PAPER

New Restricted Isometry Condition Using Null Space Constant for Compressed Sensing

Haiyang ZOU[†a)] ***and*** **Wengang ZHAO**[††b)], ***Nonmembers***

SUMMARY It has been widely recognized that in compressed sensing, many restricted isometry property (RIP) conditions can be easily obtained by using the null space property (NSP) with its null space constant (NSC) $0 < \theta \leq 1$ to construct a contradicted method for sparse signal recovery. However, the traditional NSP with $\theta = 1$ will lead to conservative RIP conditions. In this paper, we extend the NSP with $0 < \theta < 1$ to a scale NSP, which uses a factor τ to scale down all vectors belonged to the Null space of a sensing matrix. Following the popular proof procedure and using the scale NSP, we establish more relaxed RIP conditions with the scale factor τ, which guarantee the bounded approximation recovery of all sparse signals in the bounded noisy through the constrained ℓ_1 minimization. An application verifies the advantages of the scale factor in the number of measurements.

key words: *compressed sensing, null space property, null space constant, restricted isometry property, ℓ_1-norm minimization*

1. Introduction

Compressed sensing (CS) [4], [5] has attracted much attention over the last decade due to the numerous applications in imaging processing [6], compressive phase retrieval [7], and statistics [8]. In particular, without the need of full signal recovery CS has also been applied into several signal processing aspects, such as estimation, detection and classification [9]–[13]. A central problem of CS is to recover an unknown sparse signal $\mathbf{x} \in \mathbb{R}^p$ from the observed data $\mathbf{y} \in \mathbb{R}^n$ where

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{z}. \tag{1}$$

Here $\mathbf{A} \in \mathbb{R}^{n\times p} (n \ll p)$ is a measurement matrix, and $\mathbf{z} \in \mathbb{R}^n$ is a measurement error. It is well-known that the sparse recovery based on ℓ_0-norm minimization is a NP-hard problem. Thus, the constrained ℓ_1-norm minimization has been well studied to provide an efficient method for sparse signal recovery. Generally speaking, the approach is to find the sparsest signal by solving the following problem [14]:

$$\widehat{\mathbf{x}} = \min_{\mathbf{x}} \{\|\mathbf{x}\|_1 \text{ s.t. } \mathbf{A}\mathbf{x} - \mathbf{y} \in \mathcal{B}\}, \tag{2}$$

where $\mathcal{B}$ is a bounded set determined by the noise structure, particularly, $\mathcal{B} = \{0\}$ means the noiseless case. In the problem (2), the most well-known properties, such as the null space property (NSP) [3], and the restricted isometry property (RIP) [14], have been introduced and studied to recover the sparse signals in the literatures [4], [5], [15]–[20]. For k-sparse signal recovery, NSP satisfies an inequality that the ℓ_1 norm of the largest k entries is less than that of the rests in null space of the sensing matrix, while RIP considers the smallest number to estimate the reconstruction bound by using k-sparse vectors. Based on NSP, there is a popular proof procedure to obtain the RIP conditions (upper bounds of the smallest number) by using contradicted methods to prove the inequality of NSP [1], [2], [21]–[25]. For discussing the proof procedure, we first introduce the NSP and RIP definitions as follows.

1.1 NSP and RIP

Some basic notations are given. A vector $\mathbf{v} \in \mathbb{R}^p$ is said to be k-sparse if $S = |\text{supp}(\mathbf{v})| < k$, where $S = \{i, \mathbf{v}_i \neq 0\}$ is the support of $\mathbf{v}$. $\mathbf{v}_{\max(k)}$ is a vector $\mathbf{v}$ with all but the largest k entries in absolute value set to zero and $\mathbf{v}_{-\max(k)} = \mathbf{v} - \mathbf{v}_{\max(k)}$. The kernel of matrix $\mathbf{A}$ is denoted as $\ker(\mathbf{A}) = \{\mathbf{v} : \mathbf{A}\mathbf{v} = 0\}$.

One of the most widely used frameworks is the NSP with a null space constant (NSC) θ $(0 < \theta \leq 1)$ [24], [26]–[34] for compressed sensing. Furthermore, θ can be tested by using semidefinite programming [28] and sandwiching algorithm [30]. Now we give the (k, θ)-NSP by following the definition 2.1 [24].

Definition 1: [24] A matrix $\boldsymbol{A}$ satisfies the null space property of order k if $\forall \boldsymbol{v} \in \ker(\boldsymbol{A}) \setminus \{0\}$ and $\forall |S| < k$,

$$\|\boldsymbol{v}_S\|_1 < \theta \|\boldsymbol{v}_{S^c}\|_1, \tag{3}$$

where S^c is the complement set of S, and $\boldsymbol{v}_S$ is the vector that has the same entry as $\boldsymbol{v}$ on S, but 0 everywhere.

Based on the above NSP with $\theta = 1$, there has a useful characterization [35] of the matrix $\mathbf{A}$ that (2) with $\mathcal{B} = \{0\}$ will produce the solution of (1) with $\mathbf{z} = 0$ if and only if $\forall \mathbf{v} \in \ker(\mathbf{A})\setminus\{0\}$ and $\forall |S| < k$, $\|\mathbf{v}_S\|_1 < \|\mathbf{v}_{S^c}\|_1$. Of course, we need not to check this inequality for all subsets S; checking the subset with the k largest (in absolute value) elements of $\mathbf{v}$ is sufficient [35]. Thus we have the following lemma, which is from [1].

Lemma 1 [1]: Using (2) with $\mathcal{B} = \{0\}$ one can recover all k-sparse signal $\boldsymbol{x}$ if and only if $\forall \boldsymbol{v} \in \ker(\boldsymbol{A}) \setminus \{0\}$,

$$\|\boldsymbol{v}_{\max(k)}\|_1 < \|\boldsymbol{v}_{-\max(k)}\|_1. \tag{4}$$

Manuscript received December 23, 2021.
Manuscript revised March 25, 2022.
Manuscript publicized June 20, 2022.

[†]The author is with the School of Computer Science, China West Normal University, China.
[††]The author is with the School of Economics and Management, Southwest University of Science and Technology, China.
a) E-mail: haiyangzou@126.com
b) E-mail: zglx0615@126.com (Corresponding author)

DOI: 10.1587/transfun.2021EAP1175

Another one of the most widely used frameworks is the RIP property [1], [2], [21], [22], [24] for compressed sensing. Here, we follow the definition [14] and give the RIP property as follows.

Definition 2 (RIP): [1], [14] Let $\mathbf{A}$ *be a measurement matrix and let* $1 \le k, k_1, k_2 \le p$ *be integers. We define the k-restricted isometry constant (RIC) to be the smallest number* $\delta_k^{\mathbf{A}}$ *such that*

$$\left(1-\delta_k^{\mathbf{A}}\right)\|\boldsymbol{c}\|_2^2 \le \|\mathbf{A}\boldsymbol{c}\|_2^2 \le \left(1+\delta_k^{\mathbf{A}}\right)\|\boldsymbol{c}\|_2^2,$$

for all k-sparse vectors $\boldsymbol{c}$*. Restricted orthogonality constant (ROC) of order* (k_1, k_2) *is defined to be the smallest non-negative number* $\delta_{k_1,k_2}^{\mathbf{A}}$ *such that*

$$|\langle \mathbf{A}\boldsymbol{a}, \mathbf{A}\boldsymbol{b}\rangle| \le \delta_{k_1,k_2}^{\mathbf{A}}\|\boldsymbol{a}\|_2\|\boldsymbol{b}\|_2,$$

for all k_1*-sparse vectors* $\boldsymbol{a}$ *and* k_2*-sparse vectors* $\boldsymbol{b}$ *with disjoint supports.*

Simply, the RIC $\delta_k^{\mathbf{A}}$ and ROC $\delta_{k_1,k_2}^{\mathbf{A}}$ in an orthonormal system are used to measure how far subsets of cardinality of columns of are. It is obvious that δ_k and ROC δ_{k_1,k_2} are increasing in each of their indices [1].

Following the inequality (4), the RIP conditions are obtained by using an assumptive inequality $\|\mathbf{v}_{\max(k)}\|_1 \ge \|\mathbf{v}_{-\max(k)}\|_1$ to result in contradictions for building upper bounds in terms of the RIC and ROC, such as, $\delta_{2k}^{\mathbf{A}} < 1/\sqrt{2}$ [24], $\delta_{1.25k}^{\mathbf{A}} + \delta_{k,1.25k}^{\mathbf{A}} < 1$ [21] and $\delta_k^{\mathbf{A}} + \delta_{2k}^{\mathbf{A}} + \delta_{3k}^{\mathbf{A}} < 1$ [36]. Recently, Cai et al. developed new (sharp) RIP conditions to ensure the sparse recovery, for example, $\delta_k^{\mathbf{A}} + \delta_{k,k}^{\mathbf{A}} < 1$ [1] and $\delta_{tk}^{\mathbf{A}} \le \sqrt{(t-1)/t}$ $(t \ge 4/3)$ [2]. Note that the above RIP conditions obtained by using NSP is different from the hybrid of RIP and NSP in the definition 1.1 [24] because this hybrid requires a matrix to satisfy both the NSP and RIP conditions (the inequalities (3) and (4)).

1.2 Challenge

To further exploit the NSP property with the NSC θ, the inequality (4) is rewritten as:

$$\|\mathbf{v}_{\max(k)}\|_1 < \theta\|\mathbf{v}_{-\max(k)}\|_1, \tag{5}$$

with $0 < \theta \le 1$. Clearly, the inequality (5) is stronger than the standard inequality (4) in the NSP property for sparse recovery since $\|\mathbf{v}_{\max(k)}\|_1 < \theta\|\mathbf{v}_{-\max(k)}\|_1 \le \|\mathbf{v}_{-\max(k)}\|_1$, where $\mathbf{v}_{\max(k)}$, $\mathbf{v}_{-\max(k)}$ and $\mathbf{v}$ are shown in the orange, dark blue, and green lines of Fig. 1, respectively. It obviously holds a distance $(1-\theta)\|\mathbf{v}_{-\max(k)}\|_1$ shown in the red line of Fig. 1. However, the inequality (5) results in conservative RIP conditions as $0 < \theta < 1$ reduces the upper bounds of RIC and ROC by directly using the assumptive inequalities $\|\mathbf{v}_{\max(k)}\|_1 \ge \theta\|\mathbf{v}_{-\max(k)}\|_1$, for example, the proposition 1 in Sect. 2.1. Therefore, this leads to a natural but crucial challenge:

Challenge: *How to improve the RIP conditions by using the distance* $(1-\theta)\|\boldsymbol{v}_{-\max(k)}\|_1$*?*

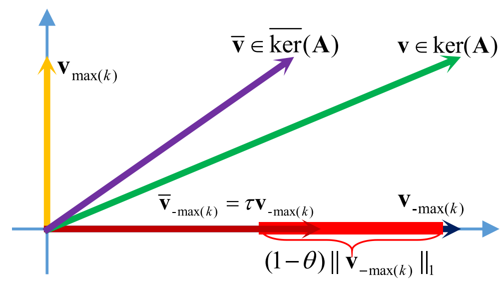

Fig. 1 The geometrical relationship of the distance τ, $\mathbf{v} \in \overline{\ker}(\mathbf{A}) \setminus \{0\}$, $\mathbf{v}_{\max(k)}$, $\mathbf{v}_{-\max(k)}$, $\overline{\mathbf{v}} \in \overline{\ker}(\mathbf{A}) \setminus \{0\}$, and $\overline{\mathbf{v}}_{-\max(k)}$, where $\theta \le \tau \le 1$, $\mathbf{v}_{\max(k)}$ is a vector $\mathbf{v}$ with all but the largest k entries in absolute value set to zero, $\mathbf{v}_{-\max(k)} = \mathbf{v} - \mathbf{v}_{\max(k)}$, and $\overline{\ker}(\mathbf{A})$ is defined in (10).

1.3 Motivation and Contributions

To address this important problem, the key idea is that the inequality (5) is considered as a NSP characterization to expand the inequality (4) for the improvement of the RIP conditions. In this paper we introduce a factor τ to scale down the vector $\mathbf{v}_{-\max(k)}$ into $\overline{\mathbf{v}}_{-\max(k)} = \tau\mathbf{v}_{-\max(k)}$ with $\theta \le \tau \le 1$. This leads to a scale kernel with a noise $\|\mathbf{A}\overline{\mathbf{v}}\|_2 < (1-\tau)\rho\|\mathbf{v}_{-\max(k)}\|_2$ $(\rho = \sqrt{\lambda_{\max}(\mathbf{A}^T\mathbf{A})})$ for all vectors $\overline{\mathbf{v}} = \mathbf{v}_{\max(k)} + \overline{\mathbf{v}}_{-\max(k)}$. Based on the scale kernel, the inequality (5) is transformed into a pseudo-standard inequality $\|\mathbf{v}_{\max(k)}\|_1 < \|\overline{\mathbf{v}}_{-\max(k)}\|_1$. Similarly, the inequality (4) is also converted into novel inequality $\|\mathbf{v}_{\max(k)}\|_1 < \|\mathbf{v}_{-\max(k)}\|_1 = \frac{1}{\tau}\|\overline{\mathbf{v}}_{-\max(k)}\|_1$, which is better than the pseudo-standard inequality. According to the proof, therefore, this novel inequality will conduct a new assumption $\|\mathbf{v}_{\max(k)}\|_1 \ge \frac{1}{\tau}\|\overline{\mathbf{v}}_{-\max(k)}\|_1$, which leads to an expansion of the upper bounds of RIC $\delta_k^{\mathbf{A}}$ and ROC $\delta_{k,k}^{\mathbf{A}}$, for studying more relaxed RIP conditions. In summary, our contributions include the following two parts.

- To the best of our knowledge, we are the first to introduce a **scale factor** τ $(\theta \le \tau \le 1)$ to breakthrough the upper bound of the previous RIP conditions in the bounded recovery as it is reached by proving a novel inequality $\|\mathbf{v}_{\max(k)}\|_1 < \frac{1}{\tau}\|\overline{\mathbf{v}}_{-\max(k)}\|_1$ in the CS problem (2).
- We establish new upper bounds of the RIP conditions for sparse signal recovery in Table 1. More specifically, it is shown that there exists a scale factor τ such that $\delta_k^{\mathbf{A}} + \delta_{k,k}^{\mathbf{A}} < 1 + (1-\sqrt{\tau})\delta_{k,k}^{\mathbf{A}}$ and $\delta_{tk}^{\mathbf{A}} < \left(1 + \frac{1-\tau^2}{\tau^2-1+2t}\right)\sqrt{\frac{t-1}{t}}$ $(\theta \le \tau \le 1$ and $t \ge \frac{4}{3})$ guarantee the bounded approximation recovery of all k-sparse signals in the noiseless or noise cases.

Note that compared to the RIP conditions [1], [2], our results have $\varepsilon_1 = (1-\sqrt{\tau})\delta_{k,k}^{\mathbf{A}}$ and $\varepsilon_2 = \left(\frac{1-\tau^2}{\tau^2-1+2t}\right)\sqrt{\frac{t-1}{t}}$ improvements, respectively. However, the Remark 2.1 [1] and the Theorem 2.2 [2] show that for any $\varepsilon > 0$, the conditions $\delta_k^{\mathbf{A}} + \delta_{k,k}^{\mathbf{A}} < 1 + \varepsilon$ and $\delta_{tk}^{\mathbf{A}} \le \sqrt{(t-1)/t} + \varepsilon$ are not sufficient to ensure stably recovery using any method [1], [2] in the bounded noisy set $\mathcal{B}$. In fact, they are special cases $\tau = 1$ of

Table 1 Comparison on NSP and RIP conditions. Our RIP results are more relaxed conditions than [1] and [2] in the bounded recovery.

Methods	Previous results	Our results	Improvement
NSP	$\|\mathbf{v}_{\max(k)}\|_1 < \|\mathbf{v}_{-\max(k)}\|_1$ [3]	$\|\mathbf{v}_{\max(k)}\|_1 < \frac{1}{\tau}\|\overline{\mathbf{v}}_{-\max(k)}\|_1$	\
RIP	$\delta_k^{\mathbf{A}} + \delta_{k,k}^{\mathbf{A}} < 1$ [1]	$\delta_k^{\mathbf{A}} + \delta_{k,k}^{\mathbf{A}} < 1 + (1-\sqrt{\tau})\delta_{k,k}^{\mathbf{A}}$	$(1-\sqrt{\tau})\delta_{k,k}^{\mathbf{A}}$
RIP	$\delta_{tk}^{\mathbf{A}} < \sqrt{\frac{t-1}{t}}$ [2]	$\delta_{tk}^{\mathbf{A}} < \left(1 + \frac{1-\tau^2}{\tau^2-1+2t}\right)\sqrt{\frac{t-1}{t}}$	$\left(\frac{1-\tau^2}{\tau^2-1+2t}\right)\sqrt{\frac{t-1}{t}}$
$\theta \le \tau \le 1$, $t \ge 4/3$, $\mathbf{v}$ belongs to the Null space of $\mathbf{A}$, $\mathbf{v}_{\max(k)}$ is regarded as $\mathbf{v}$ with all but the largest k entries in absolute value set to zero, and $\mathbf{v}_{-\max(k)} = \mathbf{v} - \mathbf{v}_{\max(k)}$. $\delta_k^{\mathbf{A}}$ and $\delta_{k,k}^{\mathbf{A}}$ are respectively regarded as restricted isometry constant (RIC) and restricted orthogonality constant (ROC).			

our results. When $\tau = 1$, $\varepsilon_1 = 0$ and $\varepsilon_2 = 0$, that means, we have not any improvement. As long as $\theta \le \tau \le 1$, that is, there is a nonzero distance $(1-\theta)\|\mathbf{v}_{-\max(k)}\|_1 > 0$, we can use the distance to breakthrough the upper bound of the previous RIP condition [1], [2]. Although there exists an additional noise ϵ, the more relaxed RIP conditions will lead to the less number of measurements in practice when all $\mathbf{v}_{-\max(k)}$ satisfy $\mathbf{A}\mathbf{v}_{-\max(k)} = 0$ or the recovery error can be tolerated.

2. Main Results

In this section, we first introduce a scale factor τ to scale down $\mathbf{v}_{-\max(k)}$ in the inequality (4) for improving the RIP conditions. Based on the scale factor τ, we propose two more relaxed RIP conditions to ensure the bounded approximation recovery of all k-sparse signals. Moreover, we also establish two RIP conditions to expand the the sparsity order k.

2.1 Scale Factor τ

Before introducing the scale factor τ, we first show the poor RIP conditions by proving the inequality (5) with θ as follows.

Proposition 1. Consider the signal recovery model $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{x}$ is a k-sparse vector. If

$$\delta_k^{\mathbf{A}} + \delta_{k,k}^{\mathbf{A}} < 1 + \left(\frac{\theta - 1}{\theta}\right)\delta_{k,k}^{\mathbf{A}}, \quad or \tag{6}$$

$$\delta_{tk}^{\mathbf{A}} \le \left(1 + \frac{\theta^2 - 1}{1 - \theta^2 + 2\theta^2 t}\right)\sqrt{\frac{t-1}{t}}, \tag{7}$$

for some $0 < \theta \le 1$ and $t \ge \frac{4}{3}$, then $\widehat{\boldsymbol{x}}$ recovers $\boldsymbol{x}$ exactly, where $\widehat{\boldsymbol{x}}$ is the ℓ_1 norm minimizer of (2) with $\mathcal{B} = \{0\}$.

Unfortunately, Proposition 1 shows that the RIP conditions (6) and (7) are less than $\delta_k^{\mathbf{A}} + \delta_{k,k}^{\mathbf{A}} < 1$ [1] and $\delta_{tk}^{\mathbf{A}} < \sqrt{(t-1)/t}$ [2], respectively. Compared to the inequality (4), however, there is a distance $(1-\theta)\|\mathbf{v}_{-\max(k)}\|_1 \ge 0$ in the inequality (5). Thus, this distance leads us to study the crucial challenge in the Sect. 1.2.

We propose a solution to transform the distance into the inequality (4) for improving the RIP conditions. To keep the orientation of $\mathbf{v}_{-\max(k)}$ and construct a relationship between the distance and $\mathbf{v}_{-\max(k)}$, we introduce a **scale factor** τ to scale down $\mathbf{v}_{-\max(k)}$ into $\overline{\mathbf{v}}_{-\max(k)}$, which is defined as:

$$\overline{\mathbf{v}}_{-\max(k)} = \tau\mathbf{v}_{-\max(k)}. \tag{8}$$

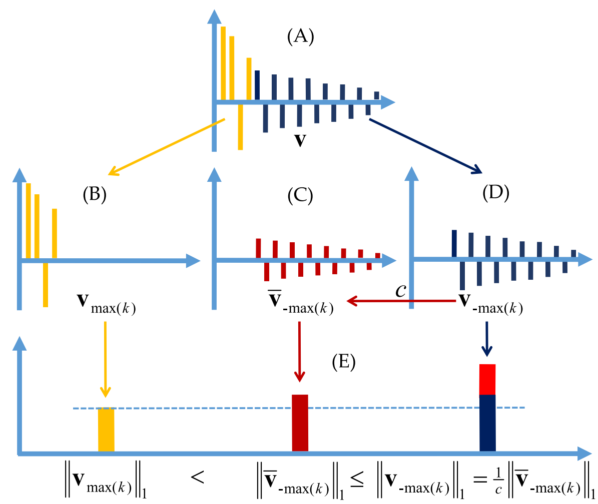

Fig. 2 Scale factor τ ($\theta \le \tau \le 1$). (A), (B), (C) and (D) respectively show $\mathbf{v}$, $\mathbf{v}_{\max(k)}$, $\overline{\mathbf{v}}_{-\max(k)}$ and $\mathbf{v}_{-\max(k)}$, where $\overline{\mathbf{v}}_{-\max(k)} = \tau\mathbf{v}_{-\max(k)}$. (E) shows $\|\mathbf{v}_{\max(k)}\|_1$, $\|\overline{\mathbf{v}}_{-\max(k)}\|_1$ and $\|\mathbf{v}_{-\max(k)}\|_1$, and the red bar exhibits a compressed distance between $\|\mathbf{v}_{\max(k)}\|_1$ and $\|\mathbf{v}_{-\max(k)}\|_1$.

In order to satisfy the inequality (5), it still holds $\theta\|\mathbf{v}_{-\max(k)}\|_1 \le \|\overline{\mathbf{v}}_{-\max(k)}\|_1 = \tau\|\mathbf{v}_{-\max(k)}\|_1 \le \|\mathbf{v}_{-\max(k)}\|_1$. Clearly, $\theta \le \tau \le 1$. Note that the computation of the scale factor τ is suspected to be NP-hard since it depends on the smallest NSC θ, which is a NP-hard computation [18]. Our aim is that the inequality (5) is transformed into a pseudo-standard inequality $\|\mathbf{v}_{\max(k)}\|_1 < \|\overline{\mathbf{v}}_{-\max(k)}\|_1$ to explore the distance by expanding the inequality (4). The whole process is shown in Fig. 2. When $\mathbf{v}$ is divided into $\mathbf{v}_{\max(k)}$ and $\mathbf{v}_{-\max(k)}$, Fig. 2(A) is changed into Fig. 2(B) and Fig. 2(D). $\mathbf{v}_{-\max(k)}$ is transformed into $\overline{\mathbf{v}}_{-\max(k)}$ by the scale factor τ, that is, Fig. 2(D) is changed into Fig. 2(C). Figure 2(E) shows the ℓ_1 norms of $\mathbf{v}_{\max(k)}$, $\overline{\mathbf{v}}_{-\max(k)}$ and $\mathbf{v}_{-\max(k)}$, and the red bar depicts the transformed distance.

Correspondingly, we have a new vector

$$\overline{\mathbf{v}} = \mathbf{v}_{\max(k)} + \overline{\mathbf{v}}_{-\max(k)}, \tag{9}$$

where $\overline{\mathbf{v}}$ is shown in purple line of Fig. 1. Clearly, $\overline{\mathbf{v}} \notin \ker(\mathbf{A})$. There has a recovery noise $\|\mathbf{A}\overline{\mathbf{v}}\|_2 = \|\mathbf{A}\mathbf{v} - \mathbf{A}\overline{\mathbf{v}}\|_2 = (1-\tau)\|\mathbf{A}\mathbf{v}_{-\max(k)}\|_2 \le (1-\tau)\|\mathbf{A}\|_2\|\mathbf{v}_{-\max(k)}\|_2 = (1-\tau)\sqrt{\lambda_{\max}(\mathbf{A}^T\mathbf{A})}\|\mathbf{v}_{-\max(k)}\|_2$. Thus, $\overline{\mathbf{v}}$ belongs to a **scale kernel** defined as:

$$\overline{\ker}(\mathbf{A}) = \{\overline{\mathbf{v}} : \|\mathbf{A}\overline{\mathbf{v}}\|_2 < (1-\tau)\rho\|\mathbf{v}_{-\max(k)}\|_2 \& \mathbf{A}\mathbf{v} = 0\}, \tag{10}$$

where $\rho = \sqrt{\lambda_{\max}(\mathbf{A}^T\mathbf{A})}$.

After the scale transformation in (8), we follow the Lemma 1 [1], and have that using (2) with $\mathcal{B} = \{0\}$ one can recover all k-sparse signal $\mathbf{x}$ if $\forall \mathbf{v} \in \overline{\ker}(\mathbf{A}) \setminus \{0\}$,

$$\|\mathbf{v}_{\max(k)}\|_1 < \|\mathbf{v}_{-\max(k)}\|_1 = \frac{1}{\tau}\|\overline{\mathbf{v}}_{-\max(k)}\|_1. \tag{11}$$

The above inequality (11) shows that the scale factor τ is successfully joined into the inequality (4).

In the popular proof procedure, therefore, the RIP conditions can be obtained by employing a new assumption $\|\mathbf{v}_{\max(k)}\|_1 \geq \frac{1}{\tau}\|\overline{\mathbf{v}}_{-\max(k)}\|_1$ to conduct contradictions for all $\overline{\mathbf{v}} \in \overline{\ker}(\mathbf{A}) \setminus \{0\}$. When $\tau = 1$, $\overline{\ker}(\mathbf{A})$ is same to $\ker(\mathbf{A})$. When $\theta \leq \tau < 1$, there exists a recovery error ρ in $\overline{\ker}(\mathbf{A})$. Although $\overline{\ker}(\mathbf{A})$ has the error ρ, proving the inequality (11) will lead to more relaxed RIP conditions due to $\frac{1}{\tau} > 1$. More importantly, these conditions will conduct the less number of measurements in practice when all $\mathbf{v}_{-\max(k)}$ satisfy $\mathbf{A}\mathbf{v}_{-\max(k)} = 0$ or the recovery error $(1-\tau)\rho$ can be tolerated. Thus, we will make use of the inequality (11) in the following subsection.

2.2 New RIP Conditions by RIC and ROC

In this subsection, we study new RIP conditions by RIC and ROC given in Definition 2. The following theorem shows a sufficient condition guarantees the stable sparse recovery in the noisy case through the constrained ℓ_1 minimization. We denote that for $1 \leq a \leq k$ and $b \geq 1$,

$$C_{a,b,k,\tau} = \max\left\{\frac{(1+\tau)k-a}{\sqrt{ab}}, \sqrt{\frac{(1+\tau)k-a}{a}}\right\}, \tag{12}$$

where $0 < \theta \leq \tau \leq 1$, and θ is given in (5).

We consider a bounded noise setting: $\|\mathbf{z}\|_2 < \epsilon$, and a recovery noise setting: $\|\mathbf{A}\overline{\mathbf{v}}\|_2 < (1-\tau)\rho$. The Gaussian noise, which is significant interest in statistics, can be essentially reduced to the bounded noise case.

Theorem 1. Denote that $\overline{\boldsymbol{v}}_{-\max(k)} = \tau\boldsymbol{v}_{-\max(k)}$ *with* $\theta \leq \tau \leq 1$, *and* $\|\boldsymbol{A}\overline{\boldsymbol{v}}\|_2 < (1-\tau)\rho$ *for all* $\overline{\boldsymbol{v}} \in \overline{\ker}(\boldsymbol{A}) \setminus \{0\}$. *Consider the signal recovery model* (1) *with* $\|\boldsymbol{z}\|_2 \leq \epsilon$. *Let* $\widehat{\boldsymbol{x}}$ *be the minimizer of* (2) *with* $\mathcal{B} = \{\boldsymbol{z} : \|\boldsymbol{z}\|_2 \leq \eta\}$ *for some* $\eta \geq \epsilon$. *If for some positive integers* $1 \leq a \leq k$ *and* $b \geq 1$,

$$\delta_a^{\boldsymbol{A}} + C_{a,b,k,\tau}\delta_{a,b}^{\boldsymbol{A}} < 1, \tag{13}$$

with a scale factor $\theta \leq \tau \leq 1$, *then*

$$\begin{aligned}\left\|\widehat{\boldsymbol{x}} - \boldsymbol{x}\right\|_2 \leq &\frac{\sqrt{2\left(1+\delta_a^{\boldsymbol{A}}\right)k/a}}{1-\delta_a^{\boldsymbol{A}} - C_{a,b,k,\tau}\delta_{a,b}^{\boldsymbol{A}}}(\epsilon + \eta + (1-\tau)\rho)\\ &+ 2\tau\chi\|\boldsymbol{x}_{-\max(k)}\|_1,\end{aligned} \tag{14}$$

where $\chi = \left(\frac{\sqrt{2k}C_{a,b,k,\tau}\delta_{a,b}^{\boldsymbol{A}}}{\left(1-\delta_a^{\boldsymbol{A}}-C_{a,b,k,\tau}\delta_{a,b}^{\boldsymbol{A}}\right)((1+\tau)k-a)} + \frac{1}{\sqrt{k}}\right)$ *and* ρ *is defined in* (10).

Remark 1. Although there is an additional noise $(1-\tau)\rho\frac{\sqrt{2(1+\delta_a^{\mathbf{A}})k/a}}{1-\delta_a^{\mathbf{A}}-C_{a,b,k,\tau}\delta_{a,b}^{\mathbf{A}}}$ to increase the upper bound of recovery error, $C_{a,b,k,\tau} \leq C_{a,b,k}$ and τ are used to decrease the first and second terms of the right of the Eq. (14). Compared to the upper bound of $\|\widehat{\beta} - \beta\|_2$ in the Eq. (19) in the literature [1], thus, it also has some additional gain in the Eq. (14). Considering $a = b = k$, $C_{a,b,k,\tau} = \sqrt{\tau}$, $C_{a,b,k} = 1$ and $\rho = \sqrt{1+\delta_a^{\mathbf{A}}}\|\mathbf{v}_{-\max(k)}\|_2$, we get a better recovery error $\left\|\widehat{\mathbf{x}} - \mathbf{x}\right\|_2 \leq \|\widehat{\beta} - \beta\|_2$ if it holds the following condition

$$\|\mathbf{x}_{-\max(k)}\|_1 \geq \frac{\sqrt{k}\left(1+\delta_k^{\mathbf{A}}\right)}{\sqrt{2}(1-\delta_k^{\mathbf{A}} - \sqrt{\tau}\delta_{k,k}^{\mathbf{A}})}\|\mathbf{x}_{-\max(k)}\|_2. \tag{15}$$

We now consider another bounded noise setting $\|\mathbf{A}^T\mathbf{z}\|_\infty < \epsilon$, and the recovery noise setting $\|\mathbf{A}^T\mathbf{A}\overline{\mathbf{v}}\|_\infty < (1-\tau)\rho$. This case is inspired by the Dantzig Selector method [16] for the Gaussian noise case.

Corollary 1. Denote that $\overline{\boldsymbol{v}}_{-\max(k)} = \tau\boldsymbol{v}_{-\max(k)}$ *with* $\theta \leq \tau \leq 1$, *and* $\|\boldsymbol{A}^T\boldsymbol{A}\overline{\boldsymbol{v}}\|_\infty < (1-\tau)\rho$ *for all* $\overline{\boldsymbol{v}} \in \overline{\ker}(\boldsymbol{A}) \setminus \{0\}$. *Consider the signal recovery model* (1) *with* $\|\boldsymbol{A}^T\boldsymbol{z}\|_\infty < \epsilon$. *Let* $\widehat{\boldsymbol{x}}$ *be the minimizer of* (2) *with* $\mathcal{B} = \{\boldsymbol{z} : \|\boldsymbol{A}^T\boldsymbol{z}\|_\infty \leq \eta\}$ *for some* $\eta \geq \epsilon$. *If the condition* (13) *holds for some positive integers* $1 \leq a \leq k$, $b \geq 1$, *then*

$$\begin{aligned}\left\|\widehat{\boldsymbol{x}} - \boldsymbol{x}\right\|_2 \leq &\frac{\sqrt{2k/a}}{1-\delta_a^{\boldsymbol{A}} - C_{a,b,k,\tau}\delta_{a,b}^{\boldsymbol{A}}}(\epsilon + \eta + (1-\tau)\rho)\\ &+ 2\tau\chi\|\boldsymbol{x}_{-\max(k)}\|_1,\end{aligned}$$

where $C_{a,b,k,\tau}$ *and* χ *are defined in* (12) *and* (14), *respectively.*

Theorem 1 and Corollary 1 show an important relationship between the stably bounded recovery accuracy and our proposed scale factor τ defined in (8). Our RIP condition is better than [1]. Specifically, when $a = k$ and $b = k$, it comes to $\delta_k^{\mathbf{A}} + \max\{\tau, \sqrt{\tau}\}\delta_{k,k}^{\mathbf{A}} < 1$. Due to $\theta \leq \tau \leq 1$, it is further rewritten as $\delta_k^{\mathbf{A}} + \sqrt{\tau}\delta_{k,k}^{\mathbf{A}} < 1$, that is, $\delta_k^{\mathbf{A}} + \delta_{k,k}^{\mathbf{A}} < \gamma = 1 + (1-\sqrt{\tau})\delta_{k,k}^{\mathbf{A}}$. Given $\theta = 0.6$ and $\delta_{k,k}^{\mathbf{A}} = 0.7$, Fig. 3 plots the changed trend γ with different τ. Moreover, the more relaxed condition shown in the red line of Fig. 3 can be held when $\mathbf{A}\overline{\mathbf{v}} = 0$ or the errors can be tolerated.

Remark 2. Theorems 2.6 and 2.7 in the literature [1] are an our special case, that is, the scale factor τ is equivalent to one. When $\tau = 1$, it shows that there is no change in scale, that is, $\overline{\mathbf{v}}_{-\max(k)} = \mathbf{v}_{-\max(k)}$. Thus we have the same RIP condition. In fact, Theorem 1 reveals the condition (13) can ensure stably recovery of all k-sparse signals in the CS problem (2).

Remark 3. Following the bounded Gaussian noise case [1], [37], we can immediately yield the corresponding results. Let $\mathbf{z} \sim \mathcal{N}(0, \sigma^2\mathbf{I}_n/2)$ and $\mathbf{A} = [\mathbf{a}_{ij}]_{n\times p}$, where $\mathbf{a}_{ij} \overset{iid}{\sim} \mathcal{N}\left(0, \frac{\sigma^2}{2(1-\tau)^2\nu^2(p-k)}\right)$ and $\nu = \max\{\widehat{\max}\{\mathbf{v}_{-\max(k)}\} : \mathbf{v} \in \ker(\mathbf{A})\}$, and $\widehat{\max}\{\mathbf{v}_{-\max(k)}\}$ denotes the largest entry in absolute value of $\mathbf{v}_{-\max(k)}$. Then $\mathbf{A}\overline{\mathbf{v}} \sim \mathcal{N}(0, \sigma^2\mathbf{I}_n/2)$. Define $\mathcal{B}^{\ell_2} = \{\mathbf{A}\overline{\mathbf{v}} + \mathbf{z} : \|\mathbf{A}\overline{\mathbf{v}} + \mathbf{z}\|_2 \leq \sigma\sqrt{n + 2\sqrt{n\log n}}\}$ and $\mathcal{B}^{DS} = \{\mathbf{A}\overline{\mathbf{v}} + \mathbf{z} : \|\mathbf{A}^T(\mathbf{A}\overline{\mathbf{v}} + \mathbf{z})\|_2 \leq$

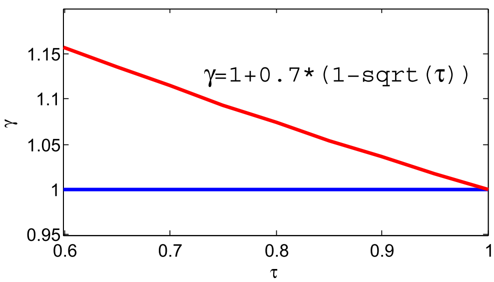

Fig. 3 More relaxed RIP condition on $\delta_k^{\mathbf{A}} + \delta_{k,k}^{\mathbf{A}} < \gamma = 1 + (1 - \sqrt{\tau})\delta_{k,k}^{\mathbf{A}}$. The green curve plots γ as a function of τ with $\delta_{k,k}^{\mathbf{A}} = 0.7$ and $\theta = 0.6 \le \tau \le 1$. When $\tau = 0.6$, the red line shows our upper boundary $\gamma = 1.157$, which is better than Cai's result $\gamma = 1$ [1] at blue line.

$\sigma\sqrt{2\log n}\}$. Based on the Theorem 1 and Corollary 1, the constraint minimizer $\widehat{\mathbf{x}}^{\ell_2}$ defined in (2) with $\mathcal{B}^{\ell_2}$ satisfies

$$\left\|\widehat{\mathbf{x}}^{\ell_2} - \mathbf{x}\right\|_2 \le \frac{2\sqrt{2}\sqrt{(1+\delta_a^{\mathbf{A}})\,k/a}}{1-\delta_a^{\mathbf{A}} - C_{a,b,k,\tau}\delta_{a,b}^{\mathbf{A}}}\sigma\sqrt{n + 2\sqrt{n\log n}} + 2\tau\chi\|\mathbf{x}_{-\max(k)}\|_1,$$

with probability at least $1-1/n$, and the constraint minimizer $\widehat{\mathbf{x}}^{DS}$ defined in (2) with $\mathcal{B}^{DS}$ satisfies

$$\left\|\widehat{\mathbf{x}}^{DS} - \mathbf{x}\right\|_2 \le \frac{2\sqrt{2}}{1-\delta_a^{\mathbf{A}} - C_{a,b,k,\tau}\delta_{a,b}^{\mathbf{A}}}\sigma\sqrt{2k\log p} + 2\tau\chi\|\mathbf{x}_{-\max(k)}\|_1,$$

with probability at least $1 - 1/\sqrt{\pi\log p}$, where χ is defined in (14).

Next, we turn to the noiseless case, that is, $\|\mathbf{z}\|_2 = 0$. Based on Theorem 1, we have the following corollary.

Corollary 2. Denote that $\bar{\boldsymbol{v}}_{-\max(k)} = \tau\boldsymbol{v}_{-\max(k)}$ *with* $\theta \le \tau \le 1$, *and* $\|\boldsymbol{A}\boldsymbol{v}_{-\max(k)}\|_2 < \rho$ *for all* $\boldsymbol{v} \in \ker(\boldsymbol{A})\backslash\{0\}$. *Consider the signal recovery model* (1) *with* $\|z\|_2 = 0$. *Let* $\widehat{\boldsymbol{x}}$ *be the minimizer of* (2) *with* $C = \{z : \|z\|_2 = 0\}$. *If the condition* (13) *holds for some positive integers* $1 \le a \le k$, $b \ge 1$, *then*

$$\left\|\widehat{\boldsymbol{x}} - \boldsymbol{x}\right\|_2 \le \frac{\sqrt{2\,(1+\delta_a^{A})\,k/a}}{1-\delta_a^{A} - C_{a,b,k,\tau}\delta_{a,b}^{A}}(1-\tau)\rho,$$

where $C_{a,b,k,\tau}$ *are defined in* (14).

Remark 4. Compared to Theorem 2.1 and 2.5 in literature [1], Corollary 1 shows that when $a = k$ and $b = k$, we can get the $(1-\sqrt{\tau})\delta_{k,k}^{\mathbf{A}}$ improvement in the noiseless case because we propose a scale factor to promote the RIP condition. But, it produces a recovery error $\left\|\mathbf{A}\bar{\mathbf{v}}\right\|_2 \le (1-\tau)\rho$. Fortunately, the more relaxed condition will lead to more extensive applications if the errors can be tolerated. This also shows the "no free lunch" theorem.

2.3 New RIP Conditions by Higher-Order RIC

In this subsection, we establish new RIP conditions on the high-order RIC [38] for sparse signal recovery by using the Polytope technique[†], since it represents a non-sparse vector by the sparse ones, which provides a bridge between general vectors and the RIP conditions [2].

Similar to Theorem 1 and Corollary 1, we consider two bounded noise setting: $\|\mathbf{z}\|_2 < \epsilon$ with $\|\mathbf{A}\bar{\mathbf{v}}\|_2 < (1-\tau)\rho$, and $\|\mathbf{A}^T\mathbf{z}\|_\infty < \epsilon$ with $\|\mathbf{A}^T\mathbf{A}\bar{\mathbf{v}}\|_\infty < (1-\tau)\rho$, respectively.

Theorem 2. Denote that $\bar{\boldsymbol{v}}_{-\max(k)} = \tau\boldsymbol{v}_{-\max(k)}$ *with* $\theta \le \tau \le 1$, *and* $\|\boldsymbol{A}\bar{\boldsymbol{v}}\|_2 < (1-\tau)\rho$ *for all* $\bar{\boldsymbol{v}} \in \overline{\ker}(\boldsymbol{A})\backslash\{0\}$. *Consider the signal recovery model* (1) *with* $\|z\|_2 \le \epsilon$. *Let* $\widehat{\boldsymbol{x}}$ *be the minimizer of* (2) *with* $\mathcal{B} = \{z : \|z\|_2 \le \eta\}$ *for some* $\eta \ge \epsilon$. *If*

$$\delta_{tk}^{A} < \left(1 + \frac{1-\tau^2}{\tau^2 - 1 + 2t}\right)\sqrt{\frac{t-1}{t}}, \tag{16}$$

for some $t \ge \frac{4}{3}$, *then*

$$\left\|\widehat{\boldsymbol{x}} - \boldsymbol{x}\right\|_2 \le \frac{\sqrt{2t(t-1)(1+\delta_{tk}^{A})}}{\Omega}(\epsilon + \eta + (1-\tau)\rho) + \frac{2\|\boldsymbol{x}_{-\max(k)}\|_1}{\sqrt{k}}\left(\frac{\sqrt{2}\tau^2\delta_{tk}^{A} + \tau\sqrt{\Omega\delta_{tk}^{A}}}{\Omega} + 1\right), \tag{17}$$

where $\Omega = \sqrt{t(t-1)} - (t + (\tau^2-1)/2)\delta_{tk}^{A}$.

Remark 5. Similar to the Remark 1, it still has some additional gain in the Eq. (17) compared to the upper bound of $\|\widehat{\beta} - \beta\|_2$ in the Eq. (10) in the literature [2] as Ω and τ are used to decrease the second terms of the right of the Eq. (17). If we have the condition

$$\|\mathbf{x}_{-\max(k)}\|_1 \ge \frac{\sqrt{kt(t-1)}(1+\delta_{tk}^{\mathbf{A}})}{2(1+\tau)\delta_{tk}^{\mathbf{A}} + \sqrt{2\Omega\delta_{tk}^{\mathbf{A}}}}\|\mathbf{x}_{-\max(k)}\|_2, \tag{18}$$

then it get a better recovery error $\left\|\widehat{\mathbf{x}} - \mathbf{x}\right\|_2 \le \|\widehat{\beta} - \beta\|_2$.

Corollary 3. Denote that $\bar{\boldsymbol{v}}_{-\max(k)} = \tau\boldsymbol{v}_{-\max(k)}$ *with* $\theta \le \tau \le 1$, *and* $\|\boldsymbol{A}\bar{\boldsymbol{v}}\|_\infty < (1-\tau)\rho$ *for all* $\bar{\boldsymbol{v}} \in \overline{\ker}(\boldsymbol{A})\backslash\{0\}$. *Consider the signal recovery model* (1) *with* $\|\boldsymbol{A}z\|_\infty < \epsilon$. *Let* $\widehat{\boldsymbol{x}}$ *be the minimizer of* (2) *with* $\mathcal{B} = \{z : \|\boldsymbol{A}z\|_\infty < \epsilon\}$ *for some* $\eta \ge \epsilon$. *If the condition* (16) *holds for some* $t \ge \frac{4}{3}$, *then*

$$\left\|\widehat{\boldsymbol{x}} - \boldsymbol{x}\right\|_2 \le \frac{\sqrt{2t^2(t-1)k}}{\Omega}(\epsilon + \eta + (1-\tau)\rho) + \frac{2\|\boldsymbol{x}_{-\max(k)}\|_1}{\sqrt{k}}\left(\frac{\sqrt{2}\tau^2\delta_{tk}^{A} + \tau\sqrt{\Omega\delta_{tk}^{A}}}{\Omega} + 1\right),$$

where Ω *is defined in* (17).

Theorem 2 and Corollary 3 also show that the condition (16) is sufficient for the exactly and bounded recovery of sparse signals via the constrained ℓ_1 minimization by employing the higher-order RIC. Figure 4 plots the changed trend of $\delta_{tk}^{\mathbf{A}}$ with different τ and t. We observe that the red surface is higher than the green surface. This

[†]It shows an elementary geometric fact: any point in a polytope can be represented as a convex combination of sparse vectors [2].

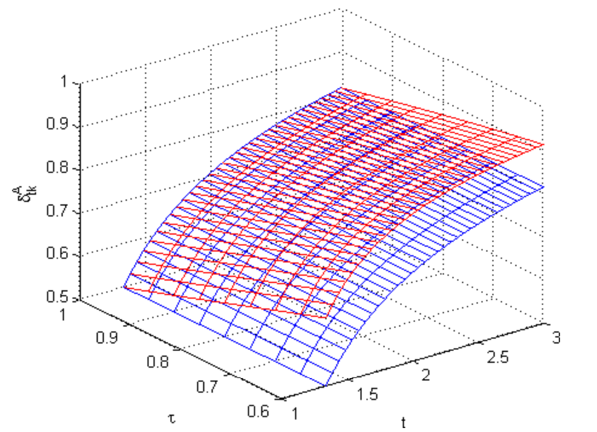

Fig. 4 More relaxed RIP condition on $\delta^{\mathbf{A}}_{tk} < \left(1 + (1-\tau^2)/(\tau^2 - 1 + 2t)\right)\sqrt{(t-1)/t}$. The red surface plots $\delta^{\mathbf{A}}_{tk}$ as a function of τ and t with $0.6 \leq \tau \leq 1$ and $t \geq 4/3$. This shows our upper boundary is better than the bule surface $\delta^{\mathbf{A}}_{tk} < \sqrt{(t-1)/t}$ [2].

shows our condition (16) is better than the previous condition $\delta^{\mathbf{A}}_{tk} < \sqrt{(t-1)/t}$ [2].

Remark 6. Theorem 2 and Corollary 3 also confirm the result of Theorem 2.2 [2], which shows that for some $t \geq 4/3$ and any ε, the condition $\delta^{\mathbf{A}}_{tk} < \sqrt{(t-1)/t} + \varepsilon$ is not sufficient to ensure exactly recovery of all k-sparse signals in the noiseless and noisy cases. Let $\varepsilon = \left((1-\tau^2)/(\tau^2 - 1 + 2t)\right)\sqrt{(t-1)/t}$. Clearly, ε is zero at $\tau = 1$, that is, there is no any improvement. In particular, we observe that when $t = 2$ and $\tau = 1$, $\delta^{\mathbf{A}}_{2k} < \rho = 1/\sqrt{2} = 0.7071$ is drawn in the dotted blue line in Fig. 4. Clearly, it is same to $\delta^{\mathbf{A}}_{2k} < 1/\sqrt{2} = 0.7071$ [2]. When $\theta \leq \tau < 1$, $\varepsilon > 0$, that means, we have a new upper bound of the RIP condition. This seems it contradicts to the Theorem 2.2 [2]. However, our more relaxed condition produce a recovery error $\left\|\mathbf{A}\bar{\mathbf{v}}\right\|_2 \leq (1-\tau)\rho$. Fortunately, the results will lead to the less number of measurements in many extensive applications when $\mathbf{A}\bar{\mathbf{v}} = 0$ or we can tolerate the errors.

Remark 7. Following the settings of Remark 3, we also consider the bounded Gaussian noise case [2], [37]. According to the Theorem 2 and Corollary 3, the constraint minimizer $\widehat{\mathbf{x}}^{\ell_2}$ defined in (2) with $\mathcal{B}^{\ell_2}$ satisfies

$$\left\|\widehat{\mathbf{x}}^{\ell_2} - \mathbf{x}\right\|_2 \leq \frac{2\sqrt{2t(t-1)(1+\delta^{\mathbf{A}}_{tk})}}{\Omega}\sigma\sqrt{n + 2\sqrt{n\log n}} + \frac{2\|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{k}}\left(\frac{\sqrt{2}\tau^2\delta^{\mathbf{A}}_{tk} + \tau\sqrt{\Omega\delta^{\mathbf{A}}_{tk}}}{\Omega} + 1\right),$$

with probability at least $1 - 1/n$, and the constraint minimizer $\widehat{\mathbf{x}}^{DS}$ defined in (2) with $\mathcal{B}^{DS}$ satisfies

$$\left\|\widehat{\mathbf{x}}^{DS} - \mathbf{x}\right\|_2 \leq \frac{4\sqrt{2t^2(t-1)}}{\Omega}\sigma\sqrt{k\log p} + \frac{2\|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{k}}\left(\frac{\sqrt{2}\tau^2\delta^{\mathbf{A}}_{tk} + \tau\sqrt{\Omega\delta^{\mathbf{A}}_{tk}}}{\Omega} + 1\right),$$

with probability at least $1 - 1/\sqrt{\pi\log p}$, where Ω is defined in (17).

Next, we turn to the noiseless case, that is, $\|\mathbf{z}\|_2 = 0$. Based on Theorem 2, we have the following corollary.

Corollary 4. Denote that $\bar{\boldsymbol{v}}_{-\max(k)} = \tau\boldsymbol{v}_{-\max(k)}$ with $\theta \leq \tau \leq 1$, and $\|\boldsymbol{A}\boldsymbol{v}_{-\max(k)}\|_2 < \rho$ for all $\boldsymbol{v} \in \ker(\boldsymbol{A})\setminus\{0\}$. Consider the signal recovery model (1) *with $\|z\|_2 = 0$. Let $\widehat{\boldsymbol{x}}$ be the minimizer of* (2) *with $C = \{z : \|z\|_2 = 0\}$. If the condition* (16) *holds for some $t \geq \frac{4}{3}$, then*

$$\left\|\widehat{\boldsymbol{x}} - \boldsymbol{x}\right\|_2 \leq \frac{\sqrt{2t(t-1)(1+\delta^{\boldsymbol{A}}_{tk})}}{\Omega}(1-\tau)\rho,$$

where Ω is defined in (17).

2.4 An Expansion of Sparsity Order by (5)

Since sparsity order k is crucial to choose the appropriate number of measurements [39], we will explore the distance $(1-\theta)\|\mathbf{v}_{-\max(k)}\|_1$ to expand the sparsity order k in this subsection. We denote a middle vector to increase k as

$$\mathbf{v}_{(m)} = \mathbf{v}_{-\max(k)} - \mathbf{v}_{-\max(k+m)}. \tag{19}$$

Now, the relationship between m and θ is given in the following lemma.

Lemma 1. If the following inequity is satisfied

$$\|\boldsymbol{v}_{(m)}\|_1 \leq \frac{1-\theta}{1+\theta}\|\boldsymbol{v}_{-\max(k+m)}\|_1, \tag{20}$$

where θ is given in (5), *then a new NSP property $\|\boldsymbol{v}_{\max(k+m)}\|_1 < \|\boldsymbol{v}_{-\max(k+m)}\|_1$ holds.*

Based on the Lemma 1 and the RIP conditions [1], [2], we have two corollaries to expand the sparsity order from k to $k+m$.

Corollary 5. Consider the signal recovery model (1) *with $\|z\|_2 \leq \epsilon$ or $\|\boldsymbol{A}z\|_\infty < \epsilon$. Let $\widehat{\boldsymbol{x}}$ be the minimizer of* (2) *with $\mathcal{B} = \{z : \|z\|_2 \leq \eta\}$ or $\mathcal{B} = \{z : \|\boldsymbol{A}z\|_\infty < \epsilon\}$ for some $\eta \geq \epsilon$. If there exists m such that the condition* (20) *and $\delta^{\boldsymbol{A}}_a + C_{a,b,k+m}\delta^{\boldsymbol{A}}_{a,b} < 1$ for some positive integers $1 \leq a \leq k+m$ and $b \geq 1$, then*

$$\left\|\widehat{\boldsymbol{x}} - \boldsymbol{x}\right\|_2 \leq \frac{\Gamma(\epsilon+\eta)}{1 - \delta^{\boldsymbol{A}}_a - C_{a,b,k+m}\delta^{\boldsymbol{A}}_{a,b}} + 2\chi\|\boldsymbol{x}_{-\max(k+m)}\|_1,$$

where $\Gamma = \sqrt{2(1+\delta^{\boldsymbol{A}}_a)(k+m)/a}$ or $\sqrt{2(k+m)}$, $\chi = \frac{\sqrt{2(k+m)}C_{a,b,k+m}\delta^{\boldsymbol{A}}_{a,b}}{\left(1-\delta^{\boldsymbol{A}}_a - C_{a,b,k+m}\delta^{\boldsymbol{A}}_{a,b}\right)(k+m-a)} + \frac{1}{\sqrt{k+m}}$ and $C_{a,b,k+m} = \max\left\{\frac{k+m-a}{\sqrt{ab}}, \sqrt{\frac{k+m-a}{a}}\right\}$.

Corollary 6. Consider the signal recovery model (1) *with $\|z\|_2 \leq \epsilon$ or $\|\boldsymbol{A}z\|_\infty < \epsilon$. Let $\widehat{\boldsymbol{x}}$ be the minimizer of* (2) *with $\mathcal{B} = \{z : \|z\|_2 \leq \eta\}$ or $\mathcal{B} = \{z : \|\boldsymbol{A}z\|_\infty < \epsilon\}$ for some $\eta \geq \epsilon$. If there exists m such that the condition* (20) *and $\delta^{\boldsymbol{A}}_{t(k+m)} < \sqrt{\frac{t-1}{t}}$ for some $t \geq \frac{4}{3}$, then*

$$\left\|\widehat{\boldsymbol{x}} - \boldsymbol{x}\right\|_2 \leq \frac{\sqrt{2t(t-1)\Theta}}{\overline{\Omega}}(\epsilon+\eta)$$

$$+\frac{2\|\boldsymbol{x}_{-\max(k+m)}\|_1}{\sqrt{k+m}}\left[\frac{\sqrt{2}\delta^{\mathbf{A}}_{t(k+m)}+\sqrt{\overline{\Omega}\delta^{\mathbf{A}}_{t(k+m)}}}{\overline{\Omega}}+1\right],$$

where $\Theta = (1+\delta^{\mathbf{A}}_{t(k+m)})$ *or* $t(k+m)$, *and* $\overline{\Omega} = \sqrt{t(t-1)} - t\delta^{\mathbf{A}}_{t(k+m)}$.

The proofs of Corollaries 5 and 6 are similar to [1], [2] by using $k+m$ instead of k. They show the important relationships between the bounded recovery accuracy and the expanded sparsity order m, which satisfies the Eq. (19). Our RIP conditions have higher sparsity order than [1], [2]. Specifically, when $a = k$ and $b = k$, they come to $\delta^{\mathbf{A}}_{k+m} + \delta^{\mathbf{A}}_{k+m,k+m} < 1$ in the Corollary 5 and $\delta^{\mathbf{A}}_{t(k+m)} < \sqrt{(t-1)/t}$ for some $t \geq 4/3$ in the Corollary 6.

3. Application

In this section, we will show the number of measurements n is changed with the scale factor τ under i.i.d. Gaussian or Bernoulli random matrices to verify our more relaxed RIP condition for efficiently sparse recovery.

Take for example i.i.d. Gaussian or Bernoulli random matrices. Theorem 5.2 in [40] shows that if a random sensing matrix satisfies $\mathbf{A} = [\mathbf{a}_{ij}]_{n\times p}$, where

$$\mathbf{a}_{ij} \overset{iid}{\sim} \mathcal{N}\left(0, \frac{1}{n}\right),$$

$$\mathbf{a}_{ij} \overset{iid}{\sim} \begin{cases} \frac{1}{\sqrt{n}}, & \text{with probability } \frac{1}{2} \\ -\frac{1}{\sqrt{n}}, & \text{with probability } \frac{1}{2} \end{cases},$$

$$\mathbf{a}_{ij} \overset{iid}{\sim} \begin{cases} \sqrt{\frac{3}{n}}, & \text{with probability } \frac{1}{6} \\ 0, & \text{with probability } \frac{2}{3} \\ -\sqrt{\frac{3}{n}}, & \text{with probability } \frac{1}{6} \end{cases},$$

then for any positive integer $s < n$ and $0 < \alpha < 1$, the RIC $\delta^{\mathbf{A}}_s$ of the matrix $\mathbf{A}$ satisfies

$$P\left(\delta^{\mathbf{A}}_s < \alpha\right) \geq 1 - 2\left(\frac{12ep}{s\alpha}\right)^s \exp\left(-n\left(\frac{\alpha^2}{16} - \frac{\alpha^3}{48}\right)\right). \tag{21}$$

Our conditions (13) and (16) with $\theta \leq \tau \leq 1$ are sufficient for the recovery of sparse signals in Theorems 1, 2, Corollaries 1 and 2 when the error $\Upsilon = \epsilon + \eta + (1-\tau)\rho$ can be tolerated.

Suppose $\theta = 0.5$ and for some given $0 < \varepsilon < 1$ one wishes the sensing matrix $\mathbf{A}$ to satisfy the RIP condition (13) with probability at least $1-\varepsilon$. When $a = k$ and $b = k$, it comes to $\delta^{\mathbf{A}}_k + \sqrt{\tau}\delta^{\mathbf{A}}_{k,k} < 1$, which is easily implied by $\delta^{\mathbf{A}}_k + \sqrt{\tau}\delta^{\mathbf{A}}_{2k} < 1$. Based on the Eq. (21) and the conditions $\delta^{\mathbf{A}}_k < 1 - 0.7\sqrt{\tau}$ and $\delta^{\mathbf{A}}_{2k} < 0.7$ in Theorem 1, we have

$$P\left(\delta^{\mathbf{A}}_k < 1 - 0.7\sqrt{\tau}\right) \geq 1 - 2\exp\Big(k\left(\log\left(12e/(1-0.7\sqrt{\tau}\right)\right) + \log(p/k)\Big) - n\left((1-0.7\sqrt{\tau})^2/16 - (1-0.7\sqrt{\tau})^3/48\right)\Big),$$

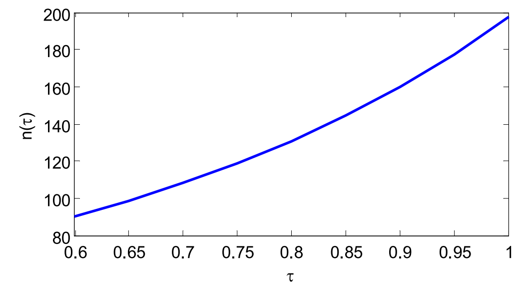

Fig. 5 Plot of $n(\tau)$ as a function of τ.

and

$$P\left(\delta^{\mathbf{A}}_{2k} < 0.7\right) \geq 1 - 2\exp\left(2k\left(\log(60e/7) + \log(p/k)\right) - n\left(49/1600 - 343/48000\right)\right).$$

Note that for given k and p, $n \geq n_1$ with

$$n_1(\tau) = \frac{1}{\left((1-0.7\sqrt{\tau})^2/16 - (1-0.7\sqrt{\tau})^3/48\right)} \left[k(\log(k/p) + \log(12/(1-0.7\sqrt{\tau})) + 1) - \log(\varepsilon/4)\right],$$

guarantees $\delta^{\mathbf{A}}_k < 1 - 0.7\sqrt{\tau}$ with probability at least $1-\varepsilon/2$, and $n \geq n_2$ with

$$n_2 = 85.2\left[k(\log(k/p) + 3.15) - \log(\varepsilon/4)/2\right],$$

ensures $\delta^{\mathbf{A}}_{2k} < 0.7$ with probability at least $1-\varepsilon/2$. Figure 5 plots the function $n(\tau)$. Thus, $\delta^{\mathbf{A}}_k + \sqrt{\tau}\delta^{\mathbf{A}}_{k,k} < 1$ satisfies with probability at least $1-\varepsilon/2$ if the number of measurements n satisfies $n \geq \max\{n_1(\tau), n_2\}$ with $0.6 < \tau \leq 1$. Therefore, for large k, p and $0.6 < c \leq 1$, the required number of measurements n holds

$$n \geq \max\left\{\left((1-0.7\sqrt{\tau})^2/16 - (1-0.7\sqrt{\tau})^3/48\right), 85.2\right\},$$

to ensure $\delta^{\mathbf{A}}_k + \sqrt{\tau}\delta^{\mathbf{A}}_{k,k} < 1$. When τ is set to 0.6, $n \geq \max\{90.1, 85.2\} = 90.1$. Therefore, compared to 115.4 [1], the required number of measurements to ensure $\delta^{\mathbf{A}}_k + \sqrt{\tau}\delta^{\mathbf{A}}_{k,k} < 1$ is less than 78.1% (90.1/115.4).

Based on Corollary 2, since $\tau = 0.6$, $\sqrt{2}(1-\tau) \approx 0.5657$ and $\sqrt{\tau} \approx 0.7746$ and $\rho = \sqrt{1+\delta^{\mathbf{A}}_k}\|\alpha_{-\max(k)}\|_2$, there has an upper bound for the error of recovery, where

$$\left\|\widehat{\mathbf{x}} - \mathbf{x}\right\|_2 \leq \frac{0.5657\left(1+\delta^{\mathbf{A}}_k\right)}{1 - \delta^{\mathbf{A}}_k - 0.7746\delta^{\mathbf{A}}_{k,k}}\|\alpha_{-\max(k)}\|_2. \tag{22}$$

Based on the Eq. (21) and the condition (16) in Theorem 2, we have for $t \geq 4/3$,

$$P\left(\delta^{\mathbf{A}}_{tk} < \left(1 + \frac{1-\tau^2}{\tau^2 - 1 + 2t}\right)\sqrt{\frac{t-1}{t}}\right) \geq 1 - 2\exp\left(tk\left(\log\left(\frac{6e(\tau^2-1+2t)}{t\sqrt{t(t-1)}}\right) + \log\left(\frac{p}{k}\right)\right) - n\left(\frac{t(t-1)}{4(\tau^2-1+2t)^2} - \frac{(t(t-1))^{3/2}}{6(\tau^2-1+2t)^3}\right)\right).$$

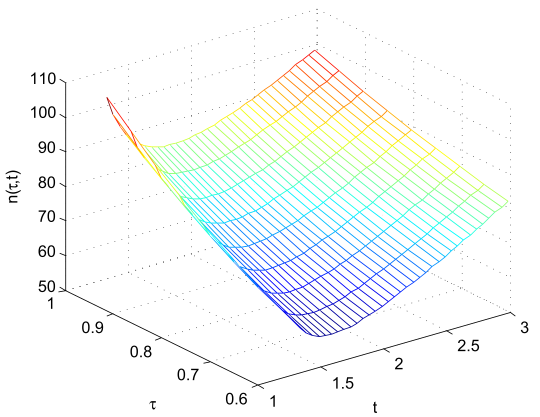

Fig. 6 Plot of $n(\tau, t)$ as a function of τ and t.

It is easy to see when p, k, and $p/k \to \infty$, the lower bound of n to ensure $\delta^{\mathbf{A}}_{tk} < \left(1 + \frac{1-\tau^2}{\tau^2-1+2t}\right)\sqrt{\frac{t-1}{t}}$ to hold with high probability is $n \geq k\log(p/k)n(t,\tau)$, where

$$n(t,\tau) \triangleq t/\left(\frac{t(t-1)}{4(\tau^2-1+2t)^2} - \frac{(t(t-1))^{3/2}}{6(\tau^2-1+2t)^3}\right).$$

Figure 6 plots the function $n(\tau, t)$. When $t = 1.58$ and $\tau = 0.6$, $n(\tau, t)$ has minimum 58.7, which is less than 70.5% (58.7/83.2) [2].

Based on Corollary 4, we set $\tau = 0.6$, $t = 1.58$, $\sqrt{t(t-1)} \approx 0.9573$, $\sqrt{t(t-1)} \approx 0.9573$, $(1-\tau)\sqrt{2t(t-1)} \approx 0.5415$ and $(t + (\tau^2 - 1)/2) \approx 0.4700$, it has an upper bound for the error of recovery, where

$$\left\|\widehat{\mathbf{x}} - \mathbf{x}\right\|_2 \leq \frac{0.5415\sqrt{(1+\delta^{\mathbf{A}}_{1.58k})(1+\delta^{\mathbf{A}}_{k})}}{0.9573 - 0.4700\delta^{\mathbf{A}}_{1.58k}}\|\alpha_{-\max(k)}\|_2. \tag{23}$$

In addition, our results can also be used for certain theoretical analysis in signal processing. Similar to [1], [11], we consider a finite window of a band-limited signal $h(t)$ as

$$h(t) = \Phi(\alpha) = \sum_{j=1}^{p} \alpha_j \phi_j(t)$$

where $\phi_j(t) = e^{i2\pi jt}$ (j is the imaginary unit) are the Fourier basis functions, and $\alpha = [\alpha_1, \cdots, \alpha_p]$ is k sparse. Assume that the measurements $y_1, \cdots, y_n$ can be represented by

$$y_i = \langle \varphi_i(t), h(t)\rangle = \sum_{l=1}^{p} \alpha_l \langle \varphi_i(t), \phi_l(t)\rangle \triangleq \sum_{l=1}^{p} r_{il}\alpha_l.$$

Then it can be written as $y = R\alpha$. As discussed above, when $R = [r_{il}]$ with r_{il} i.i.d. Gaussian or Bernoulli, as discussed above, the measurement matrix satisfies the RIP condition of order k or $2k$ with high probability provided that $n \gtrsim \kappa_0 k\log(p/k)$, where $\kappa_0 \geq 90.1$ or 58.7. Although there is a lower number of the measurements, it also has upper bounds for the error of recovery, which are similar to the Eqs. (22) and (23).

4. Conclusion

In this paper, we proposed a factor τ to scale down $\mathbf{v}_{-\max(k)}$ into $\overline{\mathbf{v}}_{-\max(k)}$. By using this scale factor to transform the distance between $\|\mathbf{v}_{\max(k)}\|_1$ and $\|\mathbf{v}_{-\max(k)}\|$ into the scale NSP, we established two more relaxed RIP conditions to guarantee the bounded approximation recovery of all k-sparse signals in the noiseless and noisy cases. In fact, although the scale factor τ created an error $(1-\tau)\rho$, we obtained an improvement to breakthrough the upper bound of the previous RIP conditions. Our theoretical results led to more extensive applications when the error is tolerated. In addition, we also explored the distance to expand the sparsity order k. An application verified our proposed RIP conditions needed a smaller number of measurements.

5. Proofs

Following the popular proof procedure [1], [2], we prove new RIP conditions by extending the NSP to the scale NSP.

5.1 Proof of Proposition 1

The proof of Proposition 1 is basically the same to Theorem 2.1 [1] and Theorem 1.1 [2]. In particular, we only need to use the inequality $\|\mathbf{v}_{\max(k)}\|_1 \leq \theta\|\mathbf{v}_{-\max(k)}\|_1$ instead of $\|\mathbf{v}_{\max(k)}\|_1 \leq \|\mathbf{v}_{-\max(k)}\|_1$ in the Theorem 2.1 [1] and the Theorem 1.1 [2].

5.2 Proof of Theorem 1

We set $\mathbf{v} = \widehat{\mathbf{x}} - \mathbf{x}$. We usually use the contradiction method to prove the following well-known inequality [1]: for all $\mathbf{v} \in \ker(\mathbf{A}) \setminus \{0\}$,

$$\|\mathbf{v}_{-\max(k)}\|_1 \leq \|\mathbf{v}_{\max(k)}\|_1 + 2\|\mathbf{x}_{-\max(k)}\|_1. \tag{24}$$

We introduce the scale factor τ to describe $\overline{\mathbf{v}}_{-\max(k)} = \tau\mathbf{v}_{-\max(k)}$ with $\theta \leq \tau \leq 1$ in (8), and denote $\overline{\mathbf{v}} = \mathbf{v}_{\max(k)} + \overline{\mathbf{v}}_{-\max(k)}$. In fact $\overline{\mathbf{v}}_{\max(k)} = \mathbf{v}_{\max(k)}$. To replace $\mathbf{v}_{-\max(k)}$ by $\overline{\mathbf{v}}_{-\max(k)}$, the inequality (24) is be equivalent to the following inequality, for all $\overline{\mathbf{v}} \in \overline{\ker}(\mathbf{A}) \setminus \{0\}$,

$$\|\overline{\mathbf{v}}_{-\max(k)}\|_1 \leq \tau\|\mathbf{v}_{\max(k)}\|_1 + 2\tau\|\mathbf{x}_{-\max(k)}\|_1. \tag{25}$$

Following the proof of Theorem 2.2 [1], thus, we only need to prove (25). By the boundedness of $\mathbf{z}$ and the definition of the feasible set for $\widehat{\mathbf{x}}$, it has

$$\begin{aligned}
\|\mathbf{A}\overline{\mathbf{v}}\|_2 =& \|\mathbf{A}\left(\mathbf{v}_{\max(k)} + \tau\mathbf{v}_{-\max(k)}\right)\|_2 \\
=& \|\mathbf{A}\left(\mathbf{v}_{\max(k)} + \mathbf{v}_{-\max(k)} - \mathbf{v}_{-\max(k)} + \tau\mathbf{v}_{-\max(k)}\right)\|_2 \\
=& \|\mathbf{A}\left(\mathbf{v} - (1-\tau)\mathbf{v}_{-\max(k)}\right)\|_2 \\
\leq& \|\mathbf{A}\mathbf{v} - (1-\tau)\mathbf{A}\mathbf{v}_{-\max(k)}\|_2 \\
\leq& \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 + \|\mathbf{A}\widehat{\mathbf{x}} - \mathbf{y}\|_2 + (1-\tau)\|\mathbf{A}\mathbf{v}_{-\max(k)}\|_2 \\
\leq& \epsilon + \eta + (1-\tau)\rho.
\end{aligned} \tag{26}$$

On the other hand, let $\mathbf{v} = \sum_{i=1}^{p} d_i \mathbf{u}_i$, where $\{d_i\}_{i=1}^{p}$ is a non-negative and non-increasing sequence and $\{\mathbf{u}_i\}_{i=1}^{p}$ are indicator vectors with different supports in $\mathbb{R}^p$. Clearly, $\mathbf{v}_{\max(k)} = \sum_{i=1}^{k} d_i \mathbf{u}_i$ and $\mathbf{v}_{-\max(k)} = \sum_{i=k+1}^{p} d_i \mathbf{u}_i$. According to $\overline{\mathbf{v}}_{-\max(k)} = \tau \mathbf{v}_{-\max(k)}$ in (8), we have $\overline{\mathbf{v}} = \mathbf{v}_{\max(k)} + \overline{\mathbf{v}}_{-\max(k)} = \sum_{i=k+1}^{p} d_i \mathbf{u}_i + \sum_{i=k+1}^{p} \overline{d}_i \mathbf{u}_i$, where $\overline{d}_i = \tau d_i$. Then we obtain

$$\sum_{i=k+1}^{p} \overline{d}_i \leq \tau \sum_{i=1}^{k} d_i + 2\tau \|\mathbf{x}_{-\max(k)}\|_1. \tag{27}$$

Hence, when $1 \leq a \leq k$, we have $\overline{\mathbf{v}}_{\max(a)} = \mathbf{v}_{\max(a)}$,

$$\begin{aligned} &\|\overline{\mathbf{v}}_{-\max(a)}\|_\infty \\ &= d_{a+1} \leq \frac{\sum_{i=1}^{a} d_i}{a} = \frac{\|\mathbf{v}_{\max(a)}\|_1}{a} \\ &\leq \frac{\|\mathbf{v}_{\max(a)}\|_1}{a} + \frac{2\tau \|\mathbf{x}_{-\max(k)}\|_1}{(1+\tau)k - a}, \end{aligned} \tag{28}$$

and

$$\begin{aligned} &\|\overline{\mathbf{v}}_{-\max(a)}\|_1 \\ &= \sum_{i=a+1}^{k} d_i + \sum_{i=k+1}^{p} \overline{d}_i \\ &\leq \frac{k-a}{k} \sum_{i=1}^{a} d_i + \tau \sum_{i=1}^{k} d_i + 2\tau \|\mathbf{x}_{-\max(k)}\|_1 \\ &\leq \frac{k-a}{a} \sum_{i=1}^{a} d_i + \tau \frac{k}{a} \sum_{i=1}^{a} d_i + 2\tau \|\mathbf{x}_{-\max(k)}\|_1 \\ &= \frac{(1+\tau)k - a}{a} \|\mathbf{v}_{\max(a)}\|_1 + 2\tau \|\mathbf{x}_{-\max(k)}\|_1. \end{aligned} \tag{29}$$

Then we set $\lambda = \frac{\|\mathbf{v}_{\max(a)}\|_1}{a} + \frac{2\tau\|\mathbf{x}_{-\max(k)}\|_1}{(1+\tau)k-a}$, $k_1 = a$ and $k_2 = (1+\tau)k - a$. By the Lemma 5.1 in [1], it follows that

$$\begin{aligned} &\left|\left\langle \mathbf{A}\left(\mathbf{v}_{\max(a)}\right), \mathbf{A}\left(\overline{\mathbf{v}}_{-\max(a)}\right)\right\rangle\right| \\ &\leq \delta^{\mathbf{A}}_{a,(1+\tau)k-a} \sqrt{(1+\tau)k - a} \left\|\mathbf{v}_{\max(a)}\right\|_2 \\ &\quad \times \left(\frac{\|\mathbf{v}_{\max(a)}\|_1}{a} + \frac{2\tau \|\mathbf{x}_{-\max(k)}\|_1}{(1+\tau)k - a} \right). \end{aligned} \tag{30}$$

Because $\left\|\mathbf{A}\left(\overline{\mathbf{v}}\right)\right\|_2 < \epsilon + \eta + (1-\tau)\rho$ when considering the noise, we have

$$\begin{aligned} &\left|\left\langle \mathbf{A}\left(\overline{\mathbf{v}}\right), \mathbf{A}\left(\mathbf{v}_{\max(a)}\right)\right\rangle\right| \\ &\leq \left\|\mathbf{A}\left(\overline{\mathbf{v}}\right)\right\|_2 \left\|\mathbf{A}\left(\mathbf{v}_{\max(a)}\right)\right\|_2 \\ &\leq (\epsilon + \eta + (1-\tau)\rho) \sqrt{1 + \delta^{\mathbf{A}}_a} \left\|\mathbf{v}_{\max(a)}\right\|_2. \end{aligned} \tag{31}$$

On the other hand, we have

$$\begin{aligned} &(\epsilon + \eta + (1-\tau)\rho) \sqrt{1 + \delta^{\mathbf{A}}_a} \left\|\mathbf{v}_{\max(a)}\right\|_2 \\ &\geq \left|\left\langle \mathbf{A}\left(\overline{\mathbf{v}}\right), \mathbf{A}\left(\mathbf{v}_{\max(a)}\right)\right\rangle\right| \\ &\geq \left|\left\langle \mathbf{A}\left(\mathbf{v}_{\max(a)}\right), \mathbf{A}\left(\mathbf{v}_{\max(a)}\right)\right\rangle\right| - \\ &\quad \left|\left\langle \mathbf{A}\left(\overline{\mathbf{v}}_{-\max(a)}\right), \mathbf{A}\left(\mathbf{v}_{\max(a)}\right)\right\rangle\right| \\ &\geq \left\|\mathbf{A}\left(\mathbf{v}_{\max(a)}\right)\right\|_2^2 - \left|\left\langle \mathbf{A}\left(\overline{\mathbf{v}}_{-\max(a)}\right), \mathbf{A}\left(\mathbf{v}_{\max(a)}\right)\right\rangle\right| \\ &\geq \left(1 - \delta^{\mathbf{A}}_a\right) \left\|\mathbf{v}_{\max(a)}\right\|_2^2 - \\ &\quad \delta^{\mathbf{A}}_{a,(1+\tau)k-a} \sqrt{(1+\tau)k - a} \left\|\mathbf{v}_{\max(a)}\right\|_2 \times \\ &\quad \left(\frac{\|\mathbf{v}_{\max(a)}\|_1}{a} + \frac{2\tau \|\mathbf{x}_{-\max(k)}\|_1}{(1+\tau)k - a} \right) \\ &\geq \left(1 - \delta^{\mathbf{A}}_a - \delta^{\mathbf{A}}_{a,(1+\tau)k-a} \sqrt{\frac{(1+\tau)k - a}{a}}\right) \left\|\mathbf{v}_{\max(a)}\right\|_2^2 \\ &\quad - \delta^{\mathbf{A}}_{a,(1+\tau)k-a} \left\|\mathbf{v}_{\max(a)}\right\|_2 \frac{2\tau \|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{(1+\tau)k - a}}. \end{aligned} \tag{32}$$

Hence,

$$\begin{aligned} &\left\|\mathbf{v}_{\max(a)}\right\|_2 \\ &\leq \frac{(\epsilon + \eta + (1-\tau)\rho) \sqrt{1 + \delta^{\mathbf{A}}_a}}{1 - \delta^{\mathbf{A}}_a - \delta^{\mathbf{A}}_{a,(1+\tau)k-a} \sqrt{\frac{(1+\tau)k-a}{a}}} + \\ &\quad \frac{\delta^{\mathbf{A}}_{a,(1+\tau)k-a}}{1 - \delta^{\mathbf{A}}_a - \delta^{\mathbf{A}}_{a,(1+\tau)k-a} \sqrt{\frac{(1+\tau)k-a}{a}}} \frac{2\tau \|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{(1+\tau)k - a}}. \end{aligned} \tag{33}$$

It follows from the Lemma 5.4 in [1] that

$$\begin{aligned} \delta^{\mathbf{A}}_{a,(1+\tau)k-a} &\leq \sqrt{\frac{(1+\tau)k - a}{\min\{b, (1+\tau)k - a\}}} \delta^{\mathbf{A}}_{a,\min\{b,(1+\tau)k-a\}} \\ &\leq \max\left\{ \sqrt{\frac{(1+\tau)k - a}{b}}, 1 \right\} \delta^{\mathbf{A}}_{a,b} \\ &= \sqrt{\frac{a}{(1+\tau)k - a}} C_{a,b,k,\tau} \delta^{\mathbf{A}}_{a,b}. \end{aligned} \tag{34}$$

Then

$$\begin{aligned} \left\|\mathbf{v}_{\max(a)}\right\|_2 &\leq \frac{(\epsilon + \eta + (1-\tau)\rho) \sqrt{1 + \delta^{\mathbf{A}}_a}}{1 - \delta^{\mathbf{A}}_a - C_{a,b,k,\tau} \delta^{\mathbf{A}}_{a,b}} \\ &\quad + \frac{\delta^{\mathbf{A}}_{a,(1+\tau)k-a}}{1 - \delta^{\mathbf{A}}_a - C_{a,b,k,\tau} \delta^{\mathbf{A}}_{a,b}} \frac{2\tau \|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{(1+\tau)k - a}} \\ &\leq \frac{\mathcal{Y}}{1 - \delta^{\mathbf{A}}_a - C_{a,b,k,\tau} \delta^{\mathbf{A}}_{a,b}}, \end{aligned} \tag{35}$$

where $\mathcal{Y} = (\epsilon + \eta + (1-\tau)\rho) \sqrt{1 + \delta^{\mathbf{A}}_a} + \frac{2\tau \delta^{\mathbf{A}}_{a,(1+\tau)k-a} \|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{(1+\tau)k-a}}$.

Since $0 < \theta \leq \tau \leq 1$, $1 \leq a \leq k$, $\|\mathbf{x}_{-\max(k)}\|_1 > 0$, $\rho > 0$, $\epsilon > 0$, $\eta > 0$, $\delta^{\mathbf{A}}_a \geq 0$ and $\delta^{\mathbf{A}}_{a,(1+\tau)k-a} \geq 0$, it has $\mathcal{Y} > 0$. If it holds the condition $\delta^{\mathbf{A}}_a + C_{a,b,k,\tau} \delta^{\mathbf{A}}_{a,b} < 1$, then the inequality (35) implies $\left\|\mathbf{v}_{\max(a)}\right\|_2 < 0$, which contradicts a basic fact that $\left\|\mathbf{v}_{\max(a)}\right\|_2 \geq 0$. Clearly, the inequality (35) contradicts the condition that $\delta^{\mathbf{A}}_a + C_{a,b,k,\tau} \delta^{\mathbf{A}}_{a,b} < 1$. Therefore, we obtain this more relaxed RIP condition. Next, we prove the approximation errors.

By applying the Lemma 5.3 [22] with $\alpha = 2$ and $\lambda = 2\|\mathbf{x}_{-\max(k)}\|_1$, we obtain

$$
\begin{aligned}
\|\mathbf{v}\|_2 =& \sqrt{\sum_{i=1}^{k} d_i^2 + \sum_{i=k+1}^{p} d_i^2} \\
\leq& \sqrt{\sum_{i=1}^{k} d_i^2 + k\left(\sqrt{\frac{\sum_{i=1}^{k} d_i^2}{k}} + \frac{2\tau\|\mathbf{x}_{-\max(k)}\|_1}{k}\right)^2} \\
\leq& \sqrt{\sum_{i=1}^{k} d_i^2 + \left(\sqrt{\sum_{i=1}^{k} d_i^2} + \frac{2\tau\|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{k}}\right)^2} \\
\leq& \sqrt{2\sum_{i=1}^{k} d_i^2} + \frac{2\tau\|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{k}} \\
\leq& \sqrt{\frac{2k}{a}}\sqrt{\sum_{i=1}^{a} d_i^2} + \frac{2\tau\|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{k}} \\
\leq& \frac{(\epsilon+\eta+(1-\tau)\rho)\sqrt{\frac{2(1+\delta_a^{\mathbf{A}})k}{a}}}{1-\delta_a^{\mathbf{A}}-C_{a,b,k,\tau}\delta_{a,b}^{\mathbf{A}}} + \\
& 2c\left(\frac{\sqrt{\frac{2k}{a}}\delta_{a,(1+\tau)k-a}^{\mathbf{A}}}{1-\delta_a^{\mathbf{A}}-C_{a,b,k,\tau}\delta_{a,b}^{\mathbf{A}}}\frac{1}{\sqrt{(1+\tau)k-a}} + \frac{1}{\sqrt{k}}\right) \\
& \|\mathbf{x}_{-\max(k)}\|_1. \qquad (36)
\end{aligned}
$$

Finally, it holds (14). The proof of Theorem 1 is complete. □

5.3 Proof of Corollary 1

The proof of Corollary 1 is basically the same to Theorem 1. Particularly, we only need to use the inequalities

$$
\begin{aligned}
\|\mathbf{A}^T\mathbf{A}\overline{\mathbf{v}}\|_\infty \leq& \|\mathbf{A}^T(\mathbf{A}\mathbf{v}+(1-\tau)\mathbf{A}\mathbf{v}_{-\max(k)})\|_\infty \\
\leq& \|\mathbf{A}^T(\mathbf{A}\mathbf{x}-\mathbf{y})\|_\infty + \|\mathbf{A}^T(\mathbf{A}\widehat{\mathbf{x}}-\mathbf{y})\|_\infty \\
& + (1-\tau)\|\mathbf{A}^T\mathbf{A}\mathbf{v}_{-\max(k)}\|_\infty \\
\leq& \epsilon+\eta+(1-\tau)\rho, \qquad (37)
\end{aligned}
$$

and

$$
\begin{aligned}
& \left|\langle \mathbf{A}(\overline{\mathbf{v}}), \mathbf{A}(\mathbf{v}_{\max(a)})\rangle\right| \\
\leq& \left\|\mathbf{v}_{\max(a)}^T\mathbf{A}^T\mathbf{A}(\overline{\mathbf{v}})\right\|_\infty \\
\leq& (\epsilon+\eta+(1-\tau)\rho)\sqrt{a}\left\|\mathbf{v}_{\max(a)}\right\|_2 \qquad (38)
\end{aligned}
$$

instead of (26) and (31).

5.4 Proof of Corollary 2

The proof of Corollary 2 is almost the same to Theorem 1. Particularly, we only need to use the inequalities $\|\overline{\mathbf{v}}_{-\max(k)}\|_1 \leq \tau\|\mathbf{v}_{\max(k)}\|_1$ instead of (25).

5.5 Proof of Theorem 2

Based on the proof procedure of the Theorem 1, we also use the contradiction method to prove the inequality (25), that is, $\|\overline{\mathbf{v}}_{-\max(k)}\|_1 \leq \tau\|\mathbf{v}_{\max(k)}\|_1 + 2\tau\|\mathbf{x}_{-\max(k)}\|_1$, for all $\mathbf{v} \in \overline{\ker}(\mathbf{A}) \setminus \{0\}$.

Similar to the proof method [2], we suppose tk is an integer and set $\mathbf{v} = \widehat{\mathbf{x}} - \mathbf{x}$. By the boundedness of $\mathbf{z}$, it also has (26), that is, $\|\mathbf{A}\overline{\mathbf{v}}\|_2 \leq \epsilon + \eta + (1-\tau)\rho$.

We set $\alpha = \frac{\|\mathbf{v}_{\max(k)}\|_1 + 2\|\mathbf{x}_{-\max(k)}\|_1}{k}$, and divide $\overline{\mathbf{v}}_{-\max(k)}$ into two parts, $\overline{\mathbf{v}}_{-\max(k)} = \overline{\mathbf{v}}^{(1)} + \overline{\mathbf{v}}^{(2)}$, where

$$\overline{\mathbf{v}}^{(1)} = \overline{\mathbf{v}}_{-\max(k)} \cdot \mathbf{1}_{\{i \| \overline{\mathbf{v}}_{-\max(k)}(i)| > \frac{\tau\alpha}{t-1}\}}, \qquad (39)$$

$$\overline{\mathbf{v}}^{(2)} = \overline{\mathbf{v}}_{-\max(k)} \cdot \mathbf{1}_{\{i \| \overline{\mathbf{v}}_{-\max(k)}(i)| \leq \frac{\tau\alpha}{t-1}\}}. \qquad (40)$$

Denote $|\mathrm{supp}(\overline{\mathbf{v}}^{(1)})| = \|\overline{\mathbf{v}}^{(1)}\|_0 = m$. Since $\|\overline{\mathbf{v}}_{-\max(k)}\|_1 \leq \tau\alpha k$ and $|\overline{\mathbf{v}}^{(1)}(i)| \geq \frac{\tau\alpha}{t-1}$ for any $\overline{\mathbf{v}}^{(1)}(i) \neq 0$, we obtain

$$
\begin{aligned}
\tau\alpha k \geq \|\overline{\mathbf{v}}^{(1)}\|_1 =& \sum_{i\in \mathrm{supp}(\overline{\mathbf{v}}^{(1)})} |\overline{\mathbf{v}}^{(1)}(i)| \\
\geq& \sum_{i\in \mathrm{supp}(\overline{\mathbf{v}}^{(1)})} \frac{\tau\alpha}{t-1} = \frac{m\tau\alpha}{t-1}. \qquad (41)
\end{aligned}
$$

Then $m \leq k(t-1)$. On the other hand, we have

$$
\begin{aligned}
\|\overline{\mathbf{v}}^{(2)}\|_1 =& \|\overline{\mathbf{v}}_{-\max(k)}\|_1 - \|\overline{\mathbf{v}}^{(1)}\|_1 \\
\leq& (k(t-1)-m)\frac{\tau\alpha}{t-1}, \qquad (42)
\end{aligned}
$$

$$\|\overline{\mathbf{v}}^{(2)}\|_\infty \leq \frac{\tau\alpha}{t-1}. \qquad (43)$$

Besides, $\|\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\|_0 = k + m \leq tk$, it has

$$
\begin{aligned}
& \left|\left\langle \mathbf{A}\left(\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right), \mathbf{A}\overline{\mathbf{v}}\right\rangle\right| \\
\leq& \left\|\mathbf{A}\left(\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right)\right\|_2 \left\|\mathbf{A}\overline{\mathbf{v}}\right\|_2 \\
\leq& \sqrt{1+\delta_k^{\mathbf{A}}}(\epsilon+\eta+(1-\tau)\rho)\left\|\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right\|_2. \qquad (44)
\end{aligned}
$$

By applying the Lemma 1.1 [2] with $s = k(t-1) - m$, $\overline{\mathbf{v}}^{(2)}$ can be expressed as a convex combination of sparse vectors $\overline{\mathbf{v}}^{(2)} = \sum_{i=1}^{N} \lambda_i \overline{\mathbf{u}}_i$, where s-sparse $\overline{\mathbf{u}}_i$ satisfies $\mathrm{supp}(\overline{\mathbf{u}}_i) \subseteq \mathrm{supp}\left(\overline{\mathbf{v}}^{(2)}\right)$, $\|\overline{\mathbf{u}}_i\|_1 = \|\overline{\mathbf{v}}^{(2)}\|_1$, $\|\overline{\mathbf{u}}_i\|_\infty \leq \frac{\tau\alpha}{t-1}$. Set $x = \|\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\|_2$ and $P = \frac{2\|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{k}}$. Hence,

$$
\begin{aligned}
\|\overline{\mathbf{u}}_i\|_2 \leq& \sqrt{\|\overline{\mathbf{u}}_i\|_0}\|\overline{\mathbf{u}}_i\|_\infty \leq \sqrt{k(t-1)-m}\frac{\tau\alpha}{t-1} \leq \sqrt{\frac{k}{t-1}}\tau\alpha \\
\leq& \tau\left(\frac{\left\|\mathbf{v}_{\max(k)}\right\|_2}{\sqrt{t-1}} + \frac{2\|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{k(t-1)}}\right) \\
\leq& \tau\left(\frac{\left\|\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right\|_2}{\sqrt{t-1}} + \frac{2\|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{k(t-1)}}\right) \\
=& \frac{\tau(x+P)}{\sqrt{t-1}}. \qquad (45)
\end{aligned}
$$

For some $1 \geq \mu \geq 0, c \geq 0, 1 \geq \tau \geq \theta$, we denote that $\beta_i = \mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)} + \mu\overline{\mathbf{u}}_i$, and have

$$\sum_{j=1}^{N}\lambda_j\beta_j - c\beta_i = \mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)} + \mu\overline{\mathbf{v}}^{(2)} - c\beta_i$$
$$= (1-\mu-c)\left(\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right) - c\mu\overline{\mathbf{u}}_i + \mu\overline{\mathbf{v}}. \quad (46)$$

Then $\sum_{j=1}^{N}\lambda_j\beta_j - c\beta_i - \mu\overline{\mathbf{v}}$ are are tk-sparse vectors as $\mathbf{v}_{\max(k)}, \overline{\mathbf{v}}^{(1)}, \overline{\mathbf{u}}_i$ are $k, m, (k(t-1)-m)$-sparse respectively. We check the ℓ_2 identity [2] as follows:

$$\sum_{i=1}^{N}\lambda_i\left\|\mathbf{A}\left(\sum_{j=1}^{N}\lambda_j\beta_j - c\beta_i\right)\right\|_2^2 + (1-2c)\sum_{1\le i,j\le N}\lambda_i\lambda_j\|\mathbf{A}(\beta_i-\beta_j)\|_2^2$$
$$= \sum_{i=1}^{N}\lambda_i(1-c)^2\|\mathbf{A}\beta_i\|_2^2. \quad (47)$$

We set $c = \frac{1}{2}$ and compute the left hand side of (47) as follow:

$$\sum_{i=1}^{N}\lambda_i\left\|\mathbf{A}\left(\sum_{j=1}^{N}\lambda_j\beta_j - c\beta_i\right)\right\|_2^2$$
$$= \sum_{i=1}^{N}\lambda_i\left\|\mathbf{A}\left((\frac{1}{2}-\mu)\left(\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right) - \frac{\mu}{2}\overline{\mathbf{u}}_i\right)\right\|_2^2 +$$
$$\mu(1-\mu)\left\langle\mathbf{A}\left(\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right), \mathbf{A}\overline{\mathbf{v}}\right\rangle \quad (48)$$

It also has that β_i, $(\frac{1}{2}-\mu)\left(\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right) - \frac{\mu}{2}\mathbf{u}_i$ are tk-sparse vectors because $\mathbf{v}_{\max(k)}, \overline{\mathbf{v}}^{(1)}, \mathbf{u}_i$ are $k, m, (k(t-1)-m)$-sparse vectors, respectively. By substituting (44), (45), and (48) into (47), we set $\Upsilon = \epsilon + \eta + (1-\tau)\rho$, $\mu = \sqrt{t(t-1)} - (t-1)$, and have

$$0 = \sum_{i=1}^{N}\lambda_i\left\|\mathbf{A}\left((\frac{1}{2}-\mu)\left(\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right) - \frac{\mu}{2}\overline{\mathbf{u}}_i\right)\right\|_2^2 +$$
$$\mu(1-\mu)\left\langle\mathbf{A}\left(\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right), \mathbf{A}\overline{\mathbf{v}}\right\rangle$$
$$- \sum_{i=1}^{N}\lambda_i(1-p)^2\left\|\mathbf{A}\left(\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)} + \mu\overline{\mathbf{u}}_i\right)\right\|_2^2$$
$$\le\left(1+\delta_{tk}^{\mathbf{A}}\right)\cdot\sum_{i=1}^{N}\lambda_i\left((\frac{1}{2}-\mu)\|\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\|_2^2 + \frac{\mu^2}{4}\|\overline{\mathbf{u}}_i\|_2^2\right)$$
$$+ \mu(1-\mu)\sqrt{1+\delta_{tk}^{\mathbf{A}}}\Upsilon\left\|\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right\|_2$$
$$- \left(1-\delta_{tk}^{\mathbf{A}}\right)\cdot\sum_{i=1}^{N}\frac{\lambda_i}{4}\left(\|\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\|_2^2 + \mu^2\|\overline{\mathbf{u}}_i\|_2^2\right)$$
$$= \sum_{i=1}^{N}\lambda_i\left[\left(\left(1+\delta_{tk}^{\mathbf{A}}\right)(\frac{1}{2}-\mu) - \left(1+\delta_{tk}^{\mathbf{A}}\right)\frac{1}{4}\right)\right.$$
$$\left.\left\|\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right\|_2^2 + \frac{1}{2}\delta_{tk}^{\mathbf{A}}\mu^2\|\overline{\mathbf{u}}_i\|_2^2\right]$$
$$+ \mu(1-\mu)\sqrt{1+\delta_{tk}^{\mathbf{A}}}\Upsilon\left\|\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right\|_2$$
$$\le\left[(\mu^2-\mu) + \delta_{tk}^{\mathbf{A}}\left(\frac{1}{2} - \mu + \mu^2 + \frac{\tau^2}{2(t-1)}\mu^2\right)\right]x^2$$
$$+ \left[\mu(1-\mu)\sqrt{1+\delta_{tk}^{\mathbf{A}}}\Upsilon + \frac{\delta_{tk}^{\mathbf{A}}\mu^2\tau^2 P}{t-1}\right]x + \frac{\delta_{tk}^{\mathbf{A}}\mu^2\tau^2 P^2}{2(t-1)}$$
$$= -t\left((2t-1) - 2\sqrt{t(t-1)}\right)\left(\sqrt{\frac{t-1}{t}} - \delta_{tk}^{\mathbf{A}}\frac{2t+\tau^2-1}{2t}\right)x^2$$
$$+ \left(\mu^2\sqrt{\frac{t}{t-1}}\sqrt{1+\delta_{tk}^{\mathbf{A}}}\Upsilon + \frac{\delta_{tk}^{\mathbf{A}}\mu^2\tau^2 P}{t-1}\right)x + \frac{\delta_{tk}^{\mathbf{A}}\mu^2\tau^2 P^2}{2(t-1)}$$
$$= \frac{\mu^2}{t-1}\left[-t\left(\sqrt{\frac{t-1}{t}} - \delta_{tk}^{\mathbf{A}}\frac{2t+\tau^2-1}{2t}\right)x^2 + \left(\sqrt{t(t-1)(1+\delta_{tk}^{\mathbf{A}})}\Upsilon + \delta_{tk}^{\mathbf{A}}\tau^2 P\right)x + \frac{\delta_{tk}^{\mathbf{A}}\tau^2 P^2}{2}\right]$$

Clearly, it is an second-order inequality for x. By solving above inequality we denote $\Omega = t\left(\sqrt{\frac{t-1}{t}} - \delta_{tk}^{\mathbf{A}}\frac{2t+\tau^2-1}{2t}\right)$, and have

$$x \le \frac{1}{2\Omega}\left\{\left(\sqrt{t(t-1)(1+\delta_{tk}^{\mathbf{A}})}\Upsilon + \delta_{tk}^{\mathbf{A}}\tau^2 P\right) + \left[\left(\sqrt{t(t-1)(1+\delta_{tk}^{\mathbf{A}})}\Upsilon + \delta_{tk}^{\mathbf{A}}\tau^2 P\right)^2 + 2\Omega\delta_{tk}^{\mathbf{A}}\tau^2 P^2\right]^{1/2}\right\}$$
$$\le \frac{\sqrt{t(t-1)(1+\delta_{tk}^{\mathbf{A}})}}{\Omega}\Upsilon + \frac{2\tau^2\delta_{tk}^{\mathbf{A}} + \tau\sqrt{2\Omega\delta_{tk}^{\mathbf{A}}}}{2\Omega}P$$

This contradicts the following fact that

$$\delta_{tk}^{\mathbf{A}} \le \left(1 + \frac{1-\tau^2}{\tau^2-1+2t}\right)\sqrt{\frac{t-1}{t}},$$
$$\|\mathbf{u}_i\|_2 \le \sqrt{\frac{k}{t-1}}\alpha \le \frac{\|\mathbf{v}_{\max(k)}\|_2}{\sqrt{t-1}}$$
$$\le \frac{\|\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\|_2}{\sqrt{t-1}}.$$

Finally, based on $\|\mathbf{v}_{-\max(k)}\|_1 \le \|\mathbf{v}_{\max(k)}\|_1 + P\sqrt{k}$, we have $\|\mathbf{v}_{-\max(k)}\|_2 \le \|\mathbf{v}_{\max(k)}\|_2 + P$. Following the (36), we get

$$\|\mathbf{v}\|_2 = \sqrt{\|\mathbf{v}_{\max(k)}\|_2^2 + \|\mathbf{v}_{-\max(k)}\|_2^2}$$
$$\le \sqrt{\|\mathbf{v}_{\max(k)}\|_2^2 + (\|\mathbf{v}_{\max(k)}\|_2 + P)^2}$$
$$\le \sqrt{2\|\mathbf{v}_{\max(k)}\|_2^2} + P \le \sqrt{2}x + P$$
$$\le \frac{\sqrt{2t(t-1)(1+\delta_{tk}^{\mathbf{A}})}}{\Omega}\Upsilon + \frac{2\|\mathbf{x}_{-\max(k)}\|_1}{\sqrt{k}}\left[\frac{\sqrt{2}\tau^2\delta_{tk}^{\mathbf{A}} + \tau\sqrt{\Omega\delta_{tk}^{\mathbf{A}}} + \Omega}{\Omega}\right]$$

When tk is not an integer, note $t' = \lceil tk\rceil/k$, then $t' > t$, $t'k$ is an integer. It can be deduced to the same result.

The proof of Theorem 2 is complete. □

5.6 Proof of Corollary 3

The proof of Corollary 3 is basically the same to Theorem 2. Particularly, we only need to use the inequalities (37), that is, $\|\mathbf{A}^T\mathbf{A}\overline{\mathbf{v}}\|_\infty \leq \epsilon + \eta + (1-\tau)\rho$, and

$$\begin{aligned}
&\left|\left\langle \mathbf{A}\left(\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right), \mathbf{A}\overline{\mathbf{v}}\right\rangle\right| \\
&\leq \left|\left\langle \mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}, \mathbf{A}^T\mathbf{A}\overline{\mathbf{v}}\right\rangle\right| \\
&\leq \left\|\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right\|_1 (\epsilon + \eta + (1-\tau)\rho) \\
&\leq \sqrt{tk}\Upsilon \left\|\mathbf{v}_{\max(k)} + \overline{\mathbf{v}}^{(1)}\right\|_2 .
\end{aligned}$$

instead of $\|\mathbf{A}\overline{\mathbf{v}}\|_2 \leq \epsilon + \eta + (1-\tau)\rho$ and (31).

5.7 Proof of Corollary 4

The proof of Corollary 4 is almost the same to Theorem 2. In particular, we only need to use the inequalities $\|\overline{\mathbf{v}}_{-\max(k)}\|_1 \leq \tau\|\mathbf{v}_{\max(k)}\|_1 + 2\tau\|\mathbf{x}_{-\max(k)}\|_1$ and $\alpha = \frac{\|\mathbf{v}_{\max(k)}\|_1 + 2\|\mathbf{x}_{-\max(k)}\|_1}{k}$ instead of $\|\overline{\mathbf{v}}_{-\max(k)}\|_1 \leq \tau\|\mathbf{v}_{\max(k)}\|_1$ and $\alpha = \frac{\|\mathbf{v}_{\max(k)}\|_1}{k}$.

5.8 Proof of Lemma 1

If $\|\mathbf{v}_{(m)}\|_1 \leq \frac{1-\theta}{1+\theta}\|\mathbf{v}_{-\max(k+m)}\|_1$, then it holds that

$$\begin{aligned}
&\|\mathbf{v}_{(m)}\|_1 \leq \frac{1-\theta}{1+\theta}\|\mathbf{v}_{-\max(k+m)}\|_1 \\
&\Rightarrow 2\|\mathbf{v}_{(m)}\|_1 - (1-\theta)\|\mathbf{v}_{(m)}\|_1 \leq (1-\theta)\|\mathbf{v}_{-\max(k+m)}\|_1 \\
&\Rightarrow 2\|\mathbf{v}_{(m)}\|_1 \leq (1-\theta)\left(\|\mathbf{v}_{(m)}\|_1 + \|\mathbf{v}_{-\max(k+m)}\|_1\right) \\
&\Rightarrow \|\mathbf{v}_{(m)}\|_1 \leq (1-\theta)\|\mathbf{v}_{-\max(k)}\|_1/2
\end{aligned} \tag{49}$$

There is a fact that

$$\begin{aligned}
&\|\mathbf{v}_{\max(k)}\|_1 < \theta\|\mathbf{v}_{-\max(k)}\|_1 \leq \|\mathbf{v}_{-\max(k)}\|_1 \\
&\Rightarrow \|\mathbf{v}_{\max(k)}\|_1 + (1-\theta)\|\mathbf{v}_{-\max(k)}\|_1/2 \\
&\quad < \|\mathbf{v}_{-\max(k)}\|_1 - (1-\theta)\|\mathbf{v}_{-\max(k)}\|_1/2 \\
&\quad \text{By using (49)} \\
&\Rightarrow \|\mathbf{v}_{\max(k)}\|_1 + \|\mathbf{v}_{(m)}\|_1 < \|\mathbf{v}_{-\max(k)}\|_1 - \|\mathbf{v}_{(m)}\|_1 \\
&\Rightarrow \|\mathbf{v}_{\max(k+m)}\|_1 < \|\mathbf{v}_{-\max(k+m)}\|_1 .
\end{aligned}$$

References

[1] T. Cai and A. Zhang, "Compressed sensing and affine rank minimization under restricted isometry," IEEE Trans. Signal Process., vol.61, no.7, pp.3279–3290, 2013.
[2] T. Cai and A. Zhang, "Sparse representation of a polytope and recovery of sparse signals and low-rank matrices," IEEE Trans. Inf. Theory, vol.60, no.1, pp.122–132, 2014.
[3] D. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," IEEE Trans. Inf. Theory, vol.47, no.7, pp.2845–2862, 2001.
[4] E. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," IEEE Trans. Inf. Theory, vol.52, no.12, pp.5406–5425, 2006.
[5] D. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory, vol.52, no.4, pp.1289–1306, 2006.
[6] J. Romberg, "Imaging via compressive sampling," IEEE Signal Process. Mag., vol.25, no.2, pp.14–20, 2008.
[7] S. Bahmani and J. Romberg, "Efficient compressive phase retrieval with constrained sensing vectors," NIPS, 2015.
[8] E. Candes and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," Ann. Statist., vol.35, no.6, pp.2313–2351, 2007.
[9] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," ICML, 2010.
[10] M.A. Davenport, P.T. Boufounos, M.B. Wakin, and R.G. Baraniuk, "Signal processing with compressive measurements," IEEE J. Sel. Topics Signal Process., vol.4, no.2, pp.445–460, 2010.
[11] M. Davenport, J. Laska, J. Treichler, and R. Baraniuk, "The pros and cons of compressive sensing for wideband signal acquisition: Noise folding versus dynamic range," IEEE Trans. Signal Process., vol.60, no.9, pp.4628–4642, 2012.
[12] S.K. Sharma, E. Lagunas, S. Chatzinotas, and B. Ottersten, "Application of compressive sensing in cognitive radio communications: A survey," IEEE Commun. Surveys Tuts., vol.18, no.3, pp.1838–1860, 2016.
[13] J. Li, H. Liu, and Y. Fu, "Predictive coding machine for compressed sensing and image denoising," Proc. AAAI Conf. on Artif. Intell., pp.3506–3513, 2018.
[14] E. Candes and T. Tao, "Decoding by linear programming," IEEE Trans. Inf. Theory, vol.51, no.12, pp.4203–4215, 2005.
[15] B. Olshausen and D. Field, "Sparse coding with an overcomplete basis set: A strategy employed by V1?," Vision Research, vol.37, no.23, pp.3311–3325, 1997.
[16] E. Candes and T. Tao, "The Dantzig selector: Statistical estimation when is much larger than n," Ann. Statist., vol.35, no.6, pp.2313–2351, 2007.
[17] A. Bruckstein, D. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," SIAM Review, vol.51, no.1, pp.34–81, 2009.
[18] A. Tillmann and M. Pfetsch, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," IEEE Trans. Inf. Theory, vol.60, no.2, pp.1248–1259, 2014.
[19] J. Wang and B. Shim, "Exact recovery of sparse signals using orthogonal matching pursuit: How many iterations do we need?," IEEE Trans. Signal Process., vol.64, no.16, pp.4194–4202, 2016.
[20] J. Wang, S. Kwon, P. Li, and B. Shim, "Recovery of sparse signals via generalized orthogonal matching pursuit: A new analysis," IEEE Trans. Signal Process., vol.64, no.4, pp.1076–1089, 2016.
[21] T. Cai, L. Wang, and G. Xu, "Shifting inequality and recovery of sparse signals," IEEE Trans. Signal Process., vol.58, no.3, pp.1300–1308, 2010.
[22] T. Cai and A. Zhang, "Sharp RIP bound for sparse signal and low-rank matrix recovery," Applied and Computational Harmonic Analysis, vol.35, no.1, pp.74–93, 2013.
[23] S. Dirksen, G. Lecue, and H. Rauhut, "On the gap between restricted isometry properties and sparse recovery conditions," IEEE Trans. Inf. Theory, vol.64, no.8, pp.5478–5487, 2018.
[24] J. Cahill, X. Chen, and R. Wang, "The gap between the null space property and the restricted isometry property," Linear Algebra and its Applications, vol.501, pp.363–375, 2016.
[25] M. Ahsen, N. Challapalli, and M. Vidyasagar, "Two new approaches to compressed sensing exhibiting both robust sparse recovery and the grouping effect," J. Machine Learning Research, vol.18, pp.1–24, 2017.
[26] M. Elad and A. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," IEEE Trans. Inf. Theory, vol.48, no.9, pp.2558–2567, 2002.
[27] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," IEEE Trans. Inf. Theory, vol.49, no.12, pp.3320–3325, 2003.
[28] L.G.A. d'Aspremont, "Testing the nullspace property using semidef-

inite programming," Math. Program., Series B, vol.127, no.1, pp.123–144, 2011.
[29] Q. Sun, "Sparse approximation property and stable recovery of sparse signals from noisy measurements," IEEE Trans. Signal Process., vol.59, no.10, pp.5086–5090, 2011.
[30] M. Cho and W. Xu, "New algorithms for verifying the null space conditions in compressed sensing," Asilomar Conference on Signals, Systems and Computers, pp.1038–1042, 2013.
[31] A. Bourrier, M. Davies, T. Peleg, P. Pérez, and R. Gribonval, "Fundamental performance limits for ideal decoders in high-dimensional linear inverse problems," IEEE Trans. Inf. Theory, vol.60, no.12, pp.7928–7946, 2014.
[32] H. Mansour and R. Saab, "Weighted one-norm minimization with inaccurate support estimates: Sharp analysis via the null-space property," ICASSP, pp.3317–3321, 2015.
[33] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best k-term approximation," J. Am. Math. Soc., vol.22, no.1, pp.211–231, 2009.
[34] M. Davies and R. Gribonval, "On Lp minimisation, instance optimality, and restricted isometry constants for sparse approximation," Proc. Sampling Theory and Applications (SampTA), 2009.
[35] M. Stojnic, W. Xu, and B. Hassibi, "Compressed sensing-probabilistic analysis of a null-space characterization," IEEE Internat. Conf. on Acoustics, Speech and Signal Processing, ICASSP, 2008.
[36] D. Ba, B. Babadi, P. Purdon, and E. Brown, "Exact and stable recovery of sequences of signals with sparse increments via differential ℓ_1-minimization," NIPS, 2012.
[37] T. Cai, G. Xu, and J. Zhang, "On recovery of sparse signals via ℓ_1 minimization," IEEE Trans. Inf. Theory, vol.55, no.7, pp.3388–3397, 2009.
[38] J. Blanchard and A. Thompson, "On support sizes of restricted isometry constants," Applied and Computational Harmonic Analysis, vol.29, no.3, pp.382–390, 2010.
[39] S.K. Sharma, S. Chatzinotas, and B. Ottersten, "Compressive sparsity order estimation for wideband cognitive radio receiver," IEEE Trans. Signal Process., vol.62, no.19, pp.4984–4996, 2014.
[40] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," Constr. Approx., vol.28, pp.253–263, 2008.

Haiyang Zou received the B.S. and M.S. degrees in Computer Application and Technology from Heilongjiang University and China West Normal University in 2002 and 2009 respectively. He is currently a assistant professor with the School of Computer Science, China West Normal University, Nanchong, China. His research interests include database application, and signal processing. He has published more than 20 papers in journals.

Wengang Zhao received the B.S. and M.S. degrees in Communication Engineering from Chengdu University of Information Technology in 2006 and Computer Application and Technology from China West Normal University in 2009, respectively. He is currently a MBA advisor with the School of Economics and Management, Southwest University of Science and Technology, Mianyang, China. His research interests include economic information technology, artificial intelligence, and signal processing.