

# The high-order block RIP for non-convex block-sparse compressed sensing

Jianwen Huang<sup>1</sup>, Feng Zhang<sup>2,3</sup>, Xinling Liu<sup>2,3</sup>, Hailing Wang<sup>2,3</sup>, Jinyao Hou<sup>2,3</sup>,  
Wendong Wang<sup>2</sup>

<sup>1</sup>*School of Mathematics & Statistics, Tianshui Normal University, Tianshui 741001 China*

<sup>2</sup>*School of Mathematic & Statistics, Southwest University, Chongqing 400715 China*

<sup>3</sup>*College of Artificial Intelligence, Southwest University, Chongqing 400715 China*

**Abstract.** This paper concentrates on the recovery of block-sparse signals, which is not only sparse but also nonzero elements are arrayed into some blocks (clusters) rather than being arbitrary distributed all over the vector, from linear measurements. We establish high-order sufficient conditions based on block RIP to ensure the exact recovery of every block  $s$ -sparse signal in the noiseless case via mixed  $l_2/l_p$  minimization method, and the stable and robust recovery in the case that signals are not accurately block-sparse in the presence of noise. Additionally, a lower bound on necessary number of random Gaussian measurements is gained for the condition to be true with overwhelming probability. Furthermore, the numerical experiments conducted demonstrate the performance of the proposed algorithm.

**Key words.** Compressed sensing; block restricted isometry property; block sparsity; mixed  $l_2/l_p$  minimization

## 1 Introduction

Block-sparse signal recovery (BSR) appears in some fields of sparse modelling and machine learning, including color imaging [13], equalization of sparse communication channels [2], multi-response linear regression [3] and imagine annotation [4] and so forth. Essentially, the important problem in BSR is how to reconstruct a block-sparse or approximately block-sparse signal from a linear system. Commonly, one thinks over the below model:

$$y = \Phi x + e,$$

where  $y \in \mathbb{R}^n$  is the observation measurement,  $\Phi \in \mathbb{R}^{n \times N}$  is a known measurement matrix (or sensing matrix) with  $n < N$ , and  $e \in \mathbb{R}^n$  is a vector of measurement errors. Generally, the conventional compressed sensing (CS) simply thinks out the sparsity of the signal to be recovered, however it doesn't consider any additional structure, i.e., non-zero elements appear in blocks (or clusters) rather than being arbitrarily spread all over

the vector. We call these signals as the block-sparse signals. In order to define block sparsity, we need to give several additional notations. Suppose that the signal  $x$  over the block index set  $\mathcal{I} = \{d_1, d_2, \dots, d_M\}$ , then we can describe the signal  $x$  as

$$x = \underbrace{[x_1, \dots, x_{d_1}]}_{x[1]}, \underbrace{[x_{d_1+1}, \dots, x_{d_1+d_2}]}_{x[2]}, \dots, \underbrace{[x_{N-d_M+1}, \dots, x_N]}_{x[M]}^\top, \quad (1.1)$$

where  $x[i]$  represents the  $i$ th block of  $x$  and  $N = \sum_{i=1}^M d_i$ . We call a vector  $x$  as block  $s$ -sparse over  $\mathcal{I} = \{d_1, d_2, \dots, d_M\}$  if  $x[i]$  is non-zero for no more than  $s$  indices  $i$  [5]. In fact, if the block-sparse structure of signal is neglected, the conventional CS doesn't efficiently treat such structured signal. To reconstruct block-sparse signal, researchers [5] [6] proposed the following mixed  $l_2/l_1$ -minimization:

$$\hat{x} = \arg \min_{\tilde{x} \in \mathbb{R}^N} \|\tilde{x}\|_{2,\mathcal{I}} \text{ s.t. } y - \Phi \tilde{x} \in \mathcal{B}, \quad (1.2)$$

where  $\|x\|_{2,\mathcal{I}} = \sum_{i=1}^M \|x[i]\|_2$  is the mixed  $l_2/l_1$  norm of a vector  $x$ . The set  $\mathcal{B}$  stands for some noise structure,

$$\mathcal{B}^{l_2}(\rho) := \{e : \|e\|_2 \leq \rho\} \quad (1.3)$$

and

$$\mathcal{B}^{DS}(\rho) := \{e : \|\Phi^\top e\|_\infty \leq \rho\}, \quad (1.4)$$

where  $\Phi^\top$  represents the conjugate transpose of the matrix  $\Phi$ . (1.2) is a convex optimization issue and could be converted into a second-order cone program, so can be solved efficiently.

To study the theoretical performance of mixed  $l_2/l_1$ -minimization, Eldar and Mishali [5] proposed the definition of block restricted isometry property (block RIP).

**Definition 1.1.** (Block RIP [5]) Given a matrix  $\Phi \in \mathbb{R}^{n \times N}$ , for every block  $s$ -sparse  $x \in \mathbb{R}^N$  over  $\mathcal{I} = \{d_1, d_2, \dots, d_M\}$ , there is a positive number  $\delta \in (0, 1)$ , if

$$(1 - \delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta)\|x\|_2^2, \quad (1.5)$$

then the matrix  $\Phi$  obeys the  $s$ -order block RIP over  $\mathcal{I}$ . Define the block RIP constant (RIC)  $\delta_{s|\mathcal{I}}$  as the smallest positive constant  $\delta$  such that (1.5) holds for all  $x \in \mathbb{R}^N$  that are block  $s$ -sparse.

For the reminder of this paper, for simplicity,  $\delta_s$  represents the block RIP constant  $\delta_{s|\mathcal{I}}$ . Eldar and Mishali [5] showed that the mixed  $l_2/l_1$ -minimization method can exactly reconstruct any block  $s$ -sparse signal when the measurement matrix  $\Phi$  fulfills the block RIP with  $\delta_{2s} < 0.414$ . Later, Lin and Li [7] improved the sufficient condition on  $\delta_{2s}$  to 0.4931, and builded condition  $\delta_s < 0.307$  for accurate recovery. In 2019, the conclusions of literature [8] and [9] together present a complete characterization to the block RIP condition on  $\delta_{t_s}$  that the mixed  $l_2/l_1$  minimization method guarantees the block-sparse signal recovery in the field of block-sparse compressed sensing.

Recently, a lot of researchers [10] [11] [12] have revealed that  $l_p$  ( $0 < p < 1$ ) minimization not only constantly needs less constrained the RIP requirements, but also could ensure exact recovery for smaller  $p$  compared with

the  $l_1$  minimization. In the present paper, we are interested in investigating the stable reconstruction of block-sparse signals by the mixed  $l_2/l_p$  ( $0 < p < 1$ ) minimization as follows:

$$\hat{x} = \arg \min_{\tilde{x} \in \mathbb{R}^N} \|\tilde{x}\|_{2,p}^p \text{ s.t. } y - \Phi\tilde{x} \in \mathcal{B}, \quad (1.6)$$

where  $\|x\|_{2,p}^p = \sum_{i=1}^M \|x[i]\|_2^p$ . Simulation experiments [13] [14] indicated that fewer linear measurements are needed for accurate reconstruction when  $0 < p < 1$  than when  $p = 1$ . More related work can be found in literature [15] [16] [17] [18] [19] [20] [21].

In this paper, we further investigate the high-order block RIP conditions for the exact and stable reconstruction for (approximate) block-sparse signals by mixed  $l_2/l_p$  minimization. The crux is extend sparse representation of an  $l_p$ -polytope (Lemma 2.2 [22]) to the block scenario. With this technique, we obtain a sufficient condition on RIC  $\delta_{ts}$  that guarantees the exact and stable reconstruction of approximate block-sparse signals via mixed  $l_2/l_p$  minimization, and establish error bounds between the solution to (1.6) and the signal  $x$  to be recovered. Obviously, when  $x$  is accurately block-sparse and  $\mathcal{B} = \{0\}$  (i.e.,  $y = \Phi x$ ), we will derive the accurate reconstruction condition. Particularly, we will determine how many random Gaussian measurements suffice for the condition to hold with high probability.

The remainder of the paper is constructed as follows. In Section 2, we will provide some notations and a few lemmas. In Section 3, we will present the main results, and the associating proofs are given in Section 5. In Section 4, a series of numerical experiments are presented to support our theoretical results. Lastly, the conclusion is drawn in Section 6.

## 2 Preliminaries

Throughout this article, we use the below notations unless special mentioning. For a subset  $T$  in  $\mathbb{R}^M$ ,  $T^c$  denotes the complement of  $T$  in  $\mathbb{R}^M$ . For any vector  $x \in \mathbb{R}^N$ ,  $x_T$  represents a vector which is equal to  $x$  on block indices  $T$  and displaces other blocks with zero. Denote  $T_0$  by block indices of the  $s$  largest block in  $l_2$  norm of the vector  $x$ , i.e.,  $\|x[i]\|_2 \geq \|x[j]\|_2$  for any  $i \in T_0$  and  $j \in T_0^c$ . We represent  $x_{\max(s)}$  as  $x$  with all but the largest  $s$  blocks in  $l_2$  norm set to zero. Henceforth, we invariably choose that  $h = \hat{x} - x_{\max(s)}$ , where  $\hat{x}$  is the minimizer of (1.6).

In order to prove our main results, it is necessary to present the below lemma which is a crucial technical tool. Factually, we extend sparse expression of an  $l_p$ -polytope proposed by [22] to the block context.

**Lemma 2.1.** *For a positive integer  $s$ , a positive number  $\alpha$  and given  $p \in (0, 1]$ , define the block  $l_p$ -polytope  $T(\alpha, s, p) \in \mathbb{R}^N$  by*

$$T(\alpha, s, p) = \{x \in \mathbb{R}^N : \|x\|_{2,p}^p \leq s\alpha^p, \|x\|_{2,\infty} \leq \alpha\}.$$

*Then any  $x \in T(\alpha, s, p)$  can be expressed as the convex combination of block  $s$ -sparse vectors, i.e.,*

$$x = \sum_i \lambda_i u_i.$$

Here  $\lambda_i > 0$  and  $\sum_i \lambda_i = 1$  and  $\|u_i\|_{2,0} \leq s$ . In addition,

$$\sum_i \lambda_i \|u_i\|_{2,2}^2 \leq \alpha^p \|x\|_{2,2-p}^{2-p}. \quad (2.7)$$

*Proof.* We can prove the assertion holds by induction. If  $x$  is block  $s$ -sparse, we can set  $u_1 = x$  and  $\lambda_1 = 1$ , then  $\|u_1\|_{2,2}^2 = \|x\|_{2,2}^2 \leq \alpha^p \|x\|_{2,2-p}^{2-p}$ . Suppose that assertion holds for all block  $(l-1)$ -sparse vectors  $x$  ( $l-1 \geq s$ ). Then for any block  $l$ -sparse vectors  $x$  such that  $\|x\|_{2,p}^p \leq s\alpha^p$  and  $\|x\|_{2,\infty} \leq \alpha$ , without loss of generality suppose that  $x$  is not block  $(l-1)$ -sparse (otherwise the statement is naturally true by assumption of  $l-1$ ). Besides,  $x$  can be represented as  $x = \sum_{i=1}^l c_i E_i$ , where  $c_1 \geq c_2 \geq \dots \geq c_l > 0$ ,  $c_1$  is equal to the largest  $\|x[i]\|_2$  for every  $i \in \{1, 2, \dots, M\}$ ,  $c_2$  is equal to the next largest  $\|x[i]\|_2$ , etc. Here  $E_i$  denotes a unit vector in  $\mathbb{R}^N$ , which is equal to  $x/c_i$  on the  $i$ th largest block of  $x$  and zero other places. Set  $c_0 = \alpha$ , and fix  $a = (a_1, a_2, \dots, a_l) \in \mathbb{R}_+^l$ , where  $a_i = c_i^{p-1}$ ,  $i = 0, 1, \dots, l$ . Then we have  $\sum_{i=1}^l a_i c_i \leq s\alpha^p$  and  $\alpha^p = a_0 c_0 \geq a_1 c_1 \geq \dots \geq a_l c_l$ .

Denote the set

$$\Gamma = \{1 \leq j \leq l-1 : \sum_{i=j}^l a_i c_i \leq (l-j)a_{j-1}c_{j-1}\}. \quad (2.8)$$

It is easy to see that  $1 \in \Gamma$ , hence  $\Gamma$  is not empty. Then we note that  $j = \max \Gamma$ , which implies

$$\begin{aligned} \sum_{i=j}^l a_i c_i &\leq (l-j)a_{j-1}c_{j-1}, \\ \sum_{i=j+1}^l a_i c_i &> (l-j-1)a_j c_j. \end{aligned} \quad (2.9)$$

It follows that

$$(l-j)a_j c_j < \sum_{i=j}^l a_i c_i \leq (l-j)a_{j-1}c_{j-1}.$$

Set

$$\begin{aligned} y_w &= \sum_{i=1}^{j-1} c_i E_i + \frac{\sum_{i=j}^l a_i c_i}{l-j} \sum_{i=j, i \neq w}^l a_i^{-1} E_i, \\ \xi_w &= 1 - \frac{l-j}{\sum_{i=j}^l a_i c_i} a_w c_w, \end{aligned}$$

where  $w = j, j+1, \dots, l$ . Then by simple calculations, we obtain  $\sum_{w=j}^l \xi_w = 1$  and  $x = \sum_{w=j}^l \xi_w y_w$ , where  $y_w$  is block  $(l-1)$ -sparse for all  $w = j, j+1, \dots, l$ . Finally, since  $y_w$  is block  $(l-1)$ -sparse, under the induction assumption, we have  $y_w = \sum_i \mu_{w,i} u_{w,i}$ , where  $u_{w,i}$  is block  $s$ -sparse, and  $\mu_{w,i} \in [0, 1]$ ,  $\sum_i \mu_{w,i} = 1$ . Hence,  $x = \sum_i \sum_{w=j}^l \mu_{w,i} u_{w,i}$ , which implies that statement is true for  $l$ .  $\square$

We will utilize the below lemma in the process of proving the main conclusions, which is a useful important inequality.

**Lemma 2.2.** (Lemma 5.3 [23]) Suppose that  $M \geq s$ ,  $a_1 \geq a_2 \geq \dots \geq a_M \geq 0$ ,  $\sum_{i=1}^s a_i \geq \sum_{i=s+1}^M a_i$ , then for all  $\alpha \geq 1$ ,

$$\sum_{j=s+1}^M a_j^\alpha \leq \sum_{i=1}^s a_i^\alpha.$$

More generally, assume that  $a_1 \geq a_2 \geq \dots \geq a_M \geq 0$ ,  $\lambda \geq 0$  and  $\sum_{i=1}^s a_i + \lambda \geq \sum_{i=s+1}^M a_i$ , then for all  $\alpha \geq 1$ ,

$$\sum_{j=s+1}^M a_j^\alpha \leq s \left( \sqrt[\alpha]{\frac{\sum_{i=1}^s a_i^\alpha}{s}} + \frac{\lambda}{s} \right)^\alpha.$$

In view of the definition of  $h$ ,  $\hat{x}$  and  $x_{\max(s)}$ , we get the below lemma.

**Lemma 2.3.** Recall that  $h = \hat{x} - x_{\max(s)}$ , where  $\hat{x}$  is the solution to (1.6). It holds that

$$\|h_{-\max(s)}\|_{2,p}^p \leq \|h_{\max(s)}\|_{2,p}^p.$$

*Proof.* Assume that  $T_0$  is the block index set over  $s$  blocks with largest  $l_2$  norm of the vector  $x$ . Therefore,  $x_{T_0} = x_{\max(s)}$ . By applying the minimality of the solution  $\hat{x}$  and the reverse triangular inequality of  $\|\cdot\|_{2,p}$ , we get

$$\begin{aligned} \|x_{T_0}\|_{2,p}^p &\geq \|\hat{x}\|_{2,p}^p = \|x_{T_0} + h_{T_0}\|_{2,p}^p + \|h_{T_0^c}\|_{2,p}^p \\ &\geq \|x_{T_0}\|_{2,p}^p - \|h_{T_0}\|_{2,p}^p + \|h_{T_0^c}\|_{2,p}^p, \end{aligned}$$

which implies

$$\|h_{T_0^c}\|_{2,p}^p \leq \|h_{T_0}\|_{2,p}^p.$$

Note that  $\|h_{-\max(s)}\|_{2,p}^p \leq \|h_{T_0^c}\|_{2,p}^p$  and  $\|h_{T_0}\|_{2,p}^p \leq \|h_{\max(s)}\|_{2,p}^p$ . Combining with the above inequalities, the desired result can be derived.  $\square$

**Lemma 2.4.** (Lemma 5.1 [25]) Let  $\Phi \in \mathbb{R}^{n \times N}$  be a random matrix whose entries obey one of the distributions given by (3.18) and that fulfills the concentration inequality

$$\mathbb{P}(\|\Phi x\|_2^2 - \|x\|_2^2 \geq \epsilon \|x\|_2^2) \leq 2e^{-nc_0(\epsilon)}, \quad \epsilon \in (0, 1), \quad (2.10)$$

where the probability is taken over all  $n \times N$  matrices  $\Phi$  and  $c_0(\epsilon)$  is a constant relying merely on  $\epsilon$  and such that for all  $\epsilon \in (0, 1)$ ,  $c_0(\epsilon) > 0$ . Suppose that  $1 \leq s \leq n$ . Then, for any  $\delta \in (0, 1)$ , we have

$$(1 - \delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta)\|x\|_2^2 \quad (2.11)$$

for all  $s$ -sparse vectors  $x \in \mathbb{R}^N$  with probability

$$\geq 1 - 2 \left( \frac{12}{\delta} \right)^s e^{-c_0(\frac{\delta}{2})n}. \quad (2.12)$$

### 3 Main Results

Based on the knowledge prepared above, we present the main results in this part—a high-order block RIP condition for the robust reconstruction of arbitrary signals with block structure via mixed  $l_2/l_p$  minimization. In the case that the signal to be recovered is block-sparse, the condition can respectively guarantee the accurate construction and stable recovery in the noise-free case and in the noise situation. When the original signal  $x$  is not block-sparse and the linear measurement is corrupted by noise, the below result presents a sufficient condition for recovery of structured signals.

**Theorem 3.1.** *Let  $y = \Phi x + e$  be noisy measurements of a signal  $x \in \mathbb{R}^N$  with  $y, e \in \mathbb{R}^n$ ,  $\Phi \in \mathbb{R}^{n \times N}$  ( $n < N$ ) and  $\|e\|_2 \leq \rho$ . Assume that  $\mathcal{B} = \mathcal{B}^{l_2}(\varepsilon)$  with  $\rho + \sigma(\Phi)\|x_{-\max(s)}\|_2 \leq \varepsilon$  in (1.6). If  $\Phi$  satisfies the block RIP with*

$$\delta_{ts} < \frac{\mu}{\frac{2-p}{t-1} - \mu} := \phi(t, p) \quad (3.13)$$

for some  $1 < t \leq 2$ , where  $\mu \in [(\sqrt{1+2p-p^2}-1)/p, (1-(t-\sqrt{t^2-t})p)/(t-1)]$  is the sole positive solution of the equation

$$g(\mu, p) = \frac{p}{2}\mu^{\frac{2}{p}} + \mu - \frac{2-p}{2(t-1)}. \quad (3.14)$$

Then the solution  $\hat{x}^{l_2}$  to (1.6) fulfills

$$\|\hat{x}^{l_2} - x\|_2 \leq C_1(\varepsilon + \rho) + C_2\|x_{-\max(s)}\|_2, \quad (3.15)$$

where

$$C_1 = \sqrt{2} \left( \frac{\phi(t, p)}{\phi(t, p) - \delta_{ts}} \frac{(2-p)(1-(t-1)\mu)}{2-p-(t-1)\mu} \sqrt{1+\delta_{ts}} + \phi(t, p) \sqrt{\frac{1-p}{\phi(t, p) - \delta_{ts}}} \right),$$

$$C_2 = \sqrt{2}\sigma(\Phi) \left( \frac{\phi(t, p)}{\phi(t, p) - \delta_{ts}} \frac{(2-p)(1-(t-1)\mu)}{2-p-(t-1)\mu} \sqrt{1+\delta_{ts}} + \phi(t, p) \sqrt{\frac{1-p}{\phi(t, p) - \delta_{ts}}} \right) + 1.$$

**Remark 3.2.** *In the case of  $d_i = 1$ ,  $i = 1, 2, \dots, M$ , (3.15) is the same as Theorem 2 in [24].*

**Theorem 3.3.** *Let  $y = \Phi x + e$  be noisy measurements of a signal  $x \in \mathbb{R}^N$  with  $y, e \in \mathbb{R}^n$ ,  $\Phi \in \mathbb{R}^{n \times N}$  ( $n < N$ ) and  $\|\Phi^\top e\|_\infty \leq \rho$ . Assume that  $\mathcal{B} = \mathcal{B}^{DS}(\varepsilon)$  with  $\rho + \sigma^2(\Phi)\|x_{-\max(s)}\|_2 \leq \varepsilon$  in (1.6). If  $\Phi$  satisfies the block RIP with  $\delta_{ts} < \phi(t, p)$  for some  $1 < t \leq 2$ , then the solution  $\hat{x}^{DS}$  to (1.6) fulfills*

$$\|\hat{x}^{DS} - x\|_2 \leq D_1(\varepsilon + \rho) + D_2\|x_{-\max(s)}\|_2, \quad (3.16)$$

where

$$D_1 = \frac{\sqrt{2ds}\phi(t, p)}{\phi(t, p) - \delta_{ts}} \left( \frac{(2-p)(1-(t-1)\mu)}{2-p-(t-1)\mu} + (1 + \sqrt{N-ds})(1-p)\phi(t, p) \right),$$

$$D_2 = \frac{\sqrt{2ds}\phi(t, p)\sigma^2(\Phi)}{\phi(t, p) - \delta_{ts}} \left( \frac{(2-p)(1-(t-1)\mu)}{2-p-(t-1)\mu} + (1 + \sqrt{N-ds})(1-p)\phi(t, p) \right) + 1.$$

**Remark 3.4.** *In the case of  $d_i = 1$ ,  $i = 1, 2, \dots, M$ , we obtain the same results as Theorem 3 in [24].*

**Corollary 3.5.** *Under the same conditions as in Theorem 3.1, suppose that  $e = 0$  and  $x$  is block  $s$ -sparse. Then  $x$  can be accurately reconstructed via*

$$\hat{x} = \arg \min_{\tilde{x} \in \mathbb{R}^N} \|\tilde{x}\|_{2,p}^p \text{ s.t. } y = \Phi \tilde{x}. \quad (3.17)$$

In the following, we will decide how many random Gaussian measurements are demanded for (3.13) to be fulfilled with overwhelming probability. In this sequel, let  $(\Omega, \tau)$  be a probability measure space and  $z$  be a random variable which follows one of the following probability distributions:

$$z \sim \mathcal{N}(0, 1/n), \quad z \sim \begin{cases} 1/\sqrt{n}, & \text{w.p. } 1/2, \\ -1/\sqrt{n}, & \text{w.p. } 1/2, \end{cases} \quad \text{or } z \sim \begin{cases} \sqrt{3/n}, & \text{w.p. } 1/6, \\ 0, & \text{w.p. } 2/3, \\ -\sqrt{3/n}, & \text{w.p. } 1/6. \end{cases} \quad (3.18)$$

Given  $n$  and  $N$ , the random matrices  $\Phi$  can be produced by making choice of the elements  $\Phi_{i,j}$  as independent copies of  $z$ . This generates the random matrices  $\Phi$ .

**Theorem 3.6.** *Let  $\Phi$  be an  $n \times N$  matrix with  $n < N$  whose elements are i.i.d. random variables defined by (3.18). If*

$$n \geq \frac{ts \log \frac{N}{ds}}{\frac{\phi^2(t,p)}{16} - \frac{\phi^3(t,p)}{48}},$$

the next assertion holds with probability more than

$1 - 2 \exp \left\{ ts \left( d \log \frac{12}{\phi(t,p)} + \log \frac{e}{t} + \log \frac{M}{s} \right) - n \left( \frac{\phi^2(t,p)}{16} - \frac{\phi^3(t,p)}{48} \right) \right\}$ : for any block  $s$ -sparse signal  $x \in \mathbb{R}^N$  over  $\mathcal{I} = \{d_1 = d, d_2 = d, \dots, d_M = d\}$  with  $Md = N$ ,  $x$  is the unique solution to (3.17) when the matrix  $\Phi$  satisfies  $\delta_{ts} < \phi(t, p)$ .

## 4 Numerical experiments

In this section, we carry out a few numerical simulations to hold out the application of our theoretical results. We could transform the constrained optimization problem (1.6) into an alternative unconstrained form below:

$$\min_{\tilde{x} \in \mathbb{R}^N} \lambda \|\tilde{x}\|_{2,p}^p + \frac{1}{2} \|y - \Phi \tilde{x}\|_2^2. \quad (4.19)$$

Solving the problem (4.19), we adopt the standard Alternating Direction Method of Multipliers (ADMM)[27][28][26].

Utilizing a auxiliary variable  $v \in \mathbb{R}^N$ , we can rewrite the formulation (4.19) as

$$\min_{\tilde{x}, v \in \mathbb{R}^N} \lambda \|v\|_{2,p}^p + \frac{1}{2} \|y - \Phi \tilde{x}\|_2^2 \text{ s.t. } \tilde{x} - v = 0. \quad (4.20)$$

The augmented Lagrangian function of the above problem is

$$\mathcal{L}_\gamma(\tilde{x}, v, z) = \lambda \|v\|_{2,p}^p + \frac{1}{2} \|y - \Phi \tilde{x}\|_2^2 + \langle z, \tilde{x} - v \rangle + \frac{\gamma}{2} \|\tilde{x} - v\|_2^2, \quad (4.21)$$

where  $z \in \mathbb{R}^n$  is a Lagrangian multiplier, and  $\gamma > 0$  is a penalty parameter. Then, ADMM composes of the below three steps:

$$\tilde{x}^{k+1} = \arg \min \frac{1}{2} \|y - \Phi \tilde{x}\|_2^2 + \frac{\gamma}{2} \|z^k + \tilde{x} - v^k\|_2^2 \quad (4.22)$$

$$v^{k+1} = \arg \min \lambda \|v\|_{2,p}^p + \frac{\gamma}{2} \|z^k + \tilde{x}^{k+1} - v\|_2^2 \quad (4.23)$$

$$z^{k+1} = z^k + \tilde{x}^{k+1} - v^{k+1}. \quad (4.24)$$

The solution of problem (4.22) is explicitly provided by

$$\tilde{x}^{k+1} = (\Phi^\top \Phi + \gamma I_n)^{-1} (\Phi^\top y - \gamma(z^k - v^k)). \quad (4.25)$$

To use the existing conclusions on the proximity operator of  $l_p$ -norm ( $0 < p < 1$ ), by employing the inequality:  $\|x\|_2 < \|x\|_p$  for  $x \in \mathbb{R}^n$  and  $0 < p < 1$ , the optimization problem (4.23) can be converted into

$$v^{k+1} = \arg \min \sum_{i=1}^M \lambda \|v[i]\|_p^p + \frac{\gamma}{2} \|z^k[i] + \tilde{x}^{k+1}[i] - v[i]\|_2^2. \quad (4.26)$$

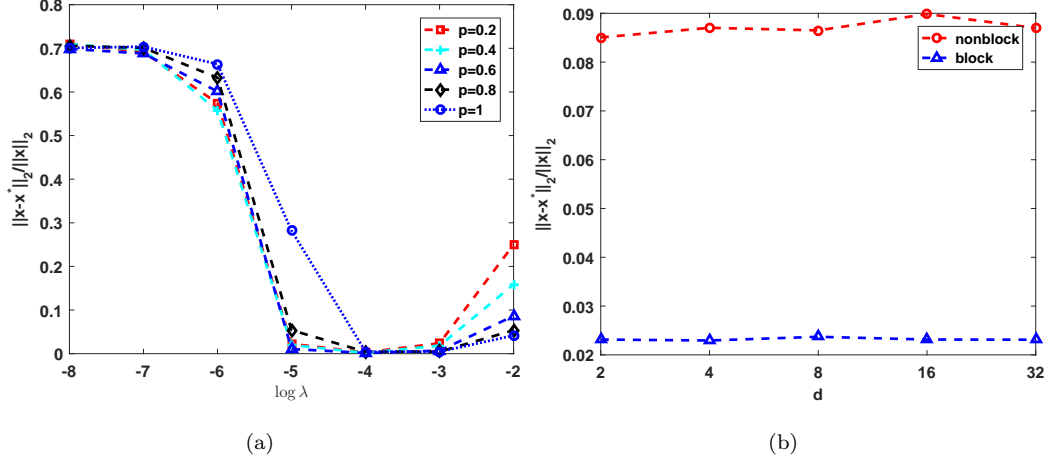
In our experiments, without loss of generality, we think over the block-sparse signal  $x$  with even block size, i.e.,  $d_i = d$ ,  $i = 1, \dots, M$ , and take the signal length  $N = 1024$ . For each experiment, first of all, we randomly produce block-sparse signal  $x$  with the amplitude of each nonzero entry generated according to the Gaussian distribution. We use an  $n \times N$  orthogonal Gaussian random matrix as the measurement matrix  $\Phi$ . We set the number of random measurement  $m = 128$  unless otherwise specified. With  $x$  and  $\Phi$ , we generate the linear measurement  $y$  by  $y = \Phi x + e$ , where  $e$  is the Gaussian noise vector. Each given experimental result is an average over 100 independent trails.

In Fig. 4.1a, we produce the signals by making choice of 64 blocks uniformly at random with  $n = 64$ ,  $N = 128$ , i.e., the block size  $d = 2$ . The relative error of recovery  $\|x - x^*\|_2 / \|x\|_2$  is plotted versus the regularization parameter  $\lambda$  for the different values of  $p$ , i.e.,  $p = 0.2, 0.4, 0.6, 0.8, 1$ . The  $\lambda$  ranges from  $10^{-8}$  to  $10^{-2}$ . From the figure, the parameter  $\lambda = 10^{-4}$  is a proper choice. Fig. 4.1b presents experimental results regarding the performance of the non-block algorithm and the block algorithm with  $p = 0.4$ . Two curves of relative error are provided via mixed  $l_2/l_p$  minimization and orthogonal greedy algorithm (OGA) [29]. Fig. 4.1b reveals the signal construction is quite significant in signal recovery.

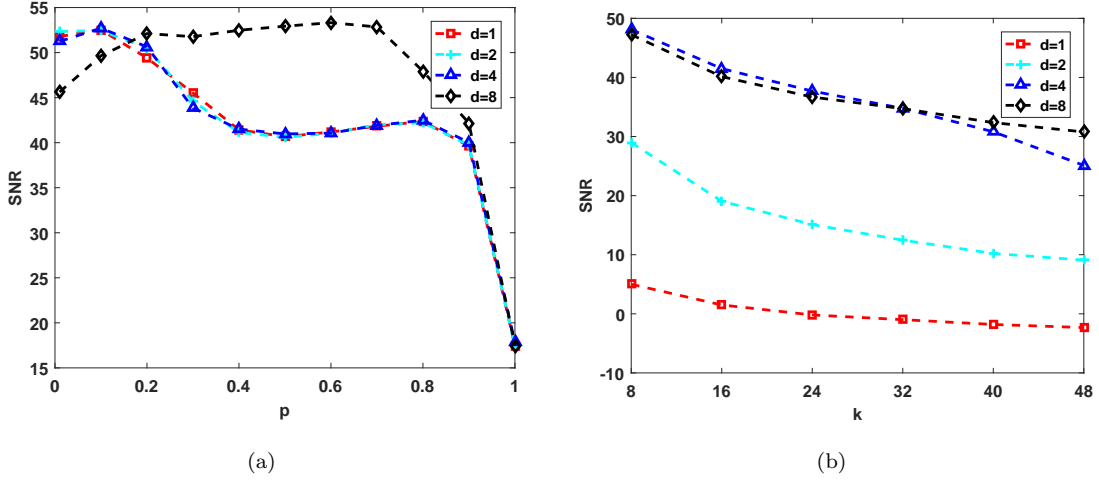
Signal-to-noise ratio (SNR,  $\text{SNR} = 20 \log_{10}(\|x\|_2 / \|x - x^*\|_2)$ ) versus the values of  $p$  and the nonzero entries  $k$ , the results are respectively given in Fig. 4.2a and b. In Fig. 4.2a, the values of  $p$  vary from 0.01 to 1, and in Fig. 4.2b, the number of nonzero entries  $k$  ranges from 8 to 48. Figs. 4.2a and b evidences that mixed  $l_2/l_p$  minimization performs better than that of standard  $l_p$  minimization. Figs. 4.2a and b provide the relationship between the relative error and the number of measurements  $n$  in different block sizes  $d = 1, 2, 4, 8$  and the values of  $p = 0.2, 0.4, 0.6, 0.8, 1$ .

Eventually, we compare the performance of Group-Lp for  $p = 0.4$  with other typical algorithms consisting of Block-OMP algorithm[30], Block-SL0 algorithm [31] for  $l_2/l_0$  solver and Block-ADM algorithm [32]. We exploit SNR to weigh the algorithm efficiency. In Fig. 4.4a, we select signals whose block size  $d = 2, 4, 8, 16, 32$  with





**Fig. 4.1:** (a) Recovery performance of mixed  $l_2/l_p$  minimization versus  $\lambda$  for block size  $d = 2$ , (b) Recovery performance of mixed  $l_2/l_p$  minimization with  $p = 0.4$  and OGA and the number of nonzero entries  $k = 64$



**Fig. 4.2:** SNR versus the values of  $p$  and the number of nonzero entries  $k$  in (a) and (b) respectively  
(a) For  $k = 8$ , (b) For  $p = 0.4$ .

the number of nonzero entries  $k = 32$  as the test signals, and in Fig. 4.4b, the block size  $d = 2$ . One can see that, overall, the performance of Group-Lp ( $p = 0.4$ ) is much better than that of other three algorithms.

## 5 The proofs of Main Results

**Proof of Theorem 3.1.** First of all, suppose that  $ts$  is an integer. Recall that  $h = \hat{x}^{l_2} - x_{\max(s)}$ , where  $\hat{x}^{l_2}$  is the solution to (1.6) with  $\mathcal{B} = \mathcal{B}^{l_2}(\varepsilon)$ . Then, by employing the condition of Theorem 3.1, we have

$$\|y - \Phi x_{\max(s)}\|_2 \leq \|y - \Phi x\|_2 + \|\Phi(x - x_{\max(s)})\|_2 \leq \rho + \sigma(\Phi)\|x - x_{\max(s)}\|_2 \leq \varepsilon,$$

that is,  $y - \Phi x_{\max(s)} \in \mathcal{B}^{l_2}(\varepsilon)$ .

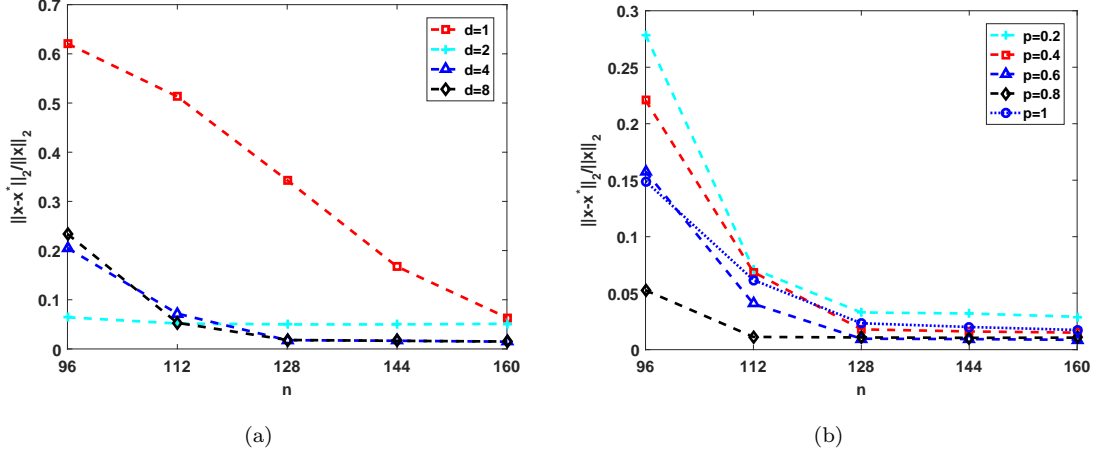


Fig. 4.3: Relative error versus the number of measurements  $n$  for (a)  $p = 0.4$ ,  $k = 16$  and (b)  $d = 2$ ,  $k = 16$

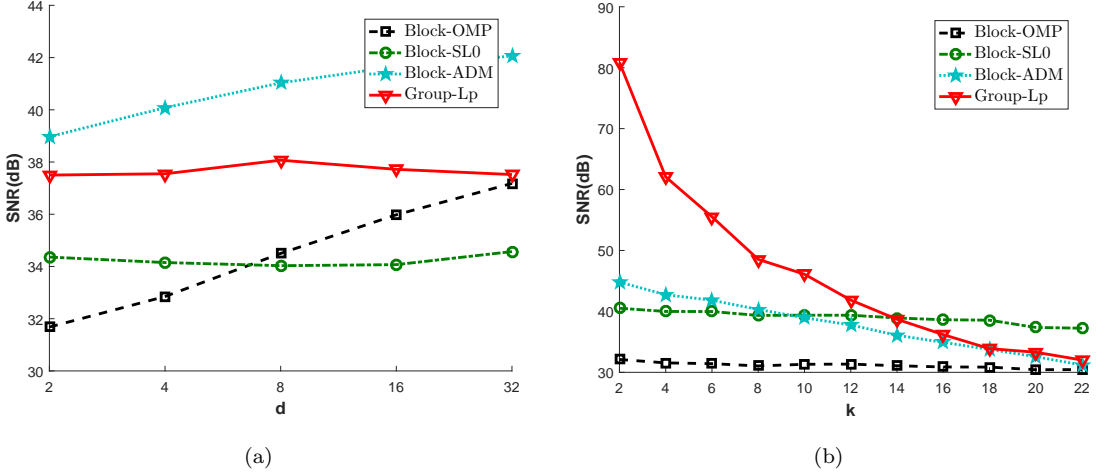


Fig. 4.4: Comparison recovery performance with respect to SNR (a) Number of nonzero coefficients  $k = 32$  and (b) block size  $d = 2$

Denote  $\alpha^p = \|h_{\max(s)}\|_{2,p}^p/s$ . Then, by utilizing Lemma 2.3, we have

$$\begin{aligned} \|h_{-\max(s)}\|_{2,\infty} &\leq \alpha \leq \frac{\alpha}{(t-1)^{1/p}}, \text{ and} \\ \|h_{-\max(s)}\|_{2,p}^p &\leq s(t-1) \left( \frac{\alpha}{(t-1)^{1/p}} \right)^p. \end{aligned} \quad (5.27)$$

Through applying Lemma 2.1 and combining with (5.27), we have  $h_{-\max(s)} = \sum_{i=1}^N \lambda_i u_i$ , where  $u_i$  is block  $(t-1)s$ -sparse,  $\sum_{i=1}^N \lambda_i = 1$  with  $\lambda_i \in [0, 1]$ , and

$$\sum_i \lambda_i \|u_i\|_{2,2}^2 \leq \frac{\alpha^p}{t-1} \|h_{-\max(s)}\|_{2,2-p}^{2-p}. \quad (5.28)$$

Hence,

$$\begin{aligned}
\sum_i \lambda_i \|u_i\|_{2,2}^2 &\stackrel{(a)}{\leq} \frac{\alpha^p}{t-1} (\|h_{-\max(s)}\|_{2,2}^2)^{\frac{2(1-p)}{2-p}} (\|h_{-\max(s)}\|_{2,p}^p)^{\frac{p}{2-p}} \\
&\leq \frac{\alpha^p}{t-1} (\|h_{-\max(s)}\|_{2,2}^2)^{\frac{2(1-p)}{2-p}} (s\alpha^p)^{\frac{p}{2-p}} \\
&\stackrel{(b)}{\leq} \frac{1}{t-1} (\|h_{-\max(s)}\|_{2,2}^2)^{\frac{2(1-p)}{2-p}} (\|h_{\max(s)}\|_{2,2}^2)^{\frac{p}{2-p}} \\
&\stackrel{(c)}{\leq} \frac{1}{t-1} (\|h_{-\max(s)}\|_2^2)^{\frac{2(1-p)}{2-p}} (\|h_{\max(s)}\|_2^2)^{\frac{p}{2-p}}, \tag{5.29}
\end{aligned}$$

where (a) follows from Hölder's inequality, (b) is due to the fact that  $\|x\|_p \leq \|x\|_q \leq n^{1/q-1/p} \|x\|_p$ ,  $x \in \mathbb{R}^n$  and given  $0 < q < p \leq \infty$ , and (c) is from the fact that  $\|x\|_{2,2}^2 = \sum_{i=1}^M \|x[i]\|_2^2 = \sum_{i=1}^N x_i^2 = \|x\|_2^2$ ,  $x \in \mathbb{R}^N$ .

From Cauchy-Schwartz inequality and the definition of block RIP, we get

$$\langle \Phi h_{\max(s)}, \Phi h \rangle \leq \|\Phi h_{\max(s)}\|_2 \|\Phi h\|_2 \leq \sqrt{1 + \delta_{ts}} \|h_{\max(s)}\|_2 \|\Phi h\|_2. \tag{5.30}$$

Since

$$\begin{aligned}
\|\Phi h\|_2 &\leq \|\Phi(\hat{x}^{l_2} - x_{\max(s)})\|_2 \leq \|\Phi \hat{x}^{l_2} - y\|_2 + \|\Phi x_{\max(s)} - y\|_2 \\
&\leq \varepsilon + \rho + \sigma(\Phi) \|x_{-\max(s)}\|_2, \tag{5.31}
\end{aligned}$$

therefore

$$\langle \Phi h_{\max(s)}, \Phi h \rangle \leq \sqrt{1 + \delta_{ts}} (\varepsilon + \rho + \sigma(\Phi) \|x_{-\max(s)}\|_2) \|h_{\max(s)}\|_2. \tag{5.32}$$

Take  $\beta_i = h_{\max(s)} + (t-1)\mu u_i$ . Then,

$$\sum_{j=1}^N \lambda_j \beta_j - \frac{p}{2} \beta_i - (t-1)\mu h = \left[1 - (t-1)\mu - \frac{p}{2}\right] h_{\max(s)} - \frac{p}{2}(t-1)\mu u_i. \tag{5.33}$$

Furthermore, both  $\sum_{j=1}^N \lambda_j \beta_j - \frac{p}{2} \beta_i - (t-1)\mu h$  and  $\beta_i - \beta_j = (t-1)\mu(u_i - u_j)$  are block  $ts$ -sparse, because  $h_{\max(s)}$  is block  $s$ -sparse, and  $u_i$  is block  $(t-1)s$ -sparse.

Observe that the identity below [22]

$$\begin{aligned}
&\sum_i \lambda_i \left\| \Phi \left( \sum_j \lambda_j \beta_j - \frac{p}{2} \beta_i \right) \right\|_2^2 + \frac{1-p}{2} \sum_{i,j} \lambda_i \lambda_j \|\Phi(\beta_i - \beta_j)\|_2^2 \\
&= \left(1 - \frac{p}{2}\right)^2 \sum_i \lambda_i \|\Phi \beta_i\|_2^2, \tag{5.34}
\end{aligned}$$

where  $\sum_i \lambda_i = 1$ .

First of all, we determine the left hand side (LHS) of (5.34). Putting (5.33) into LHS of (5.34) and combining

with (5.32) and the concept of block RIP, we get

$$\begin{aligned}
LHS &= \sum_i \lambda_i \|\Phi[(1 - (t-1)\mu - \frac{p}{2})h_{\max(s)} - \frac{p}{2}(t-1)\mu u_i + (t-1)\mu h]\|_2^2 \\
&\quad + \frac{1-p}{2} \sum_{i,j} \lambda_i \lambda_j (t-1)^2 \mu^2 \|\Phi(u_i - u_j)\|_2^2 \\
&= \sum_i \lambda_i \|\Phi[(1 - (t-1)\mu - \frac{p}{2})h_{\max(s)} - \frac{p}{2}(t-1)\mu u_i]\|_2^2 \\
&\quad + 2(1 - \frac{p}{2})(1 - (t-1)\mu)(t-1)\mu \langle \Phi h_{\max(s)}, \Phi h \rangle + (1-p)(t-1)^2 \mu^2 \|\Phi h\|_2^2 \\
&\quad + \frac{1-p}{2} (t-1)^2 \mu^2 \sum_{i,j} \lambda_i \lambda_j (t-1)^2 \mu^2 \|\Phi(u_i - u_j)\|_2^2 \\
&\leq (1 + \delta_{ts}) [\sum_i \lambda_i \|(1 - (t-1)\mu - \frac{p}{2})h_{\max(s)} - \frac{p}{2}(t-1)\mu u_i\|_2^2 \\
&\quad + \frac{1-p}{2} (t-1)^2 \mu^2 \sum_{i,j} \lambda_i \lambda_j (t-1)^2 \mu^2 \|u_i - u_j\|_2^2] \\
&\quad + 2(1 - \frac{p}{2})(1 - (t-1)\mu)(t-1)\mu \sqrt{1 + \delta_{ts}} (\varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2) \|h_{\max(s)}\|_2 \\
&\quad + (1-p)(t-1)^2 \mu^2 (\varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2)^2 \\
&= (1 + \delta_{ts}) [1 - (t-1)\mu - \frac{p}{2}]^2 \|h_{\max(s)}\|_2^2 + (1 - \frac{p}{2})^2 (t-1)^2 \mu^2 (1 + \delta_{ts}) \sum_i \lambda_i \|u_i\|_2^2 \\
&\quad - (1-p)(t-1)^2 \mu^2 (1 + \delta_{ts}) \|h_{-\max(s)}\|_2^2 \\
&\quad + 2(1 - \frac{p}{2})(1 - (t-1)\mu)(t-1)\mu \sqrt{1 + \delta_{ts}} (\varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2) \|h_{\max(s)}\|_2 \\
&\quad + (1-p)(t-1)^2 \mu^2 (\varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2)^2. \tag{5.35}
\end{aligned}$$

Then again, by exploiting the representation of  $\beta_i$  and the notion of block RIP, we get

$$\begin{aligned}
RHS &= (1 - \frac{p}{2})^2 \sum_i \lambda_i \|\Phi \beta_i\|_2^2 \\
&= (1 - \frac{p}{2})^2 \sum_i \lambda_i \|\Phi(h_{\max(s)} + (t-1)\mu u_i)\|_2^2 \\
&\geq (1 - \frac{p}{2})^2 (1 - \delta_{ts}) \sum_i \lambda_i \|h_{\max(s)} + (t-1)\mu u_i\|_2^2 \\
&= (1 - \frac{p}{2})^2 (1 - \delta_{ts}) \|h_{\max(s)}\|_2^2 + (1 - \frac{p}{2})^2 (t-1)^2 \mu^2 (1 - \delta_{ts}) \sum_i \lambda_i \|u_i\|_2^2. \tag{5.36}
\end{aligned}$$

A combination of (5.35) and (5.36), we obtain

$$\begin{aligned}
&\{(1 + \delta_{ts})[1 - (t-1)\mu - \frac{p}{2}]^2 - (1 - \frac{p}{2})^2 (1 - \delta_{ts})\} \|h_{\max(s)}\|_2^2 \\
&+ 2(1 - \frac{p}{2})^2 (t-1)^2 \mu^2 \delta_{ts} \sum_i \lambda_i \|u_i\|_2^2 - (1-p)(t-1)^2 \mu^2 (1 + \delta_{ts}) \|h_{-\max(s)}\|_2^2 \\
&+ 2(1 - \frac{p}{2})(1 - (t-1)\mu)(t-1)\mu \sqrt{1 + \delta_{ts}} (\varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2) \|h_{\max(s)}\|_2 \\
&+ (1-p)(t-1)^2 \mu^2 (\varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2)^2 \geq 0. \tag{5.37}
\end{aligned}$$

Substituting (5.29) into (5.37), we get

$$\begin{aligned}
& \{(1 + \delta_{ts})[1 - (t-1)\mu - \frac{p}{2}]^2 - (1 - \frac{p}{2})^2(1 - \delta_{ts})\} \|h_{\max(s)}\|_2^2 \\
& + 2(1 - \frac{p}{2})^2(t-1)\mu^2\delta_{ts} (\|h_{-\max(s)}\|_2^2)^{\frac{2(1-p)}{2-p}} (\|h_{\max(s)}\|_2^2)^{\frac{p}{2-p}} \\
& - (1-p)(t-1)^2\mu^2(1 + \delta_{ts})\|h_{-\max(s)}\|_2^2 \\
& + 2(1 - \frac{p}{2})(1 - (t-1)\mu)(t-1)\mu\sqrt{1 + \delta_{ts}}(\varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2)\|h_{\max(s)}\|_2 \\
& + (1-p)(t-1)^2\mu^2(\varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2)^2 \geq 0.
\end{aligned} \tag{5.38}$$

Regarding the LHS as the function of  $\|h_{-\max(s)}\|_2^2$ , we consider its extremum problem. Then,

$$\begin{aligned}
& \{(1 + \delta_{ts})[1 - (t-1)\mu - \frac{p}{2}]^2 - (1 - \frac{p}{2})^2(1 - \delta_{ts})\} \|h_{\max(s)}\|_2^2 \\
& + \frac{p}{2}(t-1)^2\mu^2(1 + \delta_{ts})\left(\frac{(2-p)\delta_{ts}}{(t-1)(1 + \delta_{ts})}\right)^{\frac{2-p}{p}} \|h_{\max(s)}\|_2^2 \\
& + 2(1 - \frac{p}{2})(1 - (t-1)\mu)(t-1)\mu\sqrt{1 + \delta_{ts}}(\varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2)\|h_{\max(s)}\|_2 \\
& + (1-p)(t-1)^2\mu^2(\varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2)^2 \geq 0.
\end{aligned} \tag{5.39}$$

Noting that (3.13) and (3.14), we get

$$\begin{aligned}
& [2 - p - (t-1)\mu]^2\left(\delta_{ts} - \frac{\mu}{\frac{2-p}{t-1} - \mu}\right)\|h_{\max(s)}\|_2^2 \\
& + 2(1 - \frac{p}{2})(1 - (t-1)\mu)(t-1)\mu\sqrt{1 + \delta_{ts}}(\varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2)\|h_{\max(s)}\|_2 \\
& + (1-p)(t-1)^2\mu^2(\varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2)^2 \geq 0.
\end{aligned} \tag{5.40}$$

The condition  $\delta_{ts} < \phi(t, p)$  guarantees that the inequality (5.40) is a second-order inequality for  $\|h_{\max(s)}\|_2$ , and the quadratic coefficient is less than zero. Consequently, we have

$$\begin{aligned}
\|h_{\max(s)}\|_2 & \leq \frac{1}{2[2 - p - (t-1)\mu]^2(\phi(t, p) - \delta_{ts})} \{2(1 - \frac{p}{2})(1 - (t-1)\mu)(t-1)\mu\sqrt{1 + \delta_{ts}}\theta \\
& + \{[2(1 - \frac{p}{2})(1 - (t-1)\mu)(t-1)\mu\sqrt{1 + \delta_{ts}}\theta]^2 \\
& + 4[2 - p - (t-1)\mu]^2(\phi(t, p) - \delta_{ts})(1-p)(t-1)^2\mu^2\theta^2\}^{\frac{1}{2}} \\
& \stackrel{(a)}{\leq} \left\{ \frac{\phi(t, p)}{\phi(t, p) - \delta_{ts}} \frac{(2-p)(1 - (t-1)\mu)}{2 - p - (t-1)\mu} \sqrt{1 + \delta_{ts}} + \phi(t, p) \sqrt{\frac{1-p}{\phi(t, p) - \delta_{ts}}} \right\} \theta,
\end{aligned} \tag{5.41}$$

where (a) follows from the fact that  $(u + v)^{1/2} \leq u^{1/2} + v^{1/2}$  for  $u, v \geq 0$ , and  $\theta = \varepsilon + \rho + \sigma(\Phi)\|x_{-\max(s)}\|_2$ . Combining with Lemmas 2.2 and 2.3, it follows that  $\|h_{-\max(s)}\|_2 \leq \|h_{\max(s)}\|_2$ .

Accordingly, it is not difficult to check that

$$\begin{aligned}
\|\hat{x}^{l_2} - x\|_2 & \leq \|\hat{x}^{l_2} - x_{\max(s)}\|_2 + \|x - x_{\max(s)}\|_2 \\
& \leq \sqrt{2}\|h_{\max(s)}\|_2 + \|x_{-\max(s)}\|_2 \\
& \leq \sqrt{2}\left\{ \frac{\phi(t, p)}{\phi(t, p) - \delta_{ts}} \frac{(2-p)(1 - (t-1)\mu)}{2 - p - (t-1)\mu} \sqrt{1 + \delta_{ts}} + \phi(t, p) \sqrt{\frac{1-p}{\phi(t, p) - \delta_{ts}}} \right\} \theta + \|x_{-\max(s)}\|_2.
\end{aligned} \tag{5.42}$$

If  $ts$  is not an integer, we represent  $t' = \lceil ts \rceil / s$ , then  $t's$  is an integer and  $t < t'$ . Due to  $\frac{\partial \phi(t,p)}{\partial t} > 0$ , the function  $\phi(t,p)$  is growing with  $t \in (1, 2]$ . Thus, we have  $\delta_{t's} = \delta_{ts} < \phi(t,p) < \phi(t',p)$ . Analogous to the above proof, we can prove the result by working on  $\delta_{t's}$ . □

**Proof of Theorem 3.3.** Similar to the proof of the former noisy situation, for the case of noise type  $\mathcal{B} = \mathcal{B}^{DS}(\rho)$ , define  $h = \hat{x}^{DS} - x_{\max(s)}$ . We can derive

$$\|\Phi^\top(y - \Phi x_{\max(s)})\|_\infty \leq \|\Phi^\top(y - \Phi x)\|_\infty + \|\Phi^\top(\Phi x - \Phi x_{\max(s)})\|_\infty \leq \rho + \sigma^2(\Phi)\|x_{-\max(s)}\|_2 \leq \varepsilon,$$

which reveals that  $y - \Phi x_{\max(s)} \in \mathcal{B}^{DS}(\varepsilon)$ . From the proof of Theorem 3.1, we have  $\|h_{-\max(s)}\|_2 \leq \|h_{\max(s)}\|_2$ . By using the inequalities  $\|x\|_p \leq \|x\|_q \leq n^{1/q-1/p}\|x\|_p$ ,  $x \in \mathbb{R}^n$  and given  $0 < q < p \leq \infty$ , we obtain  $\|h_{-\max(s)}\|_1 \leq \sqrt{N-ds}\|h_{\max(s)}\|_1$ . Hence,

$$\|h\|_1 \leq (1 + \sqrt{N-ds})\|h_{\max(s)}\|_1 \leq (1 + \sqrt{N-ds})\sqrt{ds}\|h_{\max(s)}\|_2. \quad (5.43)$$

Then,

$$\begin{aligned} \|\Phi h\|_2^2 &= \langle h, \Phi^\top \Phi h \rangle \leq \|h\|_1 \|\Phi^\top \Phi h\|_\infty \\ &\leq \|h\|_1 (\|\Phi^\top(\Phi \hat{x}^{DS} - y)\|_\infty + \|\Phi^\top(\Phi x_{\max(s)} - y)\|_\infty) \\ &\leq (1 + \sqrt{N-ds})\sqrt{ds}\|h_{\max(s)}\|_2 (\varepsilon + \rho + \sigma^2(\Phi)\|x_{-\max(s)}\|_2) \end{aligned} \quad (5.44)$$

and

$$\begin{aligned} \langle \Phi h_{\max(s)}, \Phi h \rangle &\leq \langle h_{\max(s)}, \Phi^\top \Phi h \rangle \leq \|h\|_1 \|\Phi^\top \Phi h\|_\infty \\ &\leq \sqrt{ds}\|h_{\max(s)}\|_2 (\varepsilon + \rho + \sigma^2(\Phi)\|x_{-\max(s)}\|_2). \end{aligned} \quad (5.45)$$

Then, it follows that the below second-order inequality for  $\|h_{\max(s)}\|_2$

$$\begin{aligned} &[2 - p - (t-1)\mu]^2 (\phi(t,p) - \delta_{ts}) \|h_{\max(s)}\|_2^2 \\ &- 2(1 - \frac{p}{2})(1 - (t-1)\mu)(t-1)\mu\sqrt{ds}\|h_{\max(s)}\|_2 (\varepsilon + \rho + \sigma^2(\Phi)\|x_{-\max(s)}\|_2) \\ &- (1 + \sqrt{N-ds})\sqrt{ds}(1-p)(t-1)^2\mu^2\|h_{\max(s)}\|_2 (\varepsilon + \rho + \sigma^2(\Phi)\|x_{-\max(s)}\|_2) \leq 0. \end{aligned} \quad (5.46)$$

Therefore,

$$\begin{aligned} \|h_{\max(s)}\|_2 &\leq \frac{\phi(t,p)}{\phi(t,p) - \delta_{ts}} \left\{ \frac{(2-p)(1-(t-1)\mu)\sqrt{ds}}{2-p-(t-1)\mu} + (1 + \sqrt{N-ds})\phi(t,p)(1-p)\sqrt{ds} \right\} \\ &\quad \times (\varepsilon + \rho + \sigma^2(\Phi)\|x_{-\max(s)}\|_2). \end{aligned}$$

The remaining proof is similar with the case of noise type  $\mathcal{B}^{l2}$ , we omit it here. □

**Proof of Theorem 3.6.** Similar to the proof of Theorem 5.2 [25], from Lemma 2.4 and the union bound, for fixed  $\delta \in (0, 1)$ , the measurement matrix  $\Phi$  satisfies block RIP (1.5) over  $\mathcal{I} = \{d_1 = d, d_2 = d, \dots, d_M = d\}$  with

probability  $\geq 1 - 2\left(\frac{12}{\delta}\right)^{sd} \binom{M}{s} e^{-c_0(\delta/2)n}$ , where  $c_0(\delta/2) = \delta^2/16 - \delta^3/48$  and  $N = Md$ . Therefore, we have

$$\mathbb{P}(\delta_s < \delta) \geq 1 - 2\left(\frac{12}{\delta}\right)^{sd} \binom{M}{s} e^{-c_0(\frac{\delta}{2})n} \quad (5.47)$$

Corollary 3.5 shows that in the case of free-noise, the guarantee to exactly recover block  $s$ -sparse signals is  $\delta_{ts} < \phi(t, p)$  ( $1 < t \leq 2$ ). Hence, for  $t \in (1, 2]$ ,  $\delta_{ts} < \phi(t, p)$  with probability

$$\begin{aligned} \mathbb{P}(\delta_{ts} < \phi(t, p)) &\geq 1 - 2\left(\frac{12}{\phi(t, p)}\right)^{tsd} \binom{M}{ts} e^{-n\left(\frac{\phi^2(t, p)}{16} - \frac{\phi^3(t, p)}{48}\right)} \\ &\stackrel{(a)}{\leq} 1 - 2e^{ts\left(d \log \frac{12}{\phi(t, p)} + \log \frac{\varepsilon}{t} + \log \frac{M}{s}\right) - n\left(\frac{\phi^2(t, p)}{16} - \frac{\phi^3(t, p)}{48}\right)}, \end{aligned} \quad (5.48)$$

where (a) follows from the inequality  $\binom{u}{v} \leq (eu/v)^v$  for integers  $u > v > 0$ . When  $M/s \rightarrow \infty$ , to ensure that  $\delta_{ts} < \phi(t, p)$  with overwhelming probability, the number of measurements must satisfy  $n \geq ts \log \frac{N}{ds} / \left(\frac{\phi^2(t, p)}{16} - \frac{\phi^3(t, p)}{48}\right)$ .  $\square$

## 6 Conclusion

In recent years, the research of non-convex block-sparse compressed sensing has become a hot topic. This paper mainly discusses non-convex block-sparse compressed sensing by employing block RIP. We establish a sufficient condition that guarantee the stable and robust signal reconstruction via mix  $l_2/l_p$  minimization method. Meanwhile, we present the upper bound estimation of recovery error. In addition, we give the number of samples needed to satisfy the sufficient conditions with high probability. Besides, we conduct a series of numerical experiments to show the verifiability of our results, and generally speaking, compared with other representative algorithms, the performance of Group-Lp algorithm is much better.

## Acknowledgements

This work was supported by Natural Science Foundation of China (Grant Nos. 61673015, 61273020), Fundamental Research Funds for the Central Universities (Grant Nos. XDJK2015A007, XDJK2018C076, SWU1809002), Graduate Student Scientific Research Innovation Projects in Chongqing (No. CYB19083), Youth Science and technology talent development project (Grant No. Qian jiao he KY zi [2018]313).

## References

- [1] A. Majumdar , R. Ward ,Compressed sensing of color images, Signal Process. 90 (2010) 3122-3127.
- [2] F. Parvaresh , H. Vikalo , S. Misra , B. Hassibi , Recovering sparse signals using sparse measurement matrices in compressed DNA microarrays, IEEE J. Sel. Top- ics Signal Process. 2 (3) (2008) 275-285 .
- [3] S. Cotter , B. Rao , Sparse channel estimation via matching pursuit with applica- tion to equalization, IEEE Trans. Commun. 50 (3) (2002) 374-377 .

- [4] J. Huang , X. Huang , D. Metaxas , Learning with Dynamic Group Sparsity, in: IEEE 12th International Conference on Computer Vision, 2009, pp. 64-71 .
- [5] Y. Eldar , M. Mishali , Robust recovery of signals from a structured union of sub- spaces, IEEE Trans. Inf. Theory 55 (11) (2009) 5302-5316 .
- [6] Y. Eldar , P. Kuppinger , H. Bolcskei , Block-sparse signals: uncertainty relations and efficient recovery, IEEE Trans. Signal Process. 58 (6) (2010) 3042-3054 .
- [7] J. Lin , S. Li , Block sparse recovery via mixed  $l_2/l_1$  minimization, Acta Mathe- matica Sinica 29 (7) (2013) 1401-1412 .
- [8] Li Y, Chen W. The high order block RIP condition for signal recovery[J]. Journal of Computational Mathematics, 2019, 37(1): 61-75.
- [9] Huang J, Wang J, Wang W, et al. Sharp sufficient condition of block signal recovery via  $l_2/l_1$ -minimisation[J]. IET Signal Processing, 2019, 13(5): 495-505.
- [10] R. Chartrand , V. Staneva , Restricted isometry properties and nonconvex compressive sensing, Inverse Probl. 24 (2008) 1-14 .
- [11] Y. Shen , S. Li , Restricted  $p$ -isometry property and its application for nonconvex compressive sensing, Adv. Comput. Math. 37 (2012) 441-452 .
- [12] M. Lai , Y. Xu , W. Yin , Improved iteratively reweighted least squares for unconstrained smoothed  $l_q$  minimization, SIAM J. Numer. Anal. 51 (2) (2013) 927-957 .
- [13] A. Majumdar , R. Ward ,Compressed sensing of color images, Signal Process. 90 (2010) 3122-3127 .
- [14] Y. Wang , J. Wang , Z. Xu , On recovery of block-sparse signals via mixed  $l_2/l_q$  ( $0 < q \leq 1$ ) norm minimization, EURASIP J. Adv. Signal Process. 76 (2013) 1-17 .
- [15] Y. Wang , J. Wang , Z. Xu , Restricted  $p$ -isometry properties of nonconvex block-sparse compressed sensing, Signal Process. 104 (2014) 188-196 .
- [16] Wen J, Zhou Z, Liu Z, et al. Sharp sufficient conditions for stable recovery of block sparse signals by block orthogonal matching pursuit[J]. Applied and Computational Harmonic Analysis, 2019, 47(3): 948-974.
- [17] Gao Y, Peng J, Yue S. Stability and robustness of the  $l_2/l_q$ -minimization for block sparse recovery[J]. Signal Processing, 2017, 137: 287-297.
- [18] Wang J, Huang J, Zhang F, et al. Group sparse recovery in impulsive noise via alternating direction method of multipliers[J]. Applied and Computational Harmonic Analysis, 2019.
- [19] Ge H, Chen W. Recovery of signals by a weighted  $l_2/l_1$  minimization under arbitrary prior support information[J]. Signal Processing, 2018, 148: 288-302.



- [20] Haifeng Li and Jinming Wen. A New Analysis for Support Recovery With Block Orthogonal Matching Pursuit, *IEEE Signal Processing Letters*, 26(2019), 247-251.
- [21] Wang W, Wang J, Zhang Z. Block-sparse signal recovery via  $l_2/l_{1-2}$  minimisation method[J]. *IET Signal Processing*, 2017, 12(4): 422-430.
- [22] Zhang R, Li S. Optimal RIP bounds for sparse signals recovery via  $l_p$  minimization[J]. *Applied and Computational Harmonic Analysis*, 47(2019)566-584.
- [23] T.T. Cai, A. Zhang, Sharp RIP bound for sparse signal and low-rank matrix recovery, *Appl. Comput. Harmon. Anal.*, 35 (2013), 74-93.
- [24] Chen B, Wan A. General RIP bounds of  $\delta_{tk}$  for sparse signals recovery by  $l_p$  minimization[J]. *Neurocomputing*, 2019, 363: 306-312.
- [25] Baraniuk R, Davenport M, DeVore R, et al. A simple proof of the restricted isometry property for random matrices[J]. *Constructive Approximation*, 2008, 28(3): 253-263.
- [26] Lu C, Feng J, Lin Z, et al. Exact low tubal rank tensor recovery from Gaussian measurements[C]. *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. AAAI Press, 2018: 2504-2510.
- [27] Wen F, Liu P, Liu Y, et al. Robust Sparse Recovery in Impulsive Noise via  $\ell_p$ - $\ell_1$  Optimization[J]. *IEEE Transactions on Signal Processing*, 2017, 65(1): 105-118.
- [28] Wen F, Pei L, Yang Y, et al. Efficient and robust recovery of sparse signal and image using generalized nonconvex regularization[J]. *IEEE Transactions on Computational Imaging*, 2017, 3(4): 566-579.
- [29] Petukhov A. Fast implementation of orthogonal greedy algorithm for tight wavelet frames[J]. *Signal processing*, 2006, 86(3): 471-479.
- [30] Eldar Y C, Kuppinger P, Bolcskei H. Block-sparse signals: Uncertainty relations and efficient recovery[J]. *IEEE Transactions on Signal Processing*, 2010, 58(6): 3042-3054.
- [31] Ghalehjeh S H, Babaie-Zadeh M, Jutten C. Fast block-sparse decomposition based on SL0[C]. *International Conference on Latent Variable Analysis and Signal Separation*. Springer, Berlin, Heidelberg, 2010: 426-433.
- [32] Deng W, Yin W, Zhang Y. Group sparse optimization by alternating direction method[C]. *Wavelets and Sparsity XV*. International Society for Optics and Photonics, 2013, 8858: 88580R.