Truncated sparse approximation property and truncated q-norm minimization

CHEN Wen-gu[1]   LI Peng[2*]

Abstract. This paper considers approximately sparse signal and low-rank matrix's recovery via truncated norm minimization $\min_x \|x_T\|_q$ and $\min_X \|X_T\|_{S_q}$ from noisy measurements. We first introduce truncated sparse approximation property, a more general robust null space property, and establish the stable recovery of signals and matrices under the truncated sparse approximation property. We also explore the relationship between the restricted isometry property and truncated sparse approximation property. And we also prove that if a measurement matrix A or linear map $\mathcal{A}$ satisfies truncated sparse approximation property of order k, then the first inequality in restricted isometry property of order k and of order $2k$ can hold for certain different constants δ_k and δ_{2k}, respectively. Last, we show that if $\delta_{s(k+|T^c|)} < \sqrt{(s-1)/s}$ for some $s \geq 4/3$, then measurement matrix A and linear map $\mathcal{A}$ satisfy truncated sparse approximation property of order k. It should be pointed out that when $T^c = \emptyset$, our conclusion implies that sparse approximation property of order k is weaker than restricted isometry property of order sk.

§1   Introduction

The problem of sparse signal recovery naturally arises in genetics, communications and image processing. Prominent examples include DNA microarrays [23, 34], wireless communications [27, 39], magnetic resonance imaging [32, 40], and more. In such contexts, we often require to recover an unknown signal $x \in \mathbb{R}^n$ from an underdetermined system of linear equations

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

where $b \in \mathbb{R}^m$ are available measurements, the matrix $A \in \mathbb{R}^{m \times n}$ $(m < n)$ models the linear measurement process and $z \in \mathbb{R}^m$ is a vector of measurement errors. The constrained l_q

minimization method estimates the signal x by

$$\hat{x} = \arg \min_{x \in \mathbb{R}^n} \{ \|x\|_q^q : \text{ subject to } b - Ax \in \mathcal{B} \} \tag{1.2}$$

where $0 < q \leq 1$ and $\mathcal{B}$ is a set determined by the noise structure. In this paper, we consider two types of bounded noises [6]. One is l_p bounded noises [21], i.e.,

$$\mathcal{B}^{l_p}(\eta) = \{ z : \|z\|_p \leq \eta \} \tag{1.3}$$

for some constant η; and the other is motivated by *Dantzig Selector* procedure [14], where

$$\mathcal{B}^{DS}(\eta) = \{ z : \|A^* z\|_\infty \leq \eta \}. \tag{1.4}$$

In particular, when $\mathcal{B} = \{0\}$, it is the noiseless case.

For the l_q-minimization problem (1.2), there are many works under the restricted isometry property [7–9, 12, 13, 16, 17] and under the null space property [17, 20, 24, 26, 37]. Let x_k be the best k-sparse approximation vector of x and $\sigma_k(x)_q$ be the error of best k-term approximation to the vector x in l_q-norm. Specially, we should point out that in 2011, Sun [37] introduced the sparse approximation property

$$\|x_k\|_r^q \leq \beta k^{q/r-1} \sigma_k(x)_q^q + D\|Ax\|_p^q$$

where $0 < q, p, r \leq \infty$, which was called sparse Riesz property in [38]. And Sun [37] showed that the sparse approximation property is an appropriate condition on a measurement matrix to consider stable recovery of any compressible signal from its noisy measurements. We should point out that when $q = 1$ and $p = 2$, the sparse approximation property is equivalent to the l_r-robust null space property [24, 26]

$$\|x_K\|_r \leq \beta k^{1/r-1} \sigma_k(x)_1 + D\|Ax\|_2,$$

where K is any subset of $\subset \{1, \dots, n\}$ with $|K| \leq k$. Here and what follows, x_K is the vector equal to x on K and to zero on K^c; K^c denotes the complementary set of K.

Although Basis Pursuit ((1.2) for $q = 1$) is much easy to solve, it requires significantly more measurements. In [41], Wang and Yin proposed an iterative support detection (ISD) method that runs as fast as the best BP algorithms but requires significantly fewer measurements. From an incorrect reconstruction, support detection identifies an index set I containing some elements of $\text{supp}(\bar{x}) = \{ j : x_j \neq 0 \}$, and signal reconstruction solves

$$\min_x \|x_T\|_1 \text{ subject to } Ax = b, \tag{1.5}$$

where $T = I^c$. In fact, $I^{(s+1)} = \text{supp}(x^{(s)})$ in the s-th iteration of ISD algorithm. In [41], Wang and Yin also generalized the null space property to the truncated null space property

$$\|v_K\|_1 \leq \beta \|v_{T \cap K^c}\|_1$$

where $T \subset \{1, \dots, n\}$ is any set with $|T| = t$, $K \subset T$ with $|K| \leq k$ and $v \in \mathcal{N}(A)$, and introduced an efficient implementation of ISD for recovering signals with fast decaying distributions of nonzeros from compressive sensing measurements. Later, in her Dissertation, Zhang [43] considered the model (1.5) in generalized casetruncated l_q-minimization with noise

$$\min_x \|x_T\|_q^q \text{ subject to } \|Ax - b\|_2 \leq \varepsilon, \tag{1.6}$$

where $0 < q \leq 1$. Via restricted isometry property (RIP), she got a stable recovery of x by (1.6).

By the fact that the sparse approximation property is an additional strengthening of null space property in noisy case and is weaker than restricted isometry property, therefore we consider the following question.

Question 1.1. *Whether can we introduce the truncated sparse approximation property for matrix A and get a stable recovery of any compressible signal from its noisy measurements via the model*

$$\min_x \ \|x_T\|_q^q \ \ subject \ \ to \ \ b - Ax \in \mathcal{B}, \tag{1.7}$$

where $0 < q \leq 1$.

In this paper, we give affirmative answer to this question. We introduce the truncated sparse approximation property with l_p-norm constraint $\|b - Ax\|_p$ and Dantzig selector constraint $\|A^*(b - Ax)\|_\infty$, and get a stable recovery of signal x via (1.7).

On the other hand, a closely related problem to compressed sensing is the low-rank matrix recovery, which aims to recover an unknown low-rank matrix based on its affine transformation

$$b = \mathcal{A}(X) + z, \tag{1.8}$$

where $X \in \mathbb{R}^{m \times n}$ is the decision variable and the linear map $\mathcal{A} : \mathbb{R}^{m \times n} \to \mathbb{R}^l$, $z \in \mathbb{R}^l$ is the measurement error and vector $b \in \mathbb{R}^l$ is given. Low-rank matrices arise in an incredibly wide range of settings throughout science and applied mathematics. To name just a few examples, we commonly encounter low-rank matrices in contexts as varied as: ensembles of signals [1, 19], system identification [31], adjacency matrices [30], distance matrices [4, 5, 36] and machine learning [2, 22, 33].

Assume that the singular value decomposition of X is $X = U\Lambda V^T = U\mathrm{diag}(\lambda(X))V^T$ and $l = \min\{m, n\}$, where $\lambda(X) = (\lambda_1, \ldots, \lambda_l)$ is the vector of singular values of the matrix X with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l \geq 0$. Let $\|X\|_{S_p} := \|\lambda(X)\|_p = \left(\sum_{j=1}^l |\lambda_j(X)|^p \right)^{1/p}$ denote the Schatten p-norm of matrix X, We should point out that $\|X\|_{S_1} = \|X\|_*$ the nuclear norm, $\|X\|_{S_2} = \|X\|_F$ the Frobenius norm and $\|X\|_{S_\infty} = \|X\|_{2 \to 2}$ the operator norm. In 2010, Recht, Fazel and Parrilo [35] generalized the restricted isometry property (RIP) from vectors to matrices and showed that if certain restricted isometry property holds for the linear transformation $\mathcal{A}$, the minimum-rank solution can be recovered by solving the minimization of the nuclear norm

$$\min_X \|X\|_* \ \ subject \ to \ \mathcal{A}(X) = b.$$

Later, the low-rank matrix recovery problem has been studied by many scholars, and readers can refer to [7–10, 18, 45, 46] under matrix restricted isometry property and [26, 28] under rank null space property.

Now, a corresponding question is as follows.

Question 1.2. *Whether can we introduce a truncated rank null space property, even a truncated rank sparse approximation property for linear map $\mathcal{A}$ and obtain a stable recovery of the approximately low-rank matrix X from its noisy measurements via the model*

$$\min_X \ \|X_T\|_{S_q}^q \ \ subject \ \ to \ \ b - \mathcal{A}(X) \in \mathcal{B}, \tag{1.9}$$

where $\|X_T\|_{S_q} = \left(\sum_{j \in T} |\lambda_j(X)|^q \right)^{1/q} = \left(\sum_{j=l-t+1}^l |\lambda_j(X)|^q \right)^{1/q}$ and $0 < q \leq 1$.

Our answer to this question is also affirmative.

Our paper is organized as follows. In Section 2, we introduce truncated sparse approximation property for vector and matrix, respectively, and establish a stable recovery of approximately sparse signals and low-rank matrices. In Section 3, we show that if a measurement matrix A satisfies truncated sparse approximation property of order k, then the first inequality in restricted isometry property of order k and of order $2k$ hold for certain different constants δ_k and δ_{2k}, respectively. And, we also show that if measurement matrix A satisfies restricted isometry property with $\delta_{s(k+|T^c|)} < \sqrt{(s-1)/s}$, then A also satisfies truncated sparse approximation property of order k. We note that when $T^c = \emptyset$, this conclusion implies that the $(l_2, l_1) - l_2$ sparse approximation property of order k and and $(l_2, l_1) -$ Dantzig selector sparse approximation property of order k is weaker than restricted 2-isometry property of order sk. And these relationships between the restricted isometry property and truncated sparse approximation property also hold for linear map $\mathcal{A}$. Finally, we summarize our conclusions and discuss the drawbacks of our results and file out some proper directions for coming study in Section 4.

§2 Truncated sparse approximation property and Stable Recovery

In this section, we will introduce t-truncated $(l_r, l_q) - l_p$ sparse approximation property and t-truncated $(l_r, l_q) -$ Dantzig selector sparse approximation property of order k for measurement matrix A, and get stable recovery of signal x from (1.7) for $\mathcal{B} = \mathcal{B}^{l_p}(\eta)$ and $\mathcal{B} = \mathcal{B}^{DS}(\eta)$. And we also introduce truncated rank sparse approximation property and establish the same stable recovery for approximately low-rank matrices. Firstly, we consider the vector case.

2.1 Vector Case

We introduce the following t-truncated sparse approximation property.

Definition 2.1. *Let $0 < p, r \leq \infty$ and $0 < q < \infty$. An $m \times n$ matrix A satisfies t-truncated $(l_r, l_q) - l_p$ sparse approximation property of order k with constants D and β if*

$$\|x_K\|_r^q \leq D\|Ax\|_p^q + \beta k^{q/r-1}\sigma_k(x_T)_q^q, \tag{2.1}$$

holds for all sets $T \subset \{1, \ldots, n\}$ with $|T| = t$, and all sets $K \subset T$-the index set of the k largest (in magnitude) coefficients of x_T.

And an $m \times n$ matrix A satisfies the t-truncated $(l_r, l_q) -$ Dantzig selector sparse approximation property of order k with constants D and β if

$$\|x_K\|_r^q \leq D\|A^*Ax\|_\infty^q + \beta k^{q/r-1}\sigma_k(x_T)_q^q \tag{2.2}$$

holds for all sets $T \subset \{1, \ldots, n\}$ with $|T| = t$, and all sets $K \subset T$-the index set of the k largest (in magnitude) coefficients of x_T.

Next, we give out a proposition and a remark about truncated sparse approximation property.

Proposition 2.2. *Let $0 < p, r \leq \infty$ and $0 < q < \infty$ and matrix A satisfies t-truncated $(l_r, l_q) - l_p$ sparse approximation property of order k with $t = |T|$, constants D and β.*

(1) For any $\tilde{T} \supseteq T$ with $|\tilde{T}| = \tilde{t} \geq t$, matrix A also satisfies $\tilde{t}$-truncated $(l_r, l_q) - l_p$ sparse approximation property of order k with constants D and β.

(2) For any $\tilde{k} \leq k$, matrix A also satisfies t-truncated $(l_r, l_q) - l_p$ sparse approximation property of order $\tilde{k}$ with $t = |T|$, constants D and β.

And for t-truncated $(l_r, l_q) -$ Dantzig selector sparse approximation property of order k, we have same results.

Proof. Assume that matrix A satisfies t-truncated $(l_r, l_q) - l_p$ sparse approximation property of order k with $t = |T|$, constants D and β, then by the definition, we have
$$\|x_K\|_r^q \leq D\|Ax\|_p^q + \beta k^{q/r-1}\sigma_k(x_T)_q^q,$$
holds for all sets $T \subset \{1, \ldots, n\}$ with $|T| = t$, and all sets $K \subset T$-the index set of the k largest (in magnitude) coefficients of x_T. For any $|\tilde{T}| = \tilde{t} \geq t$ with $\tilde{T} \subseteq T$, one can get
$$\|x_K\|_r^q \leq D\|Ax\|_p^q + \beta k^{q/r-1}\sigma_k(x_T)_q^q \leq D\|Ax\|_p^q + \beta k^{q/r-1}\sigma_k(x_{\tilde{T}})_q^q,$$
which implies that matrix A also satisfies $\tilde{t}$-truncated $(l_r, l_q) - l_p$ sparse approximation property of order k with constants D and β.

And for any $\tilde{k} \leq k$, let $\tilde{K} \subset T$-the index set of the $\tilde{k}$ largest (in magnitude) coefficients of x_T, then we have
$$\|x_{\tilde{K}}\|_r^q \leq \|x_K\|_r^q \leq D\|Ax\|_P^q + \beta k^{q/r-1}\sigma_k(x_T)_q^q = D\|Ax\|_p^q + \beta\Big(\frac{\sigma_k(x_T)_q}{k^{1/q-1/r}}\Big)^q$$
$$\leq D\|Ax\|_p^q + \beta\Big(\frac{\sigma_{\tilde{k}}(x_T)_q}{\tilde{k}^{1/q-1/r}}\Big)^q = D\|Ax\|_p^q + \beta\tilde{k}^{q/r-1}\sigma_{\tilde{k}}(x_T)_q^q.$$

And for for t-truncated $(l_r, l_q) -$ Dantzig selector sparse approximation property of order k, the proof is similar and we omit it here. $\square$

Remark 2.3. *If $x \in \mathcal{N}(A)$, then (2.1) and (2.2) become*
$$\|x_K\|_r^q \leq \beta k^{q/r-1}\sigma_k(x_T)_q^q. \tag{2.3}$$
When $r = q = 1$, it is equivalent to the t-truncated null space property which was introduced in [41].

Next, we consider the model (1.1) in the setting where the observations contain noise and the signal is not exactly k-sparse.

Theorem 2.4. *Consider the signal model (1.1) with $\|z\|_p \leq \varepsilon$ and suppose $\hat{x}^{l_q}$ is the minimizer of (1.7) with $\mathcal{B} = \mathcal{B}^{l_p}(\eta)$ defined in (1.3) for some $\eta \geq \varepsilon$. Let $0 < q \leq 1$, $q \leq r \leq \infty$, $1 \leq p \leq \infty$. If measurement matrix A satisfies t-truncated $(l_r, l_q) - l_p$ sparse approximation property of order k with constants $D \in (0, \infty)$ and $\beta \in (0, 1)$, then*
$$\|\hat{x}^{l_q} - x\|_r^q \leq \bigg(\max\Big\{ \Big(\frac{|T^c|}{k}\Big)^q, 1\Big\} + 1 + \Big(\frac{|T| - k}{k}\Big)^{q/r} \bigg)\frac{D}{1 - \beta}(\varepsilon + \eta)^q$$
$$+ \bigg(\max\Big\{ \Big(\frac{|T^c|}{k}\Big)^q, 1\Big\} + 1 + \Big(\frac{|T| - k}{k}\Big)^{q/r} \bigg)\frac{2\beta}{1 - \beta}k^{q/r-1}\sigma_k(x_T)_q^q, \tag{2.4}$$

and

$$\|\hat{x}^{l_q} - x\|_q^q \leq \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} \left(\frac{|T^c|}{k}\right)^{1-q/r} + 2 \right) \frac{D}{1-\beta} k^{1-q/r} (\varepsilon + \eta)^q$$

$$+ \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} \left(\frac{|T^c|}{k}\right)^{1-q/r} \frac{2\beta}{1-\beta} + \frac{2(1+\beta)}{1-\beta} \right) \sigma_k(x_T)_q^q \qquad (2.5)$$

if $q < r$, and

$$\|\hat{x}^{l_q} - x\|_q^q \leq \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} + 2 \right) \frac{D}{1-\beta} (\varepsilon + \eta)^q$$

$$+ \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} \frac{2\beta}{1-\beta} + \frac{2(1+\beta)}{1-\beta} \right) \sigma_k(x_T)_q^q \qquad (2.6)$$

if $q = r$.

Consider the signal model (1.1) with $\|A^* z\|_\infty \leq \varepsilon$ and suppose $\hat{x}^{DS}$ is the minimizer of (1.7) with $\mathcal{B} = \mathcal{B}^{DS}(\eta)$ defined in (1.4) for some $\eta \geq \varepsilon$. Let $0 < q \leq 1$ and $q \leq r \leq \infty$. If measurement matrix A satisfies t-truncated $(l_r, l_q) -$ Dantzig selector sparse approximation property of order k with constants $D \in (0, \infty)$ and $\beta \in (0, 1)$, then we have

$$\|\hat{x}^{DS} - x\|_r^q \leq \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} + 1 + \left(\frac{|T| - k}{k}\right)^{q/r} \right) \frac{D}{1-\beta} (\varepsilon + \eta)^q$$

$$+ \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} + 1 + \left(\frac{|T| - k}{k}\right)^{q/r} \right) \frac{2\beta}{1-\beta} k^{q/r-1} \sigma_k(x_T)_q^q,$$

and

$$\|\hat{x}^{DS} - x\|_q^q \leq \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} \left(\frac{|T^c|}{k}\right)^{1-q/r} + 2 \right) \frac{D}{1-\beta} k^{1-q/r} (\varepsilon + \eta)^q$$

$$+ \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} \left(\frac{|T^c|}{k}\right)^{1-q/r} \frac{2\beta}{1-\beta} + \frac{2(1+\beta)}{1-\beta} \right) \sigma_k(x_T)_q^q$$

if $q < r$, and

$$\|\hat{x}^{DS} - x\|_q^q \leq \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} + 2 \right) \frac{D}{1-\beta} (\varepsilon + \eta)^q$$

$$+ \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} \frac{2\beta}{1-\beta} + \frac{2(1+\beta)}{1-\beta} \right) \sigma_k(x_T)_q^q$$

if $q = r$.

Before proving Theorem 2.4, we first state an auxiliary lemma.

Lemma 2.5. *Suppose $x, \hat{x} \in \mathbb{R}^n$, $v = \hat{x} - x$. If $\|\hat{x}_T\|_q^q \leq \|x_T\|_q^q$, then we have*
$$\|v_{T \cap K^c}\|_q^q \leq 2\sigma_k(x_T)_q^q + \|v_{T \cap K}\|_q^q,$$
where K is the index set of k largest (in magnitude) coefficients of x_T.

Proof. Since $\|\hat{x}_T\|_q^q \leq \|x_T\|_q^q$, one can get
$$\|x_T\|_q^q \geq \|\hat{x}_T\|_q^q = \|(v + x)_T\|_q^q = \|(v + x)_{T \cap K}\|_q^q + \|(v + x)_{T \cap K^c}\|_q^q$$
$$\geq \left(\|x_{T \cap K}\|_q^q - \|v_{T \cap K}\|_q^q \right) + \left(\|v_{T \cap K^c}\|_q^q - \|x_{T \cap K^c}\|_q^q \right),$$
therefore,
$$\|v_{T \cap K^c}\|_q^q \leq \left(\|x_T\|_q^q - \|x_{T \cap K}\|_q^q + \|x_{T \cap K^c}\|_q^q \right) + \|v_{T \cap K}\|_q^q$$
$$\leq 2\sigma_k(x_T)_q^q + \|v_{T \cap K}\|_q^q.$$

$\square$

Proof of Theorem 2.4. We finish our proof via the following seven steps.

Step 1. Tube constraint and cone constraint

Let $v = \hat{x}^{l_q} - x$, we have

$$\|Av\|_p \leq \|A\hat{x} - b\|_p + \|Ax - b\|_p \leq \eta + \varepsilon. \tag{2.7}$$

Let K denote the index set of k largest (in magnitude) coefficients of x_T. By Lemma 2.5, we have

$$\|v_{T \cap K^c}\|_q^q \leq 2\sigma_k(x_T)_q^q + \|v_{T \cap K}\|_q^q. \tag{2.8}$$

Step 2. Partition the set $\{1, \ldots, n\}$.

Let $\{1, \ldots, n\} = T^c \cup T = T^c \cup (T \cap K) \cup (T \cap K^c)$, for any $q \leq \tilde{r} \leq r$, we have

$$\|v\|_{\tilde{r}}^q = \left(\|v_{T^c}\|_{\tilde{r}}^{\tilde{r}} + \|v_{T \cap K}\|_{\tilde{r}}^{\tilde{r}} + \|v_{T \cap K^c}\|_{\tilde{r}}^{\tilde{r}}\right)^{q/\tilde{r}} \leq \|v_{T^c}\|_{\tilde{r}}^q + \|v_{T \cap K}\|_{\tilde{r}}^q + \|v_{T \cap K^c}\|_{\tilde{r}}^q. \tag{2.9}$$

Next, we estimate v_{T^c}, $v_{T \cap K}$ and $v_{T \cap K^c}$, respectively.

Step 3. Estimate $v_{T \cap K}$.

By the t-truncated $(l_r, l_q) - l_p$ sparse approximation property, we have

$$\|v_{T \cap K}\|_r^q \leq D\|Av\|_p^q + \beta k^{q/r-1}\sigma_k(v_T)_q^q \leq D\|Av\|_p^q + \beta k^{q/r-1}\|v_{T \cap K^c}\|_q^q$$
$$\leq D(\varepsilon + \eta)^q + \beta k^{q/r-1}\left(\|v_{T \cap K}\|_q^q + 2\sigma_k(x_T)_q^q\right)$$
$$\leq D(\varepsilon + \eta)^q + \beta k^{q/r-1}\left(|T \cap K|^{1/q-1/r}\|v_{T \cap K}\|_r\right)^q + 2\beta k^{q/r-1}\sigma_k(x_T)_q^q$$
$$\leq D(\varepsilon + \eta)^q + \beta\|v_{T \cap K}\|_r^q + 2\beta k^{q/r-1}\sigma_k(x_T)_q^q,$$

where the third inequality follows by (2.8). Therefore, we obtain

$$\|v_{T \cap K}\|_r^q \leq \frac{D}{1-\beta}(\varepsilon + \eta)^q + \frac{2\beta}{1-\beta}k^{q/r-1}\sigma_k(x_T)_q^q. \tag{2.10}$$

by $0 < \beta < 1$.

Step 4. Estimate $v_{T \cap K^c}$.

Combining (2.8) with (2.10) yields

$$\|v_{T \cap K^c}\|_q^q \leq \|v_{T \cap K}\|_q^q + 2\sigma_k(x_T)_q^q \leq \left(|T \cap K|^{1/q-1/r}\|v_{T \cap K}\|_r\right)^q + 2\sigma_k(x_T)_q^q$$
$$\leq k^{1-q/r}\left(\frac{D}{1-\beta}(\varepsilon + \eta)^q + \frac{2\beta}{1-\beta}k^{q/r-1}\sigma_k(x_T)_q^q\right) + 2\sigma_k(x_T)_q^q$$
$$= \frac{D}{1-\beta}k^{1-q/r}(\varepsilon + \eta)^q + \frac{2}{1-\beta}\sigma_k(x_T)_q^q. \tag{2.11}$$

And by (2.10), we have

$$\|v_{T \cap K^c}\|_r^q \leq \left(|T \cap K^c| \max_{j \in T \cap K^c} |v_j|^r\right)^{q/r}$$
$$\leq \left(|T \cap K^c| \frac{\sum_{j \in T \cap K} |v_j|^r}{|T \cap K|}\right)^{q/r} = \left(\frac{|T| - k}{k}\right)^{q/r}\|v_{T \cap K}\|_r^q$$
$$\leq \left(\frac{|T| - k}{k}\right)^{q/r}\left(\frac{D}{1-\beta}(\varepsilon + \eta)^q + \frac{2\beta}{1-\beta}k^{q/r-1}\sigma_k(x_T)_q^q\right). \tag{2.12}$$

Step 5. Estimate v_{T^c}.

Our idea of proving this part comes from [41, Theorem 3.3]. Let $K_1 = T^c$, $K_2 = T \cap K$ and $K_3 = T \cap K^c$. Next, we deal with "$|K_1| \leq k$" and "$|K_1| > k$", respectively.

Case 1: $|K_1| \leq k$.

We can find $K' \subset K_2$ such that $|K_1 \cup K'| = k$ and $|K'| = k - (n - t)$. Let $\tilde{T} = K^c \cup K'$, then $|\tilde{T}| = |K^c| + |K'| = (n - k) + (k - (n - t)) = t$, $(K_1 \cup K') \subset \tilde{T}$ and $(K_1 \cup K')^c \cap \tilde{T} = T \cap K^c = K_3$, therefore from the t-truncated $(l_r, l_q) - l_p$ sparse approximation property, we have

$$\|v_{T^c}\|_r^q \leq \|v_{K_1 \cup K'}\|_r^q \leq D\|Av\|_p^q + \beta k^{q/r-1} \sigma_k(v_{\tilde{T}})_q^q$$
$$\leq D\|Av\|_p^q + \beta k^{q/r-1}\|v_{K_3}\|_q^q$$
$$\leq D(\varepsilon + \eta)^q + \beta k^{q/r-1}\|v_{T \cap K^c}\|_q^q.$$

Case 2: $|K_1| > k$.

Let $K'' \subset K_1$ denote the set of indices corresponding to the largest k entries of v_{K_1}. Take $\tilde{\tilde{T}} = K_3 \cup K''$, then $|\tilde{\tilde{T}}| = |K_3| + |K''| = t - k + k = t$, $K'' \subset \tilde{\tilde{T}}$ and $K'' \cap \tilde{\tilde{T}} = K_3$. Then t-truncated $(l_r, l_q) - l_p$ sparse approximation property leads us to conclude that

$$\|v_{T^c}\|_r^q = \|v_{K_1}\|_r^q \leq \left(\frac{|K_1|}{k}\|v_{K''}\|_r\right)^q$$
$$\leq \left(\frac{|K_1|}{k}\right)^q \left(D\|Av\|_p^q + \beta k^{q/r-1} \sigma_k(v_{\tilde{\tilde{T}}})_q^q\right)$$
$$\leq \left(\frac{|K_1|}{k}\right)^q \left(D(\varepsilon + \eta)^q + \beta k^{q/r-1}\|v_{T \cap K^c}\|_q^q\right).$$

Combining **Case 1** and **Case 2**, and using (2.11), we have

$$\|v_{T^c}\|_r^q \leq \max\left\{\left(\frac{|K_1|}{k}\right)^q, 1\right\}\left(D(\varepsilon + \eta)^q + \beta k^{q/r-1}\|v_{T \cap K^c}\|_q^q\right)$$
$$\leq \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(D(\varepsilon + \eta)^q + \beta k^{q/r-1}\left(\frac{D}{1-\beta}k^{1-q/r}(\varepsilon + \eta)^q + \frac{2}{1-\beta}\sigma_k(x_T)_q^q\right)\right)$$
$$\leq \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(\frac{D}{1-\beta}(\varepsilon + \eta)^q + \frac{2\beta}{1-\beta}k^{q/r-1}\sigma_k(x_T)_q^q\right). \tag{2.13}$$

Step 6. Obtain conclusions for $\mathcal{B} = \mathcal{B}^{l_p}(\eta)$.

For $\tilde{r} = r$, an application (2.13) (2.10) and (2.12) to (2.9) yields

$$\|v\|_r^q \leq \|v_{T^c}\|_r^q + \|v_{T \cap K}\|_r^q + \|v_{T \cap K^c}\|_r^q$$
$$\leq \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(\frac{D}{1-\beta}(\varepsilon + \eta)^q + \frac{2\beta}{1-\beta}k^{q/r-1}\sigma_k(x_T)_q^q\right)$$
$$+ \left(\frac{D}{1-\beta}(\varepsilon + \eta)^q + \frac{2\beta}{1-\beta}k^{q/r-1}\sigma_k(x_T)_q^q\right)$$
$$+ \left(\frac{|T| - k}{k}\right)^{q/r}\left(\frac{D}{1-\beta}(\varepsilon + \eta)^q + \frac{2\beta}{1-\beta}k^{q/r-1}\sigma_k(x_T)_q^q\right)$$
$$\leq \left(\max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\} + 1 + \left(\frac{|T| - k}{k}\right)^{q/r}\right)\frac{D}{1-\beta}(\varepsilon + \eta)^q$$
$$+ \left(\max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\} + 1 + \left(\frac{|T| - k}{k}\right)^{q/r}\right)\frac{2\beta}{1-\beta}k^{q/r-1}\sigma_k(x_T)_q^q,$$

which obtain the inequality (2.4).

For $\tilde{r} = q < r$, by (2.13), (2.10) and (2.11), we have

$$\|v\|_q^q \le \|v_{T^c}\|_q^q + \|v_{T \cap K}\|_q^q + \|v_{T \cap K^c}\|_q^q$$

$$\le \left(|T^c|^{1/q-1/r}\|v_{T^c}\|_r\right)^q + \left(k^{1/q-1/r}\|v_{T \cap K}\|_r\right)^q + \|v_{T \cap K^c}\|_q^q$$

$$\le |T^c|^{1-q/r} \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(\frac{D}{1-\beta}(\varepsilon + \eta)^q + \frac{2\beta}{1-\beta}k^{q/r-1}\sigma_k(x_T)_q^q\right)$$

$$+ k^{1-q/r}\left(\frac{D}{1-\beta}(\varepsilon + \eta)^q + \frac{2\beta}{1-\beta}k^{q/r-1}\sigma_k(x_T)_q^q\right)$$

$$+ \left(\frac{D}{1-\beta}k^{1-q/r}(\varepsilon + \eta)^q + \frac{2}{1-\beta}\sigma_k(x_T)_q^q\right)$$

$$= \left(\max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(\frac{|T^c|}{k}\right)^{1-q/r} + 2\right)\frac{D}{1-\beta}k^{1-q/r}(\varepsilon + \eta)^q$$

$$+ \left(\max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(\frac{|T^c|}{k}\right)^{1-q/r}\frac{2\beta}{1-\beta} + \frac{2(1+\beta)}{1-\beta}\right)\sigma_k(x_T)_q^q,$$

which finish the inequality (2.5).

Last, we turn our attention to prove inequality (2.6). For $r = q$, applying (2.13) (2.10) and (2.11) to (2.9), one can get

$$\|v\|_q^q \le \|v_{T^c}\|_q^q + \|v_{T \cap K}\|_q^q + \|v_{T \cap K^c}\|_q^q$$

$$\le \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(\frac{D}{1-\beta}(\varepsilon + \eta)^q + \frac{2\beta}{1-\beta}\sigma_k(x_T)_q^q\right)$$

$$+ \left(\frac{D}{1-\beta}(\varepsilon + \eta)^q + \frac{2\beta}{1-\beta}\sigma_k(x_T)_q^q\right) + \left(\frac{D}{1-\beta}(\varepsilon + \eta)^q + \frac{2}{1-\beta}\sigma_k(x_T)_q^q\right)$$

$$\le \left(\max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\} + 2\right)\frac{D}{1-\beta}(\varepsilon + \eta)^q$$

$$+ \left(\max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\frac{2\beta}{1-\beta} + \frac{2(1+\beta)}{1-\beta}\right)\sigma_k(x_T)_q^q.$$

Step 7. Inequalities for Dantzig selector.

In the case of noise type $\mathcal{B}^{DS}$, define $v = \hat{x}^{DS} - x$. we can obtain

$$\|A^*Av\|_\infty \le \|A*(A\hat{x} - b)\|_\infty + \|A^*(Ax - b)\|_\infty \le \eta + \varepsilon \tag{2.14}$$

and (2.2) instead of (2.7) and (2.1), respectively. The following steps are similar with the case of noise type $\mathcal{B}^{l_p}$, we omit it here. □

For $r = q$, we have a sufficient and necessary condition for guaranteeing truncated sparse approximation property, which is similar to the conclusion of classic robust l_q-null space property, see [26, Chapter 4.3]. We omit the proof details.

Theorem 2.6. *(1) The $m \times n$ measurement matrix A satisfies t-truncated $(l_q, l_q) - l_p$ sparse approximation property of order k with $t = |T|$, constants $D \in (0, \infty)$ and $\beta \in (0, 1)$ if and only if*

$$\|(y - x)_T\|_q^q \le \frac{1+\beta}{1-\beta}\left(\|y_T\|_q^q - \|x_T\|_q^q + 2\sigma_k(x_T)_q^q\right) + \frac{2D}{1-\beta}\|A(y-x)\|_p^q \tag{2.15}$$

holds for all vectors $y, x \in \mathbb{C}^n$.

(2) The $m \times n$ measurement matrix A satisfies t-truncated $(l_q, l_q) - $ Dantzig selector sparse

approximation property of order k with $t = |T|$, constants $D \in (0, \infty)$ and $\beta \in (0,1)$ if and only if

$$\|(y-x)_T\|_q^q \leq \frac{1+\beta}{1-\beta}\left(\|y_T\|_q^q - \|x_T\|_q^q + 2\sigma_k(x_T)_q^q\right) + \frac{2D}{1-\beta}\|A^*A(y-x)\|_\infty^q \qquad (2.16)$$

holds for all vectors $y, x \in \mathbb{C}^n$.

Last, we will show that if there exists a stable recovery for compressible signal x, then the measurement matrix A also posses the truncated sparse approximation property for $r = p$.

Theorem 2.7. *Let $0 < q, p \leq \infty$. If the error between the given vector x and the solution $\hat{x}$ of (1.7) satisfies*

$$\|\hat{x} - x\|_p^q \leq B_1(\varepsilon + \eta)^q + B_2 k^{q/p-1}\sigma_k(x_T)_q^q, \qquad (2.17)$$

where B_1, B_2 are positive constants independent of ε, η and x, and $|T| = t$, then

$$\|x\|_p^q \leq B_1\|Ax\|_p^q + B_2 k^{q/p-1}\sigma_k(x_T)_q^q,$$

or

$$\|x\|_p^q \leq B_1\|A^*Ax\|_\infty^q + B_2 k^{q/p-1}\sigma_k(x_T)_q^q,$$

and hence A satisfies the t-truncated $(l_p, l_q) - l_p$ sparse approximation property of order k with constants B_1 and B_2, or the t-truncated $(l_p, l_q) - l_p$ sparse approximation property of order k with constants B_1 and B_2.

Proof. We observe that zero vector is the solution of truncated l_q-minimization problem (1.7) with $\eta = 0$ and $\varepsilon = \|Ax\|_p$ or $\varepsilon = \|A^*Ax\|_\infty$, and $b = Ax$ for any $x \in \mathbb{R}^n$. Therefore the truncated sparse approximation property follows from (2.17) immediately. $\square$

Remark 2.8. *In [41], Wang and Yin proposed an iterative support detection (ISD) method to compute truncated ℓ_1 minimization (1.5). And in her Dissertation, Zhang [43] proposed a ℓ_q-iterative support detection method to compute truncated ℓ_q minimization (1.7) with $\mathcal{B} = \mathcal{B}^{l_2}(\eta)$. Now another important question is that wether there exists a algorithm to compute truncated ℓ_q minimization (1.7) with $\mathcal{B} = \mathcal{B}^{DS}(\eta)$ or $\mathcal{B} = \mathcal{B}^{l_q}(\eta)$ for $0 < q \leq 1$? This is one direction of our future research.*

2.2 Matrix Cases

Now, we consider the matrix case. First, we introduce t-truncated rank sparse approximation property as follows.

Definition 2.9. *Let $0 < p, r \leq \infty$ and $0 < q < \infty$. A linear map $\mathcal{A}$ satisfies t-truncated $(l_r, l_q) - l_p$ rank sparse approximation property of order k with constants D and β if*

$$\|X_K\|_{S_r}^q \leq D\|\mathcal{A}(X)\|_p^q + \beta k^{q/r-1}\sigma_k(X_T)_q^q, \qquad (2.18)$$

holds for all sets $T \subset \{1, \ldots, \min\{m, n\}\}$ with $|T| = t$, and all sets $K \subset T$-the index set of the k largest (in magnitude) coefficients of $\lambda(X)_T$.

And a linear map $\mathcal{A}$ satisfies the t-truncated $(l_r, l_q) - $ Dantzig selector rank sparse approximation property of order k with constants D and β if

$$\|X_K\|_{S_r}^q \leq D\|\mathcal{A}^*\mathcal{A}(X)\|_{S_\infty}^q + \beta k^{q/r-1}\sigma_k(X_T)_q^q \qquad (2.19)$$

holds for all sets $T \subset \{1, \ldots, \min\{m, n\}\}$ with $|T| = t$, and all sets $K \subset T$-the index set of the k largest (in magnitude) coefficients of $\lambda(X)_T$.

Let's consider the matrix recovery model (1.8) in the setting where the observations contain noise and matrix is not exactly k-low-rank.

Theorem 2.10. *Consider the matrix model (1.8) with $\|z\|_p \leq \varepsilon$ and suppose $\hat{X}^{l_q}$ is the minimizer of (1.9) with $\mathcal{B} = \mathcal{B}^{l_p}(\eta)$ defined in (1.3) for some $\eta \geq \varepsilon$. Let $0 < q \leq 1$, $q \leq r \leq \infty$, $1 \leq p \leq \infty$. If linear mapping $\mathcal{A}$ satisfies t-truncated $(l_r, l_q) - l_p$ sparse approximation property of order k with $t = |T|$, constants $D \in (0, \infty)$ and $\beta \in (0, 1)$, then*

$$\|\hat{X}^{l_q} - X\|_{S_r}^q \leq \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} + 1 + \left(\frac{|T| - k}{k}\right)^{q/r} \right) \frac{D}{1 - \beta}(\varepsilon + \eta)^q$$
$$+ \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} + 1 + \left(\frac{|T| - k}{k}\right)^{q/r} \right) \frac{2\beta}{1 - \beta} k^{q/r-1} \sigma_k(X_T)_q^q, \quad (2.20)$$

and

$$\|\hat{X}^{l_q} - X\|_{S_q}^q \leq \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} \left(\frac{|T^c|}{k}\right)^{1-q/r} + 2 \right) \frac{D}{1 - \beta} k^{1-q/r}(\varepsilon + \eta)^q$$
$$+ \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} \left(\frac{|T^c|}{k}\right)^{1-q/r} \frac{2\beta}{1 - \beta} + \frac{2(1 + \beta)}{1 - \beta} \right) \sigma_k(X_T)_q^q \quad (2.21)$$

if $q < r$, and

$$\|\hat{X}^{l_q} - X\|_{S_q}^q \leq \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} + 2 \right) \frac{D}{1 - \beta}(\varepsilon + \eta)^q$$
$$+ \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} \frac{2\beta}{1 - \beta} + \frac{2(1 + \beta)}{1 - \beta} \right) \sigma_k(X_T)_q^q \quad (2.22)$$

if $q = r$.

Consider the matrix model (1.8) with $\|\mathcal{A}^*(z)\|_{S_\infty} \leq \varepsilon$ and suppose $\hat{x}^{DS}$ is the minimizer of (1.9) with $\overline{\mathcal{B}} = \overline{\mathcal{B}}^{DS}(\eta) := \{z \in \mathbb{R}^l : \|\mathcal{A}^*(z)\|_{S_\infty} \leq \eta\}$ for some $\eta \geq \varepsilon$. Let $0 < q \leq 1$ and $q \leq r \leq \infty$. If linear mapping $\mathcal{A}$ satisfies t-truncated $(l_r, l_q) -$ Dantzig selector sparse approximation property of order k with $t = |T|$, constants $D \in (0, \infty)$ and $\beta \in (0, 1)$, then we have

$$\|\hat{X}^{DS} - X\|_{S_r}^q \leq \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} + 1 + \left(\frac{|T| - k}{k}\right)^{q/r} \right) \frac{D}{1 - \beta}(\varepsilon + \eta)^q$$
$$+ \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} + 1 + \left(\frac{|T| - k}{k}\right)^{q/r} \right) \frac{2\beta}{1 - \beta} k^{q/r-1} \sigma_k(X_T)_q^q,$$

and

$$\|\hat{X}^{DS} - X\|_{S_q}^q \leq \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} \left(\frac{|T^c|}{k}\right)^{1-q/r} + 2 \right) \frac{D}{1 - \beta} k^{1-q/r}(\varepsilon + \eta)^q$$
$$+ \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} \left(\frac{|T^c|}{k}\right)^{1-q/r} \frac{2\beta}{1 - \beta} + \frac{2(1 + \beta)}{1 - \beta} \right) \sigma_k(X_T)_q^q$$

if $q < r$, and

$$\|\hat{X}^{DS} - X\|_{S_q}^q \leq \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} + 2 \right) \frac{D}{1 - \beta}(\varepsilon + \eta)^q$$
$$+ \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} \frac{2\beta}{1 - \beta} + \frac{2(1 + \beta)}{1 - \beta} \right) \sigma_k(X_T)_q^q$$

if $q = r$.

To prove the results of matrix recovery, we need following lemma, which $q = 1$ comes from [35, Lemma 2.3], and $0 < q < 1$ comes from [29, Lemma 2.2] and [44, Lemma 2.1].

Lemma 2.11. *Let $0 < q \leq 1$. Let $X, Y \in \mathbb{R}^{m \times n}$ be matrices with $X^T Y = O$ and $XY^T = O$, then the following holds:*

(1) $\|X + Y\|_{S_q}^q = \|X\|_{S_q}^q + \|Y\|_{S_q}^q$;

(2) $\|X + Y\|_{S_q} \geq \|X\|_{S_q} + \|Y\|_{S_q}$.

In order to get the cone constraint for matrix's Schatten norm, we also need the following lemma which was gave by Yue and So in [42] and Audenaert [3].

Lemma 2.12. *Let $X, Y \in \mathbb{R}^{m \times n}$ be given matrices. Suppose that $f : \mathbb{R}_+ \to \mathbb{R}_+$ is a concave function satisfying $f(0) = 0$. Then, for any $k \in \{1, \ldots, \min\{m, n\}\}$, we have*

$$\sum_{j=1}^{k} |f(\lambda_j(X)) - f(\lambda_j(Y))| \leq \sum_{j=1}^{k} f(\lambda_j(X - Y)).$$

Lemma 2.13. *Suppose $X, \hat{X} \in \mathbb{R}^{m \times n}$, $V = \hat{X} - X$. If $\|\hat{X}_T\|_{S_q}^q \leq \|X_T\|_{S_q}^q$, then we have*

$$\|V_{T \cap K^c}\|_{S_q}^q \leq 2\sigma_k(\lambda(X)_T)_q^q + \|V_{T \cap K}\|_{S_q}^q,$$

where K is the index set of k largest (in magnitude) coefficients of $\lambda(X)_T$

Proof. By $\|\hat{X}_T\|_{S_q}^q \leq \|X_T\|_{S_q}^q$ and Lemma 2.11, one can get

$$\|\lambda(X)_T\|_q^q = \|X_T\|_{S_q}^q \geq \|\hat{X}_T\|_{S_q}^q = \|(V + X)_T\|_{S_q}^q = \|\lambda(V + X)_{T \cap K}\|_q^q + \|\lambda(V + X)_{T \cap K^c}\|_q^q.$$

Since $x \to |x|^q$ is concave on $\mathbb{R}_+$ for any $q \in (0, 1]$, by taking $f(\cdot) = (\cdot)^q$ in above Lemma 2.12, we immediately obtain

$$\|\lambda(X)_T\|_q^q \geq \left(\|\lambda(X)_{T \cap K}\|_q^q - \|\lambda(V)_{T \cap K}\|_q^q\right) + \left(\|\lambda(V)_{T \cap K^c}\|_q^q - \|\lambda(X)_{T \cap K^c}\|_q^q\right),$$

Therefore,

$$\|V_{T \cap K^c}\|_{S_q}^q = \|\lambda(V)_{T \cap K^c}\|_q^q \leq \left(\|\lambda(X)_T\|_q^q - \|\lambda(X)_{T \cap K}\|_q^q + \|\lambda(X)_{T \cap K^c}\|_q^q\right) + \|\lambda(V)_{T \cap K}\|_q^q$$

$$\leq 2\sigma_k(\lambda(X)_T)_q^q + \|\lambda(V)_{T \cap K}\|_q^q = 2\sigma_k(\lambda(X)_T)_q^q + \|V_{T \cap K}\|_{S_q}^q.$$

$\square$

The proof of these results of matrix case are similar to the vector case and we omit them here.

Conversely, if there exists a stable recovery of approximately low-rank matrix X, then the linear map $\mathcal{A}$ also posses the truncated sparse approximation property of $r = p$.

Theorem 2.14. *Let $0 < q, p \leq \infty$. If the error between the given matrix X and the solution $\hat{X}$ of (1.9) satisfies*

$$\|\hat{X} - X\|_{S_p}^q \leq B_1(\varepsilon + \eta)^q + B_2 k^{q/p-1} \sigma_k(X_T)_q^q, \tag{2.23}$$

where B_1, B_2 are positive constants independent of ε, η and x, and $|T| = t$, then

$$\|X\|_{S_p}^q \leq B_1 \|\mathcal{A}(X)\|_p^q + B_2 k^{q/p-1} \sigma_k(X_T)_q^q,$$

or

$$\|X\|_{S_p}^q \le B_1 \|\mathcal{A}^* \mathcal{A}(X)\|_{S_\infty}^q + B_2 k^{q/p-1} \sigma_k(X_T)_q^q,$$

and hence $\mathcal{A}$ satisfies the t-truncated $(l_p, l_q) - l_p$ rank sparse approximation property of order k with constants B_1 and B_2, or the t-truncated $(l_p, l_q)-$Dantzig-selector rank sparse approximation property of order k with constants B_1 and B_2.

§3   Truncated sparse approximation property and Restricted Isometry Property

In this section, we will consider the relationship between the restricted isometry property and truncated sparse approximation property. First, we recall the definition of restricted p-isometry property, which was introduced in [12] for $p = 2$ and in [16] for $0 < p \le 1$.

Definition 3.1. *For an matrix A, $k > 0$, and $0 < p \le 1$ or $p = 2$, the k-th restricted p-isometry constants $\delta_k = \delta_k(A)$ is the smallest numbers such that*

$$(1 - \delta_k)\|x\|_2^p \le \|Ax\|_p^p \le (1 + \delta_k)\|x\|_2^p$$

for all x such that $\|x\|_0 \le k$.

And the matrix-restricted 2-isometry property was first introduced by Recht, Fazel and Parrilo in [35]. Later, it was extended to matrix-restricted p-isometry property for $0 < p \le 1$ by Zhang, Huang and Zhang in [44].

Definition 3.2. *Let $0 < p \le 1$ or $p = 2$, $\mathcal{A} : \mathbb{R}^{m \times n} \to \mathbb{R}^l$ be a linear map. Without loss of generality, assume $m \le n$. For every integer k with $1 \le k \le \min\{m, n\}$, define the matrix k-th restricted p-isometry constant to be the smallest number $\delta_k = \delta_k(\mathcal{A})$ such that*

$$(1 - \delta_k)\|X\|_{S_2}^p \le \|\mathcal{A}(X)\|_p^p \le (1 + \delta_k)\|X\|_{S_2}^p$$

holds for all matrices X of rank at most k.

3.1   Truncated sparse approximation property implies the first inequality in Restricted Isometry Property

Firstly, we show that if a measurement matrix satisfies truncated sparse approximation property of order k, then the first inequality in restricted isometry property of order k and of order $2k$ hold for certain different constants δ_k and δ_{2k}, respectively.

Theorem 3.3. *Let $0 < q \le r \le \infty$, $0 < p \le \infty$.*

(1) If a matrix A satisfies t-truncated $(l_r, l_q) - l_p$ sparse approximation property of order k with $t = |T|$, $2k \le t$ and constants $D \in (0, \infty)$ and $\beta \in (0, 1)$, then for all $x \in \Sigma_k(T) = \{x \in \mathbb{R}^n : \|x_T\|_0 \le k\}$,

$$\frac{1}{C_1}\|x\|_r^q \le \|Ax\|_p^q, \tag{3.1}$$

where

$$C_1 = \left(\max\left\{ \left(\frac{|T^c|}{k}\right)^q, 1 \right\} + 1 \right)D,$$

and for all $x \in \Sigma_{2k}(T)$,

$$\frac{1}{C_2}\|x\|_r^q \le \|Ax\|_p^q, \tag{3.2}$$

where

$$C_2 = \left(\frac{2}{1-\beta} + \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(\left(\frac{|T \cap K^c|}{k}\right)^{1-q/r}\frac{2\beta}{1-\beta} + 1\right)\right)D.$$

(2) If a matrix A satisfies t-truncated $(l_r, l_q) -$ Dantzig-selsector sparse approximation property of order k with $t = |T|$, $2k \le t$ and constants $D \in (0, \infty)$ and $\beta \in (0, 1)$, then for all $x \in \Sigma_k(T)$,

$$\frac{1}{C_1}\|x\|_r^q \le \|A^*Ax\|_\infty^q, \tag{3.3}$$

and for all $x \in \Sigma_{2k}(T)$,

$$\frac{1}{C_2}\|x\|_r^q \le \|A^*Ax\|_\infty^q, \tag{3.4}$$

where C_1 and C_2 are the constants in above (1).

Proof. We only prove part (1). Let K denotes the index set of the k largest (in magnitude) coefficients of x_T.

(a). First, we deal with k-sparse vector. Take a k-sparse vector $x_T \in \mathbb{R}^n$, then $x_T = x_K$ and $\sigma_k(x_T)_q = \|x_{T \cap K^c}\|_q = 0$. From t-truncated $(l_r, l_q) - l_p$ sparse approximation property, one can get

$$\|x_T\|_r^q = \|x_K\|_r^q \le D\|Ax\|_p^q.$$

Similar to the **Step 5** of the proof of Theorem 2.4, we can get

$$\|x_{T^c}\|_r^q \le \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(D\|Ax\|_p^q + \beta k^{q/r-1}\|x_{T \cap K^c}\|_q^q\right)$$

$$= \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}D\|Ax\|_p^q.$$

Therefore,

$$\|x\|_r^q \le \|x_T\|_r^q + \|x_{T^c}\|_r^q \le D\|Ax\|_p^q + \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}D\|Ax\|_p^q$$

$$= \left(\max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\} + 1\right)D\|Ax\|_p^q$$

$$:= C_1\|Ax\|_p^q.$$

(b). Next, we consider the $2k$-sparse vector. Take a $2k$-sparse vector $x_T \in \mathbb{R}^n$, we can write

$$x_T = x_{K_1} + x_{K_2},$$

where $K_1, K_2 \subset T$ with $\|x_{K_j}\|_0 \le k$ for $j = 1, 2$ and $K_1 \cap K_2 = \emptyset$, then we have

$$\|x_T\|_r^q \le \|x_{K_1}\|_r^q + \|x_{K_2}\|_r^q. \tag{3.5}$$

The t-truncated $(l_r, l_q) - l_p$ sparse approximation property leads us to conclude that

$$\|x_{K_i}\|_r^q \le D\|Ax\|_p^q + \beta k^{q/r-1}\|x_{K_j}\|_q^q \le D\|Ax\|_p^q + \beta\|x_{K_j}\|_r^q, \tag{3.6}$$

where $i, j \in \{1, 2\}$ and $i \ne j$. Therefore, we can get

$$\|x_{K_1}\|_r^q + \|x_{K_2}\|_r^q \le 2D\|Ax\|_p^q + \beta(\|x_{K_1}\|_r^q + \|x_{K_2}\|_r^q),$$

which implies that

$$\|x_T\|_r^q \leq \|x_{K_1}\|_r^q + \|x_{K_2}\|_r^q \leq \frac{2D}{1-\beta}\|Ax\|_p^q. \tag{3.7}$$

On the other hand, similar to the **Step 5** of the proof of Theorem 2.4, we also can get

$$\|x_{T^c}\|_r^q \leq \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(D\|Ax\|_p^q + \beta k^{q/r-1}\|x_{T\cap K^c}\|_q^q\right),$$

and by (3.7), we get

$$\|x_{T\cap K^c}\|_q^q \leq |T \cap K^c|^{1-q/r}\|x_{T\cap K^c}\|_r^q \leq |T \cap K^c|^{1-q/r}\|x_T\|_r^q$$

$$\leq \frac{2D}{1-\beta}|T \cap K^c|^{1-q/r}\|Ax\|_p^q.$$

Therefore, we can obtain a upper bound for $\|x_{T^c}\|_r^q$ as follows

$$\|x_{T^c}\|_r^q \leq \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(\left(\frac{|T\cap K^c|}{k}\right)^{1-q/r}\frac{2\beta}{1-\beta} + 1\right)D\|Ax\|_p^q. \tag{3.8}$$

Therefore, combining (3.7) and (3.8), we can get

$$\|x\|_r^q \leq \|x_T\|_r^q + \|x_{T^c}\|_r^q$$

$$\leq \frac{2D}{1-\beta}\|Ax\|_p^q + \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(\left(\frac{|T\cap K^c|}{k}\right)^{1-q/r}\frac{2\beta}{1-\beta} + 1\right)D\|Ax\|_p^q$$

$$\leq \left(\frac{2}{1-\beta} + \max\left\{\left(\frac{|T^c|}{k}\right)^q, 1\right\}\left(\left(\frac{|T\cap K^c|}{k}\right)^{1-q/r}\frac{2\beta}{1-\beta} + 1\right)\right)D\|Ax\|_p^q$$

$$:= C_2\|Ax\|_p^q.$$

$\square$

Remark 3.4. *When $p = 2$ or $0 < p \leq 1$, and $r = 2$, we can obtain from Theorem 3.3 that if measurement matrix A has the t-truncated $(l_r, l_q) - l_p$ sparse approximation property or t-truncated $(l_r, l_q) - $Dantzig-selsector sparse approximation property of order k with $t = |T|$, constants $D \in (0, \infty)$ and $\beta \in (0, 1)$, then the first inequality in the restricted isometry property of order k and of order $2k$ hold for constants $\delta = 1 - 1/(C_1)^{p/q}$ and $\delta = 1 - 1/(C_2)^{p/q}$, respectively.*

3.2   Restricted Isometry Property implies Truncated sparse approximation property

In this subsection, we will show that the truncated sparse approximation property of order k can be deduced from the restricted 2-isometry property with $\delta_{sk} < \sqrt{(s-1)/s}$ for some $s \geq 4/3$. And especially when $T^c = \emptyset$-the classic case, it shows us that the $(l_2, l_1) - l_2$ sparse approximation property of order k and and $(l_2, l_1) - $Dantzig selector sparse approximation property of order k is weaker than restricted 2-isometry property of order sk. Our main result can be stated as follows.

Theorem 3.5. *If A has restricted 2-isometry property of order $s(k + |T^c|)$ with $\delta_{s(k+|T^c|)} <$*

$\sqrt{(s-1)/s}$ *for some* $s \geq 4/3$, *then*

$$\|x_K\|_2 \leq \frac{2\sqrt{1+\delta}}{1-\delta^2}\|Ax\|_2 + \frac{\delta}{\sqrt{(s-1)(1-\delta^2)}}\frac{\sigma_k(x_T)_1}{\sqrt{k+|T|^c}}$$

$$=: D_1\|Ax\|_2 + \beta\frac{\sigma_k(x_T)_1}{\sqrt{k+|T|^c}}$$

and

$$\|x_K\|_2 \leq \frac{2\sqrt{2(k+|T^c|)}}{1-\delta^2}\|A^*Ax\|_\infty + \frac{\delta}{\sqrt{(s-1)(1-\delta^2)}}\frac{\sigma_k(x_T)_1}{\sqrt{k+|T|^c}}$$

$$=: D_2\|A^*Ax\|_\infty + \beta\frac{\sigma_k(x_T)_1}{\sqrt{k+|T|^c}},$$

where $T \subset \{1, \ldots, n\}$ *is the any set with* $|T| = t$, $K \subset T$ *is the index set of the* k *largest (in magnitude) coefficients of* x_T.

Proof. Our proof adapts Fourcat's [25, Theorem 5] approach. Let $K_0 := K \cup T^c$ and

$$\alpha = \frac{\|x_{T \cap K^c}\|_1}{k+|T^c|},$$

where $k + |T^c| = |K_0|$. We partition the set $T \cap K^c$-complement of $K \cup T^c$ as

$$T \cap K^c = K' \cup K'',$$

where

$$K' := \left\{j \in T \cap K^c : |x_j| > \frac{\alpha}{s-1}\right\},$$

$$K'' := \left\{j \in T \cap K^c : |x_j| \leq \frac{\alpha}{s-1}\right\}.$$

Denote that $k' = \|v_{K'}\|_0$, we can derive that $k' \leq (s-1)(k+|T^c|)$ from

$$\alpha(k+|T^c|) = \|x_{T \cap K^c}\|_1 \geq \|x_{K'}\|_1 \geq k'\frac{\alpha}{s-1}.$$

By

$$\|x_{K''}\|_\infty \leq \frac{\alpha}{s-1}$$

and

$$\|x_{K''}\|_1 = \|x_{T \cap K^c}\|_1 - \|x_{K'}\|_1 \leq \alpha(k+|T^c|) - \frac{k'\alpha}{s-1} = ((s-1)(k+|T^c|) - k')\frac{\alpha}{s-1},$$

we know that the vector $x_{K''}$ belongs to a scaled version of the polytope $\left((s-1)(k+|T^c|) - k'\right)\mathcal{B}^{l_1}(\alpha/(s-1)) \cap \mathcal{B}^{l_\infty}(\alpha/(s-1))$. By sparse representation of a polytope in [9, Lemma 1.1], this polytope can be represented as the convex hull of $\left((s-1)(k+|T^c|) - k'\right)$-sparse vectors:

$$x_{K''} = \sum_{j=1}^{\tau} \rho_j u^j,$$

where u^j is $\left((s-1)(k+|T^c|) - k'\right)$-sparse and

$$\sum_{j=1}^{\tau}\rho_j = 1, \ 0 \leq \rho_j \leq 1, j = 1, \ldots, \tau$$

$$\text{supp}(u^j) \subset \text{supp}(x_{K''})$$

$$\|u^j\|_1 = \|x_{K''}\|_1, \|u^j\|_\infty \leq \frac{\alpha}{s-1}.$$

Hence,

$$\|u^j\|_2 \le \|u^j\|_\infty \sqrt{\|u^j\|_0} \le \frac{\alpha}{s-1}\sqrt{(s-1)(k+|T^c|)-k'} \le \sqrt{\frac{k+|T^c|}{s-1}}\alpha. \qquad (3.9)$$

We observe that

$$\langle A(x_{K_0}+x_{K'}), Ax\rangle \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (3.10)$$

$$= \frac{1}{4\delta}\sum_{j=1}^{\tau}\rho_j\Big(\|A\big((1+\delta)(x_{K_0}+x_{K'})+\delta u^j\big)\|_2^2 - \|A\big((1-\delta)(x_{K_0}+x_{K'})-\delta u^j\big)\|_2^2\Big).$$

$$(3.11)$$

Owing to

$$\|(1+\delta)(x_{K_0}+x_{K'})+\delta u^j\|_0 = \|(1-\delta)(x_{K_0}+x_{K'})-\delta u^j\|_0$$
$$\le (k+|T^c|)+k'+\big((s-1)(k+|T^c|)-k'\big) = s(k+|T^c|),$$

therefore we take $\delta = \delta_{s(k+|T^c|)}$.

By $\|x_{K_0}+x_{K'}\|_0 \le (k+|T^c|)+k' \le s(k+|T^c|)$, we give a upper bound estimate for the left-hand side of (3.10) as follows

$$\langle A(x_{K_0}+x_{K'}), Ax\rangle \le \|A(x_{K_0}+x_{K'})\|_2\|Ax\|_2 \le \sqrt{1+\delta}\|Ax\|_2\|x_{K_0}+x_{K'}\|_2$$

or

$$\langle A(x_{K_0}+x_{K'}), Ax\rangle \le \|x_{K_0}+x_{K'}\|_1\|A^*Ax\|_\infty \le \sqrt{s(k+|T^c|)}\|A^*Ax\|_\infty\|x_{K_0}+x_{K'}\|_2.$$

On the other hand, by restricted isometry property, the right-hand side of (3.10) is bounded from below by

$$\frac{1}{4\delta}\sum_{j=1}^{\tau}\rho_j\Big((1-\delta)\big((1+\delta)^2\|x_{K_0}+x_{K'}\|_2^2+\delta^2\|u^j\|_2^2\big) - (1+\delta)\big((1-\delta)^2\|x_{K_0}+x_{K'}\|_2^2+\delta^2\|u^j\|_2^2\big)\Big)$$

$$= \frac{1}{4\delta}\sum_{j=1}^{N}\rho_j\big(2\delta(1+\delta)(1-\delta)\|x_{K_0}+x_{K'}\|_2^2 - 2\delta^3\|u^j\|_2^2\big)$$

$$\ge \frac{1-\delta^2}{2}\|x_{K_0}+x_{K'}\|_2^2 - \frac{\delta^2}{2}\Big(\sqrt{\frac{k+|T^c|}{s-1}}\alpha\Big)^2$$

$$= \frac{1-\delta^2}{2}\|x_{K_0}+x_{K'}\|_2^2 - \frac{\delta^2}{2}\frac{\|x_{T\cap K^c}\|_1^2}{(s-1)(k+|T^c|)},$$

where the last inequality follows by (3.9).

Combining the two bounds gives

$$\|x_{K_0}+x_{K'}\|_2^2 - \frac{\delta^2}{(s-1)(1-\delta^2)}\Big(\frac{\|x_{T\cap K^c}\|_1}{\sqrt{k+|T^c|}}\Big)^2 \le \frac{2\sqrt{1+\delta}}{1-\delta^2}\|Ax\|_2\|x_{K_0}+x_{K'}\|_2$$

or

$$\|x_{K_0}+x_{K'}\|_2^2 - \frac{\delta^2}{(s-1)(1-\delta^2)}\Big(\frac{\|x_{T\cap K^c}\|_1}{\sqrt{k+|T^c|}}\Big)^2 \le \frac{2\sqrt{2(k+|T^c|)}}{1-\delta^2}\|A^*Ax\|_\infty\|x_{K_0}+x_{K'}\|_2,$$

which is equivalent to

$$\Big(\|x_{K_0}+x_{K'}\|_2 - \frac{\sqrt{1+\delta}}{1-\delta^2}\|Ax\|_2\Big)^2 \le \frac{\delta^2}{(s-1)(1-\delta^2)}\Big(\frac{\|x_{T\cap K^c}\|_1}{\sqrt{k+|T^c|}}\Big)^2 + \frac{1+\delta}{(1-\delta^2)^2}\|Ax\|_2^2$$

or

$$\left(\|x_{K_0} + x_{K'}\|_2 - \frac{\sqrt{2(k + |T^c|)}}{1 - \delta^2} \|A^*Ax\|_\infty \right)^2$$

$$\leq \frac{\delta^2}{(s - 1)(1 - \delta^2)} \left(\frac{\|x_{T \cap K^c}\|_1}{\sqrt{k + |T^c|}} \right)^2 + \frac{2(k + |T^c|)}{(1 - \delta^2)^2} \|A^*Ax\|_2^2.$$

As a result, we obtain

$$\|x_K\|_2 \leq \|x_{K_0}\|_2 \leq \|x_{K_0} + x_{K'}\|_2 \leq \frac{2\sqrt{1 + \delta}}{1 - \delta^2} \|Ax\|_2 + \frac{\delta}{\sqrt{(s - 1)(1 - \delta^2)}} \frac{\sigma_k(x_T)_1}{\sqrt{k + |T^c|}}$$

$$\leq \frac{2\sqrt{1 + \delta}}{1 - \delta^2} \|Ax\|_2 + \frac{\delta}{\sqrt{(s - 1)(1 - \delta^2)}} \frac{\sigma_k(x_T)_1}{\sqrt{k + |T^c|}}$$

$$=: D_1 \|Ax\|_2 + \beta \frac{\sigma_k(x_T)_1}{\sqrt{k + |T^c|}}$$

and

$$\|x_K\|_2 \leq \|x_{K_0}\|_2 \leq \|x_{K_0} + x_{K'}\|_2 \leq \frac{2\sqrt{2(k + |T^c|)}}{1 - \delta^2} \|A^*Ax\|_\infty + \frac{\delta}{\sqrt{(s - 1)(1 - \delta^2)}} \frac{\sigma_k(x_T)_1}{\sqrt{k + |T^c|}}$$

$$\leq \frac{2\sqrt{2(k + |T^c|)}}{1 - \delta^2} \|A^*Ax\|_\infty + \frac{\delta}{\sqrt{(s - 1)(1 - \delta^2)}} \frac{\sigma_k(x_T)_1}{\sqrt{k + |T^c|}}$$

$$=: D_2 \|A^*Ax\|_2 + \beta \frac{\sigma_k(x_T)_1}{\sqrt{k + |T^c|}},$$

which are the desired inequalities with

$$\beta = \frac{\delta}{\sqrt{(s - 1)(1 - \delta^2)}} < 1$$

when $\delta_{s(k+|T^c|)} = \delta < \sqrt{\frac{s-1}{s}}$, and with

$$D_1 = \frac{2\sqrt{1 + \delta}}{1 - \delta^2}, D_2 = \frac{2\sqrt{2(k + |T^c|)}}{1 - \delta^2}.$$

$\square$

From Theorems 3.5 and Theorem 2.4, we can obtain that under the condition $\delta_{s(k+|T^c|)} < \sqrt{\frac{s-1}{s}}$, the stable recovery holds as follows.

Theorem 3.6. *Consider the signal model (1.1) with $\|z\|_2 \leq \varepsilon$ and let $\hat{x}^{l_2}$ is the minimizer of (1.7) with $\mathcal{B} = \mathcal{B}^{l_2}(\eta)$ defined in (1.3) for some $\eta \geq \varepsilon$. If measurement matrix A has restricted isometry property of order $s(k + |T^c|)$ with $\delta_{s(k+|T^c|)} < \sqrt{(s - 1)/s}$ for some $s \geq 4/3$, then*

$$\|\hat{x}^{l_2} - x\|_2 \leq \left(\max \left\{ \left(\frac{|T^c|}{k} \right), 1 \right\} + 1 + \left(\frac{|T| - k}{k} \right)^{1/2} \right) \frac{D_1}{1 - \beta} (\varepsilon + \eta)$$

$$+ \left(\max \left\{ \left(\frac{|T^c|}{k} \right), 1 \right\} + 1 + \left(\frac{|T| - k}{k} \right)^{1/2} \right) \frac{2\beta}{1 - \beta} k^{-1/2} \sigma_k(x_T)_1,$$

where D_1 is the constant in above Theorem 3.5.

Consider the signal model (1.1) with $\|A^*z\|_\infty \leq \varepsilon$ and let $\hat{x}^{DS}$ is the minimizer of (1.7) with $\mathcal{B} = \mathcal{B}^{DS}(\eta)$ defined in (1.4) for some $\eta \geq \varepsilon$. If measurement matrix A has restricted isometry

property of order $s(k + |T^c|)$ with $\delta_{s(k+|T^c|)} < \sqrt{(s-1)/s}$ for some $s \geq 4/3$, then

$$\|\hat{x}^{l_2} - x\|_2 \leq \left(\max\left\{ \left(\frac{|T^c|}{k}\right), 1 \right\} + 1 + \left(\frac{|T| - k}{k}\right)^{1/2} \right) \frac{D_1}{1 - \beta}(\varepsilon + \eta)$$

$$+ \left(\max\left\{ \left(\frac{|T^c|}{k}\right), 1 \right\} + 1 + \left(\frac{|T| - k}{k}\right)^{1/2} \right) \frac{2\beta}{1 - \beta} k^{-1/2} \sigma_k(x_T)_1,$$

where D_2 is the constant in above Theorem 3.5.

Remark 3.7. *It should be point out that Theorem 3.5 and Theorem 3.6 also hold for $1 < s < 4/3$ with the same proof. However the bound $\sqrt{(s-1)/s}$ is not sharp for $1 < s < 4/3$. See Cai and Zhang's [9] for further discussions.*

Remark 3.8. *We point out that if we take $T^c = \emptyset$, Theorem 3.5 implies that if A has restricted 2-isometry property of order sk with $\delta_{sk} < \sqrt{(s-1)/s}$ for some $s \geq 4/3$, then A satisfies $(l_2, l_1) - l_2$ sparse approximation property of order k certain constants $0 < \beta < 1$ and $D_1 > 0$, and $(l_2, l_1) -$ Dantzig selector sparse approximation property of order k with certain constants $0 < \beta < 1$ and $D_2 > 0$, i.e l_2-robust null space property of order k [24]. This shows that the $(l_2, l_1) - l_2$ sparse approximation property of order k and and $(l_2, l_1) -$ Dantzig selector sparse approximation property of order k is weaker than restricted 2-isometry property of order tk.*

Combining this result with [37, Theorem 1.1], we get following robust estimate, which is same as [9, Theorem 2.1] but with different constants.

Theorem 3.9. *Consider the signal model (1.1) with $\|z\|_2 \leq \varepsilon$ and let $\hat{x}^{l_2}$ is the minimizer of (1.2) with $\mathcal{B} = \mathcal{B}^{l_2}(\eta)$ defined in (1.3) for some $\eta \geq \varepsilon$. If measurement matrix A has restricted 2-isometry property of order tk with $\delta_{sk} < \sqrt{(s-1)/s}$ for some $s \geq 4/3$, then*

$$\|\hat{x}^{l_2} - x\|_2 \leq \frac{2\sqrt{1 + \delta}\left(3\sqrt{(s-1)(1-\delta^2)} + \delta\right)}{(1 - \delta^2)(\sqrt{(s-1)(1-\delta^2)} - \delta)}(\varepsilon + \eta)$$

$$+ \frac{\left(\sqrt{(s-1)(1-\delta^2)} + \delta\right)^2}{\sqrt{(s-1)(1-\delta^2)}(\sqrt{(s-1)(1-\delta^2)} - \delta)} \frac{2\sigma_k(x)_1}{\sqrt{k + |T|^c}}.$$

Consider the signal model (1.1) with $\|A^*z\|_\infty \leq \varepsilon$ and let $\hat{x}^{DS}$ is the minimizer of (1.2) with $\mathcal{B} = \mathcal{B}^{DS}(\eta)$ defined in (1.4) for some $\eta \geq \varepsilon$. If measurement matrix A has restricted 2-isometry property of order tk with $\delta_{sk} < \sqrt{(s-1)/s}$ for some $s \geq 4/3$, then

$$\|\hat{x}^{DS} - x\|_2 \leq \frac{2\sqrt{2k}\left(3\sqrt{(s-1)(1-\delta^2)} + \delta\right)}{(1 - \delta^2)(\sqrt{(s-1)(1-\delta^2)} - \delta)}(\varepsilon + \eta)$$

$$+ \frac{\left(\sqrt{(s-1)(1-\delta^2)} + \delta\right)^2}{\sqrt{(s-1)(1-\delta^2)}(\sqrt{(s-1)(1-\delta^2)} - \delta)} \frac{2\sigma_k(x)_1}{\sqrt{k + |T|^c}}.$$

Remark 3.10. *It also should be point out that Zhang and Li [45] gave a complete answer to the conjecture on restricted isometry property constants δ_{sk} $(0 < s < 4/3)$ which was proposed by Cai and Zhang [9]. They showed that when $0 < s < 4/3$, the condition $\delta_{sk} < s/(4 - s)$ is sufficient to guarantee the exact recovery for all k-sparse signals in the noiseless case via the constrained ℓ_q-norm minimization for $0 < q \leq 1$ and these bounds are also sharp.*

We propose the following conjecture.

Conjecture 3.11. *If A has restricted 2-isometry property of order sk with $\delta_{sk} < (4-s)/s$ for some $s \in (0, 4/3)$, then A satisfies $(l_2, l_1) - l_2$ sparse approximation property of order k with certain constants $0 < \gamma < 1$ and $0 < D_3 < \infty$, and $(l_2, l_1) - $ Dantzig selector sparse approximation property of order k with certain constants $0 < \gamma < 1$ and $0 < D_4 < \infty$.*

Remark 3.12. *We should point out that the above Theorem 3.5, Theorem 3.6, Theorem 3.9 and Theorem 3.3 also hold for matrix case.*

§4   Conclusions and Discussion

To solve truncated norm minimization (1.7) or (1.9) with noise, we introduce truncated sparse approximation property, and obtain the stable recovery of signals and matrices under the truncated sparse approximation property (see Theorem 2.4 and Theorem 2.10). When we investigate the relationship between the truncated sparse approximation property and restricted isometry property, we find that if a measurement matrix satisfies truncated sparse approximation property of order k, then the first inequality in restricted isometry property of order k and of order $2k$ hold for certain different constants δ_k and δ_{2k}, respectively (see Theorem 3.3). And we also find that the truncated sparse approximation property of order k can be deduced from the restricted isometry property with $\delta_{sk} < \sqrt{(s-1)/s}$ for some $s \geq 4/3$ (see Theorem 3.5). And especially when $T^c = \emptyset$-the classic case, it shows us that the $(l_2, l_1) - l_2$ sparse approximation property of order k and and $(l_2, l_1) - $ Dantzig selector sparse approximation property of order k is weaker than restricted 2-isometry property of order sk. And combining this result with [37, Theorem 1.1], we get the stable recovery of signals and matrices under the restricted isometry property of order tk with $\delta_{sk} < \sqrt{(s-1)/s}$ for some $s \geq 4/3$ (see Theorem 3.9), which is the same as [9, Theorem 2.1] but with different constants. Therefore, these results in our paper may guide the practitioners to apply robust null space property in compressed sensing, low-rank matrix recovery and sparse phase retrieval, especially in truncated norm minimization problem.

However, Zhang and Li [45] showed that for $0 < s < 4/3$, the condition $\delta_{sk} < s/(4-s)$ is sufficient to guarantee the exact recovery for all k-sparse signals in the noiseless case via the constrained l_q-norm minimization and these bounds are also sharp. Can the truncated sparse approximation property of order k be deduced from the restricted isometry property with $\delta_{sk} < s/(4-s)$ for some $0 < s < 4/3$? But our method in Theorem 3.5 may not fit for this question. This is one direction of our future research (see Conjecture 3.11). And another question is whether there exists any algorithm for computing truncated ℓ_q minimization (1.7) with $\mathcal{B} = \mathcal{B}^{DS}(\eta)$ or $\mathcal{B} = \mathcal{B}^{l_q}(\eta)$ for $0 < q \leq 1$ (see Remark 2.8)? This is anther direction of our future research.

References

[1] A Ahmed, J Romberg. *Compressive multiplexing of correlated signals*, IEEE Trans Inform Theory, 2015, 61(1): 479-498.

[2] A Argyriou, T Evgeniou, M Pontil. *Convex multi-task feature learning*, Mach Learn, 2008, 73(3): 243-272.

[3] K M R Audenaert. *A generalisation of Mirskys singular value inequalities*, available at http://arxiv.org/abs/1410.4941, 2014.

[4] P Biswas, T-C Lian, T-C Wang, Y Ye. *Semidefinite programming based algorithms for sensor network localization*, ACM Transactions on Sensor Networks, 2006, 2(2): 188-220.

[5] I Borg, P Groenen. *Modern Multidimensional Scaling*, Springer Press, New York, 2010.

[6] T T Cai, L Wang. *Orthogonal matching pursuit for sparse signal recovery with noise*, IEEE Trans Inform Theory, 2011, 57(7): 4680-4688.

[7] T T Cai, A Zhang. *Compressed sensing and affine rank minimization under restricted isometry*, IEEE Trans Inform Theory, 2013, 61: 3279-3290.

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

[9] T T Cai, A Zhang. *Sparse representation of a polytope and recovery of sparse signals and low-rank matrices*, IEEE Trans Inform Theory, 2014, 60(1): 122-132.

[10] T T Cai, A Zhang. *ROP: Matrix recovery via rank-one projections*, Ann Statist, 2015, 43(1): 102-138.

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

[12] E J Candès, T Tao. *Decoding by linear programming*, IEEE Trans Inform Theory, 2005, 51(12): 4203-4215.

[13] E J Candès, T Tao. *Near optimal signal recovery from random projections: universal encoding strategies?*, IEEE Trans Inform Theory, 2006, 52(12): 5406-5425.

[14] E J Candès, T Tao. *The dantzig selector: Statistical estimation when p is much larger than n*, Ann Statist, 2007, 35(6): 2313-2351.

[15] A Chambolle, T Pock. *A first-order primal-dual algorithm for convex problems with applications to imaging*, J Math Imaging Vision, 2011, 40(1): 120-145.

[16] R Chartrand, V Staneva. *Restricted isometry properties and nonconvex compressive sensing*, Inverse Problems, 2008, 24(3): 035020-1-035020-14.

[17] A Cohen, W Dahmen, R A Devore. *Compressed sensing and best k-term approximation*, J Amer Math Soc, 2009, 22(1): 211-231.

[18] P Combettes, J-C Pesquet. *Proximal splitting methods in signal processing*, In Fixed- Point Algorithms for Inverse Problems in Science and Engineering, ed by H Bauschke, R Burachik, P Combettes, V Elser, D Luke, H Wolkowicz (Springer, New York, 2011), 185-212.

[19] M E Davies, Y C Eldar. *Rank awareness in joint sparse recovery*, IEEE Trans Inform Theory, 2012, 58(2): 1135-1146.

[20] D L Donoho, M Elad. *Optimally sparse representations in general (nonorthogonal) dictionaries via ℓ_1 minimization*, Proc Natl Acad Sci USA, 2003, 100(5): 2197-2202.

[21] D L Donoho, M Elad, V N Temlyakov. *Stable recovery of sparse overcomplete representations in the presence of noise*, IEEE Trans Inform Theory, 2006, 52(1): 6-18.

[22] S T Dumais. *Latent semantic analysis*, Ann Rev Inf Sci Tech, 2004, 38(1): 188-230.

[23] S Erickson, C Sabatti. *Empirical Bayes estimation of a sparse vector of gene expression changes*, Statist Appl Genetics Mol Biol, 2005, 4(1): 1-25.

[24] S Foucart. *Stability and robustness of ℓ_1-minimizations with Weibull matrices and redundant dictionaries*, Linear Algebra Appl., 2014, 441: 4-21.

[25] S Fourcat. *Flavors of Compressive Sensing, International Conference Approximation Theory*, Approximation Theory XV: San Antonio 2016, ed by Fasshauer, G E, Schumaker, L L, 61-104.

[26] S Foucart, H Rauhut. *A mathematical Introduction to Compressive Sensing*, Birkhäuser/Springer Press, New York, 2013.

[27] M Herman, T Strohmer. *High-resolution radar via compressed sensing*, IEEE Trans Signal Process, 2009, 57(6): 2275-2284.

[28] M Kabanava, R Kueng, H Rauhut, U Terstiege. *Stable low-rank matrix recovery via null space properties*, Information and Inference, 2016, 5(4): 405-441.

[29] L Kong, N Xiu. *Exact low-rank matrix recovery via nonconvex Schatten p-minimization*, Asia-Pac J Oper Res, 2013, 30(03): 1340010-1-1340010-13.

[30] N Linial, E London, Y Rabinovich. *The geometry of graphs and some of its algorithmic applications*, Combinatorica, 1995, 15(2): 215-245.

[31] Z Liu, L Vandenberghe. *Interior-point method for nuclear norm approximation with application to system identification*, SIAM J Matrix Anal Appl, 2009, 31(3): 1235-1256.

[32] M Lustig, D Donoho, J M Pauly. *Sparse MRI: The application of compressed sensing for rapid MR imaging*, Magnetic Resonance in Medicine, 2007, 58: 1182-1195.

[33] G Obozinski, B Taskar, M I Jordan. *Joint covariate selection and joint subspace selection for multiple classification problems*, Stat Comput, 2010, 20(2): 231-252.

[34] F Parvaresh, H Vikalo, S Misra, B Hassibi. *Recovering sparse signals using sparse measurement matrices in compressed DNA microarrays*, IEEE J Sel Top Signal Process, 2008, 2(3): 275-285.

[35] B Recht, M Fazel, P A Parrilo. *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Rev, 2010, 52(3): 471-501.

[36] A M-C So, Y Ye. *Theory of semidefinite programming for sensor network localization*, Math Program Ser B, 2007, 109: 367-384.

[37] Q Sun. *Sparse approximation property and stable recovery of sparse signals from noisy measurements*, IEEE Trans Signal Process, 2011, 59(10): 5086-5090.

[38] Q Sun, W-S Tang. *Nonlinear frames and sparse reconstructions in Banach spaces*, J Fourier Anal Appl, 2017, 23(5): 1118-1152.

[39] G Tauböck, F Hlawatsch, D Eiwen, H Rauhut. *Compressive estimation of doubly selective channels in multicarrier systems: leakage effects and sparsity-enhancing processing*, IEEE J Sel Top Sig Process, 2010, 4(2): 255-271.

[40] S Vasanawala, M Alley, B Hargreaves, R Barth, J Pauly, M Lustig. *Improved pediatric MR imaging with compressed sensing*, Radiology, 2010, 256(2): 607-616.

[41] Y Wang, W Yin. *Sparse signal reconstruction via iterative support detection*, SIAM J Imag Sci, 2010, 3(3): 462-491.

[42] M-C Yue, A M-C So. *A perturbation inequality for concave functions of singular values and its applications in low-rank matrix recovery*, Appl Comput Harmon Anal, 2016, 40(2): 396-416.

[43] X Zhang. *On compressive sensing related algorithms and application* (in Chinese), Ph.D, Zhejiang University, Hangzhou, China, 2013.

[44] M Zhang, Z-H Huang, Y Zhang. *Restricted p-isometry properties of nonconvex matrix recovery*, IEEE Trans Inform Theory, 2013, 59(7): 4316-4323.

[45] R Zhang, S Li. *A proof of conjecture on restricted isometry property constants δ_{tk} $(0 < t < \frac{4}{3})$*, IEEE Trans Inform Theory, 2018, 64(3): 1699-1705.

[46] R Zhang, S Li. *Optimal RIP bounds for sparse signals recovery via ℓ_p minimization*, Appl Comput Harmon Anal, 2017, doi:10.1016/j.acha.2017.10.004.

1 Institute of Applied Physics and Computational Mathematics, Beijing, 100088, China.
    Email: chenwg@iapcm.ac.cn
2 Graduate School, China Academy of Engineering Physics, Beijing 100088, China.
    Email: lipeng16@gscaep.ac.cn