Analysis non-sparse recovery for non-convex relaxed ℓ_q minimization

Jianwen Huang[a*], Feng Zhang[b], Xinling Liu[c], Jianjun Wang[b]

[a] *School of Mathematics and Statistics, Tianshui Normal University, Tianshui 741001 China*

[b] *School of Mathematic and Statistics, Southwest University, Chongqing 400715 China*

[c] *College of Mathematics and Information, China West Normal University, Nanchong 637009, China*

Abstract. This paper studies construction of signals, which are sparse or nearly sparse with respect to a tight frame D from underdetermined linear systems. In the paper, we propose a non-convex relaxed $\ell_q(0 < q \leq 1)$ minimization for sparse dictionary recovery. Based on the ℓ_q robust D-Null Space Property, we derive the sparse or non-sparse solution to the non-convex relaxed ℓ_q minimization problem and the associating performance bound in which the $\|D^*D\|_{1,1}$ in the noise bound constant is removed. Additionally, we show that our method can stably recover sparse or approximately sparse signals with respect to a tight frame provided that the measurement matrix A fulfills a properly adapted restricted isometry property. As byproduct, when choose $\rho \to \infty$, we obtain the recovery guarantee and the corresponding error estimation via the unconstrained non-convex ℓ_q minimization.

Key words. Compressed sensing; ℓ_q robust D-Null Space Property; Non-convex relaxed ℓ_q minimization method; Restricted isometry property adapted D; Sparse recovery.

1 Introduction

One of the central goals for compressed sensing is to estimate sparse signals from what was before deemed to be undersampled data. In compressed sensing, one thinks over the linear model as follows:

$$b = Ax + w, \tag{1.1}$$

in which $A \in \mathbb{R}^{m\times n}$ is a known measurement matrix ($m \ll n$) and $w \in \mathbb{R}^m$ denotes a noise vector. The aim is to recover the unknown signal x based on $(b,\ A)$.

When x to be recovered signal is a sparse vector, one can utilize some sparse recovery approaches, see, e.g. [1–11], to stably reconstruct x provided that the measurement matrix A satisfies a few appropriate assumptions.

*E-mail: hjw1303987297@126.com

Motivated by a variety of practical applications including signal modeling in array processing, Gabor frames in radar and sonar, images with curves, and so forth [12, 13], researchers have generalized the sparse construction theory to the context of sparse dictionary construction, see e.g. [14, 15] and the references therein. Normally, sparsity is represented not with respect to an orthogonal basis but with respect to a redundant dictionary $D \in \mathbb{R}^{n\times p}$ $(n \leq p)$, that is, the signal x is now represented as $x = Df$, in which f is (approximately) sparse.

In the present paper, it is assumed that D is a tight frame. Formally, $D \in \mathbb{R}^{n\times p}$ $(n \leq p)$ constructs a tight frame for $\mathbb{R}^n$, when

$$x = \sum_{j=1}^{p} \langle x, D_k \rangle D_k,$$

for all $x \in \mathbb{R}^n$, in which D_j $(1 \leq j \leq p)$ stand for the columns of D, and $\langle \cdot, \cdot \rangle$ represents the Euclidean inner product. The restricted isometry property adapted to D (D-RIP) for sparse dictionary recovery is one of the most extensively utilized frameworks, which was firstly introduced by [16] and is a natural generalization of standard RIP.

Definition 1.1. *(D-RIP). Let D be an $n \times p$ matrix. A measurement matrix A is said to satisfy the restricted isometry property adapted to D (for short D-RIP) of order k with constant δ if*

$$(1-\delta)\|Du\|_2^2 \leq \|ADu\|_2^2 \leq (1+\delta)\|Du\|_2^2 \tag{1.2}$$

holds for all k sparse vectors $u \in \mathbb{R}^p$. The D-RIP constant δ_k is defined as the smallest number δ such that (1.2) holds for all k sparse vectors $u \in \mathbb{R}^p$.

For recovering the signal x from (1.1), researcher [17] has considered the following analysis LASSO model (written as ALASSAO):

$$\min_{\tilde{x}\in\mathbb{R}^n} \lambda\|D^*\tilde{x}\|_1 + \frac{1}{2}\|A\tilde{x} - b\|_2^2, \tag{1.3}$$

in which λ is a tuning parameter containing the tolerance to the noise term. Numerous sufficient conditions have been proposed, which can guarantee the stable reconstruction of the signal x via the method (1.3). These consists of $\delta_{3k} < 1/4$ [18], $\delta_{2k} < 0.1907$ [19], and $\delta_{2k} < 0.2$ [19].

Some works [19, 21] have proposed numerical algorithms to solve the analysis LASSO. Especially, the monotone version of fast iterative shrinkage thresholding algorithm (MFISTA) has been deemed by [19] to solve the problem (1.3). However, it is extremely difficult to compute the proximal operator of $\|D^*x\|_1$, therefore, they proposed the relaxed ALASSO model (abbreviated as RALASSO) which is based on decomposition technique as follows:

$$\min_{\tilde{x}\in\mathbb{R}^n,\ \tilde{z}\in\mathbb{R}^p} \lambda\|\tilde{z}\|_1 + \frac{1}{2}\|A\tilde{x} - b\|_2^2 + \frac{\rho}{2}\|D^*\tilde{x} - \tilde{z}\|_2^2. \tag{1.4}$$

Some sufficient conditions ensure that the signal x can be stably recovered via the approach (1.4) have been established, e.g. $\delta_{2k} < 0.1907$ [19], $\delta_{2k} < 0.2$ [20] and $\delta_{tk} < \sqrt{(t-1)/(t+8)}$ for $t > 1$ [22].

Among the recent studies in non-convex compressed sensing, it has been showed [23–26] that ℓ_q minimization $(0 < q < 1)$ requires remarkably fewer linear measurements for reconstruction of sparse signals than that by ℓ_1

minimization. Consequently, in this paper, we are interesting in the following non-convex relaxed $\ell_q(0 < q \leq 1)$ minimization problem:

$$\min_{\tilde{x}\in\mathbb{R}^n,\ \tilde{z}\in\mathbb{R}^p} \lambda\|\tilde{z}\|_q + \frac{1}{2}\|A\tilde{x}-b\|_2^2 + \frac{\rho}{2}\|D^*\tilde{x}-\tilde{z}\|_2^2. \tag{1.5}$$

Particularly, one can see that $\tilde{z} = D^*\tilde{x}$ due to the third term and the problem (1.5) is equivalent to the following unconstrained non-convex ℓ_q minimization problem provided that the parameter ρ tends to ∞.

$$\min_{\tilde{x}\in\mathbb{R}^n} \lambda\|D^*\tilde{x}\|_q + \frac{1}{2}\|A\tilde{x}-b\|_2^2. \tag{1.6}$$

However generally, compared to the problem (1.6), it is not hard to solve the problem (1.5) in applications.

In this paper, recovery guarantees for the non-convex relaxed $\ell_q(0 < q \leq 1)$ minimization model (1.5) are presented. One of the main results is that the sufficient condition for recovering sparse or non-sparse signals via the method (1.5) and the associating performance bound are established by exploiting the ℓ_q robust D-Null Space Property (ℓ_q D-NSP), in which the term $\|D^*D\|_1$ in the noise bound constant is removed. Additionally, with the D-RIP, we derive a new condition for recovery of sparse signals via the method (1.5) and the corresponding error estimation of upper bound. In particular, when $q = 1$, our result is the same as that of Theorem IV.1 [19]. Furthermore, in the case of $\rho \to \infty$, as by-product, we gain the recovery bound for the model (1.6) from the mentioned before results.

We use the following notation throughout this paper. We call the set of indices of nonzero elements for a vector x as the support of x represented as $\mathrm{supp}(x)$. A vector x is called as k-sparse if $|\mathrm{supp}(x)| \leq k$. For a given index set $T \in \{1, 2, \cdots, n\}$, T^c stands for the complement of T in $\{1, 2, \cdots, n\}$. For a given matrix $A \in \mathbb{R}^{m\times n}$, A_T denotes the sub-matrix of A that is constructed from the columns of A indexed by T, as well as putting all other columns to zero. Let A^* indicate the conjugate transpose of a matrix A, so that A_T^* denotes $(A_T)^*$. For a vector $x \in \mathbb{R}^n$ and $q > 0$, define $\|x\|_q = (\sum_{i=1}^n |x_i|^q)^{\frac{1}{q}}$, and $\|x\|_\infty = \max_{1\leq i\leq n} |x_i|$. $\|A\|_{p,q}$ stands for the norm of A from ℓ_p to ℓ_q.

The paper is constructed as follows. In Section 2, we present the main results. The auxiliary lemmas that are used in the proof of main results are given in Section 3. In Section 4, we prove the main results.

2 Main results

One can easily see that the authors [19] proved recovery guarantees for the model (1.5) as $q = 1$ with D-RIP provided that A satisfies the D-RIP $\delta_{2k} < 0.1907$. Unfortunately, the performance bound, which was given by them, excessively relies on the $\|D^*D\|_1$. As a result, to a certain extent, the error bound to the model (1.5) for $q = 1$ is not better. Hence, based on the ℓ_q robust D-Null Space Property, we derive the performance bound on the model (1.5) which doesn't involve the term $\|D^*D\|_1$. Let D_T^*x denote the best k-sparse approximation of D^*x. The definition of ℓ_q D-NSP is as follows:

Definition 2.1. *(ℓ_q D-NSP [27–29]) Let D be a matrix with $n \times p$. For any $x \in \mathbb{R}^n$, there exist constants $\beta > 0$*

and $0 < \gamma < 3^{1-\frac{2}{q}}$ *such that*

$$\|D_T^* x\|_2 \leq \beta \|Ax\|_2 + \frac{\gamma}{k^{\frac{1}{q}-\frac{1}{2}}} \|D^* x - D_T^* x\|_q,$$

then the matrix A *is said to fulfill the* ℓ_q *robust* D*-Null Space Property of order* k *with constants* β *and* γ*.*

Now, it is assumed that A obeys ℓ_q D-NSP with constants β and γ, we study the recovery condition on the model (1.5). Our result is stated as follows.

Theorem 2.2. *Let* $A \in \mathbb{R}^{m\times n}$ *be a measurement matrix,* $D \in \mathbb{R}^{n\times p}$ *be a tight frame, and let* A *obeys* ℓ_q *D-NSP with constants* β *and* γ*. It is assumed that* $\hat{x}$ *is the solution to the problem (1.5). Think about the measurement* $b = Ax + w$*, in which* w *is noise fulfilling* $\|D^* A^* w\|_\infty \leq \frac{\lambda}{2}$*. Then*

$$\|\hat{x} - x\|_2 \leq C_0 \lambda k^{1-\frac{q}{2}} + C_1 \|D_{T^c}^{\tilde{*}} x\|_q + C_2 \tilde{C}(\lambda, \rho), \tag{2.7}$$

in which the definitions of $\|D_{T^c}^{\tilde{*}} x\|_q$ *and* $\tilde{C}(\lambda,\rho)$ *are provided in the proof of Theorem 2.2,* C_0 *is constant which relies on* $\beta,\ \gamma,\ q$*,* $C_1,\ C_2$ *are constants which count on* $\beta,\ \gamma,\ q,\ k$*.*

In the following, based on D-RIP, we discuss the recovery condition that can guarantee a sparse signal to be stably reconstructed via the method (1.5). The result is described as follows.

Theorem 2.3. *Let* A *be an* $m \times n$ *measurement matrix,* D *be an* $n \times p$ *tight frame, and suppose that* A *satisfies the D-RIP with* $\delta_{2k} < 1/(1+3^{1/q}\sqrt{2})$*. Let* $b = Ax + w$*, in which* w *is noise that meet* $\|D^* A^* w\|_\infty \leq \frac{\lambda}{2}$*. Assume that* $\hat{x}$ *is the solution to the problem (1.5). Then*

$$\|\hat{x} - x\|_2 \leq \tilde{C}_1 \lambda \sqrt{k} + \tilde{C}_2 \|D^* x - D_T^* x\|_q + \tilde{C}_3 C^{\frac{1}{q}}(\lambda, \rho), \tag{2.8}$$

where $\tilde{C}_1$ *is constant relying on* $\delta_{2k},\ q,\ \|D^* D\|_{1,1}$*, and* $\tilde{C}_2,\ \tilde{C}_3$ *are constants relying on* $\delta_{2k},\ q,\ k$*.*

Remark 2.4. *In the case of* $q = 1$*, our sufficient condition reduces to that of Theorem IV.1 in [19].*

Remark 2.5. *We show that different choices of* q *can lead to different conditions. For example, see Table 2.1.*

q	Recovery condition
0.2	0.002901461
0.5	0.072844236
0.8	0.151891909
0.9	0.172608221
1	0.19074357

Table 2.1: Different sufficient conditions.

Selecting ρ converges to ∞ in the model (1.5) for in which $\tilde{z} = D^* \tilde{x}$ results in the problem (1.6). Consequently, the below result holds.

Proposition 2.6. *Let A be a measurement matrix with $m \times n$, D be a tight frame with $n \times p$, and it is assumed that A fulfills the D-RIP with $\delta_{2k} < 1/(1+3^{1/q}\sqrt{2})$. Take the measurement $b = Ax + w$ into account, in which w is noise satisfying $\|D^*A^*w\|_\infty \le \frac{\lambda}{2}$. Let $\hat{x}$ be the optimal solution of (1.6). Then*

$$\|\hat{x} - x\|_2 \le \tilde{C}_1 \lambda \sqrt{k} + \tilde{C}_2 \|D^*x - D_T^*x\|_q. \tag{2.9}$$

Here $\tilde{C}_1$ is constant depending on $\delta_{2k},\ q,\ \|D^*D\|_{1,1}$, and $\tilde{C}_2$ is constants counting on $\delta_{2k},\ q,\ k$.

3 Technical lemmas

Before proving the main results, we first provide some auxiliary lemmas which will be utilized in the proof of main results.

Lemma 3.1. *The solution $\hat{x}$ to the problem (1.5) fulfills the following inequalities,*
(a)

$$\|Ah\|_2^2 + \lambda \|D_{T^c}^*h\|_q^q \le 3\lambda \|D_T^*h\|_q^q + 4\lambda \|D_{T^c}^*x\|_q^q + 2\lambda C(\lambda, \rho), \tag{3.10}$$

(b)

$$\|D_{T^c}^*h\|_q^q \le \|D_T^*h\|_q^q + 2\|D_{T^c}^*x\|_q^q + C(\lambda, \rho) + \frac{1}{2}\|D^*h\|_1, \tag{3.11}$$

where $C(\lambda,\rho) = p(\frac{\lambda}{\rho})^{\frac{q}{2-q}} \left(q^{\frac{q}{2-q}} - \frac{1}{2} q^{\frac{2}{2-q}} \right)$.

Lemma 3.2. *Let D be a tight frame, and $h = \hat{x} - x$ be recovery error to the problem (1.5). Suppose that A fulfills the D-RIP with δ_{2k}. Then,*

$$\left\langle Ah, ADD_{T_{01}}^*h \right\rangle \ge (1-\delta_{2k})\|D_{T_{01}}^*h\|_2^2 - \sqrt{2}k^{\frac{1}{2}-\frac{1}{q}}\sigma_{2k}\|D_{T_{01}}^*h\|_2 \|D_{T^c}^*h\|_q. \tag{3.12}$$

Lemma 3.3. *Let $\hat{x}$ be the solution to the problem (1.5). Then,*

$$\|D^*A^*Ah\|_\infty \le \left(\frac{1}{2} + \|D^*D\|_{1,1} \right) \lambda. \tag{3.13}$$

4 Proofs

Let $T_0 = T$ be the set of indices of the k largest entries of D^*x in magnitude. The set T^c is decomposed into sets of size k. Those sets are stood as $T_1, T_2, \cdots, T_J$ (only the number of elements of T_J is probably less than k), in which T_1 represents the indices of the k largest entries of $D_{T^c}^*x$ in magnitude, T_2 represents the indices of the next k largest elements of $D_{T^c}^*x$ in magnitude, etc. Set $T_{01} = T_0 \bigcup T_1$. For convenient, denote $h = \hat{x} - x$ the reconstruction error with $\hat{x}$ being the optimal solution to (1.5). In the proof, we use the below essential inequalities related to the l_q (quasi)norm. For any vectors $u, v \in \mathbb{R}^n$, $\|u\|_{q_2} \le \|v\|_{q_1} \le n^{1/p_1 - 1/p_2}\|u\|_{q_2}$, $0 < p_1 \le p_2 \le \infty$. The triangle inequality on $\|\cdot\|_q^q$ for $0 < q \le 1$: $\|u+v\|_q^q \le \|u\|_q^q + \|v\|_q^q$.

The proof of Theorem 2.2. By (4.32), it results in

$$\sum_{j=1}^{J}\|D_{T_j}^*h\|_2^q \le \frac{\sum_{j=1}^{J}\|D_{T_j}^*h\|_q^q+\|D_T^*h\|_q^q}{k^{1-\frac{q}{2}}}\le \frac{\|D_{T^c}^*h\|_q^q+\|D_T^*h\|_q^q}{k^{1-\frac{q}{2}}},$$

so which combines with the inequality that $\|u\|_q\le 2^{\frac{1}{q}-1}\|u\|_1$ for any $u\in\mathbb{R}^2$, implies

$$\begin{aligned}\|D_{T^c}^*h\|_2 \le&\sum_{j=1}^{J}\|D_{T_j}^*h\|_2\le\left(\sum_{j=1}^{J}\|D_{T_j}^*h\|_2^q\right)^{\frac{1}{q}}\le\frac{\left(\|D_{T^c}^*h\|_q^q+\|D_T^*h\|_q^q\right)^{\frac{1}{q}}}{k^{\frac{1}{q}-\frac{1}{2}}}\\ \le&\frac{2^{\frac{1}{q}-1}(\|D_{T^c}^*h\|_q+\|D_T^*h\|_q)}{k^{\frac{1}{q}-\frac{1}{2}}}.\end{aligned}\tag{4.14}$$

By Lemma 3.1(a), we get

$$\|D_{T^c}^*h\|_q^q\le 3\|D_T^*h\|_q^q+4\|D_{T^c}^*x\|_q^q+2C(\lambda,\rho).\tag{4.15}$$

So that

$$\begin{aligned}\|D_{T^c}^*h\|_q\le&\left(3\|D_T^*h\|_q^q+4\|D_{T^c}^*x\|_q^q+2C(\lambda,\rho)\right)^{\frac{1}{q}}\\ \le&3^{\frac{1}{q}-1}2^{\frac{1}{q}}C^{\frac{1}{q}}(\lambda,\rho)+3^{\frac{2}{q}-1}\|D_T^*h\|_q+3^{\frac{1}{q}-1}2^{\frac{2}{q}}\|D_{T^c}^*h\|_q.\end{aligned}\tag{4.16}$$

Plugging into (4.14), it follows that

$$\|D_{T^c}^*h\|_2\le\frac{2^{\frac{1}{q}-1}}{k^{\frac{1}{q}-\frac{1}{2}}}\left(3^{\frac{1}{q}-1}2^{\frac{1}{q}}C^{\frac{1}{q}}(\lambda,\rho)+(3^{\frac{2}{q}-1}+1)\|D_T^*h\|_q+3^{\frac{1}{q}-1}2^{\frac{2}{q}}\|D_{T^c}^*h\|_q\right).\tag{4.17}$$

By (4.17), we get

$$\begin{aligned}\|h\|_2=&\|D^*h\|_2=\sqrt{\|D_{T^c}^*h\|_2^2+\|D_T^*h\|_2^2}\\ \le&\left(\|D_{T^c}^*h\|_2^2+\frac{2^{\frac{2}{q}-2}}{k^{\frac{2}{q}-1}}\left(3^{\frac{1}{q}-1}2^{\frac{1}{q}}C^{\frac{1}{q}}(\lambda,\rho)+(3^{\frac{2}{q}-1}+1)\|D_T^*h\|_q+3^{\frac{1}{q}-1}2^{\frac{2}{q}}\|D_{T^c}^*h\|_q\right)^2\right)^{\frac{1}{2}}\\ \le&\left[1+2^{\frac{1}{q}-1}(3^{\frac{2}{q}-1}+1)\right]\|D_T^*h\|_2+\frac{2^{\frac{1}{q}-1}}{k^{\frac{1}{q}-\frac{1}{2}}}\left(3^{\frac{1}{q}-1}2^{\frac{1}{q}}C^{\frac{1}{q}}(\lambda,\rho)+3^{\frac{1}{q}-1}2^{\frac{2}{q}}\|D_{T^c}^*h\|_q\right).\end{aligned}\tag{4.18}$$

In the following, we bound the term $\|D_T^*h\|_2$. By utilizing the ℓ_q D-NSP of the measurement matrix A, it implies that

$$\begin{aligned}\|D_T^*x\|_2\le&\beta\|Ah\|_2+\frac{\gamma}{k^{\frac{1}{q}-\frac{1}{2}}}\|D^*h-D_T^*h\|_q\\ \overset{(a)}{\le}&\beta\|Ah\|_2+\frac{\gamma}{k^{\frac{1}{q}-\frac{1}{2}}}\left(3^{\frac{1}{q}-1}2^{\frac{1}{q}}C^{\frac{1}{q}}(\lambda,\rho)+3^{\frac{2}{q}-1}\|D_T^*h\|_q+3^{\frac{1}{q}-1}2^{\frac{2}{q}}\|D_{T^c}^*x\|_q\right)\\ \overset{(b)}{\le}&\beta\|Ah\|_2+\frac{3^{\frac{1}{q}-1}2^{\frac{1}{q}}C^{\frac{1}{q}}(\lambda,\rho)\gamma}{k^{\frac{1}{q}-\frac{1}{2}}}+\gamma 3^{\frac{2}{q}-1}\|D_T^*h\|_2+\frac{3^{\frac{1}{q}-1}2^{\frac{2}{q}}\gamma}{k^{\frac{1}{q}-\frac{1}{2}}}\|D_{T^c}^*x\|_q,\end{aligned}\tag{4.19}$$

where (a) is from (4.16), and (b) is due to the inequality that $\|u\|_q\le k^{\frac{1}{q}-\frac{1}{2}}\|u\|_2$ for any $u\in\mathbb{R}^k$. Thereupon,

$$\|D_T^*x\|_2\le\frac{1}{1-\gamma 3^{\frac{2}{q}-1}}\left(\beta\|Ah\|_2+\frac{3^{\frac{1}{q}-1}2^{\frac{1}{q}}C^{\frac{1}{q}}(\lambda,\rho)\gamma}{k^{\frac{1}{q}-\frac{1}{2}}}+\frac{3^{\frac{1}{q}-1}2^{\frac{2}{q}}\gamma}{k^{\frac{1}{q}-\frac{1}{2}}}\|D_{T^c}^*x\|_q\right).\tag{4.20}$$

By making use of Lemma 3.1(a) again, it leads to

$$\begin{aligned}\|Ah\|_2^2 \leq & 3\lambda\|D_T^*h\|_q^q + 4\lambda\|D_{T^c}^*x\|_q^q + 2\lambda C(\lambda,\rho) \\ \leq & 3\lambda k^{1-\frac{q}{2}}\|D_T^*h\|_2^q + 4\lambda\|D_{T^c}^*x\|_q^q + 2\lambda C(\lambda,\rho) \\ \leq & \frac{3\lambda k^{1-\frac{q}{2}}}{(1-\gamma 3^{\frac{2}{q}-1})^q}\left(\beta\|Ah\|_2 + \frac{3^{\frac{1}{q}-1}2^{\frac{1}{q}}C^{\frac{1}{q}}(\lambda,\rho)\gamma}{k^{\frac{1}{q}-\frac{1}{2}}} + \frac{3^{\frac{1}{q}-1}2^{\frac{2}{q}}\gamma}{k^{\frac{1}{q}-\frac{1}{2}}}\|D_{T^c}^*x\|_q\right)^q + +4\lambda\|D_{T^c}^*x\|_q^q + 2\lambda C(\lambda,\rho) \\ \leq & \frac{3\lambda k^{1-\frac{q}{2}}\beta^q}{(1-\gamma 3^{\frac{2}{q}-1})^q}\|Ah\|_2^q + 2\lambda\left(\frac{3^{2-q}\gamma^q}{(1-\gamma 3^{\frac{2}{q}-1})^q}+1\right)C(\lambda,\rho) + 4\lambda\left(\frac{3^{2-q}\gamma^q}{(1-\gamma 3^{\frac{2}{q}-1})^q}+1\right)\|D_{T^c}^*x\|_q^q,\end{aligned}$$

which deduces

$$\|Ah\|_2 \leq \frac{3\lambda k^{1-\frac{q}{2}}\beta^q}{(1-\gamma 3^{\frac{2}{q}-1})^q} + \frac{2[(1-\gamma 3^{\frac{2}{q}-1})^q + 3^{2-q}\gamma^q]}{3k^{1-\frac{q}{2}}\beta^q}[C(\lambda,\rho) + 2\|D_{T^c}^*x\|_q^q]. \tag{4.21}$$

Putting (4.21) into (4.20), we get

$$\begin{aligned}\|D_T^*x\|_2 \leq & \frac{3\lambda k^{1-\frac{q}{2}}\beta^{q+1}}{(1-\gamma 3^{\frac{2}{q}-1})^{q+1}} + \frac{1}{1-\gamma 3^{\frac{2}{q}-1}}\left[\frac{2[(1-\gamma 3^{\frac{2}{q}-1})^q + 3^{2-q}\gamma^q]}{3k^{1-\frac{q}{2}}\beta^{q-1}} + \frac{3^{\frac{1}{q}-1}2^{\frac{1}{q}}\gamma}{k^{\frac{1}{q}-\frac{1}{2}}}\right]\tilde{C}(\lambda,\rho) \\ & + \frac{1}{1-\gamma 3^{\frac{2}{q}-1}}\left[\frac{2^2[(1-\gamma 3^{\frac{2}{q}-1})^q + 3^{2-q}\gamma^q]}{3k^{1-\frac{q}{2}}\beta^{q-1}} + \frac{3^{\frac{1}{q}-1}2^{\frac{2}{q}}\gamma}{k^{\frac{1}{q}-\frac{1}{2}}}\right]\|\tilde{D_{T^c}^*x}\|_q,\end{aligned} \tag{4.22}$$

in which $\tilde{C}(\lambda,\rho) = \max\{C(\lambda,\rho), C^{\frac{1}{q}}(\lambda,\rho)\}$ and $\|\tilde{D_{T^c}^*x}\|_q = \max\{\|D_{T^c}^*x\|_q^q, \|D_{T^c}^*x\|_q\}$. Plugging (4.22) to (4.18), we obtain

$$\begin{aligned}\|h\|_2 \leq & \frac{3\lambda k^{1-\frac{q}{2}}\beta^{q+1}}{(1-\gamma 3^{\frac{2}{q}-1})^{q+1}}[1 + 2^{\frac{1}{q}-1}(3^{\frac{2}{q}-1}+1)] \\ & + \left\{\frac{1 + 2^{\frac{1}{q}-1}(3^{\frac{2}{q}-1}+1)}{1-\gamma 3^{\frac{2}{q}-1}}\left[\frac{2[(1-\gamma 3^{\frac{2}{q}-1})^q + 3^{2-q}\gamma^q]}{3k^{1-\frac{q}{2}}\beta^{q-1}} + \frac{3^{\frac{1}{q}-1}2^{\frac{1}{q}}\gamma}{k^{\frac{1}{q}-\frac{1}{2}}}\right] + \frac{2^{\frac{2}{q}-1}3^{\frac{1}{q}-1}}{k^{\frac{1}{q}-\frac{1}{2}}}\right\}\tilde{C}(\lambda,\rho) \\ & + \left\{\frac{1 + 2^{\frac{1}{q}-1}(3^{\frac{2}{q}-1}+1)}{1-\gamma 3^{\frac{2}{q}-1}}\left[\frac{2^2[(1-\gamma 3^{\frac{2}{q}-1})^q + 3^{2-q}\gamma^q]}{3k^{1-\frac{q}{2}}\beta^{q-1}} + \frac{3^{\frac{1}{q}-1}2^{\frac{2}{q}}\gamma}{k^{\frac{1}{q}-\frac{1}{2}}}\right] + \frac{2^{\frac{3}{q}-1}3^{\frac{1}{q}-1}}{k^{\frac{1}{q}-\frac{1}{2}}}\right\}\|\tilde{D_{T^c}^*x}\|_q.\end{aligned}$$

The proof is completed. □

The proof of Theorem 2.3. By Lemma 3.2, it follows that

$$\langle Ah, ADD_{T_{01}}^*h\rangle \geq (1-\delta_{2k})\|D_{T_{01}}^*h\|_2^2 - \sqrt{2}k^{\frac{1}{2}-\frac{1}{q}}\sigma_{2k}\|D_{T_{01}}^*h\|_2\|D_{T^c}^*h\|_q. \tag{4.23}$$

Let $\alpha = \frac{1}{2} + \|D^*D\|_{1,1}$. By employing the Hölder inequality, we get

$$\langle Ah, ADD_{T_{01}}^*h\rangle = \langle D^*A^*Ah, D_{T_{01}}^*h\rangle \leq \|D^*A^*Ah\|_\infty\|D_{T_{01}}^*h\|_1 \overset{(a)}{\leq} \sqrt{2k}\alpha\lambda\|D_{T_{01}}^*h\|_2, \tag{4.24}$$

where (a) follows from the Cauchy-Schwartz inequality.

Now combining with (4.23) and (4.24), and by a basic calculation, it leads to

$$\|D_{T_{01}}^*h\|_2 \leq \frac{\sqrt{2k}\alpha\lambda + \sqrt{2}k^{\frac{1}{2}-\frac{1}{q}}\sigma_{2k}\|D_{T^c}^*h\|_q}{1-\sigma_{2k}}. \tag{4.25}$$

From the inequality that $\|u\|_q \le k^{\frac{1}{q}-\frac{1}{2}}\|u\|_2$ for any $u \in \mathbb{R}^k$ together with (4.25), we get

$$\begin{aligned}\|D_T^* h\|_q \le & k^{\frac{1}{q}-\frac{1}{2}}\|D_T^* h\|_2 \le k^{\frac{1}{q}-\frac{1}{2}}\|D_{T_{01}}^* h\|_2 \\ \le & \frac{\sqrt{2}k^{\frac{1}{q}}\alpha\lambda + \sqrt{2}\sigma_{2k}\|D_{T^c}^* h\|_q}{1-\sigma_{2k}}.\end{aligned} \tag{4.26}$$

By substituting (4.26) into (3.11), utilizing the inequality that $(a+b)^q \le a^q + b^q$ for any $a,\ b \ge 0$ and by an elementary computation, we get

$$\|D_{T^c}^* h\|_q^q \le \frac{2^{\frac{q}{2}}3k\alpha^q\lambda^q}{(1-\sigma_{2k})^q} + \frac{2^{\frac{q}{2}}3\sigma_{2k}^q\|D_{T^c}^* h\|_q^q}{(1-\sigma_{2k})^q} + 2C(\lambda,\rho) + 4\|D_{T^c}^* x\|_q^q. \tag{4.27}$$

Since $\delta_{2k} < 1/(1+3^{1/q}\sqrt{2})$, we get the below estimation

$$\|D_{T^c}^* h\|_q^q \le \left(1 - \frac{2^{\frac{q}{2}}3\sigma_{2k}^q}{(1-\sigma_{2k})^q}\right)^{-1}\left(\frac{2^{\frac{q}{2}}3k\alpha^q\lambda^q}{(1-\sigma_{2k})^q} + 2C(\lambda,\rho) + 4\|D_{T^c}^* x\|_q^q\right).$$

The above inequality together with the inequality that $\|u\|_q \le k^{\frac{1}{q}-1}\|u\|_1$ for any $u \in \mathbb{R}^k$ implies

$$\|D_{T^c}^* h\|_q \le \left(1 - \frac{2^{\frac{q}{2}}3\sigma_{2k}^q}{(1-\sigma_{2k})^q}\right)^{-\frac{1}{q}} 3^{\frac{1}{q}-1}\left(\frac{2^{\frac{1}{2}}3^{\frac{1}{q}}k^{\frac{1}{q}}\alpha\lambda}{1-\sigma_{2k}} + 2^{\frac{1}{q}}C^{\frac{1}{q}}(\lambda,\rho) + 2^{\frac{2}{q}}\|D_{T^c}^* x\|_q\right). \tag{4.28}$$

Now we establish the bound regarding the recovery error. Notice that

$$\|h\|_2 = \|D^* h\|_2 \le \|D_{T_{01}}^* h\|_2 + \sum_{j\ge 2}\|D_{T_j}^* h\|_2.$$

By utilizing (4.25) and (4.33) to the above inequality, we have

$$\|h\|_2 \le \frac{\sqrt{2k}\alpha\lambda}{1-\sigma_{2k}} + \frac{[(\sqrt{2}-1)\sigma_{2k}+1]k^{\frac{1}{2}-\frac{1}{q}}\|D_{T^c}^* h\|_q}{1-\sigma_{2k}}.$$

Introducing (4.28) to the above, we gain

$$\begin{aligned}\|h\|_2 \le & \left[1 + \frac{[(\sqrt{2}-1)\sigma_{2k}+1]3^{\frac{2}{q}-1}}{[(1-\sigma_{2k})^q - 2^{\frac{q}{2}}3\sigma_{2k}^q]^{\frac{1}{q}}}\right]\frac{\sqrt{2k}\alpha\lambda}{1-\sigma_{2k}} \\ & + \frac{[(\sqrt{2}-1)\sigma_{2k}+1]3^{\frac{1}{q}-1}2^{\frac{1}{q}}k^{\frac{1}{2}-\frac{1}{q}}}{[(1-\sigma_{2k})^q - 2^{\frac{q}{2}}3\sigma_{2k}^q]^{\frac{1}{q}}}\left(C^{\frac{1}{q}}(\lambda,\rho) + 2^{\frac{1}{q}}\|D_{T^c}^* x\|_q\right).\end{aligned}$$

We complete the proof. □

The proof of Lemma 3.1. Since the proof of (a) is similar to that of (b), we only give the proof of (b).

Note that $(\hat{x},\ \hat{z})$ is minimum to (1.5), for any group $(x,\ z)$ that obeys $z = D^* x$, so it leads to

$$\frac{1}{2}\|A\hat{x} - b\|_2^2 + \lambda\|\hat{z}\|_q^q + \frac{\rho}{2}\|\hat{z} - D^*\hat{x}\|_2^2 \le \frac{1}{2}\|Ax - b\|_2^2 + \lambda\|D^* x\|_q^q.$$

Due to $b = Ax + w$ and $h = \hat{x} - x$, it results in

$$\frac{1}{2}\|Ah\|_2^2 - < Ah, w > \le \lambda\|D^* x\|_q^q - \lambda\|\hat{z}\|_q^q - \frac{\rho}{2}\|\hat{z} - D^*\hat{x}\|_2^2.$$

By the concept of tight frame, and then employing the Hölder inequality and the assumption $\|D^*A^*w\|_\infty \le \frac{\lambda}{2}$, we get

$$< Ah, w >=< D^*h, D^*A^*w >\le \|D^*A^*w\|_\infty \|D^*h\|_1 \le \frac{\lambda}{2}\|D^*h\|_1.$$

Thereby,

$$\frac{1}{2}\|Ah\|_2^2 - \frac{\lambda}{2}\|D^*h\|_1 \le \lambda\|D^*x\|_q^q - \lambda\|\hat{z}\|_q^q - \frac{\rho}{2}\|\hat{z} - D^*\hat{x}\|_2^2.$$

By the first-order optimality condition of (1.5) for the variable $\tilde{z}$, we get

$$\lambda\theta + \rho(\hat{z} - D^*\hat{x}) = 0,$$

in which θ is a sub-gradient of the function $\|z\|_q^q$ and $\theta_i = q|z_i|^{q-1}$.

Set $C(\lambda,\rho) = p(\frac{\lambda}{\rho})^{\frac{q}{2-q}}\left(q^{\frac{q}{2-q}} - \frac{1}{2}q^{\frac{2}{2-q}}\right)$. Substituting the term $\hat{z}$ with $\hat{z} = -\frac{\lambda}{\rho}\theta + D^*\hat{x}$, we get

$$\begin{aligned}\frac{1}{2}\|Ah\|_2^2 - \frac{\lambda}{2}\|D^*h\|_1 \le&\lambda\|D^*x\|_q^q - \lambda\| - \frac{\lambda}{\rho}\theta + D^*\hat{x}\|_q^q - \frac{\rho}{2}\| - \frac{\lambda}{\rho}\theta\|_2^2\\ \le&\lambda\|D^*x\|_q^q - \lambda\|D^*\hat{x}\|_q^q + \frac{\lambda^2}{\rho}\|\theta\|_q^q - \frac{\lambda^2}{2\rho}\|\theta\|_2^2\\ \overset{(a)}{\le}&\lambda\|D^*x\|_q^q - \lambda\|D^*\hat{x}\|_q^q + \lambda C(\lambda,\rho),\end{aligned}$$

where (a) follows from the fact that the function $\frac{\lambda^{q+1}}{\rho^q}\theta_i^q - \frac{\lambda^2}{2\rho}\theta_i^2$ reaches the maximum at $\theta_i = q^{\frac{1}{2-q}}(\frac{\lambda}{\rho})^{\frac{q-1}{2-q}}$.

It accordingly follows that

$$\|D^*\hat{x}\|_q^q \le C(\lambda,\rho) + \frac{1}{2}\|D^*h\|_1 + \|D^*x\|_q^q. \tag{4.29}$$

Because of $h = \hat{x} - x$, we get

$$\begin{aligned}&\|D^*\hat{x}\|_q^q\\ =&\|D^*h + D^*x\|_q^q\\ =&\|D_T^*h + D_T^*x\|_q^q + \|D_{T^c}^*h + D_{T^c}^*x\|_q^q\\ \ge&\|D_T^*x\|_q^q - \|D_T^*h\|_q^q + \|D_{T^c}^*h\|_q^q - \|D_{T^c}^*x\|_q^q.\end{aligned} \tag{4.30}$$

Observe that

$$\begin{aligned}&C(\lambda,\rho) + \frac{1}{2}\|D^*h\|_1 + \|D^*x\|_q^q\\ &\le C(\lambda,\rho) + \frac{1}{2}\|D^*h\|_1 + \|D_T^*x\|_q^q + \|D_{T^c}^*x\|_q^q.\end{aligned} \tag{4.31}$$

Combining with (4.29), (4.30) and (4.31), we get

$$\|D_{T^c}^*h\|_q^q \le \|D_T^*h\|_q^q + 2\|D_{T^c}^*x\|_q^q + C(\lambda,\rho) + \frac{1}{2}\|D^*h\|_1.$$

□

The proof of Lemma 3.2. For each $t \in T_j$ and $s \in T_{j-1}$, $|D^*_{T_j} h(t)| \leq |D^*_{T_{j-1}} h(s)|$, then

$$|D^*_{T_j} h(t)|^q \leq \frac{\|D^*_{T_{j-1}} h\|_q^q}{k}.$$

It follows that

$$|D^*_{T_j} h(t)|^2 \leq \frac{\|D^*_{T_{j-1}} h\|_q^2}{k^{\frac{2}{q}}},$$
$$\|D^*_{T_j} h\|_2^2 \leq \frac{\|D^*_{T_{j-1}} h\|_q^2}{k^{\frac{2}{q}-1}},$$
$$\|D^*_{T_j} h\|_2^q \leq \frac{\|D^*_{T_{j-1}} h\|_q^q}{k^{1-\frac{q}{2}}}.$$

Hence,

$$\sum_{j=2}^{J} \|D^*_{T_j} h\|_2^q \leq \frac{\sum_{j=1}^{J} \|D^*_{T_j} h\|_q^q}{k^{1-\frac{q}{2}}} \leq \frac{\|D^*_{T^c} h\|_q^q}{k^{1-\frac{q}{2}}}. \tag{4.32}$$

(4.32) combines with the inequality that $\|u\|_1 \leq \|u\|_q$, for any $u \in \mathbb{R}^p$ implies

$$\sum_{j=2}^{J} \|D^*_{T_j} h\|_2 \leq \left(\sum_{j=2}^{J} \|D^*_{T_j} h\|_2^q \right)^{\frac{1}{q}} = \frac{\|D^*_{T^c} h\|_q}{k^{\frac{1}{q}-\frac{1}{2}}}. \tag{4.33}$$

By the proof of Lemma IV.2 [19], it leads to

$$< Ah, ADD^*_{T_{01}} h >\geq (1-\delta_{2k})\|D^*_{T_{01}} h\|_2^2 - \sqrt{2}\sigma_{2k} \|D^*_{T_{01}} h\|_2 \sum_{j=2}^{J} \|D^*_{T_j} h\|_2. \tag{4.34}$$

A combination of (4.33) and (4.34), the desired result follows. □

5 Acknowledgements

The work of J. Huang was supported in part by the National Natural Science Foundation of China (Grant No. 12101454), and in part by the Fuxi Scientific Research Innovation Team of Tianshui Normal University (No. FXD2020-03). The work of J. Wang was supported in part by the National Natural Science Foundation of China (Grant No. 12071380). The work of F. Zhang was supported in part by the National Natural Science Foundation of China (No. Grant 12101512), in part by Fundamental Research Funds for the Central Universities (Grant No. SWU120078), and China Postdoctoral Science Foundation (Grant No. 2021M692681). The work of J. Jia was supported in part by the National Natural Science Foundation of China (Grant No. 62063031), in part by the Natural Science Foundation of Gansu Province (No. 21JR1RE292), and in part by the College Innovation Ability Promotion Project of Gansu Province (No. 2019A-100). The work of Z. Chang was supported in part by Natural Science Foundation of Gansu Province(No. 21JR7RE172).

References

[1] Candes E J, Romberg J K, Tao T. Stable signal recovery from incomplete and inaccurate measurements[J]. Communications on Pure and Applied Mathematics, 2006, 59(8): 1207-1223.

[2] Donoho D L. Compressed sensing[J]. IEEE Transactions on information theory, 2006, 52(4): 1289-1306.

[3] Cai T T, Zhang A. Sparse representation of a polytope and recovery of sparse signals and low-rank matrices[J]. IEEE transactions on information theory, 2013, 60(1): 122-132.

[4] Wen J, Zhang R, Yu W. Signal-Dependent Performance Analysis of Orthogonal Matching Pursuit for Exact Sparse Recovery[J]. IEEE Transactions on Signal Processing, 2020, 68: 5031-5046.

[5] Wen J, Zhou Z, Liu Z, et al. Sharp sufficient conditions for stable recovery of block sparse signals by block orthogonal matching pursuit[J]. Applied and Computational Harmonic Analysis, 2019, 47(3): 948-974.

[6] Zhang F, Wang J, Wang W, et al. Low-tubal-rank plus sparse tensor recovery with prior subspace information[J]. IEEE transactions on pattern analysis and machine intelligence, 2021, 43(10): 3492-3507.

[7] Wang H, Zhang F, Wang J, et al. Generalized Nonconvex Approach for Low-Tubal-Rank Tensor Recovery[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, DOI: 10.1109/TNNLS.2021.3051650.

[8] Li S, Lin J, Liu D, et al. Iterative hard thresholding for compressed data separation[J]. Journal of Complexity, 2020, 59: 101469.

[9] Liu D, Li S, Shen Y. One-bit compressive sensing with projected subgradient method under sparsity constraints[J]. IEEE Transactions on Information Theory, 2019, 65(10): 6650-6663.

[10] Ge H, Chen W, Ng M K. New Restricted Isometry Property Analysis for $\ell_1 - \ell_2$ Minimization Methods[J]. SIAM Journal on Imaging Sciences, 2021, 14(2): 530-557.

[11] Ge H, Chen W, Ng M K. On Recovery of Sparse Signals With Prior Support Information via Weighted ℓ_p-Minimization[J]. IEEE Transactions on Information Theory, 2021, 67(11): 7579-7595.

[12] Feichtinger, Hans G., and Thomas Strohmer, eds. Gabor analysis and algorithms: Theory and applications. Springer Science & Business Media, 2012.

[13] Cai J F, Dong B, Osher S, et al. Image restoration: total variation, wavelet frames, and beyond[J]. Journal of the American Mathematical Society, 2012, 25(4): 1033-1089.

[14] Candes E, Demanet L, Donoho D, et al. Fast discrete curvelet transforms[J]. Multiscale modeling & simulation, 2006, 5(3): 861-899.

[15] Candes E J, Donoho D L. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities[J]. Communications on Pure and Applied Mathematics, 2004, 57(2): 219-266.

[16] Candes E J, Eldar Y C, Needell D, et al. Compressed sensing with coherent and redundant dictionaries[J]. Applied and Computational Harmonic Analysis, 2011, 31(1): 59-73.

[17] Elad M, Milanfar P, Rubinstein R. Analysis versus synthesis in signal priors[J]. Inverse problems, 2007, 23(3): 947.

[18] Lin J, Li S. Sparse recovery with coherent tight frames via analysis Dantzig selector and analysis LASSO[J]. Applied and Computational Harmonic Analysis, 2014, 37(1): 126-139.

[19] Tan Z, Eldar Y C, Beck A, et al. Smoothing and decomposition for analysis sparse recovery[J]. IEEE Transactions on Signal Processing, 2014, 62(7): 1762-1774.

[20] Shen Y, Han B, Braverman E. Stable recovery of analysis based approaches[J]. Applied and Computational Harmonic Analysis, 2015, 39(1): 161-172.

[21] Boyd S, Parikh N, Chu E. Distributed optimization and statistical learning via the alternating direction method of multipliers[M]. Now Publishers Inc, 2011.

[22] Ge H, Wen J, Chen W, et al. Stable sparse recovery with three unconstrained analysis based approaches. 2018, http://alpha.math.uga.edu/ mjlai/papers/20180126.pdf.

[23] Chartrand R. Exact reconstruction of sparse signals via nonconvex minimization[J]. IEEE Signal Processing Letters, 2007, 14(10): 707-710.

[24] Foucart S, Lai M J. Sparsest solutions of underdetermined linear systems via lq-minimization for $0 < q \leq 1$[J]. Applied and Computational Harmonic Analysis, 2009, 26(3): 395-407.

[25] Wan A. Uniform RIP Conditions for Recovery of Sparse Signals by ℓ_p $(0 < p \leq 1)$ Minimization[J]. IEEE Transactions on Signal Processing, 2020, 68: 5379-5394.

[26] Zhang R, Li S. Optimal RIP bounds for sparse signals recovery via lp minimization[J]. Applied and Computational Harmonic Analysis, 2019, 47(3): 566-584.

[27] Lin J, Li S. Restricted q-Isometry Properties Adapted to Frames for Nonconvex l_q-Analysis[J]. IEEE Transactions on Information Theory, 2016, 62(8): 4733-4747.

[28] Foucart S. Stability and robustness of ℓ_1-minimizations with Weibull matrices and redundant dictionaries[J]. Linear Algebra and its Applications, 2014, 441: 4-21.

[29] Gao Y, Han X, Ma M. Recovery of low-rank matrices based on the rank null space properties[J]. International Journal of Wavelets, Multiresolution and Information Processing, 2017, 15(04): 1750032.