Analysis in Theory and Applications
Anal. Theory Appl., Vol. **37**, No. 3 (2021), pp. 289-310

DOI: 10.4208/ata.2021.lu80.02
September 2021

Weighted ℓ_p-Minimization for Sparse Signal Recovery under Arbitrary Support Prior

Yueqi Ge[1], Wengu Chen[2,*], Huanmin Ge[3] and Yaling Li[4]

[1] *Liwan Primary School, Shenzhen, Guangdong 518000, China*
[2] *Institute of Applied Physics and Computational Mathematics, Beijing 100088, China*
[3] *The Sports Engineering College, Beijing Sport University, Beijing 100088, China*
[4] *Department of Mathematics, Zhejiang University of Science and Technology, Hangzhou, Zhejiang 310023, China*

Received 28 July 2020; Accepted (in revised version) 21 July 2021

Dedicated to Prof. Shanzhen Lu with admiration on the occasion of his 80th birthday

Abstract. Weighted ℓ_p $(0 < p \leq 1)$ minimization has been extensively studied as an effective way to reconstruct a sparse signal from compressively sampled measurements when some prior support information of the signal is available. In this paper, we consider the recovery guarantees of k-sparse signals via the weighted ℓ_p $(0 < p \leq 1)$ minimization when arbitrarily many support priors are given. Our analysis enables an extension to existing works that assume only a single support prior is used.

Key Words: Adaptive recovery, compressed sensing, weighted ℓ_p minimization, sparse representation, restricted isometry property.

AMS Subject Classifications: 90C26, 90C30, 94A20

1 Introduction

Compressed sensing [2, 5] is a new data acquisition paradigm, which reliably recovers a high dimensional sparse signal $\boldsymbol{x} \in \mathbb{R}^n$ (a signal is called k-sparse if the number of its nonzero entries has at most $k \ll n$) from significantly fewer linear observations

$$\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{e}, \tag{1.1}$$

where $\boldsymbol{\Phi} \in \mathbb{R}^{m\times n}$ is a measurement matrix and $\boldsymbol{e} \in \mathbb{R}^m$ denotes additive noise that satisfies $\|\boldsymbol{e}\|_2 \leq \epsilon$ for some known $\epsilon \geq 0$. Compressed sensing is nonadaptive because the measurement matrix $\boldsymbol{\Phi}$ does not depend on the signal being measured. But, some

*Corresponding author. *Email addresses:* `geyueqi123@163.com` (Y. Ge), `chenwg@iapcm.ac.cn` (W. Chen), `gehuanmin@163.com` (H. Ge), `leeyaling@126.com` (Y. Li)

prior information of the signal $\boldsymbol{x}$ may be included in the estimates of the support of $\boldsymbol{x}$ or some estimates of largest coefficients of $\boldsymbol{x}$ in some settings. For example, video and audio signals exhibit strong correlation over temporal frames, which can be used to estimate a portion of the support based on previously decoded frames (see [6]). Therefore, the recovery of the signal $\boldsymbol{x}$ incorporating prior support information has received much attention including the weighted ℓ_1-minimization [3, 4, 6, 14, 16, 17, 19], the weighted ℓ_p $(0 < p < 1)$-minimization [10, 11, 13, 18] and the greedy algorithm with partial support information [7, 12, 15].

This paper considers the recovery of the signal $\boldsymbol{x}$ from (1.1) and is devoted to new RIP bounds for the exact and stable recovery of sparse signals with arbitrary many support priors via the weighted ℓ_p-minimization:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{x}\|_{p,\mathbf{w}}^p \quad \text{subject to} \ \ \|\boldsymbol{\Phi}\boldsymbol{x}-\boldsymbol{y}\|_2 \le \varepsilon, \tag{1.2}$$

where $\mathbf{w} \in [0,1]^n$ is a weight vector and

$$\|\boldsymbol{x}\|_{p,\mathbf{w}} = \Big(\sum_{i=1}^{n} \mathrm{w}_i |x_i|^p\Big)^{\frac{1}{p}}.$$

The main idea inherited in the weighted ℓ_p $(0 < p \le 1)$-minimization is to make the entries of $\boldsymbol{x}$, which are "expected" to be large, be penalized less in the weighted objective function in (1.2) by the effect of the weight $\mathbf{w}$.

As $p = 1$, the method (1.2) reduces to the weighted ℓ_1-minimization:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{x}\|_{1,\mathbf{w}} \quad \text{subject to} \ \ \|\boldsymbol{\Phi}\boldsymbol{x}-\boldsymbol{y}\|_2 \le \varepsilon. \tag{1.3}$$

The rest of the paper is organized as follows. In Section 2, we recall a recently established RIP bound for signal recovery by virtue of the weighted ℓ_p-minimization with a single weight. In Section 3, we respectively present sufficient conditions for the recovery of sparse signals by weighted ℓ_p-minimization with non-uniform weights in both the noiseless and ℓ_2 bounded noise. Section 4 is devoted to the proofs of the main results.

2 Weighted ℓ_p-minimization with a single weight

Let $\widetilde{T} \subseteq [n] = \{1, 2, \cdots, n\}$ be a known single support estimate of $\boldsymbol{x}$. The weight vector $\mathbf{w}$ in this case is taken by

$$\mathrm{w}_i = \begin{cases} \omega, & i \in \widetilde{T}, \\ 1, & i \in \widetilde{T}^c, \end{cases} \tag{2.1}$$

for some fixed $\omega \in [0,1]$ and $i \in [n]$.

The restricted isometry property (RIP) is one of the main tools used to evaluate the recovery performance via a variety of efficient algorithms. The RIP notion introduced by Candès et al. in [2], is the most widely used framework in compressed sensing.

Definition 2.1. *For a matrix $\mathbf{\Phi} \in \mathbb{R}^{m\times n}$ and an integer $1 \le k \le n$, $\mathbf{\Phi}$ is said to satisfy the RIP of order k if there exists a constant $\delta_k \in [0,1)$ such that*

$$(1-\delta_k)\|\boldsymbol{x}\|_2^2 \le \|\mathbf{\Phi}\boldsymbol{x}\|_2^2 \le (1+\delta_k)\|\boldsymbol{x}\|_2^2 \tag{2.2}$$

holds for all k-sparse signals $\boldsymbol{x} \in \mathbb{R}^n$. The smallest constant δ_k is called the restricted isometry constant (RIC) of order k for $\mathbf{\Phi}$.

When k is not an integer, δ_k is defined as $\delta_{\lceil k\rceil}$ in [1], where $\lceil k\rceil$ denotes an integer satisfying $k \le \lceil k\rceil < k+1$.

The main result of [9] generalizes the recovery condition from [21] to the weighted ℓ_p-minimization (1.2) where the weight vector $\mathbf{w}$ is specified in (2.1).

Theorem 2.1 below states the main result of [9] which presents a sufficient condition for the exact recovery of sparse signal $\boldsymbol{x}$ from $\boldsymbol{y} = \mathbf{\Phi}\boldsymbol{x}$.

Theorem 2.1. *Let $\boldsymbol{x}$ be an arbitrary k-sparse vector in $\mathbb{R}^n$ with $T = supp(\boldsymbol{x})$ and $\boldsymbol{y} = \mathbf{\Phi}\boldsymbol{x}$. Let $\widetilde{T} \subseteq [n]$ be an arbitrary set and $\rho \ge 0$ and $0 \le \alpha \le 1$ with $\alpha\rho \le 1$ such that $|\widetilde{T}| = \rho k$ and $|\widetilde{T}\cap T| = \alpha\rho k$. Given the weight $\omega \in [0,\,1]$ and $0 < p \le 1$, define some important parameters somehow depending on the weight ω, and the size and the overlap of the true signal support T and the prior support estimate $\widetilde{T}$, and p as follows*

- *The constant ζ:*

$$\zeta = \Big(\omega + (1-\omega)(1+\rho-2\alpha\rho)^{\frac{2-p}{2}}\Big)^{\frac{2}{2-p}}, \tag{2.3}$$

- *the constant d:*

$$d = \begin{cases} 1, & \omega = 1, \\ 1+(\max\{0,1-2\alpha\})\rho, & 0 \le \omega < 1, \end{cases} \tag{2.4}$$

- *the parameter Θ is defined by*

$$\Theta = \frac{\zeta}{t-d}, \tag{2.5}$$

- *for $\Theta > 0$, the quantity $\delta(p,\Theta)$ is defined by*

$$\delta(p,\Theta) = \begin{cases} \dfrac{1}{\sqrt{p^2+(2-p)^2\Theta}-(1-p)}, & \Theta \ge \Theta_0 = \dfrac{2+p}{2-p}, \\ \dfrac{z_0}{(2-p)\Theta - z_0}, & \Theta < \Theta_0, \end{cases} \tag{2.6}$$

where $z_0 \in ((1-p)\Theta, \min(1, \frac{2-p}{2}\Theta)$ is the only positive solution of the equation

$$\frac{p}{2}z^{\frac{2}{p}} + z - \frac{(2-p)\Theta}{2} = 0. \tag{2.7}$$

Moreover, for $\Theta = \frac{\zeta}{t-d} = 0$, we define $\delta(p,\Theta) = 1$.

If the measurement matrix $\mathbf{\Phi}$ *satisfies RIP with*

$$\delta_{tk} < \delta(p,\Theta) \tag{2.8}$$

for $d < t \leq 2d$*, then the weighted* ℓ_p*-minimization* (1.2) *with the weight vector* $\mathbf{w}$ *defined in* (2.1) *and* $0 < p \leq 1$ *recovers* $\boldsymbol{x}$ *exactly.*

3 Weighted ℓ_p-minimization with non-uniform weights

In this section, we present our main results for generalizing the weighted ℓ_p-minimization theory of [9], to allow for arbitrary weight assignments.

We consider the weighted ℓ_p-minimization with L distinct weights, where $1 \leq L \leq n$. Let $\widetilde{T}_j \subseteq [n]$ be arbitrary L disjoint sets and denote $\rho_j \geq 0$ and $0 \leq \alpha_j \leq 1$ such that $|\widetilde{T}_j| = \rho_j k$ and $|\widetilde{T}_j \cap T| = \alpha_j \rho_j k$, $j = 1, \cdots, L$, where $\rho_j \geq 0$ and $0 \leq \alpha_j \leq 1$ are called the relative size and accurary for each $j = 1, \cdots, L$. Define $\widetilde{T} = \cup_{j=1}^{L} \widetilde{T}_j$. The weight vector $\mathbf{w}$ in this general case is chosen in the following way

$$\mathrm{w}_i = \begin{cases} \omega_j, & i \in \widetilde{T}_j, \\ 1, & i \in \widetilde{T}^c, \end{cases} \tag{3.1}$$

for $i \in [n]$ and $\omega_j \in [0,1]$, $j = 1, \cdots, L$ are given weights.

We first provide a recovery guarantee for the weighted ℓ_p-minimization with L distinct weights in noiseless case.

Theorem 3.1. *For* $0 < p \leq 1$ *and* $\boldsymbol{y} = \mathbf{\Phi}\boldsymbol{x}$*, suppose that* $\boldsymbol{x}$ *be k-sparse with* $T = supp(\boldsymbol{x})$*. Let* $\widetilde{T}_i \subseteq [n]$ *be arbitrary L disjoint sets and* $\rho_i \geq 0$ *and* $0 \leq \alpha_i \leq 1$ *such that* $|\widetilde{T}_i| = \rho_i k$ *and* $|\widetilde{T}_i \cap T| = \alpha_i \rho_i k$, $i = 1, \cdots, L$*. Without loss of generality, assume that the weights in* (3.1) *are ordered so that* $0 \leq \omega_L \leq \cdots \leq \omega_1 \leq 1$*. Let*

$$\beta_i = \max\left\{ \sum_{j=i}^{L} \alpha_j \rho_j, \sum_{j=i}^{L} (1-\alpha_j)\rho_j \right\},$$

$$b_i = \begin{cases} 1, & i = 1, \\ sgn(\omega_{i-1} - \omega_i), & i = 2, \cdots, L, \end{cases}$$

and

$$d=\begin{cases}1, & \omega_1=\omega_2=\cdots=\omega_L=1,\\ \max\limits_{i\in\{1,2\cdots,L\}}\Big\{b_i\Big(1-\sum\limits_{j=i}^{L}\alpha_j\rho_j+\beta_i\Big)\Big\}, & 0\le\prod\limits_{i=1}^{L}\omega_i<1,\end{cases}\tag{3.2a}$$

$$\begin{aligned}\gamma_L=&\,\omega_L+(1-\omega_1)\Big(1+\sum_{i=1}^{L}\rho_i-2\sum_{i=1}^{L}\alpha_i\rho_i\Big)^{\frac{2-p}{2}}\\&+\sum_{i=2}^{L}(\omega_{i-1}-\omega_i)\Big(1+\sum_{j=i}^{L}\rho_j-2\sum_{j=i}^{L}\alpha_j\rho_j\Big)^{\frac{2-p}{2}}.\end{aligned}\tag{3.2b}$$

If the measurement matrix $\mathbf{\Phi}$ *satisfies RIP and*

$$\delta_{tk}<\delta(t,\,p,\Theta),\tag{3.3}$$

where $d<t\le 2d$*, and for*

$$\Theta=\frac{\gamma_L^{2/(2-p)}}{t-d}>0,\tag{3.4}$$

$\delta(t,p,\Theta)$ *is defined by*

$$\delta(t,p,\Theta)=\begin{cases}\dfrac{1}{\sqrt{p^2+(2-p)^2\Theta}-(1-p)}, & \Theta\ge\Theta_0=\dfrac{2+p}{2-p},\\ \dfrac{z_0}{(2-p)\Theta-z_0}, & \Theta<\Theta_0,\end{cases}\tag{3.5}$$

where $z_0\in\big((1-p)\Theta,\min(1,\frac{2-p}{2}\Theta)\big)$ *is the only positive solution of the equation*

$$\frac{p}{2}z^{\frac{2}{p}}+z-\frac{2-p}{2}\Theta=0,\tag{3.6}$$

and

$$\delta(t,p,\Theta)=1\quad \textit{if}\ \Theta=\frac{\gamma_L^{2/(2-p)}}{t-d}=0,$$

then the weighted ℓ_p*-minimization* (1.2) *recovers* $\boldsymbol{x}$ *exactly.*

As $p=1$, Theorem 3.1 presents a sufficient condition of the weighted ℓ_1-minimization (1.3) for the exact recovery of $\boldsymbol{x}$, which improves the theory of [17]. See the following Corollary 3.1.

Corollary 3.1. *If* $p=1$ *and* $\mathbf{\Phi}$ *satisfies RIP with*

$$\delta_{tk}<\frac{1}{\sqrt{1+\Theta}},\tag{3.7}$$

where $d < t \leq 2d$ and

$$\begin{aligned}\Theta =&(t-d)^{-1}\Bigg(\omega_L + (1-\omega_1)\sqrt{1+\sum_{i=1}^{L}\rho_i - 2\sum_{i=1}^{L}\alpha_i\rho_i}\\&+\sum_{i=2}^{L}(\omega_{i-1}-\omega_i)\sqrt{1+\sum_{j=i}^{L}\rho_j - 2\sum_{j=i}^{L}\alpha_j\rho_j}\Bigg)^2,\end{aligned}$$

then the weighted ℓ_1-minimization (1.3) *exactly recover $\boldsymbol{x}$.*

Remark 3.1. Note that the sufficient condition (3.7) is identical to the condition (3.1) in [8], since

$$\delta_{tk} < \frac{1}{\sqrt{1+\Theta}} = \sqrt{\frac{t-d}{t-d+\gamma_L^2}},$$

where the equality is from $\Theta = \frac{\gamma_L^2}{t-d}$ and

$$\begin{aligned}\gamma_L =&\omega_L + (1-\omega_1)\sqrt{1+\sum_{i=1}^{L}\rho_i - 2\sum_{i=1}^{L}\alpha_i\rho_i}\\&+\sum_{i=2}^{L}(\omega_{i-1}-\omega_i)\sqrt{1+\sum_{j=i}^{L}\rho_j - 2\sum_{j=i}^{L}\alpha_j\rho_j}.\end{aligned}\tag{3.8}$$

In noisy case, we have the following theorem.

Theorem 3.2. *For $0 < p \leq 1$ and $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{e}$, suppose that $\hat{\boldsymbol{x}}$ is a minimizer of the weighted ℓ_p-minimization* (1.2) *with $\|\boldsymbol{e}\|_2 \leq \varepsilon$. If $\boldsymbol{\Phi}$ satisfies RIP with*

$$\delta_{tk} < \delta(t, p, \Theta)\tag{3.9}$$

for some $d < t \leq 2d$, where $\delta(t, p, \Theta)$ is defined in (3.5) *for $\Theta > 0$. Then*

$$\begin{aligned}&\|\boldsymbol{x}-\hat{\boldsymbol{x}}\|_2\\ \leq&\sqrt{2}\cdot\left[\frac{4(2-p)\eta(1-\eta)\sqrt{1+\delta_{tk}}+2\eta\sqrt{2(2-p)(1-p)(2-p-\eta)(\delta(t,p,\Theta)-\delta_{tk})}}{(2-p)(2-p-\eta)(\delta(t,p,\Theta)-\delta_{tk})}\right]\varepsilon,\end{aligned}$$

where

$$\eta = \begin{cases}\dfrac{2-p}{\sqrt{p^2+(2-p)^2\Theta}+p}, & \Theta \geq \Theta_0 = \dfrac{2+p}{2-p},\\ \dfrac{z_0}{\Theta}, & \Theta < \Theta_0,\end{cases}\tag{3.10}$$

and γ_L, z_0 are defined as in Theorem 3.1.

4 Proofs of the main results

4.1 Sparse representation and technical lemmas

The original work in [2] triggers an RIP analysis for signal recovery via l_1 minimization. The RIP analysis in [1] and [22] attains the summit for sparse signal recovery via l_1 minimization. The results in [1] and [22] depend on a key tool established in [20] and [1] independently, which represents points in a polytope

$$V = \{v \in \mathbb{R}^n,\ \|v\|_1 \le k\alpha,\ \|v\|_\infty \le \alpha \ \text{ for some } \alpha > 0\}$$

by convex combinations of $k-$sparse vectors. Zhang and Li [21] developed the tool, which extends the sparse representation of a polytope in [1] and [20] adapted to l_p, $(0 < p \le 1)$ case.

Lemma 4.1 ([21, Lemma 2.2]). *For $\boldsymbol{x} \in \mathbb{R}^n$ which satisfies $|\mathrm{supp}(\boldsymbol{x})| = K$, $\|\boldsymbol{x}\|_p^p \le L\rho^p$ and $\|\boldsymbol{x}\|_\infty \le \rho$ with $L \le K$ being a positive integer, ρ being a positive constant and $0 < p \le 1$, then $\boldsymbol{x}$ can be represented as the convex combination of L-sparse vectors, i.e.,*

$$\boldsymbol{x} = \sum_i \lambda_i \boldsymbol{u}_i,$$

where $\lambda_i > 0$, $\sum_i \lambda_i = 1$ and $\|\boldsymbol{u}_i\|_0 \le L$. Furthermore,

$$\sum_i \lambda_i \|\boldsymbol{u}_i\|_2^2 \le \min\Big\{\frac{n}{L}\|\boldsymbol{x}\|_2^2, \rho^p \|\boldsymbol{x}\|_{2-p}^{2-p}\Big\}. \tag{4.1}$$

For the weighted ℓ_p-minimization (1.2) with L distinct weights, the cone constraint inequality can be stated as follows.

Lemma 4.2. *If $\|\hat{\boldsymbol{x}}\|_{p,\mathbf{w}}^p \le \|\boldsymbol{x}\|_{p,\mathbf{w}}^p$ and $\boldsymbol{h} = \hat{\boldsymbol{x}} - \boldsymbol{x}$, then for any index set $\Gamma \subseteq [n]$,*

$$\begin{aligned}\|\boldsymbol{h}_{\Gamma^c}\|_p^p \le & \omega_L \|\boldsymbol{h}_\Gamma\|_p^p + (1-\omega_1)\|\boldsymbol{h}_{(\Gamma\cup\bigcup_{i=1}^L \widetilde{T}_i)\setminus(\bigcup_{i=1}^L \widetilde{T}_i\cap\Gamma)}\|_p^p \\ & + \sum_{j=2}^L (\omega_{j-1}-\omega_j)\|\boldsymbol{h}_{(\Gamma\cup\bigcup_{i=j}^L \widetilde{T}_i)\setminus(\bigcup_{i=j}^L \widetilde{T}_i\cap\Gamma)}\|_p^p \\ & + 2\Big(\omega\|\boldsymbol{x}_{\Gamma^c}\|_p^p + (1-\omega)\|\boldsymbol{x}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p - \sum_{i=1}^L(\omega-\omega_i)\|\boldsymbol{x}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p\Big),\end{aligned} \tag{4.2}$$

where

$$\widetilde{T} = \cup_{i=1}^L \widetilde{T}_i \quad \textit{and} \quad \omega = \sum_{i=1}^L \omega_i.$$

Proof. By $\hat{\boldsymbol{x}} = \boldsymbol{x} + \boldsymbol{h}$ and the choice of the weights in (3.1),

$$\|\hat{\boldsymbol{x}}\|_{p,\mathbf{w}}^p = \|\boldsymbol{x}+\boldsymbol{h}\|_{p,\mathbf{w}}^p \le \|\boldsymbol{x}\|_{p,\mathbf{w}}^p$$

implies

$$\sum_{i=1}^{L}\omega_i\|\boldsymbol{x}_{\widetilde{T}_i}+\boldsymbol{h}_{\widetilde{T}_i}\|_p^p+\|\boldsymbol{x}_{\widetilde{T}^c}+\boldsymbol{h}_{\widetilde{T}^c}\|_p^p\le\sum_{i=1}^{L}\omega_i\|\boldsymbol{x}_{\widetilde{T}_i}\|_p^p+\|\boldsymbol{x}_{\widetilde{T}^c}\|_p^p.$$

Furthermore, we have

$$\begin{aligned}&\sum_{i=1}^{L}(\omega_i\|\boldsymbol{x}_{\widetilde{T}_i\cap\Gamma}+\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma}\|_p^p+\omega_i\|\boldsymbol{x}_{\widetilde{T}_i\cap\Gamma^c}+\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p)\\&\quad+\|\boldsymbol{x}_{\widetilde{T}^c\cap\Gamma}+\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p+\|\boldsymbol{x}_{\widetilde{T}^c\cap\Gamma^c}+\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p\\&\le\sum_{i=1}^{L}(\omega_i\|\boldsymbol{x}_{\widetilde{T}_i\cap\Gamma}\|_p^p+\omega_i\|\boldsymbol{x}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p)+\|\boldsymbol{x}_{\widetilde{T}^c\cap\Gamma}\|_p^p+\|\boldsymbol{x}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p.\end{aligned}$$

Next, we use the reverse triangle inequality to get

$$\begin{aligned}&\sum_{i=1}^{L}\omega_i\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p+\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p\\&\le\sum_{i=1}^{L}\omega_i\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma}\|_p^p+\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p+2\Big(\sum_{i=1}^{L}\omega_i\|\boldsymbol{x}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p+\|\boldsymbol{x}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p\Big).\end{aligned}\tag{4.3}$$

Now, we can write

$$\|\boldsymbol{h}_{\Gamma^c}\|_p^p=\sum_{i=1}^{L}\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p+\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p.$$

Let us add and subtract $\omega_i\|\boldsymbol{h}_{\widetilde{T}_j\cap\Gamma^c}\|_p^p$ for all pairs of i and j such that $i,j=1,\cdots,L$ and $i\neq j$, and $\omega_i\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p$ for $i=1,\cdots,L$ to the left side of (4.3). Then the left side of (4.3) becomes

$$\begin{aligned}&\sum_{i=1}^{L}\omega_i\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p+\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p+\sum_{i,j,i\neq j}\omega_i\|\boldsymbol{h}_{\widetilde{T}_j\cap\Gamma^c}\|_p^p-\sum_{i\neq j}\omega_i\|\boldsymbol{h}_{\widetilde{T}_j\cap\Gamma^c}\|_p^p\\&\quad+\sum_{i=1}^{L}\omega_i\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p-\sum_{i=1}^{L}\omega_i\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p\\&=\sum_{i=1}^{L}\omega_i\Big(\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p+\sum_{j\neq i}\|\boldsymbol{h}_{\widetilde{T}_j\cap\Gamma^c}\|_p^p\Big)-\sum_{i\neq j}\omega_i\|h_{\widetilde{T}_j\cap\Gamma^c}\|_p^p+(1-\omega)\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p+\sum_{i=1}^{L}\omega_i\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p\\&=\omega\|\boldsymbol{h}_{\Gamma^c}\|_p^p+(1-\omega)\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p-\sum_{j=1}^{L}\Big(\sum_{i\neq j}\omega_i\Big)\|\boldsymbol{h}_{\widetilde{T}_j\cap\Gamma^c}\|_p^p\\&=\omega\|\boldsymbol{h}_{\Gamma^c}\|_p^p+(1-\omega)\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p-\sum_{j=1}^{L}(\omega-\omega_j)\|\boldsymbol{h}_{\widetilde{T}_j\cap\Gamma^c}\|_p^p.\end{aligned}$$

Similarly, we can write

$$\|\boldsymbol{h}_\Gamma\|_p^p = \sum_{i=1}^{L} \|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma}\|_p^p + \|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p.$$

Let us add and subtract $\omega_i\|\boldsymbol{h}_{\widetilde{T}_j\cap\Gamma}\|_p^p$ for all pairs of i and j such that $i,j=1,\cdots,L$ and $i\neq j$, and $\omega_i\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p$ for $i=1,\cdots,L$ to the right side of (4.3), as well as $\omega_i\|\boldsymbol{x}_{\widetilde{T}_j\cap\Gamma^c}\|_p^p$ for $i=1,\cdots,L$ and $i\neq j$, and $\omega_i\|\boldsymbol{x}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p$ for $i=1,\cdots,L$. Then the right side of (4.3) becomes

$$\begin{aligned}&\omega\|\boldsymbol{h}_\Gamma\|_p^p + (1-\omega)\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p - \sum_{i=1}^{L}(\omega-\omega_i)\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma}\|_p^p\\ &\quad + 2\Big(\omega\|\boldsymbol{x}_{\Gamma^c}\|_p^p + (1-\omega)\|\boldsymbol{x}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p - \sum_{i=1}^{L}(\omega-\omega_i)\|\boldsymbol{x}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p\Big).\end{aligned}$$

Let

$$D = \omega\|\boldsymbol{x}_{\Gamma^c}\|_p^p + (1-\omega)\|\boldsymbol{x}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p - \sum_{i=1}^{L}(\omega-\omega_i)\|\boldsymbol{x}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p.$$

Putting these together, we have

$$\begin{aligned}&\omega\|\boldsymbol{h}_{\Gamma^c}\|_p^p + (1-\omega)\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p - \sum_{i=1}^{L}(\omega-\omega_i)\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p\\ \le&\omega\|\boldsymbol{h}_\Gamma\|_p^p + (1-\omega)\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p - \sum_{i=1}^{L}(\omega-\omega_i)\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma}\|_p^p + 2D.\end{aligned} \tag{4.4}$$

But, we can also write $\|\boldsymbol{h}_{\Gamma^c}\|_p^p$ as

$$\|\boldsymbol{h}_{\Gamma^c}\|_p^p = \omega\|\boldsymbol{h}_{\Gamma^c}\|_p^p + \sum_{i=1}^{L}(1-\omega)\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p + (1-\omega)\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p.$$

Solving for $\omega\|\boldsymbol{h}_{\Gamma^c}\|_p^p$ and substituting into (4.4) gives

$$\begin{aligned}&\|\boldsymbol{h}_{\Gamma^c}\|_p^p - \sum_{i=1}^{L}(1-\omega)\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p - (1-\omega)\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p\\ &\quad + (1-\omega)\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma^c}\|_p^p - \sum_{i=1}^{L}(\omega-\omega_i)\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p\\ \le&\omega\|\boldsymbol{h}_\Gamma\|_p^p + (1-\omega)\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p - \sum_{i=1}^{L}(\omega-\omega_i)\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma}\|_p^p + 2D.\end{aligned}$$

Simplifying, we get

$$\begin{aligned}
\|\boldsymbol{h}_{\Gamma^c}\|_p^p \le& \sum_{i=1}^{L}(1-\omega)\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p + \sum_{i=1}^{L}(\omega-\omega_i)\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p + \omega\|\boldsymbol{h}_\Gamma\|_p^p \\
&+ (1-\omega)\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p - \sum_{i=1}^{L}(\omega-\omega_i)\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma}\|_p^p + 2D \\
=& \sum_{i=1}^{L}(1-\omega_i)\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p + \omega\|\boldsymbol{h}_\Gamma\|_p^p + \|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p \\
&- \sum_{i=1}^{L}\omega_i\Big(\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p + \sum_{j=1,j\neq i}^{L}\|\boldsymbol{h}_{\widetilde{T}_j\cap\Gamma}\|_p^p\Big) + 2D \\
=& \sum_{i=1}^{L}(1-\omega_i)\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p + \omega\|\boldsymbol{h}_\Gamma\|_p^p + \|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p - \sum_{i=1}^{L}\omega_i\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p \\
&+ \sum_{i=1}^{L}\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p - \sum_{i=1}^{L}\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p + 2D \qquad (4.5) \\
=& \omega\|\boldsymbol{h}_\Gamma\|_p^p + \|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p - \sum_{i=1}^{L}\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p + \sum_{i=1}^{L}(1-\omega_i)\Big(\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p + \|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p\Big) + 2D \\
=& (\omega-(L-1))\|\boldsymbol{h}_\Gamma\|_p^p + \sum_{i=1}^{L}(1-\omega_i)\Big(\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p + \|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma}\|_p^p\Big) + 2D, \qquad (4.6)
\end{aligned}$$

where in (4.5) we have added zero and observed that

$$\|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p + \sum_{j=1,j\neq i}^{L}\|\boldsymbol{h}_{\widetilde{T}_j\cap\Gamma}\|_p^p = \|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p$$

and in (4.6), we have observed that

$$\sum_{i=1}^{L}\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p = (L-1)\|\boldsymbol{h}_\Gamma\|_p^p + \|\boldsymbol{h}_{\widetilde{T}^c\cap\Gamma}\|_p^p.$$

Then assuming, without loss of generality, $\omega_1 \ge \omega_2 \ge \cdots \ge \omega_L$, and writing $1-\omega_i = 1-\omega_1+\omega_1-\omega_i$ for $i>1$, we have

$$\begin{aligned}
\|\boldsymbol{h}_{\Gamma^c}\|_p^p \le& (\omega-(L-1))\|\boldsymbol{h}_\Gamma\|_p^p + (1-\omega_1)\sum_{i=1}^{L}\Big(\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p + \|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p\Big) \\
&+ \sum_{i=2}^{L}(\omega_1-\omega_i)\Big(\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p + \|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p\Big) + 2D. \qquad (4.7)
\end{aligned}$$

Next, write $\omega_1-\omega_i=\omega_1-\omega_2+\omega_2-\omega_i$ for $i>2$. Then we have

$$\begin{aligned}\|\boldsymbol{h}_{\Gamma^c}\|_p^p \le&(\omega-(L-1))\|\boldsymbol{h}_\Gamma\|_p^p+(1-\omega_1)\sum_{i=1}^{L}\Big(\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p+\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p\Big)+(\omega_1-\omega_2)\\ &\times\sum_{i=2}^{L}\Big(\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p+\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p\Big)+\sum_{i=3}^{L}(\omega_2-\omega_i)\Big(\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p+\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p\Big)\\ &+2D. \end{aligned}\tag{4.8}$$

Continuing in this way gives us

$$\begin{aligned}\|\boldsymbol{h}_{\Gamma^c}\|_p^p \le&(\omega-(L-1))\|\boldsymbol{h}_\Gamma\|_p^p+(1-\omega_1)\sum_{i=1}^{L}\Big(\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p+\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p\Big)\\ &+\sum_{j=2}^{L}(\omega_{j-1}-\omega_j)\sum_{i=j}^{L}\Big(\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p+\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p\Big)+2D. \end{aligned}\tag{4.9}$$

Noting

$$\begin{aligned}&\|\boldsymbol{h}_{\widetilde{T}_i^c\cap\Gamma}\|_p^p=\sum_{j=1,j\neq i}^{L}\|\boldsymbol{h}_{\widetilde{T}_j\cap\Gamma}\|_p^p+\|\boldsymbol{h}_{\Gamma\cap\bigcap_{j=1}^{L}\widetilde{T}_j^c}\|_p^p,\\ &\|\boldsymbol{h}_\Gamma\|_p^p=\sum_{i=1}^{L}\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma}\|_p^p+\|\boldsymbol{h}_{\Gamma\cap\bigcap_{j=1}^{L}\widetilde{T}_j^c}\|_p^p,\\ &\sum_{i=j}^{L}\|\boldsymbol{h}_{\widetilde{T}_i\cap\Gamma^c}\|_p^p+\|\boldsymbol{h}_{\Gamma\cap\bigcap_{i=j}^{L}\widetilde{T}_i^c}\|_p^p=\|\boldsymbol{h}_{\Gamma\cup\bigcup_{i=j}^{L}\widetilde{T}_i\setminus\bigcup_{i=j}^{L}(\widetilde{T}_i\cap\Gamma)}\|_p^p,\end{aligned}$$

for any $j=1,2,\cdots,L$, the above inequality can also be expressed as

$$\begin{aligned}\|h_{\Gamma^c}\|_p^p \le&(\omega-(L-1))\|\boldsymbol{h}_\Gamma\|_p^p+(1-\omega_1)\Big((L-1)\|\boldsymbol{h}_\Gamma\|_p^p+\|\boldsymbol{h}_{\Gamma\cup\bigcup_{i=1}^{L}\widetilde{T}_i\setminus\bigcup_{i=1}^{L}(\widetilde{T}_i\cap\Gamma)}\|_p^p\Big)\\ &+\sum_{j=2}^{L}(\omega_{j-1}-\omega_j)\Big((L-j)\|\boldsymbol{h}_\Gamma\|_p^p+\|\boldsymbol{h}_{\Gamma\cup\bigcup_{i=j}^{L}\widetilde{T}_i\setminus\bigcup_{i=j}^{L}(\widetilde{T}_i\cap\Gamma)}\|_p^p\Big)+2D. \end{aligned}\tag{4.10}$$

Combining the coefficients of $\|\boldsymbol{h}_\Gamma\|_p^p$, we have

$$\begin{aligned}&\sum_{i=1}^{L}\omega_i-(L-1)+(1-\omega_1)(L-1)+\sum_{j=2}^{L}(\omega_{j-1}-\omega_j)(L-j)\\ =&\sum_{i=1}^{L}\omega_i-(L-1)\omega_1+(L-2)\omega_1+\sum_{j=2}^{L-1}(L-(j+1))\omega_j-\sum_{j=2}^{L-1}(L-j)\omega_j\\ =&\sum_{i=2}^{L}\omega_i-\sum_{j=2}^{L-1}\omega_j=\omega_L.\end{aligned}$$

Finally, we obtain that

$$\begin{aligned}\|\boldsymbol{h}_{\Gamma^c}\|_p^p \leq & \omega_L \|\boldsymbol{h}_\Gamma\|_p^p + (1-\omega_1)\|\boldsymbol{h}_{\Gamma \cup \bigcup_{i=1}^L \widetilde{T}_i \setminus \bigcup_{i=1}^L (\widetilde{T}_i \cap \Gamma)}\|_p^p \\ & + \sum_{j=2}^{L} (\omega_{j-1} - \omega_j) \|\boldsymbol{h}_{\Gamma \cup \bigcup_{i=j}^L \widetilde{T}_i \setminus \bigcup_{i=j}^L (\widetilde{T}_i \cap \Gamma)}\|_p^p + 2D.\end{aligned}$$

Thus, we complete the proof. □

The following two technical lemmas will be used to simplify the proof of our main results.

Lemma 4.3 ([9, Lemma V.1]). *Let p and q be two positive numbers. Then*

(a) $\|\boldsymbol{x}\|_p \leq \|\boldsymbol{x}\|_2 |\mathrm{supp}(\boldsymbol{x})|^{\frac{2-p}{2p}}$, *if* $0 < p < 2$,

(b) $\|\boldsymbol{x}\|_p^p \leq (\|\boldsymbol{x}\|_2^2)^{\frac{1}{q}} (\|\boldsymbol{x}\|_{p_1}^{p_1})^{1-\frac{1}{q}}$, *if* $pq > 2$ *and* $q > 1$, *where* $p_1 = (p - \frac{2}{q})(\frac{q}{q-1})$.

Lemma 4.4 ([9, Lemma V.2]). *For $0 < p \leq 1$ and $\Lambda > 0$, the function*

$$g(z) = \frac{p}{2} z^{\frac{2}{p}} + z - \frac{2-p}{2}\Lambda$$

is monotone increasing in $(0, \infty)$. In addition, the following statements hold:

(I) If $0 < \Lambda \leq \frac{2}{2-p}$, there exists a unique point $z_0 \in ((1-p)\Lambda, (1-\frac{p}{2})\Lambda) \subseteq (0,1)$ such that $g(z_0) = 0$.

(II) If $\frac{2}{2-p} < \Lambda < \frac{2+p}{2-p}$, there exists a unique point $z_0 \in ((1-p)\Lambda, 1) \subseteq (0,1)$ such that $g(z_0) = 0$.

(III) If $\Lambda \geq \frac{2+p}{2-p}$, there does not exist a point $z_0 \in (0,\, 1)$ such that $g(z_0) = 0$.

4.2 Proof of Theorem 3.1

Proof. We assume that tk is an integer. When tk is not an integer, it can be treated as in [1] and [9]. Let $\boldsymbol{h} = \hat{\boldsymbol{x}} - \boldsymbol{x}$, where $\hat{\boldsymbol{x}}$ is a minimizer of the weighted ℓ_p-minimization problem (1.2) with $\epsilon = 0$. Then

$$\boldsymbol{\Phi h} = \boldsymbol{0}. \tag{4.11}$$

We prove $\boldsymbol{h} = \boldsymbol{0}$ to show that $\boldsymbol{x}$ could be recovered exactly via the weighted ℓ_p-minimization (1.2).

On the contrary, we suppose here that $\boldsymbol{h} \neq \boldsymbol{0}$, then $\boldsymbol{h}_{\max(dk)} \neq \boldsymbol{0}$, where $\boldsymbol{h}_{\max(dk)}$ is the best dk-term approximation of $\boldsymbol{h}$ and we define

$$\boldsymbol{h}_{-\max(dk)} = \boldsymbol{h} - \boldsymbol{h}_{\max(dk)}.$$

Since T is the support set of the k-sparse vector $\boldsymbol{x}$, we know that $|T| \leq k$. Recall the definition of d in (3.2a),

$$d = \begin{cases} 1, & \omega_1 = \cdots = \omega_L = 1, \\ \max\limits_{i \in \{1,2\cdots,L\}} \Big\{ b_i \Big(1 - \sum\limits_{j=i}^{L} \alpha_j \rho_j + \beta_i\Big) \Big\}, & 0 \leq \prod\limits_{i=1}^{L} \omega_i < 1, \end{cases} \tag{4.12}$$

where

$$\beta_i = \max\Big\{ \sum_{j=i}^{L} \alpha_j \rho_j, \sum_{j=i}^{L} (1-\alpha_j)\rho_j \Big\},$$
$$b_i = \begin{cases} 1, & i = 1, \\ sgn(\omega_{i-1} - \omega_i), & i = 2, \cdots, L. \end{cases}$$

It is clear that $d \geq 1$ and dk is an integer. Thus,

$$\begin{aligned} &\|\boldsymbol{h}_{-\max(dk)}\|_p^p \leq \|\boldsymbol{h}_{T^c}\|_p^p \\ \leq &\omega_L \|\boldsymbol{h}_T\|_p^p + (1-\omega_1) \|\boldsymbol{h}_{T \cup \bigcup_{i=1}^{L} \widetilde{T}_i \setminus \bigcup_{i=1}^{L} (\widetilde{T}_i \cap T)}\|_p^p \\ &\quad + \sum_{j=2}^{L} (\omega_{j-1} - \omega_j) \|\boldsymbol{h}_{T \cup \bigcup_{i=j}^{L} \widetilde{T}_i \setminus \bigcup_{i=j}^{L} (\widetilde{T}_i \cap T)}\|_p^p & (4.13) \\ \leq &\begin{cases} \|\boldsymbol{h}_T\|_p^p, & \omega_1 = \cdots = \omega_L = 1, \\ \omega_L \|\boldsymbol{h}_T\|_p^p + (1-\omega_L)\|\boldsymbol{h}_{\max(dk)}\|_p^p, & 0 \leq \prod\limits_{i=1}^{L} \omega_i < 1, \end{cases} & (4.14) \end{aligned}$$

where the first inequality is from $d \geq 1$ and $|T| \leq k$, the second inequality follows from Lemma 4.2 with $\Gamma = T$ and the last inequality is due to

$$\Big| \Big(T \cup \bigcup_{j=i}^{L} \widetilde{T}_j\Big) \Big\backslash \bigcup_{j=i}^{L} \Big(T \cap \widetilde{T}_j\Big) \Big| \leq k + \sum_{j=i}^{L} \rho_j k - 2 \sum_{j=i}^{L} \alpha_j \rho_j k = k\Big(1 + \sum_{j=i}^{L} \rho_j - 2\sum_{j=i}^{L} \alpha_j \rho_j\Big) \leq dk$$

with

$$\beta_i = \max\Big\{ \sum_{j=i}^{L} \alpha_j \rho_j, \sum_{j=i}^{L} (1-\alpha_j)\rho_j \Big\}.$$

Let

$$\nu = \Big(\frac{\omega_L \|\boldsymbol{h}_T\|_p^p + (1-\omega_1) \|\boldsymbol{h}_{T \cup \bigcup_{i=1}^{L} \widetilde{T}_i \setminus \bigcup_{i=1}^{L} (\widetilde{T}_i \cap T)}\|_p^p + \sum_{j=2}^{L} (\omega_{j-1} - \omega_j) \|\boldsymbol{h}_{T \cup \bigcup_{i=1}^{L} \widetilde{T}_i \setminus \bigcup_{i=1}^{L} (\widetilde{T}_i \cap \mathbb{T})}\|_p^p}{k(t-d)} \Big)^{\frac{1}{p}}. \tag{4.15}$$

Then $\nu \geq 0$. First, we suppose that $\nu = 0$, then we have $\|\boldsymbol{h}_{T^c}\|_p^p = 0$ by (4.13), which implies $\boldsymbol{h}$ is k-sparse. Since the sensing matrix $\boldsymbol{\Phi}$ satisfies the RIP of order tk with $t > d \geq 1$ and (4.11), we have $\boldsymbol{h} = \boldsymbol{0}$. Therefore, $\boldsymbol{x}$ is exactly recovered by (1.2) with $\epsilon = 0$.

For $\nu>0$, we divide the vector $\boldsymbol{h}_{-\max(dk)}$ into two parts, i.e.,

$$\boldsymbol{h}_{-\max(dk)}=\boldsymbol{h}^{(1)}+\boldsymbol{h}^{(2)}, \tag{4.16}$$

where

$$\boldsymbol{h}^{(1)}=\boldsymbol{h}_{-\max(dk)}\cdot\chi_{\{i||\boldsymbol{h}_{-\max(dk)}(i)|>\nu\}}, \tag{4.17a}$$

$$\boldsymbol{h}^{(2)}=\boldsymbol{h}_{-\max(dk)}\cdot\chi_{\{i||\boldsymbol{h}_{-\max(dk)}(i)|\le\nu\}}. \tag{4.17b}$$

Then

$$\|\boldsymbol{h}^{(1)}\|_p^p\le\|\boldsymbol{h}_{-\max(dk)}\|_p^p\le k(t-d)\nu^p$$

by (4.13) and (4.15). Denote $|\mathrm{supp}(\boldsymbol{h}^{(1)})|=\|\boldsymbol{h}^{(1)}\|_0=m$. Since all non-zero entries of $\boldsymbol{h}^{(1)}$ have absolute value larger than ν, we have

$$(t-d)k\nu^p\ge\|\boldsymbol{h}_{-\max(dk)}\|_p^p\ge\|\boldsymbol{h}^{(1)}\|_p^p=\sum_{i\in\mathrm{supp}(\boldsymbol{h}^{(1)})}|\boldsymbol{h}^{(1)}(i)|^p\ge m\nu^p. \tag{4.18}$$

By (4.18) and $\nu\neq0$, one has

$$|\mathrm{supp}(\boldsymbol{h}^{(1)})|=m\le(t-d)k$$

and

$$|\mathrm{supp}(\boldsymbol{h}_{\max(dk)})+\mathrm{supp}(\boldsymbol{h}^{(1)})|\le dk+|\mathrm{supp}(\boldsymbol{h}^{(1)})|\le dk+(t-d)k=tk. \tag{4.19}$$

Moreover,

$$\|\boldsymbol{h}^{(2)}\|_\infty\overset{(a)}{\le}\nu,\quad \|\boldsymbol{h}^{(2)}\|_p^p\overset{(b)}{=}\|\boldsymbol{h}_{-\max(dk)}\|_p^p-\|\boldsymbol{h}^{(1)}\|_p^p\overset{(c)}{\le}((t-d)k-m)\nu^p, \tag{4.20}$$

where (a) is from (4.17b), (b) is due to (4.16) and (c) follows from (4.18). Applying Lemma 4.1 with $L=k(t-d)-m$ and $\rho=\nu$, we can express $\boldsymbol{h}^{(2)}$ as a convex combination of $(k(t-d)-m)$-sparse vectors, i.e., $\boldsymbol{h}^{(2)}=\sum_i\lambda_i\boldsymbol{u_i}$, where $\lambda_i>0$, $\sum_i\lambda_i=1$, $\boldsymbol{u_i}$ is $(k(t-d)-m)$-sparse and $\mathrm{supp}(\boldsymbol{u_i})\subseteq\mathrm{supp}(\boldsymbol{h}^{(2)})$. By (4.16), we have

$$\langle\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)},\boldsymbol{u_i}\rangle=0. \tag{4.21}$$

Furthermore, by (4.1),

$$\begin{aligned}\Sigma_i\lambda_i\|\boldsymbol{u_i}\|_2^2&\le\min\left\{\frac{n}{L}\|\boldsymbol{h}^{(2)}\|_2^2,\nu^p\|\boldsymbol{h}^{(2)}\|_{2-p}^{2-p}\right\}\le\nu^p\|\boldsymbol{h}^{(2)}\|_{2-p}^{2-p}\\&\le\nu^p(\|\boldsymbol{h}^{(2)}\|_2^2)^{\frac{2-2p}{2-p}}(\|\boldsymbol{h}^{(2)}\|_p^p)^{\frac{p}{2-p}}\\&\le\nu^p(\|\boldsymbol{h}^{(2)}\|_2^2)^{\frac{2-2p}{2-p}}\Big(((t-d)k-m)\nu^p\Big)^{\frac{p}{2-p}}\\&\le(\|\boldsymbol{h}^{(2)}\|_2^2)^{\frac{2-2p}{2-p}}\Big(k(t-d)\nu^2\Big)^{\frac{p}{2-p}},\end{aligned} \tag{4.22}$$

where the third inequality is from Lemma 4.3(b), and the fourth inequality follows from (4.20). By (4.15), we have

$$\begin{aligned}
&k(t-d)\nu^2\\
=&(k(t-d))^{1-\frac{2}{p}}\Big(\omega_L\|\boldsymbol{h}_{\mathbb{T}}\|_p^p+(1-\omega_1)\|\boldsymbol{h}_{T\cup\bigcup_{i=1}^L\widetilde{T}_i\setminus\bigcup_{i=1}^L(\widetilde{T}_i\cap T)}\|_p^p\\
&\quad+\sum_{j=2}^L(\omega_{j-1}-\omega_j)\|\boldsymbol{h}_{T\cup\bigcup_{i=1}^L\widetilde{T}_i\setminus\bigcup_{i=1}^L(\widetilde{T}_i\cap T)}\|_p^p\Big)^{\frac{2}{p}}\\
\le&(k(t-d))^{1-\frac{2}{p}}\Big(\omega_L|T|^{\frac{2-p}{2}}\|\boldsymbol{h}_T\|_p^p\\
&\quad+(1-\omega_1)|T\cup\bigcup_{i=1}^L\widetilde{T}_i\setminus\bigcup_{i=1}^L(\widetilde{T}_i\cap T)|^{\frac{2-p}{2}}\|\boldsymbol{h}_{\mathbb{T}}\|_p^p+(1-\omega_1)\|\boldsymbol{h}_{T\cup\bigcup_{i=1}^L\widetilde{T}_i\setminus\bigcup_{j=i}^L(\widetilde{T}_i\cap T)}\|_2^p\\
&\quad+\sum_{i=2}^L(\omega_{i-1}-\omega_i)|\bigcup_{j=i}^L\widetilde{T}_i\setminus\bigcup_{j=i}^L(\widetilde{T}_j\cap T)|^{\frac{2-p}{2}}\|\boldsymbol{h}_T\|_p^p+(1-\omega_1)\|\boldsymbol{h}_{T\cup\bigcup_{j=i}^L\widetilde{T}_j\setminus\bigcup_{j=i}^L(\widetilde{T}_j\cap T)}\|_2^p)\Big)^{\frac{2}{p}}\\
\le&(k(t-d))^{1-\frac{2}{p}}k^{\frac{2-p}{p}}\Big(\omega_L+(1-\omega_1)\Big(1+\sum_{j=1}^L\rho_j-2\sum_{j=1}^L\alpha_j\rho_j\Big)^{\frac{2-p}{2}}\\
&\quad+\sum_{i=2}^L(\omega_{i-1}-\omega_i)\Big(1+\sum_{j=1}^L\rho_j-2\sum_{j=1}^L\alpha_j\rho_j\Big)^{\frac{2-p}{2}}\Big)^{\frac{2}{p}}\|\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}\|_2^2\\
=&(t-d)^{1-\frac{2}{p}}\Big(\omega_L+(1-\omega_1)\Big(1+\sum_{j=1}^L\rho_j-2\sum_{j=1}^L\alpha_j\rho_j\Big)^{\frac{2-p}{2}}\\
&\quad+\sum_{i=2}^L(\omega_{i-1}-\omega_i)\Big(1+\sum_{j=1}^L\rho_j-2\sum_{j=1}^L\alpha_j\rho_j\Big)^{\frac{2-p}{2}}\Big)^{\frac{2}{p}}\|\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}\|_2^2,
\end{aligned}\tag{4.23}$$

where the first inequality is due to $0<p\le 1$ and Lemma 4.3(a) and the second inequality is from $|T|\le k$ and

$$\begin{aligned}
&\Big|T\cup\bigcup_{j=i}^L\tilde{T}_j\Big\backslash\bigcup_{j=i}^L(T\cap\tilde{T}_j)\Big|\\
=&k+\sum_{j=i}^L\rho_jk-2\sum_{j=i}^L\alpha_j\rho_jk=k\Big(1+\sum_{j=i}^L\rho_j-2\sum_{j=i}^L\alpha_j\rho_j\Big)\le dk.
\end{aligned}$$

Then, by (4.22) and (4.23),

$$\begin{aligned}
\sum_i\lambda_i\|\boldsymbol{u_i}\|_2^2\le&(\|\boldsymbol{h}^{(2)}\|_2^2)^{\frac{2-2p}{2-p}}(t-d)^{-1}\Big(\omega_L+(1-\omega_1)\Big(1+\sum_{j=1}^L\rho_j-2\sum_{j=1}^L\alpha_j\rho_j\Big)^{\frac{2-p}{2}}\\
&+\sum_{i=2}^L(\omega_{i-1}-\omega_i)\Big(1+\sum_{j=1}^L\rho_j-2\sum_{j=1}^L\alpha_j\rho_j\Big)^{\frac{2-p}{2}}\Big)^{\frac{2}{2-p}}(\|\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}\|_2^2)^{\frac{p}{2-p}}
\end{aligned}$$

$$= \Theta\mu^{\frac{2-2p}{2-p}} \|\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}\|_2^2, \tag{4.24}$$

where the equality is due to (3.4) and

$$\mu = \frac{\|\boldsymbol{h}^{(2)}\|_2^2}{\|\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}\|_2^2}. \tag{4.25}$$

We have $0 \leq \mu \leq 1$ since

$$\begin{aligned} \|\boldsymbol{h}^{(2)}\|_2^2 \leq & \|\boldsymbol{h}^{(2)}\|_\infty^{2-p} \|\boldsymbol{h}^{(2)}\|_p^p \\ \leq & \|\boldsymbol{h}^{(2)}\|_\infty^{2-p} \|\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}\|_p^p \\ \leq & \min_{i\in \operatorname{supp}(\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)})} |h_i|^{2-p} \|\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}\|_p^p \\ \leq & \|\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}\|_2^2, \end{aligned}$$

where the second inequality is from (4.14), $|T| \leq k \leq dk$ with $d \geq 1$.

For $\eta \in \mathbb{R}$, let

$$\boldsymbol{\theta_i} = \boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)} + \eta \boldsymbol{u_i},$$

then

$$\begin{aligned} \sum_j \lambda_j \boldsymbol{\theta_j} - \frac{p}{2}\boldsymbol{\theta_i} =& \Big(1 - \frac{p}{2}\Big)(\boldsymbol{h}_{\max(dk} + \boldsymbol{h}^{(1)}) + \eta \sum_j \lambda_j \boldsymbol{u_j} - \frac{p}{2}\eta \boldsymbol{u_i} \\ \overset{(a)}{=} & \Big(1 - \frac{p}{2}\Big)(\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}) + \eta \boldsymbol{h}^{(2)} - \frac{p}{2}\eta \boldsymbol{u_i} \\ \overset{(b)}{=} & \Big(1 - \frac{p}{2} - \eta\Big)(\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}) + \eta \boldsymbol{h} - \frac{p}{2}\eta \boldsymbol{u_i}, \end{aligned} \tag{4.26}$$

i.e.,

$$\sum_j \lambda_j \boldsymbol{\theta_j} - \frac{p}{2}\boldsymbol{\theta_i} - \eta \boldsymbol{h} = \Big(1 - \frac{p}{2} - \eta\Big)\Big(\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}\Big) - \frac{p}{2}\eta \boldsymbol{u_i},$$

where (a) is due to $\boldsymbol{h}^{(2)} = \sum_i \lambda_i \boldsymbol{u_i}$, and (b) is from

$$\boldsymbol{h} = \boldsymbol{h}_{\max(dk)} + \boldsymbol{h}_{\max(dk)^c} \quad \text{and} \quad \boldsymbol{h}_{\max(dk)^c} = \boldsymbol{h}^{(1)} + \boldsymbol{h}^{(2)}.$$

Due to

$$\|\boldsymbol{u_i}\|_0 \leq k(t-d) - |\operatorname{supp}(\boldsymbol{h}^{(2)})|$$

and the definition of $\boldsymbol{h}_{\max(dk)}$, the vectors $\boldsymbol{\theta_i}$,

$$\sum_j \lambda_j \boldsymbol{\theta_j} - \frac{p}{2}\boldsymbol{\theta_i} - \eta \boldsymbol{h} \quad \text{and} \quad \Big(1 - \frac{p}{2} - \eta\Big)(\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}) - \frac{p}{2}\eta \boldsymbol{u_i}$$

are all tk-sparse. By (4.11) and (4.26), we have

$$\begin{aligned}
&\sum_i \lambda_i \Big\|\boldsymbol{\Phi}\Big(\sum_j \lambda_j \boldsymbol{\theta}_j - \frac{p}{2}\boldsymbol{\theta}_i\Big)\Big\|_2^2 \\
=&\sum_i \lambda_i \Big\|\boldsymbol{\Phi}\Big(\Big(1-\frac{p}{2}-\eta\Big)(\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}) - \frac{p}{2}\eta \boldsymbol{u}_i\Big)\Big\|_2^2 \\
\leq&(1+\delta_{tk})\sum_i \lambda_i \Big\|\Big(1-\frac{p}{2}-\eta\Big)(\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}) - \frac{p}{2}\eta \boldsymbol{u}_i\Big\|_2^2 \\
=&(1+\delta_{tk})\Big[\Big(1-\frac{p}{2}-\eta\Big)^2 \|\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}\|_2^2 + \frac{p^2\eta^2}{4}\sum_i \lambda_i \|\boldsymbol{u}_i\|_2^2\Big],
\end{aligned} \tag{4.27}$$

where the first inequality is from

$$\Big(1-\frac{p}{2}-\eta\Big)(\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}) - \frac{p}{2}\eta \boldsymbol{u}_i$$

is tk-sparse and the last equality is due to (4.21). Since $\boldsymbol{\theta}_i$ is a tk-sparse vectors, we have

$$\begin{aligned}
&\frac{1-p}{2}\sum_{i,j}\lambda_i\lambda_j \|\boldsymbol{\Phi}(\boldsymbol{\theta}_i-\boldsymbol{\theta}_j)\|_2^2 \\
=&\eta^2\frac{1-p}{2}\sum_{i,j}\lambda_i\lambda_j \|\boldsymbol{\Phi}(\boldsymbol{u}_i-\boldsymbol{u}_j)\|_2^2 \\
\leq&(1+\delta_{tk})\eta^2\frac{1-p}{2}\sum_{i,j}\lambda_i\lambda_j \|\boldsymbol{u}_i-\boldsymbol{u}_j\|_2^2 \\
=&(1+\delta_{tk})\eta^2(1-p)\Big(\sum_i \lambda_i\|\boldsymbol{u}_i\|_2^2 - \Big\|\sum_i \lambda_i \boldsymbol{u}_i\Big\|_2^2\Big) \\
=&(1+\delta_{tk})\eta^2(1-p)\Big(\sum_i \lambda_i\|\boldsymbol{u}_i\|_2^2 - \|\boldsymbol{h}^{(2)}\|_2^2\Big),
\end{aligned} \tag{4.28}$$

where the inequality is from that $\boldsymbol{u}_i$ is $(k(t-d)-m)$-sparse and $d<t\leq 2d$. $\boldsymbol{u}_i-\boldsymbol{u}_j$ is tk-sparse as $d<t\leq 2d$ since

$$tk-2(k(t-d)-m)=k(2d-t)+m\geq 0.$$

Since $\boldsymbol{\theta}_i$ is tk-sparse, it follows that

$$\begin{aligned}
&\Big(1-\frac{p}{2}\Big)^2\sum_i \lambda_i\|\boldsymbol{\Phi}\boldsymbol{\theta}_i\|_2^2 \geq (1-\delta_{tk})\Big(1-\frac{p}{2}\Big)^2\sum_i \lambda_i\|\boldsymbol{\theta}_i\|_2^2 \\
=&(1-\delta_{tk})\Big(1-\frac{p}{2}\Big)^2\Big(\|\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}\|_2^2 + \eta^2\sum_i \lambda_i\|\boldsymbol{u}_i\|_2^2\Big),
\end{aligned} \tag{4.29}$$

where the equality is from the definition of $\boldsymbol{\theta}_i$ and (4.21).

By (4.27)-(4.29) and the following identity (see [21, (21)])

$$\sum_i \lambda_i \Big\| \boldsymbol{\Phi}\Big(\sum_j \lambda_j \boldsymbol{\theta_j} - \frac{p}{2}\boldsymbol{\theta_i}\Big)\Big\|_2^2 + \frac{1-p}{2} \sum_{i,j} \lambda_i \lambda_j \|\boldsymbol{\Phi}(\boldsymbol{\gamma_i} - \boldsymbol{\theta_j})\|_2^2$$
$$- \Big(1-\frac{p}{2}\Big)^2 \sum_i \lambda_i \|\boldsymbol{\Phi}\boldsymbol{\theta}_i\|_2^2 = 0, \tag{4.30}$$

we have

$$\begin{aligned}
0 \leq &(1+\delta_{tk}) \Big[\Big(1-\frac{p}{2}-\eta\Big)^2 \|\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}\|_2^2 \\
&+ \eta^2 \Big(\frac{p^2}{4} + (1-p)\Big) \sum_i \lambda_i \|\boldsymbol{u_i}\|_2^2 - \eta^2 (1-p) \|\boldsymbol{h}^{(2)}\|_2^2 \Big] \\
&- (1-\delta_{tk}) \Big(1-\frac{p}{2}\Big)^2 \Big(\|\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}\|_2^2 + \eta^2 \sum_i \lambda_i \|\boldsymbol{u_i}\|_2^2 \Big) \\
= &(1+\delta_{tk}) \Big[\Big(1-\frac{p}{2}-\eta\Big)^2 \|\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}\|_2^2 - \eta^2(1-p)\|\boldsymbol{h}^{(2)}\|_2^2 \Big] \\
&- (1-\delta_{tk}) \Big(1-\frac{p}{2}\Big)^2 \|\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}\|_2^2 \\
&+ 2\delta_{tk} \Big(1-\frac{p}{2}\Big)^2 \eta^2 \sum_i \lambda_i \|\boldsymbol{u_i}\|_2^2.
\end{aligned}$$

From (4.25), (4.24) and the above inequality, it follows that

$$\begin{aligned}
0 \leq &\bigg((1+\delta_{tk}) \Big(\Big(1-\frac{p}{2}-\eta\Big)^2 - \eta^2(1-p)\mu \Big) - (1-\delta_{tk}) \Big(1-\frac{p}{2}\Big)^2 \\
&+ 2\delta_{tk} \Big(1-\frac{p}{2}\Big)^2 \eta^2 \Theta \mu^{\frac{2-2p}{2-p}} \bigg) \|\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}\|_2^2 \\
= &\Big[(\eta^2 - (2-p)\eta - \eta^2(1-p)\mu) + \delta_{tk} \Big(\Big(1-\frac{p}{2}-\eta\Big)^2 + \Big(1-\frac{p}{2}\Big)^2 \\
&+ 2\Big(1-\frac{p}{2}\Big)^2 \eta^2 \Theta \mu^{\frac{2-2p}{2-p}} - \eta^2(1-p)\mu \Big) \Big] \|\boldsymbol{h}_{\max(dk)} + \boldsymbol{h}^{(1)}\|_2^2.
\end{aligned} \tag{4.31}$$

Next, let the arbitrary vector η satisfies

$$\eta = \frac{2-p}{\sqrt{(1-(1-p)\mu)^2 + (2-p)^2 \Theta \mu^{\frac{2-2p}{2-p}}} + 1 - (1-p)\mu}. \tag{4.32}$$

By $0 < p \leq 1$ and $0 \leq \mu \leq 1$, it is clear that $0 < \eta < \frac{2-p}{1-(1-p)\mu}$. Moreover, we have

$$\begin{aligned}
\eta^2 - (2-p)\eta - \eta^2(1-p)\mu = &\eta^2 \Big(1 - (1-p)\mu - (2-p)\frac{1}{\eta} \Big) \\
\overset{(a)}{=} &- \eta^2 \sqrt{(1-(1-p)\mu)^2 + (2-p)^2 \Theta \mu^{\frac{2-2p}{2-p}}}
\end{aligned}$$

and

$$
\begin{aligned}
&\Big(1-\frac{p}{2}-\eta\Big)^2+\Big(1-\frac{p}{2}\Big)^2+2\Big(1-\frac{p}{2}\Big)^2\eta^2\Theta\mu^{\frac{2-2p}{2-p}}-\eta^2(1-p)\mu\\
=&\eta^2\Big(1-(1-p)\mu+\frac{1}{2}(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}+\frac{(2-p)^2}{2\eta^2}-(2-p)\frac{1}{\eta}\Big)\\
\overset{(b)}{=}&\eta^2\Big(1-(1-p)\mu+\frac{1}{2}(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}+\frac{1}{2}\Big(\sqrt{(1-(1-p)\mu)^2+(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}}\\
&\quad+1-(1-p)\mu\Big)^2-\Big(\sqrt{(1-(1-p)\mu)^2+(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}}+1-(1-p)\mu\Big)\Big)\\
=&\eta^2\sqrt{(1-(1-p)\mu)^2+(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}}\Big(\sqrt{(1-(1-p)\mu)^2+(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}}-(1-p)\mu\Big),
\end{aligned}
$$

where (a) and (b) are from (4.32). Therefore, from (4.31), it follows that

$$
\begin{aligned}
&-\eta^2\sqrt{(1-(1-p)\mu)^2+(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}}\Big[1-\delta_{tk}\Big(\sqrt{(1-(1-p)\mu)^2+(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}}\\
&\quad-(1-p)\mu\Big)\Big]\|\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}\|_2^2\geq 0. \qquad (4.33)
\end{aligned}
$$

Define a function

$$
f(\mu)=\sqrt{(1-(1-p)\mu)^2+(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}}-(1-p)\mu,
$$

where $0\leq\mu\leq 1$. If $\Theta=0$, then $f(\mu)=1-2(1-p)\mu\leq 1$. In this case, (4.33) is a contradiction from $\delta_{tk}<1$. In the following, we assume that $\Theta>0$. By some elementary calculation, we have

$$
\begin{aligned}
f'(\mu)=&\frac{-2(1-p)(2-p)\Theta\mu^{-\frac{2p}{2-p}}}{\sqrt{(1-(1-p)\mu)^2+(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}}}\\
&\cdot\Big[\frac{\frac{p}{2}\mu^{\frac{p}{2-p}\frac{2}{p}}+\mu^{\frac{p}{2-p}}-\frac{2-p}{2}\Theta}{(-1+(1-p)\mu)+(2-p)\Theta\mu^{\frac{-p}{2-p}}+\sqrt{(1-(1-p)\mu)^2+(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}}}\Big]\\
=&\frac{-2(1-p)(2-p)\Theta\mu^{-\frac{2p}{2-p}}}{\sqrt{(1-(1-p)\mu)^2+(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}}}\\
&\cdot\Big[\frac{g(\mu^{\frac{p}{2-p}})}{(-1+(1-p)\mu)+(2-p)\Theta\mu^{\frac{-p}{2-p}}+\sqrt{(1-(1-p)\mu)^2+(2-p)^2\Theta\mu^{\frac{2-2p}{2-p}}}}\Big],
\end{aligned}
$$

where

$$
g(z)=\frac{p}{2}z^{\frac{2}{p}}+z-\frac{2-p}{2}\Theta.
$$

We will use Lemma 4.4 with $z=\mu^{\frac{p}{2-p}}$ to analyze the extreme value of $g(z)$ according to the value of Θ.

(I) For $0<\Theta<\frac{2+p}{2-p}$, by Lemma 4.4 with $z=\mu^{\frac{p}{2-p}}$, a unique point

$$z_0\in\Big((1-p)\Theta,\min\Big(\Big(1-\frac{p}{2}\Big)\Theta,1\Big)\Big)$$

satisfies

$$\begin{cases} g(z)<0, & 0\le z<z_0,\\ g(z)=0, & z=z_0,\\ g(z)>0, & z_0<z\le 1,\end{cases}$$

which implies that

$$\begin{cases} f'(\mu)>0, & 0\le \mu<z_0^{\frac{2-p}{p}},\\ f'(\mu)=0, & \mu=z_0^{\frac{2-p}{p}},\\ f'(\mu)>0, & z_0^{\frac{2-p}{p}}<\mu\le 1.\end{cases}$$

Therefore, when $\mu=z_0^{\frac{2-p}{p}}$, the function $f(\mu)$ achieves its maximal value that

$$\begin{aligned} f(z_0^{\frac{2-p}{p}}) =&\sqrt{\Big(1-(1-p)z_0^{\frac{2-p}{p}}\Big)^2+(2-p)^2\Theta\Big(z_0^{\frac{2-p}{p}}\Big)^{\frac{2-2p}{2-p}}}-(1-p)z_0^{\frac{2-p}{p}}\\ =&\frac{(2-p)\Theta-z_0}{z_0}. \end{aligned}\tag{4.34}$$

By (3.5), (4.34) and (4.33), there is a contradiction under the hypothesis

$$\|\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}\|_2\neq 0.$$

Then

$$\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}=\mathbf{0}.$$

Due to the definition of $\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}$, we have

$$\boldsymbol{h}=\mathbf{0}.$$

(II) For $\Theta\ge\frac{2+p}{2-p}$, by Lemma 4.4 with $z=\mu^{\frac{p}{2-p}}$, $g(z)<0$ for $0\le\mu<1$, which means that $f'(\mu)>0$. Therefore, when $\mu=1$, $f(\mu)$ achieves its maximal value that

$$f_{\max}(1)=\sqrt{p^2+(2-p)^2\Theta}-(1-p).\tag{4.35}$$

By (3.5), (4.35) and (4.33), there is a contradiction under the hypothesis

$$\|\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}\|\neq 0.$$

Then

$$\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}=\mathbf{0}.$$

Due to the definition of $\boldsymbol{h}_{\max(dk)}+\boldsymbol{h}^{(1)}$, we have $\boldsymbol{h}=\mathbf{0}$. In conclusion, we complete the proof of Theorem 3.1. $\square$

4.3 Proof of Corollary 3.1

Proof. By $p=1$ and (3.4),

$$\begin{aligned}\Theta =&(t-d)^{-1}\Big(\omega_L+(1-\omega_1)\Big(1+\sum_{j=1}^{L}\rho_j-2\sum_{j=1}^{L}\alpha_j\rho_j\Big)^{\frac{2-p}{2}}\\ &+\sum_{i=2}^{L}(\omega_{i-1}-\omega_i)\Big(1+\sum_{j=1}^{L}\rho_j-2\sum_{j=1}^{L}\alpha_j\rho_j\Big)^{\frac{2-p}{2}}\Big)^2.\end{aligned}$$

On one hand, the only positive solution z_0 of Eq. (3.6) with $p=1$ is $-1+\sqrt{1+\Theta}$.
From $p=1$ and $z_0=-1+\sqrt{1+\Theta}$, it follows that

$$\frac{z_0}{(2-p)\Theta-z_0}=\frac{-1+\sqrt{1+\Theta}}{\Theta-(-1+\sqrt{1+\Theta})}=\frac{1}{\sqrt{1+\Theta}}.$$

On the other hand, for $p=1$,

$$\frac{1}{\sqrt{p^2+(2-p)^2\Theta}-(1-p)}=\frac{1}{\sqrt{1+\Theta}}.$$

By Theorem 3.1, the condition (3.7) guarantees the exact recovery of x. $\square$

4.4 Proof of Theorem 3.2

Proof. Theorem 3.2 can be proved by following the routine proofs of Theorem III.10 in [9] and Theorem 3.1 in this paper. We omit the details. $\square$

Acknowledgements

This work is in honor of eighties birthday of Professor Shanzhen Lu and supported by the NSF of China (Nos. 11871109, 11901037 and 11801509) and NSAF (Grant No. U1830107), CAEP Foundation (Grant No. CX20200027).

References

[1] T. T. Cai, and A. Zhang, Sparse representation of a polytope and recovery of sparse signals and low-rank matrices, IEEE Trans. Inf. Theory, 60(1) (2014), 122–132.
[2] E. J. Candès, J. Romberg and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Commun. Pure Appl. Math., 59(8) (2006), 1207–1223.
[3] W. Chen, and Y. Li, Recovery of signals under the condition on RIC and ROC via prior support information, Appl. Comput. Harmon. Anal., 46 (2019), 417–430.
[4] W. Chen, Y. Li and G. Wu, Recovery of signals under the high order RIP condition via prior support information, Signal Process., 153 (2018), 83–94.

[5] D. Donoho, Compressed sensing, IEEE Trans. Inf. Theory, 52(4) (2006), 1289–1306.

[6] M. P. Friedlander, H. Mansour, R. Saab and O. Yilmaz, Recoverying compressively sampled signals using partial support information, IEEE Trans. Inf. Theory, 58(2) (2012), 1122–1134.

[7] H. Ge, and W. Chen, A sharp recovery condition for sparse signals with partial support information via orthogonal matching pursuit, Circuits, Systems, and Signal Processing, 38 (2019), 3295–3320.

[8] H. Ge and W. Chen, Recovery of signals by a weighted ℓ_2/ℓ_1 minimization under arbitrary prior support information, Signal Process., 148 (2018), 288–302.

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

[10] N. Ghadermarzy, H. Mansour and O. Yilmaz, Non-convex compressed sensing using partial support information, Sampling Theory Signal Image Process., 13(3) (2014), 251–272.

[11] S. He, Y. Wang, J. Wang, and Z. Xu, Block-sparse compressed sensing with partially known signal support via non-convex minimisation, IET Signal Processing, 10(7) (2016), 717–723.

[12] C. Herzet, C. Soussen, J. Idier and R. Gribonval, Exact recovery conditions for sparse representations with partial support information, IEEE Trans. Inf. Theory, 59(11) (2013), 7509–7524.

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

[14] L. Jacques, A short note compressed sensing with partially known signal support, Signal Process., 90(12) (2010), 3308–3312.

[15] N. B. Karahanoglu and H. Erdogan, Online Recovery Guarantees and Analytical Results for OMP, Mathematics, 2012.

[16] M. A. Khajehnejad, W. Xu, A. S. Avestimehr and B. Hassibi, Weighted ℓ_1-minimization for sparse recovery with prior information, IEEE Int. Symp. Inf. Theory, ISIT, (2009), 483–487.

[17] D. Needell, R. Saab and T. Woolf, Weighted ℓ_1-minimization for sparse recovery under arbitrary prior information, Information and Inference: A J. IMA, 6 (2007), 284–309.

[18] A. A. Saleh, F. Alajaji, W. Y. Chan, Compressed sensing with non-gaussian noise and partial support information, IEEE Signal Processing Lett., 20(10) (2015), 1703–1707.

[19] N. Vaswani and W. Lu, Modified-CS: modifying compressive sensing for problems with partially known support, IEEE Trans. Signal Process., 58(9) (2010), 4595–4607.

[20] G. Xu and Z. Xu, On the ℓ_1-norm invariant convex k-sparse decomposition of signals, Journal of the Operations Research Society of China, 1(4) (2013), 537–541.

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

[22] R. Zhang, and S. Li, A proof of conjecture on restricted isometry property constants $\delta_{tk}(0 < t < \frac{4}{3})$, IEEE Trans. Inf. Theory, 64(3) (2018), 1699–1705.