

[15] Zhang M, Huang Z H, Zhang Y. Restricted p-Isometry Properties of Nonconvex Matrix Recovery[J]. IEEE Transactions on Information Theory, 2013, 59(7): 4316-4323.

[16] L. C. Kong and N. H. Xiu, "Exact low-rank matrix recovery via nonconvex Schatten p-minimization," *Asia Pac. J. Oper. Res.*, vol. 30, no. 3, pp. 1340010, Jun. 2013.

[17] W. G. Chen and Y. L. Li, "Stable recovery of low-rank matrix via nonconvex Schatten p-minimization," *Sci. China Math.*, vol. 58, no. 12, pp 2643-2654, Dec. 2015.

[18] T.T. Cai and A.R. Zhang, Sparse representation of a polytope and recovery of sparse signals and low-rank matrices, IEEE Trans. Inf. Theory, vol. 60, no. 1, pp. 122-132, Jan. 2014.

[19] Wang, WD, Zhang, F, Wang, JJ. Low-rank matrix recovery via regularized nuclear norm minimization. 2018, [online] Available: https://www.researchgate.net/publication/328600364_Low-rank_ matrix_recovery_-via_regularized_nuclear_norm_minimization

[20] A. C. Gilbert, M. J. Strauss, J. A. Tropp, and R. Vershynin, One sketch for all: Fast algorithms for compressed sensing, in Proc. 39th ACM Symp. Theory Comput. (STOC), San Diego, CA, Jun. 2007.

[21] Kabanava M, Kueng R, Rauhut H, et al. Stable low-rank matrix recovery via null space properties. Information and Inference: A Journal of the IMA, 2016, 5(4): 405-441.

[22] Foucart, S. & Rauhut, H. (2013) A mathematical introduction to compressive sensing. Applied and Numerical Harmonic Analysis (J. S. Benedetto ed.). New York: Birkhäuser/Springer, 2013.

[23] Horn, R. & Johnson, C. Topics in Matrix Analysis. Cambridge: Cambridge University Press, 1991.