New sufficient conditions of robust recovery for low-rank matrices *

Jianwen Huang[a], Jianjun Wang[a,b], Feng Zhang[a], Wendong Wang[a]
[a]School of Mathematics and Statistics, Southwest University, Chongqing, 400715, China
[b]Research Center for Artificial Intelligence & Education Big Data, Southwest University, Chongqing, 400715, China

Abstract. In this paper we investigate the reconstruction conditions of nuclear norm minimization for low-rank matrix recovery from a given linear system of equality constraints. Sufficient conditions are derived to guarantee the robust reconstruction in bounded l_2 and Dantzig selector noise settings ($\epsilon \neq 0$) or exactly reconstruction in the noiseless context ($\epsilon = 0$) of all rank r matrices $X \in \mathbb{R}^{m\times n}$ from $b = \mathcal{A}(X) + z$ via nuclear norm minimization. Furthermore, we not only show that when $t = 1$, the upper bound of δ_r is the same as the result of Cai and Zhang [9], but also demonstrate that the gained upper bounds concerning the recovery error are better. Finally, we prove that the restricted isometry property condition is sharp.

Keywords. Low-rank matrix recovery; nuclear norm normalization; restricted isometry property condition; compressed sensing; convex optimization.

AMS Classification(2010): 90C25, 90C59, 15A52, 52A41, 49M30, 65K10, 65J22

1 Introduction

Suppose that $X \in \mathbb{R}^{m\times n}$ is an unknown low rank matrix, $\mathcal{A} : \mathbb{R}^{m\times n} \to \mathbb{R}^q$ is a known linear map, $b \in \mathbb{R}^q$ is a given observation and $z \in \mathbb{R}^q$ is measurement error. The rank minimization problem is defined as follows:

$$\min_X \ \text{rank}(X) \ \text{s.t.} \ \|\mathcal{A}(X) - b\|_2 \leq \epsilon, \tag{1.1}$$

where $b = \mathcal{A}(X) + z$ and ϵ stands for the noise level. Since the problem (1.1) is NP-hard in general, Recht et al. [1] introduced a convex relaxation, which minimizes nuclear norm (also known as the Schatten 1-norm or trace norm)

$$\min_X \ \|X\|_* \ \text{s.t.} \ \|\mathcal{A}(X) - b\|_2 \leq \epsilon, \tag{1.2}$$

where $\|X\|_* = \sum_i^{\min\{m,n\}} \sigma_i(X)$ and $\sigma_i(X)$ are the singular values. The problem (1.2) is convex, thus there are a large number of approaches which can be used for solving it. For efficient algorithms

*Corresponding author. E-mail address: wjj@swu.edu.cn (J. Wang).

solving broad-scale instants, a great deal of researchers have developed them, such as Tanner and Wei [2], Zhang and Li [3], Lu et al [4] and Lin et al. [5].

When $m = n$ and the matrix $X = \mathrm{diag}(x)$ ($x \in \mathbb{R}^m$) is diagonal matrix, the problems (1.1) and (1.2) degenerate to the l_0-minimization and l_1-minimization, respectively, which belong to the main optimization problems in the compressed sensing (CS).

In order to study the relationship between l_1-minimization and the nuclear norm minimization problems, Recht et al. [1] extended the notion of restricted isometry constant proposed by Candès [6] to low-rank matrix recovery case. The concept is as follows:

Definition 1.1. *Let $\mathcal{A} : \mathbb{R}^{m\times n} \to \mathbb{R}^q$ be a linear map. For any integer r ($1 \le r \le \min\{m, n\}$), the restricted isometry constant (RIC) of order r is defined as the smallest positive number δ_r that satisfies*

$$(1 - \delta_r)\|X\|_F^2 \le \|\mathcal{A}(X)\|_2^2 \le (1 + \delta_r)\|X\|_F^2 \tag{1.3}$$

for all r-rank matrices X (i.e., the rank of X is at most r), where $\|X\|_F^2 = \langle X, X\rangle = Tr(X^\top X)$ is the Frobenius norm of X, which is also equal to the sum of the square of singular values and the inner product in $\mathbb{R}^{m\times n}$ as $\langle X, Y\rangle = Tr(X^\top Y) = \sum_{i=1}^m \sum_{j=1}^n X_{ij}Y_{ij}$ for matrices X and Y of the same dimensions.

By the aforementioned definition, it is easy to see that if $r_1 \le r_2$, then $\delta_{r1} \le \delta_{r2}$.

Although it is not easy to examine the restricted isometry property for a given linear map, it is one of the central notions in low-rank matrix recovery. In fact, it has been showed [1] that Gaussian or sub-Gaussian random measurement map $\mathcal{A}$ fulfills the restricted isometry property with high probability.

Recht et al. [1] showed that for the noiseless case (i.e., $\epsilon = 0$), if $\delta_{5r} < 1/10$, then the minimum-rank solution to (1.2) can be recovered by solving a convex optimization problem. Candès and Plan [7] proved that when $\delta_{4r} < \sqrt{2} - 1$, a low-rank matrix can be robustly recovered by nuclear norm minimization (1.2). Mohan and Fazel [8] improved the upper bound of RIC to $\delta_{4r} < 0.558$. Cai and Zhang [9] presented the sharp condition $\delta_r < 1/3$ ($\delta_k < 1/3$) for low-rank matrix (sparse signal) recovery. Wang and Li [10] showed that the upper bounds $\delta_r < 1/3$ and $\delta_{2r} < \sqrt{2}/2$ are optimal. Kong and Xiu [11] obtained a uniform bound on RIC $\delta_{4r} < \sqrt{2} - 1$ for any $p \in (0, 1]$ for low-rank matrix recovery via Schatten p-minimization. Chen and Li [12] showed that for any given $\delta_{4r} \in [\sqrt{3}/2, 1)$, $p \in (0, 2(1 - \delta_{4r})]$ suffices for the robust recovery of all r-rank matrices via Schatten p-minimization.

Cai and Zhang [13] showed that for any given $t \ge 4/3$, $\delta_{tr} < \sqrt{(t-1)/t}$ ensures the exact reconstruction for all matrices with rank no more than r in the noise-free case via the constrained nuclear norm minimization (1.2). Furthermore, for any $\varepsilon > 0$, $\delta_{tr} < \sqrt{(t-1)/t}+\varepsilon$ doesn't suffice to make sure the exact recovery of all r-rank matrices for large r. Besides, they showed that condition $\delta_{tr} < \sqrt{(t-1)/t}$ suffices for robust reconstruction of nearly low-rank matrices in the noisy case.

Motivated by the aforementioned papers, we further discuss the upper bounds of δ_{tr} associated with some linear map $\mathcal{A}$ as $0 < t < 4/3$. Sufficient conditions regarding δ_{tr} with $0 < t < 4/3$ are established to guarantee the robust reconstruction ($\epsilon \ne 0$) or ($\epsilon = 0$) of all r-rank matrices $X \in \mathbb{R}^{m\times n}$ satisfying $b = \mathcal{A}(X) + z$ with $\|z\|_2 \le \epsilon$ and $\|\mathcal{A}^*(z)\| \le \epsilon$, respectively. Thereby, combined with [13], a complete description for sharp restricted isometry property (RIP) constants for all $t > 0$ is established to ensure the exact reconstruction of all matrices with rank no more than r via nuclear norm minimization.

2 Preliminaries

We begin by introducing basic notations. We also gather a few lemmas needed for the proofs of main results.

For any matrix $X \in \mathbb{R}^{m\times n}$ $(m \leq n)$, the singular value decomposition (SVD) of X is represented by

$$X = U\mathrm{diag}(\sigma(X))V^\top,$$

where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times m}$ are orthogonal matrices, and $\sigma(X) = (\sigma_1(X), \cdots, \sigma_m(X))^\top$ indicates the vector of the singular values of X. Assume that $\sigma_1(X) \geq \sigma_2(X) \geq \cdots \geq \sigma_m(X)$. Consequently, the best r-rank approximation to the matrix X is

$$X^{(r)} = U \begin{bmatrix} \mathrm{diag}(\sigma^r(X)) & 0 \\ 0 & 0 \end{bmatrix} V^\top,$$

where $\sigma^r(X) = (\sigma_1(X), \cdots, \sigma_r(X))^\top$.

For a linear map $\mathcal{A}:\ \mathbb{R}^{m\times n} \to \mathbb{R}^q$, denote by its adjoint operator $\mathcal{A}^*:\ \mathbb{R}^q \to \mathbb{R}^{m\times n}$. Then, for all $X \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^q$, $\langle X, \mathcal{A}^*(b)\rangle = \langle \mathcal{A}(X), b\rangle$.

Without loss of generality, let X be the original matrix that we want to find and X^* be an optimal solution to the problem (1.2). Let $Z = X - X^*$. Let SVD of $U^\top ZV \in \mathbb{R}^{m\times m}$ be provided by

$$U^\top ZV = U_0 \begin{bmatrix} \mathrm{diag}\left(\sigma_T(U^\top ZV)\right) & 0 \\ 0 & \mathrm{diag}\left(\sigma_{T^c}(U^\top ZV)\right) \end{bmatrix} V_0^\top$$

where $U_0,\ V_0 \in \mathbb{R}^{m\times m}$ are orthogonal matrices, $\sigma_T(U^\top ZV) = \left(\sigma_1(U^\top ZV), \cdots, \sigma_r(U^\top ZV)\right)^\top$, $\sigma_{T^c}(U^\top ZV) = \left(\sigma_{r+1}(U^\top ZV), \cdots, \sigma_m(U^\top ZV)\right)^\top$, and we suppose that $\sigma_1(U^\top ZV) \geq \cdots \geq \sigma_r(U^\top ZV) \geq \sigma_{r+1}(U^\top ZV) \geq \cdots \geq \sigma_m(U^\top ZV)$. Therefore, the matrix Z is decomposed as

$$Z = Z^{(r)} + Z_c^{(r)},$$

where

$$Z^{(r)} = UU_0 \begin{bmatrix} \mathrm{diag}\left(\sigma_T(U^\top ZV)\right) & 0 \\ 0 & 0 \end{bmatrix} V_0^\top V^\top$$

and

$$Z_c^{(r)} = UU_0 \begin{bmatrix} 0 & 0 \\ 0 & \mathrm{diag}\left(\sigma_{T^c}(U^\top ZV)\right) \end{bmatrix} V_0^\top V^\top.$$

It is not hard to see that $X^{(r)}(Z_c^{(r)})^\top = 0$ and $(X^{(r)})^\top Z_c^{(r)} = 0$.

In order to show the main results, we need some elementary identities, which were given in [14] (see Lemma 1).

Lemma 2.1. *Give matrices $\{V_i :\ i \in T\}$ in a matrix space $\mathcal{V}$ with inner product $\langle\cdot\rangle$, where T denotes the index set with $|T| = r$. Select all subsets $T_i \subset T$ with $|T_i| = k$, $i \in I$ and $|I| = \binom{r}{k}$, then we get*

$$\sum_{i\in I}\sum_{p\in T_i} V_p = \binom{r-1}{k-1} \sum_{p\in T} V_p \ (k \geq 1), \tag{2.1}$$

and

$$\sum_{i\in I}\sum_{p\neq q\in T_i}\langle V_p, V_q\rangle = \binom{r-2}{k-2}\sum_{p\neq q\in T}\langle V_p, V_q\rangle \ \ (k\geq 2). \tag{2.2}$$

Cai and Zhang developed a new elementary technique which states an elementary geometric fact: Any point in a polytope can be represented as a convex combination of sparse vectors (see Lemma 1.1 in [13]). It gives a crucial technical tool for the proof of our main results. It is also the special case $p=1$ of Zhang and Li's result (see Lemma 2.2 in [15]).

Lemma 2.2. *Let $r\leq m$ be an integer, and α be a positive real number. We can represent any vector x in the set*

$$V=\{x\in\mathbb{R}^m:\ \|x\|_1\leq r\alpha,\ \|x\|_\infty\leq\alpha\},$$

as a convex combination of r-sparse vectors, i.e.,

$$x=\sum_i \lambda_i u_i$$

where $\sum_i\lambda_i=1$ with $\lambda_i\geq 0$, $|\mathrm{sup}(u_i)|\leq r$, $\mathrm{sup}(u_i)\subset\mathrm{sup}(x)$ and $\sum_i\lambda_i\|u_i\|_2^2\leq r\alpha^2$.

Lemma 2.3. *(Lemma 2.3 in [1]) Let $X,\ Y$ be the matrices of same dimensions. If $XY^\top=0$ and $X^\top Y=0$, then*

$$\|X+Y\|_*=\|X\|_*+\|Y\|_*. \tag{2.3}$$

Lemma 2.4. *We have*

$$\|Z_c^{(r)}\|_*\leq\|Z^{(r)}\|_*+2\|X-X^{(r)}\|_*. \tag{2.4}$$

Proof. Since X^* is the optimal solution to the problem (1.2), we get

$$\|X\|_*\geq\|X^*\|_*=\|X-Z\|_*. \tag{2.5}$$

Applying the reverse inequality to (2.5), we get

$$\begin{aligned}\|X-Z\|_*&=\|(X^{(r)}-Z_c^{(r)})+(X-X^{(r)}-Z^{(r)})\|_*\\&\geq\|X^{(r)}-Z_c^{(r)}\|_*-\|X-X^{(r)}-Z^{(r)}\|_*.\end{aligned} \tag{2.6}$$

By Lemma 2.3 and the forward inequality, we get

$$\begin{aligned}&\|X^{(r)}+(-Z_c^{(r)})\|_*-\|X-X^{(r)}+(-Z^{(r)})\|_*\\&\geq\|X^{(r)}\|_*+\|Z_c^{(r)}\|_*-\|X-X^{(r)}\|_*-\|Z^{(r)}\|_*.\end{aligned} \tag{2.7}$$

Combining with (2.5), (2.6) and (2.7), we get

$$\begin{aligned}\|Z_c^{(r)}\|_*&\leq\|X\|_*-\|X^{(r)}\|_*+\|X-X^{(r)}\|_*+\|Z^{(r)}\|_*\\&\leq\|Z^{(r)}\|_*+2\|X-X^{(r)}\|_*.\end{aligned}$$

The proof of the lemma is completed.

□

Select positive integers a and b satisfying $a+b=tr$ and $b\le a\le r$. We use $T_i,\ S_j$ to represent all possible index set contained in $\{1,2,\cdots,r\}$ (i.e., $T_i,\ S_j\subset\{1,\cdots,r\}$) and $|T_i|=a$, $|S_j|=b$, where $i\in A$ and $j\in B$ with $|A|=\binom{r}{a}$ and $|B|=\binom{r}{b}$. Define

$$Z_{T_i}^{(r)}=UU_0\begin{bmatrix}\operatorname{diag}\left(\sigma_{T_i}(U^\top ZV)\right) & 0\\ 0 & 0\end{bmatrix}V_0^\top V^\top,$$

and

$$Z_{S_j}^{(r)}=UU_0\begin{bmatrix}\operatorname{diag}\left(\sigma_{S_j}(U^\top ZV)\right) & 0\\ 0 & 0\end{bmatrix}V_0^\top V^\top.$$

Here $\sigma_{T_i}(U^\top ZV)$ ($\sigma_{S_j}(U^\top ZV)$) denotes the vector that equals to $\sigma_T(U^\top ZV)$ on T_i (S_j), and zero elsewhere.

Lemma 2.5. *We have*

$$Z_c^{(r)}=\sum_k\mu_kU_k,\ Z_c^{(r)}=\sum_k\nu_kV_k,\ Z_c^{(r)}=\sum_k\tau_kW_k,$$

where $\sum_k\mu_k=\sum_k\nu_k=\sum_k\tau_k=1$ *with* $\nu_k,\ \mu_k,\ \tau_k\ge0$, $U_k,\ V_k,\ W_k$ *are* b*-rank,* a*-rank and* $(t-1)r$*-rank* $(t>1)$ *with*

$$\sum_k\mu_k\|U_k\|_F^2\le\frac{r^2}{b}\alpha^2, \tag{2.8}$$

$$\sum_k\nu_k\|V_k\|_F^2\le\frac{r^2}{a}\alpha^2, \tag{2.9}$$

and

$$\sum_k\tau_k\|W_k\|_F^2\le\frac{r^2}{t-1}\alpha^2. \tag{2.10}$$

Proof. Set

$$\alpha=\frac{\|Z^{(r)}\|_*+2\|X-X^{(r)}\|_*}{r}.$$

By Lemma 2.4, then

$$\|Z_c^{(r)}\|_*\le r\alpha.$$

By the definition of $Z_c^{(r)}$, we get

$$\|\sigma_{T^c}(U^\top ZV)\|_1\le r\alpha\le b\frac{r}{b}\alpha. \tag{2.11}$$

By the decomposition of Z, we get

$$\begin{aligned}\|\sigma_{T^c}(U^\top ZV)\|_\infty&\le\frac{\|\sigma_T(U^\top ZV)\|_1}{r}\\&\le\frac{\|Z^{(r)}\|_*+2\|X-X^{(r)}\|_*}{r}\end{aligned}$$

$$\leq \alpha \leq \frac{r}{b}\alpha. \tag{2.12}$$

Combining with Lemma 2.2, (2.11) and (2.12), $\sigma_{T^c}(U^\top ZV)$ is decomposed into the convex combination of b-sparse vectors, i.e., $\sigma_{T^c}(U^\top ZV) = \sum_k \mu_k u_k$ with

$$\sum_k \mu_k \|u_k\|_2^2 \leq \frac{r^2}{b}\alpha^2.$$

Define

$$U_k = UU_0 \begin{bmatrix} 0 & 0 \\ 0 & \mathrm{diag}(u_k) \end{bmatrix} V_0^\top V^\top.$$

It is easy to see that U_k is b-rank. Therefore, $Z_c^{(r)}$ is decomposed as $Z_c^{(r)} = \sum_k \mu_k U_k$ with

$$\sum_k \mu_k \|U_k\|_F^2 = \sum_k \mu_k \|u_k\|_2^2 \leq \frac{r^2}{b}\alpha^2.$$

Likewise, $Z_c^{(r)}$ can also be denoted by

$$Z_c^{(r)} = \sum_k \nu_k V_k, \;\; Z_c^{(r)} = \sum_k \tau_k W_k,$$

where V_k is a-rank, W_k is $(t-1)r$-rank $(t>1)$ with

$$\sum_k \nu_k \|V_k\|_F^2 \leq \frac{r^2}{a}\alpha^2,$$

and

$$\sum_k \tau_k \|V_k\|_F^2 \leq \frac{r^2}{t-1}\alpha^2.$$

□

One can easily check that $\left\langle Z_{T_i}^{(r)}, U_k \right\rangle = 0$, $\left\langle Z_{S_j}^{(r)}, V_k \right\rangle = 0$ and $\langle Z^{(r)}, W_k \rangle = 0$.

Lemma 2.6. *We have that for $0 < t < 1$,*

$$\begin{aligned}
&\frac{\rho_{a,b}(t)}{\binom{r}{a}\binom{r-a}{b}} \sum_{T_i \bigcap S_j = \emptyset} \left[\left\| \mathcal{A}\left(Z_{T_i}^{(r)} + Z_{S_j}^{(r)}\right) \right\|_2^2 - \frac{r-a-b}{abr} \left\| \mathcal{A}\left(bZ_{T_i}^{(r)} - aZ_{S_j}^{(r)}\right) \right\|_2^2 \right] \\
&= -2t^2(2-t)ab \left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z \right\rangle + t\Delta_{a,b},
\end{aligned} \tag{2.13}$$

and for $1 \leq t < 4/3$,

$$\begin{aligned}
&\rho_{a,b}(t) \sum_k \tau_k \left[\left\| \mathcal{A}\left(Z^{(r)} + (t-1)W_k\right) \right\|_2^2 - \left\| (t-1)\mathcal{A}\left(Z^{(r)} - W_k\right) \right\|_2^2 \right] \\
&= -2t^3[ab - (t-1)r^2] \left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z \right\rangle + (4-3t)\Delta_{a,b},
\end{aligned} \tag{2.14}$$

where

$$\rho_{a,b}(t) = (a+b)^2 - 2ab(4-t),$$

and

$$\Delta_{a,b} = \frac{r-b}{a\binom{r}{a}} \sum_{i\in A,\ k} \mu_k \left[a^2 \left\| \mathcal{A}\left(Z_{T_i}^{(r)} + \frac{b}{r} U_k \right) \right\|_2^2 - b^2 \left\| \mathcal{A}\left(Z_{T_i}^{(r)} - \frac{a}{r} U_k \right) \right\|_2^2 \right]$$
$$+ \frac{r-a}{b\binom{r}{b}} \sum_{j\in B,\ k} \nu_k \left[b^2 \left\| \mathcal{A}\left(Z_{S_j}^{(r)} + \frac{a}{r} V_k \right) \right\|_2^2 - a^2 \left\| \mathcal{A}\left(Z_{S_j}^{(r)} - \frac{b}{r} V_k \right) \right\|_2^2 \right]. \tag{2.15}$$

Proof. The proof takes advantage of the ideas from [9], [14]. By Lemma 2.1, we get

$$\begin{aligned}
\Delta_{a,b} =& (a^2-b^2)\left[\frac{r-b}{a\binom{r}{a}} \sum_{i\in A} \|\mathcal{A}Z_{T_i}^{(r)}\|_2^2 - \frac{r-a}{b\binom{r}{b}} \sum_{j\in B} \|\mathcal{A}Z_{S_j}^{(r)}\|_2^2 \right] \\
&+ \frac{2(a^2b+ab^2)}{r} \left\langle \frac{r-b}{a\binom{r}{a}} \sum_{i\in A} \mathcal{A}Z_{T_i}^{(r)} + \frac{r-a}{b\binom{r}{b}} \sum_{j\in B} \mathcal{A}Z_{S_j}^{(r)}, \mathcal{A}Z_c^{(r)} \right\rangle \\
&= (a^2-b^2)\left(\frac{r-b}{a\binom{r}{a}} \tbinom{r-1}{a-1} \|\mathcal{A}Z^{(r)}\|_2^2 - \frac{r-a}{b\binom{r}{b}} \tbinom{r-1}{b-1} \|\mathcal{A}Z^{(r)}\|_2^2 \right) \\
&+ \frac{2ab(a+b)}{r} \left\langle \frac{r-b}{a\binom{r}{a}} \tbinom{r-1}{a-1} \mathcal{A}Z^{(r)} + \frac{r-a}{b\binom{r}{b}} \tbinom{r-1}{b-1} \mathcal{A}Z^{(r)}, \mathcal{A}Z_c^{(r)} \right\rangle \\
&= (a^2-b^2)\frac{a-b}{r} \|\mathcal{A}Z^{(r)}\|_2^2 + 2abt \frac{2r-a-b}{r} \left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z_c^{(r)} \right\rangle \\
&= t\rho_{a,b}(t) \|\mathcal{A}Z^{(r)}\|_2^2 + 2abt(2-t) \left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z \right\rangle .
\end{aligned} \tag{2.16}$$

where the first equality follows from Lemma 2.5, i.e., $Z_c^{(r)}$ has the convex decomposition, and in the second equality, we used the identity (2.1).

As $0 < t < 1$, by Lemma 2 [14], we get

$$\begin{aligned}
LHS =& \rho_{a,b}(t) \left(\frac{a+b}{r} \right)^2 \|\mathcal{A}Z^{(r)}\|_2^2 \\
=& \rho_{a,b}(t) t^2 \|\mathcal{A}Z^{(r)}\|_2^2.
\end{aligned} \tag{2.17}$$

Substituting (2.16) to the right hand side of (2.13), we get

$$\begin{aligned}
RHS =& t\left[t\rho_{a,b}(t)\|\mathcal{A}Z^{(r)}\|_2^2 + 2abt(2-t) \left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z \right\rangle \right] \\
&- 2t^2(2-t)ab \left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z \right\rangle \\
=& LHS.
\end{aligned}$$

Accordingly, the identity (2.13) holds.

As $1 \le t < 4/3$, we get

$$\begin{aligned}
LHS =& \rho_{a,b}(t) \left\{ [1-(t-1)^2] \|\mathcal{A}Z^{(r)}\|_2^2 + 2(t-1)t \left\langle \mathcal{A}Z^{(r)}, \sum_k \tau_k \mathcal{A}W_k \right\rangle \right\} \\
=& \rho_{a,b}(t) \left\{ [1-(t-1)^2] \|\mathcal{A}Z^{(r)}\|_2^2 + 2(t-1)t \left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z_c^{(r)} \right\rangle \right\}
\end{aligned}$$

$$= \rho_{a,b}(t)\left\{ (4t-3t^2)\|\mathcal{A}Z^{(r)}\|_2^2 + 2(t-1)t\left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z \right\rangle \right\}.$$

We have

$$\begin{aligned}
RHS =&(4-3t)\left[t\rho_{a,b}(t)\|\mathcal{A}Z^{(r)}\|_2^2 + 2abt(2-t)\left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z \right\rangle \right] \\
&- 2t^3[ab-(t-1)r^2]\left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z \right\rangle \\
=&(4t-3t^2)\rho_{a,b}(t)\|\mathcal{A}Z^{(r)}\|_2^2 + 2t\left\{ab(2-t)(4-3t) - t^2[ab-(t-1)r^2]\right\}\left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z \right\rangle \\
=&LHS.
\end{aligned}$$

Therefore, the identity (2.14) holds.

□

Lemma 2.7. *Let X, Y be matrices with X, $Y \in \mathbb{R}^{m\times n}$. If $XY^\top = 0$ and $X^\top Y = 0$, then the following holds:*

$$\|X+Y\|_F^2 = \|X\|_F^2 + \|Y\|_F^2. \tag{2.18}$$

Proof. Due to $XY^\top = 0$ and $X^\top Y = 0$, combining with the proof of Lemma 2.3 [1], then there are matrices $[U_X\ U_Y\ U_Z]$ and $[V_X\ V_Y\ V_Z]$ such that the singular value decompositions of X and Y are as follows:

$$X = [U_X\ U_Y\ U_Z]\begin{bmatrix} \mathrm{diag}(\sigma(X)) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}[V_X\ V_Y\ V_Z]^\top,$$

and

$$Y = [U_X\ U_Y\ U_Z]\begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathrm{diag}(\sigma(Y)) & 0 \\ 0 & 0 & 0 \end{bmatrix}[V_X\ V_Y\ V_Z]^\top,$$

where $[U_X\ U_Y\ U_Z]$ and $[V_X\ V_Y\ V_Z]$ are orthogonal matrices. As a consequence, we have the SVD of $X+Y$ as follows:

$$X+Y = [U_X\ U_Y\ U_Z]\begin{bmatrix} \mathrm{diag}(\sigma(X)) & 0 & 0 \\ 0 & \mathrm{diag}(\sigma(Y)) & 0 \\ 0 & 0 & 0 \end{bmatrix}[V_X\ V_Y\ V_Z]^\top.$$

Therefore,

$$\|X+Y\|_F^2 = \left\| \begin{bmatrix} \mathrm{diag}(\sigma(X)) & 0 & 0 \\ 0 & \mathrm{diag}(\sigma(Y)) & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\|_F^2 = \|X\|_F^2 + \|Y\|_F^2.$$

The desired result are derived.

□

Lemma 2.8. *It holds that*

$$\{[(a+b)^2 - 4ab]t - [(a+b)^2 - 2ab](2-t)\delta_{tr}\}\|Z^{(r)}\|_F^2 - 2abr\delta_{tr}\alpha^2(2-t) \le \Delta_{a,b}. \tag{2.19}$$

Proof. Note that the ranks of matrices $U_k,\ Z_{S_j}^{(r)}$ are no more than b, the ranks of matrices $V_k,\ Z_{T_i}^{(r)}$ are at most a and $a+b=tr$. By the tr-order restricted isometry property, we get

$$\begin{aligned}\Delta_{a,b} \geq & \frac{r-b}{a\binom{r}{a}} \sum_{i\in A,\ k} \mu_k \left[a^2(1-\delta_{tr}) \left\| Z_{T_i}^{(r)} + \frac{b}{r}U_k \right\|_F^2 - b^2(1+\delta_{tr}) \left\| Z_{T_i}^{(r)} - \frac{a}{r}U_k \right\|_F^2 \right] \\ & + \frac{r-a}{b\binom{r}{b}} \sum_{j\in B,\ k} \nu_k \left[b^2(1-\delta_{tr}) \left\| Z_{S_j}^{(r)} + \frac{b}{r}V_k \right\|_F^2 - a^2(1+\delta_{tr}) \left\| Z_{S_j}^{(r)} - \frac{b}{r}V_k \right\|_F^2 \right].\end{aligned}$$

Since the inner product of $Z_{T_i}^{(r)}$ ($Z_{S_j}^{(r)}$) and U_k (V_k) equals to zero, by some elementary calculation, we get

$$\begin{aligned}\Delta_{a,b} \geq & (a^2-b^2) \left[\frac{r-b}{a\binom{r}{a}} \sum_{i\in A} \|Z_{T_i}^{(r)}\|_F^2 - \frac{r-a}{b\binom{r}{b}} \sum_{j\in B} \|Z_{S_j}^{(r)}\|_F^2 \right] \\ & - (a^2+b^2)\delta_{tr} \left[\frac{r-b}{a\binom{r}{a}} \sum_{i\in A} \|Z_{T_i}^{(r)}\|_F^2 + \frac{r-a}{b\binom{r}{b}} \sum_{j\in B} \|Z_{S_j}^{(r)}\|_F^2 \right] \\ & - \frac{2ab^2(r-b)\delta_{tr}}{r^2} \sum_k \mu_k \|U_k\|_F^2 - \frac{2a^2b(r-a)\delta_{tr}}{r^2} \sum_k \nu_k \|V_k\|_F^2. \end{aligned} \tag{2.20}$$

By Lemma 2.1, we get

$$\sum_{i\in A} \|Z_{T_i}^{(r)}\|_F^2 = \tbinom{r-1}{a-1} \|Z^{(r)}\|_F^2, \tag{2.21}$$

and

$$\sum_{j\in B} \|Z_{S_j}^{(r)}\|_F^2 = \tbinom{r-1}{b-1} \|Z^{(r)}\|_F^2. \tag{2.22}$$

Substituting (2.21) and (2.22) into (2.20) and combining with inequalities (2.8) and (2.9), we get

$$\begin{aligned}\Delta_{a,b} \geq & \frac{(a-b)^2(a+b)}{r} \|Z^{(r)}\|_F^2 - (a^2+b^2)\delta_{tr}(2-t)\|Z^{(r)}\|_F^2 \\ & - \frac{2ab^2(r-b)\delta_{tr}}{r^2}\frac{r^2\alpha^2}{b} - \frac{2a^2b(r-a)\delta_{tr}}{r^2}\frac{r^2\alpha^2}{a} \\ & = \{[(a+b)^2-4ab]t - [(a+b)^2-2ab](2-t)\delta_{tr}\}\|Z^{(r)}\|_F^2 - 2abr\delta_{tr}\alpha^2(2-t).\end{aligned}$$

□

Lemma 2.9. *It holds that*

$$\|\mathcal{A}Z\|_2 \leq 2\epsilon \tag{2.23}$$

Proof. Due to the feasibility of X^*, we get

$$\|\mathcal{A}Z\|_2 = \|\mathcal{A}X - \mathcal{A}X^*\|_2 \leq \|\mathcal{A}X - b\|_2 + \|\mathcal{A}X^* - b\|_2 \leq 2\epsilon.$$

□

Lemma 2.10. *(Lemma 4.1 in [9]) For all linear maps $\mathcal{A}:\ \mathbb{R}^{m\times n}\to\mathbb{R}^q$ and $r\geq 2$, $s\geq 2$, we have*

$$\delta_{sr}\leq (2s-1)\delta_r. \tag{2.24}$$

Lemma 2.11. *It holds that for $0<t<1$,*

$$\begin{aligned}&\frac{\rho_{a,b}(t)}{\binom{r}{a}\binom{r-a}{b}}\sum_{T_i\bigcap S_j=\emptyset}\left[\left\|\mathcal{A}\left(Z_{T_i}^{(r)}+Z_{S_j}^{(r)}\right)\right\|_2^2-\frac{r-a-b}{abr}\left\|\mathcal{A}\left(bZ_{T_i}^{(r)}-aZ_{S_j}^{(r)}\right)\right\|_2^2\right]\\&\leq \rho_{a,b}(t)t[t-(2-t)\delta_{tr}]\|Z^{(r)}\|_F^2,\end{aligned} \tag{2.25}$$

and for $1\leq t<4/3$,

$$\begin{aligned}&\rho_{a,b}(t)\sum_k\tau_k\left[\left\|\mathcal{A}\left(Z^{(r)}+(t-1)W_k\right)\right\|_2^2-\left\|(t-1)\mathcal{A}\left(Z^{(r)}+W_k\right)\right\|_2^2\right]\\&\leq \rho_{a,b}(t)\left\{\left[t(2-t)-(t^2-2t+2)\delta_{tr}\right]\|Z^{(r)}\|_F^2-2r\alpha^2\delta_{tr}(t-1)\right\},\end{aligned} \tag{2.26}$$

where

$$\rho_{a,b}(t)=(a+b)^2-2ab(4-t).$$

Proof. We first consider the case of $0<t<1$. As tr equals to even, we can fix $a=b=tr/2$; And as tr equals to odd, we can set $a=b+1=(tr+1)/2$; For both cases, one can easily prove that $\rho_{a,b}(t)<0$. Since $Z_{T_i}^{(r)}$, $Z_{S_j}^{(r)}$ are a-rank and b-rank, respectively, by utilizing tr-order RIP, we get

$$\begin{aligned}&\frac{\rho_{a,b}(t)}{\binom{r}{a}\binom{r-a}{b}}\sum_{T_i\bigcap S_j=\emptyset}\left[\left\|\mathcal{A}\left(Z_{T_i}^{(r)}+Z_{S_j}^{(r)}\right)\right\|_2^2-\frac{r-a-b}{abr}\left\|\mathcal{A}\left(bZ_{T_i}^{(r)}-aZ_{S_j}^{(r)}\right)\right\|_2^2\right]\\&\leq\frac{\rho_{a,b}(t)}{\binom{r}{a}\binom{r-a}{b}}\sum_{T_i\bigcap S_j=\emptyset}\left[(1-\delta_{tr})\left\|Z_{T_i}^{(r)}+Z_{S_j}^{(r)}\right\|_F^2-\frac{r-a-b}{abr}(1+\delta_{tr})\left\|bZ_{T_i}^{(r)}-aZ_{S_j}^{(r)}\right\|_F^2\right]\\&=\frac{\rho_{a,b}(t)}{\binom{r}{a}\binom{r-a}{b}}\left\{(1-\delta_{tr})\left[\binom{r-a}{b}\sum_{i\in A}\|Z_{T_i}^{(r)}\|_F^2+\binom{r-b}{a}\sum_{j\in B}\|Z_{S_j}^{(r)}\|_F^2\right]\right.\\&\left.-\frac{1-t}{ab}(1+\delta_{tr})\left[b^2\binom{r-a}{b}\sum_{i\in A}\|Z_{T_i}^{(r)}\|_F^2+a^2\binom{r-b}{a}\sum_{j\in B}\|Z_{S_j}^{(r)}\|_F^2\right]\right\}\\&=\frac{\rho_{a,b}(t)}{\binom{r}{a}\binom{r-a}{b}}\left\{(1-\delta_{tr})\left[\binom{r-a}{b}\binom{r-1}{a-1}+\binom{r-b}{a}\binom{r-1}{b-1}\right]\|Z^{(r)}\|_F^2\right.\\&\left.-(1+\delta_{tr})\frac{1-t}{ab}\left[b^2\binom{r-a}{b}\binom{r-1}{a-1}+a^2\binom{r-b}{a}\binom{r-1}{b-1}\right]\|Z^{(r)}\|_F^2\right\}\\&=\rho_{a,b}(t)t[t-(2-t)\delta_{tr}]\|Z^{(r)}\|_F^2,\end{aligned} \tag{2.27}$$

where we made use of Lemma 2.1 to the second equality.

Next, we discuss the case of $1\leq t<4/3$.

Observe that $Z^{(r)}$, W_k are r-rank and $(t-1)r$-rank, respectively. Under the assumption of $\rho_{a,b}(t)<0$, combining with tr-order RIP, we get

$$\rho_{a,b}(t)\sum_k\tau_k\left[\left\|\mathcal{A}\left(Z^{(r)}+(t-1)W_k\right)\right\|_F^2-\left\|(t-1)\mathcal{A}\left(Z^{(r)}+W_k\right)\right\|_F^2\right]$$

$$
\begin{aligned}
&\leq \rho_{a,b}(t)\sum_k \tau_k\left[(1-\delta_{tr})\left\|Z^{(r)}+(t-1)W_k\right\|_F^2-(t-1)^2(1+\delta_{tr})\left\|Z^{(r)}+W_k\right\|_F^2\right]\\
&=\rho_{a,b}(t)\sum_k \tau_k\Bigg\{(1-\delta_{tr})\left[\|Z^{(r)}\|_F^2+(t-1)^2\|W_k\|_F^2\right]\\
&-(t-1)^2(1+\delta_{tr})\left(\|Z^{(r)}\|_F^2+\|W_k\|_F^2\right)\Bigg\}\\
&=\rho_{a,b}(t)\left\{\left[(1-\delta_{tr})-(t-1)^2(1+\delta_{tr})\right]\|Z^{(r)}\|_F^2-2\delta_{tr}(t-1)^2\sum_k \tau_k\|W_k\|_F^2\right\}\\
&\leq \rho_{a,b}(t)\left\{\left[t(2-t)-(t^2-2t+2)\delta_{tr}\right]\|Z^{(r)}\|_F^2-2r\alpha^2\delta_{tr}(t-1)\right\},
\end{aligned}
\tag{2.28}
$$

where the first equality follows from the fact that $\langle Z^{(r)}, W_k\rangle = 0$, and the last inequality, we used the inequality (2.10).

□

3 Main results

Theorem 3.1. *Consider rank minimization problem $b = \mathcal{A}X + z$ with $\|z\|_2 \leq \epsilon$. If $\delta_{tr} < t/(4-t)$ with $0 < t < 4/3$, then the solution X^* to the nuclear norm minimization problem (1.2) fulfils*

$$
\|X - X^*\|_F \leq C_1\epsilon + C_2\|X - X^{(r)}\|_*, \tag{3.1}
$$

where

$$
C_1 = \frac{2\sqrt{2(1+\delta_{tr})}\kappa}{\frac{t}{4-t}-\delta_{tr}}, \tag{3.2}
$$

and

$$
C_2 = \frac{2\sqrt{2}}{\sqrt{r}}\left\{\frac{1}{4}+\frac{2\delta_{tr}+\sqrt{\delta_{tr}(4-t)\left(\frac{t}{4-t}-\delta_{tr}\right)}}{\frac{t}{4-t}-\delta_{tr}}\right\} \tag{3.3}
$$

with

$$
\kappa = \max\left\{\frac{t}{4-t}, \frac{\sqrt{t}}{4-t}\right\}.
$$

Similarly, Consider rank minimization problem $b = \mathcal{A}X + z$ with z such that $\|\mathcal{A}^*(z)\| \leq \epsilon$. If $\delta_{tr} < t/(4-t)$ with $0 < t < 4/3$, then the solution X° to the nuclear norm minimization problem $\min_X\ \|X\|_*$ *s.t.* $\|\mathcal{A}^*(z)\| \leq \epsilon$ *fulfils*

$$
\|X - X^\circ\|_F \leq D_1\epsilon + C_2\|X - X^{(r)}\|_*, \tag{3.4}
$$

where

$$
D_1 = \frac{2\sqrt{2r}\kappa}{\frac{t}{4-t}-\delta_{tr}}, \tag{3.5}
$$

and C_2 is given by (3.3).

Remark 3.1. *As $t = 1$, the upper bound $\delta_r < 1/3$ is coincident with Theorems 3.7 and 3.8 of [9]. Furthermore, the upper bounds of error estimates $\|X - X^*\|_F$ ($\|X - X^\circ\|_F$) are smaller than the results of [9]. In theory, the recovered precision is given by our results is higher than that of theirs.*

Corollary 3.1. *Assume that $X \in \mathbb{R}^{m\times n}$ is a r-rank matrix. Let $b = \mathcal{A}X$. If*

$$\delta_{tr} < t/(4-t) \tag{3.6}$$

for $0 < t < 4/3$, then the solution X^* to the nuclear norm minimization problem (1.2) in the noiseless case (i.e., $\epsilon = 0$) reconstructs X exactly.

Remark 3.2. *As $t = 1$, the upper bound $\delta_r < 1/3$ is the same as Theorem 3.5 of [9].*

The Gaussian noise situation is of special interest in statistics and image processing. Note that the Gaussian random variables are essentially bounded. The results given in Theorem 3.1 regarding the bounded noise situation are immediately applied to the Gaussian noise situation, which employs the similar discussion as that in [16].

Theorem 3.2. *Assume that the low-rank recovery model $b = \mathcal{A}X + z$ with $z \sim N_q(0, \sigma^2 I)$. $\delta_{tr} < t/(4-t)$ for some $0 < t < 4/3$. Let X^* represent the minimizer of $\min_X \|X\|_*$ s.t. $\|z\|_2 \le \sigma\sqrt{q + 2\sqrt{q\log q}}$ and let X° be the minimizer of $\min_X \|X\|_*$ s.t. $\|\mathcal{A}^*(z)\| \le 2\sigma\sqrt{\log n}$. We have with probability at least $1 - 1/q$,*

$$\begin{aligned}\|X - X^*\|_F \le{}& \frac{2\sqrt{2(1+\delta_{tr})}\kappa}{\frac{t}{4-t} - \delta_{tr}}\sigma\sqrt{q + 2\sqrt{q\log q}}\\ &+ 2\sqrt{2}\left\{\frac{1}{4} + \frac{2\delta_{tr} + \sqrt{\delta_{tr}(4-t)\left(\frac{t}{4-t} - \delta_{tr}\right)}}{\frac{t}{4-t} - \delta_{tr}}\right\}\frac{\|X - X^{(r)}\|_*}{\sqrt{r}},\end{aligned}$$

and probability at least $1 - 1/\sqrt{\pi\log n}$,

$$\begin{aligned}\|X - X^\circ\|_F \le{}& \frac{4\sqrt{2r}\kappa}{\frac{t}{4-t} - \delta_{tr}}\sigma\sqrt{\log n}\\ &+ 2\sqrt{2}\left\{\frac{1}{4} + \frac{2\delta_{tr} + \sqrt{\delta_{tr}(4-t)\left(\frac{t}{4-t} - \delta_{tr}\right)}}{\frac{t}{4-t} - \delta_{tr}}\right\}\frac{\|X - X^{(r)}\|_*}{\sqrt{r}},\end{aligned}$$

where κ is defined in Theorem 3.1.

Theorem 3.3. *Let $1 \le r \le m/2$. There is a linear map $\mathcal{A}: \mathbb{R}^{m\times m} \to \mathbb{R}^q$ with $\delta_{tr} < t/(4-t) + \varepsilon$ with $0 < t < 4/3$, $\varepsilon > 0$ such that for some r-rank matrices Y_1, $Y_2 \in \mathbb{R}^{m\times m}$ with $Y_1 \neq Y_2$, $\mathcal{A}Y_1 = \mathcal{A}Y_1$. Hence, there don't exist any approach to exactly reconstruct all r-rank matrices X based on $(\mathcal{A}, z)$.*

Remark 3.3. *Theorems 3.1 and 3.3 jointly indicate the condition $\delta_{tr} < t/(4-t)$ with $0 < t < 4/3$ is sharp.*

4 Proofs of main results

With above preparation, we present the proof of main result.

Proof of Theorem 3.1. By the definition of α and notice that the rank of $Z^{(r)}$ is at most r, we get

$$\begin{aligned}\alpha^2 &= \frac{\|Z^{(r)}\|_*^2 + 4\|Z^{(r)}\|_*\|X - X^{(r)}\|_* + 4\|X - X^{(r)}\|_*^2}{r^2} \\ &\leq \frac{\|Z^{(r)}\|_F^2}{r} + \frac{4\|Z^{(r)}\|_F\|X - X^{(r)}\|_*}{r\sqrt{r}} + \frac{4\|X - X^{(r)}\|_*^2}{r^2},\end{aligned} \tag{4.1}$$

where in the last step, we used the fact that for any $X \in \mathbb{R}^{m\times n}$ $(m \leq n)$ and $p \in (0, 1]$,

$$m^{\frac{1}{p}-\frac{1}{2}}\|X\|_F \geq \|X\|_p \tag{4.2}$$

with $\|X\|_p = (\sum_i \sigma_i^p(X))^{1/p}$.

In the situation of $0 < t < 1$, by Lemma 2.10, we have

$$\begin{aligned}\left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z\right\rangle &\leq \|\mathcal{A}Z^{(r)}\|_2\|\mathcal{A}Z\|_2 \\ &\leq \sqrt{1+\delta_r}\|Z^{(r)}\|_F\|\mathcal{A}Z\|_2 \\ &= \sqrt{1+\delta_{\frac{1}{t}(tr)}}\|Z^{(r)}\|_F\|\mathcal{A}Z\|_2 \\ &\leq \sqrt{1+\left(\frac{2}{t}-1\right)\delta_{tr}}\|Z^{(r)}\|_F\|\mathcal{A}Z\|_2 \\ &\leq \sqrt{\frac{1+\delta_{tr}}{t}}\|Z^{(r)}\|_F\|\mathcal{A}Z\|_2,\end{aligned} \tag{4.3}$$

where in the first inequality, we used Cauchy-Schwarz inequality, and the second inequality follows from RIP of r-order.

Plugging (2.23) to (4.3), it follows that

$$\left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z\right\rangle \leq 2\epsilon\sqrt{\frac{1+\delta_{tr}}{t}}\|Z^{(r)}\|_F. \tag{4.4}$$

Combining with equation (2.13) and inequalities (2.19), (2.25) and (4.4), we have

$$\begin{aligned}&\rho_{a,b}(t)t[t-(2-t)\delta_{tr}]\|Z^{(r)}\|_F^2 + 4ab\epsilon t^2(2-t)\sqrt{\frac{1+\delta_{tr}}{t}}\|Z^{(r)}\|_F \\ &- t\left\{\{[(a+b)^2-4ab]t - [(a+b)^2-2ab](2-t)\delta_{tr}\}\|Z^{(r)}\|_F^2 - 2abr\delta_{tr}\alpha^2(2-t)\right\} \geq 0.\end{aligned}$$

Applying inequality (4.1) to above equality, we get

$$\begin{aligned}&2abt(t-2)\left[(4-t)\left(\frac{t}{4-t}-\delta_{tr}\right)\|Z^{(r)}\|_F^2\right. \\ &\left.-\left[2\epsilon\sqrt{(1+\delta_{tr})t} + \frac{4\delta_{tr}\|X-X^{(r)}\|_*}{\sqrt{r}}\right]\|Z^{(r)}\|_F - \frac{4\delta_{tr}\|X-X^{(r)}\|_*^2}{r}\right] \geq 0.\end{aligned} \tag{4.5}$$

In the situation of $1 \le t < 4/3$, due to the monotonicity of RIC δ_{tr}, it implies that

$$\begin{aligned}\left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z \right\rangle &\le \sqrt{1+\delta_r}\|Z^{(r)}\|_F\|\mathcal{A}Z\|_2 \\ &\le \sqrt{1+\delta_{tr}}\|Z^{(r)}\|_F\|\mathcal{A}Z\|_2 \\ &\le 2\epsilon\sqrt{1+\delta_{tr}}\|Z^{(r)}\|_F. \end{aligned} \tag{4.6}$$

It is easy to check that

$$\begin{aligned} ab &\ge \left(\frac{tr}{2}\right)^2 - \frac{1}{4} = \frac{(2-t)^2r^2-1}{4} - (1-t)r^2 \\ &> -(1-t)r^2. \end{aligned} \tag{4.7}$$

Combining with equation (2.14) and inequalities (2.19), (2.26) and (4.6), it holds that

$$\begin{aligned} &\rho_{a,b}(t)\left\{\left[t(2-t)-(t^2-2t+2)\delta_{tr}\right]\|Z^{(r)}\|_F^2 - 2r\alpha^2\delta_{tr}(t-1)\right\} \\ &+ 4\epsilon\sqrt{1+\delta_{tr}}t^3[ab-(t-1)r^2]\|Z^{(r)}\|_F \\ &-(4-3t)\left\{\{[(a+b)^2-4ab]t-[(a+b)^2-2ab](2-t)\delta_{tr}\}\|Z^{(r)}\|_F^2 - 2abr\delta_{tr}\alpha^2(2-t)\right\} \ge 0. \end{aligned} \tag{4.8}$$

Due to inequality (4.1), by fundamental calculation, we get

$$\begin{aligned} &2[(t-1)r^2-ab]t^2\left[(4-t)\left(\frac{t}{4-t}-\delta_{tr}\right)\|Z^{(r)}\|_F^2\right. \\ &\left.-\left[2\epsilon\sqrt{1+\delta_{tr}}t+\frac{4\delta_{tr}\|X-X^{(r)}\|_*}{\sqrt{r}}\right]\|Z^{(r)}\|_F - \frac{4\delta_{tr}\|X-X^{(r)}\|_*^2}{r}\right] \ge 0. \end{aligned} \tag{4.9}$$

Thereby, two second-order inequalities concerning $\|Z^{(r)}\|_F$ are established. Under the condition of $\delta_{tr} < t/(4-t)$, applying quadratic formula and some elementary compute, we have

$$\begin{aligned} \|Z^{(r)}\|_F \le &\frac{1}{2(4-t)(\frac{t}{4-t}-\delta_{tr})}\left[\frac{4\delta_{tr}\|X-X^{(r)}\|_*}{\sqrt{r}} + 2\epsilon\sqrt{1+\delta_{tr}}(4-t)\kappa\right. \\ &+\left[\left(\frac{4\delta_{tr}\|X-X^{(r)}\|_*}{\sqrt{r}}+2\epsilon\sqrt{1+\delta_{tr}}(4-t)\kappa\right)^2\right. \\ &\left.\left.+\frac{16\delta_{tr}\|X-X^{(r)}\|_*^2(4-t)}{r}\left(\frac{t}{4-t}-\delta_{tr}\right)\right]^{\frac{1}{2}}\right] \\ \le &\frac{1}{2(4-t)(\frac{t}{4-t}-\delta_{tr})}\left[\frac{8\delta_{tr}\|X-X^{(r)}\|_*}{\sqrt{r}} + 4\epsilon\sqrt{1+\delta_{tr}}(4-t)\kappa\right. \\ &\left.+\frac{4\|X-X^{(r)}\|_*}{\sqrt{r}}\sqrt{(4-t)(\frac{t}{4-t}-\delta_{tr})\delta_{tr}}\right] \\ = &\frac{2\sqrt{1+\delta_{tr}}\kappa}{\frac{t}{4-t}-\delta_{tr}}\epsilon \\ &+\frac{4\delta_{tr}+2\sqrt{(4-t)(\frac{t}{4-t}-\delta_{tr})\delta_{tr}}}{(4-t)(\frac{t}{4-t}-\delta_{tr})\delta_{tr}}\frac{\|X-X^{(r)}\|_*}{\sqrt{r}} \end{aligned} \tag{4.10}$$

with κ is defined in Theorem 3.1, where the second inequality follows from the fact that for any vector $x \in \mathbb{R}^n$, $\|x\|_2 \leq \|x\|_1$.

Then,

$$\begin{aligned}
\|Z_c^{(r)}\|_F &= \left(\sum_{i \geq r+1} \sigma_i^2(U^\top Z V) \right)^{1/2} \\
&\leq \left(\max_{i \geq r+1} \{\sigma_i(U^\top Z V)\} \sum_{i \geq r+1} \sigma_i(U^\top Z V) \right)^{1/2} \\
&= \|Z_c^{(r)}\|^{1/2} \|Z_c^{(r)}\|_*^{1/2} \\
&\leq \frac{\|Z^{(r)}\|_*^{1/2}}{\sqrt{r}} (\|Z^{(r)}\|_* + 2\|X - X^{(r)}\|_*)^{1/2} \\
&\leq \left(\|Z^{(r)}\|_F^2 + \frac{2\|X - X^{(r)}\|_* \|Z^{(r)}\|_F}{\sqrt{r}} \right)^{1/2}, \qquad (4.11)
\end{aligned}$$

where in the second inequality, we used Lemma 2.4, and the third inequality follows from the fact that for any r-rank matrix X, $\|X\|_* \leq \sqrt{r}\|X\|_F$.

A combination of (4.10) and (4.11) implies that

$$\begin{aligned}
\|Z\|_F &= \left(\|Z_c^{(r)}\|_F^2 + \|Z^{(r)}\|_F^2 \right)^{1/2} \\
&\leq \left(2\|Z^{(r)}\|_F^2 + \frac{2\|X - X^{(r)}\|_* \|Z^{(r)}\|_F}{\sqrt{r}} \right)^{1/2} \\
&\leq \sqrt{2}\|Z^{(r)}\|_F + \frac{\|X - X^{(r)}\|_*}{\sqrt{2r}} \\
&\leq \frac{2\sqrt{2(1+\delta_{tr})}\kappa\epsilon}{\frac{t}{4-t} - \delta_{tr}} \\
&\quad + \frac{2\sqrt{2}}{\sqrt{r}} \left[\frac{1}{4} + \frac{2\delta_{tr} + \sqrt{(4-t)(\frac{t}{4-t} - \delta_{tr})\delta_{tr}}}{\frac{t}{4-t} - \delta_{tr}} \right] \|X - X^{(r)}\|_*.
\end{aligned}$$

In the situation of the error bound $\|\mathcal{A}^*(z)\| \leq \epsilon$, set $Z = X - X^\circ$. It holds that

$$\begin{aligned}
\|\mathcal{A}^*\mathcal{A}Z\| &= \|\mathcal{A}^*(\mathcal{A}X - b) - \mathcal{A}^*(\mathcal{A}X^\circ - b)\| \\
&\leq \|\mathcal{A}^*(\mathcal{A}X - b)\| + \|\mathcal{A}^*(\mathcal{A}X^\circ - b)\| \\
&\leq 2\epsilon.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\left\langle \mathcal{A}Z^{(r)}, \mathcal{A}Z \right\rangle &= \left\langle Z^{(r)}, \mathcal{A}^*\mathcal{A}Z \right\rangle \\
&\leq \|Z^{(r)}\|_* \cdot 2\epsilon \\
&\leq 2\epsilon\sqrt{r}\|Z^{(r)}\|_F.
\end{aligned}$$

The rest of steps are similar with the situation of the error bound $\|z\|_2 \leq \epsilon$. The proof of Theorem 3.1 is completed.

$\square$

Proof of Theorem 3.3. Let $E = \text{diag}(x) \in \mathbb{R}^{m\times m}$ with

$$x = \frac{1}{\sqrt{2r}}(\underbrace{1, \cdots, 1}_{2r}, 0, \cdots, 0).$$

Define $\mathcal{A}: \mathbb{R}^{m\times m} \to \mathbb{R}^q$ as

$$\mathcal{A}X = \frac{2}{\sqrt{4-t}} \left(\sigma(X) - \langle \sigma(X), \sigma(E)\rangle \sigma(E)\right).$$

Applying the Cauchy-Schwarz inequality, for all $\lceil tr \rceil$-rank matrices X, we get

$$\begin{aligned} |\langle \sigma(X), \sigma(E)\rangle| &\leq \|\sigma(X)\|_2 \|\sigma(E) \cdot 1_{\sup(\sigma(X))}\|_2 \\ &\leq \sqrt{\frac{\lceil tr \rceil}{2r}} \|X\|_F, \end{aligned}$$

and

$$\begin{aligned} \|\mathcal{A}X\|_2^2 &= \frac{4}{4-t} \left\langle \sigma(X) - \langle \sigma(X), \sigma(E)\rangle \sigma(E), \sigma(X) - \langle \sigma(X), \sigma(E)\rangle \sigma(E)\right\rangle \\ &= \frac{4}{4-t} \left[\|X\|_F^2 - |\langle \sigma(X), \sigma(E)\rangle|^2\right]. \end{aligned}$$

Therefore,

$$\begin{aligned} \|\mathcal{A}X\|_2^2 &\leq \left(1 + \frac{t}{4-t}\right) \|X\|_F^2 \\ &\leq \left(1 + \frac{t}{4-t} + \varepsilon\right) \|X\|_F^2. \end{aligned} \tag{4.12}$$

For $r > 1/\varepsilon$, we get

$$\begin{aligned} \|\mathcal{A}X\|_2^2 &\geq \frac{4}{4-t}\left(1 - \frac{\lceil tr \rceil}{2r}\right) \|X\|_F^2 \\ &\geq \frac{4}{4-t}\left(1 - \frac{tr}{2r} - \frac{1}{2r}\right) \|X\|_F^2 \\ &\geq \frac{4}{4-t}\left(1 - \frac{tr}{2r} - \frac{\varepsilon}{2}\right) \|X\|_F^2 \\ &\geq \left(1 - \frac{t}{4-t} - \varepsilon\right) \|X\|_F^2. \end{aligned}$$

Accordingly, by Definition 1.1, we obtain $\delta_{tr} = \delta_{\lceil tr \rceil} = \frac{t}{4-t} + \varepsilon$. Suppose $Y_1 = \text{diag}(y_1)$, $Y_2 = \text{diag}(y_2) \in \mathbb{R}^{m\times m}$ with

$$y_1 = (\underbrace{1, \cdots, 1}_{r}, 0, \cdots, 0)$$

and

$$y_2 = (\underbrace{0, \cdots, 0}_{r}, \underbrace{-1, \cdots, -1}_{r}, 0, \cdots, 0).$$

It is easy to verify that Y_1 and Y_2 are both matrices of rank r such that $Y_1 - Y_2 \in \mathcal{N}(\mathcal{A})$, i.e., $\mathcal{A}Y_1 = \mathcal{A}Y_2$. Consequently, it is not possible to reconstruct both Y_1 and Y_2 based on $(z, \mathcal{A})$.

$\square$

5 Conclusions

In this paper, we establish sufficient conditions which ensure the stable recovery or exactly recovery of any r-rank matrix satisfying a given linear system of equality constraints via solving a convex optimization problem, i.e., nuclear norm minimization. When the parameter t is equal to 1, the bound of RIC δ_r coincide with the result of [9]. Meanwhile, the derived upper bounds regarding the reconstruction error are better than those of [9]. Besides, the restricted isometry property condition is proved sharp.

Acknowledgments

This work was supported by Natural Science Foundation of China (No. 61673015, 61273020, 11661084), Fundamental Research Funds for the Central Universities (No. XDJK2015A007, SWU1809002), Science Computing and Intelligent Information Processing of GuangXi higher education key laboratory (No. GXSCIIP201702), Science and Technology Plan Project of Guizhou Province(No. LH[2015]7053, LH[2015]7055), 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. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. SIAM Rev., 14: 1-46 (2009)

[2] Tanner J, Wei K. Low rank matrix completion by alternating steepest descent methods. Applied & Computational Harmonic Analysis, 2016, 40(2):417-429.

[3] Zhang N, Li Q. On optimal solutions of the constrained ℓ_0 regularization and its penalty problem. Inverse Problems, 2016, 33(2).

[4] Lu C, Tang J, Yan S, et al. Nonconvex Nonsmooth Low-Rank Minimization via Iteratively Reweighted Nuclear Norm. IEEE Trans Image Process, 2015, 25(2):829-839.

[5] Lin Z, Xu C, Zha H. Robust Matrix Factorization by Majorization Minimization. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2018, 40(1):208-220.

[6] Candes, E.J. The restricted isometry property and its implications for compressed sensing. C. R. Acad. Sci. Paris. Ser.1, 346: 589-592 (2008)

[7] Candes E J, Plan Y. Tight Oracle Inequalities for Low-Rank Matrix Recovery From a Minimal Number of Noisy Random Measurements. IEEE Transactions on Information Theory, 2011, 57(4):2342-2359.

[8] Mohan K, Fazel M. New restricted isometry results for noisy low-rank recovery. In: IEEE International Symposium on Information Theory Proceedings. Seattle: IEEE, 2010, 1573-1577

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

[10] Wang H M, Li S. The bounds of restricted isometry constants for low rank matrices recovery. Sci China Math, 2013, 56: 1117-1127

[11] Kong L C, Xiu N H. Exact low-rank matrix recovery via nonconvex Schatten p-minimization. Asia-Pac J Oper Res, 2013, 30: 1-13

[12] Chen W G, Li Y L. Stable recovery of low-rank matrix via nonconvex Schatten p-minimization. Science China Mathematics, 2015, 58(12):2643-2654.

[13] Cai T T, Zhang A. Sparse representation of a polytope and recovery of sparse signals and low-rank matrices. IEEE Trans Inform Theory, 2014, 60: 122-132

[14] R. Zhang, S. Li, "A proof of conjecture on restricted isometry property constants δ_{tk} $(0 < t < 4/3)$," *IEEE T. Inform. Theory*, vol. 64, no. 3, pp. 1699-1705, Mar. 2018.

[15] Zhang R, Li S. Optimal RIP Bounds for Sparse Signals Recovery via l_p Minimization. Applied & Computational Harmonic Analysis, 2018, doi.org/10.1016/j.acha.2017.10.004

[16] Cai T T, Wang L, Xu G. Shifting Inequality and Recovery of Sparse Signals. IEEE Transactions on Signal Processing, 2010, 58(3): 1300-1308.