IET Signal Processing

Research Article

Block-sparse compressed sensing with partially known signal support via non-convex minimisation

ISSN 1751-9675
Received on 6th October 2015
Revised on 17th March 2016
Accepted on 1st April 2016
doi: 10.1049/iet-spr.2015.0425
www.ietdl.org

Shiying He[1,2], Yao Wang[1,2] ✉, Jianjun Wang[3], Zongben Xu[1]

[1]*School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, People's Republic of China*
[2]*State Key laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Science, Shenyang, People's Republic of China*
[3]*School of Mathematics and Statistics, Southwest University, Beibei, Chongqing, People's Republic of China*
✉ *E-mail: yao.s.wang@gmail.com*

Abstract: The mixed l_2/l_p $(0 < p \le 1)$ norm minimisation method with partially known support for recovering block-sparse signals is studied. The authors mainly extend this work on block-sparse compressed sensing by incorporating some known part of the block support information as a priori and establish sufficient restricted p-isometry property (p-RIP) conditions for exact and robust recovery. The authors' theoretical results show it is possible to recover the block-sparse signals via l_2/l_p minimisation from reduced number of measurements by applying the partially known support. The authors also derive a lower bound on necessary random Gaussian measurements for the p-RIP conditions to hold with high possibility. Finally, a series of numerical experiments are carried out to illustrate that fewer measurements with smaller p are needed to reconstruct the signal.

1 Introduction

The aim of compressed sensing (CS) is to recover an unknown signal by exploiting its key property, i.e. the sparsity. Mathematically, denoting by x an unknown signal which has only a few non-zero entries, we observe y through a measurement matrix $\boldsymbol{\Phi}$ according to $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{e}$. Here, $\boldsymbol{\Phi}$ is an $M \times N$ matrix with $M \ll N$ and $\boldsymbol{e}$ indicates noise in the measurement satisfying $\|\boldsymbol{e}\| < \epsilon$. Compressed sensing [1, 2] is concerned with the recovery of an unknown signal from this undetermined system of linear equations by using sparsity prior of the desired signal which can restrict the set of possible solutions. In order to obtain better recovery performance, recently several new trends in this field are formed to explore the structures of signals and further incorporate some prior information from compressive measurements.

Considering the structure of the sparse signal, one of those techniques is to assume the unknown sparse signal $\boldsymbol{x}$ has the block structure which means the support elements lie in clusters or some fixed blocks. In this paper, we consider the block-sparse signal $\boldsymbol{x} \in \mathbb{R}^N$ over the index set $\mathcal{I} = \{d_1, d_2, \ldots, d_m\}$, which can be indicated as the following form

$$\boldsymbol{x} = [\underbrace{\boldsymbol{x}_1 \cdots \boldsymbol{x}_{d_1}}_{x[1]} \underbrace{\boldsymbol{x}_{d_1+1} \cdots \boldsymbol{x}_{d_1+d_2}}_{x[2]} \cdots \underbrace{\boldsymbol{x}_{N-d_{m-1}+1} \cdots \boldsymbol{x}_N}_{x[m]}]^{\mathrm{T}},$$

where $\boldsymbol{x}[i]$ denotes the ith block of $\boldsymbol{x}$ and $N = \sum_{i=1}^{m} d_i$. Accordingly, the term block K-sparse refers to that there exist at most K blocks of $\boldsymbol{x}$ with non-zero Euclidean norm. In fact, there are many examples related to the block-sparse structure in practice, such as the electrocardiography (ECG) signal, which has a clear block-sparse structure and the entries in the same block are highly related in amplitude. Block sparsity also plays an important role in the study of gene expression [3], multi-band signals [4], colour imaging [5] and more.

In the conventional CS framework, due to the non-convexity of l_0 norm, l_1 minimisation is a commonly used strategy for sparse signal recovery. To exploit the structure information that the non-zero entries of the original signal lies in consecutive positions, an approach of mixed l_2/l_1 norm minimisation has been put forward based on the basic pursuit (BP) method used in traditional CS. Specifically, it has been shown in [6] that one can recover the block-sparse signal $\boldsymbol{x}$ with small observation error by solving the following convex optimisation problem

$$\begin{aligned} \min \quad & \|\boldsymbol{x}\|_{2,1} \\ \text{s.t} \quad & \|\boldsymbol{y} - \boldsymbol{\Phi x}\|_2 \le \epsilon \end{aligned} \tag{1}$$

Here, $\|\boldsymbol{x}\|_{2,1} = \sum_{j=1}^{J} \|\boldsymbol{x}[j]\|_2$ where $\|\boldsymbol{x}[j]\|_2$ denotes the l_2 norm of signal with all entries assigned to the jth group, the noise measurements is denoted by $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{e}$ where $\boldsymbol{e}$ is the error and the measurement matrix is presumed to satisfy some properties ensuring the successful reconstruction.

An efficient algorithm based on l_2/l_1 minimisation are studied in [6, 7] where it is solved by recasting it as a second-order cone-programming problem. Besides, greedy algorithms have also been proposed to deal with the problem as presented in [8] where a block orthogonal matching pursuit algorithm is extended from the orthogonal matching pursuit algorithm [9] for recovering block-sparse signals. Non-convex minimisation models, such as l_p minimisation for conventional sparse signals and mixed l_2/l_p method for the block-sparse case, are put forward which help to improve the performance of the recovery methods. In [10], the authors present a sufficient condition for exact and robust recovery via l_2/l_p minimisation and propose an iteratively reweighted least square algorithm for solving the resulting non-convex problem.

Another variation of the studies on CS is to incorporate the compressive measurements with some part of the support of the signal known as a priori. Driven by the real-time dynamic MRI reconstruction and video compression problem, it is very useful to consider this pattern of sparse signal recovery. Vaswani and Lu [11] and Lu and Vaswani [12] have showed that it is possible to recover sparse signals with reduced number of linear measurements when the partially known support is given.

Considering the problem of reconstruction of a K-sparse signal $\boldsymbol{x}$ with partial known support from measurements $\boldsymbol{y}$, the support of $\boldsymbol{x}$, denoted as S, can be represented as $S = T \cup \Delta$, where T is the known part with size $|T|$ and Δ is the unknown part with size $|\Delta|$. Furthermore, with partially known support T, the candidates for the K-sparse signal $\boldsymbol{x}$ are restricted in a signal space smaller than that in traditional CS. Since the signal $\boldsymbol{x}$ is known to be non-zero at some locations, CS with partially known support allows us to minimise the number of non-zero elements at other locations, i.e outside the support T, when solving $\boldsymbol{y} = \boldsymbol{\Phi x}$ for a sparse solution. The known support is denoted as T with size —T—, and the unknown part $T^c = \text{support}(\boldsymbol{x})\backslash T$. This procedure can be formulated as the following minimisation problem

$$\begin{aligned} \min \quad & \|\boldsymbol{x}_{T^c}\|_0 \\ \text{s.t} \quad & \|\boldsymbol{y} - \boldsymbol{\Phi x}\|_2 \le \epsilon \end{aligned}$$

Taking computational complexity into consideration, one usually consider the following model instead

$$\begin{aligned} \min \quad & \|\boldsymbol{x}_{T^c}\|_1 \\ \text{s.t} \quad & \|\boldsymbol{y} - \boldsymbol{\Phi x}\|_2 \le \epsilon \end{aligned}$$

where $\boldsymbol{x}_{T^c}$ denotes the signal outside the known support. Several works presented in [13–16] have studied this case in detail.

2 Block-sparse recovery with partially known signal support via non-convex minimisation

Several studies on non-convex CS have shown that compared with l_1 minimisation, l_p method can lead to successful recovery of the signal with reduced number of measurements [17–19]. This strategy has also been extended to the block-sparse case via the l_2/l_p minimisation. To investigate the theoretical performance of l_2/l_p minimisation, we first present theories on the measurement matrix $\boldsymbol{\Phi}$ satisfying the restricted p-isometry property (block p-RIP) condition.

Definition 1 ([20]): An $M \times N$ measurement matrix $\boldsymbol{\Phi}$ is said to have the block p-RIP over $\mathcal{I} = \{d_1, d_2, \ldots, d_n\}$ of order K with positive constant $\delta_{K|\mathcal{I}}$ if

$$(1 - \delta'_{K|\mathcal{I}})\|\boldsymbol{x}\|_2^p \le \|\boldsymbol{\Phi x}\|_p^p \le (1 + \delta'_{K|\mathcal{I}})\|\boldsymbol{x}\|_2^p, \tag{2}$$

for all $\boldsymbol{x} \in \mathbb{R}^N$ that is block K-sparse over block index set $\mathcal{I}$.

As we can see from the definitions described above, block K-sparse over $\mathcal{I}$ is equivalent to k-sparse in the conventional situation where k is the sum of the K largest entries in $\mathcal{I}$. For the rest of this paper, we use δ_K to represent the modified block RIP constant (p-RIC) instead of $\delta'_{K|\mathcal{I}}$ without specifically mentioning.

In the rest of this section, we will show that it is possible to recover block-sparse signals using the following constrained mixed l_2/l_p minimisation with partially known information incorporated. To be more clear, assume that we have known some of the exact block positions from K blocks that have the largest l_2 norm and we derive the following problem

$$\begin{aligned} \min \quad & \|\boldsymbol{x}_{T^c}\|_{2,p} \\ \text{s.t} \quad & \|\boldsymbol{y} - \boldsymbol{\Phi x}\|_2 \le \epsilon \end{aligned} \tag{3}$$

where T^c is the complement of the partially known set of the signal, $\|\boldsymbol{x}\|_{2,p} = (\sum_{i=1}^{k} \|\boldsymbol{x}[i]\|_2^p)^{1/p}$ $(0 < p < 1)$ and ϵ controls the noise level.

First, we present the sufficient recovery conditions theoretically based on the p-RIP of the measurement matrix.

Theorem 2: Suppose $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{e}$ is the observed measurements of a block-sparse signal $\boldsymbol{x}$, with the noise level $\|\boldsymbol{e}\|_2 \le \epsilon$, the block p-RIP constant δ_K with respect to the measurement matrix $\boldsymbol{\Phi}$ satisfies

$$b\delta_{s+(a+1)K} + \delta_{aK} < b - 1 \tag{4}$$

for $b > 1$, $a = b^{2/(2-p)}$, which is rounded up so that aK is an integer, then the solution $\boldsymbol{x}^*$ to (3) obeys

$$\|\boldsymbol{x}^* - \boldsymbol{x}\|_2 < C\frac{\|\boldsymbol{r} - \boldsymbol{r_K}\|_{2,p}}{K^{(1/p)-(1/2)}} + D\epsilon \tag{5}$$

where $\boldsymbol{x} \in \mathbb{R}^N$ is a given nearly block K-sparse signal, T is the partially known block set with size $|T| = s$, $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{x}_T$ and $\boldsymbol{r_K}$ is the best K-block term approximation to $\boldsymbol{r}$. The constants C and D are determined explicitly as

$$C = \frac{2^{(2/p)-1}[(1 - \delta_{s+(a+1)K})^{(1/p)} + (1 + \delta_{aK})^{(1/p)}]}{(b - b\delta_{s+(a+1)K} - 1 - \delta_{aK})^{(1/p)}} \tag{6}$$

$$D = \frac{2^{(1/p)}M^{(1/p)-(1/2)}(1 + b^{(1/p)})}{(b - b\delta_{s+(a+1)K} - 1 - \delta_{aK})^{(1/p)}}. \tag{7}$$

Proof: Let T be the known block support set, T_0 is the locations of K blocks with K-largest l_2 norm of the unknown of x with $T \cap T_0 = \emptyset$. Furthermore, we suppose $(T \cup T_0)^c$ is divided into disjoint sets $T_1, T_2, \ldots, T_J$ and each set contains aK blocks except possibly the last T_J. Considering that $\boldsymbol{x}^*$ is the solution to (3) and $h = \boldsymbol{x}^* - \boldsymbol{x}$, our aim is to bound $\|\boldsymbol{h}\|_2$ based on the triangle inequality for $\|\cdot\|_2$.

(i) Since $\|\boldsymbol{x}^*\|_p = \|\boldsymbol{x} + \boldsymbol{h}\|_p$ where $\boldsymbol{x}^*$ is the minimum, we have

$$\begin{aligned} \|\boldsymbol{x}_{T^c}\|_{2,p}^p &\ge \|\boldsymbol{x}_{T^c} + \boldsymbol{h}_{T^c}\|_{2,p}^p \\ &= \|\boldsymbol{x}_{T_0} + \boldsymbol{x}_{T_0}\|_{2,p}^p + \|\boldsymbol{x}_{\bar{T}_0^c} + \boldsymbol{h}_{\bar{T}_0^c}\|_{2,p}^p \\ &\ge \|\boldsymbol{x}_{T_0}\|_{2,p}^p - \|\boldsymbol{h}_{T_0}\|_{2,p}^p - \|\boldsymbol{x}_{\bar{T}_0^c}\|_{2,p}^p + \|\boldsymbol{h}_{\bar{T}_0^c}\|_{2,p}^p. \end{aligned}$$

then

$$\|\boldsymbol{h}_{\bar{T}_0^c}\|_{2,p}^p \le \|\boldsymbol{h}_{T_0}\|_{2,p}^p + 2\|\boldsymbol{x}_{\bar{T}_0^c}\|_{2,p}^p \tag{8}$$

Arrange the elements of $\bar{T}_0^c$ in order of decreasing magnitude of $|\boldsymbol{h}|$, which means the block indices are in the form of $\|\boldsymbol{h}_{T_j}[1]\|_2 \ge \|\boldsymbol{h}_{T_j}[2]\|_2 \ge \ldots \ge \|\boldsymbol{h}_{T_j}[aK]\|_2 \ge \|\boldsymbol{h}_{T_{j+1}}[1]\|_2 \ge \ldots$ for any $j \ge 1$. Note that for each i, k, $1 \le i, k \le aK$, we have

$$\|\boldsymbol{h}_{T_j}[i]\|_2 \le \|\boldsymbol{h}_{T_{j-1}}[k]\|_2.$$

Thus, $\|\boldsymbol{h}_{T_j}[i]\|_2^p \le (1/aK)\|\boldsymbol{h}_{T_{j-1}}\|_{2,p}^p$ and $\|\boldsymbol{h}_{T_j}[i]\|_2^2 \le (1/aK)^{(2/p)}\|\boldsymbol{h}_{T_{j-1}}\|_{2,p}^2$holds.

For $\|\boldsymbol{h}_{T_j}\|_{2,2}^2 = \sum_i \|\boldsymbol{h}_{T_j}[i]\|_2^2$, we have

$$\|\boldsymbol{h}_{T_j}\|_{2,2}^p \le (aK)^{(p/2)-1}\|\boldsymbol{h}_{T_{j-1}}\|_{2,p}^p \tag{9}$$

Then, we obtain the following

$$\begin{aligned} \|\boldsymbol{h}_{\bar{T}_{01}^c}\|_{2,2}^p &\le \sum_{j\ge 2}\|\boldsymbol{h}_{T_j}\|_{2,2}^p \\ &\le (aK)^{(p/2)-1}\sum_{j\ge 1}\|\boldsymbol{h}_{T_j}\|_{2,p}^p = (aK)^{(p/2)-1}\|\boldsymbol{h}_{\bar{T}_0^c}\|_{2,p}^p \\ &\le (aK)^{(p/2)-1}(\|\boldsymbol{h}_{T_0}\|_{2,p}^p + 2\|\boldsymbol{x}_{\bar{T}_0^c}\|_{2,p}^p) \end{aligned} \tag{10}$$

For convenience, we denote $\bar{T}_0 = T \cup T_0$ and $\bar{T}_{01} = T \cup T_0 \cup T_1$.

(ii) By the theory and properties of block p-RIP, we decompose $\boldsymbol{\Phi h}$

$$
\begin{aligned}
\|\boldsymbol{\Phi h}\|_p^p &= \|\boldsymbol{\Phi h}_{\bar{T}_{01}} + \boldsymbol{\Phi h}_{\bar{T}_{01}^c}\|_p^p \\
&\geq \|\boldsymbol{\Phi h}_{\bar{T}_{01}}\|_p^p - \|\boldsymbol{\Phi h}_{\bar{T}_{01}^c}\|_p^p \\
&\geq (1-\delta_{s+(a+1)K})\|\boldsymbol{h}_{\bar{T}_{01}}\|_2^p - (1+\delta_{aK})\sum_{i\geq 2}\|\boldsymbol{h}_{T_i}\|_2^p \\
&\geq (1-\delta_{s+(a+1)K})\|\boldsymbol{h}_{\bar{T}_{01}}\|_2^p - (1+\delta_{aK})(aK)^{(p/2)-1} \\
&\quad \times (2\|\boldsymbol{x}_{\bar{T}_0^c}\|_{2,p}^p + \|\boldsymbol{h}_{T_0}\|_{2,p}^p) \\
&\geq (1-\delta_{s+(a+1)K} - (1+\delta_{aK})a^{(p/2)-1})\|\boldsymbol{h}_{\bar{T}_{01}}\|_2^p \\
&\quad - 2(1+\delta_{aK})(aK)^{(p/2)-1}\|\boldsymbol{x}_{\bar{T}_0^c}\|_{2,p}^p
\end{aligned} \tag{11}
$$

where the last inequality holds from the fact that

$$
\|\boldsymbol{h}_{T_0}\|_{2,p}^p \leq K^{1-(p/2)}\|\boldsymbol{h}_{T_0}\|_2^p \leq K^{1-(p/2)}\|\boldsymbol{h}_{\bar{T}_{01}}\|_2^p
$$

Note the facts that

$$
\|\boldsymbol{\Phi h}\|_2 \leq \|\boldsymbol{\Phi x}^* - \boldsymbol{y}\|_2 + \|\boldsymbol{\Phi x} - \boldsymbol{y}\|_2 \leq 2\epsilon,
$$

and that

$$
\begin{aligned}
\|\boldsymbol{\Phi h}\|_p^p &\leq \left(\sum_{i=1}^{M}(|(\boldsymbol{\Phi h})_i|^p)^{(2/p)}\right)^{(p/2)}\left(\sum_{i=1}^{M}1\right)^{1-(p/2)} \\
&= M^{1-(p/2)}\|\boldsymbol{\Phi h}\|_2^p,
\end{aligned}
$$

according to Hölder's inequality, then we have

$$
\|\boldsymbol{\Phi h}\|_p^p \leq M^{1-(p/2)}\|\boldsymbol{\Phi h}\|_2^p \leq M^{1-(p/2)}(2\epsilon)^p. \tag{12}
$$

Combining (11) and (12), we can obtain the following inequality by setting $a = b^{2/(2-q)}$ if $b\delta_{s+(a+1)K} + \delta_{aK} < b - 1$ (see (13))

Denoting $\boldsymbol{r}$ as $\boldsymbol{x} - \boldsymbol{x}_T$ and $\boldsymbol{r_K}$ the best K-block term approximation to $\boldsymbol{r}$, we have obtained the conclusion of the theorem. □

Corollary 3: Let x be an exact block-sparse signal and suppose the partially known set is given, then the solution $\boldsymbol{x}^*$ obeys

$$
\|\boldsymbol{x} - \boldsymbol{x}^*\|_2 < \boldsymbol{D\epsilon}
$$

where D is given in (7); Let $\boldsymbol{y} = \boldsymbol{\Phi x}$ be measurements of a block-sparse signal $\boldsymbol{x}$ and the p-RIP constant δ_K with respect to the measurement matrix $\boldsymbol{\Phi}$ satisfies (4),then the solution $\boldsymbol{x}^*$ to (3) obeys

$$
\|\boldsymbol{x} - \boldsymbol{x}^*\|_2 < C\frac{\|\boldsymbol{r} - \boldsymbol{r}_K\|_{2,p}}{k^{(1/p)-(1/2)}}
$$

and C is given in (6).

Remark 4: Theorem 2 and the corollary above offer necessary conditions for recovering block-sparse signals in the method of l_2/l_p minimisation with partially known support information. The constants C and D determine the upper bounds to estimate the recovery error.

Definition 5: Given a constant $c > 0$, we define δ_K^c, which leads to a slightly stronger version as the smallest number such that

$$
(1-\delta_K^c)\|\boldsymbol{x}\|_2^p \leq \frac{1}{c}\|\boldsymbol{\Phi x}\|_p^p \leq (1+\delta_K^c)\|\boldsymbol{x}\|_2^p \tag{14}
$$

holds. Then, we have the exact recovery condition as

$$
b\delta_{s+(a+1)K}^c + \delta_{aK}^c < b - 1 \tag{15}
$$

since the isometry constants are not scale invariant, while the sufficient condition is, which is analysed similarly to [20].

3 p-RIP for random, Gaussian matrices

In this part, we will derive a lower bound on the necessary number of Gaussian measurements for the p-RIP (14) to hold with high probability. Here, $\boldsymbol{\Phi}$ is an $M \times N$ matrix whose elements are i.i.d. random variables distributed normally with mean zero and variance σ^2, where $M < N$. For a given p, set $\mu_p := \sigma^p 2^{(p/2)}\Gamma(p + (1/2))/\sqrt{\pi}$ as in [21].

To prove Theorem 7 we need the following Lemma 6 [21].

Lemma 6: Let $0 < p \leq 1$ and Ψ be an $M \times L$ submatrix of $\boldsymbol{\Phi}$. Suppose $\delta > 0$. Choose $\eta,\ \tau > 0$ such that $(\eta + \tau^p/1 - \tau^p) \leq \delta$. Then

$$
(1-\delta)M\mu_p\|\boldsymbol{x}\|_2^p \leq \|\Psi\boldsymbol{x}\|_p^p \leq (1+\delta)M\mu_p\|\boldsymbol{x}\|_2^p \tag{16}
$$

holds uniformly for $\boldsymbol{x} \in \mathbb{R}^L$ with probability exceeding $1 - 2(1 + (2/\tau))^L e^{-\left((\eta^2 M)/(2pc_p^2)\right)}$, where

$$
c_p = (31/40)^{1/4}\left[1.13 + \sqrt{p}\left(\frac{\Gamma((p+1/2))}{\sqrt{\pi}}\right)^{-(1/p)}\right]. \tag{17}
$$

$$
\begin{aligned}
\|\boldsymbol{h}\|_2 &\leq \|\boldsymbol{h}_{\bar{T}_{01}}\|_2 + \|\boldsymbol{h}_{\bar{T}_{01}^c}\|_2 \\
&\leq 2^{(1/p)-1}\left(\frac{2^{(1/p)}(1+\delta_{aK})^{(1/p)}K^{(1/2)-(1/p)}\|\boldsymbol{x}_{\bar{T}_0^c}\|_{2,p}}{(b - b\delta_{s+(a+1)K} - 1 - \delta_{aK})^{(1/p)}} + \frac{2b^{(1/p)}M^{(1/p)-(1/2)}\boldsymbol{\epsilon}}{(b - b\delta_{s+(a+1)K} - 1 - \delta_{aK})^{(1/p)}}\right) \\
&\quad + 2^{(1/p)-1}\left(\frac{2^{(1/p)}(1-\delta_{s+(a+1)K})^{(1/p)}K^{(1/2)-(1/p)}\|\boldsymbol{x}_{\bar{T}_0^c}\|_{2,p}}{(b - b\delta_{s+(a+1)K} - 1 - \delta_{aK})^{(1/p)}} + \frac{2M^{(1/p)-(1/2)}\boldsymbol{\epsilon}}{(b - b\delta_{s+(a+1)k} - 1 - \delta_{ak})^{(1/p)}}\right) \\
&= \frac{2^{(2/p)-1}[(1-\delta_{s+(a+1)K})^{(1/p)} + (1+\delta_{aK})^{(1/p)}]}{(b - b\delta_{s+(a+1)K} - 1 - \delta_{aK})^{(1/p)}}\frac{\|\boldsymbol{x}_{\bar{T}_0^c}\|_{2,q}}{k^{(1/p)-(1/2)}} \\
&\quad + \frac{2^{(1/p)}M^{(1/p)-(1/2)}(1+b^{(1/p)})}{(b - b\delta_{s+(a+1)K} - 1 - \delta_{aK})^{(1/p)}}\boldsymbol{\epsilon} \\
&= C\frac{\|\boldsymbol{x}_{\bar{T}_0^c}\|_{2,p}}{k^{(1/p)-(1/2)}} + D\boldsymbol{\epsilon}.
\end{aligned} \tag{13}
$$

Theorem 7: Suppose $\Phi = (\phi_{i,j})_{M\times N}$ is an $M \times N$ $(M < N)$ matrix with i.i.d Gaussian random entries. Specifically, the elements are with $\phi_{i,j} \sim \mathcal{N}(0, \sigma^2)$. Then there exist constants $C_1(p)$ and $C_2(p)$ such that whenever $0 < p \leq 1$ and

$$M \geq C_1(p)(K + [b^{2/(2-p)}](K - s)) \cdot d + pC_2(p)(K + [b^{2/(2-p)}](K - s)) \ln\left(\frac{m}{K + [b^{2/(2-p)}](K - s)}\right) \tag{18}$$

the following is true with probability exceeding $1 - 2e^{-((\eta^2 M)/(4pc_p^2))}$: For any block K-sparse signal $x \in \mathbb{R}^N$ with known support size $|T| = s$ over $\mathcal{I} = \{d_1, d_2, \ldots, d_m\}$ and $d = \max_{i\in[1,m]} d_i$, x can be exactly recovered (when the measurements are obtained in the way of $y = \Phi x$).

Proof: The proof is similar to [20, 22]. By (14), we set $c = M\mu_p$. We will determine what M must to be for $\delta^c_{s+(a+1)(K-s)} = \delta^c_{K+(K-s)a} < (b - 1/b + 1)$ to hold with high probability, i.e the failure probability is at most $2e^{-(\eta^2 M/4pc_p^2)}$. Leaving $b > 1$ undetermined, we set $L = (s + (a+1)(K-s))d = (K + [b^{2/(2-p)}](K-s))d$. Choose $\eta = (r(b-1)/b+1)$ for $r \in (0, 1)$ and $\tau^p = ((1-r)(b-1)/2b) < 1$ to satisfy

$$\frac{\eta + \tau^p}{1 - \tau^p} \leq \delta_{K+(K-s)a}{}^c \leq \frac{b-1}{b+1}.$$

From Lemma 6 and the union bound, we can conclude that Ψ fails to satisfy (14) with probability lower than

$$\binom{m}{k + [b^{2/(2-p)}](k-s)} 2\left(1 + \frac{2}{\tau}\right)^L e^{-\eta^2 M/2pc_p^2},$$

where c_p is given in (17). Thus, the upper bound of this quantity is $2e^{-(\eta^2 M/4pc_p^2)}$. Then the following holds,

$$\begin{aligned} M \geq& \frac{4pc_p^2}{\eta^2}\Bigg[(k + [b^{2/(2-p)}](k-s))d(\ln 3 - \ln\tau) \\ &+ (K + [b^{2/(2-p)}](K-s))\left(\ln\frac{m}{K + [b^{2/(2-p)}](K-s)} + 1\right)\Bigg] \\ =& \frac{4c_p^2(b+1)^2}{r^2(b-1)^2}\left(K + [b^{2/(2-p)}](K-s))d\left(p\ln 3 + \ln\frac{2b}{(1-r)(b-1)}\right)\right. \\ &\left. + p\frac{4c_p^2(b+1)^2}{r^2(b-1)^2}(K + [b^{2/(2-p)}](K-s)\right) \\ &\times \left(ln\frac{m}{K + [b^{2/(2-p)}](K-s)} + 1\right) \end{aligned}$$

We can substitute any $b > 1$ and $r \in (0, 1)$ since these are free parameters, which can be chosen independently for each p. Therefore, the lower bound of the measurement size M can be expressed in the form

$$M \geq C_1(p)(K + [b^{2/(2-p)}](K - s)) \cdot d + pC_2(p)(K + [b^{2/(2-p)}](K - s)) \ln\left(\frac{m}{K + [b^{2/(2-p)}](K - s)}\right) \tag{19}$$

to yield the p-RIP condition (14) with probability exceeding $1 - 2e^{-(\eta^2 M/4pc_p^2)}$. □

Remark 8: Theorem 7 implies when decreasing p and incorporating known support, that is increasing s, allow fewer measurements to be sufficient for (3) to successfully recover block-sparse signals. Thus, in order to satisfy the block p-RIP condition, roughly $M \simeq (K + [b^{2/(2-p)}](K-s))\ln(m/(K + [b^{2/(2-p)}](K-s)))$ measurements are needed. To be specific, for a given $p \in (0, 1]$, there exist $C_1(p)$ and $C_2(p)$ which are finite constants, the second term of (18) predominant and it vanishes when $p \to 0$. When $p = 1$, (18) has the form of

$$M \geq C_1(1)(K + b^2(K-s)) \cdot d + C_2(1)(K + b^2(K - s)) \ln\left(\frac{m}{K + b^2(K-s)}\right).$$

Note the $(K–s)$ which refers to the information of the known support. Thus, it leads to fewer measurements than traditional l_2/l_1 minimisation [7].

Remark 9: The proof is conducted in a general case that the block-sparse signal has an uneven size blocks. We can easily adapt the Theorem 7 to a special case in which $d_i = d$ as what we do in the following numerical sections.

4 Numerical experiments

In this section, we conduct several numerical experiments to illustrate the effectiveness of (3) as well as to validate the theoretical results for both exactly block-sparse and block-compressible cases. Since problem (3) is concerned with l_2/l_p minimisation with partially known support and $p \in (0, 1)$, it has several local minimas on the feasible set due to the non-convexity. Here, we modify the iteratively reweighted least squared approach for block-sparse signal recovery (block-IRLS) [20, 23] by incorporating partial knowledge about the support.

The modified block-IRLS is based on the rewritten form of (3)

$$\min_x \|x\|^{2,p}_{\epsilon,p,w} + \frac{1}{2\lambda}\|\Phi x - y\|_2^2 \tag{20}$$

where $\|x\|^{2,p}_{\epsilon,p,w} = \sum_{i=1}^m w_i^{(p/2)}(\epsilon^2 + \|x[i]\|_2^2)^{(p/2)}$,

$$w_i = \begin{cases} 1, & i \in T^c \\ w, & i \in T \end{cases},$$

$0 < w < 1$ and λ is the regularisation parameter and ϵ is a small enough smooth parameter. We begin with $x^{(0)}$ satisfying $\Phi x^{(0)} = y$ and set $\epsilon_0 = 1$. Then, we update $x^{(t+1)}$ via the following equation

$$(\lambda(W^{(t)})^2 + \Phi^T\Phi)x^{(t+1)} = \Phi^T y \tag{21}$$

where the weighting matrix $W^{(t)}$ is defined as

$$W_i^{(t)} = \text{diag}(p^{(1/2)}(w_i(\epsilon_t^2 + \|x[i]\|_2^2))^{(p/4)-(1/2)})$$

for the ith block in the tth iteration. The value of ϵ is decreased in the pattern of $\epsilon_{t+1} = 0.99\epsilon_t$ and the iteration continues until ϵ becomes very small.

The rest of this section presents the simulation results related to the former theorems and the performance of the partially known support algorithm. We create synthetic block-sparse signals with non-zero entries from Gaussian distribution randomly, setting the length of the signal to $N = 512$ and the even block size $d = 4$. The signals are sampled using measurement matrix Φ that has i.i.d. entries drawn from a standard Gaussian distribution with normalised columns. We average 50 repetitions of each experiment varying the amplitude of the signals. The influence of the number of measurements and the support size on the recovery result is

IET Signal Process., 2016, Vol. 10, Iss. 7, pp. 717–723

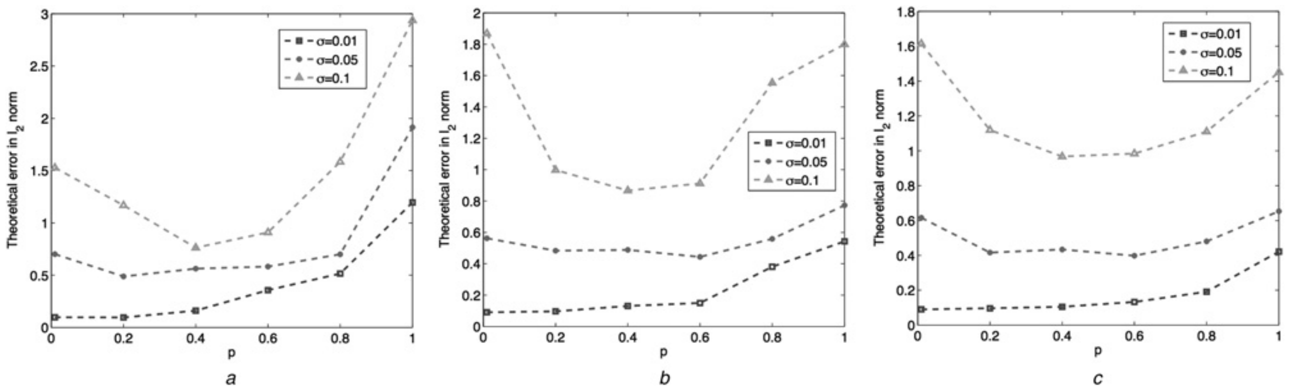

Fig. 1 *Theoretical error against p via mixed l_2/l_p minimisation with partially known support for different cases. Different noise levels are considered in the observation process for σ = 0.01, 0.05 and 0.1*

a $s = 4$, $M = 100$
b $s = 8$, $M = 100$
c $s = 8$, $M = 120$

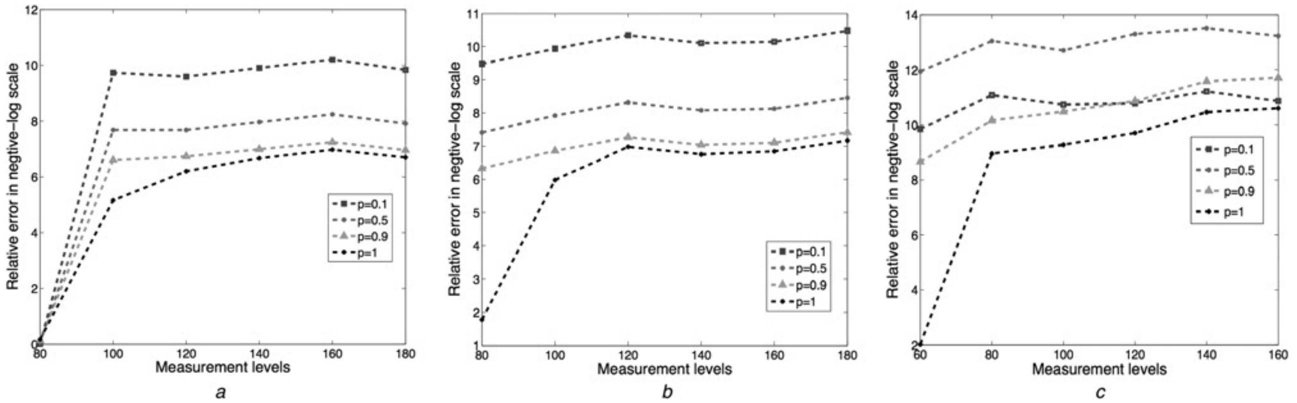

Fig. 2 *Performance of l_2/l_p minimisation with partially known support in terms of relative error in negative-log scale for strict block-sparse signal, varying the number of measurements for*

a $s = 2$
b $s = 6$
c $s = 10$

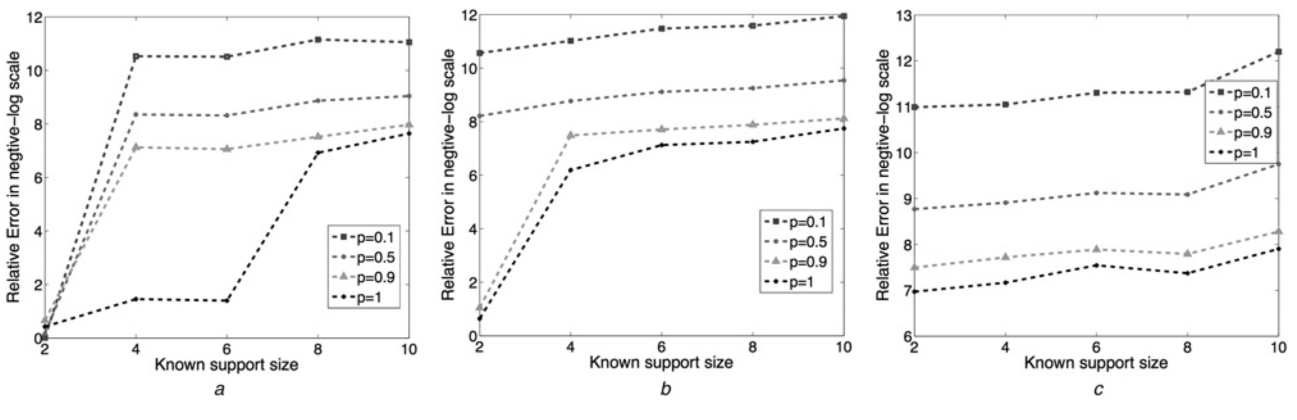

Fig. 3 *Performance of l_2/l_p minimisation with partially known support in terms of relative error in negative-log scale for strict block-sparse signal, varying the support size s for*

a $M = 80$
b $M = 100$
c $M = 120$

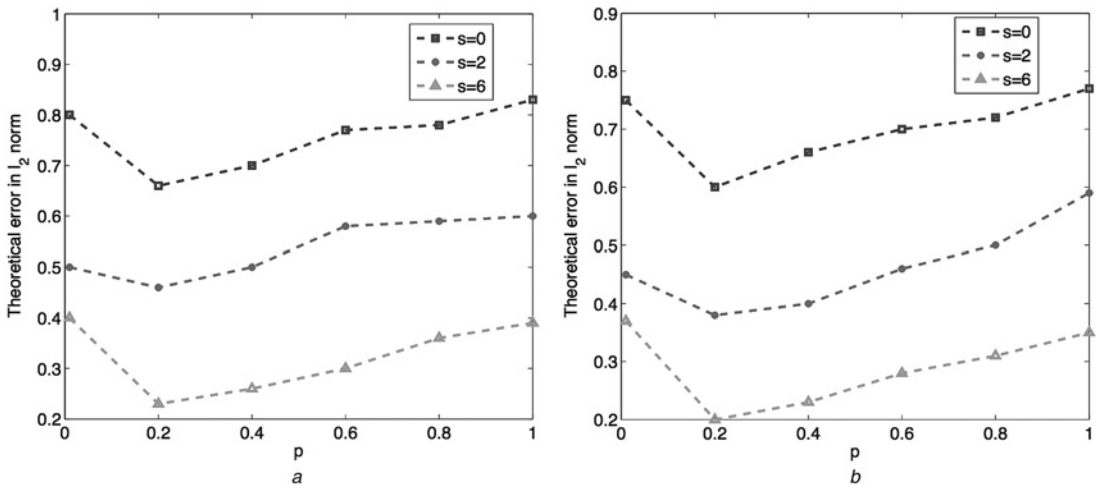

Fig. 4 *Theoretical error against p via l_2/l_p minimisation with partially known support for different block-compressible signals. The number of measurements is fixed to 60*

a $\tau = 2.5$, $d = 4$
b $\tau = 3.5$, $d = 4$

illustrated in detail. The simulation results related to the theorems are also studied. Since the constants C and D in (6) and (7) depend on the partially known support level, on the degree of the compressibility of the signal, on p determined by the recovery algorithm and on the number of the measurements M, we test the two constants separately by studying the exact block-sparse case for D and compressible case for C.

4.1 Exactly block-sparse case

In Fig. 1, we generate exact block-sparse signals with 12 non-zero blocks and then observe the noisy measurement $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{\sigma z}$ by setting $\sigma = 0.01$, 0.05 and 0.1 regarding the noise level while the term $\boldsymbol{z}$ denotes the Gaussian white noise. In order to discuss the behaviour of the constant D in (5), we set the error bound of $\|A\boldsymbol{x} - \boldsymbol{y}\|_2^2$, i.e $\epsilon = \sigma\sqrt{M + \lambda\sqrt{2M}}$ and let $\lambda = 2$ as studied in [24]. We plot the curves of the theoretical recovery errors measured by l_2 norm versus different values of p. It can be easily calculated that $\epsilon = 0.113$, 0.566 and 1.133 which further implies that when incorporating known support, the constant D is roughly not beyond 2.5. This illustrates that for a wide range of p, (3) guarantees a stable recovery of block-sparse signals in the presence of noise.

For the next set of experiments, the signals are generated with 48 non-zero entries of which the locations are selected randomly using standard Gaussian distribution. We observe $\boldsymbol{y}$ by $\boldsymbol{y} = \boldsymbol{\Phi x}$. The relative error between the original signal and the reconstructive signal in negative-log scale, i.e $-\log(\|\boldsymbol{x}^* - \boldsymbol{x}\|_2/\|\boldsymbol{x}\|_2)$, is used as the performance measure. In Figs. 2*a–c*, we compare the performance of l_2/l_p for different values of p by plotting the relative error in negative-log scale versus measurement level M choosing different known block support size $s = 2$, 6 and 10, respectively. It is clear that compared with $p = 1$, smaller p means one need fewer measurements for exact recovery. In Figs. 3*a–c*, we set the measurement level to $M = 80{,}100$ and 120, respectively, for different values of p varying the number of known block support. We can see from the curves that more known support information contributes to the improvement of l_2/l_p minimisation.

4.2 Compressible case

Fig. 4 is based on the compressible signals with the l_2 norm of blocks decay like $j^{-\tau}$ where $j \in \{1, \ldots, m\}$ and $\tau > 1$. The measurements $\boldsymbol{y}$ is obtained by $\boldsymbol{y} = \boldsymbol{\Phi x}$ and the parameter is set by $\tau = 2.5$ and 3.5 in purpose of testing the constant C in Theorem 2. Note that the term $\|\boldsymbol{r} - \boldsymbol{r_K}\|$ in Theorem 2 can be bounded by the best K-block approximation error of the signal. Thus for nearly nine-block-sparse signals ($\tau = 2.5$) and nearly six-block-sparse signals ($\tau = 3.5$), it is enough for us to check the nine-block and six-block approximation errors ,respectively, which are 0.239 and 0.039 on average. One can

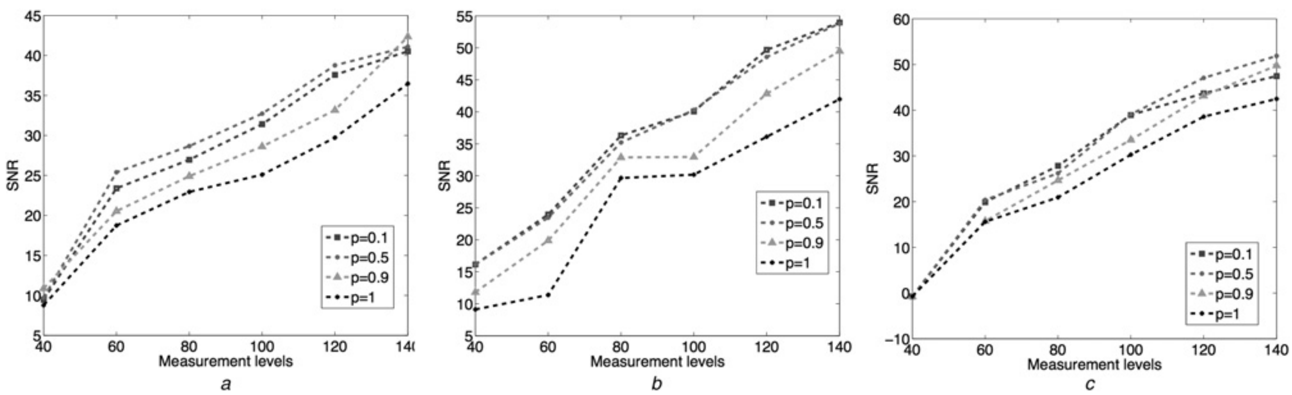

Fig. 5 *Performance of l_2/l_p minimisation with partially known support in terms of SNR for block-compressible signal, varying the number of measurements for*

a $s = 2$
b $s = 6$
c $s = 10$

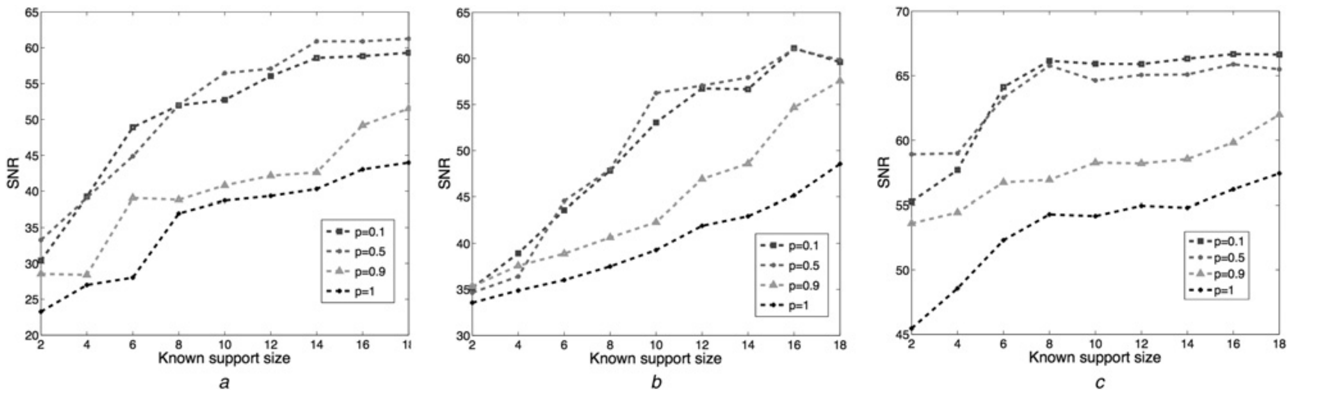

Fig. 6 *Performance of l_2/l_p minimisation with partially known support in terms of SNR for block-compressible signal, varying the number of measurements for*
a $M = 100$
b $M = 120$
c $M = 160$

deduce from Fig. 4 that the constant C is <10 and hence the method proposed in this paper achieves robust recovery.

In the following experiments, we fix $\tau = 2.5$ and plot reconstruction signal-to-noise ratio (SNR) varying the measurement level for $s = 2$, 6 and 10. SNR represents the average reconstruction SNR which is calculated as $\text{SNR} = 20\log_{10}\big((\|\boldsymbol{x}\|_2/\|\boldsymbol{x}-\boldsymbol{x}^*\|_2)\big)$. As is shown in Fig. 5, in all cases, the l_2/l_p minimisation with partially known support gives better performance when M, the measurement level, is lower. Similar to the results in the exactly sparse case, one can improve the recovery by decreasing p value when meeting low measurement level. To analyse the effect of partially known information for different number of the block support size s, SNR is shown in Figs. 6*a*–*c* for $M = 100$,120 and 160. It is obvious that SNR increases with larger size of known block support. From a practical perspective, incorporating more exact known support information can bring better recovery result.

5 Conclusion

In this paper, we modify the l_2/l_p minimisation by incorporating partially known support in the process of recovering block-sparse signals. Theoretical stability guarantees of this non-convex method are established and upper bounds of the reconstruction error are given by combing the p-RIP condition and partially known information. Particularly, we derive a theorem to determine the number of Gaussian random measurements to recover the signal with high probability. Numerical experiments are conducted to verify the theoretical results and further demonstrate that the modified l_2/l_p minimisation could improve its performance, accordingly needing fewer measurements to yield a good approximate reconstruction of the original signal.

6 Acknowledgments

This work was partially supported by the National Natural Science Foundation of China under Grant Nos. 11501440, 61273020.

7 References

1 Donoho, D.L.: 'Compressed sensing', *IEEE Trans. Inf. Theory*, 2006, **52**, (4), pp. 1289–1306
2 Candes, E., Romberg, J., Tao, T.: 'Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information', *IEEE Trans. Inf. Theory*, 2006, **52**, (2), pp. 489–509
3 Erickson, S., Sabatti, C.: 'Empirical Bayes estimation of a sparse vector of gene expression changes', *Stat. Appl. Genet. Mol. Biol.*, 2005, **4**, (22), pp. 1–35
4 Mishali, M., Eldar, Y.C.: 'Blind multi-band signal reconstruction: compressed sensing for analog signals', *IEEE Trans. Signal Process.*, 2009, **57**, (3), pp. 993–1009
5 Majumdar, A., Ward, R.K.: 'Compressed sensing of color images', *Signal Process.*, 2012, **90**, (12), pp. 3122–3127
6 Stojnic, M., Parvaresh, F., Hassibi, B.: 'On the reconstruction of block-sparse signals with an optimal number of measurements', *IEEE Trans. Signal Process.*, 2009, **57**, (8), pp. 3075–3085
7 Eldar, Y., Mishali, M.: 'Robust recovery of signals from a structured union of subspaces', *IEEE Trans. Inf. Theory*, 2009, **55**, (11), pp. 5302–5316
8 Eldar, Y., Kuppinger, P., Blcskei, H.: 'Block-sparse signals: uncertainty relations and efficient recovery', *IEEE Trans. Signal Process.*, 2010, **58**, (6), pp. 3042–3054
9 Tropp, J.A., Gilbert, A.C.: 'Signal recovery from random measurements via orthogonal matching pursuit', *IEEE Trans. Inf. Theory*, 2007, **53**, (12), pp. 4655–4666
10 Wang, Y., Wang, J., Xu, Z.: 'On recovery of block-sparse signals via mixed $l_2 = l_q(0 < q \le 1)$ norm minimization', *EURASIP J. Adv. Signal Process.*, 2013, **76**, pp. 1–17
11 Vaswani, N., Lu, W.: 'Modified-CS: modifying compressive sensing for problems with partially known support', *IEEE Trans. Signal Process.*, 2010, **58**, (9), pp. 4595–4607
12 Lu, W., Vaswani, N.: 'Exact reconstruction conditions and error bounds for regularized modified basis pursuit (reg-modified-BP)'. 2010 Conf. Record of the Forty Fourth Asilomar Conf. on Signals, Systems and Computers (ASILOMAR), 2010, pp. 763–767
13 Jacques, L.: 'A short note on compressed sensing with partially known signal support', *Signal Process.*, 2010, **90**, (12), pp. 3308–3312
14 Borries, R., Miosso, C., Potes, C.: 'Compressed sensing using prior information'. Second IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2007, pp. 121–124
15 Khajehnejad, M.A., Xu, W., Avestimehr, A.S., *et al.*: 'Weighted l_1 minimization for sparse recovery with prior information'. Proc. of the 2009 IEEE Int. Symp. on Information Theory, 2009, pp. 483–487
16 Friedlander, M.P., Mansour, H., Saab, R., *et al.*: 'Recovering compressively sampled signals using partial support information', *IEEE Trans. Inf. Theory*, 2012, pp. 1122–1134
17 Chartrand, R.: 'Exact reconstruction of sparse signals via non convex minimization', *IEEE Signal Process. Lett.*, 2007, **14**, (10), pp. 707–710
18 Foucart, S., Lai, M.: 'Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q \le 1$', *Appl. Comput. Harmon. Anal.*, 2009, **26**, (3), pp. 395–407
19 Xu, Z., Zhang, H., Wang, Y., *et al.*: '$l_{1/2}$ regularization.', *Sci. China Inf.*, 2010, **53**, (6), pp. 1159–1169
20 Wang, Y., Wang, J., Xu, Z.: 'Restricted p-isometry properties of nonconvex block-sparse compressed sensing', *Signal Process.*, 2014, **104**, pp. 188–196
21 Chartrand, R., Staneva, V.: 'Restricted isometry properties and nonconvex compressive sensing', *Inverse Probl.*, 2008, **24**, (3), pp. 1–14
22 Shen, Y., Li, S.: 'Restriced p-isometry property and its application for nonconvex compressive sensing', *Adv. Comput. Math.*, 2012, **37**, pp. 441–452
23 Daubechies, I., DeVore, R., Fornasier, M., *et al.*: 'Iteratively reweighted least squares minimization for sparse recovery', *Commun. Pure Appl. Math*, 2010, **63**, pp. 1–38
24 Cands, E., Romberg, J., Tao, T.: 'Stable signal recovery from incomplete and inaccurate measurements', *Commun. Pure Appl. Math.*, 2006, **59**, (8), pp. 1207–1223