Perturbations of Compressed Data Separation with Redundant Tight Frames

FENG ZHANG[1], JIANJUN WANG[1], YAO WANG[2], JIANWEN HUANG[1] AND WENDONG WANG[1]

[1]School of Mathematics and Statistics, Southwest University, Chongqing, 400715, China

[2]School of Mathematics and Statistics, Xi'an JiaoTong University, Xi'an, 710049, China

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

This work was supported by Natural Science Foundation of China (No. 61673015, 61273020), Fundamental Research Funds for the Central Universities(No. XDJK2015A007), Science Computing and Intelligent Information Processing of GuangXi higher education key laboratory (No. GXSCIIP201702).

.

Abstract—In the era of big data, the multi-modal data can be seen everywhere. Research on such data has attracted extensive attention in the past few years. In this paper, we investigate perturbations of compressed data separation with redundant tight frames via $\tilde{\mathbf{\Phi}}$-ℓ_q-minimization. By exploiting the properties of the redundant tight frame and the perturbation matrix, i.e., mutual coherence, null space property and restricted isometry property, the condition on reconstruction of sparse signal with redundant tight frames is established and the error estimation between the local optimal solution and the original signal is also provided. Numerical experiments are carried out to show that $\tilde{\mathbf{\Phi}}$-ℓ_q-minimization are robust and stable for the reconstruction of sparse signal with redundant tight frames. To our knowledge, our works may be the first study concerning perturbations of the measurement matrix and the redundant tight frame for compressed data separation.

Index Terms—Compressed data separation, perturbation, null space property, restricted isometry property.

1 Introduction

COMPRESSED sensing [1], [2], [3] is a novel signal processing technique for efficiently reconstructing a signal by solving underdetermined linear systems. The basic principle is that a sparse or compressible signal can be reconstructed from far fewer samples than that is required by the Shannon-Nyquist sampling theorem. Compressed sensing is being extensively applied in various fields of science and engineering, including compressive imaging [4], medical imaging [5], pattern recognition [6], image processing [7], etc.

Suppose that we observe

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{f} + \boldsymbol{z},$$

where $\boldsymbol{f} \in \mathbb{R}^n$ is an unknown signal to be reconstructed, $\boldsymbol{A}$ is an $m \times n$ measurement matrix with $m \ll n$, $\boldsymbol{y} \in \mathbb{R}^m$ are available measurements, and $\boldsymbol{z} \in \mathbb{R}^m$ is a simple additive noise with level ε ($\|\boldsymbol{z}\|_2 \leq \varepsilon$). The problem is of course ill-posed but suppose now that $\boldsymbol{f}$ is known to be sparse or nearly sparse in the sense that it depends on a smaller number of unknown parameters. However, in reality, the common signals are not necessarily sparse, and even these signals can not be sparsely represented in some orthogonal basis. Naturally, the above model can not be directly applied to the reconstruction of this kind of signals. Recently, there are some literature showing that some signals can be sparsely represented in certain redundant tight frames $\boldsymbol{D} \in \mathbb{R}^{n\times d}$ ($n \leq d, \boldsymbol{D}\boldsymbol{D}^* = \boldsymbol{D}_n$) [8], [9]. That is $\boldsymbol{f} = \boldsymbol{D}\boldsymbol{x}$, where $\boldsymbol{x} \in \mathbb{R}^d$ is (approximately) sparse. Following this, the above problem can be regarded as the $\boldsymbol{D}$-ℓ_0-minimization:

$$\min_{\bar{\boldsymbol{f}}\in\mathbb{R}^n} \|\boldsymbol{D}^*\bar{\boldsymbol{f}}\|_0 \;\; s.t. \;\; \|\boldsymbol{A}\bar{\boldsymbol{f}} - \boldsymbol{y}\|_2 \leq \varepsilon, \tag{1.1}$$

where $\boldsymbol{D}^*$ is the conjugate of the transpose of $\boldsymbol{D}$ and $\|\boldsymbol{D}^*\boldsymbol{f}\|_0$ represents the number of nonzero elements of $\boldsymbol{D}^*\boldsymbol{f}$. We call such a signal $\boldsymbol{D}^*\boldsymbol{f}$ s-sparse, if $\|\boldsymbol{D}^*\boldsymbol{f}\|_0 \leq s$. However, (1.1) is a NP problem that can not be effectively solved in practice. Relaxation methods replace ℓ_0-norm by the following convex objective function:

$$\min_{\bar{\boldsymbol{f}}\in\mathbb{R}^n} \|\boldsymbol{D}^*\bar{\boldsymbol{f}}\|_1 \;\; s.t. \;\; \|\boldsymbol{A}\bar{\boldsymbol{f}} - \boldsymbol{y}\|_2 \leq \varepsilon, \tag{1.2}$$

where $\|\boldsymbol{D}^*\bar{\boldsymbol{f}}\|_1 = \sum_{i=1}^{d} |(\boldsymbol{D}^*\bar{\boldsymbol{f}})_i|$.

Since (1.2) is a convex optimization problem, it can be transformed into an equivalent quadratic optimization problem that can be very effectively solved. However, the obtained solution by this method is not necessarily the most sparse solution. Notice that the ℓ_0-norm is the limit of the ℓ_q-norm as $q \to 0$:

$$\|\boldsymbol{f}\|_0 = \lim_{q\to 0} \|\boldsymbol{f}\|_q^q = \lim_{q\to 0} \sum_j |f_j|^q.$$

Naturally, many researchers have utilized ℓ_q-norm with $0 < q \leq 1$ to replace ℓ_1-norm, see [10], [11], [12], [13], [14]. Therefore, the following $\boldsymbol{D}$-ℓ_q-minimization problem is proposed to solve problem (1.1):

$$\min_{\bar{\boldsymbol{f}}\in\mathbb{R}^n} \|\boldsymbol{D}^*\bar{\boldsymbol{f}}\|_q^q \;\; s.t. \;\; \|\boldsymbol{A}\bar{\boldsymbol{f}} - \boldsymbol{y}\|_2 \leq \varepsilon,$$

where $\|\boldsymbol{D}^*\bar{\boldsymbol{f}}\|_q^q = \sum_{i=1}^{d} |(\boldsymbol{D}^*\bar{\boldsymbol{f}})_i|^q$.

In [12], Li and Lin have conducted a detailed analysis for $\boldsymbol{D}$-ℓ_q-minimization. The authors obtained the sufficient condition for robust and stable reconstruction of the original signal, and established an upper bound estimation of approximation error between the reconstructive signal and the true signal. Along this line, a few of scholars had paid great efforts [15], [13].

However, in the real world, we often encounter with some complex data such as: multi-frequency acoustic data (data from the superposition of different instruments) [16], neurobiology image data [17], and radar data [18] have wide application. These data show some special structures different from the traditional one, for example multiple modes, i.e., being composed of distinct subcomponents. For these data, one can try to separate it into suitable single components for convenient analysis. In [19], [20], [21], [22], typical instances consist of the texture separation from cartoon images, blind source separation and separation of sinusoids and spikes. The problem is referred as compressed data separation. In view of mathematical point, we consider splitting the signal $\boldsymbol{f} = \boldsymbol{f}_1 + \boldsymbol{f}_2$ into its constituents $\boldsymbol{f}_1 \in \mathbb{R}^n$ and $\boldsymbol{f}_2 \in \mathbb{R}^n$, which are assumed to be sparse in redundant tight frames $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$, respectively. By using linear, nonadaptive, and noisy measurements $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{f} + \boldsymbol{z}$ and $\boldsymbol{A}$, we try to reconstruct the unknown constituents $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$. In 2013, Donoho and Kutyniok [23] proposed the following $\boldsymbol{D}$-ℓ_1-separation:

$$(\hat{\boldsymbol{f}}_1, \hat{\boldsymbol{f}}_2) = \underset{\bar{\boldsymbol{f}}_1, \bar{\boldsymbol{f}}_2 \in \mathbb{R}^n}{\arg\min} \|\boldsymbol{D}_1^*\bar{\boldsymbol{f}}_1\|_1 + \|\boldsymbol{D}_2^*\bar{\boldsymbol{f}}_2\|_1 \quad s.t.\ \ \boldsymbol{f} = \bar{\boldsymbol{f}}_1 + \bar{\boldsymbol{f}}_2.$$

As we know, for the measurements $\boldsymbol{y}$, the simple additive noise $\boldsymbol{z}$ was uncorrelated with signal $\boldsymbol{f}$. However, the signal $\boldsymbol{f}$ may be polluted due to the influence of the measurement matrix and the dictionary. So, it is necessary to consider the multiplicative noise which is closely related to the signal $\boldsymbol{f}$. This kind of noise is usually generated by non-ideal measurement devices and reconstruction devices as well as the computational limitations. In order to simulate the real situation and interpret the precision errors of the measurement and reconstruction process, one should introduce the multiplicative noise into compressed data separation [24], [25]. Here, we consider the following complex case by respectively incorporating perturbations $\boldsymbol{E}$, $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ to the matrix $\boldsymbol{A}$, tight frames $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$:

$$\tilde{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{E},\ \ \tilde{\boldsymbol{D}}_1 = \boldsymbol{D}_1 + \boldsymbol{E}_1,\ \ \tilde{\boldsymbol{D}}_2 = \boldsymbol{D}_2 + \boldsymbol{E}_2,$$

where $\boldsymbol{E} \in \mathbb{R}^{m\times n}$, $\boldsymbol{E}_1 \in \mathbb{R}^{n\times d_1}$ and $\boldsymbol{E}_2 \in \mathbb{R}^{n\times d_2}$. These perturbations can be quantified with the following relative bounds:

$$\frac{\|\boldsymbol{E}\|_2}{\|\boldsymbol{A}\|_2} \le \varepsilon_{\boldsymbol{A}},\ \ \frac{\|\boldsymbol{E}_1\|_2}{\|\boldsymbol{D}_1\|_2} \le \varepsilon_{\boldsymbol{D}_1},\ \ \frac{\|\boldsymbol{E}_2\|_2}{\|\boldsymbol{D}_2\|_2} \le \varepsilon_{\boldsymbol{D}_2},$$

where $\varepsilon_{\boldsymbol{A}}$, $\varepsilon_{\boldsymbol{D}_1}$ and $\varepsilon_{\boldsymbol{D}_2}$ are perturbation levels of the measurement matrix $\boldsymbol{A}$ and the redundant tight frames $\boldsymbol{D}_1$, $\boldsymbol{D}_2$, respectively. Meanwhile, considering the merits of ℓ_q-norm $(0 < q \le 1)$ with characterizing sparsity, we adopt $\tilde{\boldsymbol{D}}$-ℓ_q-split analysis with perturbations to recover the constituents as follows:

$$(\hat{\boldsymbol{f}}_1, \hat{\boldsymbol{f}}_2) = \underset{\bar{\boldsymbol{f}}_1, \bar{\boldsymbol{f}}_2 \in \mathbb{R}^n}{\arg\min} \|\tilde{\boldsymbol{D}}_1^*\bar{\boldsymbol{f}}_1\|_q^q + \|\tilde{\boldsymbol{D}}_2^*\bar{\boldsymbol{f}}_2\|_q^q \quad s.t.\ \ \|\tilde{\boldsymbol{A}}(\bar{\boldsymbol{f}}_1 + \bar{\boldsymbol{f}}_2) - \boldsymbol{y}\|_2 \le \varepsilon, \tag{1.3}$$

where $\boldsymbol{y} = \boldsymbol{A}(\boldsymbol{f}_1 + \boldsymbol{f}_2) + \boldsymbol{z} \in \mathbb{R}^m$ and ε is a mixed noise level of measurement noise $\boldsymbol{z}$ and matrix perturbation $\boldsymbol{E}$. In general, these perturbations are more difficult to analyze than simple additive noise $\boldsymbol{z}$ since they are correlated with constituents $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ of interest. To see this, simply calculate as:

$$\tilde{\boldsymbol{A}}(\boldsymbol{f}_1 + \boldsymbol{f}_2) = \boldsymbol{A}(\boldsymbol{f}_1 + \boldsymbol{f}_2) + \boldsymbol{E}(\boldsymbol{f}_1 + \boldsymbol{f}_2),$$
$$\tilde{\boldsymbol{D}}_1^*\boldsymbol{f}_1 = \boldsymbol{D}_1^*\boldsymbol{f}_1 + \boldsymbol{E}_1^*\boldsymbol{f}_1,\ \ \tilde{\boldsymbol{D}}_2^*\boldsymbol{f}_2 = \boldsymbol{D}_2^*\boldsymbol{f}_2 + \boldsymbol{E}_2^*\boldsymbol{f}_2,$$

there will be three extra noise terms $\boldsymbol{E}(\boldsymbol{f}_1 + \boldsymbol{f}_2)$, $\boldsymbol{E}_1\boldsymbol{f}_1$ and $\boldsymbol{E}_2\boldsymbol{f}_2$. To facilitate the problem, we demand for simplifying (1.3) and initially assume the following set-up:

- $\boldsymbol{A}$ is an $m \times n$ measurement matrix.
- $\tilde{\boldsymbol{A}}$ is an $m \times n$ full rank measurement matrix (perturbation matrix of the true matrix $\boldsymbol{A}$).
- $\boldsymbol{D}_1 \in \mathbb{R}^{n\times d_1}$ and $\boldsymbol{D}_2 \in \mathbb{R}^{n\times d_2}$ are two redundant tight frames.
- $\boldsymbol{D}_1^*\boldsymbol{f}_1$ and $\boldsymbol{D}_2^*\boldsymbol{f}_2$ are approximately s_1-sparse and s_2-sparse, respectively.
- $\tilde{\boldsymbol{D}}_1 \in \mathbb{R}^{n\times d_1}$ and $\tilde{\boldsymbol{D}}_2 \in \mathbb{R}^{n\times d_2}$ are two perturbation dictionaries of $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$, respectively.
- $d = d_1 + d_2$, $\boldsymbol{D} = [\boldsymbol{D}_1|\boldsymbol{D}_2]_{n\times d}$, $\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{D}_1 & 0 \\ 0 & \boldsymbol{D}_2 \end{bmatrix}_{2n\times d}$, $\tilde{\boldsymbol{\Phi}} = \begin{bmatrix} \tilde{\boldsymbol{D}}_1 & 0 \\ 0 & \tilde{\boldsymbol{D}}_2 \end{bmatrix}_{2n\times d}$, $\boldsymbol{f} = \begin{pmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \end{pmatrix}_{2n\times 1}$, $\mathring{\boldsymbol{A}} = \tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^* \in \mathbb{R}^{m\times 2n}$.
- $\boldsymbol{\Phi}^*\boldsymbol{f}$ is approximately s-sparse, where $s = s_1 + s_2$.

Then, we can rewrite (1.3) as the following $\tilde{\boldsymbol{\Phi}}$-$\ell_q$-minimization problem:

$$\hat{\boldsymbol{f}} = \underset{\bar{\boldsymbol{f}} \in \mathbb{R}^{2n}}{\arg\min} \|\tilde{\boldsymbol{\Phi}}^*\bar{\boldsymbol{f}}\|_q^q \ \ s.t.\ \ \|\mathring{\boldsymbol{A}}\bar{\boldsymbol{f}} - \boldsymbol{y}\|_2 \le \varepsilon. \tag{1.4}$$

Taking into account the special case of ℓ_1-minimization and non-perturbation, in 2013, Lin, Li, et al. [26] have done some valuable work that investigated compressed data separation using the model

$$\hat{\boldsymbol{f}} = \underset{\bar{\boldsymbol{f}} \in \mathbb{R}^{2n}}{\arg\min} \|\boldsymbol{\Phi}^*\bar{\boldsymbol{f}}\|_1 \ \ s.t.\ \ \|\boldsymbol{A}\boldsymbol{D}\boldsymbol{\Phi}^*\bar{\boldsymbol{f}} - \boldsymbol{y}\|_2 \le \varepsilon,$$

they obtained sufficient conditions for the robust and stable reconstruction of the signal and gave an upper bound on the estimation error

$$\|\hat{\boldsymbol{f}} - \boldsymbol{f}\|_2 \le C_0\varepsilon + C_1\frac{\|\boldsymbol{\Phi}^*\boldsymbol{f} - (\boldsymbol{\Phi}^*\boldsymbol{f})_{[s]}\|_1}{\sqrt{s}},$$

where $\|\boldsymbol{\Phi}^*\boldsymbol{f} - (\boldsymbol{\Phi}^*\boldsymbol{f})_{[s]}\|_1$ is the best s-term ℓ_1 approximation error [27]. This influential result has far-reaching significance for the research of the compressed data separation. Considering the importance of the above problem, we conduct a deep investigation and provide two important results that show $\tilde{\boldsymbol{\Phi}}$-$\ell_q$-split analysis is robust and stable with regard to measurement noise and perturbation of the measurement matrix $\boldsymbol{A}$, tight frames $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$.

In short summary, our contributions are as follows:

- We first investigate the perturbations of the measurement matrix and the redundant tight frame for compressed data separation.
- We establish two sufficient conditions for the robust and stable reconstruction of the original signal.
- We obtain the estimation of upper bound on error between the reconstructive signal and the true signal.
- We perform a series of experiments to verify the reconstruction effects of $\tilde{\boldsymbol{\Phi}}$-$\ell_q$-minimization method.

The paper is organized as follows. In Section 2, we give the main result of this paper. With respect to the main theorem, we will present some meaningful remarks. In Section 3, we carry out some numerical simulation experiments on signal reconstruction. The conclusion is addressed in section 4. Finally, proofs of Theorem 2.1 and Theorem 2.2 are presented in Appendix A and Appendix B, respectively.

2 Main Result

In this section, we present our two main contributions.

2.1 Reconstruction error estimation with $\boldsymbol{\Phi}$-NSP$_q$

One of our main results is to get the upper bound of reconstruction error by using $\boldsymbol{\Phi}$-NSP$_q$ and $\tilde{\boldsymbol{\Phi}}$-$\ell_q$-split analysis with perturbations. The $\boldsymbol{\Phi}$-NSP$_q$, analogous to the null space property, is imposed on the measurement matrix and its definition as follows.

Definition 2.1 ($\boldsymbol{\Phi}$-NSP$_q$ [28])**.** *Let $\boldsymbol{\Phi} \in \mathbb{R}^{2n\times d}$ be a dictionary matrix as in the previous setting, if there exists $0 < c < 1$ such that*

$$\forall \bar{\boldsymbol{f}} \in \ker \boldsymbol{A},\ \forall |T| \le s \ \ \|\boldsymbol{\Phi}_T^* \bar{\boldsymbol{f}}\|_q^q \le c\|\boldsymbol{\Phi}_{T^c}^* \bar{\boldsymbol{f}}\|_q^q,$$

where $|T|$ is the cardinality for the index set $T \subset \{1, 2, \cdots, d\}$, T^c is its complementary index set and $\boldsymbol{\Phi}_T^* \bar{\boldsymbol{f}} = (\boldsymbol{\Phi}^* \bar{\boldsymbol{f}})_T$ is the restriction of $\boldsymbol{\Phi}^* \bar{\boldsymbol{f}}$ on T, then matrix $\boldsymbol{A}$ satisfies the ℓ_q null space property of order s relative to $\boldsymbol{\Phi}$ ($\boldsymbol{\Phi}$-NSP$_q$), and the smallest constant c is named as the null space constant (NSC).

We are now prepared to state our first main result.

Theorem 2.1. *Suppose that a tight frame $\boldsymbol{\Phi} \in \mathbb{R}^{2n\times d}$ satisfies $\boldsymbol{\Phi}\boldsymbol{\Phi}^* = \boldsymbol{I}_{2n}$ and that $\tilde{\boldsymbol{\Phi}} \in \mathbb{R}^{2n\times d}$ fulfils $\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} \le \tau_1$. Moreover, suppose that the matrix $\mathring{\boldsymbol{A}} \in \mathbb{R}^{m\times 2n}$ obeys the $\boldsymbol{\Phi}$-NSP$_q$ of order s with the null space constant c ($0 < c < 1$). If the noise measurement $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f} + \boldsymbol{z}$ satisfies $\|\boldsymbol{A}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f} - \boldsymbol{y}\|_2 \le \varepsilon$, then any solution $\hat{\boldsymbol{f}}$ of (1.4) satisfies*

$$\|\hat{\boldsymbol{f}} - \boldsymbol{f}\|_2 \le C_1\varepsilon + C_2\|\boldsymbol{\Phi}^*\boldsymbol{f} - (\boldsymbol{\Phi}^*\boldsymbol{f})_{[s]}\|_q + C_3\|\boldsymbol{A} - \tilde{\boldsymbol{A}}\|_{op} + C_4,$$

where

$$\tau_1 = 5\left(\frac{1-c}{10}\right)^{\frac{1}{q}} d^{\frac{1}{2}-\frac{1}{q}},$$

and

$$C_1 = \frac{2}{\nu_{\mathring{\boldsymbol{A}}}\left(\tau_1 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\right)},\ C_2 = \frac{2^{\frac{1}{q}} d^{\frac{1}{2}-\frac{1}{q}}}{\tau_1 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}},$$

$$C_3 = \frac{\left(1 + 2^{\frac{1}{q}-\frac{1}{2}} + 2^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\right)\|\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2}{\nu_{\mathring{\boldsymbol{A}}}\left(\tau_1 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\right)},$$

$$C_4 = \frac{2^{\frac{1}{q}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\|\boldsymbol{f}\|_2}{\tau_1 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}}.$$

Proof. See Appendix A. □

The operator norm of an $m\times n$ matrix as a mapping from $(\mathbb{R}^n, \|\cdot\|_2)$ to $(\mathbb{R}^m, \|\cdot\|_2)$, denoted by $\|\cdot\|_{op}$. The smallest positive singular value of $\boldsymbol{A}$ denoted by $\nu_{\mathring{\boldsymbol{A}}}$. This constraint $\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} \le \tau_1$ can be met by controlling the disturbance level of the frame $\boldsymbol{\Phi}$ such that $\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}$ is small enough.

Remark 2.1. *Theorem 2.1 is our highlight that we first use the $\boldsymbol{\Phi}$-NSP$_q$ to deal with the reconstruction of the compressed data separation with respect to perturbations on the measurement matrix and the dictionary. From Theorem 2.1, the condition that $\mathring{\boldsymbol{A}}$ satisfies the $\boldsymbol{\Phi}$-NSP$_q$ is only a necessary condition, however, when $\boldsymbol{D}_1$, $\boldsymbol{D}_2$ are the canonical basis, the $\boldsymbol{\Phi}$-NSP$_q$ degenerates to the standard NSP$_q$ that is a necessary and sufficient condition to robustly and stably recover any (approximately) sparse signal.*

The above statement can be summarized by the following corollary.

Corollary 2.1. *Let $\boldsymbol{D}_1$, $\boldsymbol{D}_2$ are the canonical basis. The matrix $\tilde{\boldsymbol{A}} \in \mathbb{R}^{m\times 2n}$ obeys the* NSP$_q$ *of order s with the null space constant c ($0 < c < 1$) is a necessary and sufficient condition to robustly and stably recover any (approximately) sparse signal in the case of perturbations of the measurement matrix and noise measurement. If the noise measurement $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{f} + \boldsymbol{z}$ satisfies $\|\boldsymbol{A}\boldsymbol{f} - \boldsymbol{y}\|_2 \le \varepsilon$, then any solution $\hat{\boldsymbol{f}}$ of the following optimization problem*

$$\min_{\bar{\boldsymbol{f}}\in\mathbb{R}^{2n}} \|\bar{\boldsymbol{f}}\|_q^q \ \ s.t.\ \ \|\tilde{\boldsymbol{A}}\bar{\boldsymbol{f}} - \boldsymbol{y}\|_2 \le \varepsilon$$

satisfies

$$\|\hat{\boldsymbol{f}} - \boldsymbol{f}\|_2 \le C_1^{'}\varepsilon + C_2^{'}\|\boldsymbol{f} - \boldsymbol{f}_{[s]}\|_q + C_3^{'}\|\boldsymbol{A} - \tilde{\boldsymbol{A}}\|_{op},$$

where

$$C_1^{'} = \frac{2}{5\nu_{\tilde{\boldsymbol{A}}}}\left(\frac{10}{1-c}\right)^{\frac{1}{q}} d^{\frac{1}{q}-\frac{1}{2}},\ C_2^{'} = \frac{1}{5}\left(\frac{5}{1-c}\right)^{\frac{1}{q}},$$

$$C_3^{'} = \frac{1 + 2^{1/q-1/2}}{5\nu_{\tilde{\boldsymbol{A}}}}\left(\frac{10}{1-c}\right)^{\frac{1}{q}} d^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{f}\|_2.$$

Corollary 2.1 shows that *NSP$_q$*, the minimal condition on $\tilde{\boldsymbol{A}}$ for exact recovery for any sparse signal, is also sufficient for robustness and stability via ℓ_q-minimization.

2.2 Reconstruction error estimation with $\boldsymbol{D}$-RIP

The other main result of this paper is obtained via $\tilde{\boldsymbol{\Phi}}$-$\ell_q$-split analysis with perturbations under $\boldsymbol{D}$-RIP, a natural property on measurement matrix, analogous to the restricted isometry property. The definition of $\boldsymbol{D}$-RIP is as follows.

Definition 2.2 ($\boldsymbol{D}$-RIP [29]). *Let $\boldsymbol{x} \in \mathbb{R}^d$ be approximately s-sparse. $\boldsymbol{D} \in \mathbb{R}^{n\times d}$ is a matrix as the previous setting, if there exists a constant $0 < \delta_s < 1$ for all s sparse vectors $\boldsymbol{x} \in \mathbb{R}^d$ such that*

$$(1-\delta_s)\|\boldsymbol{D}\boldsymbol{x}\|_2^2 \le \|\boldsymbol{A}\boldsymbol{D}\boldsymbol{x}\|_2^2 \le (1+\delta_s)\|\boldsymbol{D}\boldsymbol{x}\|_2^2,$$

then matrix $\boldsymbol{A}$ satisfies the restricted isometry property with respect to $\boldsymbol{D}$ ($\boldsymbol{D}$-RIP) of order s, the smallest constant δ_s is referred to as the restricted isometry constant with respect to $\boldsymbol{D}$ ($\boldsymbol{D}$-RIC).

Given a deterministic matrix $\boldsymbol{A}$, it is generally NP-hard, however, to verify whether $\boldsymbol{A}$ is a $\boldsymbol{D}$-RIP matrix. Fortunately, some random matrices have been proved to satisfy $\boldsymbol{D}$-RIP with overwhelmingly high probability, such as Gaussian random matrices, Bernoulli random matrices and partial Fourier random matrices, etc.

Next, we introduce the concept of the mutual coherence to provide a measurement of incoherence between the frames $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$, which can be used to measure the morphological difference between components.

Definition 2.3 (Mutual Coherence [26]). *Let $\boldsymbol{D}_1 = (d_{1i})_{1\le i\le d_1}$ and $\boldsymbol{D}_2 = (d_{2j})_{1\le j\le d_2}$. The mutual coherence of $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ is defined as*

$$\mu = \mu(\boldsymbol{D}_1;\boldsymbol{D}_2) = \max_{i,j} |< d_{1i}, d_{2j} >|.$$

We are now ready to state our second main result.

Theorem 2.2. *Suppose that a tight frame $\boldsymbol{\Phi} \in \mathbb{R}^{2n\times d}$ satisfies $\boldsymbol{\Phi}\boldsymbol{\Phi}^* = \boldsymbol{I}_{2n}$ and that $\tilde{\boldsymbol{\Phi}} \in \mathbb{R}^{2n\times d}$ fulfils $\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} \le \tau_2$. Fix positive integers s, k with $s \le k$. Moreover, suppose that $\tilde{\boldsymbol{A}}$ obeys the $\boldsymbol{D}$-RIP with constant $\tilde{\delta}_{s+k}$ and that the $\boldsymbol{D}$-RIP constant $\tilde{\delta}_{s+k}$ and the mutual coherence μ between $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ jointly meets*

$$\tilde{\delta}_{s+k} < W(s,\mu,k,q) := \frac{2(1-\alpha^2)^2 - \mu(s+k) - 4\alpha^2}{2(1-\alpha^2)^2 - \mu(s+k) + 4\alpha^2}.$$

If the noise measurement $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f} + \boldsymbol{z}$ satisfies $\|\boldsymbol{A}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f} - \boldsymbol{y}\|_2 \le \varepsilon$, then any solution $\hat{\boldsymbol{f}}$ of (1.4) satisfies

$$\|\hat{\boldsymbol{f}} - \boldsymbol{f}\|_2 \le C_5\varepsilon + C_6\|\boldsymbol{\Phi}^*\boldsymbol{f} - (\boldsymbol{\Phi}^*\boldsymbol{f})_{[s]}\|_q + C_7\|\boldsymbol{A} - \tilde{\boldsymbol{A}}\|_{op} + C_8,$$

where

$$\tau_2 = (\frac{d}{s})^{\frac{1}{2}-\frac{1}{q}}\sqrt{\frac{V_1}{2V_3}},\ \alpha = \frac{1}{2}(\frac{4s}{k})^{\frac{1}{q}-\frac{1}{2}},$$

and

$$C_5 = \frac{(\frac{d}{s})^{\frac{1}{2}-\frac{1}{q}}\sqrt{\frac{2V_2}{V_3}}}{\tau_2 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}},\ C_6 = \frac{(d/2)^{\frac{1}{2}-\frac{1}{q}}}{\tau_2 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}},$$

$$C_7 = \frac{(\frac{d}{s})^{\frac{1}{2}-\frac{1}{q}}\sqrt{\frac{V_2}{2V_3}} + \frac{2^{\frac{1}{q}-1}}{\nu_{\tilde{\boldsymbol{A}}}} + \frac{2^{\frac{1}{q}-1}}{\nu_{\tilde{\boldsymbol{A}}}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}}{\tau_2 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}},$$

$$C_8 = \frac{2^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\|\boldsymbol{f}\|_2}{\tau_2 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}}.$$

In addition, the constants $V_i (i = 1,2,3)$ are quantified in (B.9).

Proof. See Appendix B. □

Similarly, this constraint $\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} \le \tau_2$ also can be achieved by bounding the disturbance level of the frame $\boldsymbol{\Phi}$ such that $\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}$ is small enough. There are plenty of constants in Theorem 2.2. It is difficulty to understand the whole statement for some readers. Therefore, we provide the proper choice of parameters in the step 5 of the proof of Theorem 2.2, which makes our results clearer.

Remark 2.2. *Lin, Li, et al. have explored the compressed data separation via ℓ_1-split analysis and ℓ_q-split analysis under the $\boldsymbol{D}$-RIP in literatures [26] and [12], respectively. Our works share the ℓ_q-minimization method with [12]. From Theorem 1 of [24], there is a close correlation among the perturbation, the restricted isometry constants δ and $\tilde{\delta}$ with respect $\boldsymbol{A}$ and $\tilde{\boldsymbol{A}}$, respectively. If more information on the perturbation matrix is known, then it may be possible to estimate a smaller, and more accurate value of $\boldsymbol{D}$-RIC. In view of this, therefore, there are essential differences between our works and [12], so the perturbation should not be neglected.*

In view of the common properties of Theorem 2.1 and Theorem 2.2, we provide some remarks as follows:

Remark 2.3. *By using the frame inequality, our results can be easily extended to the general frames cases and because there exists only a difference of constants in the proofs. In Theorem 2.1 and Theorem 2.2, we assume $\boldsymbol{\Phi}$ is a tight frame ($\rho_1 = \rho_2$). This means that $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ are also tight frames. It is helpful for simplifying the analysis, but is of course not necessary because the assumption does not affect the generalization of our theorems. Since the condition of the theorem can be weakened, in this situation, our theory will be more practical significance and applied values.*

Remark 2.4. *When $\boldsymbol{D}_1 = \boldsymbol{D}_2 = \boldsymbol{I}$, $\boldsymbol{\Phi}$-NSP$_q$ and $\boldsymbol{D}$-RIP will reduce to the standard NSP$_q$ and RIP, respectively. Our results show that NSP$_q$ or RIP characterizes the exact recovery of any sparse signal $\boldsymbol{f} = \boldsymbol{f}_1 + \boldsymbol{f}_2$ from its noiseless observation $\boldsymbol{y} = \boldsymbol{A}(\boldsymbol{f}_1 + \boldsymbol{f}_2)$ via $\tilde{\boldsymbol{\Phi}}$-$\ell_q$-split analysis.*

Remark 2.5. *The above theorems offer the upper bound estimation on reconstruction error, which clearly depicts relationship among reconstruction error, the best s-term approximation, noise level and q. Particularly, it shows that the reconstruction speed is proportionally controlled by the best s-term approximation, perturbation and noise level. Obviously, with no perturbations on the measurement matrix or the redundant tight frame, $\|\hat{\boldsymbol{f}} - \boldsymbol{f}\|_2 \to 0$ as $\varepsilon \to 0$, it therefore shows that any s-sparse signal can be approximated arbitrarily well, especially, when $\varepsilon = 0$, $\boldsymbol{f}$ can be exactly reconstructed.*

3 Numerical simulations

In this section, we provide an efficient algorithm and a series of numerical simulations to evaluate the performance of our $\tilde{\boldsymbol{\Phi}}$-$\ell_q$-minimization method.

3.1 An IRLS algorithm for $\tilde{\mathbf{\Phi}}$-ℓ_q-minimization problem

In order to solve the $\tilde{\mathbf{\Phi}}$-ℓ_q-minimization problem (1.4) with $0 < q \leq 1$, we first derive an efficient algorithm which can be seen a natural extension of the iterative reweighted least squares algorithm (IRLS) [30]. Similarly, the problem (1.4) can be rewritten as the following unconstrained regularization problem:

$$\min_{\bar{\boldsymbol{f}} \in \mathbb{R}^{2n}} \|\tilde{\mathbf{\Phi}}^* \bar{\boldsymbol{f}}\|_{q,\epsilon}^q + \frac{1}{2\lambda}\|\mathring{\boldsymbol{A}}\bar{\boldsymbol{f}} - \boldsymbol{y}\|_2^2, \tag{3.1}$$

where ϵ is a smoothing parameter, λ is a regularization parameter and $\|\tilde{\mathbf{\Phi}}^* \bar{\boldsymbol{f}}\|_{q,\epsilon}^q = \sum_{i=1}^d \left(\epsilon^2 + (\tilde{\mathbf{\Phi}}_{[i]}^* \bar{\boldsymbol{f}})^2\right)^{\frac{q}{2}}$. For convenience, we let $\boldsymbol{f}_0$ denote a critical point of (3.1) and it satisfies the first-order optimality condition

$$\sum_{i=1}^d \frac{q\tilde{\mathbf{\Phi}}_{[i]}\tilde{\mathbf{\Phi}}_{[i]}^*}{\left(\epsilon^2 + (\tilde{\mathbf{\Phi}}_{[i]}^* \boldsymbol{f}_0)^2\right)^{1-\frac{q}{2}}}\boldsymbol{f}_0 + \frac{1}{\lambda}\mathring{\boldsymbol{A}}^*(\mathring{\boldsymbol{A}}\boldsymbol{f}_0 - \boldsymbol{y}) = 0. \tag{3.2}$$

Because of the nonlinearity of the above system, there is no straightforward method to solve it. However, we can use the iterative method to approximate the solution of problem (3.2), and the iterative process is as follows:

$$\left\{\sum_{i=1}^d \frac{q\lambda\tilde{\mathbf{\Phi}}_{[i]}\tilde{\mathbf{\Phi}}_{[i]}^*}{\left((\epsilon^{(t)})^2 + (\tilde{\mathbf{\Phi}}_{[i]}^* \boldsymbol{f}^{(t)})^2\right)^{1-\frac{q}{2}}} + \mathring{\boldsymbol{A}}^*\mathring{\boldsymbol{A}}\right\}\boldsymbol{f}^{(t+1)} = \mathring{\boldsymbol{A}}^*\boldsymbol{y},$$

the above method is summarized as Algorithm 1:

Algorithm 1 IRLS algorithm for $\tilde{\mathbf{\Phi}}$-ℓ_q-minimization problem

1: Initialize $\boldsymbol{f}^{(0)}$ such that $\mathring{\boldsymbol{A}}\boldsymbol{f}^{(0)} = \boldsymbol{y}$, and $\epsilon^{(0)} = 1$, $0 < q \leq 1$, λ.
2: Set $t = 0$.
3: **repeat**
4: Search $\boldsymbol{f}^{(t+1)}$ by solving

$$\boldsymbol{f}^{(t+1)} = \left\{\tilde{\mathbf{\Phi}}\text{Diag}\left[\frac{q\lambda\boldsymbol{I}}{\left((\epsilon^{(t)})^2 + (\tilde{\mathbf{\Phi}}_{[i]}^* \boldsymbol{f}^{(t)})^2\right)^{1-\frac{q}{2}}}, i = 1, 2, \cdots, d\right]\tilde{\mathbf{\Phi}}^* + \mathring{\boldsymbol{A}}^*\mathring{\boldsymbol{A}}\right\}^{-1}\mathring{\boldsymbol{A}}^*\boldsymbol{y}.$$

5: Update $\epsilon^{(t+1)} = 0.9\epsilon^{(t)}$.
6: Replace t with $t + 1$.
7: **until** Any of the following stopping criterions are satisfied.
1) $\|\boldsymbol{f}^{(t+1)} - \boldsymbol{f}^{(t)}\|_2 \leq 1 \times 10^{-5}$;
2) $t \leq 100$.
8: Output $\boldsymbol{f}^{(t+1)}$ as the approximation to $\boldsymbol{f}_0$.

3.2 Experimental settings

Throughout the experiments, the measurement matrix $\boldsymbol{A}$ is generated by creating an $m \times n$ Gaussian matrix with $m = 128$ and $n = 256$, and the tight frames $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ are generated by creating two $n \times d_1$ and $n \times d_2$ DCT dictionaries with $d_1 = d_2 = 512$, respectively. The elements of perturbation matrices $\boldsymbol{E}$, $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ are subject to normal distribution, moreover $\|\boldsymbol{E}\|_2 = \varepsilon_{\boldsymbol{A}}\|\boldsymbol{A}\|_2$, $\|\boldsymbol{E}_1\|_2 = \varepsilon_{\boldsymbol{D}_1}\|\boldsymbol{D}_1\|_2$ and $\|\boldsymbol{E}_2\|_2 = \varepsilon_{\boldsymbol{D}_2}\|\boldsymbol{D}_2\|_2$, where $\varepsilon_{\boldsymbol{A}}$, $\varepsilon_{\boldsymbol{D}_1}$ and $\varepsilon_{\boldsymbol{D}_2}$ are perturbation levels of the measurement matrix $\boldsymbol{A}$ and the redundant tight frames $\boldsymbol{D}_1$, $\boldsymbol{D}_2$, respectively. As is shown in the conditions $\|\mathbf{\Phi}^* - \tilde{\mathbf{\Phi}}^*\|_{op} < \tau_1$ and $\|\mathbf{\Phi}^* - \tilde{\mathbf{\Phi}}^*\|_{op} < \tau_2$, the dictionary $\mathbf{\Phi}$ is very sensitive to the perturbation, so we make $\|\mathbf{\Phi}^* - \tilde{\mathbf{\Phi}}^*\|_{op}$ small enough by controlling $\varepsilon_{\boldsymbol{D}_1}$ and $\varepsilon_{\boldsymbol{D}_2}$, meanwhile, we keep the perturbation matrices $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ unchanged and only consider the change of $\boldsymbol{E}$ in the experiment. We set the value of the noise vector $\boldsymbol{z}$ obeying a Gaussian distribute with mean 0 and deviation 0.05. The original signal $\boldsymbol{f}$ is synthesized by using $\boldsymbol{f} = \mathbf{\Phi}\boldsymbol{x}$ where $\boldsymbol{x} \in \mathbb{R}^d$ is a s-sparse signal with $d = 1024$ and $s = 30$. The relative error between the reconstructed signal $\hat{\boldsymbol{f}}$ and the original signal $\boldsymbol{f}$ is denoted as $\|\hat{\boldsymbol{f}} - \boldsymbol{f}\|_2/\|\boldsymbol{f}\|_2$. We perform 100 times against each test and report the average result.

3.3 Experimental results

Fig 3.1 presents the relationship between the q, the perturbation level, and the relative error of signal reconstruction. The results show that the smaller the perturbation, the better the reconstruction effect of the signal. Moreover, the reconstruction effect is the best when q is around 0.5, and the reconstruction effect is the worst when $q = 1$. An instance is also presented in Fig 3.2, which carves the recovery of the signal f and its constituents f_1, f_2 via $\tilde{\mathbf{\Phi}}$-ℓ_q-minimization method with $q = 0.5$ and $\varepsilon_{\boldsymbol{A}} = 0.01$. The results show that $\tilde{\mathbf{\Phi}}$-ℓ_q-minimization method can almost accurately reconstruct the original signal.

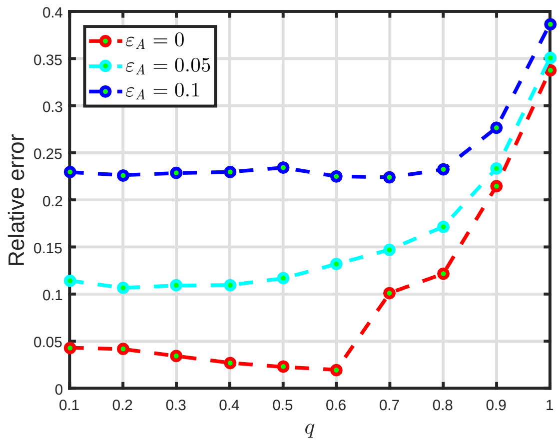
4 Conclusion

This paper mainly investigates $\tilde{\mathbf{\Phi}}$-ℓ_q-split analysis $(0 < q \leq 1)$ to recover the general signal based on the measurement matrix and the redundant tight frames with perturbations. The sufficient conditions $\mathbf{\Phi}$-NSP$_q$ and $\boldsymbol{D}$-RIP for the robust and stable reconstruction of the original signal are established, and the estimations of upper bound on error are obtained. The derived results show that the upper bound of the error is mainly controlled by q, the best s-term approximation, $\|\mathbf{\Phi}^* - \tilde{\mathbf{\Phi}}^*\|_{op}$ and $\|\boldsymbol{A} - \tilde{\boldsymbol{A}}\|_{op}$. In addition, a series of experiments are conducted to test $\tilde{\mathbf{\Phi}}$-ℓ_q-minimization method. The simulation results show that $\tilde{\mathbf{\Phi}}$-ℓ_q-minimization method has the ideal reconstruction effect. Our works are helpful in understanding and development of the compression data separation.

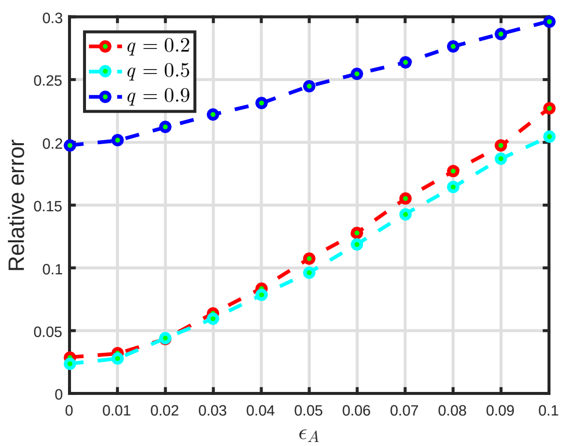
Appendix A
Proof of Theorem 2.1

In order to improve the readability of theorem proving, we initially review some inequalities used repeatedly in this paper as follows:

1) *The triangle inequality:*
$\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$, $\forall \boldsymbol{x},\ \boldsymbol{y} \in \mathbb{R}^n$.
2) *The reverse triangle inequality:*
$\|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$, $\forall \boldsymbol{x},\ \boldsymbol{y} \in \mathbb{R}^n$.



(a)



(b)

Fig. 3.1: Parameters selection for $\tilde{\mathbf{\Phi}}$-ℓ_q-minimization method. (a) for q versus relative error with different values of ε_A. (b) for ε_A versus relative error with different values of q.

3) *The frame inequality:*
$\rho_1\|\boldsymbol{f}\| \leq \|\mathbf{\Phi}^*\boldsymbol{f}\| \leq \rho_2\|\boldsymbol{f}\|,\ 0 < \rho_1 \leq \rho_2,\ \forall \boldsymbol{f} \in \mathbb{R}^d$.

4) *The quasi-norm inequality:*
$\|\boldsymbol{x}\|_p \leq n^{\frac{1}{p}-\frac{1}{q}}\|\boldsymbol{x}\|_q \leq \|\boldsymbol{x}\|_q \leq n^{\frac{1}{q}-\frac{1}{p}}\|\boldsymbol{x}\|_p$,
$0 < q \leq p \leq \infty,\ \forall \boldsymbol{x} \in \mathbb{R}^n$.
Two special cases of quasi-norm inequality:

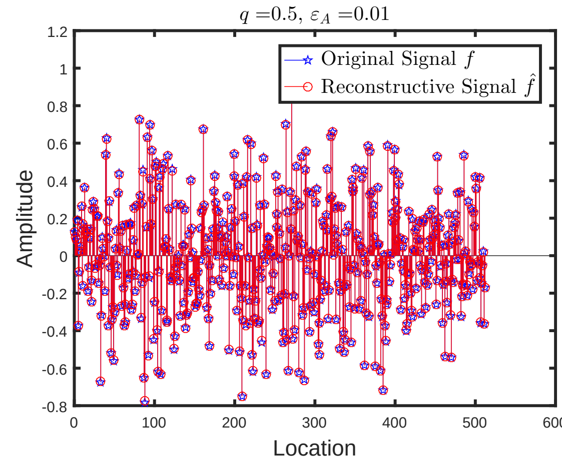
4.1) $\|\boldsymbol{x}\|_1^q \leq \|\boldsymbol{x}\|_q^q,\ 0 < q \leq 1.$
$$\Leftrightarrow \left(\sum_{i=1}^{n}|x_i|\right)^q \leq \sum_{i=1}^{n}|x_i|^q,\ 0 < q \leq 1.$$

4.2) $\|\boldsymbol{x}\|_1^t \leq n^{t-1}\|\boldsymbol{x}\|_t^t,\ t \geq 1.$
$$\Leftrightarrow \left(\sum_{i=1}^{n}|x_i|\right)^t \leq n^{t-1}\sum_{i=1}^{n}|x_i|^t,\ t \geq 1.$$

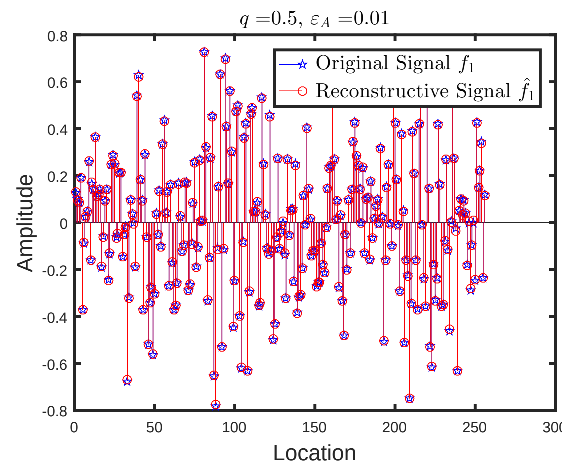
The following lemma provides a useful property deriving from the singular value decomposition.

Lemma A.1 ([10]). *Suppose $\boldsymbol{M}$ is an $m \times n$ $(m \leq n)$ matrix, then any vector $\boldsymbol{\xi} \in \mathbb{R}^n$ can be decomposed as $\boldsymbol{\xi} = \boldsymbol{\gamma} + \boldsymbol{\eta}$ with $\boldsymbol{\gamma} \in \ker \boldsymbol{M}$, $\boldsymbol{\eta} \perp \ker \boldsymbol{M}$ and $\|\boldsymbol{\eta}\| \leq \frac{1}{\nu_{\boldsymbol{M}}}\|\boldsymbol{M}\boldsymbol{\xi}\|$, where $\nu_{\boldsymbol{M}}$ is the smallest positive singular value of $\boldsymbol{M}$.*

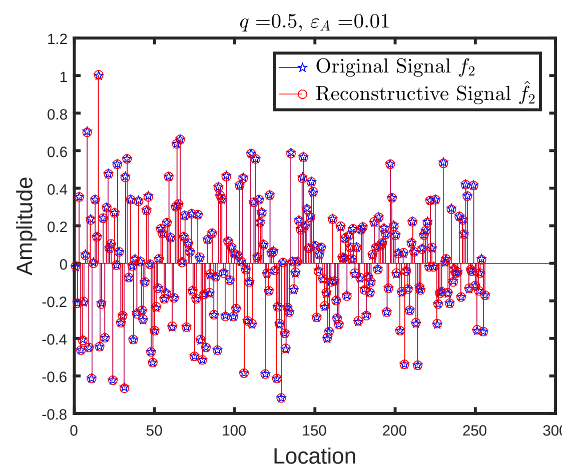
With these preparations we embark on the proof of Theorem 2.1.



(a) Signal f



(b) Signal f_1



(c) Signal f_2

Fig. 3.2: Signal reconstruction via $\tilde{\mathbf{\Phi}}$-ℓ_q-minimization method with $q = 0.5$ and $\varepsilon_{\boldsymbol{A}} = 0.01$. (a), (b) and (c) for the signal f and its constituents f_1, f_2, respectively.

Proof. Step 1: Estimation of the perturbations.

It is known that, $\|\boldsymbol{A}\boldsymbol{D}\mathbf{\Phi}^*\boldsymbol{f} - \boldsymbol{y}\|_2 \leq \varepsilon$ is valid. But $\|\tilde{\boldsymbol{A}}\boldsymbol{D}\mathbf{\Phi}^*\boldsymbol{f} - \boldsymbol{y}\|_2$ is not necessarily less than ε because $\tilde{\boldsymbol{A}}$ is a perturbation of $\boldsymbol{A}$. Moreover, because $\tilde{\boldsymbol{A}}$ is a full rank matrix, so there are some $\boldsymbol{w}$s for each $\boldsymbol{f}$ such that $\tilde{\boldsymbol{A}}\boldsymbol{D}\mathbf{\Phi}^*(\boldsymbol{w} + \boldsymbol{f}) = \boldsymbol{A}\boldsymbol{D}\mathbf{\Phi}^*\boldsymbol{f}$, that is $\tilde{\boldsymbol{A}}\boldsymbol{D}\mathbf{\Phi}^*\boldsymbol{w} = (\boldsymbol{A} - \tilde{\boldsymbol{A}})\boldsymbol{D}\mathbf{\Phi}^*\boldsymbol{f}$, which means $\|\tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^*(\boldsymbol{w}+\boldsymbol{f})-\boldsymbol{y}\|_2 \le \varepsilon$ is feasible. Moreover, among all $\boldsymbol{w}$ which satisfy this equation, there exists a unique vector of minimal ℓ_2 norm with $\boldsymbol{w}\perp\ker(\tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^*)$. Thus, by Lemma A.1, we have

$$\|\boldsymbol{w}\|_2 \le \frac{1}{\nu_{\mathring{\boldsymbol{A}}}}\|\mathring{\boldsymbol{A}}\boldsymbol{w}\|_2 = \frac{1}{\nu_{\mathring{\boldsymbol{A}}}}\|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2. \tag{A.1}$$

Since $\boldsymbol{\Phi}$ is a tight frame, using the frame inequality with $\rho_2=1$, we get $\|\boldsymbol{\Phi}^*\boldsymbol{w}\|_2 \le \|\boldsymbol{w}\|_2$, and hence

$$\begin{aligned}\|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q^q &\overset{(a)}{\le} \|\boldsymbol{\Phi}^*\boldsymbol{w}-\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q^q + \|\boldsymbol{\Phi}^*\boldsymbol{w}\|_q^q \\ &\overset{(b)}{\le} \left(d^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^*\boldsymbol{w}-\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_2\right)^q + \left(d^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^*\boldsymbol{w}\|_2\right)^q \\ &\overset{(c)}{\le} d^{1-\frac{q}{2}}\|\boldsymbol{\Phi}^*-\tilde{\boldsymbol{\Phi}}^*\|_{op}^q\|\boldsymbol{w}\|_2^q + d^{1-\frac{q}{2}}\|\boldsymbol{w}\|_2^q \\ &= d^{1-\frac{q}{2}}\|\boldsymbol{w}\|_2^q\left(\|\boldsymbol{\Phi}^*-\tilde{\boldsymbol{\Phi}}^*\|_{op}^q+1\right),\end{aligned} \tag{A.2}$$

where (a) follows from the triangle inequality, and (b) is due to the quasi-norm inequality. Notice that in (c), the operator norm of an $m\times n$ matrix as a mapping from $(\mathbb{R}^n,\|\cdot\|_2)$ to $(\mathbb{R}^m,\|\cdot\|_2)$, denoted by $\|\cdot\|_{op}$. Thus

$$\|(\boldsymbol{\Phi}^*-\tilde{\boldsymbol{\Phi}}^*)\boldsymbol{w}\|_2 \le \|\boldsymbol{\Phi}^*-\tilde{\boldsymbol{\Phi}}^*\|_{op}\|\boldsymbol{w}\|_2$$

is an immediate consequence of the definition of operator norm[1].

Taking the qth root of (A.2) and using the special case 4.2) of quasi-norm inequality, we have

$$\|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q \le (2d)^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{w}\|_2\left(\|\boldsymbol{\Phi}^*-\tilde{\boldsymbol{\Phi}}^*\|_{op}+1\right).$$

By (A.1), we have

$$\|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q \le \frac{(2d)^{\frac{1}{q}-\frac{1}{2}}}{\nu_{\mathring{\boldsymbol{A}}}}\left(\|\boldsymbol{\Phi}^*-\tilde{\boldsymbol{\Phi}}^*\|_{op}+1\right)\|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2. \tag{A.3}$$

Step 2: Consequence of the minimizer.

Since both $\hat{\boldsymbol{f}}$ and $\boldsymbol{f}+\boldsymbol{w}$ are feasible, but $\hat{\boldsymbol{f}}$ is a minimum solution of (1.4), we have

$$\begin{aligned}\|\tilde{\boldsymbol{\Phi}}^*\hat{\boldsymbol{f}}\|_q^q &\le \|\tilde{\boldsymbol{\Phi}}^*(\boldsymbol{f}+\boldsymbol{w})\|_q^q \\ &= \|\tilde{\boldsymbol{\Phi}}_T^*\boldsymbol{f}+\tilde{\boldsymbol{\Phi}}_T^*\boldsymbol{w}\|_q^q + \|\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{f}+\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{w}\|_q^q.\end{aligned} \tag{A.4}$$

Moreover, let $\boldsymbol{h}=\hat{\boldsymbol{f}}-\boldsymbol{f}$ where $\hat{\boldsymbol{f}}$ is the optimal solution of (1.4) and $\boldsymbol{f}$ is the original signal, we have

$$\begin{aligned}\|\tilde{\boldsymbol{\Phi}}^*\hat{\boldsymbol{f}}\|_q^q =&\|\tilde{\boldsymbol{\Phi}}^*(\boldsymbol{h}+\boldsymbol{f})\|_q^q \\ =&\|\tilde{\boldsymbol{\Phi}}_T^*\boldsymbol{h}+\tilde{\boldsymbol{\Phi}}_T^*\boldsymbol{f}\|_q^q + \|\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{h}+\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{f}\|_q^q \\ \ge&\|\tilde{\boldsymbol{\Phi}}_T^*\boldsymbol{f}+\tilde{\boldsymbol{\Phi}}_T^*\boldsymbol{w}\|_q^q - \|\tilde{\boldsymbol{\Phi}}_T^*\boldsymbol{h}-\tilde{\boldsymbol{\Phi}}_T^*\boldsymbol{w}\|_q^q \\ &+ \|\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{h}\|_q^q - \|\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{f}\|_q^q,\end{aligned} \tag{A.5}$$

here, the last inequality holds because of the reverse triangle inequality.

Combining (A.4) with (A.5), yields

$$\|\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{h}\|_q^q \le \|\tilde{\boldsymbol{\Phi}}_T^*\boldsymbol{h}\|_q^q + 2\|\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{f}\|_q^q + \|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q^q. \tag{A.6}$$

1. We define the operator norm of $\boldsymbol{Q}\in\mathbb{R}^{m\times n}$ as: $\|\boldsymbol{Q}\|_{op} := \sup\{\|\boldsymbol{Q}\boldsymbol{v}\|/\|\boldsymbol{v}\| : \boldsymbol{v}\in\mathbb{R}^n \text{ with } \boldsymbol{v}\ne\boldsymbol{0}\}$

Adding the term $\|\boldsymbol{\Phi}_{T^c}^*\boldsymbol{h}\|_q^q$ to both sides of (A.6), we get

$$\begin{aligned}\|\boldsymbol{\Phi}_{T^c}^*\boldsymbol{h}\|_q^q + \|\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{h}\|_q^q \le& \|\boldsymbol{\Phi}_{T^c}^*\boldsymbol{h}\|_q^q + \|\tilde{\boldsymbol{\Phi}}_T^*\boldsymbol{h}\|_q^q + 2\|\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{f}\|_q^q \\ &+\|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q^q + \|\boldsymbol{\Phi}_T^*\boldsymbol{h}\|_q^q - \|\boldsymbol{\Phi}_T^*\boldsymbol{h}\|_q^q.\end{aligned}$$

By rewriting the above inequality, we obtain

$$\begin{aligned}&\|\boldsymbol{\Phi}_{T^c}^*\boldsymbol{h}\|_q^q \\ \le&\|\boldsymbol{\Phi}_T^*\boldsymbol{h}\|_q^q + \left(\|\tilde{\boldsymbol{\Phi}}_T^*\boldsymbol{h}\|_q^q - \|\boldsymbol{\Phi}_T^*\boldsymbol{h}\|_q^q\right) + 2\|\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{f}\|_q^q \\ &+\left(\|\boldsymbol{\Phi}_{T^c}^*\boldsymbol{h}\|_q^q - \|\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{h}\|_q^q\right) + \|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q^q \\ \le&\|\boldsymbol{\Phi}_T^*\boldsymbol{h}\|_q^q + \|\boldsymbol{\Phi}_T^*\boldsymbol{h}-\tilde{\boldsymbol{\Phi}}_T^*\boldsymbol{h}\|_q^q + 2\|\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{f}\|_q^q \\ &+\|\boldsymbol{\Phi}_{T^c}^*\boldsymbol{h}-\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{h}\|_q^q + \|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q^q \\ =&\|\boldsymbol{\Phi}_T^*\boldsymbol{h}\|_q^q + \|\boldsymbol{\Phi}^*\boldsymbol{h}-\tilde{\boldsymbol{\Phi}}^*\boldsymbol{h}\|_q^q + 2\|\tilde{\boldsymbol{\Phi}}_{T^c}^*\boldsymbol{f}\|_q^q + \|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q^q,\end{aligned} \tag{A.7}$$

where the second inequality utilizes the reverse triangle inequality again.

Step 3: Consequence of $\boldsymbol{\Phi}$-NSP$_q$.

Utilizing the assumption that $\mathring{\boldsymbol{A}}$ satisfies the $\boldsymbol{\Phi}$-NSP$_q$, T is a index set with $|T|\le s$, and we decompose $\boldsymbol{h}$ as $\boldsymbol{h}=\boldsymbol{\gamma}+\boldsymbol{\eta}$ with $\boldsymbol{\gamma}\in\ker\mathring{\boldsymbol{A}}$ and $\boldsymbol{\eta}\perp\ker\mathring{\boldsymbol{A}}$, we get

$$\begin{aligned}\|\boldsymbol{\Phi}_T^*\boldsymbol{h}\|_q^q &\overset{(a)}{\le} \|\boldsymbol{\Phi}_T^*\boldsymbol{\gamma}\|_q^q + \|\boldsymbol{\Phi}_T^*\boldsymbol{\eta}\|_q^q \\ &\overset{(b)}{\le} c\|\boldsymbol{\Phi}_{T^c}^*\boldsymbol{\gamma}\|_q^q + \|\boldsymbol{\Phi}_T^*\boldsymbol{\eta}\|_q^q \\ &\le c\|\boldsymbol{\Phi}_{T^c}^*\boldsymbol{h}\|_q^q + \|\boldsymbol{\Phi}^*\boldsymbol{\eta}\|_q^q,\end{aligned} \tag{A.8}$$

where, according to the triangle inequality, (a) is definitely true; while (b) holds since by definition of $\boldsymbol{\Phi}$-NSP$_q$ with null space constant c.

Step 4: Estimation of $\|\boldsymbol{\Phi}^*\boldsymbol{\eta}\|_q$.

Since $\boldsymbol{\Phi}$ is a tight frame with $\rho_2=1$, we easily obtain

$$\|\boldsymbol{\Phi}^*\boldsymbol{\eta}\|_q \le d^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^*\boldsymbol{\eta}\|_2 \le d^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\eta}\|_2.$$

On account of $\boldsymbol{\eta}\perp\ker\mathring{\boldsymbol{A}}$, by Lemma A.1, we have

$$\|\boldsymbol{\eta}\|_2 \le \frac{1}{\nu_{\mathring{\boldsymbol{A}}}}\|\mathring{\boldsymbol{A}}\boldsymbol{h}\|_2 = \frac{1}{\nu_{\mathring{\boldsymbol{A}}}}\|\tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^*(\hat{\boldsymbol{f}}-\boldsymbol{f})\|_2.$$

Note that

$$\begin{aligned}&\|\tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^*(\hat{\boldsymbol{f}}-\boldsymbol{f})\|_2 \\ =&\|\tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^*\hat{\boldsymbol{f}}-\boldsymbol{y}+\boldsymbol{y}-\boldsymbol{A}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}+\boldsymbol{A}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}-\tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2 \\ \le&\|\tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^*\hat{\boldsymbol{f}}-\boldsymbol{y}\|_2 + \|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2 + \|\boldsymbol{A}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}-\tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2 \\ \le&2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2,\end{aligned}$$

that is because $\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2 \le \varepsilon$ follows from the assumption of Theorem 2.1; and because $\hat{\boldsymbol{f}}$ is the optimal solution of (1.4), $\hat{\boldsymbol{f}}$ satisfies the constraint condition of (1.4), that is, $\|\boldsymbol{y}-\tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^*\hat{\boldsymbol{f}}\|_2 \le \varepsilon$.

Thus, we have

$$\|\boldsymbol{\eta}\|_2 \le \frac{1}{\nu_{\mathring{\boldsymbol{A}}}}\left\{2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\right\}.$$

So, the following holds

$$\|\boldsymbol{\Phi}^*\boldsymbol{\eta}\|_q \le \frac{d^{\frac{1}{q}-\frac{1}{2}}}{\nu_{\mathring{\boldsymbol{A}}}}\left\{2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\right\}. \tag{A.9}$$

Step 5: Estimation of $\|\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}\|_q^q$.

$$\begin{aligned}
&\|\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}\|_q^q \\
\overset{(a)}{\le}&\|\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}\|_q^q + \|\boldsymbol{\Phi}^*_T\boldsymbol{f} - \tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{f}\|_q^q \\
=&\left(\|\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}\|_q^q - \|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q^q\right) + \|\boldsymbol{\Phi}^*_T\boldsymbol{f} - \tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{f}\|_q^q + \|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q^q \\
\overset{(b)}{\le}&\|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f} - \tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}\|_q^q + \|\boldsymbol{\Phi}^*_T\boldsymbol{f} - \tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{f}\|_q^q + \|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q^q \\
=&\|\boldsymbol{\Phi}^*\boldsymbol{f} - \tilde{\boldsymbol{\Phi}}^*\boldsymbol{f}\|_q^q + \|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q^q, \qquad \text{(A.10)}
\end{aligned}$$

where (a) is founded on the non-negativity of quasi-norm, that is, $\|\boldsymbol{\Phi}^*_T\boldsymbol{f} - \tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{f}\|_q^q \ge 0$, and (b) holds because of the reverse triangle inequality.

Step 6: Bounding the error.

Based on the fact that $\boldsymbol{\Phi}$ is a tight frame with $\rho_1 = 1$ and the quasi-norm inequality, we have

$$\|\boldsymbol{h}\|_2 \le \|\boldsymbol{\Phi}^*\boldsymbol{h}\|_2 \le \|\boldsymbol{\Phi}^*\boldsymbol{h}\|_q.$$

In order to get bounds on $\|\boldsymbol{h}\|_2$, we are first ready to estimate $\|\boldsymbol{\Phi}^*\boldsymbol{h}\|_q$.

By (A.7), it is easy to see that

$$\begin{aligned}
&\|\boldsymbol{\Phi}^*\boldsymbol{h}\|_q \\
=&\left(\|\boldsymbol{\Phi}^*_T\boldsymbol{h}\|_q^q + \|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{h}\|_q^q\right)^{\frac{1}{q}} \\
\le&\left(2\|\boldsymbol{\Phi}^*_T\boldsymbol{h}\|_q^q + \|\boldsymbol{\Phi}^*\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*\boldsymbol{h}\|_q^q + 2\|\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}\|_q^q + \|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q^q\right)^{\frac{1}{q}}.
\end{aligned}$$

On the other hand, associating with (A.7) and (A.8), we get

$$\begin{aligned}
\|\boldsymbol{\Phi}^*_T\boldsymbol{h}\|_q^q \le&\frac{c}{1-c}\|\boldsymbol{\Phi}^*\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*\boldsymbol{h}\|_q^q + \frac{2c}{1-c}\|\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}\|_q^q \\
&+ \frac{c}{1-c}\|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q^q + \frac{1}{1-c}\|\boldsymbol{\Phi}^*\boldsymbol{\eta}\|_q^q.
\end{aligned}$$

Hence

$$\begin{aligned}
\|\boldsymbol{h}\|_2 \le&\left(\frac{1+c}{1-c}\|\boldsymbol{\Phi}^*\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*\boldsymbol{h}\|_q^q + \frac{2+2c}{1-c}\|\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}\|_q^q\right. \\
&\left.+\frac{1+c}{1-c}\|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q^q + \frac{2}{1-c}\|\boldsymbol{\Phi}^*\boldsymbol{\eta}\|_q^q\right)^{\frac{1}{q}}.
\end{aligned}$$

Substituting (A.10) into the above inequality, we have

$$\begin{aligned}
&\|\boldsymbol{h}\|_2 \\
\le&\left(\frac{1+c}{1-c}\|\boldsymbol{\Phi}^*\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*\boldsymbol{h}\|_q^q + \frac{2+2c}{1-c}\|\boldsymbol{\Phi}^*\boldsymbol{f} - \tilde{\boldsymbol{\Phi}}^*\boldsymbol{f}\|_q^q\right. \\
&\left.+\frac{2+2c}{1-c}\|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q^q + \frac{1+c}{1-c}\|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q^q + \frac{2}{1-c}\|\boldsymbol{\Phi}^*\boldsymbol{\eta}\|_q^q\right)^{\frac{1}{q}} \\
\le&5^{\frac{1}{q}-1}\left\{\left(\frac{1+c}{1-c}\right)^{\frac{1}{q}}\left(\|\boldsymbol{\Phi}^*\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*\boldsymbol{h}\|_q + \|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q\right)\right. \\
&+\left(\frac{2+2c}{1-c}\right)^{\frac{1}{q}}\left(\|\boldsymbol{\Phi}^*\boldsymbol{f} - \tilde{\boldsymbol{\Phi}}^*\boldsymbol{f}\|_q + \|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q\right) \\
&\left.+\left(\frac{2}{1-c}\right)^{\frac{1}{q}}\|\boldsymbol{\Phi}^*\boldsymbol{\eta}\|_q\right\}.
\end{aligned}$$

In particular, since $\frac{1}{q} > 1$, so the second inequality takes advantage of the special case 4.2) of the quasi-norm inequality.

Then plugging (A.3) and (A.9) to the above inequality, we obtain

$$\begin{aligned}
&\left\{1 - \frac{1}{5}\left(\frac{10}{1-c}\right)^{\frac{1}{q}} d^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\right\}\|\boldsymbol{h}\|_2 \\
\le&\frac{2}{5\nu_{\mathring{\boldsymbol{A}}}}\left(\frac{10}{1-c}\right)^{\frac{1}{q}} d^{\frac{1}{q}-\frac{1}{2}}\varepsilon + \frac{1}{5}\left(\frac{20}{1-c}\right)^{\frac{1}{q}}\|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q \\
&+\left\{\frac{1}{5\nu_{\mathring{\boldsymbol{A}}}}\left(\frac{10}{1-c}\right)^{\frac{1}{q}} d^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\left(1 + 2^{\frac{1}{q}-\frac{1}{2}}\right.\right. \\
&\left.\left.+2^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\right)\right\}\|\boldsymbol{A} - \tilde{\boldsymbol{A}}\|_{op} \\
&+\frac{1}{5}\left(\frac{20}{1-c}\right)^{\frac{1}{q}} d^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\|\boldsymbol{f}\|_2.
\end{aligned}$$

Here, just like (A.2), we use the operator inequality for operators $(\boldsymbol{A} - \tilde{\boldsymbol{A}})$ and $(\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*)$, respectively.

Let

$$\tau_1 = 5\left(\frac{1-c}{10}\right)^{\frac{1}{q}} d^{\frac{1}{2}-\frac{1}{q}},$$

by controlling the disturbance level of the frame $\boldsymbol{\Phi}$ such that $\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} < \tau_1$, then

$$1 - \frac{1}{5}\left(\frac{10}{1-c}\right)^{\frac{1}{q}} d^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} = 1 - \frac{1}{\tau_1}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} > 0.$$

Therefore

$$\|\boldsymbol{h}\|_2 \le C_1\varepsilon + C_2\|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q + C_3\|\boldsymbol{A} - \tilde{\boldsymbol{A}}\|_{op} + C_4,$$

where

$$C_1 = \frac{2}{\nu_{\mathring{\boldsymbol{A}}}\left(\tau_1 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\right)},\ C_2 = \frac{2^{\frac{1}{q}} d^{\frac{1}{2}-\frac{1}{q}}}{\tau_1 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}},$$

$$C_3 = \frac{\left(1 + 2^{\frac{1}{q}-\frac{1}{2}} + 2^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\right)\|\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2}{\nu_{\mathring{\boldsymbol{A}}}\left(\tau_1 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\right)},$$

$$C_4 = \frac{2^{\frac{1}{q}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\|\boldsymbol{f}\|_2}{\tau_1 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}},$$

and $\|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q$ is the best s-term ℓ_q approximation error, denoted by $\|\boldsymbol{\Phi}^*\boldsymbol{f} - (\boldsymbol{\Phi}^*\boldsymbol{f})_{[s]}\|_q$. Obviously, C_i $(i = 1, 2, 3, 4)$ is positive because of $\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} < \tau_1$.

So far, the proof of theorem 2.1 is completed. □

Appendix B
Proof of Theorem 2.2

Let T be the indices of entries with s largest magnitudes in the vector $\tilde{\boldsymbol{\Phi}}^*\boldsymbol{f}$, and denote the complement of T by T^c. Setting $T_0 = T$, we decompose T_0^c into r sets of size k (to be chosen later) where T_1 corresponds to the locations of the k largest entries in $\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}$, T_2 to the next k largest entries and so on. Finally, we let $T_{01} = T_0 \bigcup T_1$ and $\boldsymbol{h} = \hat{\boldsymbol{f}} - \boldsymbol{f}$ where $\hat{\boldsymbol{f}}$ is the optimal solution of (1.4) and $\boldsymbol{f}$ is the original signal.

We now begin the proof of Theorem 2.2.

Proof. Step 1: Bounding the tail of $\boldsymbol{\Phi}^*\boldsymbol{f}$.

By construction of the T_j, we have that each coefficient of $\tilde{\boldsymbol{\Phi}}^*_{T_{j+1}}\boldsymbol{h}$, written $|\tilde{\boldsymbol{\Phi}}^*_{T_{j+1}}\boldsymbol{h}|_{(i)}$, is at most the average of those on T_j:

$$|\tilde{\boldsymbol{\Phi}}^*_{T_{j+1}}\boldsymbol{h}|_{(i)} \leq \|\tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_1/k,$$

squaring these terms and summing, and then taking the square root yields

$$\|\tilde{\boldsymbol{\Phi}}^*_{T_{j+1}}\boldsymbol{h}\|_2 \leq \|\tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_1/\sqrt{k} \leq k^{\frac{1}{2}-\frac{1}{q}}\|\tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_q,$$

that is,

$$\sum_{j\geq 2}\|\tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_2 \leq \sum_{j\geq 1} k^{\frac{1}{2}-\frac{1}{q}}\|\tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_q,$$

so

$$\sum_{j\geq 2}\|\tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_2^q \leq k^{\frac{q}{2}-1}\sum_{j\geq 1}\|\tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_q^q = k^{\frac{q}{2}-1}\|\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{h}\|_q^q. \quad \text{(B.1)}$$

Moreover

$$\begin{aligned}\sum_{j\geq 2}\|\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_2^q &= \left(\sum_{j\geq 2}\|\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_2^q\right)^{\frac{2}{q}\cdot\frac{q}{2}} \\ &\leq \left(r^{\frac{2}{q}-1}\sum_{j\geq 2}\|\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_2^2\right)^{\frac{q}{2}} \\ &= r^{1-\frac{q}{2}}\|\boldsymbol{\Phi}^*_{T_{01}^c}\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_{T_{01}^c}\boldsymbol{h}\|_2^q, \qquad \text{(B.2)}\end{aligned}$$

where the second inequality is due to the special case 4.2) of quasi-norm inequality with $r = \frac{d-s}{k}$.

Combining (B.1) with (B.2), and utilizing the triangle inequality, we have

$$\begin{aligned}\sum_{j\geq 2}\|\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h}\|_2^q &\leq \sum_{j\geq 2}\left(\|\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_2 + \|\tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_2\right)^q \\ &\overset{(a)}{\leq} \sum_{j\geq 2}\|\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_2^q + \sum_{j\geq 2}\|\tilde{\boldsymbol{\Phi}}^*_{T_j}\boldsymbol{h}\|_2^q \\ &\overset{(b)}{\leq} r^{1-\frac{q}{2}}\|\boldsymbol{\Phi}^*_{T_{01}^c}\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_{T_{01}^c}\boldsymbol{h}\|_2^q \\ &\quad + k^{\frac{q}{2}-1}\left(\|\tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{h}\|_q^q + 2\|\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}\|_q^q + \|\tilde{\boldsymbol{\Phi}}^* w\|_q^q\right),\end{aligned}$$

where (a) holds because of the special case 4.1) of quasi-norm inequality, and (b) uses the result of (A.6).

Taking the qth root of both sides for the above inequality, we get

$$\begin{aligned}\sum_{j\geq 2}\|\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h}\|_2 &\leq \left(\sum_{j\geq 2}\|\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h}\|_2^q\right)^{\frac{1}{q}} \\ &\leq 4^{\frac{1}{q}-1}\Bigg\{r^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^*_{T_{01}^c}\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_{T_{01}^c}\boldsymbol{h}\|_2 \\ &\quad + k^{\frac{1}{2}-\frac{1}{q}}\left(\|\tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{h}\|_q + 2\|\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}\|_q + \|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q\right)\Bigg\},\end{aligned}$$

where, the last inequality follows from the special case 4.2) of quasi-norm inequality. There is already the upper bound of $\|\tilde{\boldsymbol{\Phi}}^*\boldsymbol{w}\|_q$ as (A.3), so we next give a upper bound on $\|\tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{h}\|_q$ and $\|\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}\|_q$, respectively.

By the quasi-norm inequality and the triangle inequality, it is not hard to check that

$$\begin{aligned}\|\tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{h}\|_q &\leq s^{\frac{1}{q}-\frac{1}{2}}\|\tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{h}\|_2 \\ &\leq s^{\frac{1}{q}-\frac{1}{2}}\left(\|\tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{h} - \boldsymbol{\Phi}^*_T\boldsymbol{h}\|_2 + \|\boldsymbol{\Phi}^*_T\boldsymbol{h}\|_2\right), \qquad \text{(B.3)}\end{aligned}$$

and by (A.10), we have

$$\begin{aligned}\|\tilde{\boldsymbol{\Phi}}^*_{T^c}\boldsymbol{f}\|_q &\leq 2^{\frac{1}{q}-1}\left(\|\boldsymbol{\Phi}^*\boldsymbol{f} - \tilde{\boldsymbol{\Phi}}^*\boldsymbol{f}\|_q + \|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q\right) \\ &\leq 2^{\frac{1}{q}-1}\left(d^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^*\boldsymbol{f} - \tilde{\boldsymbol{\Phi}}^*\boldsymbol{f}\|_2 + \|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q\right) \\ &\leq 2^{\frac{1}{q}-1}\left(d^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\|\boldsymbol{f}\|_2 + \|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q\right). \qquad \text{(B.4)}\end{aligned}$$

Note in particular that $r = \frac{d-s}{k} \approx \frac{d}{k}$ is suitable by the partition of T_0^c. Hence, by (A.3), (B.3) and (B.4), we obtain

$$\begin{aligned}&\sum_{j\geq 2}\|\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h}\|_2 \\ \leq& 4^{\frac{1}{q}-1}\left(\frac{s}{k}\right)^{\frac{1}{q}-\frac{1}{2}}\Bigg\{\|\boldsymbol{\Phi}^*_T\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{h}\|_2 \\ &+ \left(\frac{d}{s}\right)^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^*_{T_{01}^c}\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_{T_{01}^c}\boldsymbol{h}\|_2 + \|\boldsymbol{\Phi}^*_T\boldsymbol{h}\|_2 \\ &+ 2^{\frac{1}{q}}\left(\frac{d}{s}\right)^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\|\boldsymbol{f}\|_2 + 2^{\frac{1}{q}}s^{\frac{1}{2}-\frac{1}{q}}\|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q \\ &+ \frac{(2d/s)^{\frac{1}{q}-\frac{1}{2}}}{\nu_{\mathring{\boldsymbol{A}}}}\left(\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} + 1\right)\|\boldsymbol{A} - \tilde{\boldsymbol{A}}\|_{op}\|\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\Bigg\}.\end{aligned}$$

Moreover, and based on the fact that $\left(\frac{d}{s}\right)^{\frac{1}{q}-\frac{1}{2}} \geq 1$ (due to $d \geq s$ and $0 < q \leq 1$), we have

$$\begin{aligned}&\|\boldsymbol{\Phi}^*_T\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{h}\|_2 + \left(\frac{d}{s}\right)^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^*_{T_{01}^c}\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_{T_{01}^c}\boldsymbol{h}\|_2 \\ &\leq \left(\frac{d}{s}\right)^{\frac{1}{q}-\frac{1}{2}}\left(\|\boldsymbol{\Phi}^*_T\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{h}\|_2 + \|\boldsymbol{\Phi}^*_{T_{01}^c}\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_{T_{01}^c}\boldsymbol{h}\|_2\right) \\ &\overset{(a)}{\leq} \left(\frac{d}{s}\right)^{\frac{1}{q}-\frac{1}{2}}\left\{2\left(\|\boldsymbol{\Phi}^*_T\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_T\boldsymbol{h}\|_2^2 + \|\boldsymbol{\Phi}^*_{T_{01}^c}\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*_{T_{01}^c}\boldsymbol{h}\|_2^2\right)\right\}^{\frac{1}{2}} \\ &= \sqrt{2}\left(\frac{d}{s}\right)^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^*\boldsymbol{h} - \tilde{\boldsymbol{\Phi}}^*\boldsymbol{h}\|_2,\end{aligned}$$

where (a) is from the special case 4.2) of quasi-norm inequality.

Thus

$$\sum_{j\geq 2}\|\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h}\|_2 \leq \alpha(\|\boldsymbol{\Phi}^*_T\boldsymbol{h}\|_2 + \beta), \qquad \text{(B.5)}$$

where

$$\begin{aligned}\alpha =& \frac{1}{2}\left(\frac{4s}{k}\right)^{\frac{1}{q}-\frac{1}{2}}, \\ \beta =& \sqrt{2}\left(\frac{d}{s}\right)^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\|\boldsymbol{h}\|_2 \\ &+ 2^{\frac{1}{q}}\left(\frac{d}{s}\right)^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\|\boldsymbol{f}\|_2 + 2^{\frac{1}{q}}s^{\frac{1}{2}-\frac{1}{q}}\|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q \\ &+ \frac{(2d/s)^{\frac{1}{q}-\frac{1}{2}}}{\nu_{\mathring{\boldsymbol{A}}}}\left(\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} + 1\right)\|(\boldsymbol{A} - \tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2.\end{aligned}$$

Step 2: Consequence of $\boldsymbol{D}$-RIP.

Since $\tilde{\boldsymbol{A}}$ satisfies the $\boldsymbol{D}$-RIP, by (B.5) and the fact that $\|\boldsymbol{D}\|_2 = \sqrt{\lambda_{\max}(\boldsymbol{D}\boldsymbol{D}^*)} = \sqrt{\lambda_{\max}(2\boldsymbol{I})} = \sqrt{2}$, we have

$$\begin{aligned} & 2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2 \\ \geq & \|\tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{h}\|_2 \geq \|\tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^*_{T_{01}}\boldsymbol{h}\|_2 - \sum_{j\geq 2}\|\tilde{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h}\|_2 \\ \geq & \sqrt{1-\tilde{\delta}_{s+k}}\|\boldsymbol{D}\boldsymbol{\Phi}^*_{T_{01}}\boldsymbol{h}\|_2 - \sqrt{1+\tilde{\delta}_k}\sum_{j\geq 2}\|\boldsymbol{D}\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h}\|_2 \\ \geq & \sqrt{1-\tilde{\delta}_{s+k}}\|\boldsymbol{D}\boldsymbol{\Phi}^*_{T_{01}}\boldsymbol{h}\|_2 - \alpha\sqrt{2(1+\tilde{\delta}_k)}(\|\boldsymbol{\Phi}^*_T\boldsymbol{h}\|_2 + \beta) \\ \geq & \sqrt{1-\tilde{\delta}_{s+k}}\|\boldsymbol{D}\boldsymbol{\Phi}^*_{T_{01}}\boldsymbol{h}\|_2 - \alpha\sqrt{2(1+\tilde{\delta}_k)}(\|\boldsymbol{h}\|_2 + \beta). \end{aligned}$$

Thus

$$\begin{aligned} \|\boldsymbol{D}\boldsymbol{\Phi}^*_{T_{01}}\boldsymbol{h}\|_2^2 \leq & \frac{1}{1-\tilde{\delta}_{s+k}}\Big\{2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2 \\ & + \alpha\sqrt{2(1+\tilde{\delta}_k)}(\|\boldsymbol{h}\|_2+\beta)\Big\}^2. \end{aligned} \tag{B.6}$$

Step 3: Consequence of the Mutual Coherence.

The following average inequality plays an important role and is employed repeatedly in our proof.

Lemma B.1 **([7]).** *For any values a, b, and $t > 0$, we have*

$$2ab \leq ta^2 + \frac{b^2}{t}.$$

We next set $T^1 = T \cap \{1, 2, \cdots, d_1\}, T^2 = \{j - d_1 | j \in T \setminus T^1\}$ and denote components of $\boldsymbol{h}$ corresponding to $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ by $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$, respectively. By applying Lemma B.1 with t_1 (to be chosen later), we have

$$\begin{aligned} \|\boldsymbol{\Phi}^*_{T_{01}}\boldsymbol{h}\|_2^2 = & \|\boldsymbol{D}^*_{1T^1_{01}}\boldsymbol{h}_1\|_2^2 + \|\boldsymbol{D}^*_{2T^2_{01}}\boldsymbol{h}_2\|_2^2 \\ = & < \boldsymbol{h}_1, \boldsymbol{D}_1\boldsymbol{D}^*_{1T^1_{01}}\boldsymbol{h}_1 > + < \boldsymbol{h}_2, \boldsymbol{D}_2\boldsymbol{D}^*_{2T^2_{01}}\boldsymbol{h}_2 > \\ \overset{(a)}{\leq} & \|\boldsymbol{h}_1\|_2\|\boldsymbol{D}_1\boldsymbol{D}^*_{1T^1_{01}}\boldsymbol{h}_1\|_2 + \|\boldsymbol{h}_2\|_2\|\boldsymbol{D}_2\boldsymbol{D}^*_{2T^2_{01}}\boldsymbol{h}_2\|_2 \\ \leq & \frac{t_1\|\boldsymbol{h}_1\|_2^2}{2} + \frac{\|\boldsymbol{D}_1\boldsymbol{D}^*_{1T^1_{01}}\boldsymbol{h}_1\|_2^2}{2t_1} \\ & + \frac{t_1\|\boldsymbol{h}_2\|_2^2}{2} + \frac{\|\boldsymbol{D}_2\boldsymbol{D}^*_{2T^2_{01}}\boldsymbol{h}_2\|_2^2}{2t_1}, \end{aligned} \tag{B.7}$$

here, (a) is by the triangular inequality.

We adopt the mutual coherence of $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$, analogous to the method in [26], to estimate $\|\boldsymbol{D}_1\boldsymbol{D}^*_{1T^1_{01}}\boldsymbol{h}_1\|_2^2 + \|\boldsymbol{D}_2\boldsymbol{D}^*_{2T^2_{01}}\boldsymbol{h}_2\|_2^2$. Here, in order to avoid repeated work, we give the result directly as follows:

$$\begin{aligned} & \|\boldsymbol{D}_1\boldsymbol{D}^*_{1T^1_{01}}\boldsymbol{h}_1\|_2^2 + \|\boldsymbol{D}_2\boldsymbol{D}^*_{2T^2_{01}}\boldsymbol{h}_2\|_2^2 \\ \leq & \frac{\mu(s+k)\|\boldsymbol{h}\|_2^2}{2} + \|\boldsymbol{D}\boldsymbol{\Phi}^*_{T_{01}}\boldsymbol{h}\|_2^2. \end{aligned} \tag{B.8}$$

Combining (B.6) with (B.7) and (B.8) yields

$$\begin{aligned} \|\boldsymbol{\Phi}^*_{T_{01}}\boldsymbol{h}\|_2^2 \leq & \frac{t_1}{2}\|\boldsymbol{h}\|_2^2 + \frac{1}{2t_1}\Big\{\frac{\mu(s+k)\|\boldsymbol{h}\|_2^2}{2} \\ & + \frac{1}{1-\tilde{\delta}_{s+k}}\Big(2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2 \\ & + \alpha\sqrt{2(1+\tilde{\delta}_k)}(\|\boldsymbol{h}\|_2+\beta)\Big)^2\Big\}. \end{aligned}$$

Step 4: Bounding the error.

Since $\boldsymbol{\Phi}$ is a tight frame, we have

$$\|\boldsymbol{h}\|_2^2 = \|\boldsymbol{\Phi}^*\boldsymbol{h}\|_2^2 = \|\boldsymbol{\Phi}^*_{T_{01}}\boldsymbol{h}\|_2^2 + \|\boldsymbol{\Phi}^*_{T^c_{01}}\boldsymbol{h}\|_2^2,$$

and

$$\begin{aligned} \|\boldsymbol{\Phi}^*_{T^c_{01}}\boldsymbol{h}\|_2^2 \leq & \Big(\sum_{j\geq 2}\|\boldsymbol{\Phi}^*_{T_j}\boldsymbol{h}\|_2\Big)^2 \\ \leq & \alpha^2(\|\boldsymbol{\Phi}^*_T\boldsymbol{h}\|_2+\beta)^2 \\ \leq & \alpha^2(\|\boldsymbol{h}\|_2+\beta)^2 \\ = & \alpha^2\|\boldsymbol{h}\|_2^2 + 2\alpha^2\beta\|\boldsymbol{h}\|_2 + \alpha^2\beta^2. \end{aligned}$$

Thus, by some simple calculations, we can show that

$$\begin{aligned} & \|\boldsymbol{h}\|_2^2 \\ \leq & \left\{\frac{t_1}{2} + \frac{\mu(s+k)}{4t_1} + \frac{\alpha^2(1+\tilde{\delta}_k)}{t_1(1-\tilde{\delta}_{s+k})} + \alpha^2\right\}\|\boldsymbol{h}\|_2^2 \\ & + \frac{\left\{2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\right\}^2}{2t_1(1-\tilde{\delta}_{s+k})} + \left\{\frac{\alpha^2(1+\tilde{\delta}_k)}{t_1(1-\tilde{\delta}_{s+k})} + \alpha^2\right\}\beta^2 \\ & + \frac{\alpha\sqrt{2(1+\tilde{\delta}_k)}}{2t_1(1-\tilde{\delta}_{s+k})}\cdot 2\left\{2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\right\}\beta \\ & + \frac{\alpha\sqrt{2(1+\tilde{\delta}_k)}}{t_1(1-\tilde{\delta}_{s+k})}\cdot 2\left\{2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\right\}\|\boldsymbol{h}\|_2 \\ & + \left\{\frac{\alpha^2(1+\tilde{\delta}_k)}{t_1(1-\tilde{\delta}_{s+k})} + \alpha^2\right\}\cdot 2\beta\|\boldsymbol{h}\|_2. \end{aligned}$$

Utilizing Lemma B.1 to the latter three terms of the above inequality (with constants t_2, t_3 to be chosen later), we have

$$\begin{aligned} & \|\boldsymbol{h}\|_2^2 \\ \leq & \left\{\frac{t_1}{2} + \frac{\mu(s+k)}{4t_1} + \frac{\alpha^2(1+\tilde{\delta}_k)}{t_1(1-\tilde{\delta}_{s+k})} + \alpha^2\right\}\|\boldsymbol{h}\|_2^2 \\ & + \frac{\left\{2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\right\}^2}{2t_1(1-\tilde{\delta}_{s+k})} + \left\{\frac{\alpha^2(1+\tilde{\delta}_k)}{t_1(1-\tilde{\delta}_{s+k})} + \alpha^2\right\}\beta^2 \\ & + \frac{\alpha\sqrt{2(1+\tilde{\delta}_k)}}{2t_1(1-\tilde{\delta}_{s+k})}\left\{\left(2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\right)^2 + \beta^2\right\} \\ & + \frac{\alpha\sqrt{2(1+\tilde{\delta}_k)}}{2t_1(1-\tilde{\delta}_{s+k})}\left\{\frac{\left(2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\right)^2}{t_2} + t_2\|\boldsymbol{h}\|_2^2\right\} \\ & + \left\{\frac{\alpha^2(1+\tilde{\delta}_k)}{t_1(1-\tilde{\delta}_{s+k})} + \alpha^2\right\}\left(\frac{\beta^2}{t_3} + t_3\|\boldsymbol{h}\|_2^2\right). \end{aligned}$$

Simplifying, this yields

$$V_1\|\boldsymbol{h}\|_2^2 \leq V_2\left\{2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\right\}^2 + V_3\beta^2,$$

where

$$V_1 = 1 - \frac{t_1}{2} - \frac{\mu(s+k)}{4t_1} - \frac{\alpha^2(1+\tilde{\delta}_k)}{t_1(1-\tilde{\delta}_{s+k})} - \alpha^2$$

$$
\begin{aligned}
&- \frac{t_2\alpha\sqrt{2(1+\tilde{\delta}_k)}}{2t_1(1-\tilde{\delta}_{s+k})} - \frac{t_3\alpha^2(1+\tilde{\delta}_k)}{t_1(1-\tilde{\delta}_{s+k})} - t_3\alpha^2, \\
V_2 =& \frac{1}{2t_1(1-\tilde{\delta}_{s+k})} + \frac{\alpha\sqrt{2(1+\tilde{\delta}_k)}}{2t_1(1-\tilde{\delta}_{s+k})} + \frac{\alpha\sqrt{2(1+\tilde{\delta}_k)}}{2t_1t_2(1-\tilde{\delta}_{s+k})}, \\
V_3 =& \frac{\alpha^2(1+\tilde{\delta}_k)}{t_1(1-\tilde{\delta}_{s+k})} + \alpha^2 + \frac{\alpha\sqrt{2(1+\tilde{\delta}_k)}}{2t_1(1-\tilde{\delta}_{s+k})} \\
&+ \frac{\alpha^2(1+\tilde{\delta}_k)}{t_1t_3(1-\tilde{\delta}_{s+k})} + \frac{\alpha^2}{t_3}. \qquad \text{(B.9)}
\end{aligned}
$$

Assuming $V_1 > 0$ (to be analyzed later), we obtain

$$
\|\boldsymbol{h}\|_2 \le \sqrt{\frac{V_2}{V_1}}\left\{2\varepsilon + \|(\boldsymbol{A}-\tilde{\boldsymbol{A}})\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\right\} + \sqrt{\frac{V_3}{V_1}}\beta.
$$

Introducing the expression of β and arranging yields

$$
\begin{aligned}
&\left\{1 - \left(\frac{d}{s}\right)^{\frac{1}{q}-\frac{1}{2}}\sqrt{\frac{2V_3}{V_1}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\right\}\|\boldsymbol{h}\|_2 \\
\le& 2\sqrt{\frac{V_2}{V_1}}\varepsilon + 2^{\frac{1}{q}}s^{\frac{1}{2}-\frac{1}{q}}\sqrt{\frac{V_3}{V_1}}\|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q \\
&+ \left\{\sqrt{\frac{V_2}{V_1}} + \frac{(2d/s)^{\frac{1}{q}-\frac{1}{2}}}{\nu_{\mathring{\boldsymbol{A}}}}\sqrt{\frac{V_3}{V_1}}\right. \\
&\left.+ \frac{(2d/s)^{\frac{1}{q}-\frac{1}{2}}}{\nu_{\mathring{\boldsymbol{A}}}}\sqrt{\frac{V_3}{V_1}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\right\}\|\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{f}\|_2\|\boldsymbol{A}-\tilde{\boldsymbol{A}}\|_{op} \\
&+ 2^{\frac{1}{q}}\left(\frac{d}{s}\right)^{\frac{1}{q}-\frac{1}{2}}\sqrt{\frac{V_3}{V_1}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\|\boldsymbol{f}\|_2.
\end{aligned}
$$

Let

$$
\tau_2 = \left(\frac{d}{s}\right)^{\frac{1}{2}-\frac{1}{q}}\sqrt{\frac{V_1}{2V_3}},
$$

by controlling the disturbance level of the frame $\boldsymbol{\Phi}$ such that $\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} < \tau_2$, then

$$
1 - \left(\frac{d}{s}\right)^{\frac{1}{q}-\frac{1}{2}}\sqrt{\frac{2V_3}{V_1}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} = 1 - \frac{1}{\tau_2}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} > 0.
$$

Therefore

$$
\|\boldsymbol{h}\|_2 \le C_5\varepsilon + C_6\|\boldsymbol{\Phi}^*_{T^c}\boldsymbol{f}\|_q + C_7\|\boldsymbol{A}-\tilde{\boldsymbol{A}}\|_{op} + C_8,
$$

where

$$
\begin{aligned}
C_5 &= \frac{\left(\frac{d}{s}\right)^{\frac{1}{2}-\frac{1}{q}}\sqrt{\frac{2V_2}{V_3}}}{\tau_2 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}}, \quad C_6 = \frac{(d/2)^{\frac{1}{2}-\frac{1}{q}}}{\tau_2 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}}, \\
C_7 &= \frac{\left(\frac{d}{s}\right)^{\frac{1}{2}-\frac{1}{q}}\sqrt{\frac{V_2}{2V_3}} + \frac{2^{\frac{1}{q}-1}}{\nu_{\mathring{\boldsymbol{A}}}} + \frac{2^{\frac{1}{q}-1}}{\nu_{\mathring{\boldsymbol{A}}}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}}{\tau_2 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}}, \\
C_8 &= \frac{2^{\frac{1}{q}-\frac{1}{2}}\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}\|\boldsymbol{f}\|_2}{\tau_2 - \|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op}}.
\end{aligned}
$$

Obviously, C_i $(i = 5,6,7,8)$ is positive because of $\|\boldsymbol{\Phi}^* - \tilde{\boldsymbol{\Phi}}^*\|_{op} < \tau_2$.

Step 5: The choice of the parameters.

Now we need to choose parameters to make sure that our hypothesis $V_1 > 0$ is valid. There are many parameters, i.e., $s, \mu, k, q, \tilde{\delta}_k, \tilde{\delta}_{s+k}, t_1, t_2, t_3$, in the expression of V_1 ($\alpha = \frac{1}{2}(\frac{4s}{k})^{\frac{1}{q}-\frac{1}{2}}$ is a function of s, k and q). It seems to cause trouble for our analysis. But we notice that the sparsity s and the mutual coherence μ can be small (the latter from Example II.1 in [26]). Moreover, $V_1(t_1,t_2,t_3)$ decreases as t_2, t_3 increase. Hence, we take t_2, t_3 arbitrarily small, i.e., $t_2, t_3 \to 0_+$, then $V_1(t_1,t_2,t_3)$ degenerates to

$$
V_1(t_1) = 1 - \frac{t_1}{2} - \frac{\mu(s+k)}{4t_1} - \frac{\alpha^2(1+\tilde{\delta}_k)}{t_1(1-\tilde{\delta}_{s+k})} - \alpha^2.
$$

Thus, let t_1 take the maximum point of $V_1(t_1)$, namely, $t_1 = \left\{\frac{\mu(s+k)}{2} + \frac{2\alpha^2(1+\tilde{\delta}_k)}{1-\tilde{\delta}_{s+k}}\right\}^{\frac{1}{2}}$. The remaining parameters are constrained to

$$
1 - \alpha^2 - \left\{\frac{\mu(s+k)}{2} + \frac{2\alpha^2(1+\tilde{\delta}_k)}{1-\tilde{\delta}_{s+k}}\right\}^{\frac{1}{2}} > 0, \qquad \text{(B.10)}
$$

such that $V_1 > 0$. Further mathematical derivation shows that (B.10) is equivalent to the following constraint

$$
\tilde{\delta}_{s+k} < W(s,\mu,k,q) := \frac{2(1-\alpha^2)^2 - \mu(s+k) - 4\alpha^2}{2(1-\alpha^2)^2 - \mu(s+k) + 4\alpha^2}.
$$

Specifically, we provide the choice of the parameters in the following four cases (but not all).

- Case 1: When $k = 4s$, $\alpha = \frac{1}{2}$ and $W(s,\mu,k,q)$ reduces to $W(s,\mu) = \frac{14-5\mu s}{22-5\mu s}$. If $\mu s \to 0$, then $\tilde{\delta}_{5s} < W(s,\mu) \to \frac{7}{11} \approx 0.636$.
- Case 2: when $k = 8s$ and $q = 1$, $\alpha = \frac{\sqrt{2}}{4}$ and $W(s,\mu,k,q)$ reduces to $W(s,\mu) = \frac{33-288\mu s}{65-288\mu s}$. If $\mu s \to 0$, then $\tilde{\delta}_{9s} < W(s,\mu) \to \frac{33}{65} \approx 0.508$.
- Case 3: when $k = 8s$ and $q = \frac{1}{2}$, $\alpha = \frac{\sqrt{2}}{8}$ and $W(s,\mu,k,q)$ reduces to $W(s,\mu) = \frac{1794-9216\mu s}{2050-9216\mu s}$. If $\mu s \to 0$, then $\tilde{\delta}_{9s} < W(s,\mu) \to \frac{897}{1025} \approx 0.875$.
- Case 4: when $k = 8s$ and $q \to 0$, $\alpha \to 0$ and $W(s,\mu,k,q)$ reduces to $W(s,\mu,q)$, then $\tilde{\delta}_{9s} <$ $W(s,\mu,q) = \frac{2\left\{1-(\frac{1}{2})^{\frac{2}{q}+1}\right\}^2 - 9\mu s - (\frac{1}{2})^{\frac{2}{q}-1}}{2\left\{1-(\frac{1}{2})^{\frac{2}{q}+1}\right\}^2 - 9\mu s + (\frac{1}{2})^{\frac{2}{q}-1}} \to 1.$

Up to now, this completes the proof of Theorem 2.2. □

REFERENCES

[1] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Proc. Mag.*, vol. 24, no. 4, pp. 118–121, Aug. 2007.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Mar. 2006.

[3] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2210–2219, Apr. 2008.

[4] J. Romberg, "Imaging via compressive sampling[introduction to compressive sampling and recovery via convex programming]," *IEEE Signal Proc. Mag.*, vol. 25, no. 2, pp. 14–20, Mar. 2008.

[5] G. H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (piccs)," *Medical physics*, vol. 35, no. 2, pp. 660–663, Sep. 2008.

[6] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang, and S. Yan, "Sparse representation for computer vision and pattern recognition," *Proc. IEEE.*, vol. 98, no. 6, pp. 1031–1044, Apr. 2010.

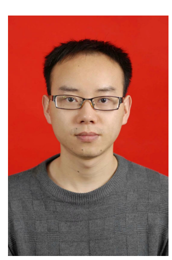
[7] R. Baraniuk and P. Steeghs, "Compressive radar imaging," in *Radar Conference, IEEE*, Boston, MA, USA, Apr. 2007, pp. 128–133.
[8] M. Elad, M. A. T. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in image processing," *Proc. IEEE*, vol. 98, no. 6, pp. 972–982, Feb. 2010.
[9] Y. Shen, B. Han, and E. Braverman, "Stable recovery of analysis based approaches," *Appl. Comput. Harmon. Anal.*, vol. 39, no. 1, pp. 161–172, Sep. 2014.
[10] A. Aldroubi, X. Chen, and A. M. Powell, "Perturbations of measurement matrices and dictionaries in compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 33, no. 2, pp. 282–291, Sep. 2012.
[11] S. Li and J. Lin, "Compressed sensing with coherent tight frames via l_q-minimization for $0 < q \leq 1$," *Inverse Probl. Imaging*, vol. 8, no. 3, pp. 761–777, Aug. 2014.
[12] J. Lin and S. Li, "Restricted q-isometry properties adapted to frames for nonconvex l_q-analysis," *IEEE Trans. Inf. Theory*, vol. 62, no. 8, pp. 4733–4747, May. 2016.
[13] Y. Wang, J. Wang, and Z. Xu, "On recovery of block-sparse signals via mixed $\ell_2/\ell_q(0 < q \leq 1)$ norm minimization," *EURASIP J. Adv. Sig. Pr.*, vol. 2013, no. 1, pp. 1–17, Apr. 2013.
[14] Z. Xu, H. Zhang, Y. Wang, X. Y. Chang, and Y. Liang, "L1/2 regularization," *Sci. China Inform. Sci.*, vol. 53, no. 6, pp. 1159–1169, Jun. 2010.
[15] Z. Han, J. Wang, J. Jing, and H. Zhang, "A simple gaussian measurement bound for exact recovery of block-sparse signals," *Discrete Dyn. Nat. Soc.*, vol. 2014, no. 3, pp. 1–8, Nov. 2014.
[16] R. J. Korneliussen, N. Diner, E. Ona, L. Berger, and P. G. Fernandes, "Proposals for the collection of multifrequency acoustic data," *ICES J. Mar. Sci.*, vol. 65, no. 6, pp. 982–994, Sep. 2008.
[17] W. Gobel and F. Helmchen, "In vivo calcium imaging of neural network function," *Physiology*, vol. 22, no. 6, pp. 358–365, Nov. 2007.
[18] J. Zeng, S. Lin, Y. Wang, and Z. Xu, "$l_{1/2}$ regularization: convergence of iterative half thresholding algorithm," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2317–2329, Feb. 2013.
[19] J. F. Cai, S. Osher, and Z. Shen, "Split bregman methods and frame based image restoration," *Siam J. Multiscale Model. Simul.*, vol. 8, no. 2, pp. 337–369, Jan. 2009.
[20] M. Elad, J. L. Starck, P. Querre, and D. L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (mca)," *Appl. Comput. Harmon. Anal.*, vol. 19, no. 3, pp. 340–358, Nov. 2005.
[21] G. Kutyniok. (2011) Data separation by sparse representations. [Online]. Available: https://arxiv.org/abs/1102.4527
[22] M. Zibulevsky and B. A. Pearlmutter, "Blind source separation by sparse decomposition in a signal dictionary," *Neural Comput.*, vol. 13, no. 4, pp. 863–882, Apr. 2001.
[23] D. Donoho and G. Kutyniok, "Microlocal analysis of the geometric separation problem," *Commun. Pure Appl. Math.*, vol. 66, no. 1, pp. 1–47, Apr. 2010.
[24] M. A. Herman and T. Strohmer, "General deviants: an analysis of perturbations in compressed sensing," *IEEE J. Sel. Top. Signa.*, vol. 4, no. 2, pp. 342–349, Feb. 2010.
[25] C. Y. Liu, J. J. Wang, W. D. Wang, and Y. Wang, "A perturbation analysis on compressed data separation with nonconvex minimization method," *Acta Electronica Sinica*, vol. 45, no. 1, pp. 37–45, Jan. 2017.
[26] J. Lin, S. Li, and Y. Shen, "Compressed data separation with redundant dictionaries," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4309–4315, Apr. 2013.
[27] R. Gribonval and M. Nielsen, "The restricted isometry property meets nonlinear approximation with redundant frames," *Journal of Approx. Theory*, vol. 165, no. 1, pp. 1–19, Jan. 2011.
[28] S. Foucart. (2009) Notes on compressed sensing. [Online]. Available: http://www.math.vanderbilt.edu/
[29] E. J. Cands, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. Comput. Harmon. Anal.*, vol. 31, no. 1, pp. 59–73, Jul. 2011.
[30] M. J. Lai and W. Yin, "Improved iteratively reweighted least squares for unconstrained smoothed ℓ_q minimization," *SIAM J. Numer. Anal.*, vol. 51, no. 2, pp. 927–957, Mar. 2013.

FENG ZHANG was born in Sichuan, China. He received the master's degree from the School of Mathematics and Statistics, Southwest University, Chongqing, China, in 2016. He is currently pursuing the Ph.D. degree with the School of Mathematics and Statistics, Southwest University. His research focus on compressed sensing, machine learning and computer vision.

JIANJUN WANG received the BS degree in mathematical education from NingXia University in 2000 and the MS degree in fundamental mathematics in 2003 from NingXia University, China. And Ph.D degree in applied mathematics was obtained from the Institute for Information and System Science, Xian Jiaotong University in Dec. 2006. He is currently a professor in the School of Mathematics & Statistics at at Southwest University of China. His research focus on machine learning, data mining, neural networks and sparse learning.

YAO WANG received the Ph.D. degree in applied mathematics from Xi'an Jiao Tong University, China, in2014. He is currently an assistant professor at the Department of Statistics in Xi'an Jiao Tong University, China. His current research interests include statistical signal processing, high dimensional statistical inference, and computational biology.

JIANWEN HUANG was born in Gansu, China. He received the master's degree from the School of Mathematics and Statistics, Southwest University, Chongqing, China, in 2013. He is currently pursuing the Ph.D. degree with the School of Mathematics and Statistics, Southwest University. His research interests include machine learning, data mining, sparse learning and extreme value theory.



WENDONG WANG received the B.S. degree from the School of Mathematics and Statistics, Southwest University, Chongqing, China, in 2014. Currently, he is pursuing the Ph.D. degree with the School of Computer and Information Science, Southwest University, Chongqing, China. His research interests include compressed sensing and machine learning.