Journal of Computational Mathematics
Vol.41, No.6, 2023, 1137–1170. http://www.global-sci.org/jcm
doi:10.4208/jcm.2207-m2022-0058

STABLE AND ROBUST RECOVERY OF APPROXIMATELY k-SPARSE SIGNALS WITH PARTIAL SUPPORT INFORMATION IN NOISE SETTINGS VIA WEIGHTED ℓ_p $(0 < p \leq 1)$ MINIMIZATION*

Biao Du and Anhua Wan[1)]
School of Mathematics, Sun Yat-sen University, Guangzhou 510275, China
Email: dubiao@mail2.sysu.edu.cn, wananhua@mail.sysu.edu.cn

Abstract

In the existing work, the recovery of strictly k-sparse signals with partial support information was derived in the ℓ_2 bounded noise setting. In this paper, the recovery of approximately k-sparse signals with partial support information in two noise settings is investigated via weighted ℓ_p $(0 < p \leq 1)$ minimization method. The restricted isometry constant (RIC) condition $\delta_{tk} < \frac{1}{p\eta^{\frac{2}{p}-1}+1}$ on the measurement matrix for some $t \in [1+\frac{2-p}{2+p}\sigma, 2]$ is proved to be sufficient to guarantee the stable and robust recovery of signals under sparsity defect in noisy cases. Herein, $\sigma \in [0,1]$ is a parameter related to the prior support information of the original signal, and $\eta \geq 0$ is determined by p, t and σ. The new results not only improve the recent work in [17], but also include the optimal results by weighted ℓ_1 minimization or by standard ℓ_p minimization as special cases.

Mathematics subject classification: 94A12, 94A15.
Key words: Signal recovery, weighted ℓ_p minimization, Approximately k-sparse signal, Noise setting, Reconstruction error bound, Restricted isometry property.

1. Introduction

As a data acquisition paradigm, compressed sensing has been a very active research area and has abundant applications [2,15,22]. Compressed sensing is particularly promising not only in applications such as hyperspectral imaging where taking measurements is costly, but also in applications such as medical and seismic imaging where the ambient dimension of the signal is very large [18].

In standard compressed sensing theory, one observes

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

where $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$ is an unknown sparse signal, $y \in \mathbb{R}^m$ is the observed signal, $A \in \mathbb{R}^{m\times n}$ is a measurement matrix with $m \ll n$, and $z \in \mathbb{R}^m$ denotes the noise in the measurement. One of the central goals of compressed sensing is to recover the original high-dimensional signal x based on the measurement matrix and the observed signal.

For signal recovery, the following noise settings

$$\mathcal{B}^{\ell_2}(\epsilon) := \left\{ z \in \mathbb{R}^m : \|z\|_2 \leq \epsilon \right\} \tag{1.2}$$

* Received March 12, 2022 / Revised version received May 23, 2022 / Accepted July 6, 2022 / Published online September 27, 2022 /
1) Corresponding author

and

$$\mathcal{B}^{DS}(\epsilon) := \left\{ z \in \mathbb{R}^m : \left\| A^T z \right\|_\infty \le \epsilon \right\} \tag{1.3}$$

are of particular interest. Herein, $\epsilon \ge 0$ denotes some known margin. The ℓ_2 bounded noise setting (1.2) was considered for example in [14], and the *DS* noise setting (1.3) was motivated by the *Dantzig Selector* procedure in [5].

Denote the support of $x = (x_1, x_2, \dots, x_n)^T$ as $\text{supp}(x) = \{i : x_i \ne 0\}$. x is called k-sparse if the number of nonzero components in x is k at most, i.e., $\|x\|_0 = |\text{supp}(x)| \le k$.

The constrained ℓ_p $(0 < p \le 1)$ minimization method estimates the signal x by

$$\hat{x} = \arg\min_{x \in \mathbb{R}^n} \left\{ \|x\|_p^p : y - Ax \in \mathcal{B} \right\}, \tag{1.4}$$

where

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

is the ℓ_p (quasi-)norm of x and $\mathcal{B} \subseteq \mathbb{R}^m$ denotes some noise structure [21, 24, 28]. When in particular $p = 1$, the ℓ_p minimization model (1.4) becomes the standard ℓ_1 minimization model [1–3].

The following restricted isometry property (RIP) is a commonly used framework for sparse recovery.

Definition 1.1 ([4]). *Suppose $A \in \mathbb{R}^{m \times n}$ is a measurement matrix, k is an integer and $1 \le k \le n$. For the measurement matrix A, the restricted isometry constant (RIC) of order k is defined as the smallest number $\delta_k \ge 0$ such that for all k-sparse vectors $x \in \mathbb{R}^n$,*

$$(1 - \delta_k) \|x\|_2^2 \le \|Ax\|_2^2 \le (1 + \delta_k) \|x\|_2^2 . \tag{1.5}$$

More generally, when k is not an integer, δ_k is defined as $\delta_{\lceil k \rceil}$, where $\lceil \cdot \rceil$ denotes the ceiling function [2].

In many practical applications, the original signal is not exactly k-sparse. As a consequence, the stable recovery of approximately sparse signals in noisy settings is of significant interest, and has been investigated under different sufficient RIC conditions by ℓ_p minimization model (1.4) [23, 26–28]. When $n \le 4k$, under the assumption $p \in (0, \frac{3+2\sqrt{2}}{2}(1 - \delta_{2k})]$ for $\delta_{2k} \in (0, 1)$, Wen, Li and Zhu [26] proved the stable recovery of approximately k-sparse signals in the ℓ_2 bounded noise case. For $p \in (0, 1]$, Zhang and Li [28] derived the sharp condition

$$\delta_{2k} < \frac{\eta}{2 - p - \eta} \tag{1.6}$$

for the stable recovery of exactly k-sparse signals in noisy cases, where $\eta \in \left(1 - p, 1 - \frac{p}{2}\right)$ is the unique positive solution of the equation

$$\frac{p}{2} \eta^{\frac{2}{p}} + \eta - 1 + \frac{p}{2} = 0. \tag{1.7}$$

In our previous work [8, 24], general condition

$$\delta_{tk} < \delta^*(p, t) := \frac{\eta}{\dfrac{2 - p}{t - 1} - \eta} \tag{1.8}$$

for some $t \in (1, 2]$ is proved to be sufficient for signal recovery in noiseless case and noisy cases, and meanwhile the original signals are not restricted to be exactly k-sparse. Herein, $\eta \in [\frac{\sqrt{1+2p-p^2}-1}{p}, \frac{1-(t-\sqrt{t^2-t})p}{t-1}]$ is the unique positive solution of the following equation:

$$\frac{p}{2}\eta^{\frac{2}{p}} + \eta - \frac{2-p}{2(t-1)} = 0. \tag{1.9}$$

Furthermore, the sharpness of the RIC condition (1.8) was proved for all $t \in [\frac{4}{2+p}, 2]$ in [25].

The constrained ℓ_p minimization method (1.4) for signal recovery is non-adaptive, since no prior information on the signal being measured is used therein. In many applications, it may be possible to draw an estimate of the support of the original signal or its largest components. For instance, video and audio signals exhibit correlation over temporal frames which can be employed to estimate a portion of the support using the previously decoded frames [16, 20]. If some prior information on the support of the original signal is provided, incorporating the prior support information into the optimization problem for signal recovery is of significant use. Therefore, the constrained ℓ_p minimization (1.4) can be modified as the following constrained weighted ℓ_p $(0 < p \le 1)$ minimization [18]

$$\hat{x} = \arg\min_{x\in\mathbb{R}^n} \left\{ \sum_{i=1}^{n} \mathrm{w}_i^p |x_i|^p : \, y - Ax \in \mathcal{B} \right\}. \tag{1.10}$$

Herein, $\mathrm{w} = (\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_n)^T$ denotes a weight vector with $\mathrm{w}_i \in [0, 1]$, $i = 1, \ldots, n$. For any given prior support estimate $\tilde{T} \subseteq \{1, \ldots, n\}$ of the original signal $x \in \mathbb{R}^n$, the "indicative" weight vector $\mathrm{w} = (\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_n)^T$ can be assigned as $\mathrm{w}_i = 1 - (1 - w)\chi_{\tilde{T}}(i)$, where $w \in [0, 1]$ and

$$\chi_{\tilde{T}}(i) = \begin{cases} 1, & i \in \tilde{T}, \\ 0, & i \notin \tilde{T} \end{cases}$$

denotes the characteristic function, i.e.,

$$\mathrm{w}_i = \begin{cases} w, & i \in \tilde{T}, \\ 1, & i \notin \tilde{T} \end{cases}, \qquad i = 1, 2, \ldots, n. \tag{1.11}$$

Therefore, $\mathrm{w} \in \{w, 1\}^n$. The improved method (1.10) for signal recovery is adaptive by exploiting some known information on the support of the original signal. As to (1.11), the main idea of the choice of the weights w_i is that the components x_i of the original signal x which are "expected" to be large in absolute value are penalized less in the weighted objective function [16].

We employ Hadamard product of the two vectors w and x, i.e.,

$$\mathrm{w} \circ x = (\mathrm{w}_1 x_1, \mathrm{w}_2 x_2, \ldots, \mathrm{w}_n x_n)^T. \tag{1.12}$$

Then we have $\sum_{i=1}^n \mathrm{w}_i^p |x_i|^p = \|\mathrm{w} \circ x\|_p^p$, and thus the constrained weighted ℓ_p $(0 < p \le 1)$ minimization (1.10) can be reformulated as

$$\hat{x} = \arg\min_{x\in\mathbb{R}^n} \left\{ \|\mathrm{w} \circ x\|_p^p : \, y - Ax \in \mathcal{B} \right\}. \tag{1.13}$$

Specifically, in the ℓ_2 bounded noise setting, (1.13) becomes

$$\hat{x} = \arg\min_{x\in\mathbb{R}^n} \left\{ \|\mathrm{w} \circ x\|_p^p : \, \|y - Ax\|_2 \le \varepsilon \right\}. \tag{1.14}$$

In the *DS* noise setting, (1.13) becomes

$$\hat{x} = \arg\min_{x\in\mathbb{R}^n} \Big\{ \|\mathrm{w} \circ x\|_p^p : \ \big\|A^T (y - Ax)\big\|_\infty \le \varepsilon \Big\}. \tag{1.15}$$

When in particular $p = 1$, (1.13) reduces to be the weighted ℓ_1 minimization

$$\hat{x} = \arg\min_{x\in\mathbb{R}^n} \Big\{ \|\mathrm{w} \circ x\|_1 : \ y - Ax \in \mathcal{B} \Big\}. \tag{1.16}$$

The recovery of sparse signals was analyzed by the weighted ℓ_1 minimization [9, 16]. The weighted $\ell_p\,(0 < p \le 1)$ minimization in ℓ_2 bounded noise setting has been studied in the literature [17, 18]. Ge, Chen and Ng [17] studied the recovery for exactly k-sparse signals with partial support information via the weighted ℓ_p minimization model of the following form:

$$\hat{x} = \arg\min_{x\in\mathbb{R}^n} \left\{ \sum_{i=1}^n \mathrm{w}_i |x_i|^p : \ y - Ax \in \mathcal{B} \right\} \tag{1.17}$$

by generalizing the condition (1.8) in [8] to

$$\delta_{tk} < \hat{\delta}(p, \Theta) := \begin{cases} \dfrac{z_0}{(2-p)\Theta - z_0}, & 0 < \Theta < \dfrac{2+p}{2-p}, \\ 1, & \Theta = 0 \end{cases} \tag{1.18}$$

for some $t \in (d + \frac{2-p}{2+p}\zeta, 2d]$, where $z_0 \in ((1-p)\Theta, \min\{1, \frac{2-p}{2}\Theta\})$ is the unique positive solution of the following equation:

$$\frac{p}{2} z^{\frac{2}{p}} + z - \frac{2-p}{2}\Theta = 0. \tag{1.19}$$

Therein, $T = \mathrm{supp}(x)$, $|T| = k$, $|\tilde{T}| = \rho k$ with $\rho \ge 0$, $|T \cap \tilde{T}| = \alpha |\tilde{T}|$ with $\alpha \in [0, 1]$ and $\alpha\rho \le 1$,

$$\zeta = \Big[w + (1-w)\,(1 + \rho - 2\alpha\rho)^{\frac{2-p}{2}} \Big]^{\frac{2}{2-p}} \quad \text{and} \quad \Theta = \frac{\zeta}{t-d}$$

with

$$d = \begin{cases} 1, & w = 1, \\ 1 + (\max\{0, 1 - 2\alpha\})\,\rho, & 0 \le w < 1. \end{cases} \tag{1.20}$$

Nevertheless, the original signal $x \in \mathbb{R}^n$ is restricted to be strictly k-sparse.

In this paper, we propose to remove the strict k-sparsity requirement on the original signal in [17], overcome the combined obstacle induced by reducing the k-sparsity assumption and incorporating prior support information into the the non-convex ℓ_p optimization problem, establish new results for the recovery of approximately k-sparse signals with partial support information in noiseless setting and two different noise settings by weighted ℓ_p minimization (1.13); and moreover, we characterize the reconstruction error bounds precisely in terms of the noise bound and the non-sparsity of the original signal together with the influence of the prior support information, and therefore derive the stable and robust recovery of approximately k-sparse signals with partial support information in noise settings.

The organization of the rest of this paper is as follows. In Section 2, some preliminary notations, propositions and lemmas are introduced. In Section 3, main theorems are derived, and a series of corollaries and comparisons are presented. The proofs of the main results are presented in Section 4. In Section 5, numerical experiments are conducted in order to demonstrate the performance of the weighted ℓ_p minimization. Finally, the conclusion of this paper is summarized in Section 6.

2. Preliminaries

For any vector $v \in \mathbb{R}^n$, denote $v_{\max(k)} \in \mathbb{R}^n$ the vector with all but the largest k components of v in absolute value set to zeros, and $v_{-\max(k)} := v - v_{\max(k)}$. Denote $v_\Gamma \in \mathbb{R}^n$ as a vector which is equal to v on the index set $\Gamma \subseteq \{1, \ldots, n\}$ and zero elsewhere, and $v_{\Gamma^c} := v - v_\Gamma$.

It is important to note the following fact related to the accuracy of support estimate of the original signal.

Proposition 2.1. *When at least* 50% *of the support estimate is accurate, the RIC condition* (1.18) *by weighted ℓ_p minimization is better compared with the previous RIC condition* (1.8) *by regular ℓ_p minimization, since*

$$\hat{\delta}(p, \Theta) \geq \delta^*(p, t) \quad \textit{for} \quad \alpha \geq 50\%. \tag{2.1}$$

Proof. (i) When $\alpha > 50\%$, we derive

$$\zeta = \left[w + (1-w)\left(1 + \rho - 2\alpha\rho\right)^{\frac{2-p}{2}} \right]^{\frac{2}{2-p}} \in [0, 1]$$

since $w \in [0, 1]$ and $p \in (0, 1]$. By (1.20), we have $d = 1$.

The subsequent discussion is divided into three cases:

(i-1) When $\zeta \in (0, 1)$, according to (1.19), we obtain $\frac{p}{2} z_0^{\frac{2}{p}} + z_0 = \frac{2-p}{2(t-d)}\zeta$. According to (1.9), we obtain $\frac{p}{2}\eta^{\frac{2}{p}} + \eta = \frac{2-p}{2(t-1)}$. Therefore, $\frac{p}{2} z_0^{\frac{2}{p}} + z_0 < \frac{p}{2}\eta^{\frac{2}{p}} + \eta$. It follows from

$$\frac{\partial\left(\frac{p}{2}\eta^{\frac{2}{p}} + \eta\right)}{\partial \eta} = \eta^{\frac{2}{p}-1} + 1 > 0 \quad \text{for} \quad \eta > 0$$

that $\frac{p}{2}\eta^{\frac{2}{p}} + \eta$ is monotonically increasing with respect to $\eta > 0$. Consequently, $z_0 < \eta$. $\eta^{\frac{2}{p}-1}$ is monotonically increasing with respect to $\eta > 0$, and thus

$$\frac{2}{2-p}\left(\frac{p}{2} z_0^{\frac{2}{p}-1} + 1\right) < \frac{2}{2-p}\left(\frac{p}{2}\eta^{\frac{2}{p}-1} + 1\right).$$

It follows from (1.19) that

$$\frac{2}{2-p}\left(\frac{p}{2} z_0^{\frac{2}{p}-1} + 1\right) = \frac{\zeta}{(t-d)z_0},$$

and it follows from (1.9) that

$$\frac{2}{2-p}\left(\frac{p}{2}\eta^{\frac{2}{p}-1} + 1\right) = \frac{1}{(t-1)\eta}.$$

Therefore, $\frac{\zeta}{(t-d)z_0} < \frac{1}{(t-1)\eta}$.

By (1.18), we have $\hat{\delta}(p, \Theta) = \frac{1}{(2-p)\frac{\zeta}{(t-d)z_0} - 1}$. By (1.8), we obtain

$$\delta^*(p, t) = \frac{1}{\frac{2-p}{(t-1)\eta} - 1}.$$

Hence, $\hat{\delta}(p, \Theta) > \delta^*(p, t)$.

(i-2) When $\zeta = 0$, by (1.18), we have $\hat{\delta}(p, \Theta) = 1$. By (1.8), we obtain $\delta^*(p, t) < 1$. Hence, $\hat{\delta}(p, \Theta) > \delta^*(p, t)$.

(i-3) When $\zeta = 1$, we derive $z_0 = \eta$, and thus $\hat{\delta}(p, \Theta) = \delta^*(p, t)$.

In summary, when $\alpha > 50\%$, $\hat{\delta}(p, \Theta) \geq \delta^*(p, t)$.

(ii) When $\alpha = 50\%$, we have $d = 1$ and

$$\zeta = \left[w + (1 - w)\left(1 + \rho - 2\alpha\rho\right)^{\frac{2-p}{2}} \right]^{\frac{2}{2-p}} = 1,$$

and thus $\frac{\zeta}{t-d} = \frac{1}{t-1}$, which yields $z_0 = \eta$. Hence, when $\alpha = 50\%$, $\hat{\delta}(p, \Theta) = \delta^*(p, t)$. □

Therefore, in this paper we will focus on the problem regarding the recovery of approximately k-sparse signals with partial support information when at least 50% of the support estimate is accurate.

In the following, we introduce some lemmas which will be employed in the proof of the main results.

Lemma 2.1 ([28]). *Assume that $x \in \mathbb{R}^n$ satisfies $\|x\|_0 = l$, $\|x\|_\infty \leq \tau$ and $\|x\|_p^p \leq k\tau^p$ with $k \leq l$ being a positive integer, $\tau > 0$ and $0 < p \leq 1$. Then x can be represented as the convex combination of k-sparse vectors u_i, i.e., $x = \sum_{i=1}^N \lambda_i u_i$, where $\lambda_i > 0$, $N \in \mathbb{N}$, $\sum_{i=1}^N \lambda_i = 1$ and $\|u_i\|_0 \leq k$. Moreover,*

$$\sum_{i=1}^{N} \lambda_i \|u_i\|_2^2 \leq \min\left\{ \frac{l}{k} \|x\|_2^2, \, \tau^p \|x\|_{2-p}^{2-p} \right\}. \tag{2.2}$$

Lemma 2.2 ([18]). *For any $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ and $\hat{x} = (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n)^T \in \mathbb{R}^n$, denote $h = \hat{x} - x$ and $T_0 = \operatorname{supp}(x_{\max(k)})$. For any $\tilde{T} \subseteq \{1, \ldots, n\}$ and $w \in [0, 1]$, define $\mathrm{w} = (\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_n)^T$ with*

$$\mathrm{w}_i = \begin{cases} w, & i \in \tilde{T}, \\ 1, & i \notin \tilde{T} \end{cases}, \; i = 1, \ldots, n.$$

If

$$\sum_{i=1}^{n} \mathrm{w}_i^p \, |\hat{x}_i|^p \leq \sum_{i=1}^{n} \mathrm{w}_i^p \, |x_i|^p, \tag{2.3}$$

then

$$\begin{aligned} \left\|h_{T_0^c}\right\|_p^p \leq{} & w^p \left\|h_{T_0}\right\|_p^p + (1 - w^p) \left\|h_{\tilde{T} \cup T_0 \setminus (\tilde{T} \cap T_0)}\right\|_p^p \\ & + 2\left[w^p \left\|x_{T_0^c}\right\|_p^p + (1 - w^p) \left\|x_{\tilde{T}^c \cap T_0^c}\right\|_p^p \right]. \end{aligned} \tag{2.4}$$

Lemma 2.3 (Hölder inequality). *Assume that $a_k \geq 0$, $b_k \geq 0 \, (k = 1, \ldots, n)$, $\frac{1}{r} + \frac{1}{s} = 1$, $r > 1$, $s > 1$. Then*

$$\sum_{k=1}^{n} a_k b_k \leq \left(\sum_{k=1}^{n} a_k^r \right)^{\frac{1}{r}} \left(\sum_{k=1}^{n} b_k^s \right)^{\frac{1}{s}}. \tag{2.5}$$

Lemma 2.4 ([1]). *Suppose that $s \geq r$, $a_1 \geq a_2 \geq \cdots \geq a_s \geq 0$, $\kappa \geq 0$ and $\sum_{i=1}^r a_i + \kappa \geq \sum_{i=r+1}^s a_i$. Then for all $\omega \geq 1$,*

$$\sum_{i=r+1}^{s} a_i^\omega \leq r \left[\left(\frac{1}{r} \sum_{i=1}^{r} a_i^\omega \right)^{\frac{1}{\omega}} + \frac{\kappa}{r} \right]^\omega. \tag{2.6}$$

3. Main Results

For signal recovery in noise settings, to be broadly applicable, the original signal which is approximately k-sparse is expected to be recovered with bounded errors. The results for stable and robust recovery of approximately k-sparse signals with partial support information in two different noise settings (1.2) and (1.3) are derived in this section.

Firstly, we consider the recovery of approximately k-sparse signals with partial support information in the ℓ_2 bounded noise setting.

Theorem 3.1. *Consider the signal recovery model* (1.1) *with* $\|z\|_2 \le \epsilon$. *Suppose that* $T_0 = \operatorname{supp}(x_{\max(k)})$, *the prior support estimate is* $\tilde{T} \subseteq \{1, \ldots, n\}$ *with cardinality* $|\tilde{T}| = \rho k$ $(\rho \ge 0)$, $|T_0 \cap \tilde{T}| = \alpha |\tilde{T}|$ *with* $\alpha \in [\frac{1}{2}, 1]$. *Suppose that* $\hat{x}^{\ell_2}$ *is the minimizer of the weighted* ℓ_p $(0 < p \le 1)$ *minimization* (1.13) *with* $\mathcal{B} = \mathcal{B}^{\ell_2}(\varepsilon)$ *for some* $\varepsilon \ge \epsilon$ *and the weight vector* $\mathrm{w} \in \{w, 1\}^n$ *is defined in* (1.11). *Denote*

$$\sigma := \left[w^p + (1 - w^p)(1 + \rho - 2\alpha\rho)^{\frac{2-p}{2}}\right]^{\frac{2}{2-p}}. \tag{3.1}$$

If the measurement matrix A *satisfies*

$$\delta_{tk} < \delta(p, t, \sigma) := \frac{1}{p\eta^{\frac{2}{p}-1} + 1} \tag{3.2}$$

for some $t \in [1 + \frac{2-p}{2+p}\sigma, 2]$, *where* $\eta \in [(1-p)\frac{\sigma}{t-1}, \min\{1, (1 - \frac{\sqrt{\frac{w^2}{t-1}+1}}{\sqrt{\frac{w^2}{t-1}+1}+1}p)\frac{\sigma}{t-1}\}]$ *is the unique nonnegative solution of the following equation:*

$$\frac{p}{2}\eta^{\frac{2}{p}} + \eta - \frac{2-p}{2(t-1)}\sigma = 0, \tag{3.3}$$

then

$$\left\|\hat{x}^{\ell_2} - x\right\|_2 \le C_1(\varepsilon + \epsilon) + C_2\left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p, \tag{3.4}$$

with $C_1 = \sqrt{1 + 2^{\frac{2}{p}-2}}D_1$ *and* $C_2 = \sqrt{D_2^2 + 2^{\frac{2}{p}-2}\left[D_2 + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\right]^2}$.

The notations of D_1 *and* D_2 *are as follows.*

(1) *When* $\sigma\delta_{tk} \ne 0$,

$$\begin{cases} D_1 = \dfrac{\left(\frac{1}{\delta(p,t,\sigma)} - 1 + p\right)\sqrt{1+\delta_{tk}} + \sqrt{\left(\frac{1}{\delta(p,t,\sigma)} - 1 + p\right)^2(1+\delta_{tk}) + 4\lambda(1-p)\beta(p,t,\sigma)}}{2\lambda\beta(p,t,\sigma)}, \\ D_2 = \left[\dfrac{2}{(k\sigma)^{\frac{2-p}{2}}}\right]^{\frac{1}{p}}\left\{\left\{\dfrac{2(1-\lambda)\beta(p,t,\sigma)}{p(1+\delta_{tk})\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)(1+\delta_{tk})}\right]^{\frac{2-p}{p}}} + 1\right\}^{\frac{p}{2}} - 1\right\}^{-\frac{1}{p}} \end{cases} \tag{3.5}$$

with $\lambda \in (0, 1)$ *and*

$$\beta(p, t, \sigma) := \frac{2\delta(p,t,\sigma) - \left(1 + \delta^2(p,t,\sigma)\right)\delta_{tk}}{2\delta^2(p,t,\sigma)} - \frac{p}{2}(1+\delta_{tk})\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)(1+\delta_{tk})}\right]^{\frac{2-p}{p}}. \tag{3.6}$$

(2) *When* $\sigma = 0$,

$$\begin{cases} D_1 = \dfrac{p\sqrt{1+\delta_{tk}} + \sqrt{p^2\left(1+\delta_{tk}\right) + 4(1-p)\left(1-\delta_{tk}\right)}}{2\left(1-\delta_{tk}\right)}, \\ D_2 = \left(\dfrac{2}{k^{\frac{2-p}{2}}}\right)^{\frac{1}{p}} \left[\dfrac{(2-p)\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{2p}} \sqrt{\dfrac{p\left(1+\delta_{tk}\right)}{2\left(1-\delta_{tk}\right)}}. \end{cases} \tag{3.7}$$

(3) *When* $\delta_{tk} = 0, \quad D_1 = 1$ and $D_2 = 0$.

Remark 3.1. In Theorem 3.1, not only the stable and robust reconstruction of approximately k-sparse signals with partial support information in ℓ_2 bounded noise setting is derived via the weighted ℓ_p minimization, but also the RIC condition (3.2) is characterized in uniform representation.

Remark 3.2. As for the recovery of approximately k-sparse signals, the new results by weighted ℓ_p minimization include the results in [8,24] as special cases by setting the weight vector $\mathrm{w} = \mathbf{1}$.

In particular, when $\mathrm{w} \circ x_{-\max(k)} = \mathbf{0}$, we have the following result for the stable and robust recovery of approximately k-sparse signals with partial support information in ℓ_2 bounded noise setting. In this case, the reconstruction error estimation can be more tightly characterized.

Corollary 3.1. *Consider the signal recovery model* (1.1) *with* $\|z\|_2 \leq \epsilon$. *Suppose that* $T_0 = \mathrm{supp}(x_{\max(k)})$, *the prior support estimate is* $\tilde{T} \subseteq \{1, \ldots, n\}$ *with* $|\tilde{T}| = \rho k$ $(\rho \geq 0)$, $|T_0 \cap \tilde{T}| = \alpha|\tilde{T}|$ *with* $\alpha \in [\frac{1}{2}, 1]$. *Suppose that* $\hat{x}^{\ell_2}$ *is the minimizer of the weighted* $\ell_p\,(0 < p \leq 1)$ *minimization* (1.13) *with* $\mathcal{B} = \mathcal{B}^{\ell_2}(\varepsilon)$ *for some* $\varepsilon \geq \epsilon$, *the weight vector* $\mathrm{w} \in \{w, 1\}^n$ *is defined in* (1.11), σ *is defined in* (3.1), *and* $\beta(p, t, \sigma)$ *is defined in* (3.6). *If* $\mathrm{w} \circ x_{-\max(k)} = \mathbf{0}$ *and the measurement matrix* A *satisfies* (3.2) *for some* $t \in [1 + \frac{2-p}{2+p}\sigma, 2]$, *then*

$$\left\|\hat{x}^{\ell_2} - x\right\|_2 \leq \sqrt{2} D_3\left(\varepsilon + \epsilon\right), \tag{3.8}$$

where

$$D_3 = \frac{\left(\frac{1}{\delta(p,t,\sigma)} - 1 + p\right)\sqrt{1+\delta_{tk}}\,2\beta(p,t,\sigma)}{2\beta(p,t,\sigma)} + \frac{\sqrt{\left(\frac{1}{\delta(p,t,\sigma)} - 1 + p\right)^2 (1+\delta_{tk}) + 4(1-p)\beta(p,t,\sigma)}}{2\beta(p,t,\sigma)}. \tag{3.9}$$

As a result of Corollary 3.1, the following result for the recovery of exactly k-sparse signals with partial support information in ℓ_2 bounded noise setting can be directly obtained, which corresponds to the main result in [17].

Corollary 3.2. *Consider the signal recovery model* (1.1) *where* $x \in \mathbb{R}^n$ *is* k*-sparse and* $\|z\|_2 \leq \epsilon$. *Denote* $T = \mathrm{supp}(x)$. *Suppose that the prior support estimate is* $\tilde{T} \subseteq \{1, \ldots, n\}$ *with* $|\tilde{T}| = \rho k$ $(\rho \geq 0)$, $|T \cap \tilde{T}| = \alpha|\tilde{T}|$ *with* $\alpha \in [\frac{1}{2}, 1]$. *Suppose that* $\hat{x}^{\ell_2}$ *is the minimizer of the weighted* $\ell_p\,(0 < p \leq 1)$ *minimization* (1.13) *with* $\mathcal{B} = \mathcal{B}^{\ell_2}(\varepsilon)$ *for some* $\varepsilon \geq \epsilon$, *the weight vector* $\mathrm{w} \in \{w, 1\}^n$ *is defined in* (1.11), *and* σ *is defined in* (3.1). *If the measurement matrix* A *satisfies* (3.2) *for some* $t \in [1 + \frac{2-p}{2+p}\sigma, 2]$, *then* $\|\hat{x}^{\ell_2} - x\|_2 \leq \sqrt{2} D_3(\varepsilon + \epsilon)$.

In noiseless setting, the recovery of approximately k-sparse signals with partial support information can be obtained via weighted ℓ_p minimization, and the reconstruction error estimation is precisely characterized, which indicates that the signal reconstruction is stable under sparsity defect.

Theorem 3.2. *Consider the signal recovery model* (1.1) *with $z=0$. Denote $T_0=\operatorname{supp}(x_{\max(k)})$. Suppose that the prior support estimate is $\tilde{T}\subseteq\{1,\ldots,n\}$ with $|\tilde{T}|=\rho k$ ($\rho\geq 0$), $|T_0\cap\tilde{T}|=\alpha|\tilde{T}|$ with $\alpha\in[\frac{1}{2},1]$.*

Suppose that $\hat{x}$ is the minimizer of the weighted $\ell_p\,(0<p\leq 1)$ minimization (1.13) *with $\mathcal{B}=\{0\}$, the weight vector $\mathrm{w}\in\{w,1\}^n$ is defined in* (1.11), *and σ is defined in* (3.1). *If the measurement matrix A satisfies* (3.2) *for some $t\in[1+\frac{2-p}{2+p}\sigma,2]$, then*

$$\|\hat{x}-x\|_2\leq\sqrt{D_4^2+\left(D_4^p+\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{2}{p}}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p, \tag{3.10}$$

where

$$D_4=\begin{cases}\left[\dfrac{2}{(k\sigma)^{\frac{2-p}{2}}}\right]^{\frac{1}{p}}\left\{\left\{\dfrac{2\delta(p,t,\sigma)-\left(1+\delta^2(p,t,\sigma)\right)\delta_{tk}}{p\left(1+\delta_{tk}\right)\delta^2(p,t,\sigma)\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)(1+\delta_{tk})}\right]^{\frac{2-p}{p}}}\right\}^{\frac{p}{2}}-1\right\}^{-\frac{1}{p}}, & \\ & \sigma\delta_{tk}\neq 0,\\ \left(\dfrac{2}{k^{\frac{2-p}{2}}}\right)^{\frac{1}{p}}\sqrt{\dfrac{p\left(1+\delta_{tk}\right)}{2\left(1-\delta_{tk}\right)}}\left[\dfrac{(2-p)\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{2p}}, & \sigma=0,\\ 0, & \delta_{tk}=0.\end{cases} \tag{3.11}$$

In the *DS* noise setting (1.3), we deduce that the stable and robust recovery of approximately k-sparse signals with partial support information can be guaranteed under the RIC condition (3.2).

Theorem 3.3. *Consider the signal recovery model* (1.1) *with $\|A^Tz\|_\infty\leq\epsilon$. Denote $T_0=$* $\operatorname{supp}(x_{\max(k)})$. *Suppose that the prior support estimate is $\tilde{T}\subseteq\{1,\ldots,n\}$ with $|\tilde{T}|=\rho k$ ($\rho\geq 0$), $|T_0\cap\tilde{T}|=\alpha|\tilde{T}|$ with $\alpha\in[\frac{1}{2},1]$. Suppose $\hat{x}^{DS}$ is the minimizer of the weighted $\ell_p\,(0<p\leq 1)$ minimization* (1.13) *with $\mathcal{B}=\mathcal{B}^{DS}(\varepsilon)$ for some $\varepsilon\geq\epsilon$, the weight vector $\mathrm{w}\in\{w,1\}^n$ is defined in* (1.11), *and σ is defined in* (3.1). *If the measurement matrix A satisfies* (3.2) *for some $t\in[1+\frac{2-p}{2+p}\sigma,2]$, then*

$$\left\|\hat{x}^{DS}-x\right\|_2\leq C_5\left(\varepsilon+\epsilon\right)+C_6\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p, \tag{3.12}$$

where $C_5=\sqrt{1+2^{\frac{2}{p}-2}}D_5$ and $C_6=\sqrt{D_6^2+2^{\frac{2}{p}-2}\left[D_6+\left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\right]^2}$.

The notations of D_5 and D_6 are as follows.

(1) *When $\sigma\delta_{tk}\neq 0$,*

$$\begin{cases}D_5=\dfrac{(1-p)\left(1+2^{\frac{1}{p}-1}\right)+\left(\frac{1}{\delta(p,t,\sigma)}-1+p\right)\sqrt{t}}{\lambda\beta(p,t,\sigma)}\sqrt{k},\\ D_6=\max\left\{\dfrac{(1-p)\left(\frac{2}{\sqrt{k}}\right)^{\frac{2}{p}-1}}{(1-p)\left(1+2^{\frac{1}{p}-1}\right)+\left(\frac{1}{\delta(p,t,\sigma)}-1+p\right)\sqrt{t}},\,D_2\right\}\end{cases} \tag{3.13}$$

with $\lambda\in(0,1)$, D_2 in (3.5) *and $\beta(p,t,\sigma)$ in* (3.6).

(2) *When* $\sigma = 0$,

$$\begin{cases} D_5 = \dfrac{(1-p)\big(1+2^{\frac{1}{p}-1}\big)+p\sqrt{t}}{1-\delta_{tk}}\sqrt{k}, \\ D_6 = \left(\dfrac{2}{k}\right)^{\frac{1}{p}-\frac{1}{2}}\left\{\dfrac{(1-p)\,2^{\frac{1}{p}-\frac{1}{2}}}{(1-p)\big(1+2^{\frac{1}{p}-1}\big)+p\sqrt{t}}+\sqrt{\dfrac{p\,(1+\delta_{tk})}{1-\delta_{tk}}}\left[\dfrac{(2-p)\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{2p}}\right\}. \end{cases} \tag{3.14}$$

(3) *When* $\delta_{tk} = 0$,

$$\begin{cases} D_5 = \left[(1-p)\big(1+2^{\frac{1}{p}-1}\big)+\big(\dfrac{1}{\delta(p,t,\sigma)}-1+p\big)\sqrt{t}\right]\delta(p,t,\sigma)\sqrt{k}, \\ D_6 = \dfrac{(1-p)\left(\frac{2}{\sqrt{k}}\right)^{\frac{2}{p}-1}}{(1-p)\big(1+2^{\frac{1}{p}-1}\big)+\left(\frac{1}{\delta(p,t,\sigma)}-1+p\right)\sqrt{t}}. \end{cases} \tag{3.15}$$

In particular, when $\mathrm{w} \circ x_{-\max(k)} = \mathbf{0}$, we have the following result for the recovery of approximately k-sparse signals with partial support information in the *DS* noise setting.

Corollary 3.3. *Consider the signal recovery model* (1.1) *with* $\|A^T z\|_\infty \le \epsilon$. *Denote* $T_0 = \mathrm{supp}(x_{\max(k)})$. *Suppose that the prior support estimate is* $\tilde{T} \subseteq \{1,\dots,n\}$ *with* $|\tilde{T}| = \rho k$ $(\rho \ge 0)$, $|T_0 \cap \tilde{T}| = \alpha|\tilde{T}|$ *with* $\alpha \in [\frac{1}{2},1]$. *Suppose that* $\hat{x}^{DS}$ *is the minimizer of the weighted* ℓ_p $(0 < p \le 1)$ *minimization* (1.13) *with* $\mathcal{B} = \mathcal{B}^{DS}(\varepsilon)$ *for some* $\varepsilon \ge \epsilon$, *the weight vector* $\mathrm{w} \in \{w,1\}^n$ *is defined in* (1.11), σ *is defined in* (3.1), *and* $\beta(p,t,\sigma)$ *is defined in* (3.6). *If* $\mathrm{w} \circ x_{-\max(k)} = \mathbf{0}$ *and the measurement matrix* A *satisfies* (3.2) *for some* $t \in [1+\frac{2-p}{2+p}\sigma, 2]$, *then*

$$\left\|\hat{x}^{DS}-x\right\|_2 \le \sqrt{2k}\,\frac{2(1-p)+\left(\frac{1}{\delta(p,t,\sigma)}-1+p\right)\sqrt{t}}{\beta(p,t,\sigma)}\,(\varepsilon+\epsilon)\,. \tag{3.16}$$

Stable recovery of k-sparse signals with partial support information in the noise setting $\mathcal{B}^{DS}(\epsilon)$ can be directly obtained from Corollary 3.3. When the original signal $x \in \mathbb{R}^n$ is k-sparse, we derive $x_{-\max(k)} = \mathbf{0}$, and thus $\mathrm{w} \circ x_{-\max(k)} = \mathbf{0}$. Then the result directly follows from Corollary 3.3.

Corollary 3.4. *Consider the signal recovery model* (1.1) *where* $x \in \mathbb{R}^n$ *is* k*-sparse and* $\|A^T z\|_\infty \le \epsilon$. *Denote* $T = \mathrm{supp}(x)$. *Suppose that the prior support estimate is* $\tilde{T} \subseteq \{1,\dots,n\}$ *with* $|\tilde{T}| = \rho k$ $(\rho \ge 0)$, $|T \cap \tilde{T}| = \alpha|\tilde{T}|$ *with* $\alpha \in [\frac{1}{2},1]$.

Suppose that $\hat{x}^{DS}$ *is the minimizer of the weighted* ℓ_p $(0 < p \le 1)$ *minimization* (1.13) *with* $\mathcal{B} = \mathcal{B}^{DS}(\varepsilon)$ *for some* $\varepsilon \ge \epsilon$, *the weight vector* $\mathrm{w} \in \{w,1\}^n$ *is defined in* (1.11), *and* σ *is defined in* (3.1). *If the measurement matrix* A *satisfies* (3.2) *for some* $t \in [1+\frac{2-p}{2+p}\sigma, 2]$, *then* (3.16) *holds.*

Furthermore, when in particular $\sigma = 1$ (For example, if $\mathrm{w} = \mathbf{1}$ or $\alpha = \frac{1}{2}$, then $\sigma = [w^p + (1-w^p)(1+\rho-2\alpha\rho)^{\frac{2-p}{2}}]^{\frac{2}{2-p}} = 1$), the RIC condition (3.2) coincides with the sharp condition (1.8) derived by ℓ_p minimization in [24], we have the following result, and the reconstruction error estimation can be further improved when $\mathrm{w} = \mathbf{1}$.

Corollary 3.5. *Consider the signal recovery model* (1.1) *where $x \in \mathbb{R}^n$ is k-sparse and* $\|A^T z\|_\infty \leq \epsilon$. *Denote* $T = \operatorname{supp}(x)$. *Suppose that the prior support estimate is* $\tilde{T} \subseteq \{1, \ldots, n\}$ *with* $|\tilde{T}| = \rho k$ $(\rho \geq 0)$, $|T \cap \tilde{T}| = \alpha |\tilde{T}|$ *with* $\alpha \in [\frac{1}{2}, 1]$.

Suppose that $\hat{x}^{DS}$ *is the minimizer of the weighted* ℓ_p $(0 < p \leq 1)$ *minimization* (1.13) *with* $\mathcal{B} = \mathcal{B}^{DS}(\varepsilon)$ *for some* $\varepsilon \geq \epsilon$, *the weight vector* $\mathrm{w} \in \{w, 1\}^n$ *defined in* (1.11), *and* σ *is defined in* (3.1). *If* $\sigma = 1$ *and the measurement matrix* A *satisfies* (3.2) *for some* $t \in [\frac{4}{2+p}, 2]$, *i.e.,* $\delta_{tk} < \delta(p, t, 1)$, *then*

$$\left\|\hat{x}^{DS} - x\right\|_2 \leq \sqrt{2k} \frac{2(1-p) + \left(\frac{1}{\delta(p,t,1)} - 1 + p\right)\sqrt{t}}{\beta(p,t,1)} (\varepsilon + \epsilon). \tag{3.17}$$

In particular, if $\mathrm{w} = \mathbf{1}$, *then*

$$\left\|\hat{x}^{DS} - x\right\|_2 \leq \sqrt{2k} \frac{1 - p + \frac{1}{\delta(p,t,1)}}{\beta(p,t,1)} (\varepsilon + \epsilon). \tag{3.18}$$

Remark 3.3. (1) Theorems 3.1 and 3.3 include our previous results in [24] for the recovery of approximately k-sparse signals by standard ℓ_p minimization method (1.4) as special cases. Actually, since prior support information of the original signal is not known or not exploited by the standard ℓ_p minimization (1.4), which corresponds to (1.13) with $\mathrm{w} = \mathbf{1} \in \mathbb{R}^n$, then $\sigma = 1$. Therefore, the RIC condition (3.2) becomes the RIC condition (1.8) derived by ℓ_p minimization in [24], i.e., $\delta(p, t, 1) = \delta^*(p, t)$.

(2) For the recovery of approximately k-sparse signals, when at least 50% of the support estimate is accurate, the RIC condition (3.2) by weighted ℓ_p minimization is better than the condition (1.8) by regular ℓ_p minimization, since for $w \in [0, 1)$, $\delta(p, t, \sigma) > \delta^*(p, t)$ for $\alpha > 50\%$ and $\delta(p, t, \sigma) = \delta^*(p, t)$ for $\alpha = 50\%$.

Remark 3.4. Theorems 3.1 and 3.3 include the main results in [9] by the weighted ℓ_1 minimization method as special cases. Actually, when in particular $p = 1$, it follows from (3.2) that $\sigma = \left[w + (1-w)\sqrt{1 + \rho - 2\alpha\rho}\right]^2$, the unique nonnegative solution the equation $\frac{\eta^2}{2} + \eta - \frac{\sigma}{2(t-1)} = 0$ is $\eta = \sqrt{1 + \frac{\sigma}{t-1}} - 1$, and thus $\delta(1, t, \sigma) = \frac{1}{\eta+1} = \sqrt{\frac{t-1}{t-1+\sigma}}$. Therefore, the RIC condition (3.2) for $p = 1$ via weighted ℓ_1 minimization is

$$\delta_{tk} < \sqrt{\frac{t-1}{t - 1 + \left[w + (1-w)\sqrt{1 + \rho - 2\alpha\rho}\right]^2}},$$

which is the RIC condition in [9].

Remark 3.5. The reconstruction error estimations (3.4) in Theorem 3.1 and (3.12) in Theorem 3.3 indicate that the obtained result for signal recovery in the two noise settings is robust under noise and stable under the non-sparsity of the original signal together with the influence of the prior support information.

4. Proofs of the Main Results

In this section, the theorems and corollaries presented in Section 3 are proved.

4.1. Proof of Theorem 3.1

(i) Firstly, we will prove that there is a unique nonnegative solution η of Eq. (3.3), i.e.,

$$\frac{p}{2}\eta^{\frac{2}{p}}+\eta-\frac{2-p}{2(t-1)}\sigma=0,$$

and the solution satisfies

$$\eta\in\left[(1-p)\frac{\sigma}{t-1},\min\left\{1,\left(1-\frac{\sqrt{\frac{w^2}{t-1}+1}}{\sqrt{\frac{w^2}{t-1}+1}+1}p\right)\frac{\sigma}{t-1}\right\}\right].$$

Denote

$$\varphi(\eta,p):=\frac{p}{2}\eta^{\frac{2}{p}}+\eta-\frac{2-p}{2(t-1)}\sigma.$$

It follows from

$$\varphi'_\eta(\eta,p)=\eta^{\frac{2}{p}-1}+1>0\quad\text{for}\quad\eta\geq 0$$

that $\varphi(\eta,p)$ is monotonically increasing for $\eta\in[0,+\infty)$.

Since $\alpha\in[\frac{1}{2},1]$, we have $1+\rho-2\alpha\rho\in[0,1]$. It follows from $w\in[0,1]$ and $p\in(0,1]$ that

$$\sigma=\left[w^p+(1-w^p)\left(1+\rho-2\alpha\rho\right)^{\frac{2-p}{2}}\right]^{\frac{2}{2-p}}\in\left[w^{\frac{2p}{2-p}},1\right].$$

For $t\in[1+\frac{2-p}{2+p}\sigma,2]$, $\frac{w^{\frac{2p}{2-p}}}{t-1}\leq\frac{\sigma}{t-1}\leq\frac{2+p}{2-p}$. Therefore,

$$\begin{aligned}\varphi(\frac{1-p}{t-1}\sigma,p)=&\frac{p\sigma}{2(t-1)}\left[(1-p)^{\frac{2}{p}}\left(\frac{\sigma}{t-1}\right)^{\frac{2}{p}-1}-1\right]\\ \leq&\frac{p\sigma}{2(t-1)}\left[(1-p)^{\frac{2}{p}}\left(\frac{2+p}{2-p}\right)^{\frac{2}{p}-1}-1\right]\leq 0.\end{aligned}\tag{4.1}$$

The above inequality follows from the fact that $(1-p)^{\frac{2}{p}}$ and $(\frac{2+p}{2-p})^{\frac{2}{p}-1}$ are monotonically decreasing with $p\in(0,1]$ and $\lim_{p\to 0^+}(1-p)^{\frac{2}{p}}(\frac{2+p}{2-p})^{\frac{2}{p}-1}=\mathrm{e}^{-2}\mathrm{e}^2=1$.

We have $\varphi(1,p)=\frac{2+p}{2}-\frac{2-p}{2(t-1)}\sigma\geq 0$ since $t\geq 1+\frac{2-p}{2+p}\sigma$.

It follows from $w\in[0,1]$ that $\frac{\sqrt{\frac{w^2}{t-1}+1}}{\sqrt{\frac{w^2}{t-1}+1}+1}\in[\frac{1}{2},\frac{\sqrt{t}}{\sqrt{t}+\sqrt{t-1}}]$. We have

$$\begin{aligned}&\varphi\left(\left(1-\frac{\sqrt{\frac{w^2}{t-1}+1}}{\sqrt{\frac{w^2}{t-1}+1}+1}p\right)\frac{\sigma}{t-1},p\right)\\ =&\frac{p\sigma}{2(t-1)}\left(1-\frac{\sqrt{\frac{w^2}{t-1}+1}}{\sqrt{\frac{w^2}{t-1}+1}+1}p\right)^{\frac{2}{p}}\left(\frac{\sigma}{t-1}\right)^{\frac{2}{p}-1}+\frac{1-\sqrt{\frac{w^2}{t-1}+1}}{\sqrt{\frac{w^2}{t-1}+1}+1}\cdot\frac{p\sigma}{2(t-1)}\\ \geq&\frac{p\sigma}{2(t-1)}\left[\left(1-\frac{\sqrt{\frac{w^2}{t-1}+1}}{\sqrt{\frac{w^2}{t-1}+1}+1}p\right)^{\frac{2}{p}}\frac{w^2}{(t-1)^{\frac{2}{p}-1}}+\frac{1-\sqrt{\frac{w^2}{t-1}+1}}{\sqrt{\frac{w^2}{t-1}+1}+1}\right]\\ \geq&\frac{p\sigma}{2(t-1)}\left[\left(1-\frac{\sqrt{\frac{w^2}{t-1}+1}}{\sqrt{\frac{w^2}{t-1}+1}+1}\right)^{2}\frac{w^2}{t-1}-\frac{\sqrt{\frac{w^2}{t-1}+1}-1}{\sqrt{\frac{w^2}{t-1}+1}+1}\right]=0.\end{aligned}\tag{4.2}$$

As a consequence, the unique nonnegative solution of (3.3) satisfies $\eta \in [(1-p)\frac{\sigma}{t-1}, \min\{1, (1-\frac{\sqrt{\frac{w^2}{t-1}+1}}{\sqrt{\frac{w^2}{t-1}+1}+1}p)\frac{\sigma}{t-1}\}]$.

(ii) In the following, we will prove stable and robust recovery of approximately k-sparse signals with partial support information in ℓ_2 bounded noise setting.

Denote $h = \hat{x}^{\ell_2} - x$. By virtue of $T_0 = \text{supp}(x_{\max(k)})$ and Lemma 2.2, we have

$$\begin{aligned}\left\|h_{-\max(k)}\right\|_p^p \le & \|h_{T_0^c}\|_p^p \qquad (4.3)\\ \le & w^p\|h_{T_0}\|_p^p + (1-w^p)\|h_{\tilde{T}\cup T_0\setminus(\tilde{T}\cap T_0)}\|_p^p + 2\left[w^p\|x_{T_0^c}\|_p^p + (1-w^p)\|x_{\tilde{T}^c\cap T_0^c}\|_p^p\right]\\ \le & w^p\left\|h_{\max(k)}\right\|_p^p + (1-w^p)\|h_{T_0\cup\tilde{T}\setminus(\tilde{T}\cap T_0)}\|_p^p + 2\left[w^p\|x_{T_0^c}\|_p^p + (1-w^p)\|x_{\tilde{T}^c\cap T_0^c}\|_p^p\right].\end{aligned}$$

Since

$$\begin{aligned}& w^p\|x_{T_0^c}\|_p^p + (1-w^p)\|x_{\tilde{T}^c\cap T_0^c}\|_p^p\\ =& w^p\|x_{\tilde{T}\cap T_0^c}\|_p^p + w^p\|x_{\tilde{T}^c\cap T_0^c}\|_p^p + (1-w^p)\|x_{\tilde{T}^c\cap T_0^c}\|_p^p]\\ =& w^p\|x_{\tilde{T}\cap T_0^c}\|_p^p + \|x_{\tilde{T}^c\cap T_0^c}\|_p^p = \left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p, \qquad (4.4)\end{aligned}$$

by (4.3) we derive

$$\left\|h_{-\max(k)}\right\|_p^p \le w^p\left\|h_{\max(k)}\right\|_p^p + (1-w^p)\|h_{T_0\cup\tilde{T}\setminus(\tilde{T}\cap T_0)}\|_p^p + 2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p. \qquad (4.5)$$

Denote

$$\nu = \left[\frac{w^p\left\|h_{\max(k)}\right\|_p^p + (1-w^p)\|h_{T_0\cup\tilde{T}\setminus(\tilde{T}\cap T_0)}\|_p^p + 2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k}\right]^{\frac{1}{p}}. \qquad (4.6)$$

Now we divide $h_{-\max(k)}$ as $h_{-\max(k)} = h_{(1)} + h_{(2)}$, where the components of $h_{(1)}$ and $h_{(2)}$ respectively satisfy

$$h_{(1)}(i) = \begin{cases} h_{-\max(k)}(i), & \left|h_{-\max(k)}(i)\right| > (t-1)^{-\frac{1}{p}}\nu,\\ 0, & \text{otherwise,}\end{cases} \qquad (4.7)$$

$$h_{(2)}(i) = \begin{cases} h_{-\max(k)}(i), & \left|h_{-\max(k)}(i)\right| \le (t-1)^{-\frac{1}{p}}\nu,\\ 0, & \text{otherwise}\end{cases} \qquad (4.8)$$

for $i = 1, \ldots, n$.

Denote $|\text{supp}(h_{(1)})| = \mathrm{r}$. Then it follows from (4.5) that

$$k\nu^p \ge \|h_{-\max(k)}\|_p^p \ge \|h_{(1)}\|_p^p \ge \mathrm{r}\frac{\nu^p}{t-1},$$

and thus $0 \le \mathrm{r} \le (t-1)k$. We derive

$$\|Ah\|_2 = \|A\hat{x}^{\ell_2} - Ax\|_2 \le \|y - A\hat{x}^{\ell_2}\|_2 + \|Ax - y\|_2 \le \varepsilon + \epsilon, \qquad (4.9)$$

$$\begin{aligned}\langle A\left(h_{\max(k)} + h_1\right), Ah\rangle \le & \|A\left(h_{\max(k)} + h_1\right)\|_2\|Ah\|_2\\ \le & \sqrt{1+\delta_{tk}}\left\|h_{\max(k)} + h_1\right\|_2(\varepsilon + \epsilon). \qquad (4.10)\end{aligned}$$

Moreover, we have $\|h_{(2)}\|_\infty \le (t-1)^{-\frac{1}{p}}\nu$ and

$$\|h_{(2)}\|_p^p = \|h_{-\max(k)}\|_p^p - \|h_{(1)}\|_p^p \le k\nu^p - \mathrm{r}\frac{\nu^p}{t-1} = [(t-1)k - \mathrm{r}]\,\frac{\nu^p}{t-1}. \tag{4.11}$$

Therefore, by Lemma 2.1, we deduce that $h_{(2)}$ can be represented as $h_{(2)} = \sum_{i=1}^N \lambda_i u_i$, where $\lambda_i > 0$, $\sum_{i=1}^N \lambda_i = 1$, u_i is $[(t-1)k-\mathrm{r}]$-sparse, and

$$\sum_{i=1}^N \lambda_i \|u_i\|_2^2 \le \frac{\nu^p}{t-1}\|h_{(2)}\|_{2-p}^{2-p}. \tag{4.12}$$

By virtue of Hölder inequality and (4.11), we obtain

$$\begin{aligned}
&\sum_{i=1}^N \lambda_i \|u_i\|_2^2 \le \frac{\nu^p}{t-1}\|h_{(2)}\|_{2-p}^{2-p} \\
\le&\frac{\nu^p}{t-1}\left(\|h_{(2)}\|_2^2\right)^{\frac{2(1-p)}{2-p}}\left(\|h_{(2)}\|_p^p\right)^{\frac{p}{2-p}} \le \frac{\nu^p}{t-1}\left(\|h_{(2)}\|_2^2\right)^{\frac{2(1-p)}{2-p}}\left\{[(t-1)k-r]\,\frac{\nu^p}{t-1}\right\}^{\frac{p}{2-p}} \\
\le&\frac{\nu^p}{t-1}\left(\|h_{(2)}\|_2^2\right)^{\frac{2(1-p)}{2-p}}(k\nu^p)^{\frac{p}{2-p}} = \frac{1}{(t-1)k}\left(\|h_{(2)}\|_2^2\right)^{\frac{2(1-p)}{2-p}}(k\nu^p)^{\frac{2}{2-p}} \\
=&\frac{\left(\|h_{(2)}\|_2^2\right)^{\frac{2(1-p)}{2-p}}}{(t-1)k}\left[w^p\left\|h_{\max(k)}\right\|_p^p + (1-w^p)\|h_{T_0\cup\tilde{T}\setminus(\tilde{T}\cap T_0)}\|_p^p + 2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p\right]^{\frac{2}{2-p}}.
\end{aligned}$$

Note that $|T_0 \cup \tilde{T} \setminus (\tilde{T}\cap T_0)| = k + \rho k - 2\alpha\rho k = (1+\rho-2\alpha\rho)k \le k$ since $\alpha \ge \frac{1}{2}$. Recall (3.1).

Therefore, we deduce

$$\begin{aligned}
&w^p\left\|h_{\max(k)}\right\|_p^p + (1-w^p)\|h_{T_0\cup\tilde{T}\setminus(\tilde{T}\cap T_0)}\|_p^p \\
\le&w^p k^{\frac{2-p}{2}}\|h_{\max(k)}\|_2^p + (1-w^p)\|h_{T_0\cup\tilde{T}\setminus(\tilde{T}\cap T_0)}\|_2^p\,[(1+\rho-2\alpha\rho)k]^{\frac{2-p}{2}} \\
\le&w^p k^{\frac{2-p}{2}}\|h_{\max(k)}\|_2^p + (1-w^p)\|h_{\max(k)}\|_2^p\,[(1+\rho-2\alpha\rho)k]^{\frac{2-p}{2}} \\
=&\left[w^p + (1-w^p)(1+\rho-2\alpha\rho)^{\frac{2-p}{2}}\right]k^{\frac{2-p}{2}}\|h_{\max(k)}\|_2^p \\
=&\sigma^{\frac{2-p}{2}}k^{\frac{2-p}{2}}\|h_{\max(k)}\|_2^p.
\end{aligned} \tag{4.13}$$

As a consequence,

$$\sum_{i=1}^N \lambda_i \|u_i\|_2^2 \le \frac{\left(\|h_{(2)}\|_2^2\right)^{\frac{2(1-p)}{2-p}}}{t-1}\left(\sigma^{\frac{2-p}{2}}\|h_{\max(k)}\|_2^p + \frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{2-p}}. \tag{4.14}$$

For any $\mu \in [0,1]$, denote $\beta_i = h_{\max(k)} + h_{(1)} + \mu u_i$ $(i = 1,\dots,N)$. Then

$$\sum_{j=1}^N \lambda_j\beta_j - \frac{p}{2}\beta_i = \left(1-\frac{p}{2}-\mu\right)\left(h_{\max(k)} + h_{(1)}\right) - \frac{p}{2}\mu u_i + \mu h.$$

It is clear that $\sum_{j=1}^{N}\lambda_j\beta_j-\frac{p}{2}\beta_j-\mu h$ and $\left(1-\frac{p}{2}-\mu\right)\left(h_{\max(k)}+h_{(1)}\right)-\frac{p}{2}\mu u_i$ are tk-sparse. In addition, u_i-u_j is $2\left[(t-1)k-\mathrm{r}\right]$-sparse, and thus u_i-u_j is tk-sparse since $2\left[(t-1)k-\mathrm{r}\right]\leq tk$ for $t\in[1+\frac{2-p}{2+p}\sigma,2]$.

It is easy to check the following identity

$$\sum_{i=1}^{N}\lambda_i\|A(\sum_{j=1}^{N}\lambda_j\beta_j-\frac{p}{2}\beta_i)\|_2^2+\frac{1-p}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j\|A(\beta_i-\beta_j)\|_2^2$$
$$=\left(1-\frac{p}{2}\right)^2\sum_{i=1}^{N}\lambda_i\|A\beta_i\|_2^2. \tag{4.15}$$

We deduce

$$\begin{aligned}
&\sum_{i=1}^{N}\lambda_i\|A(\sum_{j=1}^{N}\lambda_j\beta_j-\frac{p}{2}\beta_i)\|_2^2\\
=&\sum_{i=1}^{N}\lambda_i\left\|A[(1-\frac{p}{2}-\mu)(h_{\max(k)}+h_{(1)})-\frac{p}{2}\mu u_i+\mu h]\right\|_2^2\\
=&\sum_{i=1}^{N}\lambda_i\|A[(1-\frac{p}{2}-\mu)(h_{\max(k)}+h_{(1)})-\frac{p}{2}\mu u_i]\|_2^2\\
&+2\langle A[(1-\frac{p}{2}-\mu)(h_{\max(k)}+h_{(1)})-\frac{p}{2}\mu h_{(2)}],\mu Ah\rangle+\mu^2\|Ah\|_2^2\\
=&\sum_{i=1}^{N}\lambda_i\|A[(1-\frac{p}{2}-\mu)(h_{\max(k)}+h_{(1)})-\frac{p}{2}\mu u_i]\|_2^2\\
&+2\mu\langle A[(1-\frac{p}{2}-\mu)(h_{\max(k)}+h_{(1)})-\frac{p}{2}\mu h_{(2)}],Ah\rangle\\
&+p\mu^2\langle A(h_{\max(k)}+h_{(1)}+h_{(2)}),Ah\rangle+(1-p)\mu^2\|Ah\|_2^2\\
=&\sum_{i=1}^{N}\lambda_i\|A[(1-\frac{p}{2}-\mu)(h_{\max(k)}+h_{(1)})-\frac{p}{2}\mu u_i]\|_2^2\\
&+(2-p)\mu(1-\mu)\langle A(h_{\max(k)}+h_{(1)}),Ah\rangle+(1-p)\mu^2\|Ah\|_2^2.
\end{aligned}\tag{4.16}$$

Therefore,

$$\begin{aligned}
0=&\sum_{i=1}^{N}\lambda_i\|A[(1-\frac{p}{2}-\mu)(h_{\max(k)}+h_{(1)})-\frac{p}{2}\mu u_i]\|_2^2+(2-p)\mu(1-\mu)\langle A(h_{\max(k)}+h_{(1)}),Ah\rangle\\
&+(1-p)\mu^2\|Ah\|_2^2+\frac{1-p}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j\|A(\beta_i-\beta_j)\|_2^2-\left(1-\frac{p}{2}\right)^2\sum_{i=1}^{N}\lambda_i\|A\beta_i\|_2^2.
\end{aligned}\tag{4.17}$$

By virtue of the definition of RIC of order tk, the inequalities (4.9) and (4.10), we obtain

$$\sum_{i=1}^{N}\lambda_i\|A(\sum_{j=1}^{N}\lambda_j\beta_j-\frac{p}{2}\beta_i)\|_2^2\leq(1+\delta_{tk})\sum_{i=1}^{N}\lambda_i\left\|\left[\left(1-\frac{p}{2}-\mu\right)\left(h_{\max(k)}+h_{(1)}\right)-\frac{p}{2}\mu u_i\right]\right\|_2^2$$

$$
\begin{aligned}
&\quad + (2-p)\mu(1-\mu)\sqrt{1+\delta_{tk}}\left\|h_{\max(k)}+h_{(1)}\right\|_2(\varepsilon+\epsilon)+(1-p)\mu^2(\varepsilon+\epsilon)^2\\
&=(1+\delta_{tk})\left[\left(1-\frac{p}{2}-\mu\right)^2\left\|h_{\max(k)}+h_{(1)}\right\|_2^2+\frac{p^2}{4}\mu^2\sum_{i=1}^N\lambda_i\|u_i\|_2^2\right]\\
&\quad + (2-p)\mu(1-\mu)\sqrt{1+\delta_{tk}}\left\|h_{\max(k)}+h_{(1)}\right\|_2(\varepsilon+\epsilon)+(1-p)\mu^2(\varepsilon+\epsilon)^2.
\end{aligned}
$$

By virtue of the definition of RIC of order tk, we derive

$$
\begin{aligned}
&\sum_{i=1}^N\sum_{j=1}^N\lambda_i\lambda_j\|A(\beta_i-\beta_j)\|_2^2\\
=&\mu^2\sum_{i=1}^N\sum_{j=1}^N\lambda_i\lambda_j\|A(u_i-u_j)\|_2^2\\
\leq&\left(1+\delta_{tk}\right)\mu^2\sum_{i=1}^N\sum_{j=1}^N\lambda_i\lambda_j\|u_i-u_j\|_2^2\\
=&2\left(1+\delta_{tk}\right)\mu^2\left(\sum_{i=1}^N\lambda_i\|u_i\|_2^2-\|h_{(2)}\|_2^2\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\sum_{i=1}^N\lambda_i\|A\beta_i\|_2^2 &\geq (1-\delta_{tk})\sum_{i=1}^N\lambda_i\|h_{\max(k)}+h_{(1)}+\mu u_i\|_2^2\\
&=(1-\delta_{tk})\left(\left\|h_{\max(k)}+h_{(1)}\right\|_2^2+\mu^2\sum_{i=1}^N\lambda_i\|u_i\|_2^2\right).
\end{aligned}
$$

By (4.15), we deduce

$$
\begin{aligned}
0 &\leq (1+\delta_{tk})\Bigg[\left(1-\frac{p}{2}-\mu\right)^2\left\|h_{\max(k)}+h_{(1)}\right\|_2^2+\frac{p^2}{4}\mu^2\sum_{i=1}^N\lambda_i\|u_i\|_2^2\\
&\quad +\mu^2\left(1-p\right)\sum_{i=1}^N\lambda_i\|u_i\|_2^2-\mu^2(1-p)\|h_{(2)}\|_2^2\Bigg]\\
&\quad -(1-\delta_{tk})\left(1-\frac{p}{2}\right)^2\left(\left\|h_{\max(k)}+h_{(1)}\right\|_2^2+\mu^2\sum_{i=1}^N\lambda_i\|u_i\|_2^2\right)\\
&\quad +(2-p)\mu(1-\mu)(\varepsilon+\epsilon)\sqrt{1+\delta_{tk}}\left\|h_{\max(k)}+h_{(1)}\right\|_2+(1-p)\mu^2(\varepsilon+\epsilon)^2\\
&=(1+\delta_{tk})\left[\left(1-\frac{p}{2}-\mu\right)^2\left\|h_{\max(k)}+h_{(1)}\right\|_2^2-\mu^2(1-p)\|h_{(2)}\|_2^2\right]\\
&\quad -(1-\delta_{tk})\left(1-\frac{p}{2}\right)^2\left\|h_{\max(k)}+h_{(1)}\right\|_2^2+2\delta_{tk}\left(1-\frac{p}{2}\right)^2\mu^2\sum_{i=1}^N\lambda_i\|u_i\|_2^2\\
&\quad +(2-p)\mu(1-\mu)(\varepsilon+\epsilon)\sqrt{1+\delta_{tk}}\left\|h_{\max(k)}+h_{(1)}\right\|_2+(1-p)\mu^2(\varepsilon+\epsilon)^2,
\end{aligned}
\tag{4.18}
$$

and the estimate (4.14) for $\sum_{i=1}^N\lambda_i\|u_i\|_2^2$ yields

$$
0\leq\left[(1+\delta_{tk})\,(1-\frac{p}{2}-\mu)^2-(1-\delta_{tk})\left(1-\frac{p}{2}\right)^2\right]\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \tag{4.19}
$$

$$
\begin{aligned}
&+ (2-p)\mu(1-\mu)(\varepsilon+\epsilon)\sqrt{1+\delta_{tk}}\left\|h_{\max(k)}+h_{(1)}\right\|_2 + (1-p)\mu^2(\varepsilon+\epsilon)^2 \\
&+ \mu^2\left[2\delta_{tk}\left(1-\frac{p}{2}\right)^2 \frac{\left(\|h_{(2)}\|_2^2\right)^{\frac{2(1-p)}{2-p}}}{t-1}\left(\sigma^{\frac{2-p}{2}}\|h_{\max(k)}\|_2^p + \frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{2-p}}\right. \\
&\left. - (1-p)\left(1+\delta_{tk}\right)\|h_{(2)}\|_2^2\right].
\end{aligned}
$$

Denote $\vartheta = \left\|h_{(2)}\right\|_2^2$. By simple calculations, we derive that the following function

$$
2\delta_{tk}\left(1-\frac{p}{2}\right)^2 \frac{\vartheta^{\frac{2(1-p)}{2-p}}}{t-1}\left(\sigma^{\frac{2-p}{2}}\|h_{\max(k)}\|_2^p + \frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{2-p}} - (1-p)\left(1+\delta_{tk}\right)\vartheta \tag{4.20}
$$

for $\vartheta \geq 0$ attains its maximum at

$$
\vartheta = \left[\frac{(2-p)\,\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}}\left(\sigma^{\frac{2-p}{2}}\|h_{\max(k)}\|_2^p + \frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}}.
$$

Therefore, we deduce

$$
\begin{aligned}
&\left[\left(1+\delta_{tk}\right)\left(1-\frac{p}{2}-\mu\right)^2 - \left(1-\delta_{tk}\right)\left(1-\frac{p}{2}\right)^2\right]\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \\
&+ (2-p)\mu(1-\mu)\sqrt{1+\delta_{tk}}\left\|h_{\max(k)}+h_{(1)}\right\|_2(\varepsilon+\epsilon) + (1-p)\mu^2(\varepsilon+\epsilon)^2 \\
&+ \mu^2\left(\sigma^{\frac{2-p}{2}}\|h_{\max(k)}\|_2^p + \frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}} \cdot \left\{\frac{2\delta_{tk}\left(1-\frac{p}{2}\right)^2}{t-1}\left[\frac{(2-p)\,\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2(1-p)}{p}}\right. \\
&\left. -(1-p)\left(1+\delta_{tk}\right)\left[\frac{(2-p)\,\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}}\right\} \geq 0.
\end{aligned} \tag{4.21}
$$

Set $\mu = \frac{2-p}{p\eta^{\frac{2-p}{p}}+2}$. By (4.21), we obtain

$$
\begin{aligned}
&\left\{\left(\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\right)^2 - (2-p)\frac{2-p}{p\eta^{\frac{2-p}{p}}+2} + \delta_{tk}\left[\left(1-\frac{p}{2}-\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\right)^2 + \left(1-\frac{p}{2}\right)^2\right]\right. \\
&\left. + \frac{p}{2}\left(\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\right)^2\left(1+\delta_{tk}\right)\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}}\right\}\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \\
&+ (2-p)\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\left(1-\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\right)\sqrt{1+\delta_{tk}}\left\|h_{\max(k)}+h_{(1)}\right\|_2(\varepsilon+\epsilon) \\
&+ (1-p)\left(\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\right)^2(\varepsilon+\epsilon)^2 + \frac{p}{2}\left(\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\right)^2\left(1+\delta_{tk}\right)\left[\frac{(2-p)\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}} \\
&\cdot\left[\left(\sigma^{\frac{2-p}{2}}\left\|h_{\max(k)}+h_{(1)}\right\|_2^p + \frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}} - \sigma^{\frac{2-p}{p}}\left\|h_{\max(k)}+h_{(1)}\right\|_2^2\right] \geq 0.
\end{aligned} \tag{4.22}
$$

In view of $\delta(p,t,\sigma) = \frac{1}{p\eta^{\frac{2-p}{p}}+1}$, we deduce

$$
\left(p\eta^{\frac{2-p}{p}}+2\right)\left(1-\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\right) = \frac{1}{\delta(p,t,\sigma)} - 1 + p
$$

and

$$\begin{aligned}
&\left(\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\right)^2-(2-p)\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}+\delta_{tk}\left[\left(1-\frac{p}{2}-\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\right)^2+\left(1-\frac{p}{2}\right)^2\right]\\
&\quad+\frac{p}{2}\left(\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\right)^2(1+\delta_{tk})\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}}\\
=&\left(\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\right)^2\left\{-\left(p\eta^{\frac{2-p}{p}}+1\right)+\frac{\delta_{tk}}{2}\left[\left(p\eta^{\frac{2-p}{p}}+1\right)^2+1\right]\right.\\
&\quad\left.+\frac{p}{2}\left(1+\delta_{tk}\right)\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}}\right\}\\
=&-\left(\frac{2-p}{p\eta^{\frac{2-p}{p}}+2}\right)^2\left\{\frac{2\delta(p,t,\sigma)-\delta_{tk}\delta^2(p,t,\sigma)-\delta_{tk}}{2\delta^2(p,t,\sigma)}\right.\\
&\quad\left.-\frac{p}{2}\left(1+\delta_{tk}\right)\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}}\right\}.
\end{aligned}\tag{4.23}$$

When $\sigma=0$, it follows from Eq. (3.3) that $\eta=0$, and thus $\frac{(2-p)\sigma\delta_{tk}}{(t-1)(1+\delta_{tk})}=\eta$; when $\sigma>0$, it follows from (3.3) that $\frac{(2-p)\sigma}{t-1}=p\eta^{\frac{2}{p}}+2\eta$, and thus the RIC condition $\delta_{tk}<\frac{1}{p\eta^{\frac{2-p}{p}}+1}$ yields

$$\frac{(2-p)\sigma\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}=\left(p\eta^{\frac{2}{p}}+2\eta\right)\frac{\delta_{tk}}{1+\delta_{tk}}<\eta. \tag{4.24}$$

Hence,

$$\begin{aligned}
\beta(p,t,\sigma):=&\frac{2\delta(p,t,\sigma)-\left(1+\delta^2(p,t,\sigma)\right)\delta_{tk}}{2\delta^2(p,t,\sigma)}-\frac{p}{2}\left(1+\delta_{tk}\right)\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}}\\
\geq&\frac{2\delta(p,t,\sigma)-\left(1+\delta^2(p,t,\sigma)\right)\delta_{tk}}{2\delta^2(p,t,\sigma)}-\frac{p}{2}\left(1+\delta_{tk}\right)\eta^{\frac{2-p}{p}}\\
=&\frac{2\delta(p,t,\sigma)-\left(1+\delta^2(p,t,\sigma)\right)\delta_{tk}}{2\delta^2(p,t,\sigma)}-\frac{1}{2}\left(1+\delta_{tk}\right)\left(\frac{1}{\delta(p,t,\sigma)}-1\right)\\
=&\frac{1}{2\delta(p,t,\sigma)}\left(\frac{1}{\delta(p,t,\sigma)}+1\right)\left(\delta(p,t,\sigma)-\delta_{tk}\right)>0.
\end{aligned}$$

Therefore, the inequality (4.22) is equivalent to

$$\begin{aligned}
&\beta(p,t,\sigma)\left\|h_{\max(k)}+h_{(1)}\right\|_2^2-\left(\frac{1}{\delta(p,t,\sigma)}-1+p\right)\sqrt{1+\delta_{tk}}\left(\varepsilon+\epsilon\right)\left\|h_{\max(k)}+h_{(1)}\right\|_2\\
&\quad-(1-p)(\varepsilon+\epsilon)^2+\frac{p}{2}\left(1+\delta_{tk}\right)\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}}\left\|h_{\max(k)}+h_{(1)}\right\|_2^2\\
&\quad-\frac{p\left(1+\delta_{tk}\right)}{2}\left[\frac{(2-p)\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}}\left(\sigma^{\frac{2-p}{2}}\left\|h_{\max(k)}+h_{(1)}\right\|_2^p\right.\\
&\quad\left.+\frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}}\leq 0.
\end{aligned}\tag{4.25}$$

(i) When $\sigma\delta_{tk} \neq 0$, (4.25) is transformed to

$$\begin{aligned}
&\lambda\beta(p,t,\sigma)\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \\
&\quad-\left(\frac{1}{\delta(p,t,\sigma)}-1+p\right)\sqrt{1+\delta_{tk}}\,(\varepsilon+\epsilon)\left\|h_{\max(k)}+h_{(1)}\right\|_2-(1-p)\,(\varepsilon+\epsilon)^2 \\
&\quad+\left\{(1-\lambda)\beta(p,t,\sigma)+\frac{p}{2}\,(1+\delta_{tk})\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{p}}\right\}\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \\
&\quad-\frac{p\,(1+\delta_{tk})}{2}\left[\frac{(2-p)\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{p}}\left(\sigma^{\frac{2-p}{2}}\right\|h_{\max(k)} \\
&\quad+h_{(1)}\left\|_2^p+\frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}}\le 0 \qquad (4.26)
\end{aligned}$$

for any $\lambda \in (0,1)$.

Therefore, we derive

$$\left\|h_{\max(k)}+h_{(1)}\right\|_2 \le D_1(\varepsilon+\epsilon)+D_2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p,$$

where

$$D_1=\frac{\left(\frac{1}{\delta(p,t,\sigma)}-1+p\right)\sqrt{1+\delta_{tk}}+\sqrt{\left(\frac{1}{\delta(p,t,\sigma)}-1+p\right)^2(1+\delta_{tk})+4\lambda(1-p)\beta(p,t,\sigma)}}{2\lambda\beta(p,t,\sigma)} \qquad (4.27)$$

and

$$D_2=\left[\frac{2}{(k\sigma)^{\frac{2-p}{2}}}\right]^{\frac{1}{p}}\left\{\left\{\frac{2(1-\lambda)\beta(p,t,\sigma)}{p\,(1+\delta_{tk})\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)(1+\delta_{tk})}\right]^{\frac{2-p}{p}}}+1\right\}^{\frac{p}{2}}-1\right\}^{-\frac{1}{p}}. \qquad (4.28)$$

(ii) When $\sigma = 0$, it follows from Eq. (3.3) that $\eta = 0$. Therefore,

$$\delta(p,t,\sigma)=\frac{1}{p\eta^{\frac{2-p}{p}}+1}=1$$

and

$$\beta(p,t,\sigma)=\frac{2\delta(p,t,\sigma)-\left(1+\delta^2(p,t,\sigma)\right)\delta_{tk}}{2\delta^2(p,t,\sigma)}-\frac{p}{2}\,(1+\delta_{tk})\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{p}}=1-\delta_{tk}.$$

The inequality (4.25) becomes

$$\begin{aligned}
&(1-\delta_{tk})\left\|h_{\max(k)}+h_{(1)}\right\|_2^2-p\sqrt{1+\delta_{tk}}\,(\varepsilon+\epsilon)\left\|h_{\max(k)}+h_{(1)}\right\|_2 \qquad (4.29)\\
&\quad-(1-p)(\varepsilon+\epsilon)^2-\frac{p}{2}\,(1+\delta_{tk})\left[\frac{(2-p)\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{p}}\left(\frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}}\le 0,
\end{aligned}$$

which is a second-order inequality for $\left\|h_{\max(k)}+h_{(1)}\right\|_2$. By solving the inequality, we obtain

$$\begin{aligned}
&\left\|h_{\max(k)}+h_{(1)}\right\|_2 \qquad (4.30)\\
&\le\frac{p\sqrt{1+\delta_{tk}}\,(\varepsilon+\epsilon)}{2\,(1-\delta_{tk})}+\frac{1}{2\,(1-\delta_{tk})}\sqrt{p^2\,(1+\delta_{tk})\,(\varepsilon+\epsilon)^2+4\,(1-\delta_{tk})\left[(1-p)(\varepsilon+\epsilon)^2+B\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^2\right]}
\end{aligned}$$

with $B := \frac{p}{2}\left(\frac{2}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}}(1+\delta_{tk})\left[\frac{(2-p)\delta_{tk}}{(t-1)(1+\delta_{tk})}\right]^{\frac{2-p}{p}}$.

Since

$$\sqrt{\left[p\sqrt{1+\delta_{tk}}\,(\varepsilon+\epsilon)\right]^2+4\,(1-\delta_{tk})\left[(1-p)(\varepsilon+\epsilon)^2+B\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^2\right]}$$
$$=\sqrt{\left[p^2\,(1+\delta_{tk})+4\,(1-\delta_{tk})\,(1-p)\right](\varepsilon+\epsilon)^2+4\,(1-\delta_{tk})\,B\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^2}$$
$$\leq\sqrt{p^2\,(1+\delta_{tk})+4\,(1-\delta_{tk})\,(1-p)}(\varepsilon+\epsilon)+2\sqrt{(1-\delta_{tk})\,B}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p,$$

we derive

$$\left\|h_{\max(k)}+h_{(1)}\right\|_2\leq D_1(\varepsilon+\epsilon)+D_2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p,$$

where D_1 and D_2 are given by (3.7).

(iii) When $\delta_{tk}=0$, we have

$$\beta(p,t,\sigma)=\frac{2\delta(p,t,\sigma)-\left(1+\delta^2(p,t,\sigma)\right)\delta_{tk}}{2\delta^2(p,t,\sigma)}-\frac{p}{2}\,(1+\delta_{tk})\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{p}}=\frac{1}{\delta(p,t,\sigma)}.$$

The inequality (4.25) becomes

$$\frac{1}{\delta(p,t,\sigma)}\left\|h_{\max(k)}+h_{(1)}\right\|_2^2-\left(\frac{1}{\delta(p,t,\sigma)}-1+p\right)(\varepsilon+\epsilon)\left\|h_{\max(k)}+h_{(1)}\right\|_2$$
$$-(1-p)(\varepsilon+\epsilon)^2\leq 0. \tag{4.31}$$

We obtain $\left\|h_{\max(k)}+h_{(1)}\right\|_2\leq\varepsilon+\epsilon$, which can be rewritten as

$$\left\|h_{\max(k)}+h_{(1)}\right\|_2\leq D_1(\varepsilon+\epsilon)+D_2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\quad(D_1=1,\ D_2=0).$$

For (i), (ii) and (iii), in view of (4.5) and $|T_0\cup\tilde{T}\setminus(\tilde{T}\cap T_0)|=k+\rho k-2\alpha\rho k\leq k$ for $\alpha\in[\frac{1}{2},1]$, we obtain $\|h_{T_0\cup\tilde{T}\setminus(\tilde{T}\cap T_0)}\|_p^p\leq\left\|h_{\max(k)}\right\|_p^p$, and thus

$$\left\|h_{-\max(k)}\right\|_p^p\leq w^p\left\|h_{\max(k)}\right\|_p^p+(1-w^p)\|h_{T_0\cup\tilde{T}\setminus(\tilde{T}\cap T_0)}\|_p^p+2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p$$
$$\leq\left\|h_{\max(k)}\right\|_p^p+2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p. \tag{4.32}$$

By Lemma 2.4, we derive

$$\left\|h_{-\max(k)}\right\|_2^2\leq k\left[\left(\frac{\left\|h_{\max(k)}\right\|_2^2}{k}\right)^{\frac{p}{2}}+\frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k}\right]^{\frac{2}{p}}$$
$$=\left(\left\|h_{\max(k)}\right\|_2^p+\frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{1-\frac{p}{2}}}\right)^{\frac{2}{p}}. \tag{4.33}$$

Then it follows from Jensen inequality that

$$\left\|h_{-\max(k)}\right\|_2\leq\left(\left\|h_{\max(k)}\right\|_2^p+\frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}$$
$$\leq 2^{\frac{1}{p}-1}\left[\left\|h_{\max(k)}\right\|_2+\left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\right]. \tag{4.34}$$

Therefore,

$$
\begin{aligned}
&\|\hat{x}^{\ell_2} - x\|_2^2 = \|h\|_2^2 = \left\|h_{\max(k)}\right\|_2^2 + \left\|h_{-\max(k)}\right\|_2^2 \\
&\le \left\|h_{\max(k)}\right\|_2^2 + 2^{\frac{2}{p}-2}\left[\left\|h_{\max(k)}\right\|_2 + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\right]^2 \\
&\le \left[D_1(\varepsilon+\epsilon) + D_2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\right]^2 \\
&\quad + 2^{\frac{2}{p}-2}\left[D_1(\varepsilon+\epsilon) + D_2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\right]^2, \qquad (4.35)
\end{aligned}
$$

and thus

$$
\|\hat{x}^{\ell_2} - x\|_2 \le C_1(\varepsilon+\epsilon) + C_2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p, \tag{4.36}
$$

where

$$
C_1 = \sqrt{1+2^{\frac{2}{p}-2}}D_1, \quad C_2 = \sqrt{D_2^2 + 2^{\frac{2}{p}-2}\left[D_2 + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\right]^2}.
$$

The proof of Theorem 3.1 is therefore completed. □

Remark 4.1. (1) In particular, in the noiseless case, the coefficient C_2 can be more precisely characterized in that $\sqrt{D_2^2 + 2^{\frac{2}{p}-2}[D_2 + (\frac{2}{k^{1-\frac{p}{2}}})^{\frac{1}{p}}]^2}$ can be replaced by $\sqrt{D_2^2 + (D_2^p + \frac{2}{k^{1-\frac{p}{2}}})^{\frac{2}{p}}}$. Actually, in the noiseless case, $\|h_{\max(k)} + h_{(1)}\|_2 \le D_2\|\mathrm{w}\circ x_{-\max(k)}\|_p$, and thus

$$
\begin{aligned}
&\|\hat{x}^{\ell_2} - x\|_2^2 = \|h\|_2^2 = \left\|h_{\max(k)}\right\|_2^2 + \left\|h_{-\max(k)}\right\|_2^2 \\
&\le \left\|h_{\max(k)}\right\|_2^2 + \left(\left\|h_{\max(k)}\right\|_2^p + \frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{1-\frac{p}{2}}}\right)^{\frac{2}{p}} \\
&\le D_2^2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^2 + \left[D_2^p\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p + \frac{2}{k^{1-\frac{p}{2}}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p\right]^{\frac{2}{p}} \\
&= \left[D_2^2 + \left(D_2^p + \frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{2}{p}}\right]\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^2.
\end{aligned}
$$

Therefore, in the noiseless case,

$$
\|\hat{x}^{\ell_2} - x\|_2 \le \sqrt{D_2^2 + \left(D_2^p + \frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{2}{p}}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p.
$$

(2) In the proof of Theorem 3.1, we employ the estimation (4.34) of $\left\|h_{-\max(k)}\right\|_2$ based on the following comparison:

By (4.33) and Jensen inequality, we deduce

$$
\begin{aligned}
\left\|h_{-\max(k)}\right\|_2^2 &\le \left(\left\|h_{\max(k)}\right\|_2^p + \frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{1-\frac{p}{2}}}\right)^{\frac{2}{p}} \\
&\le 2^{\frac{2}{p}-1}\left[\left\|h_{\max(k)}\right\|_2^2 + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{2}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^2\right]. \qquad (4.37)
\end{aligned}
$$

On the other hand, by (4.34), i.e.,

$$\begin{aligned}\left\|h_{-\max(k)}\right\|_2 \leq & \left(\left\|h_{\max(k)}\right\|_2^p + \frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}} \\ & \leq 2^{\frac{1}{p}-1}\left[\left\|h_{\max(k)}\right\|_2 + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\right],\end{aligned}$$

we obtain

$$\left\|h_{-\max(k)}\right\|_2^2 \leq 2^{\frac{2}{p}-2}\left[\left\|h_{\max(k)}\right\|_2 + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\right]^2. \tag{4.38}$$

Now we compare the two different estimations (4.37) and (4.38) of $\left\|h_{-\max(k)}\right\|_2^2$. Since

$$\left\|h_{\max(k)}\right\|_2\left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p \leq \frac{1}{2}\left[\left\|h_{\max(k)}\right\|_2^2 + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{2}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^2\right],$$

we derive

$$\begin{aligned}& 2^{\frac{2}{p}-2}\left[\left\|h_{\max(k)}\right\|_2 + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\right]^2 \\ = & \ 2^{\frac{2}{p}-2}\left[\left\|h_{\max(k)}\right\|_2^2 + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{2}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^2 + 2\left\|h_{\max(k)}\right\|_2\left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\right] \\ \leq & \ 2^{\frac{2}{p}-1}\left[\left\|h_{\max(k)}\right\|_2^2 + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{2}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^2\right].\end{aligned}$$

Therefore, we conclude that the estimation (4.38) is more precise than another estimation (4.37), and thus the estimation (4.34) of $\left\|h_{-\max(k)}\right\|_2$ is employed, i.e.,

$$\left\|h_{-\max(k)}\right\|_2 \leq 2^{\frac{1}{p}-1}\left[\left\|h_{\max(k)}\right\|_2 + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\right].$$

4.2. Proof of Corollary 3.1

When in particular $\mathrm{w}\circ x_{-\max(k)} = \mathbf{0}$, the inequality (4.25) in the proof of Theorem 3.1 becomes

$$\begin{aligned}\beta(p,t,\sigma)\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 - \left(\frac{1}{\delta(p,t,\sigma)} - 1 + p\right)\sqrt{1+\delta_{tk}}\,(\varepsilon+\epsilon)\left\|h_{\max(k)}+h_{(1)}\right\|_2 \\ - (1-p)(\varepsilon+\epsilon)^2 \leq 0.\end{aligned} \tag{4.39}$$

By solving the inequality, we obtain

$$\left\|h_{\max(k)}+h_{(1)}\right\|_2 \leq D_3(\varepsilon+\epsilon),$$

where

$$D_3 = \frac{\left(\frac{1}{\delta(p,t,\sigma)} - 1 + p\right)\sqrt{1+\delta_{tk}} + \sqrt{\left(\frac{1}{\delta(p,t,\sigma)} - 1 + p\right)^2(1+\delta_{tk}) + 4(1-p)\beta(p,t,\sigma)}}{2\beta(p,t,\sigma)}. \tag{4.40}$$

Since $\mathrm{w} \circ x_{-\max(k)} = \mathbf{0}$, the inequality (4.33) in the proof of Theorem 3.1 becomes

$$\left\|h_{-\max(k)}\right\|_2^2 \leq \left\|h_{\max(k)}\right\|_2^2.$$

Therefore,

$$\left\|\hat{x}^{\ell_2} - x\right\|_2 = \sqrt{\left\|h_{\max(k)}\right\|_2^2 + \left\|h_{-\max(k)}\right\|_2^2} \leq \sqrt{2}\left\|h_{\max(k)}\right\|_2 \leq \sqrt{2} D_3(\varepsilon + \epsilon). \tag{4.41}$$

This completes the proof of Corollary 3.1. □

4.3. Proof of Corollary 3.2

Denote $T_0 = \mathrm{supp}(x_{\max(k)})$. Since the signal $x \in \mathbb{R}^n$ is k-sparse and $T = \mathrm{supp}(x)$, we derive $T_0 = T$ and $x_{-\max(k)} = \mathbf{0}$, and thus $\mathrm{w} \circ x_{-\max(k)} = \mathbf{0}$. Then the result directly follows from Corollary 3.1. □

4.4. Proof of Theorem 3.2

Denote $h = \hat{x} - x$. In the noiseless setting, $\epsilon = 0$ and $\varepsilon = 0$. Therefore, the inequality (4.25) in the proof of Theorem 3.1 becomes

$$\begin{aligned} &\frac{2\delta(p,t,\sigma) - \left(1 + \delta^2(p,t,\sigma)\right)\delta_{tk}}{2\delta^2(p,t,\sigma)} \left\|h_{\max(k)} + h_{(1)}\right\|_2^2 \\ &\quad - \frac{p\,(1+\delta_{tk})}{2}\left[\frac{(2-p)\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{p}} \left(\sigma^{\frac{2-p}{2}}\left\|h_{\max(k)}\right.\right. \\ &\quad \left.\left. + h_{(1)}\right\|_2^p + \frac{2\left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}} \leq 0. \end{aligned} \tag{4.42}$$

(i) When $\sigma\delta_{tk} \neq 0$, we obtain $\left\|h_{\max(k)} + h_{(1)}\right\|_2 \leq D_4 \left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p$, where

$$D_4 = \left[\frac{2}{(k\sigma)^{\frac{2-p}{2}}}\right]^{\frac{1}{p}} \left\{\left\{\frac{2\delta(p,t,\sigma) - \left(1+\delta^2(p,t,\sigma)\right)\delta_{tk}}{p\,(1+\delta_{tk})\,\delta^2(p,t,\sigma)\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)(1+\delta_{tk})}\right]^{\frac{2-p}{p}}}\right\}^{\frac{p}{2}} - 1\right\}^{-\frac{1}{p}}. \tag{4.43}$$

(ii) When $\sigma = 0$, it follows from the RIC condition (3.2) that $\delta(p,t,\sigma) = 1$. The inequality (4.42) becomes

$$\begin{aligned} &(1-\delta_{tk})\left\|h_{\max(k)} + h_{(1)}\right\|_2^2 \\ &\quad - \frac{p\,(1+\delta_{tk})}{2}\left[\frac{(2-p)\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{p}} \left[\frac{2\left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right]^{\frac{2}{p}} \leq 0, \end{aligned} \tag{4.44}$$

and we obtain $\left\|h_{\max(k)} + h_{(1)}\right\|_2 \leq D_4 \left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p$, where

$$D_4 = \sqrt{\frac{p\,(1+\delta_{tk})}{2\,(1-\delta_{tk})}}\left[\frac{(2-p)\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{2p}} \left(\frac{2}{k^{\frac{2-p}{2}}}\right)^{\frac{1}{p}}. \tag{4.45}$$

(iii) When $\delta_{tk} = 0$, the inequality (4.42) becomes $\frac{1}{\delta(p,t,\sigma)} \left\|h_{\max(k)} + h_{(1)}\right\|_2^2 \leq 0$. Thus, $\left\|h_{\max(k)} + h_{(1)}\right\|_2 = 0$, which can be equivalently written as $\left\|h_{\max(k)} + h_{(1)}\right\|_2 \leq D_4 \|\mathrm{w} \circ x_{-\max(k)}\|_p$ with $D_4 = 0$.

Therefore, for (i), (ii) and (iii), by (4.33) in the proof of Theorem 3.1, we deduce

$$\begin{aligned}\left\|h_{-\max(k)}\right\|_2 &\leq \left(\left\|h_{\max(k)}\right\|_2^p + \frac{2\left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p^p}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}} \\ &\leq \left(D_4^p + \frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}} \left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p, \end{aligned} \tag{4.46}$$

and thus

$$\begin{aligned}\|\hat{x}^{\ell_2} - x\|_2 =& \sqrt{\left\|h_{\max(k)}\right\|_2^2 + \left\|h_{-\max(k)}\right\|_2^2} \\ \leq& \sqrt{D_4^2 + \left(D_4^p + \frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{2}{p}}} \left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p. \end{aligned} \tag{4.47}$$

□

4.5. Proof of Theorem 3.3

Denote $h = \hat{x}^{DS} - x$. We have

$$\begin{aligned}\left\|A^T A h\right\|_\infty =& \left\|A^T(A\hat{x}^{DS} - Ax)\right\|_\infty \\ \leq& \left\|A^T(A\hat{x}^{DS} - y)\right\|_\infty + \left\|A^T(Ax - y)\right\|_\infty \leq \varepsilon + \epsilon. \end{aligned} \tag{4.48}$$

The same as in the proof of Theorem 3.1, we derive the inequality (4.5), i.e.,

$$\left\|h_{-\max(k)}\right\|_p^p \leq w^p \left\|h_{\max(k)}\right\|_p^p + (1 - w^p)\|h_{T_0 \cup \tilde{T} \setminus (\tilde{T} \cap T_0)}\|_p^p + 2\left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p^p.$$

Then it follows from $|T_0 \cup \tilde{T} \setminus (\tilde{T} \cap T_0)| = k + \rho k - 2\alpha\rho k \leq k$ for $\alpha \in [\frac{1}{2}, 1]$ that $\|h_{T_0 \cup \tilde{T} \setminus (\tilde{T} \cap T_0)}\|_p^p \leq \left\|h_{\max(k)}\right\|_p^p$, and thus

$$\left\|h_{-\max(k)}\right\|_p^p \leq \left\|h_{\max(k)}\right\|_p^p + 2\left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p^p.$$

By Lemma 2.4 and Jensen inequality, we derive

$$\begin{aligned}\left\|h_{-\max(k)}\right\|_1 \leq& k\left[\left(\frac{\left\|h_{\max(k)}\right\|_1}{k}\right)^p + \frac{2\left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p^p}{k}\right]^{\frac{1}{p}} \\ \leq& 2^{\frac{1}{p}-1} k \left[\frac{\left\|h_{\max(k)}\right\|_1}{k} + \left(\frac{2}{k}\right)^{\frac{1}{p}} \left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p\right] \\ =& 2^{\frac{1}{p}-1}\left\|h_{\max(k)}\right\|_1 + \frac{2^{\frac{2}{p}-1}}{k^{\frac{1}{p}-1}} \left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p. \end{aligned} \tag{4.49}$$

Therefore,

$$\begin{aligned}\|h\|_1 &= \left\|h_{\max(k)}\right\|_1 + \left\|h_{-\max(k)}\right\|_1 \\ &\leq \left(1 + 2^{\frac{1}{p}-1}\right) \left\|h_{\max(k)}\right\|_1 + k^{1-\frac{1}{p}} 2^{\frac{2}{p}-1} \left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p \\ &\leq \left(1 + 2^{\frac{1}{p}-1}\right) \sqrt{k} \left\|h_{\max(k)}\right\|_2 + k^{1-\frac{1}{p}} 2^{\frac{2}{p}-1} \left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p. \end{aligned} \tag{4.50}$$

We have

$$\begin{aligned}\|Ah\|_2^2 =&\langle h, A^T Ah\rangle \leq \|h\|_1 \left\|A^T Ah\right\|_\infty \\ &\leq \left[\left(1+2^{\frac{1}{p}-1}\right)\sqrt{k}\left\|h_{\max(k)}\right\|_2 + k^{1-\frac{1}{p}}2^{\frac{2}{p}-1}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\right](\varepsilon+\epsilon) \end{aligned}\tag{4.51}$$

and

$$\begin{aligned}&\langle A(h_{\max(k)}+h_{(1)}), Ah\rangle = \langle h_{\max(k)}+h_{(1)}, A^T Ah\rangle \\ \leq &\left\|h_{\max(k)}+h_{(1)}\right\|_1 \left\|A^T Ah\right\|_\infty \leq \sqrt{tk}\left\|h_{\max(k)}+h_{(1)}\right\|_2(\varepsilon+\epsilon).\end{aligned}\tag{4.52}$$

Similar to the proof of Theorem 3.1, for

$$\mu = \frac{2-p}{p\eta^{\frac{2-p}{p}}+2} = (2-p)\frac{\delta(p,t,\sigma)}{\delta(p,t,\sigma)+1},$$

by virtue of the identity (4.17), we obtain

$$\begin{aligned}&(1-p)\|h\|_1(\varepsilon+\epsilon) + \frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\langle A(h_{\max(k)}+h_{(1)}), Ah\rangle \\ &+\frac{\delta_{tk}\delta^2(p,t,\sigma)+\delta_{tk}-2\delta(p,t,\sigma)}{2\delta^2(p,t,\sigma)}\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \\ &+\frac{p(1+\delta_{tk})}{2}\left[\frac{(2-p)\delta_{tk}}{(t-1)(1+\delta_{tk})}\right]^{\frac{2-p}{p}}\left(\sigma^{\frac{2-p}{2}}\left\|h_{\max(k)}+h_{(1)}\right\|_2^p + \frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}} \geq 0,\end{aligned}\tag{4.53}$$

and then

$$\begin{aligned}&(1-p)\left[\left(1+2^{\frac{1}{p}-1}\right)\sqrt{k}\left\|h_{\max(k)}+h_{(1)}\right\|_2 + \frac{2^{\frac{2}{p}-1}}{k^{\frac{1}{p}-1}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\right](\varepsilon+\epsilon) \\ &+\frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{tk}(\varepsilon+\epsilon)\left\|h_{\max(k)}+h_{(1)}\right\|_2 \\ &+\frac{\delta_{tk}\delta^2(p,t,\sigma)+\delta_{tk}-2\delta(p,t,\sigma)}{2\delta^2(p,t,\sigma)}\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \\ &+\frac{p(1+\delta_{tk})}{2}\left[\frac{(2-p)\delta_{tk}}{(t-1)(1+\delta_{tk})}\right]^{\frac{2-p}{p}}\left(\sigma^{\frac{2-p}{2}}\left\|h_{\max(k)}+h_{(1)}\right\|_2^p + \frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}} \geq 0,\end{aligned}$$

i.e.,

$$\begin{aligned}&\frac{\delta_{tk}\delta^2(p,t,\sigma)+\delta_{tk}-2\delta(p,t,\sigma)}{2\delta^2(p,t,\sigma)}\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \\ &+\left[(1-p)\left(1+2^{\frac{1}{p}-1}\right)+\frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}\right]\sqrt{k}(\varepsilon+\epsilon)\left\|h_{\max(k)}+h_{(1)}\right\|_2 \\ &+(1-p)k^{1-\frac{1}{p}}2^{\frac{2}{p}-1}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p(\varepsilon+\epsilon) \\ &+\frac{p(1+\delta_{tk})}{2}\left[\frac{(2-p)\delta_{tk}}{(t-1)(1+\delta_{tk})}\right]^{\frac{2-p}{p}}\left(\sigma^{\frac{2-p}{2}}\left\|h_{\max(k)}+h_{(1)}\right\|_2^p + \frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}} \geq 0.\end{aligned}\tag{4.54}$$

Therefore, by the notation (3.6) of $\beta(p,t,\sigma)$, we have

$$\begin{aligned}
&\beta(p,t,\sigma)\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \\
&-\left[(1-p)\left(1+2^{\frac{1}{p}-1}\right)+\frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}\right]\sqrt{k}\,(\varepsilon+\epsilon)\left\|h_{\max(k)}+h_{(1)}\right\|_2 \\
&-(1-p)\,k^{1-\frac{1}{p}}2^{\frac{2}{p}-1}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p(\varepsilon+\epsilon) \\
&+\frac{p\,(1+\delta_{tk})}{2}\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{p}}\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \\
&-\frac{p\,(1+\delta_{tk})}{2}\left[\frac{(2-p)\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{p}}\left(\sigma^{\frac{2-p}{2}}\left\|h_{\max(k)}+h_{(1)}\right\|_2^p+\frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}}\le 0.
\end{aligned} \tag{4.55}$$

(i) When $\sigma\delta_{tk}\neq 0$, for any $\lambda\in(0,1)$,

$$\begin{aligned}
&\lambda\beta(p,t,\sigma)\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \\
&\quad-\left[(1-p)\left(1+2^{\frac{1}{p}-1}\right)+\frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}\right]\sqrt{k}\,(\varepsilon+\epsilon)\left\|h_{\max(k)}+h_{(1)}\right\|_2 \\
&\quad-(1-p)\,k^{1-\frac{1}{p}}2^{\frac{2}{p}-1}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p(\varepsilon+\epsilon) \\
&\quad+\left\{(1-\lambda)\beta(p,t,\sigma)+\frac{p\,(1+\delta_{tk})}{2}\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{p}}\right\}\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \\
&\quad-\frac{p\,(1+\delta_{tk})}{2}\left[\frac{(2-p)\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{p}}\left(\sigma^{\frac{2-p}{2}}\left\|h_{\max(k)}+h_{(1)}\right\|_2^p+\frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}}\le 0.
\end{aligned}$$

Therefore, we deduce

$$\begin{aligned}
&\left\|h_{\max(k)}+h_{(1)}\right\|_2 \\
&\le\max\left\{\frac{\left[(1-p)\left(1+2^{\frac{1}{p}-1}\right)+\frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}\right]\sqrt{k}\,(\varepsilon+\epsilon)+\sqrt{E}}{2\lambda\beta(p,t,\sigma)},\,D_2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\right\},
\end{aligned} \tag{4.56}$$

where D_2 is defined in (3.5) and

$$\begin{aligned}
E:=&\left[(1-p)\left(1+2^{\frac{1}{p}-1}\right)+\frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}\right]^2 k\,(\varepsilon+\epsilon)^2 \\
&+4\lambda\beta(p,t,\sigma)\,(1-p)\,k^{1-\frac{1}{p}}2^{\frac{2}{p}-1}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p(\varepsilon+\epsilon)\,.
\end{aligned} \tag{4.57}$$

Since

$$\begin{aligned}
\sqrt{E}\le&\left[(1-p)\left(1+2^{\frac{1}{p}-1}\right)+\frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}\right]\sqrt{k}\,(\varepsilon+\epsilon) \\
&+\frac{2\lambda\beta(p,t,\sigma)\,(1-p)\,k^{\frac{1}{2}-\frac{1}{p}}2^{\frac{2}{p}-1}}{(1-p)\left(1+2^{\frac{1}{p}-1}\right)+\frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p,
\end{aligned} \tag{4.58}$$

we obtain

$$\left\|h_{\max(k)}+h_{(1)}\right\|_2\le D_5(\varepsilon+\epsilon)+D_6\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p,$$

where

$$D_5=\frac{(1-p)\left(1+2^{\frac{1}{p}-1}\right)+\frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}}{\lambda\beta(p,t,\sigma)}\sqrt{k} \tag{4.59}$$

and

$$D_6 = \max\left\{\frac{(1-p)\,k^{\frac{1}{2}-\frac{1}{p}}2^{\frac{2}{p}-1}}{(1-p)\left(1+2^{\frac{1}{p}-1}\right)+\frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}},\,D_2\right\}. \tag{4.60}$$

(ii) When $\sigma = 0$, it follows from (3.3) that $\eta = 0$, and thus

$$\delta(p,t,\sigma) = \frac{1}{p\eta^{\frac{2-p}{p}}+1} = 1.$$

We have

$$\beta(p,t,\sigma) = \frac{2\delta(p,t,\sigma)-\left(1+\delta^2(p,t,\sigma)\right)\delta_{tk}}{2\delta^2(p,t,\sigma)} - \frac{p}{2}\left(1+\delta_{tk}\right)\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}} = 1-\delta_{tk}.$$

Therefore, the inequality (4.55) becomes

$$\begin{aligned}
&\left(1-\delta_{tk}\right)\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 - \left[(1-p)\left(1+2^{\frac{1}{p}-1}\right)+p\sqrt{t}\right]\sqrt{k}\,(\varepsilon+\epsilon)\left\|h_{\max(k)}+h_{(1)}\right\|_2\\
&\quad-(1-p)\,k^{1-\frac{1}{p}}2^{\frac{2}{p}-1}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p(\varepsilon+\epsilon)\\
&\quad-\frac{p}{2}\left(1+\delta_{tk}\right)\left[\frac{(2-p)\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}}\left(\frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}} \le 0.
\end{aligned} \tag{4.61}$$

By solving the above inequality, we obtain

$$\left\|h_{\max(k)}+h_{(1)}\right\|_2 \le \frac{\left[(1-p)\left(1+2^{\frac{1}{p}-1}\right)+p\sqrt{t}\right]\sqrt{k}\,(\varepsilon+\epsilon)+\sqrt{F}}{2\left(1-\delta_{tk}\right)}, \tag{4.62}$$

where

$$\begin{aligned}
F :=& \left[(1-p)\left(1+2^{\frac{1}{p}-1}\right)+p\sqrt{t}\right]^2 k\,(\varepsilon+\epsilon)^2\\
&+4\left(1-\delta_{tk}\right)\left\{(1-p)\,k^{1-\frac{1}{p}}2^{\frac{2}{p}-1}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p(\varepsilon+\epsilon)\right.\\
&\left.+\frac{p}{2}\left(1+\delta_{tk}\right)\left[\frac{(2-p)\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}}\left(\frac{2\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^p}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}}\right\}\\
=&\left[(1-p)\left(1+2^{\frac{1}{p}-1}\right)+p\sqrt{t}\right]^2 k\,(\varepsilon+\epsilon)^2\\
&+2\left(1-\delta_{tk}\right)(1-p)\,k^{1-\frac{1}{p}}2^{\frac{2}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p(\varepsilon+\epsilon)\\
&+2p\left(1-\delta_{tk}^2\right)\left[\frac{(2-p)\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{p}}\left(\frac{2}{k^{\frac{2-p}{2}}}\right)^{\frac{2}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p^2.
\end{aligned} \tag{4.63}$$

We derive

$$\begin{aligned}
\sqrt{F} \le& \left[(1-p)\left(1+2^{\frac{1}{p}-1}\right)+p\sqrt{t}\right]\sqrt{k}\,(\varepsilon+\epsilon)+\frac{\left(1-\delta_{tk}\right)(1-p)\,k^{1-\frac{1}{p}}2^{\frac{2}{p}}}{\left[(1-p)\left(1+2^{\frac{1}{p}-1}\right)+p\sqrt{t}\right]\sqrt{k}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p\\
&+\sqrt{2p\left(1-\delta_{tk}^2\right)}\left[\frac{(2-p)\delta_{tk}}{(t-1)\left(1+\delta_{tk}\right)}\right]^{\frac{2-p}{2p}}2^{\frac{1}{p}}k^{\frac{1}{2}-\frac{1}{p}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p,
\end{aligned} \tag{4.64}$$

and therefore

$$\left\|h_{\max(k)} + h_{(1)}\right\|_2 \le D_5(\varepsilon + \epsilon) + D_6 \left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p,$$

where

$$D_5 = \frac{(1-p)\left(1 + 2^{\frac{1}{p}-1}\right) + p\sqrt{t}}{1-\delta_{tk}}\sqrt{k} \tag{4.65}$$

and

$$D_6 = \left(\frac{2}{k}\right)^{\frac{1}{p}-\frac{1}{2}} \left\{ \frac{(1-p)\,2^{\frac{1}{p}-\frac{1}{2}}}{(1-p)\left(1+2^{\frac{1}{p}-1}\right) + p\sqrt{t}} + \sqrt{\frac{p\,(1+\delta_{tk})}{1-\delta_{tk}}} \left[\frac{(2-p)\delta_{tk}}{(t-1)\,(1+\delta_{tk})}\right]^{\frac{2-p}{2p}} \right\}. \tag{4.66}$$

(iii) When $\delta_{tk} = 0$, we have $\beta(p,t,\sigma) = \frac{1}{\delta(p,t,\sigma)}$, and the inequality (4.55) becomes

$$\begin{aligned} &\frac{1}{\delta(p,t,\sigma)} \left\|h_{\max(k)} + h_{(1)}\right\|_2^2 \\ &\quad - \left[(1-p)\left(1+2^{\frac{1}{p}-1}\right) + \frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}\right]\sqrt{k}\,(\varepsilon+\epsilon)\left\|h_{\max(k)} + h_{(1)}\right\|_2 \\ &\quad - (1-p)\,k^{1-\frac{1}{p}}2^{\frac{2}{p}-1}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p(\varepsilon+\epsilon) \le 0. \end{aligned} \tag{4.67}$$

By solving the inequality, we obtain

$$\begin{aligned} &\left\|h_{\max(k)} + h_{(1)}\right\|_2 \\ \le& \delta(p,t,\sigma)\left[(1-p)\left(1+2^{\frac{1}{p}-1}\right) + \frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}\right]\sqrt{k}\,(\varepsilon+\epsilon) \\ &+ \frac{(1-p)\,k^{\frac{1}{2}-\frac{1}{p}}2^{\frac{2}{p}-1}}{(1-p)\left(1+2^{\frac{1}{p}-1}\right) + \frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}}\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p. \end{aligned} \tag{4.68}$$

Therefore,

$$\left\|h_{\max(k)} + h_{(1)}\right\|_2 \le D_5(\varepsilon + \epsilon) + D_6 \left\|\mathrm{w} \circ x_{-\max(k)}\right\|_p,$$

where

$$D_5 = \left[(1-p)\left(1+2^{\frac{1}{p}-1}\right) + \frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}\right]\delta(p,t,\sigma)\sqrt{k} \tag{4.69}$$

and

$$D_6 = \frac{(1-p)\,k^{\frac{1}{2}-\frac{1}{p}}2^{\frac{2}{p}-1}}{(1-p)\left(1+2^{\frac{1}{p}-1}\right) + \frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}}. \tag{4.70}$$

Analogous to (4.36) in the proof of Theorem 3.1, it can be readily derived that

$$\|\hat{x}^{DS} - x\|_2 = \sqrt{\left\|h_{\max(k)}\right\|_2^2 + \left\|h_{-\max(k)}\right\|_2^2} \le C_5\,(\varepsilon+\epsilon) + C_6\left\|\mathrm{w}\circ x_{-\max(k)}\right\|_p,$$

where $C_5 = \sqrt{1+2^{\frac{2}{p}-2}}D_5$ and $C_6 = \sqrt{D_6^2 + 2^{\frac{2}{p}-2}\left[D_6 + \left(\frac{2}{k^{1-\frac{p}{2}}}\right)^{\frac{1}{p}}\right]^2}$. □

4.6. Proof of Corollary 3.3

When $\mathrm{w} \circ x_{-\max(k)} = \mathbf{0}$, the inequality (4.53) becomes

$$\begin{aligned}(1-p)\|h\|_1(\varepsilon+\epsilon) &+ \frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\langle A(h_{\max(k)}+h_{(1)}), Ah\rangle \\ &+ \frac{\delta_{tk}\delta^2(p,t,\sigma)+\delta_{tk}-2\delta(p,t,\sigma)}{2\delta^2(p,t,\sigma)}\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \\ &+ \frac{p(1+\delta_{tk})}{2}\left[\frac{(2-p)\delta_{tk}}{(t-1)(1+\delta_{tk})}\right]^{\frac{2-p}{p}}\left(\sigma^{\frac{2-p}{2}}\left\|h_{\max(k)}+h_{(1)}\right\|_2^p\right)^{\frac{2}{p}} \geq 0, \end{aligned} \tag{4.71}$$

and the inequality (4.32) becomes

$$\left\|h_{-\max(k)}\right\|_p^p \leq \left\|h_{\max(k)}\right\|_p^p.$$

Then by virtue of Lemma 2.4, we derive

$$\left\|h_{-\max(k)}\right\|_2^2 \leq k\left[\left(\frac{\left\|h_{\max(k)}\right\|_2^2}{k}\right)^{\frac{p}{2}}\right]^{\frac{2}{p}} = \left\|h_{\max(k)}\right\|_2^2, \tag{4.72}$$

$$\left\|h_{-\max(k)}\right\|_1 \leq k\left[\left(\frac{\left\|h_{\max(k)}\right\|_1}{k}\right)^{p}\right]^{\frac{1}{p}} = \|h_{\max(k)}\|_1. \tag{4.73}$$

Therefore,

$$\|h\|_1 = \left\|h_{\max(k)}\right\|_1 + \left\|h_{-\max(k)}\right\|_1 \leq 2\left\|h_{\max(k)}\right\|_1 \leq 2\sqrt{k}\left\|h_{\max(k)}\right\|_2. \tag{4.74}$$

Recall (4.48), i.e., $\left\|A^TAh\right\|_\infty \leq \varepsilon+\epsilon$. We obtain

$$\|Ah\|_2^2 = \langle h, A^TAh\rangle \leq \|h\|_1\left\|A^TAh\right\|_\infty \leq 2\sqrt{k}\left\|h_{\max(k)}\right\|_2(\varepsilon+\epsilon). \tag{4.75}$$

Recall (4.52) and (3.6), together with (4.71), we obtain

$$\begin{aligned}2(1-p)\sqrt{k}(\varepsilon+\epsilon)\left\|h_{\max(k)}+h_{(1)}\right\|_2 &+ \frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{tk}(\varepsilon+\epsilon)\left\|h_{\max(k)}+h_{(1)}\right\|_2 \qquad (4.76)\\ +\Bigg\{-\frac{2\delta(p,t,\sigma)-\left(1+\delta^2(p,t,\sigma)\right)\delta_{tk}}{2\delta^2(p,t,\sigma)} &+ \frac{p}{2}(1+\delta_{tk})\left[\frac{(2-p)\sigma\delta_{tk}}{(t-1)(1+\delta_{tk})}\right]^{\frac{2-p}{p}}\Bigg\}\left\|h_{\max(k)}+h_{(1)}\right\|_2^2 \geq 0.\end{aligned}$$

It follows from (4.76) that

$$\left\|h_{\max(k)}+h_{(1)}\right\|_2 \leq D_7(\varepsilon+\epsilon), \tag{4.77}$$

where $D_7 = \sqrt{k}\frac{2(1-p)+\frac{1-(1-p)\delta(p,t,\sigma)}{\delta(p,t,\sigma)}\sqrt{t}}{\beta(p,t,\sigma)}$. Hence,

$$\begin{aligned}\|\hat{x}^{DS}-x\|_2 =& \sqrt{\left\|h_{\max(k)}\right\|_2^2+\left\|h_{-\max(k)}\right\|_2^2} \\ \leq& \sqrt{2}\left\|h_{\max(k)}+h_{(1)}\right\|_2 \leq \sqrt{2}D_7(\varepsilon+\epsilon). \end{aligned} \tag{4.78}$$

This completes the proof of Corollary 3.3. □

4.7. Proof of Corollary 3.5

When in particular $\sigma = 1$, by virtue of Corollary 3.3, we directly derive (3.17).

If $\mathrm{w} = \mathbf{1}$, then $w = 1$, and thus $\sigma = 1$. In addition, in view of (4.6), we obtain

$$\nu = \left(\frac{\left\| h_{\max(k)} \right\|_p^p}{k} \right)^{\frac{1}{p}}. \tag{4.79}$$

Recall the definition (4.7) of $h_{(1)} \in \mathbb{R}^n$, i.e.,

$$h_{(1)}(i) = \begin{cases} h_{-\max(k)}(i), & \left| h_{-\max(k)}(i) \right| > (t-1)^{-\frac{1}{p}} \nu, \\ 0, & \text{otherwise.} \end{cases}$$

Since $w = 1$, we have

$$(t-1)^{-\frac{1}{p}} \nu \geq \nu = \left(\frac{\left\| h_{\max(k)} \right\|_p^p}{k} \right)^{\frac{1}{p}} \geq \| h_{-\max(k)} \|_\infty \geq \left| h_{-\max(k)}(i) \right| \tag{4.80}$$

for $t \in [\frac{4}{2+p}, 2]$, and thus $h_{(1)} = \mathbf{0}$. Then the inequality (4.52) becomes

$$\begin{aligned} \left\langle A(h_{\max(k)} + h_{(1)}), Ah \right\rangle =& \langle h_{\max(k)}, A^T Ah \rangle \\ \leq& \| h_{\max(k)} \|_1 \left\| A^T Ah \right\|_\infty \leq \sqrt{k} \| h_{\max(k)} \|_2 (\varepsilon + \epsilon), \end{aligned} \tag{4.81}$$

and therefore, the inequality (4.71) turns to be

$$\begin{aligned} &(1-p) \|h\|_1 (\varepsilon + \epsilon) + \frac{1 - (1-p)\delta(p,t,1)}{\delta(p,t,1)} \sqrt{k} (\varepsilon + \epsilon) \left\| h_{\max(k)} \right\|_2 \\ &+ \frac{\delta_{tk} \delta^2(p,t,1) + \delta_{tk} - 2\delta(p,t,1)}{2\delta^2(p,t,1)} \left\| h_{\max(k)} \right\|_2^2 \\ &+ \frac{p}{2} (1 + \delta_{tk}) \left[\frac{(2-p)\delta_{tk}}{(t-1)(1+\delta_{tk})} \right]^{\frac{2-p}{p}} \left\| h_{\max(k)} \right\|_2^2 \geq 0. \end{aligned} \tag{4.82}$$

Recall the notation (3.6) of $\beta(p,t,\sigma)$ and (4.74), i.e., $\|h\|_1 \leq 2\sqrt{k} \left\| h_{\max(k)} \right\|_2$. Therefore, the inequality (4.82) yields

$$\begin{aligned} &2(1-p)\sqrt{k} \left\| h_{\max(k)} \right\|_2 (\varepsilon + \epsilon) + \frac{1-(1-p)\delta(p,t,1)}{\delta(p,t,1)} \sqrt{k} (\varepsilon + \epsilon) \left\| h_{\max(k)} \right\|_2 \\ &-\beta(p,t,1) \left\| h_{\max(k)} \right\|_2^2 \geq 0. \end{aligned} \tag{4.83}$$

We obtain

$$\left\| h_{\max(k)} \right\|_2 \leq \sqrt{k} \frac{1 - p + \frac{1}{\delta(p,t,1)}}{\beta(p,t,1)} (\varepsilon + \epsilon), \tag{4.84}$$

and thus

$$\left\| \hat{x}^{DS} - x \right\|_2 \leq \sqrt{2} \left\| h_{\max(k)} \right\|_2 \leq \sqrt{2k} \frac{1 - p + \frac{1}{\delta(p,t,1)}}{\beta(p,t,1)} (\varepsilon + \epsilon). \tag{4.85}$$

This completes the proof. □

5. Numerical Experiments

In this section, we present a series of numerical experiments to illustrate the performance of recovery of approximate k-sparse signals by the weighted ℓ_p $(0 < p \leq 1)$ minimization (1.13). We adopt the iteratively reweighted least squares (IRLS) algorithm proposed in [7] to solve the nonconvex optimization problem.

First, $x^{(0)} = \arg\min_{x \in \mathbb{R}^n} \|y - Ax\|_2^2$ and $x^{(t+1)}$ is the solution of

$$\min_{x \in \mathbb{R}^n} \frac{1}{2\lambda}\|y - Ax\|_2^2 + \frac{1}{2}\left\|W^{(t)}x\right\|_2^2, \tag{5.1}$$

where $\lambda > 0$ is a regularization parameter, and the weight matrix $W_i^{(t)}$ is defined as

$$W_i^{(t)} = \operatorname{diag}\left(\sqrt{pw_i^p}\left(\tau_t^2 + \left(x_i^{(t)}\right)^2\right)^{p/4-1/2}\right)$$

for $i = 1, \ldots, n$. Then by (5.1), we obtain

$$x^{(t+1)} = \left(W^{(t)}\right)^{-1}\left(\Phi\left(W^{(t)}\right)^{-1}\right)^T\left(\Phi\left(W^{(t)}\right)^{-1} + \lambda I_n\right)^{-1}\left(\Phi\left(W^{(t)}\right)^{-1}\right)^T y.$$

$\tau_0 = 1$, $\tau_{t+1} = \min\left\{\tau_t, \gamma r(x^{(t+1)})_{\hat{k}+1}\right\}$, where $\gamma \in (0,1)$ is a constant. We set $\gamma = 0.9$ and $\lambda = 10^{(-6)}$. $r(x)$ is the rearrangement of the absolute values of x in decreasing order. $\hat{k}$ is the number of the support estimate, we set $\hat{k} = k$. When $t > 1000$ or $r(x^{(t+1)})_{\hat{k}+1} = 0$, we stop the iteration and output $x^{(t+1)}$.

We consider signals $x \in \mathbb{R}^{500}$ such that $x_j = j^{-d}$ for some $d > 1$, $i = 1, \ldots, n$. In our experiments, the measurement matrix A is generated as an $m \times 500$ matrix with entries drawing from i.i.d standard normal distribution. For a generated approximately sparse signal x, the measurements $y = Ax + z$, where A is an $m \times 500$ Gaussian matrix with n varying between 80 and 220. In the case of noisy measurements, z is standard Gaussian white noise and $\frac{\|z\|_2}{\|x\|_2} = 0.02$. In the case of noiseless measurements, $z = 0$. The recovery performance is assessed by the signal to noise ratio (SNR) denoted by

$$\mathrm{SNR} = 20\log_{10}\frac{\|x\|_2}{\|x - x^{(t+1)}\|_2}, \tag{5.2}$$

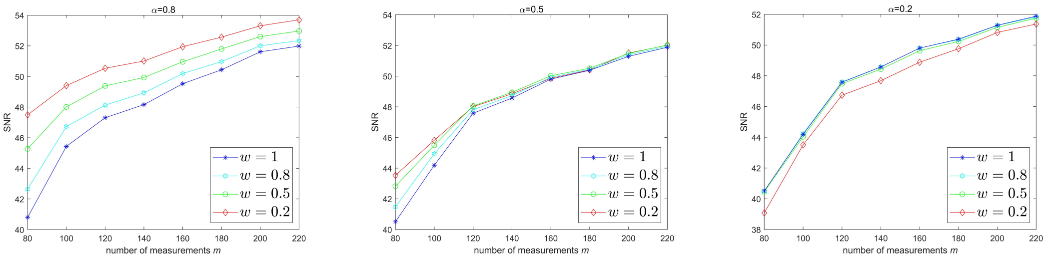
and the measure of the SNR is dB.

In each experiment, we report the average results over 30 replications and set $d = 2$.

Fig. 5.1 illustrates the recovery performance of approximate k-sparse signal under different measurements in the noiseless case. We set $p = 0.5$ and $k = 40$, i.e., the best 40-term sparse approximation. It shows that for $\alpha \geq 0.5$, smaller weight w yields better recovery performance.

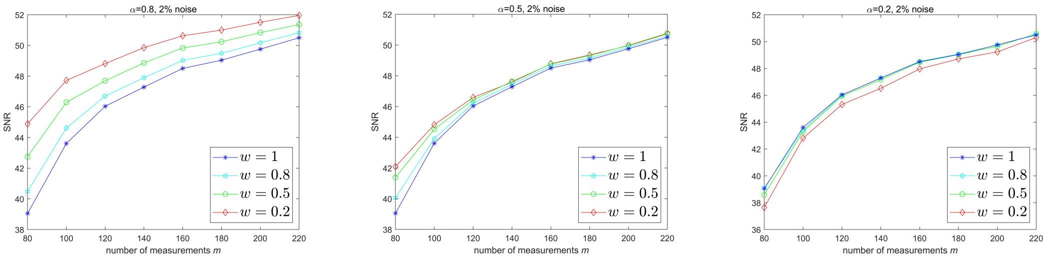
Fig. 5.2 shows the recovery performance of approximate sparse signal under different measurements in the noisy case. The results are consistent with the no-noise case.

In Fig. 5.3 and Fig. 5.4, we illustrate the impacts of $p \in (0,1]$ for both the noiseless and noisy measurements cases. We set $\alpha = 0.8$ and $\rho = 1$. The results show that smaller p always leads to better recovery performance. In some measurements, the recovery performance of $p = 0.2$ is worse than $p = 0.5$ and $p = 0.7$. This is because a smaller p makes the minimizing functional more nonconvex and thus more difficult to solve.



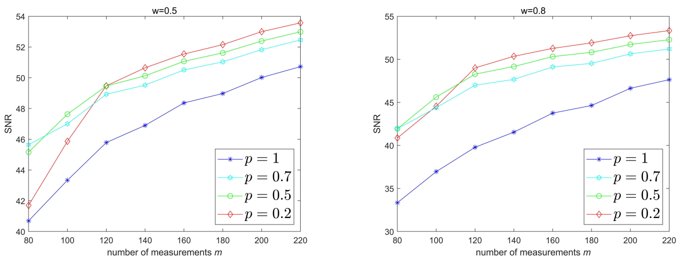
(a) (b) (c)

Fig. 5.1. In the noiseless case, for $p = 0.5$ and $\rho = 1$, the recovery performance of the weighted ℓ_p minimization when (a) $\alpha = 0.8$, (b) $\alpha = 0.5$, (c) $\alpha = 0.2$.



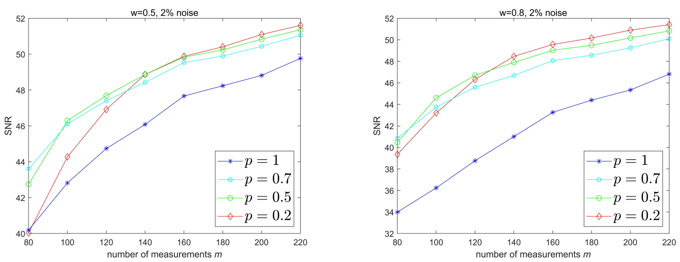
(a) (b) (c)

Fig. 5.2. In the noise case, for $p = 0.5$ and $\rho = 1$, the recovery performance of the weighted ℓ_p minimization when (a) $\alpha = 0.8$, (b) $\alpha = 0.5$, (c) $\alpha = 0.2$.



(a) (b)

Fig. 5.3. In the noiseless case, for $\alpha = 0.8$ and $\rho = 1$, the recovery performance of the weighted ℓ_p minimization when (a) $w = 0.5$, (b) $w = 0.8$.



(a) (b)

Fig. 5.4. In the noisy case, for $\alpha = 0.8$ and $\rho = 1$, the recovery performance of the weighted ℓ_p minimization when (a) $w = 0.5$, (b) $w = 0.8$.

6. Conclusions

The recovery of approximately k-sparse signal with partial support information in two different noise settings is investigated by the weighted ℓ_p $(0 < p \leq 1)$ minimization method (1.13). The newly derived theorems and corollaries indicate that approximately k-sparse signal $x \in \mathbb{R}^n$ can be stably and robustly recovered by the minimizer $\hat{x}$ of (1.13) when there is partial and possibly partly inaccurate prior support information. The obtained results not only improve the work in [17] which addressed the ℓ_2-bounded noise setting concerning the recovery of strictly k-sparse original signal, but also include the optimal results by weighted ℓ_1 minimization or by standard ℓ_p minimization as special cases.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under grant numbers 12171496 and U1811461, in part by Guangdong Basic and Applied Basic Research Foundation under grant number 2020A1515010454, and in part by the Science and Technology Program of Guangzhou under grant number 201904010374.

References

[1] T. Cai, A. Zhang, Sharp RIC bound for sparse signal and low-rank matrix recovery, *Appl. Comput. Harmon. Anal.*, **35**:1 (2013), 74–93.

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

[3] E. Candès, J. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, *Commun. Pure Appl. Math.*, **59**:8 (2006), 1207–1223.

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

[5] E. Candès, T. Tao, The Dantzig selector: Statistical estimation when p is much larger than n, *Ann. Stat.*, **35**:6 (2007), 2313–2351.

[6] R. Chartrand, V. Staneva, Restricted isometry properties and nonconvex compressive sensing, *Inverse Probl.*, **24**:3 (2008), 1–14.

[7] R. Chartrand, W. Yin, Iteratively reweighted algorithms for compressive sensing, in: Proc. International Conference on Acoustics, Speech, and Signal Processing, IEEE, Las Vegas, NV, USA, (2008), 3869–3872.

[8] B. Chen, A. Wan, General RIP bounds of δ_{tk} for sparse signal recovery by ℓ_p minimization, *Neurocomputing*, **363** (2019), 306–312.

[9] W. Chen, Y. Li, G. Wu, Recovery of signals under the high order RIP condition via prior support information, *Signal Process.*, **153** (2018), 83–94.

[10] I. Daubechies, M. Defrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Commun. Pure Appl. Math.*, **57**:11 (2004), 1413–1457.

[11] I. Daubechies, R. Devore, M. Fornasier, C. S. Güntürk, Iteratively re-weighted least squares minimization: Proof of faster than linear rate for sparse recovery, in: Conference on Information Sciences and Systems, IEEE, Princeton, NJ, USA, (2008), 26–29.

[12] I. Daubechies, R. Devore, M. Fornasier, C.S. Güntürk, Iteratively reweighted least squares minimization for sparse recovery, *Commun. Pure Appl. Math.*, **63**:1 (2010), 1–38.

[13] D.L. Donoho, Compressed sensing, *IEEE Trans. Inform. Theory*, **52**:4 (2006), 1289–1306.

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

[15] S. Foucart, H. Rauhut, A Mathematical Introduction to Compressive Sensing, Birkhäuser, Boston, 2013.

[16] M. Friedlander, H. Mansour, R. Saab, Ö Yilmaz, Recovering compressively sampled signals using partial support information, *IEEE Trans. Inform. Theory*, **58**:2 (2012) 1122–1134.

[17] H. Ge, W. Chen, M. Ng, New RIP bounds for recovery of sparse signals with partial support information via weighted ℓ_p minimization, *IEEE Trans. Inform. Theory*, **66**:6 (2020) 3914–3928.

[18] N. Ghadermarzy, H. Mansour, Ö. Yilmaz, Non-convex compressed sensing using partial support information, *Sampling Theory in Signal Image Processing*, **13**:3 (2014), 251–272.

[19] R. Gribonval, M. Nielsen, Sparse representations in unions of bases, *IEEE Trans. Inform. Theory*, **49**:12 (2003), 3320–3325.

[20] T. Ince, A. Nacaroglu, N. Watsuji, Nonconvex compressed sensing with partially known signal support, *Signal Process.*, **93**:1 (2013), 338–344.

[21] M. Lai, J. Wang, An unconstrained ℓ_q minimization with $0 < q \leq 1$ for sparse solution of underdetermined linear systems, *SIAM J. Optim.*, **21**:1 (2011), 82–101.

[22] X. Lan, A. Ma, P. Yuen, R. Chellappa, Joint sparse representation and robust feature-level fusion for multi-cue visual tracking, *IEEE Trans. Image Process.*, **24**:12 (2015), 5826–5841.

[23] Q. Sun, Recovery of sparsest signals via ℓ_q-minimization, *Appl. Comput. Harmon. Anal.*, **32**:3 (2012), 329–341.

[24] A. Wan, Stable recovery of approximately k-sparse signals in noisy cases via ℓ_p minimization, *Neurocomputing*, **406** (2020), 159–168.

[25] A. Wan, Uniform RIP conditions for recovery of sparse signals by $\ell_p\,(0 < p \leq 1)$ minimization, *IEEE Trans. Signal Process.*, **68** (2020), 5379–5394.

[26] J. Wen, D. Li, F. Zhu, Stable recovery of sparse signals via ℓ_q-minimization, *Appl. Comput. Harmon. Anal.*, **38**:1 (2015), 161–176.

[27] C. Yang, X. Shen, H. Ma, Y. Gu, H. So, Sparse recovery conditions and performance bounds for ℓ_p-minimization, *IEEE Trans. Signal Process.*, **66**:19 (2018), 5014–5028.

[28] R. Zhang, S. Li, Optimal RIP bounds for sparse signals recovery via ℓ_p minimization, *Appl. Comput. Harmon. Anal.*, **47**:3 (2019), 566–584.