Recovery analysis for weighted mixed ℓ_2/ℓ_p minimization with $0 < p \leq 1$

Zhiyong Zhou[*], Jun Yu
Department of Mathematics and Mathematical Statistics, Umeå University,
Umeå, 901 87, Sweden

September 4, 2017

Abstract: We study the recovery conditions of weighted mixed $\ell_2/\ell_p\,(0 < p \leq 1)$ minimization for block sparse signals reconstruction from compressed measurements when partial block support information is available. We show that the block p-restricted isometry property can ensure the robust recovery. Moreover, we present the sufficient and necessary condition for the recovery by using weighted block p-null space property. The relationship between the block p-RIP and the weighted block p-null space property has been established at the same time. Finally, we illustrate our results with a series of numerical experiments.
Keywords: Compressive sensing; Prior support information; Block sparse; Non-convex minimization.

1 Introduction

Since its advent [3–5, 10], Compressive Sensing (CS) has attracted considerable attentions (see the monographs [13, 18] for a comprehensive view). CS aims to recover an unknown signal with underdetermined linear measurements. Specifically, the standard CS problem consists in the reconstruction of a N-dimensional sparse or compressible signal x from a significantly fewer number of linear measurements $y = Ax + e \in \mathbb{R}^m$ with $m \ll N$ and the noise e satisfying $\|e\|_2 \leq \varepsilon$ for some known constant $\varepsilon > 0$. Then, if the measurement matrix A satisfies the restricted isometry property (RIP) condition, the robust recovery can be guaranteed by using the ℓ_1 minimization

$$\min_{z} \|z\|_1, \text{ subject to } \|y - Az\|_2 \leq \varepsilon. \tag{1}$$

To obtain better recovery performance, the structures and some prior information of signals are incorporated in the recovery algorithms. In this paper, we consider both cases. As for the structure, we assume that the unknown signal x is block sparse or nearly block sparse [11, 12, 14–16], which means that the nonzero entries of x occur in clusters. Block sparse model appears in many practical scenarios, such as when dealing with multi-band signals [27], in measurements of gene expression levels [30], and in colour imaging [24]. Moreover, block sparse model can be used to treat the problems of multiple measurement vector (MMV) [8, 14, 26] and sampling signals that lie in a union of subspaces [2, 14, 27].

[*]Corresponding author, zhiyong.zhou@umu.se. This work is supported by the Swedish Research Council grant (Reg.No. 340-2013-5342).

With $N = \sum_{i=1}^{n} d_i$, we define the i-th block $x[i]$ of a length-N vector x over $\mathcal{I} = \{d_1, \cdots, d_n\}$. The i-th block is of length d_i, and the blocks are formed sequentially so that

$$x = (\underbrace{x_1 \cdots x_{d_1}}_{x^T[1]} \underbrace{x_{d_1+1} \cdots x_{d_1+d_2}}_{x^T[2]} \cdots \underbrace{x_{N-d_n+1} \cdots x_N}_{x^T[n]})^T. \tag{2}$$

Without loss of generality, we assume that $d_1 = d_2 = \cdots = d_n = d$, implying that $N = nd$. A vector $x \in \mathbb{R}^N$ is called block k-sparse over $\mathcal{I} = \{d, \cdots, d\}$ if $x[i]$ is nonzero for at most k indices i. In other words, by denoting $\|x\|_{2,0} = \sum_{i=1}^{n} I(\|x[i]\|_2 > 0)$, a block k-sparse vector x can be defined by $\|x\|_{2,0} \leq k$. For any $p > 0$, we define the mixed ℓ_2/ℓ_p norm $\|x\|_{2,p} = (\sum_{i=1}^{n} \|x[i]\|_2^p)^{1/p}$. To make explicit use of the block structure to achieve better sparse recovery performance, the corresponding extended version of sparse representation algorithm has been developed, namely the mixed ℓ_2/ℓ_1 minimization,

$$\min_z \|z\|_{2,1}, \text{ subject to } \|y - Az\|_2 \leq \varepsilon. \tag{3}$$

It was shown in [14] that if the measurement matrix A satisfies the block RIP condition which generalizes the conventional RIP notion, then the mixed ℓ_2/ℓ_1-norm recovery algorithm is guaranteed to recover any block sparse signal, irrespectively of the locations of the nonzero blocks. Furthermore, recovery will be robust in the presence of noise and modeling errors (i.e., when the vector is not exactly block sparse). Other existing recovery algorithms include group lasso [39], adaptive lasso [23], iterative reweighted ℓ_2/ℓ_1 recovery algorithms [40], block version of the CoSaMP algorithm [12], and the extensions of the CoSaMP algorithm and of the Iterative Hard Thresholding algorithm [1].

On the other hand, we also consider the case that an estimate of the block support of the signal is available. The related literatures on signal recovery using partial support or block support information include [19–22, 28, 29, 31–33, 38]. For an arbitrary signal $x \in \mathbb{R}^N$ defined as (2), let x^k be its best approximation with k nonzero blocks, so that x^k minimizes $\|x - g\|_{2,1}$ over all block k-sparse vectors g. Let T_0 be the block support of x^k, where $T_0 \subset \{1, \cdots, n\}$ and $|T_0| \leq k$. Let $\tilde{T}$, the block support estimate, be a subset of $\{1, 2 \cdots, n\}$ with cardinality $|\tilde{T}| = \rho k$, where $0 \leq \rho \leq a$ for some $a > 1$ and $|\tilde{T} \cap T_0| = \alpha \rho k$ (for interpretation of ρ and α see Remark 1 in Section 2). To incorporate prior block support information $\tilde{T}$, the weighted mixed ℓ_2/ℓ_1 minimization

$$\min_z \sum_{i=1}^{n} w_i \|z[i]\|_2, \text{ subject to } \|y - Az\|_2 \leq \varepsilon, \text{ where } w_i = \begin{cases} \omega \in [0,1], & i \in \tilde{T} \\ 1, & i \in \tilde{T}^c \end{cases} \tag{4}$$

is adopted. The main idea is to choose ω such that the ℓ_2 norm of the blocks of x that are "expected" to be large are penalized less in the weighted objective function.

Moreover, as is shown in many literatures [6, 17, 37], ℓ_p-minimization with $0 < p < 1$ allows exact recovery with fewer measurements than that by ℓ_1-minimization. Thus, it is natural to adopt the nonconvex minimization to the setting of block sparse signal reconstruction with prior block support information. Specifically, we consider the weighted mixed $\ell_2/\ell_p\,(0 < p \leq 1)$ minimization problem:

$$\min_z \sum_{i=1}^{n} w_i \|z[i]\|_2^p, \text{ subject to } \|y - Az\|_2 \leq \varepsilon, \text{ where } w_i = \begin{cases} \omega \in [0,1], & i \in \tilde{T} \\ 1, & i \in \tilde{T}^c \end{cases}. \tag{5}$$

When there is no prior block support information ($\omega = 1$), the mixed $\ell_2/\ell_p(0 < p \leq 1)$ minimization problem has been studied in [34, 35]. And the case that there is partially known signal block support but with $\omega = 0$ was considered in [20]. We generalize the existing results to incorporating the prior known block support information with a weight $\omega \in [0, 1]$. In summary, the main contributions of our work include the following aspects. First, we provide the recovery analysis for the weighted mixed $\ell_2/\ell_p\,(0 < p \leq 1)$ minimization by using block p-RIP condition. This result extends the existing literatures [19, 20, 33–35]. Second, we propose the weighted block p-null space property and present the sufficient and necessary condition for the weighted mixed $\ell_2/\ell_p\,(0 < p \leq 1)$ minimization by this new tool. Third, we establish the relationship between block p-RIP condition and weighted block p-null space property. Finally, we illustrate our results via a series of simulations.

The paper is organized as follows. In Section 2, we present the sufficient recovery condition by using block p-RIP. In Section 3, we introduce the notion of weighted block p-null space property (NSP) and establish the sufficient and necessary condition with this new tool. In Section 4, we establish the relationship between these two conditions. In Section 5, we conduct some simulations to illustrate the theoretical results. Section 6 is devoted to the conclusion.

2 Block p-RIP

As for the weighted mixed $\ell_2/\ell_p\,(0 < p \leq 1)$ minimization, we can obtain the reconstruction guarantees by using block p-RIP. We begin with introducing the definition of block restricted p-isometry constant, which is a natural extension of the conventional restricted p-isometry constant.

Definition 1 ([7, 20, 34, 35]) *Given a measurement matrix $A \in \mathbb{R}^{m\times N}$ and $0 < p \leq 1$. Then the block p-restricted isometry constant (RIC) $\delta_{k|\mathcal{I}}$ over $\mathcal{I} = \{d_1, \cdots, d_n\}$ of order k is defined to be the smallest positive number such that*

$$(1 - \delta_{k|\mathcal{I}})\|x\|_2^p \leq \|Ax\|_p^p \leq (1 + \delta_{k|\mathcal{I}})\|x\|_2^p \tag{6}$$

for all $x \in \mathbb{R}^N$ that are block k-sparse over $\mathcal{I}$.

For convenience, we write δ_k for the block p-RIC $\delta_{k|\mathcal{I}}$ whenever there is no confusion. Then, we have the sufficient condition for the robust recovery result by using (5) with the block p-RIC δ_k.

Theorem 1 *Let $x \in \mathbb{R}^N$, and x^k be its best approximation with k nonzero blocks, supported on block index set T_0. Let $\tilde{T} \subset \{1, 2, \cdots, n\}$ be an arbitrary set. Define ρ and α as before such that $|\tilde{T}| = \rho k$ and $|\tilde{T} \cap T_0| = \alpha\rho k$. Suppose that there exists an $a \in \mathbb{Z}$, with $a \geq (1-\alpha)\rho$, $a > 1$, and the measurement matrix A satisfies*

$$\delta_{ak} + \frac{a^{1-p/2}}{\gamma}\delta_{(a+1)k} < \frac{a^{1-p/2}}{\gamma} - 1, \tag{7}$$

where $\gamma = \omega + (1-\omega)(1+\rho-2\alpha\rho)^{1-p/2}$ *for some given* $0 \le \omega \le 1$*. Then the solutions* $x^\sharp$ *of problem (5) obeys*

$$\|x^\sharp - x\|_2 \le C_1 \frac{\left(\omega\|x-x^k\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c \cap T_0^c}\|_{2,p}^p\right)^{1/p}}{k^{1/p-1/2}} + C_2\varepsilon, \tag{8}$$

for some positive constants C_1 *and* C_2*.*

Remark 1. Note that this theorem involves two important ratios: ρ determines the ratio of the estimated block support size to the actual block support of x^k (or the block support of x if x is block k-sparse), while α determines the ratio of the number of block indices in block support of x^k that were accurately estimated in $\tilde{T}$ to the size of $\tilde{T}$. Specifically, $\alpha = \frac{|\tilde{T}\cap T_0|}{|\tilde{T}|}$.

Remark 2. The constants C_1 and C_2 are explicitly given by the following expressions:

$$\begin{aligned} C_1 &= \frac{2^{2/p-1}a^{1/2-1/p}\left[(1+\delta_{ak})^{1/p} + (1-\delta_{(a+1)k})^{1/p}\right]}{[(1-\delta_{(a+1)k}) - a^{p/2-1}(1+\delta_{ak})\gamma]^{1/p}}, \\ C_2 &= \frac{2^{1/p}m^{1/p-1/2}(1+a^{1/2-1/p}\gamma^{1/p})}{[(1-\delta_{(a+1)k}) - a^{p/2-1}(1+\delta_{ak})\gamma]^{1/p}}. \end{aligned}$$

Remark 3. For Theorem 1 to be held, it is sufficient that A satisfies

$$\delta_{(a+1)k} < \delta := \frac{a^{1-p/2} - [\omega + (1-\omega)(1+\rho-2\alpha\rho)^{1-p/2}]}{a^{1-p/2} + [\omega + (1-\omega)(1+\rho-2\alpha\rho)^{1-p/2}]}. \tag{9}$$

Next, we illustrate how the slighted stronger sufficient conditions vary with α and ω for $p = 0.01, 0.5, 1$, respectively. In Figure 1, for each p, we plot, for different values of α, δ versus ω, where we set the parameters $a = 3$ and $\rho = 1$. We observe that as α increases the sufficient condition on the block p-RIC becomes weaker, allowing for a wider class of measurement matrices A. And for each p with $\alpha > 0.5$, see $\alpha = 0.7$ and 0.9 in our setting, as ω decreases, the sufficient condition becomes weaker. For instance, when $\alpha = 0.7, p = 0.5$, with $\omega = 0.2$ it suffices to have $\delta = 0.5072$, compared with $\delta = 0.3902$ for $\omega = 1$. The opposite conclusion holds for the case $\alpha < 0.5$. When $\alpha = 0.5$, the sufficient condition remains the same for different ω. From another point of view, as p decreases, the sufficient condition becomes weaker, which reflects the benefits of using nonconvex minimization.

Remark 4. By setting $\omega = 1$ and $a = b^{2/(2-p)}$ with $b > 1$, this theorem reduces to Theorem 2.1 in [35]. In addition, by setting $\omega = 0, \alpha = 0.5$ and $a = \rho = b^{2/(2-p)}$ with $b > 1$, then this theorem goes to Theorem 2 in [20] with $s = ak$.

Proof. Let $x^\sharp = x + h$ be a solution of problem (5), where x is the unknown true signal. Throughout the paper, x_T will denote the vector equal to x on the block index set T and zero elsewhere. Then, we have

$$\omega\|x^\sharp_{\tilde{T}}\|_{2,p}^p + \|x^\sharp_{\tilde{T}^c}\|_{2,p}^p \le \omega\|x_{\tilde{T}}\|_{2,p}^p + \|x_{\tilde{T}^c}\|_{2,p}^p.$$

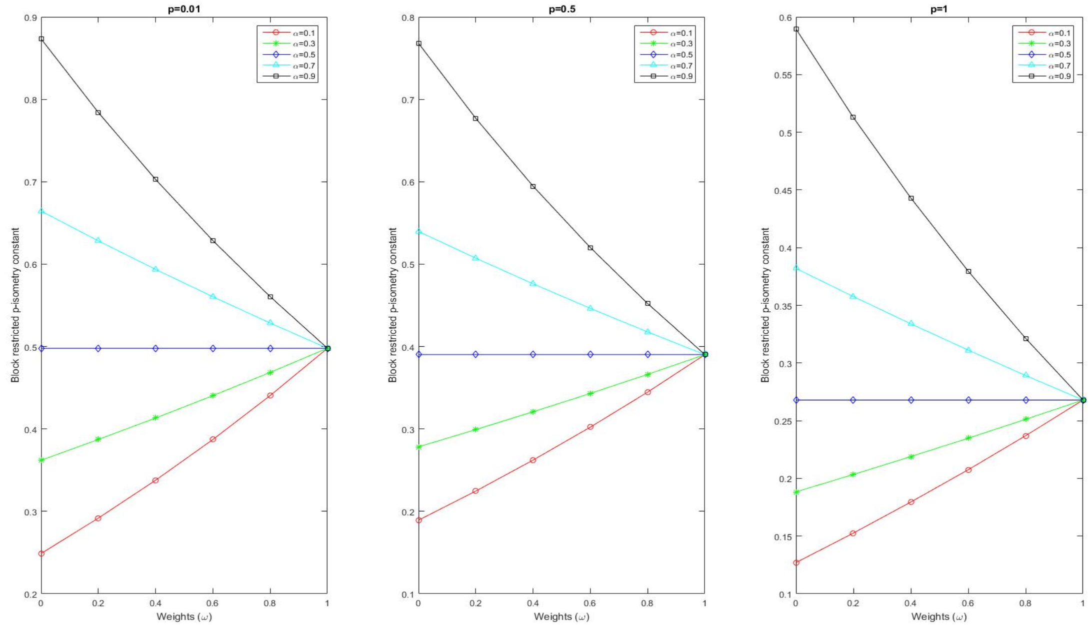

Figure 1: Comparison of the sufficient conditions for the block restricted p-isometry constants with various of α for $p = 0.01, 0.5, 1$. We set $a = 3$ and $\rho = 1$ for all the figures.

That is,

$$\omega\|x_{\tilde{T}} + h_{\tilde{T}}\|_{2,p}^p + \|x_{\tilde{T}^c} + h_{\tilde{T}^c}\|_{2,p}^p \leq \omega\|x_{\tilde{T}}\|_{2,p}^p + \|x_{\tilde{T}^c}\|_{2,p}^p.$$

Consequently,

$$\begin{aligned}&\omega\|x_{\tilde{T}\cap T_0} + h_{\tilde{T}\cap T_0}\|_{2,p}^p + \omega\|x_{\tilde{T}\cap T_0^c} + h_{\tilde{T}\cap T_0^c}\|_{2,p}^p + \|x_{\tilde{T}^c\cap T_0} + h_{\tilde{T}^c\cap T_0}\|_{2,p}^p + \|x_{\tilde{T}^c\cap T_0^c} + h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\\ &\leq \omega\|x_{\tilde{T}\cap T_0}\|_{2,p}^p + \omega\|x_{\tilde{T}\cap T_0^c}\|_{2,p}^p + \|x_{\tilde{T}^c\cap T_0}\|_{2,p}^p + \|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p.\end{aligned}$$

The forward and reverse triangle inequalities implies

$$\omega\|h_{\tilde{T}\cap T_0^c}\|_{2,p}^p + \|h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p \leq \|h_{\tilde{T}^c\cap T_0}\|_{2,p}^p + \omega\|h_{\tilde{T}\cap T_0}\|_{2,p}^p + 2\Big(\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p + \omega\|x_{\tilde{T}\cap T_0^c}\|_{2,p}^p\Big).$$

Adding and subtracting $\omega\|h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p$ on the left hand side, and $\omega\|h_{\tilde{T}^c\cap T_0}\|_{2,p}^p$ on the right, we obtain

$$\omega\|h_{T_0^c}\|_{2,p}^p + (1-\omega)\|h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p \leq \omega\|h_{T_0}\|_{2,p}^p + (1-\omega)\|h_{\tilde{T}^c\cap T_0}\|_{2,p}^p + 2\Big(\omega\|x_{T_0^c}\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\Big).$$

But we can also write

$$\|h_{T_0^c}\|_{2,p}^p = \omega\|h_{T_0^c}\|_{2,p}^p + (1-\omega)\|h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p + (1-\omega)\|h_{\tilde{T}\cap T_0^c}\|_{2,p}^p.$$

Therefore, we have

$$\|h_{T_0^c}\|_{2,p}^p \leq \omega\|h_{T_0}\|_{2,p}^p + (1-\omega)\|h_{\tilde{T}\cap T_0^c}\|_{2,p}^p + (1-\omega)\|h_{\tilde{T}^c\cap T_0}\|_{2,p}^p + 2\Big(\omega\|x_{T_0^c}\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\Big).$$

Let the set $\tilde{T}_\alpha = T_0 \cap \tilde{T}$, then we can write $\|h_{\tilde{T}\cap T_0^c}\|_{2,p}^p + \|h_{\tilde{T}^c\cap T_0}\|_{2,p}^p = \|h_{(T_0\cup\tilde{T})\setminus\tilde{T}_\alpha}\|_{2,p}^p$ and simplify the bound on

$$\|h_{T_0^c}\|_{2,p}^p \leq \omega\|h_{T_0}\|_{2,p}^p + (1-\omega)\|h_{(T_0\cup\tilde{T})\setminus\tilde{T}_\alpha}\|_{2,p}^p + 2\Big(\omega\|x_{T_0^c}\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\Big). \tag{10}$$

Next, we decompose $h_{T_0^c}$ into disjoint block index set T_j, each of T_j $(j \geq 1)$ consists of ak blocks, where $a > 1$. That is, T_1 indexes the ak blocks with largest ℓ_2 norm of $h_{T_0^c}$, T_2 indexes the second ak blocks with largest ℓ_2 norm of $h_{(T_0\cup T_1)^c}$, and so on. Then this gives $h_{T_0^c} = \sum_{j\geq 1} h_{T_j}$. For each $i \in T_j$ $(j \geq 2)$, it is easy to see that

$$\|h_{T_j}[i]\|_2^p \leq \frac{\|h_{T_{j-1}}[i]\|_2^p + \cdots + \|h_{T_{j-1}}[ak]\|_2^p}{ak} = \frac{\|h_{T_{j-1}}\|_{2,p}^p}{ak}.$$

Then

$$\|h_{T_j}[i]\|_2^2 \leq \frac{\|h_{T_{j-1}}\|_{2,p}^2}{(ak)^{2/p}}, \quad \|h_{T_j}\|_2^2 \leq \frac{ak\|h_{T_{j-1}}\|_{2,p}^2}{(ak)^{2/p}},$$
$$\|h_{T_j}\|_2^p \leq \frac{\|h_{T_{j-1}}\|_{2,p}^p}{(ak)^{1-p/2}}.$$

Thus,

$$\|h_{(T_0\cup T_1)^c}\|_2 \leq \|\sum_{j\geq 2} h_{T_j}\|_2 \leq \sum_{j\geq 2}\|h_{T_j}\|_2 \leq (ak)^{1/2-1/p}\sum_{j\geq 2}\|h_{T_{j-1}}\|_{2,p} = (ak)^{1/2-1/p}\sum_{j\geq 1}\|h_{T_j}\|_{2,p}.$$

Hence,

$$\|h_{(T_0\cup T_1)^c}\|_2^p \leq (ak)^{p/2-1}\left(\sum_{j\geq 1}\|h_{T_j}\|_{2,p}\right)^p \leq (ak)^{p/2-1}\sum_{j\geq 1}\|h_{T_j}\|_{2,p}^p = (ak)^{p/2-1}\|h_{T_0^c}\|_{2,p}^p.$$

Combining the above expression with (10), we get

$$\|h_{(T_0\cup T_1)^c}\|_2^p \leq (ak)^{p/2-1}\left[\omega\|h_{T_0}\|_{2,p}^p + (1-\omega)\|h_{(T_0\cup\tilde{T})\setminus\tilde{T}_\alpha}\|_{2,p}^p + 2\Big(\omega\|x_{T_0^c}\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\Big)\right]. \tag{11}$$

By Hölder's inequality it follows that

$$\|Ah\|_p^p \leq \left(\sum_{i=1}^m (|(Ah)_i|^p)^{2/p}\right)^{p/2} \cdot \left(\sum_{i=1}^m 1\right)^{1-p/2} = m^{1-p/2}\|Ah\|_2^p.$$

Since $x^\sharp$ is a solution of problem (5) and $\|y - Ax\|_2 = \|e\|_2 \leq \varepsilon$, so we have

$$\|Ah\|_2 = \|A(x^\sharp - x)\|_2 \leq \|Ax - y\|_2 + \|Ax^\sharp - y\|_2 \leq 2\varepsilon.$$

Thus, $\|Ah\|_p^p \leq m^{1-p/2}\|Ah\|_2^p \leq m^{1-p/2}(2\varepsilon)^p$. Therefore, we obtain

$$\|Ah_{T_0\cup T_1}\|_p^p \leq \|Ah\|_p^p + \|Ah_{(T_0\cup T_1)^c}\|_p^p \leq m^{1-p/2}(2\varepsilon)^p + \sum_{j\geq 2}\|Ah_{T_j}\|_p^p$$
$$\leq m^{1-p/2}(2\varepsilon)^p + (1+\delta_{ak})\sum_{j\geq 2}\|h_{T_j}\|_2^p.$$

Moreover, we have

$$\sum_{j\geq 2}\|h_{T_j}\|_2^p \leq (ak)^{p/2-1}\sum_{j\geq 2}\|h_{T_{j-1}}\|_{2,p}^p \leq (ak)^{p/2-1}\|h_{T_0^c}\|_{2,p}^p.$$

Thus, we have

$$\begin{aligned}\|Ah_{T_0\cup T_1}\|_p^p &\leq m^{1-p/2}(2\varepsilon)^p + (ak)^{p/2-1}(1+\delta_{ak})\|h_{T_0^c}\|_{2,p}^p \\ &\leq m^{1-p/2}(2\varepsilon)^p + 2(ak)^{p/2-1}(1+\delta_{ak})\Big(\omega\|x_{T_0^c}\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\Big) \\ &\quad + \omega(ak)^{p/2-1}(1+\delta_{ak})\|h_{T_0}\|_{2,p}^p + (1-\omega)(ak)^{p/2-1}(1+\delta_{ak})\|h_{(T_0\cup\tilde{T})\setminus\tilde{T}_\alpha}\|_{2,p}^p.\end{aligned}$$

Noting that $|(T_0\cup\tilde{T})\setminus\tilde{T}_\alpha| = (1+\rho-2\alpha\rho)k$, then $\|h_{(T_0\cup\tilde{T})\setminus\tilde{T}_\alpha}\|_{2,p}^p \leq [(1+\rho-2\alpha\rho)k]^{1-p/2}\|h_{(T_0\cup\tilde{T})\setminus\tilde{T}_\alpha}\|_2^p$. Since the set T_1 contains the ak blocks with largest ℓ_2 norm of $h_{T_0^c}$ with $a>1$, and $|\tilde{T}\setminus\tilde{T}_\alpha| = (1-\alpha)\rho k \leq ak$, then $\|h_{(T_0\cup\tilde{T})\setminus\tilde{T}_\alpha}\|_2^p \leq \|h_{T_0\cup T_1}\|_2^p$, which implies that $\|h_{(T_0\cup\tilde{T})\setminus\tilde{T}_\alpha}\|_{2,p}^p \leq [(1+\rho-2\alpha\rho)k]^{1-p/2}\|h_{T_0\cup T_1}\|_2^p$. In addition, $\|h_{T_0}\|_{2,p}^p \leq k^{1-p/2}\|h_{T_0}\|_2^p \leq k^{1-p/2}\|h_{T_0\cup T_1}\|_2^p$. Thus,

$$\begin{aligned}(1-\delta_{(a+1)k})\|h_{T_0\cup T_1}\|_2^p &\leq \|Ah_{T_0\cup T_1}\|_p^p \\ &\leq m^{1-p/2}(2\varepsilon)^p + 2(ak)^{p/2-1}(1+\delta_{ak})\Big(\omega\|x_{T_0^c}\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\Big) \\ &\quad + \omega a^{p/2-1}(1+\delta_{ak})\|h_{T_0\cup T_1}\|_2^p + (1-\omega)a^{p/2-1}(1+\rho-2\alpha\rho)^{1-p/2}(1+\delta_{ak})\|h_{T_0\cup T_1}\|_2^p.\end{aligned}$$

Therefore, if $(1-\delta_{(a+1)k}) - a^{p/2-1}(1+\delta_{ak})\gamma > 0$, that is $\delta_{ak} + \frac{a^{1-p/2}}{\gamma}\delta_{(a+1)k} < \frac{a^{1-p/2}}{\gamma} - 1$, then we have

$$\|h_{T_0\cup T_1}\|_2^p \leq \frac{m^{1-p/2}(2\varepsilon)^p + 2(ak)^{p/2-1}(1+\delta_{ak})\Big(\omega\|x_{T_0^c}\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\Big)}{(1-\delta_{(a+1)k}) - a^{p/2-1}(1+\delta_{ak})\gamma}, \tag{12}$$

where $\gamma = \omega + (1-\omega)(1+\rho-2\alpha\rho)^{1-p/2}$. According to (11), we also have

$$\begin{aligned}\|h_{(T_0\cup T_1)^c}\|_2^p &\leq a^{p/2-1}\gamma\|h_{T_0\cup T_1}\|_2^p + 2(ak)^{p/2-1}\Big(\omega\|x_{T_0^c}\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\Big) \\ &\leq \frac{a^{p/2-1}\gamma m^{1-p/2}(2\varepsilon)^p + 2(ak)^{p/2-1}(1-\delta_{(a+1)k})\Big(\omega\|x_{T_0^c}\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\Big)}{(1-\delta_{(a+1)k}) - a^{p/2-1}(1+\delta_{ak})\gamma}.\end{aligned}$$

Therefore, since $\|v\|_p \leq 2^{1/p-1}\|v\|_1$ for all $v\in\mathbb{R}^2$, we have

$$\begin{aligned}\|h\|_2 &\leq \|h_{T_0\cup T_1}\|_2 + \|h_{(T_0\cup T_1)^c}\|_2 \\ &\leq 2^{1/p-1}\left(\frac{2m^{1/p-1/2}\varepsilon + 2^{1/p}(ak)^{1/2-1/p}(1+\delta_{ak})^{1/p}\Big(\omega\|x_{T_0^c}\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\Big)^{1/p}}{[(1-\delta_{(a+1)k}) - a^{p/2-1}(1+\delta_{ak})\gamma]^{1/p}}\right) \\ &\quad + 2^{1/p-1}\left(\frac{2a^{1/2-1/p}\gamma^{1/p}m^{1/p-1/2}\varepsilon + 2^{1/p}(ak)^{1/2-1/p}(1-\delta_{(a+1)k})^{1/p}\Big(\omega\|x_{T_0^c}\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\Big)^{1/p}}{[(1-\delta_{(a+1)k}) - a^{p/2-1}(1+\delta_{ak})\gamma]^{1/p}}\right) \\ &\leq \frac{2^{2/p-1}(ak)^{1/2-1/p}\left[(1+\delta_{ak})^{1/p} + (1-\delta_{(a+1)k})^{1/p}\right]\Big(\omega\|x_{T_0^c}\|_{2,p}^p + (1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\Big)^{1/p}}{[(1-\delta_{(a+1)k}) - a^{p/2-1}(1+\delta_{ak})\gamma]^{1/p}}\end{aligned}$$

$$+\frac{2^{1/p}m^{1/p-1/2}(1+a^{1/2-1/p}\gamma^{1/p})\varepsilon}{[(1-\delta_{(a+1)k})-a^{p/2-1}(1+\delta_{ak})\gamma]^{1/p}}=C_1\frac{\left(\omega\|x_{T_0^c}\|_{2,p}^p+(1-\omega)\|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p\right)^{1/p}}{k^{1/p-1/2}}+C_2\varepsilon,$$

which completes the proof.

Corollary 1 *Let $A\in\mathbb{R}^{m\times N}$ be a measurement matrix, $x\in\mathbb{R}^N$ be a block k-sparse signal supported on block index set T_0 with $y=Ax$, and $0<p\leq 1$. Let $\tilde{T}\subset\{1,2,\cdots,n\}$ be an arbitrary set. Define ρ and α as before such that $|\tilde{T}|=\rho k$ and $|\tilde{T}\cap T_0|=\alpha\rho k$. Suppose that there exists an $a\in\mathbb{Z}$, with $a\geq(1-\alpha)\rho$, $a>1$, and the measurement matrix A satisfies*

$$\delta_{ak}+\frac{a^{1-p/2}}{\gamma}\delta_{(a+1)k}<\frac{a^{1-p/2}}{\gamma}-1,\tag{13}$$

where $\gamma=\omega+(1-\omega)(1+\rho-2\alpha\rho)^{1-p/2}$ for some given $0\leq\omega\leq 1$. Then the unique solution of problem (5) with $\varepsilon=0$ is exactly x.

Remark 5. According to the arguments in [20, 35], if we let A be an $m\times N$ matrix whose entries are i.i.d Gaussian random variables with mean zero, then there exist $C_3(p)$ and $C_4(p)$ such that whenever $0<p\leq 1$ and $m\geq C_3(p)kd+pC_4(p)k\ln(n/k)$, the block p-RIP (7) will hold and the block sparse signal x can be exactly recovered with high probability. For a given $p\in(0,1]$, $C_3(p)$ and $C_4(p)$ are finite constants, and the second term of the above expression has the dominant impact on the number of measurements in an asymptotic sense. When $p\to 0$, the condition reduces to $m\geq C_3(0)kd$. While, when $p=1$, the required number of measurements $m\geq C_3(1)kd+C_4(1)k\ln(n/k)$, which implies fewer measurements are required with smaller p for exact recovery via weighted mixed ℓ_2/ℓ_p minimization than the case that $p=1$. This is largely consistent with what we have mentioned in Remark 3, namely, as p decreases, the sufficient condition for the exact recovery becomes weaker. Though we have only considered the case that $d_1=d_2=\cdots=d_n=d$, the results here can be adapted to the case in which d_i are not equal via replacing d by the maximal block length $\max\limits_{1\leq i\leq n} d_i$.

3 Weighted Block p-NSP

In this section, we focus on the noise free case, that is $e=0$, and the signal x is exactly block k-sparse, supported on block index set T_0. We obtain the exact recovery condition for the problem (5) with $\varepsilon=0$ by using weighted block p-null space property. For an index set $V\subset\{1,\cdots,n\}$, we define

$$\Gamma_s(V):=\{U\subset\{1,\cdots,n\}:|(V\cap U^c)\cup(V^c\cap U)|\leq s\}.$$

Definition 2 *Let $T\subset\{1,\cdots,n\}$ with $|T|\leq k$ and $\tilde{T}\in\Gamma_s(T)$. Assume the block size equals d. A matrix $A\in\mathbb{R}^{m\times N}$ is said to have the weighted nonuniform block p-null space property with parameters T and $\tilde{T}$, and constant C if for any vector $h: Ah=0$, we have*

$$\omega\|h_T\|_{2,p}^p+(1-\omega)\|h_S\|_{2,p}^p\leq C\|h_{T^c}\|_{2,p}^p,\tag{14}$$

where $S=(\tilde{T}\cap T^c)\cup(\tilde{T}^c\cap T)$. In this case, we say A satisfies ω-d-p-NSP$(T,\tilde{T},C)$.

Next, we define a weighted uniform block p-null space property that leads to a sufficient and necessary condition for the exact recovery of all block k-sparse signals from compressive measurements using weighted mixed ℓ_2/ℓ_p minimization problem.

Definition 3 *Assume the block size equals d. A matrix $A \in \mathbb{R}^{m\times N}$ is said to have the weighted block p-null space property with parameters k and s, and constant C if for any vector $h : Ah = 0$, and for every block index set $T \subset \{1, \cdots, n\}$ with $|T| \leq k$ and $S \subset \{1, \cdots, n\}$ with $|S| \leq s$, we have*

$$\omega\|h_T\|_{2,p}^p + (1-\omega)\|h_S\|_{2,p}^p \leq C\|h_{T^c}\|_{2,p}^p. \tag{15}$$

*In this case, we say A satisfies ω-d-p-*NSP(k, s, C).

Remark 6. Our notations ω-d-p-NSP$(T, \tilde{T}, C)$ and ω-d-p-NSP(k, s, C) are direct extensions of the recent notations ω-NSP$(T, \tilde{T}, C)$ and ω-NSP(k, s, C) proposed by [25], which corresponds to the special case that $p = 1$ and the block size $d = 1$. That's to say ω-1-1-NSP(k, s, C) is equivalent to ω-NSP(k, s, C). Consequently, 1-1-1-NSP(k, k, C) is the standard null space property of order k, i.e, NSP(k, C).

The following theorem presents the sufficient and necessary condition for weighted mixed ℓ_2/ℓ_p minimization problem to recover all block k-sparse signals when the error in the block support estimate is of size s or less.

Theorem 2 *Assume the block size equals d. Given a matrix $A \in \mathbb{R}^{m\times N}$, every block k-sparse signal $x \in \mathbb{R}^N$, supported on block index set T_0, is the unique solution of the weighted mixed ℓ_2/ℓ_p, $0 < p \leq 1$ norm minimization problem (5) with $\varepsilon = 0$ and $\tilde{T} = \Gamma_s(T_0)$, if and only if A satisfies ω-d-p-*NSP(k, s, C) *for some positive constant $C < 1$.*

Proof. a) " $\Rightarrow$ ". Assume if A does not satisfy ω-d-p-NSP(k, s, C) for any constant $C > 1$, then there exists a vector $h : Ah = 0$ and block index set T with $|T| \leq k$ and block index set S with $|S| \leq s$, such that $Ah_T = -Ah_{T^c}$ and

$$\omega\|h_T\|_{2,p}^p + (1-\omega)\|h_S\|_{2,p}^p \geq \|h_{T^c}\|_{2,p}^p.$$

Define $\tilde{T} = (T^c \cap S) \cup (T \cap S^c)$ so that $S = (T \cap \tilde{T}^c) \cup (T^c \cap \tilde{T})$. Substituting for S and splitting the set T, we obtain

$$\begin{aligned}&\omega\left(\|h_{T\cap\tilde{T}^c}\|_{2,p}^p + \|h_{T\cap\tilde{T}}\|_{2,p}^p\right) + (1-\omega)\left(\|h_{T\cap\tilde{T}^c}\|_{2,p}^p + \|h_{T^c\cap\tilde{T}}\|_{2,p}^p\right)\\ &= \|h_{T\cap\tilde{T}^c}\|_{2,p}^p + \omega\|h_{T\cap\tilde{T}}\|_{2,p}^p + (1-\omega)\|h_{T^c\cap\tilde{T}}\|_{2,p}^p \geq \|h_{T^c}\|_{2,p}^p.\end{aligned}$$

Then, we have

$$\omega\|h_{T\cap\tilde{T}}\|_{2,p}^p + \|h_{T\cap\tilde{T}^c}\|_{2,p}^p \geq \omega\|h_{T^c\cap\tilde{T}}\|_{2,p}^p + \|h_{T^c\cap\tilde{T}^c}\|_{2,p}^p.$$

In other words, the weighted mixed ℓ_2/ℓ_p norm of the vector h_T equals or exceeds that of $-h_{T^c}$. So h_T is not the unique minimizer, which is contradictory to our condition. Thus, the necessity is verified.

b) " $\Leftarrow$ ". Let $x^\sharp$ be a minimizer of weighted mixed $\ell_2/\ell_p, 0 < p \leq 1$ norm minimization problem (5) with $\varepsilon = 0$ and $\tilde{T} = \Gamma_s(T_0)$ and define $h = x^\sharp - x$. Then, by the optimality of $x^\sharp$, we have

$$\omega\|x_{\tilde{T}} + h_{\tilde{T}}\|_{2,p}^p + \|x_{\tilde{T}^c} + h_{\tilde{T}^c}\|_{2,p}^p \leq \omega\|x_{\tilde{T}}\|_{2,p}^p + \|x_{\tilde{T}^c}\|_{2,p}^p,$$

which is equivalent to

$$\begin{aligned}
&\omega\|x_{\tilde{T}\cap T_0} + h_{\tilde{T}\cap T_0}\|_{2,p}^p + \omega\|x_{\tilde{T}\cap T_0^c} + h_{\tilde{T}\cap T_0^c}\|_{2,p}^p + \|x_{\tilde{T}^c\cap T_0} + h_{\tilde{T}^c\cap T_0}\|_{2,p}^p + \|x_{\tilde{T}^c\cap T_0^c} + h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p \\
&\leq \omega\|x_{\tilde{T}\cap T_0}\|_{2,p}^p + \omega\|x_{\tilde{T}\cap T_0^c}\|_{2,p}^p + \|x_{\tilde{T}^c\cap T_0}\|_{2,p}^p + \|x_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p.
\end{aligned}$$

Since x is strictly block k-sparse and supported on the block index set T_0, thus $x_{T_0^c} = 0$. Then we have

$$\omega\|h_{\tilde{T}\cap T_0^c}\|_{2,p}^p + \|h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p \leq \|h_{\tilde{T}^c\cap T_0}\|_{2,p}^p + \omega\|h_{\tilde{T}\cap T_0}\|_{2,p}^p.$$

Adding and subtracting $\omega\|h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p$ on the left hand side, and $\omega\|h_{\tilde{T}^c\cap T_0}\|_{2,p}^p$ on the right, we obtain

$$\begin{aligned}
&\omega\|h_{\tilde{T}\cap T_0^c}\|_{2,p}^p + \omega\|h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p + \|h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p - \omega\|h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p \\
&\leq \omega\|h_{\tilde{T}\cap T_0}\|_{2,p}^p + \omega\|h_{\tilde{T}^c\cap T_0}\|_{2,p}^p + \|h_{\tilde{T}^c\cap T_0}\|_{2,p}^p - \omega\|h_{\tilde{T}^c\cap T_0}\|_{2,p}^p.
\end{aligned}$$

Therefore,

$$\omega\|h_{T_0^c}\|_{2,p}^p + (1-\omega)\|h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p \leq \omega\|h_{T_0}\|_{2,p}^p + (1-\omega)\|h_{\tilde{T}^c\cap T_0}\|_{2,p}^p.$$

Finally, by adding $(1-\omega)\|h_{\tilde{T}\cap T_0^c}\|_{2,p}^p$ to both sides, we have

$$\begin{aligned}
\|h_{T_0^c}\|_{2,p}^p &\leq \omega\|h_{T_0}\|_{2,p}^p + (1-\omega)\|h_{\tilde{T}\cap T_0^c}\|_{2,p}^p + (1-\omega)\|h_{\tilde{T}^c\cap T_0^c}\|_{2,p}^p \\
&= \omega\|h_{T_0}\|_{2,p}^p + (1-\omega)\|h_S\|_{2,p}^p,
\end{aligned} \tag{16}$$

by setting $S = (\tilde{T}\cap T_0^c) \cup (\tilde{T}^c \cap T_0)$. Note that when $|S| \leq s$, the above inequality is in contradiction with that A satisfies ω-d-p-NSP(k, s, C) for some constant $C < 1$, unless $h = 0$. Hence, we have $x^\sharp = x$, and the sufficiency holds.

4 From Block p-RIP to Weighted Block p-NSP

In this part, we establish the relationship between the block p-RIP and weighted block p-NSP, that is block p-RIP can directly imply the weighted block p-NSP, then we can obtain the exact recovery for every block k-sparse signal via the weighted mixed $\ell_2/\ell_p, 0 < p \leq 1$ norm minimization problem (5) with $\varepsilon = 0$, as presented in the following theorem.

Theorem 3 *Assume the block size equals d and let $x \in \mathbb{R}^N$ be block k-sparse, supported on block index set T_0. Let $\tilde{T} \subset \{1, 2, \cdots, n\}$ be an arbitrary set and define ρ and α as before such that $|\tilde{T}| = \rho k$ and $|\tilde{T}\cap T_0| = \alpha\rho k$. Let $s = (1 + \rho - 2\alpha\rho)k$, then we have $\tilde{T} = \Gamma_s(T_0)$. Suppose that there exists an $a \in \mathbb{Z}$, with $a \geq (1-\alpha)\rho$, $a > 1$, and the measurement matrix A satisfies*

$$\delta_{ak} + \frac{a^{1-p/2}}{\gamma}\delta_{(a+1)k} < \frac{a^{1-p/2}}{\gamma} - 1, \tag{17}$$

where $\gamma = \omega + (1-\omega)(s/k)^{1-p/2}$ for some given $0 \leq \omega \leq 1$. Then A satisfies ω-d-p-NSP(k, s, C) for some constant $C < 1$.

Consequently, according to Theorem 2, we have x is the unique solution of problem (5) with $\varepsilon = 0$.

Proof. To show for any vector $h : Ah = 0$, and for every block index set $T \subset \{1, \cdots, n\}$ with $|T| \leq k$, and $S \subset \{1, \cdots, n\}$ with $|S| \leq s$, we have

$$\omega\|h_T\|_{2,p}^p + (1-\omega)\|h_S\|_{2,p}^p \leq C\|h_{T^c}\|_{2,p}^p, \tag{18}$$

for some constant $C < 1$, we only need to show (18) holds for $T = G_0$ and $S = H_0$, where G_0 is the block index set over the k blocks with largest ℓ_2 norm of h, H_0 is the block index set over the s blocks with largest ℓ_2 norm of h.

Next, we decompose $h_{G_0^c}$ into disjoint block index set G_j, each of G_j ($j \geq 1$) consists of ak blocks. That is, G_1 indexes the ak blocks with largest ℓ_2 norm of $h_{G_0^c}$. G_2 indexes the second ak blocks with largest ℓ_2 norm of $h_{(G_0\cup G_1)^c}$, and so on.

Since $Ah = 0$, then we have

$$\|Ah_{G_0\cup G_1}\|_p^p \leq \|Ah_{(G_0\cup G_1)^c}\|_p^p \leq \sum_{j\geq 2}\|Ah_{G_j}\|_p^p \leq (1+\delta_{ak})\sum_{j\geq 2}\|h_{G_j}\|_2^p.$$

Moreover, we have

$$\sum_{j\geq 2}\|h_{G_j}\|_2^p \leq (ak)^{p/2-1}\sum_{j\geq 2}\|h_{G_{j-1}}\|_{2,p}^p \leq (ak)^{p/2-1}\|h_{G_0^c}\|_{2,p}^p.$$

Thus,

$$(1-\delta_{(a+1)k})\|h_{G_0\cup G_1}\|_2^p \leq \|Ah_{G_0\cup G_1}\|_p^p \leq (1+\delta_{ak})(ak)^{p/2-1}\|h_{G_0^c}\|_{2,p}^p.$$

It implies that

$$\|h_{G_0\cup G_1}\|_2^p \leq \frac{(ak)^{p/2-1}(1+\delta_{ak})}{1-\delta_{(a+1)k}}\|h_{G_0^c}\|_{2,p}^p. \tag{19}$$

In addition, $\|h_{G_0}\|_{2,p}^p \leq k^{1-p/2}\|h_{G_0}\|_2^p \leq k^{1-p/2}\|h_{G_0\cup G_1}\|_2^p$. Hence,

$$\|h_{G_0}\|_{2,p}^p \leq k^{1-p/2}\|h_{G_0\cup G_1}\|_2^p \leq \frac{a^{p/2-1}(1+\delta_{ak})}{1-\delta_{(a+1)k}}\|h_{G_0^c}\|_{2,p}^p. \tag{20}$$

For the term $\|h_{H_0}\|_{2,p}^p$, we have $\|h_{H_0}\|_{2,p}^p \leq s^{1-p/2}\|h_{H_0}\|_2^p$. Moreover, if $s \leq k$, we have $H_0 \subset G_0$ and $\|h_{H_0}\|_2^p \leq \|h_{G_0}\|_2^p \leq \|h_{G_0\cup G_1}\|_2^p$. If $s > k$, then $G_0 \subset H_0$. However, since $|H_0 \setminus G_0| = s - k = (1+\rho-2\alpha\rho)k - k = (1-2\alpha)\rho k \leq ak$, then we also have $\|h_{H_0}\|_2^p \leq \|h_{G_0\cup G_1}\|_2^p$ as the block index set G_1 contains the ak blocks with largest ℓ_2 norm of $h_{G_0^c}$. As a consequence,

$$\|h_{H_0}\|_{2,p}^p \leq s^{1-p/2}\|h_{H_0}\|_2^p \leq s^{1-p/2}\|h_{G_0\cup G_1}\|_2^p \leq \frac{(ak/s)^{p/2-1}(1+\delta_{ak})}{1-\delta_{(a+1)k}}\|h_{G_0^c}\|_{2,p}^p. \tag{21}$$

Therefore, we obtain

$$\begin{aligned}\omega\|h_{G_0}\|_{2,p}^p + (1-\omega)\|h_{H_0}\|_{2,p}^p &\leq \omega\frac{a^{p/2-1}(1+\delta_{ak})}{1-\delta_{(a+1)k}}\|h_{G_0^c}\|_{2,p}^p + (1-\omega)\frac{(ak/s)^{p/2-1}(1+\delta_{ak})}{1-\delta_{(a+1)k}}\|h_{G_0^c}\|_{2,p}^p \\ &= \frac{a^{p/2-1}\gamma(1+\delta_{ak})}{1-\delta_{(a+1)k}}\|h_{G_0^c}\|_{2,p}^p.\end{aligned}$$

Let $C = \frac{a^{p/2-1}\gamma(1+\delta_{ak})}{1-\delta_{(a+1)k}}$, we have $C < 1$ under the condition of block p-RIP. The proof is completed.

5 Simulation

In this section, we conduct several simulations to illustrate our presented theoretical results. We adopt the iteratively reweighted least squares (IRLS) approach to solve the nonconvex optimization problem. We begin with $x^{(0)} = \arg\min \|y - Ax\|_2^2$, and set $\gamma_0 = 1$. Then, let $x^{(t+1)}$ be the solution of

$$\min_x \frac{1}{2\lambda}\|y - Ax\|_2^2 + \frac{1}{2}\|W^{(t)}x\|_2^2, \tag{22}$$

where $\lambda > 0$ is a regularization parameter, and the weight matrix $W^{(t)}$ is defined as $W_i^{(t)} = \mathrm{diag}((\gamma_t + \omega_i^{\frac{4}{p(p-2)}}\|\omega_i^{1/p}x^{(t)}[i]\|_2^2)^{p/4-1/2})$ for the i-th block. Then, $x^{(t+1)}$ can be given explicitly as

$$x^{(t+1)} = (W^{(t)})^{-1}\Big([A(W^{(t)})^{-1}]^T[A(W^{(t)})^{-1}] + \lambda I\Big)^{-1}[A(W^{(t)})^{-1}]^T y.$$

The value of γ is decreased according to the rule $\gamma_{t+1} = 0.1\gamma_t$, and the iteration is continued until $\|x^{(t+1)} - x^{(t)}\|_2 \leq 10^{-5}$ or the iteration times is larger than 2500. In our experiments, the measurement matrix A is generated as an $m \times N$ matrix with entries drawing from i.i.d standard normal distribution. For a generated block sparse or nearly block sparse signal x, the measurements $y = Ax + \sigma z$ with z being the standard Gaussian white noise. We consider several different parameter values to demonstrate the theoretical results. In each experiment, we report the average results over 20 replications. For IRLS, we set $\lambda = 10^{-6}$ in the noise free case ($\sigma = 0$), and $\lambda = 10^{-2}$ in the noisy case ($\sigma = 0.01$). The recovery performance is assessed by the signal to noise ratio (SNR), which is given by

$$\mathrm{SNR} = 20\log_{10}\left(\frac{\|x\|_2}{\|x - x^\sharp\|_2}\right).$$

5.1 Exactly block sparse case

We first consider the case that x is exactly block k-sparse with $k = 20$. In this set of experiments, the signal is of length $N = 400$, which is generated by choosing k blocks uniformly at random, and then choosing the nonzero values from the standard normal distribution for these k blocks.

Figure 2 illustrates the recovery performance of the weighted ℓ_2/ℓ_p minimization problem with $p = 0.5$ and $d = 2$ in both with and without noise cases. It can be observed that when $\alpha = 0.7$, the best recovery is achieved for very small ω whereas a $\omega = 1$ results in the worst SNR. On the other hand, when $\alpha \leq 0.5$, the performance of the recovery algorithms is better for large ω than that for small ω. The case $\omega = 0$ results in the worst SNR. Figure 3 shows the averaged SNR using weighted mixed ℓ_2/ℓ_p minimization for different values of the ratio parameter ρ with fixed $p = 0.5$ and $d = 2$. It shows that when $\alpha = 0.7$, using a larger support estimate results in better reconstruction. However, in both the noise free and noisy measurements cases, the recovery performance is more sensitive to the accuracy α of the block support estimate than the ratio parameter ρ.

In addition, we illustrate the impacts of p and d in Figure 4 for both the noise free and noisy measurements cases. We fix $k = 20$, $\alpha = 0.7$ and $\omega = 0.5$. It is evident that as p decreases, the recovery performance becomes better. And as d increases, more number of measurements are required to obtain good reconstructions. These are largely consistent with the theoretical results in Section 2.

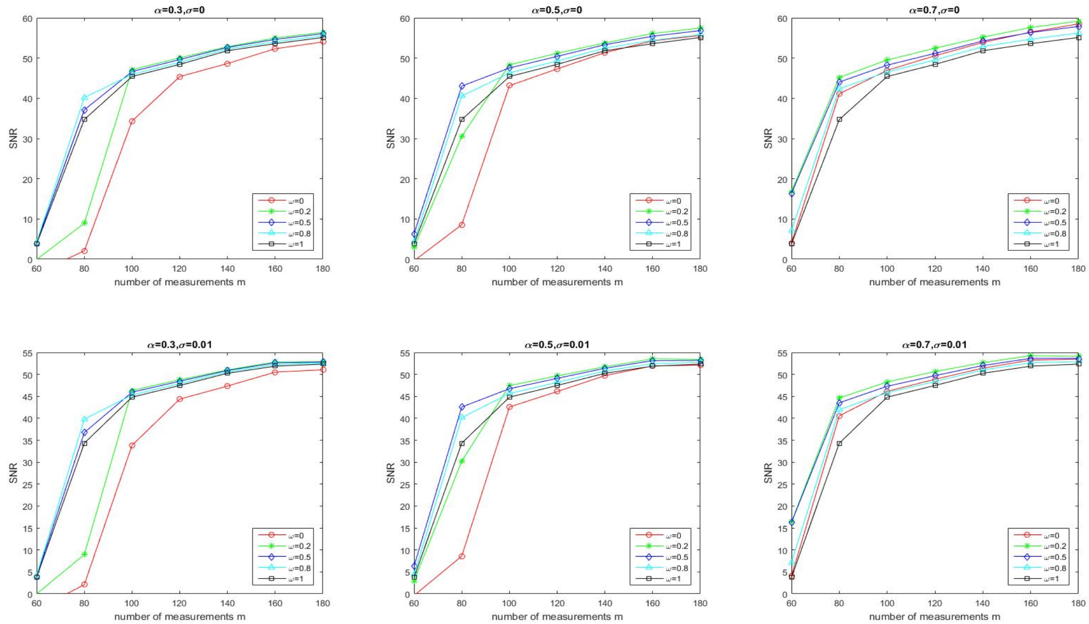

Figure 2: Performance of weighted mixed ℓ_2/ℓ_p recovery with $p = 0.5$ in terms of SNR for exactly block sparse signal x depending on ω with $k = 20, d = 2$, while varying the number of measurements m.

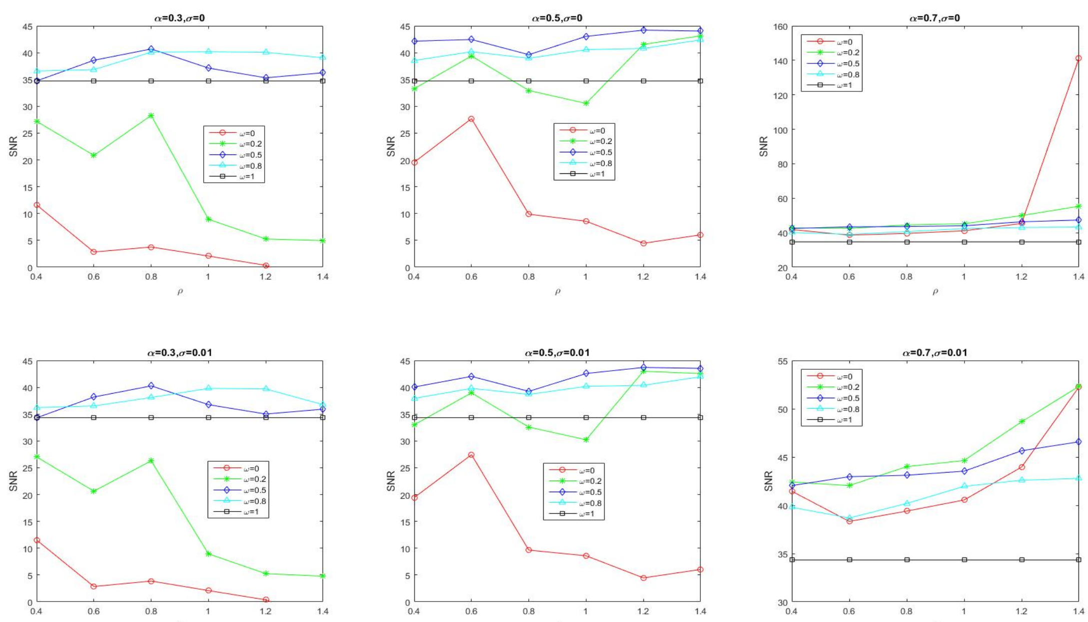

Figure 3: Performance of weighted mixed ℓ_2/ℓ_p recovery with $p = 0.5$ in terms of SNR for exactly block sparse signal x depending on ω with $k = 20, d = 2$, while varying the size of the block support estimate ρ as a proportion of k.

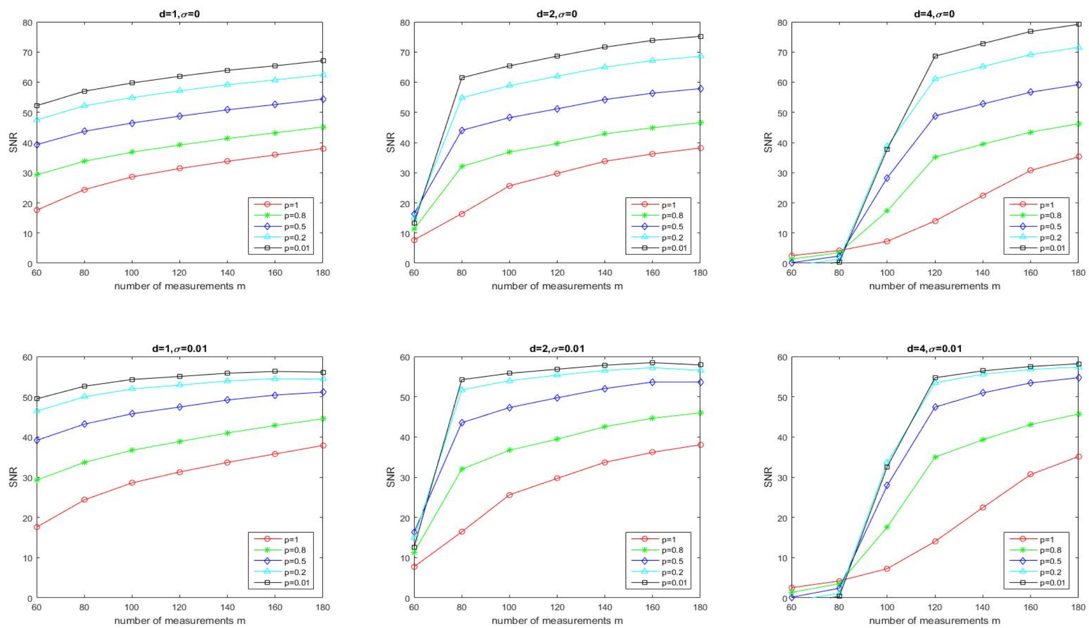

Figure 4: Performance of weighted mixed ℓ_2/ℓ_p recovery in terms of SNR for exactly block sparse signal x depending on p with $k = 20, \alpha = 0.7, \omega = 0.5$, while varying the number of measurements m.

5.2 Nearly block sparse case

Next, we generate x with the ℓ_2 norm of its blocks decay like $i^{-\theta}$ where $i \in \{1, \cdots, n\}$ and $\theta > 1$. In Figure 5, we illustrate the averaged SNR versus the size of the block support estimate ρ as a proportion of k for $\theta = 1.5$. To calculate α, we set $k = 20$, i.e., we are interested in the best 20-term block sparse approximation. It shows that mediate values of ω (e.g., $\omega = 0.2$ or 0.5) results in the best recovery. Generally, larger block support estimate favours better reconstruction result. Finally, we illustrate the impacts of θ and d for different p in Figure 6. For $\theta = 1.5$ we set $k = 20$, while for $\theta = 2$, we consider $k = 10$. It is evident that the recovery performance improves as θ increases, i.e., the block sparsity of signal increases. Moreover, decreasing p and d improves the recovery performance when other parameters are fixed.

6 Conclusion

In this paper, we presented the recovery analysis of weighted mixed $\ell_2/\ell_p\,(0 < p \leq 1)$ minimization by using both block p-RIP and weighted block p-null space property. In addition, we established the relationship between these two conditions. A series of simulations were conducted to illustrate our theoretical results.

There are some interesting issues left for future work. One is that it maybe possible to generalize ℓ_2-constrained minimization problem considered in this paper to the ℓ_q-constrained minimization problem with $2 \leq q \leq \infty$ (see [9, 32]) and with $0 \leq q < 2$ ([36]). Another issue that is not addressed here is to

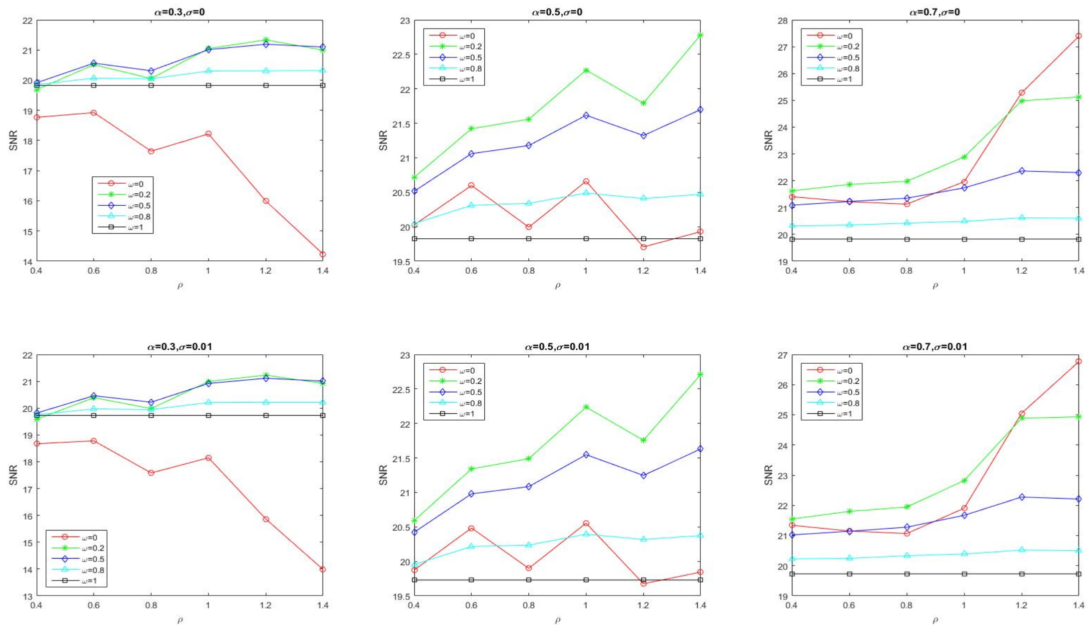

Figure 5: Performance of weighted mixed ℓ_2/ℓ_p recovery in terms of SNR for nearly block sparse signal x with $p = 0.5, m = 80, \theta = 1.5, k = 20, \alpha = 0.7, \omega = 0.5$, while varying the size of the block support estimate ρ as a proportion of k.

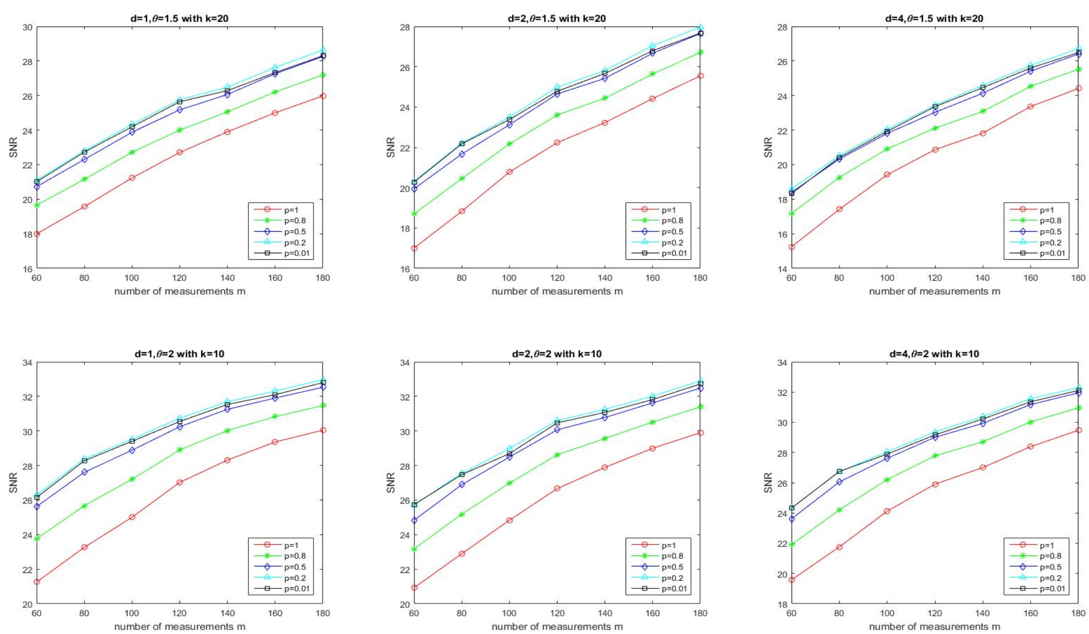

Figure 6: Performance of weighted mixed ℓ_2/ℓ_p recovery in terms of SNR for nearly block sparse signal x for different θ and d depending on p with fixed $\alpha = 0.7, \omega = 0.5, \sigma = 0.01$ for all the cases, while varying the number of measurements m.

obtain the minimum required number of measurements for perfect recovery directly via weighted block p-null space property.

References

[1] Baraniuk, R.G., Cevher, V., Duarte, M.F. and Hegde, C. (2010). Model-based compressive sensing. *IEEE Transactions on Information Theory* **56**(4) 1982–2001.

[2] Blumensath, T. and Davies, M.E. (2009). Sampling theorems for signals from the union of finite-dimensional linear subspaces. *IEEE Transactions on Information Theory* **55**(4) 1872–1882.

[3] Candes, E.J., Romberg, J.K. and Tao, T. (2006). Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics* **59**(8) 1207–1223.

[4] Candes, E.J. and Tao, T. (2005). Decoding by linear programming. *IEEE Transactions on Information Theory* **51**(12) 4203–4215.

[5] Candes, E.J. and Tao, T. (2006). Near-optimal signal recovery from random projections: Universal encoding strategies?. *IEEE Transactions on Information Theory* **52**(12) 5406–5425.

[6] Chartrand, R. (2007). Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signal Processing Letters* **14**(10) 707–710.

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

[8] Chen, J. and Huo, X. (2006). Theoretical results on sparse representations of multiple-measurement vectors. *IEEE Transactions on Signal Processing* **54**(12) 4634–4643.

[9] Dirksen, S., Lecué, G. and Rauhut, H. (2016). On the gap between restricted isometry properties and sparse recovery conditions. *IEEE Transactions on Information Theory*, (DOI: 10.1109/TIT.2016.2570244).

[10] Donoho, D.L. (2006). Compressed sensing. *IEEE Transactions on Information Theory* **52**(4) 1289–1306.

[11] Duarte, M.F. and Eldar, Y.C. (2011). Structured compressed sensing: From theory to applications. *IEEE Transactions on Signal Processing* **59**(9) 4053–4085.

[12] Eldar, Y.C., Kuppinger, P. and Bolcskei, H. (2010). Block-sparse signals: Uncertainty relations and efficient recovery. *IEEE Transactions on Signal Processing* **58**(6) 3042–3054.

[13] Eldar, Y.C. and Kutyniok, G. (2012). Compressed Sensing: Theory and Applications. Cambridge University Press.

[14] Eldar, Y.C. and Mishali, M. (2009). Robust recovery of signals from a structured union of subspaces. *IEEE Transactions on Information Theory* **55**(11) 5302–5316.

[15] Eldar, Y. C. and Rauhut, H. (2010). Average case analysis of multichannel sparse recovery using convex relaxation. *IEEE Transactions on Information Theory* **56**(1) 505–519.

[16] Elhamifar, E. and Vidal, R. (2012). Block-sparse recovery via convex optimization. *IEEE Transactions on Signal Processing* **60**(8) 4094–4107.

[17] Foucart, S. and Lai, M.J. (2009). Sparsest solutions of underdetermined linear systems via ℓ_q-minimization for $0 < q \leq 1$. *Applied and Computational Harmonic Analysis* **26**(3) 395-407.

[18] Foucart, S. and Rauhut, H. (2013). A Mathematical Introduction to Compressive Sensing. New York: Springer-Verlag.

[19] Friedlander, M.P., Mansour, H., Saab, R. and Yilmaz, O. (2012). Recovering compressively sampled signals using partial support information. *IEEE Transactions on Information Theory* **58**(2) 1122–1134.

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

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

[22] Khajehnejad, M.A., Xu, W., Avestimehr, A.S. and Hassibi, B. (2009). Weighted ℓ_1 minimization for sparse recovery with prior information. In *Information Theory, 2009. ISIT 2009. IEEE International Symposium on* (pp. 483-487).

[23] Lv, X., Bi, G. and Wan, C. (2011). The group lasso for stable recovery of block-sparse signal representations. *IEEE Transactions on Signal Processing* **59**(4) 1371–1382.

[24] Majumdar, A. and Ward, R.K. (2010). Compressed sensing of color images. *Signal Processing* **90**(12) 3122–3127.

[25] Mansour, H. and Saab, R. (2017). Recovery analysis for weighted ℓ_1-minimization using the null space property. *Applied and Computational Harmonic Analysis* **43**(1) 23–38.

[26] Mishali, M. and Eldar, Y.C. (2008). Reduce and boost: Recovering arbitrary sets of jointly sparse vectors. *IEEE Transactions on Signal Processing* **56**(10) 4692–4702.

[27] Mishali, M. and Eldar, Y.C. (2009). Blind multiband signal reconstruction: Compressed sensing for analog signals. *IEEE Transactions on Signal Processing* **57**(3) 993–1009.

[28] Needell, D., Saab, R. and Woolf, T. (2017). Weighted ℓ_1-minimization for sparse recovery under arbitrary prior information. *Information and Inference*, iaw023.

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

[30] Parvaresh, F., Vikalo, H., Misra, S. and Hassibi, B. (2008). Recovering sparse signals using sparse measurement matrices in compressed DNA microarrays. *IEEE Journal of Selected Topics in Signal Processing* **2**(3) 275–285.

[31] Rauhut, H. and Ward, R. (2016). Interpolation via weighted ℓ_1 minimization. *Applied and Computational Harmonic Analysis* **40**(2) 321–351.

[32] Saleh, A.A., Alajaji, F., and Chan, W.Y. (2015). Compressed sensing with non-Gaussian noise and partial support information. *IEEE Signal Processing Letters* **22**(10) 1703–1707.

[33] Vaswani, N. and Lu, W. (2010). Modified-CS: Modifying compressive sensing for problems with partially known support. *IEEE Transactions on Signal Processing* **58**(9) 4595-4607.

[34] Wang, Y., Wang, J. and Xu, Z. (2013). On recovery of block-sparse signals via mixed $\ell_2/\ell_q(0 < q \leq 1)$ norm minimization. *EURASIP Journal on Advances in Signal Processing* **2013**(1):76.

[35] Wang, Y., Wang, J. and Xu, Z. (2014). Restricted p-isometry properties of nonconvex block-sparse compressed sensing. *Signal Processing* **104** 188–196.

[36] Wen, F., Liu, P., Liu, Y., Qiu, R.C. and Yu, W. (2017). Robust Sparse Recovery in Impulsive Noise via ℓ_p-ℓ_1 Optimization. *IEEE Transactions on Signal Processing* **65**(1) 105–118.

[37] Xu, Z., Chang, X., Xu, F. and Zhang, H. (2012). $\ell_{1/2}$ regularization: A thresholding representation theory and a fast solver. *IEEE Transactions on Neural Networks and Learning Systems* **23**(7) 1013–1027.

[38] Yu, X. and Baek, S.J. (2013). Sufficient conditions on stable recovery of sparse signals with partial support information. *IEEE Signal Processing Letters* **20**(5) 539–542.

[39] Yuan, M. and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **68(1)** 49–67.

[40] Zeinalkhani, Z. and Banihashemi, A.H. (2015). Iterative reweighted ℓ_2/ℓ_1 recovery algorithms for compressed sensing of block sparse signals. *IEEE Transactions on Signal Processing* **63**(17) 4516–4531.