Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 11, No. 2, pp. 365–383, April 2015

CONDITIONS FOR THE EQUIVALENCE BETWEEN THE LOW-n-RANK TENSOR RECOVERY PROBLEM AND ITS CONVEX RELAXATION*

Min Zhang and Zheng-Hai Huang†

Abstract: Recently, by using the n-rank of a tensor, the low-rank matrix recovery problem has been generalized to the higher order, named as the low-n-rank tensor recovery problem. To solve this problem, one popular approach is to relax it to a tractable convex form; and then, design algorithms to resolve the convex one. Lately, several efficient algorithms have been presented, however, few articles have investigated the conditions that guarantee the equivalence between the low-n-rank tensor recovery and its convex relaxation so far. In this paper, we discuss three conditions for the equivalence, which are the null space property, the s-n-good condition and the restricted isometry property; and discuss the relationship among these three conditions. In addition, we use the quantity of the s-n-goodness to measure the error bound of the convex relaxation to the original tensor problem.

Key words: *low-n-rank tensor recovery; null space property; s-n-good condition; restricted isometry property*

Mathematics Subject Classification: *62B10, 90C26, 90C59*

1 Introduction

Since the year 2006, a new theory of Compressed Sensing (CS) has emerged, which became a hot issue in the field of signal processing, machine learning and optimization. It shows that sparse or compressible signals can be reconstructed from a surprisingly small number of linear measurements. The mathematical model of CS, called sparse signal recovery, is listed as follows:

$$\min_{x\in\mathbb{R}^n} \|x\|_0 \quad \text{s.t. } Ax = Aw, \tag{SSR}$$

where $w \in \mathbb{R}^n$ is the real signal to be recovered, $x \in \mathbb{R}^n$ is the decision variable, $\|x\|_0$ is the cardinality of x, $A \in \mathbb{R}^{M\times n}$ is the given measurement matrix, and $Aw \in \mathbb{R}^M$ is the given measurement vector (see, e.g., [1,6,7,9]). Unfortunately, problem (SSR) is provably NP-hard ([10, 24]). A well-known tractable heuristic for problem (SSR) is the l_1-norm minimization relaxation:

$$\min_{x\in\mathbb{R}^n} \|x\|_1 \quad \text{s.t. } Ax = Aw, \tag{l_1-min}$$

*This work was supported in part by the National Natural Science Foundation of China (No. 11171252, No. 11201332 and No. 11431002).

†Corresponding author.

© 2015 Yokohama Publishers

where $\|x\|_1 = \sum_{i=1}^{n} |x_i|$ is the l_1-norm of x. Problem (SSR) is generalized to the matrix space by utilizing the rank of a matrix instead of $\|\cdot\|_0$, named as low-rank matrix recovery:

$$\min_{X \in \mathbb{R}^{n_1 \times n_2}} \operatorname{rank}(X) \quad \text{s.t. } \mathcal{A}(X) = \mathcal{A}(W), \tag{LMR}$$

where $W \in \mathbb{R}^{n_1 \times n_2}$ is the original low-rank matrix to be recovered, and $\mathcal{A} : \mathbb{R}^{n_1 \times n_2} \to \mathbb{R}^M$ is a given linear transformation. Analogous to problem (SSR), problem (LMR) is also NP-hard ([8,21]) and its popular convex relaxation was proposed by Fazel, Hindi and Boyd [11], named as nuclear norm minimization:

$$\min_{X \in \mathbb{R}^{n_1 \times n_2}} \|X\|_* \quad \text{s.t. } \mathcal{A}(X) = \mathcal{A}(W), \tag{NNM}$$

where $\|X\|_*$ is the nuclear norm of the matrix X, i.e., the sum of its singular values.

It has been shown that if the measurement matrix A (respectively, the linear transformation $\mathcal{A}$) in the constraints satisfies some special conditions, which are called exact recovery conditions, one can obtain a solution to the original problem (SSR) (respectively, problem (LMR)) via its relaxation (l_1-min) (respectively, problem (NNM)). In other words, problems (SSR) and (LMR) are equivalent to their relaxations (l_1-min) and (NNM), respectively. Particularly, for Gaussian random matrices or transformations, these conditions can be met with high probability whenever the number of the measurements is large enough. There are mainly three kinds of the exact recovery conditions: the null space property, the s-goodness condition and the restricted isometry property.

One exact recovery condition is the null space property, which has been studied in the last years [25, 26, 31, 32]. The null space property is a necessary and sufficient condition for the solution of problem (l_1-min) (respectively, problem (NNM)) to coincide with the solution of problem (SSR) (respectively, problem (LMR)) in an affine space. It characterizes a particular property of the null-space of the linear map.

s-goodness condition is the other one exact recovery condition. Juditsky and Nemirovski [14] first established necessary and sufficient conditions for a sensing matrix to be "s-good" to allow for the equivalence between problem (l_1-min) and problem (SSR), which comes from the optimal condition in term of the optimization technique. Then, based on the singular value decomposition of a matrix and the partition technique, Kong et al. [17, 18] extended s-goodness conditions to the matrix case. They also demonstrated that these characteristics lead to verifiable sufficient conditions for exact s-rank matrix recovery and to computable upper bounds on those s, for which a given linear transformation is s-good.

The third popular exact recovery condition is the restricted isometry property (RIP). Restricted isometry constant δ_r of a matrix was first introduced by Candés and Tao [6]. Further works for RIP in CS were studied in [2,5,7,9]. When it comes to the matrix case, the RIP condition has been extended from those results in CS [4,15,19,22,30]. Recently, Oymak et al. [27] used a key singular value inequality to extend several classes of recovery conditions, from vectors to matrices in a simple and transparent way from a broad perspective. And lately, Cai and Zhang [3] established sharp RIP conditions for sparse signal and low-rank matrix recovery. The RIP condition which guarantees the equivalence between problem (LMR) and its nonconvex relaxation was investigated in [16, 39].

In modern world, the structure of data we encounter becomes more and more complex. Compared with the vector and matrix representations, utilizing tensors to represent the signals or images with high dimensional data can be more convenient for modeling and dealing with in many specific problems. Hence, some researchers begin to study the similar low-rank tensor recovery problem. In 2009, Liu et al. [20] first introduced the trace norm for tensors and proposed a method for solving the low-rank tensor completion problem. Later, by using the n-rank of a tensor, Gandy et al. [12] considered the low-n-rank tensor recovery, denoted by LTR, which is more general than the one in [20]. They introduced a tractable convex relaxation of the n-rank model and proposed efficient algorithms to solve this problem numerically. More recently, several efficient methods for solving LTR were proposed, for example, tensor-HC method [33], fixed point iterative method [37], splitting augmented Lagrangian method [36], hard thresholding iterative algorithm [40], and other methods [34, 35]. However, there are little investigation on the theory of the successful recovery. In the recent years, researchers have studied how many random measurements are needed to guarantee the successful recovery by LTR and its convex relaxation, and obtained some amazing results [13, 23, 29, 35]. To some extent, their work gave theoretical guidance to the application. However, there are few studies on the equivalence between LTR and its convex relaxation. In this paper, we will focus on this issue and explore what conditions can guarantee the equivalence. In 2012, we discussed this topic and obtained some preliminary results [38]; and this paper is the latest revision.

The organization of this paper is as follows. In the next section, we first introduce some preliminary knowledge, then propose the model of LTR and its convex relaxation. In Sections 3, we discuss the null space property for the equivalence between LTR and its relaxation. The s-n-good condition is explored in Section 4, and we also use the quantity of the s-n-goodness to measure the distance between the original tensor to be recovered and the υ-optimal solution of the convex relaxation problem. In Section 5, we establish a sharp RIP condition, which is the extension of that in [3]. The relationship among the three exact recovery conditions are given in Section 6; and conclusions are given in the last section.

2 Preliminaries

Throughout this paper, we denote the tensor space by $\mathbf{T} := \mathbb{R}^{n_1\times n_2\times\ldots\times n_N}$ and an N-th order tensor by $\mathcal{X} \in \mathbf{T}$. The inner product of two tensors is defined as

$$\langle \mathcal{X}, \mathcal{Y}\rangle = \sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2}\cdots\sum_{j_N=1}^{n_N} x_{j_1,j_2,\ldots,j_N} y_{j_1,j_2,\ldots,j_N}$$

and the corresponding Frobenius-norm is $\|\mathcal{X}\|_F = \sqrt{\langle \mathcal{X}, \mathcal{X}\rangle}$. In particular, when $N = 2$, the Frobenius norm of a matrix $X \in \mathbb{R}^{n_1\times n_2}$ is $\|X\|_F := \sqrt{\mathrm{trace}(X^T X)}$.

We will use the notations of fibers and unfoldings given in [12]. The mode-i fibers are all vectors $x_{j_1\ldots j_{i-1}:j_{i+1}\ldots j_N}$ obtained by fixing the indexes of $\{j_1, \ldots j_N\}\backslash j_i$, which are analogue of matrix rows and columns. The mode-i unfolding of $\mathcal{X} \in \mathbf{T}$, denoted by $\mathcal{X}_{<i>}$, arranges the mode-i fibers to be the columns of the resulting matrix. The tensor element $(j_1, j_2, \ldots, j_N)$ is mapped to the matrix element (j_i, l), where $l = 1 + \sum_{k=1,\, k\neq i}^{N}(j_k - 1)L_k$ with $L_k = \prod_{j=1,\, j\neq i}^{k-1} n_j$, which infers $\mathcal{X}_{<i>} \in \mathbb{R}^{n_i\times J_i}$, where $J_i = \prod_{k=1,k\neq i}^{N} n_k$. The n-rank of a tensor $\mathcal{X} \in \mathbf{T}$ is defined by

$$n\text{-}rank(\mathcal{X}) = rank(\mathcal{X}_{<1>}), (rank\mathcal{X}_{<2>}), \ldots, (rank\mathcal{X}_{<N>}).$$

The mode-i product of a tensor $\mathcal{X} \in \mathbf{T}$ by a matrix $U \in \mathbb{R}^{J_i\times n_i}$, denoted by $\mathcal{X} \times_i U$, is a $(n_1 \times \cdots \times n_{i-1} \times J_i \times n_{i+1} \times \cdots \times n_N)$-tensor of which the entries are given by $(\mathcal{X} \times_i U)_{k_1\ldots k_{i-1}j_ik_{i+1}\ldots k_N} = \sum_{k_i} x_{k_1\ldots k_{i-1}k_ik_{i+1}\ldots k_N} u_{j_ik_i}$.

It is easy to see that the mode-i unfolding of a tensor $\mathcal{X} \in \mathbf{T}$ can be expressed as a linear operator. Hence, for any $i \in \{1, 2, \ldots, N\}$, we define a linear operator by

$$\mathscr{B}_i : \mathbf{T} \to \mathbb{R}^{n_i\times J_i} \ \text{ with } \ \mathscr{B}_i(\mathcal{X}) := \mathcal{X}_{<i>}.$$

It is not difficult to find that the conjugate of $\mathscr{B}_i$ is the refolding of a matrix into a tensor by the mode-i for any $i \in \{1, 2, \dots, N\}$.

Proposition 2.1. *For any $i \in \{1, 2, \dots, N\}$, by defining the refolding of a matrix into a tensor by the mode-i to a linear transformation $\mathcal{G}_i$, i.e., $\mathcal{G}_i(\mathcal{X}_{<i>}) = \mathcal{X}$, we have $\mathscr{B}_i^* = \mathcal{G}_i$, where $\mathscr{B}_i^*$ denotes the conjugate of $\mathscr{B}_i$.*

Utilizing Proposition 2.1, we have the following result directly.

Lemma 2.2. *Let $\mathscr{I} : \mathbf{T} \to \mathbf{T}$ be the identity operator in the Tensor Hilbert space, then $\mathscr{B}_i^* \circ \mathscr{B}_i = \mathscr{I}$, where $\circ$ denotes the composition of two operators.*

Now, by extending the definition of an r-rank matrix, we try to character a tensor with low n-rank.

Definition 2.3. We say a tensor $\mathcal{X}$ is r-n-rank, if its every mode-i unfolding $\mathscr{B}_i(\mathcal{X})$ is an r-rank matrix, i.e., $rank(\mathscr{B}_i(\mathcal{X})) \leq r$.

Next, we consider the model of low-rank tensor recovery. Suppose that $\mathcal{W}$ is the true data to be restored and it has low n-rank, one can observe the linear measurements $\mathscr{A}(\mathcal{W})$, where $\mathscr{A} : \mathbf{T} \to \mathbb{R}^M$ is a given linear transformation. The target is to recover the true data $\mathcal{W}$ from the known facts. Replacing $rank(\cdot)$ of a matrix by the n-rank of a tensor, the model of LTR is extended from problem (LMR) naturally:

$$\min_{\mathcal{X} \in \mathbf{T}} \quad n\text{-}rank(\mathcal{X}) \quad \text{s.t. } \mathscr{A}(\mathcal{X}) = \mathscr{A}(\mathcal{W}). \tag{LTR}$$

This model is also proposed in [23] and they illustrated this model can successfully recover $\mathcal{W}$ with probability one whenever the number of measurements M is large enough, i.e., problem (LTR) recovers every r-n-rank tensor $\mathcal{X}$ with probability one whenever $M \geq (2r)^N + 2rnN + 1$. Also, Mu et. al. defined the recoverability of the original data $\mathcal{W}$ by problem (LTR) in [23], which we will use in the consequent discussions.

Definition 2.4. [Definition 1 in [23]] We call $\mathcal{W}$ recoverable by problem (LTR) if the following set is empty:

$$\{\mathcal{X}' \neq \mathcal{W} \mid \mathscr{A}(\mathcal{X}') = \mathscr{A}(\mathcal{W}),\ n\text{-}rank(\mathcal{X}') \preceq_{\mathbb{R}_+^N} n\text{-}rank(\mathcal{W})\}.$$

Problem (LTR) is a difficult non-convex multiple objective programming, however, it has a convex relaxation, which is much more tractable:

$$\min_{\mathcal{X}} \frac{1}{N} \sum_{i=1}^{N} \|\mathscr{B}_i(\mathcal{X})\|_* \quad \text{s.t. } \mathscr{A}(\mathcal{X}) = \mathscr{A}(\mathcal{W}). \tag{LTP}$$

It is also proved that problem (LTP) recovers the underlying tensor $\mathcal{W}$ when the number of measurements M is sufficiently large [35].

In the following, we will explore what conditions of $\mathscr{A}$ can guarantee the equivalence between problems (LTR) and (LTP), including the null space property, the s-n-goodness condition and the restricted isometry property. And the equivalence we say means that an r-n-rank tensor $\mathcal{W}$ is the unique minimizer of both problem (LTP) and problem (LTR).

3 Null Space Property

Null space property (NSP) is one basic condition for the equivalence between problems (LMR) and (NNM). The following lemma is one of the classic results:

Lemma 3.1 (Lemma 6 in [25])**.** *One can recover any r-rank matrix W via problem* (NNM) *if and only if the NSP holds, i.e., $2\|W\|_r < \|W\|_*$ for all $W \in Null(\mathcal{A})\backslash\{0\}$, where $\|W\|_r$ is the sum of the r largest singular values of matrix W.*

Now, we extend the above result to the tensor case.

Theorem 3.2. *Let $\mathscr{A} : \mathbf{T} \to \mathbb{R}^M$ be a linear transformation. If*

$$\frac{1}{N}\sum_{i=1}^{N}(\|\mathscr{B}_i(\mathcal{Y})\|_* - 2\|\mathscr{B}_i(\mathcal{Y})\|_r) > 0 \tag{3.1}$$

holds for any $\mathcal{Y} \in Null(\mathscr{A})\backslash\{0\}$, then one can recover any r-n-rank tensor $\mathcal{W}$ by solving problem (LTP)*.*

Proof. It is obvious that $\mathcal{W}+\mathcal{Y}$ is a feasible point of problem (LTP) for any $\mathcal{Y} \in Null(\mathscr{A})\backslash\{0\}$, where $Null(\mathscr{A}) := \{\mathcal{X} : \mathscr{A}(\mathcal{X}) = 0\}$.

If $\mathcal{W} \in \mathbf{T}$ is an r-n-rank tensor, then $\mathscr{B}_i(\mathcal{W})$ is r-rank for every $i \in \{1, 2, \ldots, N\}$. Thus, by the proof of Lemma 6 in [25], we have

$$\|\mathscr{B}_i(\mathcal{W}+\mathcal{Y})\|_* \geq \|\mathscr{B}_i(\mathcal{W})\|_* + \|\mathscr{B}_i(\mathcal{Y})\|_* - 2\|\mathscr{B}_i(\mathcal{Y})\|_r. \tag{3.2}$$

Furthermore, combining (3.2) with (3.1), we obtain

$$\frac{1}{N}\sum_{i=1}^{N}\|\mathscr{B}_i(\mathcal{W}+\mathcal{Y})\|_* > \frac{1}{N}\sum_{i=1}^{N}\|\mathscr{B}_i(\mathcal{W})\|_*. \tag{3.3}$$

Hence, $\mathcal{W}$ is the unique solution to problem (LTP). □

Note that Theorem 3.2 gives a sufficient condition for a low n-rank tensor to be the unique optimal solution of problem (LTP). However, it is not clear whether this low n-rank tensor is recoverable by problem (LTR) under the null space property (3.1). The following theorem is to give an answer, and we can see that it needs more strict condition to guarantee the equivalence between problems (LTR) and (LTP).

Theorem 3.3. *Let $\mathscr{A} : \mathbf{T} \to \mathbb{R}^M$ be a linear transformation and $\mathcal{W} \in \mathbf{T}$ be an r-n-rank tensor. If for every $i \in \{1, 2, \ldots, N\}$,*

$$\|\mathscr{B}_i(\mathcal{Y})\|_* - 2\|\mathscr{B}_i(\mathcal{Y})\|_r > 0 \tag{3.4}$$

holds for any $\mathcal{Y} \in Null(\mathscr{A})\backslash\{0\}$, then $\mathcal{W}$ is the unique optimal solution to problem (LTP)*. What's more, $\mathcal{W}$ is recoverable by problem* (LTR)*. In other words, problems* (LTR) *and* (LTP) *have the same unique minimizer.*

Proof. First, it's easy to see that (3.4) is more strict than (3.1). Thus, $\mathcal{W}$ is the unique solution to problem (LTP) since it's an r-n-rank tensor.

Second, to prove $\mathcal{W}$ is recoverable by problem (LTR), we should prove that

$$\{\mathcal{X}' \neq \mathcal{W} \mid \mathscr{A}(\mathcal{X}') = \mathscr{A}(\mathcal{W}),\ n\text{-}rank(\mathcal{X}') \preceq_{\mathbb{R}^N_+} n\text{-}rank(\mathcal{W})\} = \emptyset.$$

Suppose that there exists an $\mathcal{X}' \neq \mathcal{W}$ satisfying that $\mathscr{A}(\mathcal{X}') = \mathscr{A}(\mathcal{W})$ and $n\text{-}rank(\mathcal{X}') \preceq_{\mathbb{R}^N_+} n\text{-}rank(\mathcal{W})$, which means for every $i \in \{1, 2, \dots, N\}$,

$$\mathscr{A} \circ \mathscr{B}_i^*(\mathscr{B}_i(\mathcal{X}')) = \mathscr{A} \circ \mathscr{B}_i^*(\mathscr{B}_i(\mathcal{W})) \quad \text{and} \quad rank(\mathscr{B}_i(\mathcal{X}')) \leq rank(\mathscr{B}_i(\mathcal{W})) \leq r.$$

Note that for every $i \in \{1, 2, \dots, N\}$, it's easy to verify that

$$Null(\mathscr{A} \circ \mathscr{B}_i^*) = \{\mathscr{B}_i(\mathcal{Y}) \ : \ \mathcal{Y} \in Null(\mathcal{A})\}.$$

So, based on Lemma 3.1, if (3.4) holds, then for every $i \in \{1, 2, \dots, N\}$, both $\mathscr{B}_i(\mathcal{W})$ and $\mathscr{B}_i(\mathcal{X}')$ are the optimal solutions to the following problem:

$$\min_X \|X\|_* \quad \text{s.t.} \ (\mathscr{A} \circ \mathscr{B}_i^*)(X) = (\mathscr{A} \circ \mathscr{B}_i^*)(\mathscr{B}_i(\mathcal{W})),$$

but $\mathscr{B}_i(\mathcal{W}) \neq \mathscr{B}_i(\mathcal{X}')$. This contradicts that the above problem has one unique minimizer. □

4 s-n-Good Condition and the Error Bound

4.1 s-n-good condition

Based on the convex optimization technique, s-good condition has been proposed to be a sufficient and necessary condition for the equivalence between problems (NNM) and (LMR) [17, 18]. Given an integer $1 \leq s \leq \min\{n_1, n_2\}$, a linear transformation $\mathcal{A} : \mathbb{R}^{n_1 \times n_2} \to \mathbb{R}^M$ is called s-good if for every s-rank matrix $W \in \mathbb{R}^{n_1 \times n_2}$, W is the unique optimal solution to the nuclear norm optimization (NNM), i.e., one can recover any s-rank matrix W by problem (NNM). Furthermore, if $\mathcal{A}$ is s-good, then W is the unique solution to problem (LMR), i.e., the solution to problem (LMR) can be exactly recovered from problem (NNM).

Now, we try to explore a generalization of the s-good condition to the tensor case. First, we write the Lagrangian function of problem (LTP):

$$L(\mathcal{X}, y) = \frac{1}{N} \sum_{i=1}^{N} \|\mathscr{B}_i(\mathcal{X})\|_* - y^T(\mathscr{A}(\mathcal{X}) - \mathscr{A}(\mathcal{W})).$$

Note that $\mathcal{W}$ is the optimal solution to problem (LTP) if and only if

$$\mathcal{O} \in \partial_{\mathcal{X}} L(\mathcal{X}, y)|_{\mathcal{X}=\mathcal{W}},$$

where $\mathcal{O} \in \mathbb{T}$ with its entries being all zeros, and $\partial_{\mathcal{X}} L$ is the subdifferentiation of Lagrangian function $L(\cdot)$ with respect to $\mathcal{X}$. Let $g_i(\mathcal{X}) = \|\mathscr{B}_i(\mathcal{X})\|_*$, which can be viewed as the composition of the operator $\|\cdot\|_*$ and the operator $\mathscr{B}_i$. Thus, we have

$$\partial_{\mathcal{X}} g_i(\mathcal{X}) = \mathscr{B}_i^* \partial_{\mathscr{B}_i(\mathcal{X})} \|\mathscr{B}_i(\mathcal{X})\|_* \quad \text{and} \quad \partial_{\mathcal{X}} L(\mathcal{X}, y) = \frac{1}{N} \sum_{i=1}^{N} \partial_{\mathcal{X}} g_i(\mathcal{X}) - \mathscr{A}^* y.$$

As it is known, for a matrix $Z \in \mathbb{R}^{n_1 \times n_2}$ with its SVD being $Z = U_{n_1 \times s} Z_s V_{n_2 \times s}^T$, where the columns of $U_{n_1 \times s}$ (respectively, $V_{n_2 \times s}$) are orthogonal, $Z_s = Diag(\sigma(Z))$ and $\sigma(z) = (\sigma_1(Z), \dots, \sigma_s(Z))^T$, the subdifferentiation of the nuclear norm is:

$$\partial \|Z\|_* = \left\{ U_{n_1 \times s} V_{n_2 \times s}^T + M : \begin{array}{c} Z \text{ and } M \text{ have orthogonal row} \\ \text{and column spaces, } and \ \|M\| \leq 1 \end{array} \right\},$$

where $\|\cdot\|$ means the matrix operator norm. Hence, if $\mathcal{W}$ is s-n-rank, then the subdifferentiation of $g_i(\mathcal{X})$ at point $\mathcal{W}$ is given by

$$\partial g_i(\mathcal{X})|_{\mathcal{X}=\mathcal{W}} = \left\{ \mathscr{B}_i^*(U_i V_i^T + M_i) : \begin{array}{l} \mathscr{B}_i(\mathcal{W}) = U_i \Sigma_i V_i^T, U_i^T M_i = 0, M_i V_i = 0, \\ \text{and } \|M_i\| \leq 1 \end{array} \right\}.$$

Therefore, s-n-rank tensor $\mathcal{W}$ is the optimal solution to problem (LTP) if and only if there exists $y \in R^M$ such that

$$\mathcal{O} = \frac{1}{N} \sum_{i=1}^{N} \mathscr{B}_i^*(U_i V_i^T + M_i) - \mathscr{A}^* y, \tag{4.1}$$

where U_i, V_i, M_i are defined as the above.

Note that

$$\mathscr{A}^* y = (\mathscr{B}_i^* \circ \mathscr{B}_i)\mathscr{A}^* y = \mathscr{B}_i^*(\mathscr{A} \circ \mathscr{B}_i^*)^* y = \frac{1}{N} \sum_{i=1}^{N} \mathscr{B}_i^*(\mathscr{A} \circ \mathscr{B}_i^*)^* y,$$

equation (4.1) turns out to be

$$\sum_{i=1}^{N} \mathscr{B}_i^*[(\mathscr{A} \circ \mathscr{B}_i^*)^* y - (U_i V_i^T + M_i)] = \mathcal{O}.$$

Thus, if for every $i \in \{1, \cdots, N\}$, there exists a vector $y \in R^M$ such that

$$(\mathscr{A} \circ \mathscr{B}_i^*)^* y = U_i V_i^T + M_i, \tag{4.2}$$

then, s-n-rank tensor $\mathcal{W}$ can be recovered by problem (LTP). And a sufficient condition for (4.2) to hold is that linear transformation $\mathscr{A} \circ \mathscr{B}_i^* : \mathbb{R}^{n_i \times J_i} \to \mathbb{R}^M$ is s-good for every index $i \in \{1, 2, \ldots, N\}$. Therefore, we define the s-n-goodness in tensor case:

Definition 4.1. Let s be an integer satisfying $0 \leq s \leq r = \min_{1 \leq i \leq N}\{n_i\}$. We call a linear transformation $\mathscr{A} : \mathbf{T} \to \mathbb{R}^M$ is s-n-good, if $\mathscr{A} \circ \mathscr{B}_i^* : \mathbb{R}^{n_i \times J_i} \to \mathbb{R}^M$ is s-good for every index $i \in \{1, 2, \ldots, N\}$.

Utilizing this definition, we can obtain the equivalence between the original tensor problem (LTR) and its convex relaxation (LTP).

Theorem 4.2. *Let $\mathscr{A} : \mathbf{T} \to \mathbb{R}^M$ be a linear transformation and s be an integer satisfying $0 \leq s \leq r = \min_{1 \leq i \leq N}\{n_i\}$. If $\mathscr{A}$ is s-n-good, then one can recover any s-n-rank tensor $\mathcal{W}$ by problem* (LTP). *Furthermore, $\mathcal{W}$ is also recoverable by problem* (LTR).

Proof. The proof is similar to the one of Theorem 3.3. By the definition of s-n-goodness of a linear transformation $\mathscr{A}$ and the assumption that $rank(\mathscr{B}_i(\mathcal{W})) \leq s$ for all $i \in \{1, 2, \ldots, N\}$, it is easy to see that for every $i \in \{1, 2, \ldots, N\}$, $\mathscr{B}_i(\mathcal{W})$ is the unique solution to the following problem:

$$\min_{X \in \mathbb{R}^{n_i \times J_i}} \|X\|_* : \quad \text{s.t. } \mathscr{A} \circ \mathscr{B}_i^*(X) = \mathscr{A} \circ \mathscr{B}_i^*(\mathscr{B}_i(\mathcal{W})). \tag{4.3}$$

Thus, one can recover any s-n-rank tensor $\mathcal{W}$ by problem (LTP). Otherwise, there exists an s-n-rank tensor $\mathcal{W}' \neq \mathcal{W}$ such that $\mathscr{A}(\mathcal{W}) = \mathscr{A}(\mathcal{W}')$ and $\sum_{i=1}^{N} \|\mathscr{B}_i(\mathcal{W}')\|_* \leq \sum_{i=1}^{N} \|\mathscr{B}_i(\mathcal{W})\|_*$, which means there must exist an index i_0 such that $\mathscr{B}_i(\mathcal{W}')$ is r-rank and $\|\mathscr{B}_{i_0}(\mathcal{W}')\|_* \leq \|\mathscr{B}_{i_0}(\mathcal{W})\|_*$. This contradicts that $\mathscr{B}_{i_0}(\mathcal{W})$ is the unique solution to problem (4.3) for $i = i_0$.

Suppose that $\mathcal{W}$ is not recoverable by problem (LTR), and there exists an $\mathcal{X}' \neq \mathcal{W}$ satisfying $\mathscr{A}(\mathcal{X}') = \mathscr{A}(\mathcal{W})$ and $rank(\mathscr{B}_i(\mathcal{X}')) \leq rank(\mathscr{B}_i(\mathcal{W})) \leq s$. Using the s-n-goodness of $\mathscr{A}$, we can obtain that $\mathscr{B}_i(\mathcal{X}')$ is the unique solution to problem (4.3). This contradicts that $\mathscr{B}_i(\mathcal{W})$ is the unique solution to problem (4.3). Hence, $\mathcal{W}$ is recoverable by (LTR). □

4.2 Error bound

In [17], the authors first introduced two quantities, $\gamma_s(\mathcal{A})$ and $\hat{\gamma}_s(\mathcal{A})$, which were defined as follows:

Definition 4.3. Let $\mathcal{A} : \mathbb{R}^{n_1 \times n_2} \to \mathbb{R}^M$ be a linear transformation and s be an integer satisfying $0 \leq s \leq r = \min\{n_1, n_2\}$. Then,

(i) $\gamma_s(\mathcal{A})$ is the infimum of $\gamma \geq 0$ such that for every matrix $X \in \mathbb{R}^{n_1 \times n_2}$ with SVD $X = U_{n_1 \times s} V_{n_2 \times s}^T$(i.e., s nonzero singular values, all equal to 1), there exists a vector $y \in \mathbb{R}^M$ such that

$$\mathcal{A}^* y = U Diag(\sigma(\mathcal{A}^* y)) V^T,$$

where $U = [U_{n_1 \times s} \;\; U_{n_1 \times (r-s)}]$, $V = [V_{n_2 \times s} \;\; V_{n_2 \times (r-s)}]$ are orthogonal matrices, and

$$\sigma_i(\mathcal{A}^* y) \begin{cases} = 1, & \text{if } \sigma_i(X) = 1, \\ \in [0, \gamma], & \text{if } \sigma_i(X) = 0 \end{cases}$$

for every $i \in \{1, 2, \ldots, r\}$. If for some X as above there does not exist such y, we set $\gamma_s(\mathcal{A}) = \infty$.

(ii) $\hat{\gamma}_s(\mathcal{A})$ is the infimum of $\gamma \geq 0$ such that for every matrix $X \in \mathbb{R}^{n_1 \times n_2}$ with s nonzero singular values, all equal to 1, there exists a vector $y \in \mathbb{R}^M$ such that

$$\|\mathcal{A}^* y - X\| \leq \gamma.$$

If for some X as above there does not exist such y, we set $\hat{\gamma}_s(\mathcal{A}) = \infty$.

They used $\gamma_s(\mathcal{A})$ and $\hat{\gamma}_s(\mathcal{A})$ to characterize the s-goodness of $\mathcal{A}$, i.e., $\mathcal{A}$ is s-good if and only if $\gamma_s(\mathcal{A}) < 1$ (or $\hat{\gamma}_s(\mathcal{A}) < \frac{1}{2}$). Therefore, using this result and Definition 4.1, we can obtain the following theorem:

Theorem 4.4. *Let $\mathscr{A} : \mathbf{T} \to \mathbb{R}^M$ be a linear transformation and s be an integer satisfying $0 \leq s \leq r = \min_{1 \leq i \leq N}\{n_i\}$. Then, $\mathscr{A}$ is s-n-good if and only if $\gamma_s(\mathscr{A} \circ \mathscr{B}_i^*) < 1$ (or $\hat{\gamma}_s(\mathscr{A} \circ \mathscr{B}_i^*) < \frac{1}{2}$) for every $i \in \{1, 2, \ldots, N\}$.*

Also, $\gamma_s(\mathcal{A})$ and $\hat{\gamma}_s(\mathcal{A})$ could be used to measure the error bound between the solution of problem (LMR) and the υ-optimal solution to problem (NNM) in [17]. In this subsection, we will try to measure the error bound between the solution of problem (LTR) and the υ-optimal solution to problem (LTP).

For the sake of the proof, we first state one important property of $\hat{\gamma}_s(\mathcal{A})$ (see [17, Theorem 2.7]:

Property 4.5. $\hat{\gamma}_s(\mathscr{A}) = \max_X\{\|X\|_s : \; \|X\|_* \leq 1, \; \mathcal{A}(X) = 0\}$, where $\|X\|_s$ is the sum of the s largest singular values of X.

In what follows, let the singular value decomposition (SVD) of the matrix $W \in \mathbb{R}^{n_1 \times n_2}$ be $W = UDiag(\sigma(W))V^T$, where U and V are both orthogonal matrices, $m = \min\{n_1, n_2\}$, $\sigma(W) = (\sigma_1(W), \ldots, \sigma_m(W))^T$ and $\sigma_1(W) \geq \ldots \geq \sigma_m(W) \geq 0$ are nonzero singular values of W in nonincreasing order. Let W^s be the best s-rank approximation of W, which is actually

$$W^s = UDiag((\sigma_1(W), \ldots, \sigma_s(W), 0 \ldots, 0)^T)V^T.$$

Now, we give the error bound in the low-n-rank tensor recovery problem.

Theorem 4.6. *Let $\mathscr{A} : \mathbf{T} \to \mathbb{R}^M$ be a linear transformation and s be an integer satisfying $0 \leq s \leq r = \min_{1 \leq i \leq N}\{n_i\}$. And let $t = \min\{2s, r\}$ and $\mathscr{A}$ be $2s$-n-good. Suppose that $\mathcal{W}$ is an optimal solution of problem* (LTR) *and $\mathcal{X}$ is a υ-optimal solution to problem* (LTP), *meaning that*

$$\mathscr{A}(\mathcal{X}) = \mathscr{A}(\mathcal{W}) \text{ and } \frac{1}{N}\sum_{i=1}^{N} \|\mathscr{B}_i(\mathcal{X})\|_* \leq \mathrm{Opt} + \upsilon,$$

where Opt *is the optimal value of* (LTP). *We can obtain*

$$\frac{1}{N}\sum_{i=1}^{N} \|\mathscr{B}_i(\mathcal{X}) - \mathscr{B}_i(\mathcal{W})\|_* \leq \frac{1}{N(1 - 2\hat{\gamma}(\mathscr{A}))}\sum_{i=1}^{N}[\upsilon + 2\|\mathscr{B}_i(\mathcal{W}) - W_i^s\|_*],$$

where W_i^s is the best s-rank approximation of $\mathscr{B}_i(\mathcal{W})$ and $\hat{\gamma}(\mathscr{A}) := \max_i\{\hat{\gamma}_t(\mathscr{A} \circ \mathscr{B}_i)\}$.

Proof. Set $\mathcal{Z} := \mathcal{X} - \mathcal{W}$. For every $i \in \{1, 2, \ldots, N\}$, let $\mathscr{B}_i(\mathcal{W})$ have structure as $\mathscr{B}_i(\mathcal{W}) = W_i^s + W_i^c$, where

$$W_i^s = U_i \begin{pmatrix} D_{is} & 0 \\ 0 & 0 \end{pmatrix} V_i^T \text{ and } W_i^c = U_i \begin{pmatrix} 0 & 0 \\ 0 & D_{i2} \end{pmatrix} V_i^T;$$

and $\mathscr{B}_i(\mathcal{X})$ has the form as

$$\mathscr{B}_i(\mathcal{X}) = U_i \begin{pmatrix} X_{i1} & X_{i2} \\ X_{i3} & X_{i4} \end{pmatrix} V_i^T.$$

Then, similar to the matrix case, we can write $\mathscr{B}_i(\mathcal{Z}) = Z_i^s + Z_i^c$, where

$$Z_i^s = U \begin{pmatrix} X_{i1} - D_{is} & X_{i2} \\ X_{i3} & 0 \end{pmatrix} V^T \text{ and } Z_i^c = U \begin{pmatrix} 0 & 0 \\ 0 & X_{i4} - D_{i2} \end{pmatrix} V^T.$$

Hence, we have $rank(Z_i^s) \leq rank((X_{i1} - D_{is},\ X_{i2})) + rank(X_{i3}) \leq 2s$, and ${Z_i^s}^T W_i^c = 0$ and ${W_i^s}^T Z_i^c = 0$. So, let $t = \min\{2s, r\}$, then $rank(Z_i^s) \leq t$.

Since $\mathcal{X}$ is a υ-optimal solution to problem (LTP), we have

$$\begin{aligned}
&\tfrac{1}{N}\textstyle\sum_{i=1}^{N} \|\mathscr{B}_i(\mathcal{W})\|_* + \upsilon \\
&\geq Opt + \upsilon \geq \tfrac{1}{N}\textstyle\sum_{i=1}^{N} \|\mathscr{B}_i(\mathcal{W} + \mathcal{Z})\|_* \\
&\geq \tfrac{1}{N}\textstyle\sum_{i=1}^{N} \|W_i^s + \mathscr{B}_i(\mathcal{Z}) - Z_i^s + Z_i^s + \mathscr{B}_i(\mathcal{W}) - W_i^s\|_* \\
&\geq \tfrac{1}{N}\textstyle\sum_{i=1}^{N}[\|W_i^s + \mathscr{B}_i(\mathcal{Z}) - Z_i^s\|_* - \|Z_i^s + \mathscr{B}_i(\mathcal{W}) - W_i^s\|_*] \\
&= \tfrac{1}{N}\textstyle\sum_{i=1}^{N}[\|W_i^s\|_* - \|\mathscr{B}_i(\mathcal{W}) - W_i^s\|_* + \|\mathscr{B}_i(\mathcal{Z}) - Z_i^s\|_* - \|Z_i^s\|_*],
\end{aligned}$$

where the last equality holds because of Lemma 2.3 in [30]. Thus,

$$\frac{1}{N}\sum_{i=1}^{N}\|\mathscr{B}_i(\mathcal{Z})-Z_i^s\|_*\leq\frac{1}{N}\sum_{i=1}^{N}[\|Z_i^s\|_*+2\|\mathscr{B}_i(\mathcal{W})-W_i^s\|_*]+\upsilon.$$

Moreover, by using the fact that $\mathscr{A}\circ\mathscr{B}_i^*(\mathscr{B}_i(\mathcal{Z}))=0$ and Property 4.5, we have $\|\mathscr{B}_i(Z)\|_t\leq\hat{\gamma}_t(\mathscr{A}\circ\mathscr{B}_i^*)\|\mathscr{B}_i(\mathcal{Z})\|_*$. Also it's easy to verify that $\|Z_i^s\|_*\leq\|\mathscr{B}_i(Z)\|_t$. Thus, we have

$$\|Z_i^s\|_*\leq\|\mathscr{B}_i(Z)\|_t\leq\hat{\gamma}_t(\mathscr{A}\circ\mathscr{B}_i^*)\|\mathscr{B}_i(\mathcal{Z})\|_*.$$

Furthermore, we can obtain that

$$\begin{aligned}&\tfrac{1}{N}\textstyle\sum_{i=1}^{N}\|\mathscr{B}_i(\mathcal{Z})\|_*\\ \leq&\tfrac{1}{N}\textstyle\sum_{i=1}^{N}[\|Z_i^s\|_*+\|\mathscr{B}_i(\mathcal{Z})-Z_i^s\|_*]\\ \leq&\tfrac{1}{N}\textstyle\sum_{i=1}^{N}[2\hat{\gamma}_t(\mathscr{A}\circ\mathscr{B}_i^*)\|\mathscr{B}_i(\mathcal{Z})\|_*+2\|\mathscr{B}_i(\mathcal{W})-W_i^s\|_*+\upsilon],\end{aligned}$$

which is equivalent to

$$\tfrac{1}{N}\textstyle\sum_{i=1}^{N}(1-2\hat{\gamma}_t(\mathscr{A}\circ\mathscr{B}_i^*))\|\mathscr{B}_i(\mathcal{Z})\|_*\leq\tfrac{1}{N}\sum_{i=1}^{N}2\|\mathscr{B}_i(\mathcal{W})-W_i^s\|_*+\upsilon.$$

Since $\hat{\gamma}(\mathscr{A})=\max_i\{\hat{\gamma}_t(\mathscr{A}\circ\mathscr{B}_i)\}$ and $\mathscr{A}$ is $2s$-n-good, which means $\hat{\gamma}(\mathscr{A})<\frac{1}{2}$, we can get the desired conclusion. □

When $N=2$, since $\mathscr{B}_1(\mathcal{X})=X$ and $\mathscr{B}_2(\mathcal{X})=X^T$, we have

$$\frac{1}{2}\sum_{i=1}^{2}\|\mathscr{B}_i(\mathcal{X})\|_*=\|X\|_*\quad\text{and}\quad\frac{1}{2}\sum_{i=1}^{2}\|\mathscr{B}_i(\mathcal{X})-\mathscr{B}_i(\mathcal{W})\|_*=\|X-W\|_*.$$

Hence, we have the following corollary in the matrix case.

Corollary 4.7. *Let $\mathcal{A}:\mathbb{R}^{n_1\times n_2}\to\mathbb{R}^M$ be a linear transformation and s be an integer satisfying $0\leq s\leq r=\min\{n_1,n_2\}$. And let $t=\min\{2s,r\}$ and $\hat{\gamma}_t(\mathcal{A})<\frac{1}{2}$ (or $\gamma_t(\mathcal{A})<1$). Also let $W\in\mathbb{R}^{n_1\times n_2}$ be the unique solution of problem* (LMR)*. Suppose that X is a υ-optimal solution to problem* (NNM)*, meaning that*

$$\mathcal{A}(X)=\mathcal{A}(W)\ \textit{and}\ \|X\|_*\leq\text{Opt}+\upsilon,$$

where Opt *is the optimal value of problem* (NNM)*. We have*

$$\|X-W\|_*\leq\frac{1+\gamma_t(\mathcal{A})}{1-\gamma_t(\mathcal{A})}[\upsilon+2\|W-W^s\|_*].$$

What is remarkable is that this corollary gets an error bound in the matrix case without the "Block Assumption" which is necessary for deriving the result in [17].

5 Restricted Isometry Property

First we recall the definition of the restricted isometry constant (RIC) in the context of matrix space:

Definition 5.1. The restricted isometry constant δ_r of a linear transformation $\mathcal{A} : \mathbb{R}^{n_1\times n_2} \to \mathbb{R}^M$ is defined as the smallest positive constant such that

$$(1-\delta_r)\|X\|_F^2 \leq \|\mathcal{A}(X)\|_2^2 \leq (1+\delta_r)\|X\|_F^2 \tag{5.1}$$

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

Cai and Zhang explored the sharp RIP condition in [3], which is the latest result of RIC discussed in the literature. The meaning of 'sharp' contains the following two aspects:

- If W is an r-rank matrix, then

$$\delta_{tr} < \sqrt{(t-1)/t} \quad \text{for some} \quad t \geq 4/3 \tag{5.2}$$

  can guarantee that W is the unique optimal solution of both problem (NNM) and problem (LMR).
- For any $\epsilon > 0$, $\delta_{tr} < \sqrt{(t-1)/t} + \epsilon$ is not sufficient to guarantee the exact recovery of all r-rank matrices for large r, i.e., there exists an r-rank matrix W such that W is not the minimizer of problem (NNM).

Next, we will investigate the similar results in tensor space. First, we extend Definition 5.1 to the tensor case, which also appeared in [28] and [29].

Definition 5.2. Let $\mathbf{r} = (r_1, \ldots, r_N)$. The restricted isometry constant (RIC) of the linear transformation $\mathscr{A} : \mathbf{T} \to \mathbb{R}^M$, $\alpha_{\mathbf{r}}$, is defined as the smallest positive constant such that

$$(1-\alpha_{\mathbf{r}})\|\mathcal{X}\|_F^2 \leq \|\mathscr{A}(\mathcal{X})\|_2^2 \leq (1+\alpha_{\mathbf{r}})\|\mathcal{X}\|_F^2 \tag{5.3}$$

holds for all tensors $\mathcal{X} \in \mathbf{T}$ of n-rank at most $\mathbf{r}$, i.e., $\text{rank}(\mathcal{X}_{<i>}) \leq r_i$ for all $i \in \{1, \ldots, N\}$.

It has been proved that for $\delta, \epsilon \in (0,1)$, a random draw of a Gaussian measurement map $\mathscr{A}$ satisfies $\alpha_{\mathbf{r}} \leq \delta$ with probability at least $1-\epsilon$ whenever the number of measurements M is large enough [29].

Next, we will discuss how large δ is to make $\alpha_{\mathbf{r}} \leq \delta$ be the sharp RIP condition for the equivalence between problems (LTP) and (LTR). Before giving the main result and proof, we propose the following two lemmas which are useful in the consequent analysis: the first one is a corollary of Lemma 1.1 in [3]; and the second one characters the relationship between the best rank-r approximation Y^r of a matrix Y and the rest part $Y - Y^r$.

Lemma 5.3. *For a positive number η and a positive integer s, define the polytope $T(\eta, s) \subset \mathbb{R}^{m\times n}$ by*

$$T(\eta, s) = \{X \in \mathbb{R}^{m\times n} : \ \|X\| \leq \eta, \ \|X\|_* \leq s\eta\},$$

where $\|X\|$ means the spectral norm of a matrix, i.e., its largest singular value. For any $X \in \mathbb{R}^{m\times n}$, define the set of low-rank matrices $U(\eta, s, X) \subset \mathbb{R}^{m\times n}$ by

$$U(\eta, s, X) = \{Y : \ rank(Y) \leq s, \ \|Y\| \leq \eta, \ \|Y\|_* = \|X\|_*\}.$$

Then, any $X \in T(\eta, s)$ can be expressed as $X = \sum_{i=1}^{M'} \lambda_i Z_i$, where $0 \leq \lambda_i \leq 1$, $\sum_{i=1}^{M'} \lambda_i = 1$,and $Z_i \in U(\eta, s, X)$.

Proof. Suppose the singular value decomposition of X is $X = \sum_{i=1}^{m} \sigma_i u_i v_i^T$, where σ_i are in descending order. Let vector $h := (\sigma_1, \ldots, \sigma_m)^T$, then we have $rank(X) = \|h\|_0$, $\|X\| = \|h\|_\infty$, and $\|X\|_* = \|h\|_1$. Using Lemma 1.1 in [3], the result is obtained directly. □

Lemma 5.4. *For any matrix $Y \in \mathbb{R}^{n_1 \times n_2}$, let the best rank-$r$ approximation of Y be Y^r. Then, $Y - Y^r$ can be divided into the sum of two parts: $Y^{(1)} + Y^{(2)}$, which satisfy that for some $t \geq 4/3$,*

$$rank(Y^{(1)}) \leq r(t-1) \quad \text{and} \quad \|Y^{(2)}\|_* \leq \|Y^r\|_* - \frac{rank(Y^{(1)})}{r(t-1)}\|Y^r\|_*.$$

Proof. Let $n = \min\{n_1, n_2\}$. Suppose that the singular value decomposition of Y is $Y = \sum_{j=1}^n \sigma_j u_j v_j^T$, where the singular values σ_j are in descending order. Since Y^r is the best rank-r approximation of Y, we have

$$Y^r = \sum_{j=1}^r \sigma_j u_j v_j^T, \quad Y - Y^r = \sum_{j=r+1}^n \sigma_j u_j v_j^T, \quad \text{and } \|Y - Y^r\|_* \leq \|Y^r\|_*.$$

Set $\eta = \|Y^r\|_*/r$. We divide $Y - Y^r$ into two parts, $Y - Y^r = Y^{(1)} + Y^{(2)}$, where

$$Y^{(1)} = \textstyle\sum_{j \geq r+1, \sigma_j > \eta/(t-1)} \sigma_j u_j v_j^T, \quad Y^{(2)} = \sum_{j \geq r+1, \sigma_j \leq \eta/(t-1)} \sigma_j u_j v_j^T.$$

Then,

$$\sum_{j \geq r+1, \sigma_j > \eta/(t-1)} \sigma_j = \|Y^{(1)}\|_* \leq \|Y - Y^r\|_* \leq \|Y^r\|_* = \eta r,$$

which shows that $rank(Y^{(1)}) \leq r(t-1)$. Since $\|Y^{(2)}\| \leq \frac{\eta}{t-1}$, we have

$$\|Y^{(2)}\|_* \quad = \quad \|Y - Y^r\|_* - \|Y^{(1)}\|_* \leq \|Y^r\|_* - \frac{rank(Y^{(1)})}{r(t-1)}\|Y^r\|_*.$$

The proof is complete. □

Now, we give the RIP condition in tensor case:

Theorem 5.5. *Let $\mathscr{A} : \mathbf{T} \to \mathbb{R}^M$ be a linear transformation and $\mathcal{W}$ be an r-n-rank tensor. For some constant $t \geq 4/3$, suppose $t\mathbf{r} = (tr, \cdots, tr) \in \mathbb{R}^N$. If the RIC of $\mathscr{A}$, $\alpha_{\mathbf{r}}$, satisfies $\alpha_{t\mathbf{r}} < \sqrt{(t-1)/t}$, then the unique minimizer of problem* (LTP) *is equal to the solution to problem* (LTR)*, which is exactly $\mathcal{W}$.*

Proof. The proof is stimulated by the approach in the proof of Theorem 1.1 in [3]. First, for some constant $t \geq 4/3$, we assume that tr is an integer. Otherwise, we set $t' = \lceil tr \rceil / r$, where $\lceil x \rceil$ represents the smallest integer which is equal to or greater than x, then $t' > t$ and $t'r$ is an integer, and hence,

$$\alpha_{t'\mathbf{r}} = \alpha_{t\mathbf{r}} < \sqrt{(t-1)/t} < \sqrt{(t'-1)/t'},$$

which can be deduced to the former case.

Based on Theorem 3.3, we only need to prove that for all $\mathcal{Y} \in Null(\mathscr{A}) \backslash \{0\}$, $\|\mathscr{B}_i(\mathcal{Y})\|_* - 2\|\mathscr{B}_i(\mathcal{Y})\|_r > 0$ for every $i \in \{1, \ldots, N\}$.

Now, suppose that there exist $\mathcal{Y} \in Null(\mathscr{A}) \backslash \{0\}$ and an index i_0 such that $\|\mathscr{B}_{i_0}(\mathcal{Y})\|_* - 2\|\mathscr{B}_{i_0}(\mathcal{Y})\|_r \leq 0$. To make the symbols simple, we denote $\mathscr{B}_{i_0}(\mathcal{Y}) := Y$. Set $rank(Y^{(1)}) = m$ and $\eta = \|Y^r\|_*/r$. Then, utilizing Lemma 5.4, we have that for some $t \geq 4/3$,

$$Y - Y^r = Y^{(1)} + Y^{(2)}, \ \|Y^{(2)}\| \leq \frac{\eta}{t-1}, \text{ and } \|Y^{(2)}\|_* \leq (r(t-1) - m)\frac{\eta}{t-1}.$$

We now apply Lemma 5.3 with $s = r(t-1)-m$. Then, $Y^{(2)}$ can be expressed as a convex combination of low-rank matrices: $Y^{(2)} = \sum_{j=1}^{J} \lambda_j Z_j$, where Z_j is $(r(t-1)-m)$-rank and $\|Z_j\| \leq \frac{\eta}{t-1}$. Then,

$$\begin{aligned} \|Z_j\|_F^2 &\leq (r(t-1)-m)\|Z_j\|^2 \leq \frac{r\eta^2}{t-1} = \frac{\|Y^r\|_*^2}{r(t-1)} \\ &\leq \frac{\|Y^r\|_F^2}{t-1} \leq \frac{\|Y^r + Y^{(1)}\|_F^2}{t-1}. \end{aligned} \tag{5.4}$$

Now for any $\mu > 0$, denote $D = Y^r + Y^{(1)}$ and $B_j = D + \mu Z_j$, then B_j is tr-rank and

$$F_j := \textstyle\sum_{j=1}^{J} \lambda_j B_j - \frac{1}{2}B_j = D + \mu Y^{(2)} - \frac{1}{2}B_j = (\frac{1}{2}-\mu)D - \frac{1}{2}\mu Z_j + \mu Y. \tag{5.5}$$

Thus, $\sum_{j=1}^{m} \lambda_j B_j - \frac{1}{2}B_j - \mu Y = (\frac{1}{2}-\mu)D - \frac{1}{2}\mu Z_j$ is also a tr-rank matrix.

It is easy to prove that

$$\sum_{j=1}^{m} \lambda_j \|\mathscr{A}(\mathscr{B}_{i_0}^*(F_j))\|_2^2 - \frac{1}{4}\sum_{j=1}^{m} \lambda_j \|\mathscr{A}(\mathscr{B}_{i_0}^*(B_j))\|_2^2 = 0.$$

Then, by setting $\mu = \sqrt{t(t-1)} - (t-1)$, we can further obtain that

$$\begin{aligned} 0 &= \sum_{j=1}^{m} \lambda_j \|\mathscr{A}(\mathscr{B}_{i_0}^*((\frac{1}{2}-\mu)D - \frac{1}{2}\mu Z_j))\|_2^2 - \frac{1}{4}\sum_{j=1}^{m} \lambda_j \|\mathscr{A}(\mathscr{B}_{i_0}^*(B_j))\|_2^2 \\ &\leq (1+\alpha_{\mathbf{tr}})\sum_{j=1}^{m} \lambda_j \|(\frac{1}{2}-\mu)D - \frac{1}{2}\mu Z_j\|_F^2 \\ &\quad -\frac{1}{4}(1-\alpha_{\mathbf{tr}})\sum_{j=1}^{m} \lambda_j \|D + \mu Z_j\|_F^2 \\ &= (1+\alpha_{\mathbf{tr}})\sum_{j=1}^{m} \lambda_j \left((\frac{1}{2}-\mu)^2\|D\|_F^2 + \frac{1}{4}\mu^2\|Z_j\|_F^2\right) \\ &\quad -\frac{1}{4}(1-\alpha_{\mathbf{tr}})\sum_{j=1}^{m} \lambda_j (\|D\|_F^2 + \mu^2\|Z_j\|_F^2) \\ &\leq \left((\frac{2t-1}{2(t-1)}\mu^2 - 2\mu + \frac{1}{2})\alpha_{\mathbf{tr}} + (\mu^2 - 2\mu)\right)\|D\|_F^2 \\ &< 0, \end{aligned} \tag{5.6}$$

where the first equality follows from $\mathscr{A}(\mathscr{B}_{i_0}^* Y) = \mathscr{A}(\mathcal{Y}) = 0$, the first inequality follows from the fact that (5.3) and $\mathscr{B}_{i_0}^*((\frac{1}{2}-\mu)(D) - \frac{1}{2}\mu Z_j)$ and $\mathscr{B}_{i_0}^*(B_j)$ are tr-rank matrices, the second inequality follows from (5.4), and the last inequality follows from $\alpha_{\mathbf{tr}} < \sqrt{(t-1)/t}$. Note that (5.6) shows that $0 < 0$, which is a contradiction. So the proof is complete. □

Next, we will prove this RIP condition is sharp. In other words, for any $\epsilon > 0$, $\alpha_{\mathbf{tr}} < \sqrt{(t-1)/t} + \epsilon$ is not sufficient to guarantee the exact recovery of all r-n-rank tensors for large r.

Theorem 5.6. *Let $t \geq 4/3$. For all $\epsilon > 0$ and $r \geq 5/\epsilon$, set $t\mathbf{r}$ to be the same in Theorem 5.5, then there exists a linear transformation $\mathscr{A}$ satisfying $\alpha_{\mathbf{tr}} < \sqrt{(t-1)/t} + \epsilon$ and some r-n-rank tensor $\mathcal{W}$ such that the minimizer of problem* (LTP) *is not equal to $\mathcal{W}$.*

Proof. The proof is similar to that of Theorem 2.2 in [3], and the main approach is to construct a linear transformation $\mathscr{A}$ satisfying $\alpha_{t\mathbf{r}} < \sqrt{(t-1)/t}+\epsilon$ and an r-n-rank tensor $\mathcal{W}$, such that there exists another tensor $\mathcal{Z}$ satisfying $\mathscr{A}(\mathcal{Z}) = \mathscr{A}(\mathcal{W})$, but $\frac{1}{N}\sum_{i=1}^{N}\|\mathscr{B}_i(\mathcal{Z})\|_* < \frac{1}{N}\sum_{i=1}^{N}\|\mathscr{B}_i(\mathcal{W})\|_*$, which means r-n-rank tensor $\mathcal{W}$ can not be recovered by problem (LTP).

Step 1: Construct a linear transformation $\mathscr{A}$ satisfying $\alpha_{t\mathbf{r}} < \sqrt{(t-1)/t}+\epsilon$

For any $\epsilon > 0$ and $r \geq 5/\epsilon$, suppose that $d_1 = ((t-1)+\sqrt{t(t-1)})r$ and d is the largest integer strictly smaller than d_1. Define

$$\mathcal{X}_1 = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \cdots \times_N \mathbf{U}^{(N)},$$

where $\mathbf{U}^{(k)} = (u_1^{(k)}\ u_2^{(k)}\ \cdots\ u_{n_k}^{(k)})$ is a unitary $n_k \times n_k$ matrix; and the element of $\mathcal{S} \in \mathbf{T}$ is given by

$$\mathcal{S}_{i_1\cdots i_N} = \begin{cases} (r+\frac{dr^2}{d_1^2})^{-\frac{1}{2}}, & \text{if } 1 \leq i_1 = \cdots = i_N \leq r, \\ (r+\frac{dr^2}{d_1^2})^{-\frac{1}{2}}(-\frac{r}{d_1}), & \text{if } r+1 \leq i_1 = \cdots = i_N \leq d, \\ 0, & \text{otherwise,} \end{cases}$$

then $\|\mathcal{X}_1\|_F = 1$. Define linear transformation $\mathscr{A} : \mathbf{T} \to \mathbb{R}^M$ by

$$\mathscr{A}(\mathcal{X}) = \sqrt{1+\sqrt{\tfrac{t-1}{t}}}(\mathcal{X} - \langle \mathcal{X}_1, \mathcal{X}\rangle \mathcal{X}_1).$$

Now, we prove that the RIC of $\mathscr{A}$, $\alpha_{t\mathbf{r}}$, satisfies $\alpha_{t\mathbf{r}} < \sqrt{(t-1)/t}+\epsilon$. For all $\lceil tr\rceil$-n-rank tensor $\mathcal{X}$, it holds that

$$\|\mathscr{A}(\mathcal{X})\|_2^2 = \left(1+\sqrt{(t-1)/t}\right)(\|\mathcal{X}\|_F^2 - |\langle \mathcal{X}_1, \mathcal{X}\rangle|^2).$$

Furthermore, we have

$$\|\mathscr{A}(\mathcal{X})\|_2^2 \leq \left(1+\sqrt{\tfrac{t-1}{t}}\right)\|\mathcal{X}\|_F^2 \leq \left(1+\sqrt{\tfrac{t-1}{t}}+\epsilon\right)\|\mathcal{X}\|_F^2. \tag{5.7}$$

Since $\mathcal{X}$ is $\lceil tr\rceil$-n-rank, by Cauchy-Schwarz inequality, we have

$$\begin{aligned}
0 &\leq |\langle \mathcal{X}_1, \mathcal{X}\rangle|^2 \\
&\leq (r+\tfrac{dr^2}{d_1^2})^{-1}\left(r+(\lceil tr\rceil - r)\tfrac{r^2}{d_1^2}\right)\|\mathcal{X}\|_F^2 \\
&\leq \tfrac{d_1^2+r^2(t-1)+r}{d_1^2+d_1r} \cdot \tfrac{1}{1-\frac{r(d_1-d)}{d_1^2+d_1r}}\|\mathcal{X}\|_F^2 \\
&= \tfrac{d_1^2+r^2(t-1)}{d_1^2+d_1r} \cdot (1+\tfrac{r}{d_1^2+d_1r}) \cdot \tfrac{1}{1-\frac{r(d_1-d)}{d_1^2+d_1r}}\|\mathcal{X}\|_F^2 \\
&= 2\sqrt{t-1}(\sqrt{t}-\sqrt{t-1})(1+\tfrac{r}{d_1^2+d_1r}) \cdot \tfrac{1}{1-\frac{r(d_1-d)}{d_1^2+d_1r}}\|\mathcal{X}\|_F^2 \\
&\leq 2\sqrt{t-1}(\sqrt{t}-\sqrt{t-1})(1+\tfrac{1}{tr}) \cdot \tfrac{1}{1-\frac{1}{2r}}\|\mathcal{X}\|_F^2 \\
&\leq 2\sqrt{t-1}(\sqrt{t}-\sqrt{t-1})(1+\tfrac{5}{2r})\|\mathcal{X}\|_F^2 \\
&\leq (2\sqrt{t(t-1)}-2(t-1)+\tfrac{5}{2r})\|\mathcal{X}\|_F^2,
\end{aligned}$$

where the third inequality follows from $\lceil tr \rceil \le tr$, the forth inequality follows from $d_1 \ge r$ since $t \ge 4/3$ and $d_1 - d \le 1$, and the last two inequalities follow from $t \ge 4/3$. Therefore,

$$\|\mathscr{A}(\mathcal{X})\|_2^2 \ge \left(1 - \sqrt{\frac{t-1}{t}} - \epsilon\right) \|\mathcal{X}\|_F^2. \tag{5.8}$$

Using (5.7) and (5.8), we get $\alpha_{t\mathbf{r}} < \sqrt{(t-1)/t} + \epsilon$.

Step2: Construct $\mathcal{W}$ and $\mathcal{Z}$

We construct $\mathcal{W}$ and $\mathcal{Z}$ by

$$\mathcal{W} = \mathcal{S}^{(W)} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \cdots \times_N \mathbf{U}^{(N)}, \quad \mathcal{Z} = \mathcal{S}^{(Z)} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \cdots \times_N \mathbf{U}^{(N)},$$

where

$$\begin{aligned} \mathcal{S}^{(W)}{}_{i_1 \cdots i_N} &= \begin{cases} 1, & \text{if } 1 \le i_1 = \cdots = i_N \le r, \\ 0, & \text{otherwise;} \end{cases} \\ \mathcal{S}^{(Z)}{}_{i_1 \cdots i_N} &= \begin{cases} \frac{r}{d_1}, & \text{if } r+1 \le i_1 = \cdots = i_N \le d, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Note that both $\mathcal{W}$ and $\mathcal{Z}$ are r-n-rank and $0 = \mathscr{A}(\mathcal{X}_1) = \mathscr{A}(\mathcal{W} - \mathcal{Z})$, which means $\mathscr{A}(\mathcal{W}) = \mathscr{A}(\mathcal{Z})$. However, since $d_1 > d$,

$$\frac{1}{N} \sum_{i=1}^{N} \|\mathscr{B}_i(\mathcal{Z})\|_* = \frac{d}{d_1} r < r = \frac{1}{N} \sum_{i=1}^{N} \|\mathscr{B}_i(\mathcal{W})\|_*,$$

which implies $\mathcal{W}$ is not the minimizer of problem (LTP). The proof is complete. □

Remark: When $N = 2$, the tensor space $\mathbf{T} = \mathbb{R}^{n_1 \times n_2 \times \ldots \times n_N}$ reduces to the matrix space $\mathbb{R}^{n_1 \times n_2}$, and hence, Theorem 5.5 reduces to Proposition 3.1 in [3] in the noiseless matrix case; while Theorem 5.6 reduces to Proposition 3.2 in [3] in the noiseless matrix case.

6 Relationship Among NSP, RIP and s-n-Goodness

In this section, we will discuss the relationship among the three exact recovery conditions: the NSP, the s-n-good condition and the RIP. Based on the results in the aforementioned sections, all these three conditions can guarantee the equivalence between problems (LTR) and (LTP). However, the RIP condition is more strict than the other two from the proof of Theorem 5.5; while the NSP condition is equivalent to the s-n-good condition.

Theorem 6.1. *(NSP and s-n-good) Let $\mathscr{A} : \mathbf{T} \to \mathbb{R}^M$ be a linear transformation and r be an integer satisfying $0 \le r \le R = \min_{1 \le i \le N}\{n_i\}$. $\mathscr{A}$ is r-n-good if and only if $\|\mathscr{B}_i(\mathcal{Y})\|_* - 2\|\mathscr{B}_i(\mathcal{Y})\|_r > 0$ holds for any $\mathcal{Y} \in Null(\mathscr{A})\backslash\{0\}$, $\forall\, i \in \{1, 2, ..., N\}$.*

Proof. From Theorem 4.4, $\mathscr{A}$ is r-n-good if and only if $\hat{\gamma}_s(\mathscr{A} \circ \mathscr{B}_i^*) < \frac{1}{2}$ for every $i \in \{1, 2, \ldots, N\}$. Based on (4.5), it is easy to obtain the equivalence between the r-n-goodness and the NSP condition. □

Theorem 6.2. *Let $\mathscr{A} : \mathbf{T} \to \mathbb{R}^M$ be a linear transformation and r be an integer satisfying $0 \le r \le R = \min_{1 \le i \le N}\{n_i\}$. Set $t\mathbf{r} = (tr, \cdots, tr) \in \mathbb{R}^N$. If the RIC of $\mathscr{A}$, $\alpha_{\mathbf{r}}$, satisfies $\alpha_{t\mathbf{r}} < \sqrt{\frac{t-1}{t}}$ for some $t \ge 4/3$, then* *(i) (RIP and NSP) for all $i \in \{1, 2, ..., N\}$, the following holds:*

$$\|\mathscr{B}_i(\mathcal{Y})\|_* - 2\|\mathscr{B}_i(\mathcal{Y})\|_r > 0, \quad \forall\ \mathcal{Y} \in Null(\mathscr{A})\backslash\{0\}.$$

(ii) (RIP and s-n-good) $\mathscr{A}$ is r-n-good.

Proof. The relationship between RIP and NSP has been proved obviously through the proof of Theorem 5.5. So, the above consequence (i) holds directly. And by utilizing the result in (i) and the equivalence between NSP and s-n-good, the relationship between RIP and s-n-good in (ii) is obtained straightly. □

Next, we will give an example to show that all the three conditions, NSP, s-n-goodness and RIP, are only sufficient but not necessary. In other words, we construct a problem such that it is equivalent to the relaxation problem but the three exact recovery conditions do not hold.

Example 6.3. Let the linear transformation $\mathscr{A} : \mathbb{R}^{6\times6\times6} \to \mathbb{R}^3$ be given by

$$\mathscr{A}(\mathcal{X}) = \begin{pmatrix} \langle \mathcal{A}_1, \mathcal{X}\rangle \\ \langle \mathcal{A}_2, \mathcal{X}\rangle \\ \langle \mathcal{A}_3, \mathcal{X}\rangle \end{pmatrix},$$

where $\mathcal{A}_i$ $(i = 1, 2, 3)$ are three $6 \times 6 \times 6$ tensors and $\langle\cdot,\cdot\rangle$ is the inner product between two tensors.

$\mathcal{A}_1 = (a^{(1)}_{ijk})$: $(a^{(1)}_{jjj})^6_{j=1} = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$ and other entries are zeros.

$\mathcal{A}_2 = (a^{(2)}_{ijk})$: $(a^{(2)}_{jjj})^6_{j=1} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, 0, -\frac{1}{\sqrt{2}})$ and other entries are zeros.

$\mathcal{A}_3 = (a^{(3)}_{ijk})$: $(a^{(3)}_{jjj})^6_{j=1} = (0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ and other entries are zeros.

Let $\mathcal{W} = (w_{ijk}) \in \mathbb{R}^{6\times6\times6}$ with $w_{333} = \sqrt{3}$ and other entries being zeros is the tensor to be recover. It's easy to see that n-rank($\mathcal{W}$)=$(1, 1, 1)$, so it's a 2-n-rank tensor. And $\mathscr{A}(\mathcal{W}) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. It's not difficult to verify that $\mathcal{W}$ is the minimizer of both problems (LTR) and (LTP) with the above linear transformation $\mathscr{A}$. However, $\mathscr{A}$ does not satisfy those three conditions.

First, we check the RIP condition. By simple calculation, we have

$$\|\mathscr{A}(\mathcal{W})\|_2^2 = \sum_{j=1}^{6} w_{jjj}^2 = 1 = \|\mathcal{W}\|_F^2.$$

Thus, $\alpha_{\mathbf{2}} = 0$ which contradicts the definition of $\alpha_{\mathbf{2}}$ $(0 < \alpha_{\mathbf{2}} < 1)$. So the RIP condition does not hold for this $\mathscr{A}$.

Second, we check the NSP condition. There exists a tensor $\mathcal{Y} = (y_{ijk}) \in \mathbb{R}^{6\times6\times6}$ with $y_{222} = y_{666} = 1$ and other entries being zeros such that $\mathscr{A}(\mathcal{Y}) = 0$, i.e., $\mathcal{Y} \in Null(\mathscr{A})\backslash\{0\}$. And we can obtain that for all $i = 1, 2, 3$,

$$2\|\mathscr{B}_i(\mathcal{Y})\|_r = \|\mathscr{B}_i(\mathcal{Y})\|_*,$$

which means NSP condition does not hold. And based on Theorem 6.1, we can know that $\mathscr{A}$ is not 2-n-good.

Hence, this example shows that all these three conditions are only sufficient.

7 Conclusions

In this paper, we explored the conditions to guarantee the equivalence between the low-n-rank tensor recovery problem (LTR) and its convex relaxation (LTP). We studied three exact recovery conditions: the NSP, the s-n-goodness condition and the RIP, with the main results being regarded as extensions from the corresponding results in the matrix case. Also, we discussed the relationship among the three exact recovery conditions and gave an example to show these conditions are only sufficient for the equivalence between problem (LTR) and its convex relaxation. A further issue is to investigate some necessary and sufficient conditions.

References

[1] A. Bruckstein, M. Elad and M. Zibulevsky, On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations, *IEEE Trans. Inform. Theory.* 54 (2008) 4813–4820

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

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

[4] E.J. Candès and Y. Plan, Tight oracle bounds for low-rank matrix recovery from a minimal number of random measurements, *IEEE Trans. Inform. Theory.* 57 (2009) 2342–2359

[5] E.J. Candès, J. Romberg and T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inform. Theory.* 52 (2006) 489–509

[6] E.J. Candès and T. Tao, Decoding by linear programming, *IEEE Trans. Inform. Theory.* 51 (2005) 4203–4215

[7] E.J. Candès and T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies, *IEEE Trans. Inform. Theory.* 52 (2006) 5406–5425

[8] A.L. Chistov and D.Y. Grigoriev, Complexity of quantifier elimination in the theory of algebraically closed fields, *Math. Found. Comput. Sci. 1984.* Springer Berlin Heidelberg, 1984, pp. 17–31

[9] D.L. Donoho, Compressed sensing, *IEEE Trans. Inform. Theory.* 52 (2006) 1289–1306

[10] D.L. Donoho, For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution, *Commun. Pure Appl. Anal.* 59 (2006) 797–829

[11] M. Fazel, H. Hindi and S. Boyd, A rank minimization heuristic with application to minimum order system approximation, *In Proc. American Control Conf., IEEE* 6 (2001) 4734–4739

[12] S. Gandy, B. Recht and I. Yamada, Tensor completion and low-n-rank tensor recovery via convex optimization, *Inverse Problems.* 27 (2011) 025010

[13] Bo Huang, Cun Mu, Donald Goldfarb and John Wright, Provable low-rank tensor recovery, *Optimization-online*, 2014.

[14] A. Juditsky and A.S. Nemirovski, On verifiable sufficient conditions for sparse signal recovery via l_1 minimization, *Mathe. Program.* 127 (2011) 57–88

[15] L.C. Kong and N.H. Xiu, New bounds for restricted isometry constants in low-rank matrix recovery, *Optimization-online*, http://www.optimization-online.org/DB_FILE/2011/01/2894.pdf (2011)

[16] L.C. Kong and N.H. Xiu, Exact low-rank matrix recovery via nonconvex schatten p-minimization, *Asia-Pac. J. Oper. Res.* 30 (2013) 1340010

[17] L.C. Kong, L. Tuncel and N.H. Xiu, Sufficient conditions for low-rank matrix recovery, translated from sparse signal recovery, *CoRR abs/1106.3276* (2011)

[18] L.C. Kong, L. Tuncel and N.H. Xiu, s-goodness for low-rank matrix recovery, *Abstr. Appl. Anal.* Volume 2013, Article ID: 101974

[19] K. Lee and Y. Bresler, Guaranteed minimum rank approximation from linear observations by nuclear norm minimization with an ellipsoidal constraint, arXiv:0903.4742v1 (2009)

[20] J. Liu, P. Musialski, P. Wonka and J.P. Ye, Tensor completion for estimating missing values in visual data, *IEEE Intl. Conf. Computer Vision (ICCV), Kyoto, Japan.* (2009) 2114–2121

[21] R. Meka, P. Jain, C. Caramanis and I.S. Dhillon, Rank minimization via online learning, *ICML'08 Proc. 25th Intl. Conf. Machine Learning.* ACM Press, 2008, pp. 656-663

[22] K. Mohan and M. Fazel, New restricted isometry results for noisy low-rank matrix recovery, *Proc. Intl. Symp. Inf. Thoery (ISIT), Austin, TX.*, 2010.

[23] Cun Mu, Bo Huang, John Wright and Donald Goldfarb, Square deal: Lower bounds and improved relaxations for tensor recovery. See arXiv:1307.5870, 2013.

[24] B.K. Natarajan, Sparse approximate solutions to linear systems, *SIAM J. Comput.* 24 (1995) 227–234

[25] S. Oymak and B. Hassibi, New null space results and recovery thresholds for matrix rank minimization. See arXiv:1011.6326v1, 2010.

[26] S. Oymak, A. Khajehnejad and B. Hassibi, Improved thresholds for rank minimization, *2011 IEEE International Conference on Speech and Signal Processing (ICASSP), Acoustics*, (2011) 5988–5991

[27] S. Oymak, K. Mohan, M. Fazel and B. Hassibi, A simplified approach to recovery conditions for low rank matrices, *2011 IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2011, pp. 2318–2322

[28] H. Rauhut, R. Schneider and Z. Stojanac, Tensor recovery via iterative hard thresholding, *SampTA*, 2013.

[29] H. Rauhut, R. Schneider and Z. Stojanac, Tensor completion in hierarchical tensor representations, *Technical report*, arXiv:1404.3905v1, 2014.

[30] B. Recht, M. Fazel and P.A. Parrilo, Guaranteed minimum rank solutions of linear matrix equations via nuclear norm minimization, *SIAM Rev.* 52 (2010) 471–501

[31] B. Recht, W. Xu and B. Hassibi, Necessary and sufficient conditions for success of the nuclear norm heuristic for rank minimization, *Proceedings of the 47th IEEE Conference on Decision and Control Cancun, Mexico*, 2008, pp. 3065–3070

[32] B. Recht, W. Xu and B. Hassibi, Null space conditions and threshlods for rank minimization, *Math. Program.* 127 (2011) 175–202

[33] M. Signoretto, Q.T. Dinh, L. De. Lathauwer and J.A.K. Suykens, Learning with tensors: a framework based on convex optimization and spectral regularization, *Machine Learning*, 2013, pp. 1-49

[34] M. Signoretto, L. De. Lathauwer and J.A.K. Suykens, Nuclear norms for tensors and their use for convex multilinear estimation, *Internal report 10-186, ESAT-SISTA, K.U. Leuven, Leuven, Belgium*, Lirias number: 270741, 2010.

[35] R. Tomioka, K. Hayashi and H. Kashima, Estimation of low-rank tensors via convex optimization, See arXiv:1010.0789v2, 2011.

[36] L. Yang, Z.H. Huang and Y.F. Li, A splitting augmented Lagrangian method for low multilinear-rank tensor recovery, *Asia-Pac. J. Oper. Res.* 32 (2015) 1540008 (25 pages)

[37] L. Yang, Z.H. Huang and X.J. Shi, A fixed point iterative method for low n-rank tensor pursuit, *IEEE Trans. Signal Process.* 61 (2013) 2952–2962

[38] M. Zhang and Z.H. Huang, Exact recovery conditions for the low-n-rank tensor recovery problem via its convex relaxation, *Technique Report*, Department of Mathematics, School of Science, Tianjin University, Tianjin 300072, P.R. China, 2012.

[39] M. Zhang, Z.H. Huang and Y. Zhang, Restricted p-isometry properties of nonconvex matrix recovery, *IEEE Trans. Inform. Theory*, 59 (2013) 4316–4323

[40] M. Zhang, L. Yang and Z.H. Huang, Minimum n-rank approximation via iterative hard thresholding, *Appl. Math. Comput.* 256 (2015) 860–875.

Manuscript received 3 January 2014
revised 29 September 2014, 10 November 2014
accepted for publication 11 November 2014

Min Zhang
Department of Mathematics, School of Science
Tianjin University, Tianjin 300072, P.R. China E-mail address: minzhang@tju.edu.cn

Zheng-Hai Huang
Department of Mathematics, School of Science
Tianjin University, Tianjin 300072, P.R. China

Center for Applied Mathematics of Tianjin University
E-mail address: huangzhenghai@tju.edu.cn