Signal Processing 153 (2018) 34–46

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

On recovery of block sparse signals via block generalized orthogonal matching pursuit

Rui Qi[a,b,c], Diwei Yang[a], Yujie Zhang[a], Hongwei Li[a,c,*]

[a] School of Mathematics and Physics, China University of Geosciences, Wuhan 430074, China
[b] Department of Foundation, Naval University of Engineering, Wuhan 430033, China
[c] Hubei Subsurface Multi-scale Imaging Key Laboratory, China University of Geosciences, Wuhan 430074, China

ARTICLE INFO

Article history:
Received 30 August 2017
Revised 24 June 2018
Accepted 29 June 2018
Available online 5 July 2018

Keywords:
Compressed sensing
Generalized orthogonal matching pursuit
Restricted isometry property
Sparse recovery

ABSTRACT

As a greedy algorithm for compressed sensing reconstruction, generalized orthogonal matching pursuit (gOMP) has been proposed by Wang et al. [10] to recover sparse signals, which simply selects multiple indices without additional postprocessing operation. In this paper, we consider efficient method for the recovery of sparse signals that exhibit additional structure in the form of the non-zero coefficients occurring in clusters. A block version of gOMP is proposed, named block generalized orthogonal matching pursuit (BgOMP). Moreover, theoretical analysis based on restricted isometry property (RIP) for BgOMP is investigated. Simulation results show that BgOMP has considerable recovery performance comparable to state-of-the-art algorithms in terms of probability of exact reconstruction and running time.

1. Introduction

The framework of compressed sensing (CS) concentrates on the recovery of unknown signal from an underdetermined system of linear equations [1,2]:

$$\mathbf{y} = \mathbf{\Phi x} + \mathbf{v}, \tag{1}$$

where $\mathbf{\Phi} \in \mathbb{R}^{m\times n}(m << n)$ is the measurement matrix, $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$ denote the unknown original signal and observed signal, respectively, $\mathbf{v} \in \mathbb{R}^n$ denotes the noise signal with $\|\mathbf{v}\|_2 < \varepsilon$. Since $m < n$, the inverse-problem of recovering $\mathbf{x}$ from $\mathbf{y}$ is ill-posed. However, by imposing extra condition that $\mathbf{x}$ is sparse, one can obtain the unique sparse solution $\mathbf{x}$ through the following sparsity promoting optimization problem [1,2]:

$$\min \|\mathbf{x}\|_0 \text{ subject to } \|\mathbf{y} - \mathbf{\Phi x}\| \le \varepsilon, \tag{2}$$

where the ℓ_0 norm denotes the number of nonzero elements of $\mathbf{x}$.

However, solving ℓ_0 norm minimization problem is NP-hard since it requires a combination search among all possible solutions of (2). Hence, many researchers have devoted to investigating faster recovery algorithms with moderate complexity. Fortunately, Candes and co-authors [1,2] proved the equivalence between the ℓ_0 norm minimization solution and ℓ_1 norm minimization solution. So the above problem can be translated to the following convex optimization problem:

$$\min \|\mathbf{x}\|_1 \text{ subject to } \|\mathbf{y} - \mathbf{\Phi x}\| \le \varepsilon \tag{3}$$

where $\|\mathbf{x}\|_1$ is the ℓ_1 norm of $\mathbf{x}$ defined as $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$. One of the most successful approaches to solve (2) is basis pursuit (BP) algorithm [2,3]. The BP algorithm can be implemented by linear programming (LP) to obtain global optimum but has too much computational burden. A widely used condition of $\mathbf{\Phi}$ ensuring the exact recovery of $\mathbf{x}$ is so called restricted isometry property (RIP) [1]. The sensing matrix $\mathbf{\Phi}$ is said to satisfy the RIP of order K, if there exists a constant $\delta \in (0, 1)$ such that

$$(1-\delta)\|\mathbf{x}\|_2^2 \le \|\mathbf{\Phi x}\|_2^2 \le (1+\delta)\|\mathbf{x}\|_2^2, \tag{4}$$

for any K-sparse vector $\mathbf{x} \in \mathbb{R}^n$. The minimum of all constant δ is referred to as an isometry constant δ_K. Another faster convex algorithm is smoothed ℓ_0 norm (SL0) algorithm [4], it uses a sequence of smoothed functions (e.g. the Gaussian family of functions) to approximate the ℓ_0 norm and minimizes it according to a minimization algorithm. It has many advantages, such as the high matching degree, the short reconstruction time, the low computation complexity and no need for the sparsity of signal.

Recently, a family of iterative greedy algorithms received significant attention due to their low complexity and simple geometric interpretation, which include orthogonal pursuit (OMP) [5], regularized OMP (ROMP) [6], subspace pursuit (SP) [7], stagewise OMP (StOMP) [8] and compressive sampling matching pursuit (CoSaMP) [9], generalized OMP (gOMP) [10], etc. As the representative method in the greedy algorithms, OMP has been widely used

* Corresponding author.
E-mail address: hwli@cug.edu.cn (H. Li).

https://doi.org/10.1016/j.sigpro.2018.06.023
0165-1684/

due to its simplicity and competitive performance. In the meantime, theoretical results about above methods have been also carried out. It is worthwhile pointing out that there are lots of RIP based recovery conditions for OMP, which referred to as [11–19]. Recently, Wen et al. [20] showed that if a sensing matrix $\mathbf{\Phi}$ satisfies the RIP with $\delta_{K+1} < 1/\sqrt{k+1}$, then under some conditions on the minimum of nonzero elements of the K-sparse vector $\mathbf{x}$, the OMP can exactly reconstruct the support set of $\mathbf{x}$ from $\mathbf{y} = \mathbf{\Phi x}+\mathbf{v}$ in K iterations. Compared with existing results, the proposed conditions are sharp in terms of $\delta_{K+1} < 1/\sqrt{k+1}$.

As to the RIP based recovery condition of gOMP algorithm, many results have been proposed in recent years. Liu and Temlyakov [12] established the exact RIP based recovery condition for gOMP algorithm with $\delta_{NK} < 1/((2+\sqrt{2})\sqrt{K/N})$ in the noise-free case. The similar condition given in [10] is $\delta_{NK} < 1/(\sqrt{K/N}+3)$. Subsequently, the above results were improved as $\delta_{NK} < 1/(\sqrt{K/N}+2)$ and $\delta_{NK+1} < 1/(\sqrt{K/N}+1)$ [21]. Recently, it was further refined to $\delta_{NK} < 1/(\sqrt{K/N}+1.27)$ [22]. When considering the noisy case, sufficient conditions of the exact support recovery of K-sparse vector $\mathbf{x}$ have also been studied (see e.g., [23–25]). Particularly, it was shown in [24] that imposing certain conditions on sparse signal $\mathbf{x}$, $\delta_{NK+1} < 1/(\sqrt{K/N}+1)$ is a sufficient condition for exact support recovery which coincides with the result of [21] in the noise-free case. The above conditions were further improved as $\delta_{NK+1} < 1/\sqrt{K/N+1}$ by Wen et al. [25], which is less restrictive than the existing results.

Different from general sparse signals in conventional sense, some real-world signals exhibit some additional structures in the form of nonzero coefficients occurring in clusters. Such signals are referred to as block-sparse signals, which arise in lots of application areas, for instance, DNA microarrays [26], equalization of sparse communication channels [27], multi-band signals [28] and color imaging [29]. In [30], the block version of OMP algorithm, named BOMP, as well as its theoretical guarantees of exact recovery condition based on block coherence was established. In [31], it was proven that sufficient recovery condition is that the sensing matrix $\mathbf{\Phi}$ satisfies the block RIP of order $K+1$ with a small isometry constant $\delta_{K+1} < 1/(2\sqrt{K}+1)$. Recently, the above condition has been improved as $\delta_{K+1} < 1/\sqrt{K+1}$ [32], under which the support set of each block K-sparse signal can be recovered in K iterations when $\mathbf{v} = \mathbf{0}$ and $\mathbf{v} \neq \mathbf{0}$. Another approaches to deal with block sparse problem, which referred to as block version of CoSaMP and iterative hard thresholding (IHT) [33], block SL0 (BSL0) [34], block SP (BSP) [35] and block StOMP (BStOMP) [36] have been investigated. As shown in these papers, we find by utilizing block-sparsity can extremely yield better reconstruction performance than regarding the signal as a general sparse case.

Enlightened by the extension of OMP to the block sparse case, the generalization of gOMP to the block sparse case leads to the block version of gOMP algorithm, named block gOMP (BgOMP). Our goal is to explicitly take the block structure into account, both in terms of the recovery algorithms and exact recovery condition. The main contributions of this paper lie in two aspects:

- Compared with aforementioned works on the recovery of block sparse signals, BgOMP algorithm achieves much better recovery performance which demonstrates the effectiveness of exploiting block structure of underlying signals.
- To the best of our knowledge, no paper reports the exact recovery condition of BgOMP algorithm to date. We prove that under some conditions on minimum norm of each $\mathbf{x}[i]$ and the sensing matrix $\mathbf{\Phi}$ satisfies the block RIP of order $NK+1$ with a small isometry constant $\delta_{NK+1} < 1/\sqrt{K/N+1}$, then at least one index in the support set of $\mathbf{x}$ will be recovered at each iteration. Specially, in the noise-free case, if the sensing matrix $\mathbf{\Phi}$ satisfies the block RIP of order $NK+1$ with a small isometry constant $\delta_{NK+1} < 1/\sqrt{K/N+1}$, then exact recovery of $\mathbf{x}$ is ensured within K iteration.

This manuscript is organized as follows. First, we introduce the concept of block sparsity and block RIP. The BgOMP algorithm is introduced in Section 3. Section 4 provides numerical result of the proposed method comparable to OMP, SP, gOMP, BOMP and BSL0 methods. The sufficient recovery conditions based on RIP is given in Section 5. The last Section is the conclusion of this paper.

We summarize notations used in this paper. Vectors and matrices are denoted by boldface, lowercase and boldface, uppercase letters, respectively. $\mathbb{R}^{m\times n}$ denotes a $m\times n$ matrix, if $n=1$, $\mathbb{R}^{m\times 1}$ denotes a column vector. For a given matrix $\mathbf{\Phi}$, $\mathbf{\Phi}^{\mathrm{T}}$, $\mathbf{\Phi}^{-1}$denote its transpose, inverse, respectively. Let $\Omega = \{1, 2, \cdots, L\}$, then $S\subseteq\Omega$ denotes the true support set of block K-sparse signal $\mathbf{x}$. Obviously, $|S| \leq K$ for any block K-sparse signal $\mathbf{x}$, where $|S|$ is the cardinality of S. For $\Lambda\subseteq S$, $S-\Lambda = S\backslash(S\cap\Lambda)$ is the block index set in S but not in Λ. S^c and Λ^c denote the complement of S and Λ respectively, i.e., $S^c = \Omega - S$, $\Lambda^c = \Omega - \Lambda$. $\mathbf{x}_\Lambda$ is a restriction of the vector $\mathbf{x}$ to the elements with block indices of Λ. $\mathbf{\Phi}_\Lambda$ is a submatrix of $\mathbf{\Phi}$ containing column with block indices of Λ. If $\mathbf{\Phi}_\Lambda$ is full column rank, then $\mathbf{\Phi}_\Lambda^+ = (\mathbf{\Phi}_\Lambda^{\mathrm{T}}\mathbf{\Phi}_\Lambda)^{-1}\mathbf{\Phi}_\Lambda^{\mathrm{T}}$ is the pseudo inverse of $\mathbf{\Phi}_\Lambda$. Span($\mathbf{\Phi}_\Lambda$) is the span of columns in $\mathbf{\Phi}_\Lambda$. $\mathbf{P}_\Lambda = \mathbf{\Phi}_\Lambda\mathbf{\Phi}_\Lambda^+$ is the projection onto span($\mathbf{\Phi}_\Lambda$). $\mathbf{P}_\Lambda^\perp = \mathbf{I} - \mathbf{P}_\Lambda$ is the projection onto the orthogonal complement of span($\mathbf{\Phi}_\Lambda$), where $\mathbf{I}$ stands for the identity matrix.

2. Block sparsity and block RIP

2.1. Block sparsity

Conventional CS only considers the sparse signal with at most K nonzero elements, and it does not take into account any further structure. However, block sparse signals arise in many practical scenarios. The recovery of block sparse vector $\mathbf{x}$ from measurement $\mathbf{y} = \mathbf{\Phi x} + \mathbf{v}$ is the focus of this paper. Generally speaking, a block sparse signal $\mathbf{x} \in \mathbb{R}^n$ over block index set $I = \{d_1, d_2, \cdots, d_L\}$ can be modeled as follows:

$$\mathbf{x} = \Big[\underbrace{x_1, \cdots, x_{d_1}}_{\mathbf{x}[1]}, \underbrace{x_{d_1+1}, \cdots, x_{d_1+d_2}}_{\mathbf{x}[2]}, \cdots, \underbrace{x_{n-d_L+1}, \cdots, x_n}_{\mathbf{x}[L]}\Big]^{\mathrm{T}}, \tag{5}$$

where $\mathbf{x}[i] (i \in \Omega = \{1, 2, \cdots, L\})$ denotes the ith block of $\mathbf{x}$ and d_i is the block size for the ith block. Similarly, measurement matrix $\mathbf{\Phi}$ can be also represented as a concatenation of column blocks $\mathbf{\Phi}[i]$ of size $m \times d_i$

$$\mathbf{\Phi} = \Big[\underbrace{\boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_{d_1}}_{\mathbf{\Phi}[1]}, \underbrace{\boldsymbol{\varphi}_{d_1+1}, \cdots, \boldsymbol{\varphi}_{d_1+d_2}}_{\mathbf{\Phi}[2]}, \cdots, \underbrace{\boldsymbol{\varphi}_{n-d_L+1}, \cdots, \boldsymbol{\varphi}_n}_{\mathbf{\Phi}[L]}\Big]. \tag{6}$$

In addition, $(\mathbf{\Phi}[i])_j$ denotes the jth column of submatrix $\mathbf{\Phi}[i]$. A vector $\mathbf{x}$ is called block K-sparse if $\mathbf{x}[i]$ has nonzero Euclidean norm for at most K indices i. Denoting

$$\|\mathbf{x}\|_{2,0} = \sum_{i=1}^{L} I(\|\mathbf{x}[i]\|_2 > 0), \tag{7}$$

with the indicator function $I(\cdot)$, a block K-sparse vector $\mathbf{x}$ is defined as $\|\mathbf{x}\|_{2,0} \leq K$. When $d_1 = d_2 = \cdots = d_L = 1$, block sparsity degenerates to conventional sparsity.

To reconstruct the block sparse signal $\mathbf{x}$ from underdetermined linear measurements $\mathbf{y} = \mathbf{\Phi x}+\mathbf{v}$, the following optimization problem should be employed [37]

$$\min \|\mathbf{x}\|_{2,0} \text{ subject to } \|\mathbf{y} - \mathbf{\Phi x}\| \leq \varepsilon, \tag{8}$$

Since problem (8) is NP-hard, a convex problem was considered as the following approximation

$$\min \|\mathbf{x}\|_{2,1} \text{ subject to } \|\mathbf{y} - \mathbf{\Phi x}\| \leq \varepsilon, \tag{9}$$

Algorithm 1 The BgOMP algorithm.

Input:$\mathbf{y} \in \mathbb{R}^m$, $\mathbf{\Phi} \in \mathbb{R}^{m\times n}$, K, ε, N.
Initialize: $k=0$, $\mathbf{r}^0=\mathbf{y}$, $\Lambda^0=\emptyset$.
1: **While** $k<K$ and $\|\mathbf{r}^k\|_2>\varepsilon$ do
2: $k=k+1$;
3: Select N block indices $\{\Gamma_i\}|_{i=1,2,\cdots,N}$ corresponding to N largest entries contained in
$\|\langle \mathbf{r}^{k-1}, \mathbf{\Phi}[j]\rangle\|_2$, $j\in\Omega=\{1,2,\cdots,L\}$;
4: $\Lambda^k=\Lambda^{k-1}\cup\{\Gamma_1,\Gamma_2,\cdots,\Gamma_N\}|$;
5: $\hat{\mathbf{x}}_{\Lambda^k}=\mathbf{\Phi}^+_{\Lambda^k}\cdot\mathbf{y}$;
6: $\mathbf{r}^k=\mathbf{y}-\mathbf{\Phi}_{\Lambda^k}\hat{\mathbf{x}}_{\Lambda^k}$.
7: **end while**
Output: $\hat{\mathbf{x}}=\arg\min_{\mathbf{x}:\mathrm{supp}(\mathbf{x})=\Lambda^k}\|\mathbf{y}-\mathbf{\Phi}\mathbf{x}\|_2$.

where $\|\mathbf{x}\|_{2,1}=\sum_{i=1}^{L}\|\mathbf{x}[i]\|_2$. It's worth noting that $\|\mathbf{x}\|_{2,\infty}$ is also used in our paper, which denotes the maximum norm of each $\mathbf{x}[i]$.

In the remainder of this paper, we will pay our attention to how and in what conditions the support set of block sparse signals can be recovered exactly both in noiseless and noisy scenarios respectively. To analyze the performance of algorithms for block sparse signals, the classic RIP was extended to the block RIP.

2.2. Block RIP

In this section, we would like to introduce the definition of block RIP of a measurement matrix, which first appeared in [37] and is a natural extension to the standard definition of RIP.

Definition 1. (Block RIP) Let $\mathbf{\Phi}\in\mathbb{R}^{m\times n}$ be a given matrix. Then $\mathbf{\Phi}$ is said to have the block RIP over $I=\{d_1,d_2,\cdots,d_L\}$ with constant $\delta_{K|I}$ if for every $\mathbf{x}\in\mathbb{R}^n$ that is block K-sparse signal over I, we have that

$$\left(1-\delta_{K|I}\right)\|\mathbf{x}\|_2^2\le\|\mathbf{\Phi}\mathbf{x}\|_2^2\le\left(1+\delta_{K|I}\right)\|\mathbf{x}\|_2^2. \tag{10}$$

For convenience, in the rest of this paper, we still use δ_K, instead of $\delta_{K|I}$, to represent the block RIP constant whenever the confusion is not caused.

To be mentioned, we mainly discuss the recovery of block sparse signal with even block size in this paper, i.e., $d_1=d_2=\cdots=d_L=d$. Note that a block K-sparse signal is Kd sparse in the conventional sense. If $\mathbf{\Phi}$ satisfies the RIP of order Kd, it must hold for all block K-sparse signals. On the contrary, if $\mathbf{\Phi}$ satisfies the block RIP of order K, it may not hold for all Kd sparse signals. That is to say, the block RIP (of order K) is a less stringent requirement comparing with the standard RIP (of order Kd).

It was established in [37] that certain random matrices satisfy the block RIP with overwhelming probability, and that this probability is substantially larger than that of satisfying the standard RIP.

3. Description of BgOMP algorithm

The BgOMP is shown as Algorithm 1. At the kth iteration, the BgOMP algorithm first chooses N block indices $\{\Gamma_i\}|_{i=1,2,\cdots,N}$corresponding to N largest entries contained in $\|\langle\mathbf{r}^{k-1},\mathbf{\Phi}[j]\rangle\|_2(j\in\Omega)$, where $\Omega=\{1,2,\cdots,L\}$ denotes the whole block indices. N is a fixed positive integer determining the number of indices chosen. Specializing for the case that $N=1$, the BgOMP algorithm boils down to BOMP which selects only one block index corresponding to the maximal norm of correlations each time. When $N>1$, the BgOMP algorithm selects more than one block index each time. It is obvious that bigger N results in much more block indices, and subsequently speeds up the algorithm. After obtaining the least-square (LS) solution $\hat{\mathbf{x}}_{\Lambda^k}=\mathbf{\Phi}^+_{\Lambda^k}\cdot\mathbf{y}$, we produce a new residual by using LS fit. The algorithm is repeated until the iteration number reaches maximum $k_{\max}=K$ or the norm of residual $\|\mathbf{r}^k\|_2$ is smaller than a threshold ε.

The most important difference between BOMP algorithm and the proposed algorithm is that the BgOMP algorithm chooses N block indices at each time. Bigger N indicates more block indices will be selected, which results in shorter running time. On the contrary, smaller N results in smaller block indices at each time, which leads to longer running time. What's more, together with BOMP algorithm, BgOMP algorithm also need block sparsity to be known as a prior.

To be mentioned, the residual $\mathbf{r}^k$ in the kth iteration of the BgOMP is orthogonal to the column of $\mathbf{\Phi}_{\Lambda^k}$, i.e.,

$$\begin{aligned}\left\langle\mathbf{\Phi}_{\Lambda^k},\mathbf{r}^k\right\rangle&=\left\langle\mathbf{\Phi}_{\Lambda^k},\mathbf{P}^{\perp}_{\Lambda^k}\mathbf{y}\right\rangle=\mathbf{\Phi}^{\mathrm{T}}_{\Lambda^k}\mathbf{P}^{\perp}_{\Lambda^k}\mathbf{y}\\&=\mathbf{\Phi}^{\mathrm{T}}_{\Lambda^k}\left(\mathbf{P}^{\perp}_{\Lambda^k}\right)^{\mathrm{T}}\mathbf{y}=\left(\mathbf{P}^{\perp}_{\Lambda^k}\mathbf{\Phi}_{\Lambda^k}\right)^{\mathrm{T}}\mathbf{y}=\mathbf{0},\end{aligned} \tag{11}$$

where we have used the following result

$$\mathbf{P}^{\perp}_{\Lambda^k}\mathbf{\Phi}_{\Lambda^k}=(\mathbf{I}-\mathbf{P}_{\Lambda^k})\mathbf{\Phi}_{\Lambda^k}=\mathbf{\Phi}_{\Lambda^k}-\mathbf{\Phi}_{\Lambda^k}\mathbf{\Phi}^+_{\Lambda^k}\mathbf{\Phi}_{\Lambda^k}=\mathbf{0}. \tag{12}$$

As mentioned in BgOMP algorithm, newly selected block indices are not overlapped with previous, that is to say $|\Lambda^k|=Nk$. However, some of the block indices selected in each iteration may be wrong, hence we have $|S\cap\Lambda^k|\le Nk$. Also, convergence in a maximum of K steps indicates that at least one correct block index is chosen at each iteration, in other words, $|S\cap\Lambda^k|\ge k$.

4. Simulation results

Our simulations are performed in Mathworks MATLAB R2014a on the environment of Intel Core2 Quad CPU 2.66 GHz processor with 3.25 G of memory. Several recovery algorithms are used for comparison with respect to the probability of exact reconstruction and running time. Three greedy algorithm including OMP [5], SP [7] and gOMP [10] are taken from the respective author's site. BSL0 algorithm is a convex optimization algorithm mentioned in the paper [34]. The BOMP code is implemented based on the paper [37]. The details of the main items are as follows:

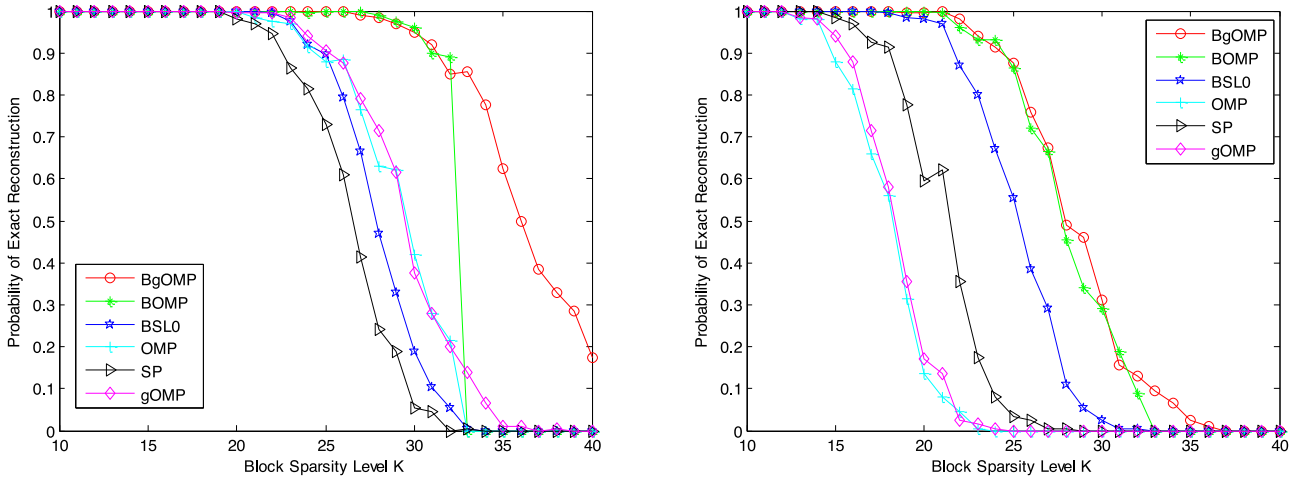
(1) Construct $m\times n$ sensing matrix $\mathbf{\Phi}$ whose entries drawn from Gaussian distribution $N(0,1)$, each column of $\mathbf{\Phi}$ is normalized to unity. The observation vector is calculated by $\mathbf{y}=\mathbf{\Phi}\mathbf{x}$, where m=128 and n=256 are fixed in experiment 4.1, 4.3 and 4.5.
(2) Generate two types of block sparse signals $\mathbf{x}$ with block sparsity K (block size $d=2$). The nonzero blocks are randomly chosen and nonzero elements are 1) drawn independently from standard Gaussian distribution $N(0,1)$, or 2) chosen randomly from the set $\{\pm1\}$, the elements of other blocks are zero. We refer to them as Gaussian or 2-ary pulse amplitude modulation (2-PAM) block sparse signals.
(3) The reconstruction is regarded as successful recovery if

$$\left\|\mathbf{x}-\hat{\mathbf{x}}\right\|_2<10^{-3}, \tag{13}$$

where $\mathbf{x}$ and $\hat{\mathbf{x}}$ denote the original and it's recovery signal, respectively.
(4) To evaluate the performance of recovery and original signal, we use a measure named Averaged Normalized Mean Squared Error (ANMSE) defined as

$$\mathrm{ANMSE}=\frac{1}{200}\sum_{i=1}^{200}10\log_{10}\frac{\left\|\mathbf{x}^{(i)}-\hat{\mathbf{x}}^{(i)}\right\|_2^2}{\left\|\mathbf{x}^{(i)}\right\|_2^2}, \tag{14}$$

where $\mathbf{x}^{(i)}$ and $\hat{\mathbf{x}}^{(i)}$ denote the original signal and it's estimation in ith trial, respectively.
(5) Each test is repeated 200 times, the values of the probability of exact reconstruction and running time are averaged.
(6) The number of indices N to be chosen per iteration is fixed to 2 except for experiment 4.3 and 4.4.



(a) Gaussian block sparse signal (b) 2-PAM block sparse signal

Fig. 1. Probability of exact reconstruction versus block sparsity.

4.1. Recovery property versus block sparsity level

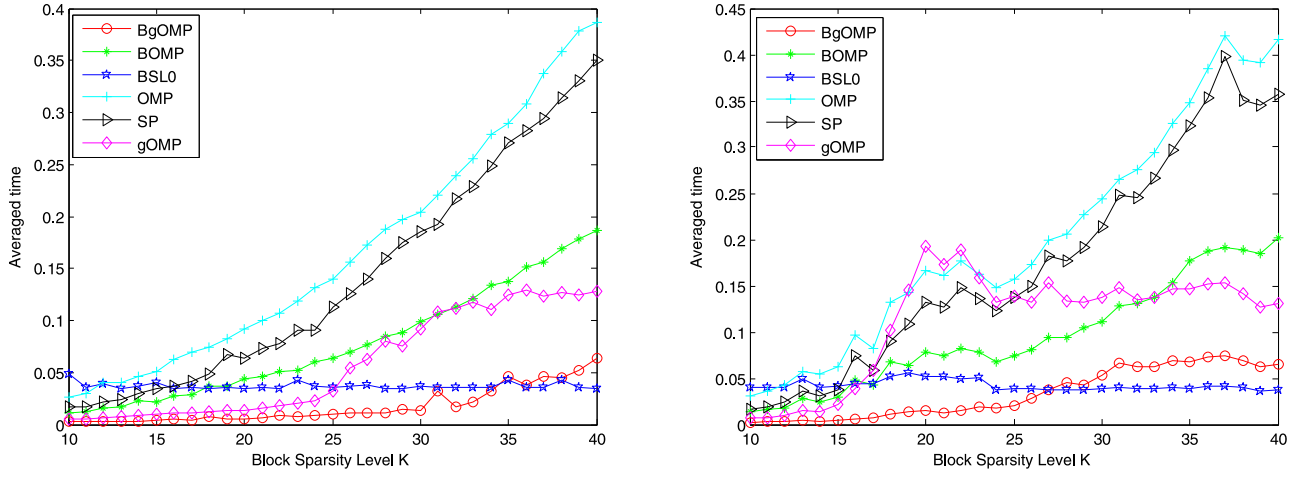
Fig. 1 depicts the probability of exact reconstruction as a function of block sparsity. As shown in Fig. 1, for both Gaussian or 2-PAM block sparse signals, the proposed BgOMP algorithm outperforms another five methods as block sparsity varies from 10 to 40. Specially, in Fig. 1(a), when block sparsity $K > 33$, all the methods fail to recover the original signal besides BgOMP algorithm. Compared with Gaussian block sparse signal in Fig. 1(a), the proposed algorithm is slightly less effective when dealing with the 2-PAM block sparse signals.

In Fig. 2, the running time of the above algorithms under different block sparsity level K is studied. It is clear from Fig. 2 that the running time of BgOMP and BSL0 algorithm is less than another four methods. To be mentioned, the running time of BSL0 algorithm stays constant versus block sparsity, the reason is that the BSL0 is a convex optimization algorithm. In addition, as the block sparsity level varies from 10 to 40, the running time of SP and OMP increases much faster than that of another algorithms.

4.2. Recovery property versus number of measurement

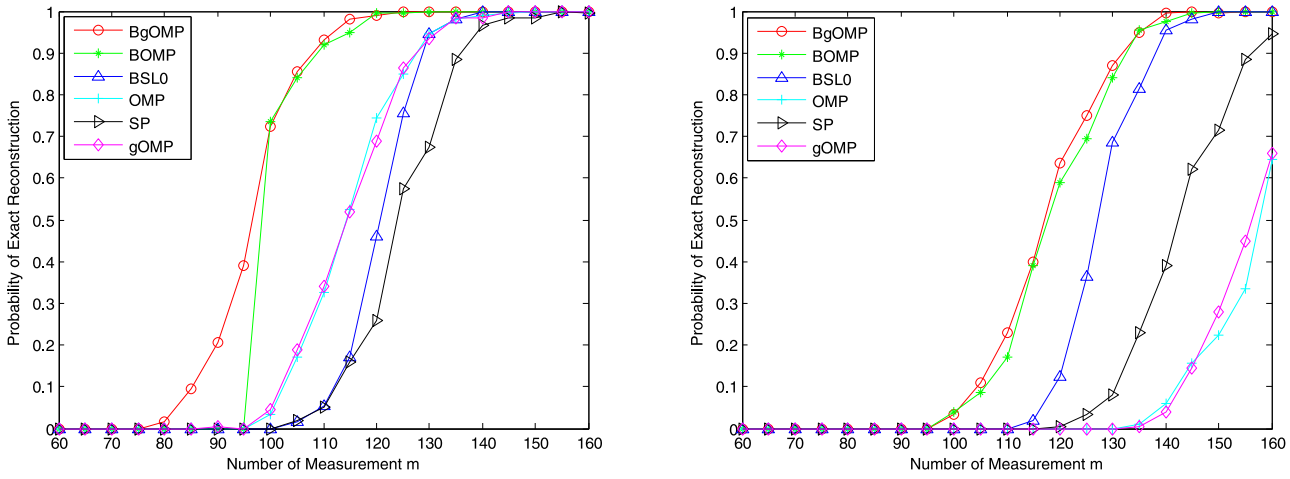
In Fig. 3, the recovery probability is studied in terms of measurement m. The number of measurement m varies from 60 to 160 while the block sparsity K is fixed to 25. As can be seen from Fig. 3(a), the BgOMP algorithm achieves much better reconstruction probability as number of measurement increases. When the number of measurement $m < 90$, the BgOMP algorithm gives considerable recovery probablity, while another five algorithms fail to recover original signal completely. As to Fig. 3(b), the BgOMP and BOMP algorithms achieve similar reconstruction probability which are higher than another four methods.

In this experiment, we compare the running time of above methods with respect to the number of measurement. As shown in Fig. 4, it is obvious that the running time of BgOMP algorithm is superior to another methods for both Gaussian and 2-PAM block sparse signals. The reason is that BgOMP can select more than one atom each time, which results in less running time.



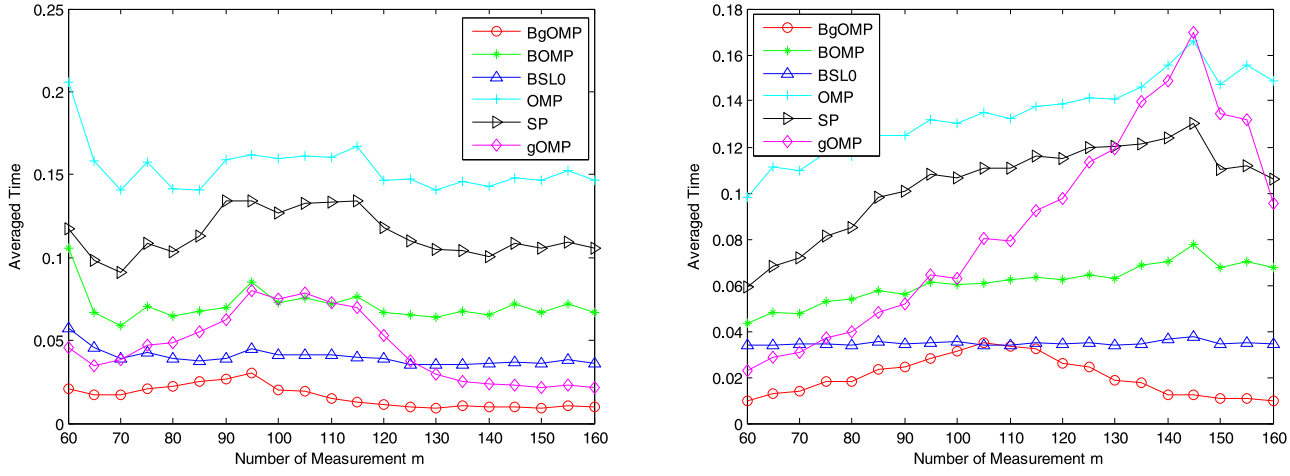
(a) Gaussian block sparse signal (b) 2-PAM block sparse signal

Fig. 2. Running time versus block sparsity.



(a) Gaussian block sparse signal (b) 2-PAM block sparse signal

Fig. 3. Probability of exact reconstruction versus number of measurement.



(a) Gaussian block sparse signal (b) 2-PAM block sparse signal

Fig. 4. Running time versus number of measurement.

4.3. Comparison of BgOMP algorithm with different N *versus block sparsity*

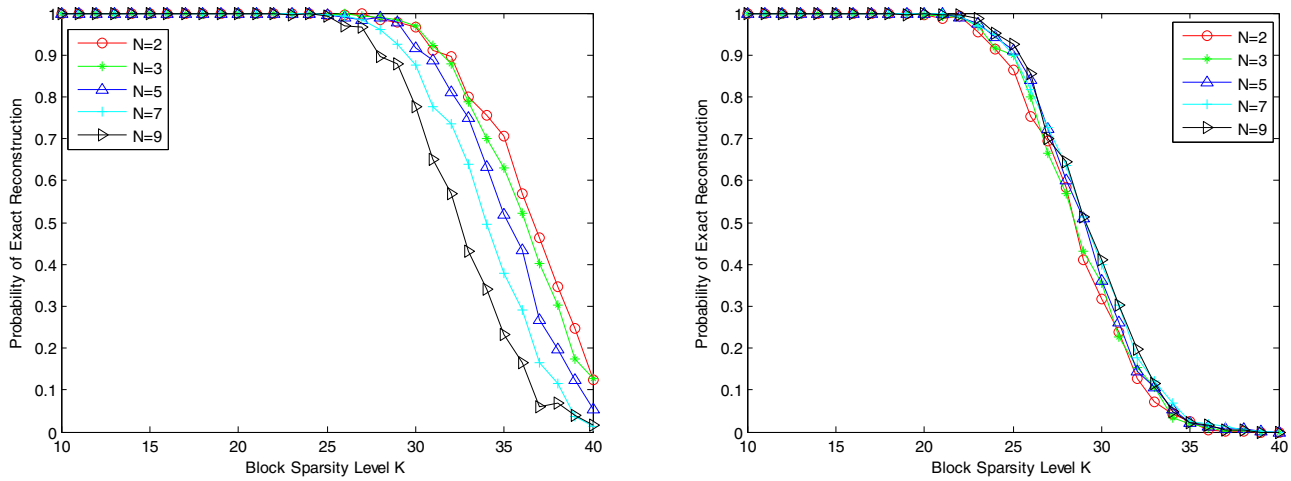
Fig. 5 shows the probability of exact reconstruction of BgOMP algorithm with different $N(N>1)$, where the block sparsity level varies from 10 to 40. As shown in Fig. 5(a) that the property of all the BgOMP algorithm with different N decreases as the block sparsity increases. For a fixed block sparsity level, we find that the recovery probability becomes worse gradually as N varies from 2 to 9. That is to say, when $N=2$, the BgOMP algorithm achieves the best recovery performance. However, from Fig. 5(b), it seems that different N does not affect the recovery property of BgOMP algorithm.

In Fig. 6, running time of BgOMP algorithm with different N versus number of measurement is studied. As displayed in Fig. 6, the running time of BgOMP algorithm increases as block sparsity varies from 10 to 40. Generally speaking, when block sparsity K is big enough, the running time changes slightly. Especially for 2-PAM block sparse signal in Fig. 6(b), it seems that when K exceeds a critical value (35 in this case), the running time remains constant and does not increase anymore. Maybe the reason is that when the block sparsity K exceeds a critical value, the BgOMP algorithm cannot successfully recover original signal, which yields the maximum of K iterations.

4.4. Comparison of BgOMP algorithm with different N *versus number of measurement*

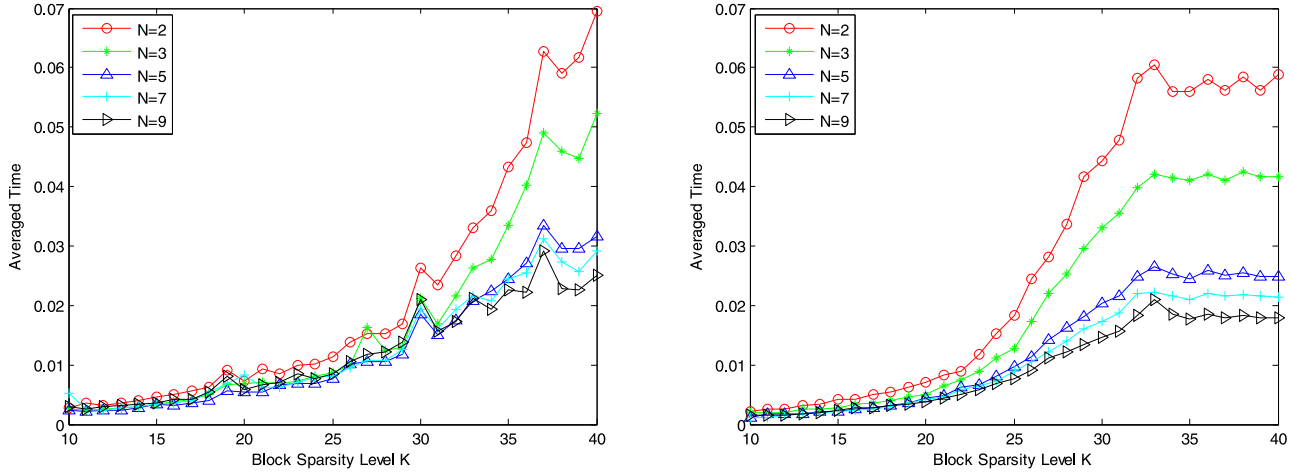
In Fig. 7, the recovery probability is observed as a function of measurement m. The number of measurement m varies from 60 to 160 as the sparsity level is fixed to 25. As can be seen from Fig. 7, the recovery probability increases as number of measurement increases. Specially, Fig. 7(a) describes that with the increase of N, the recovery property of BgOMP algorithm gets worse and worse. Similarly to the result in Fig. 5(b), it seems that different N does not affects the recovery property of BgOMP algorithm.

Fig. 8 describes the running time versus number of measurement m. The running time of BgOMP algorithm with different N is more or less similar as number of measurement varies from 60 to 160. The running time of BgOMP algorithm first increases and then decreases, it reaches the maximum as m is 95 in Fig. 8(a) and 110 in Fig. 8(b). Overall, for both Gaussian and 2-PAM block sparse signal, the BgOMP algorithm with $N=2$ requires much more running time.



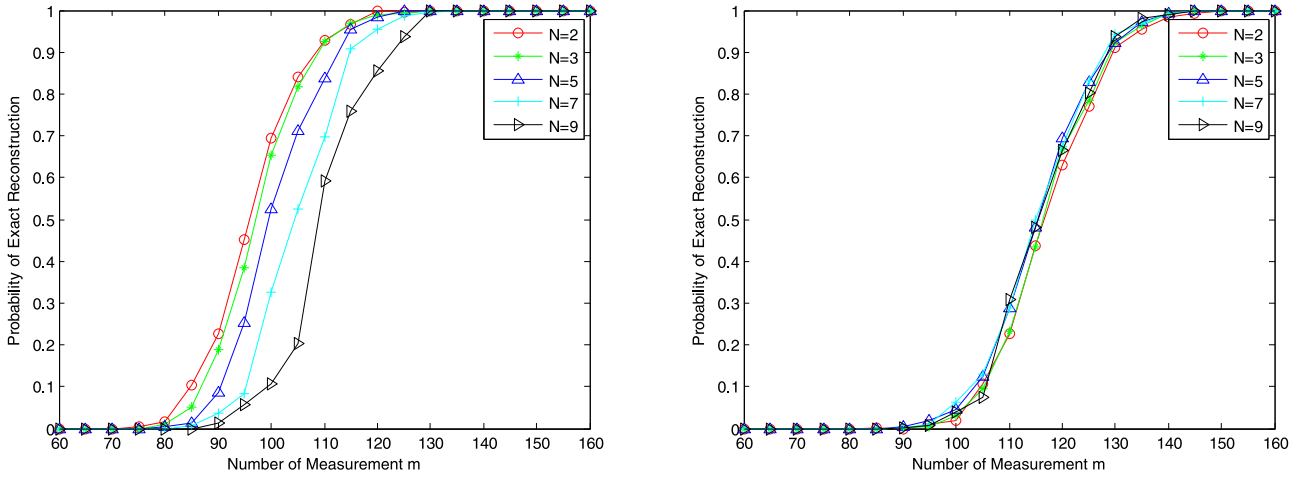
(a) Gaussian block sparse signal (b) 2-PAM block sparse signal

Fig. 5. Probability of exact reconstruction of BgOMP algorithm for different N.



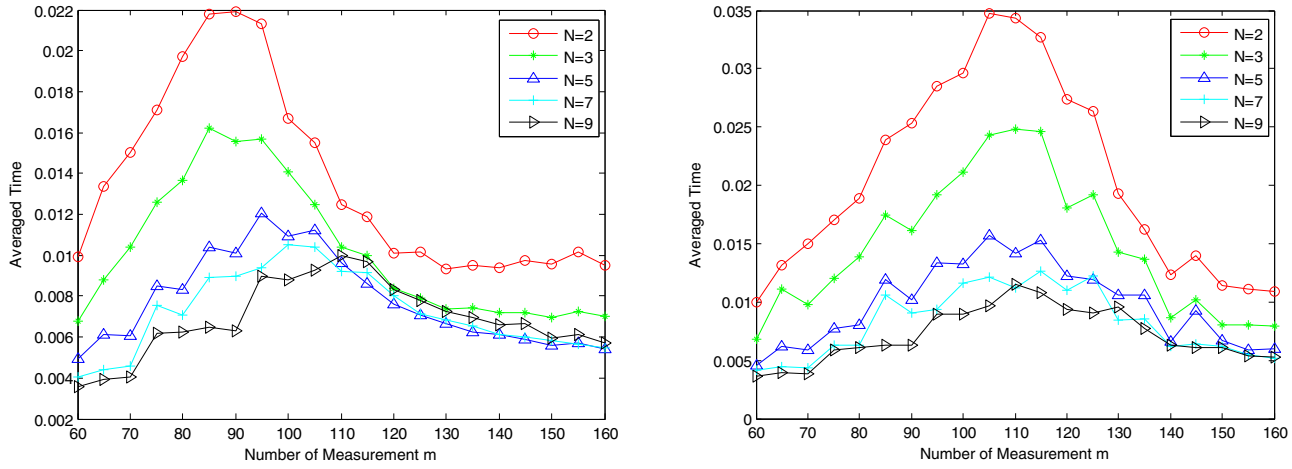
(a) Gaussian block sparse signal (b) 2-PAM block sparse signal

Fig. 6. Running time of BgOMP algorithm for different N.



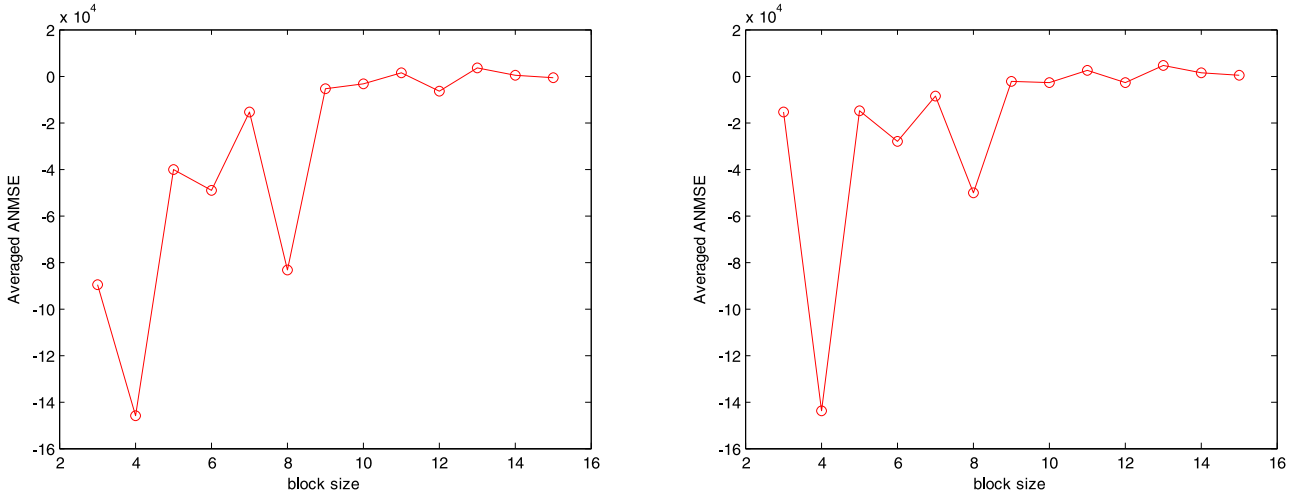
(a) Gaussian block sparse signal (b) 2-PAM block sparse signal

Fig. 7. Probability of exact reconstruction of BgOMP algorithm for different N.



(a) Gaussian block sparse signal (b) 2-PAM block sparse signal

Fig. 8. Running time of BgOMP algorithm for different N.



(a) Gaussian block sparse signal (b) 2-PAM block sparse signal

Fig. 9. Averaged ANMSE of BgOMP algorithm for different d.

4.5. Comparison of BgOMP algorithm with different block size

In order to evaluate the reconstruction property of BgOMP algorithm, different block size d is used to recover the block K-sparse signal. We generate the sources with block size $d = 4$ which is unknown in recovery procedure. In Fig. 9, the averaged ANMSE of BgOMP algorithm is observed with respect to block size d varying from 3 to 15. Note that the block sparsity of signal decreases as the block size increases, where the total sparsity of the signal remains constant, i.e., $Kd = 60$. As shown in Fig. 9, when block size chosen in recovery procedure is multiple of real block size, the performance will be better than that of other block size. Specially, as the block size chosen is exactly equal to the real block size, we will achieve the best performance. That is to say, if we know the real block size d in advance, our algorithm will obtain much better performance. Even if real block size is not known as a prior, we can also recover the sources with high considerable recovery performance.

4.6. ECG signal recovery

In this experiment, we evaluate BgOMP algorithm on real-life electrocardiography (ECG) data. The ECG data comes from the National Metrology Institute of Germany, which is online available in the Physionet [38]. The database contains 549 records from 290 patients, where each signal record contains 38,400 data points. For each of the total 200 trials, we randomly select a short-term segment which consists of $n = 1024$ data samples from the signal record of one patient. Note that the ECG signal themselves are not sparse in the time domain as shown in Fig. 10(a). We apply the orthogonal Daubechies wavelets (db1), which is reported to be the most popular wavelet family for ECG compression, to obtain the sparse signal shown in Fig. 10(b). A compressed sensing Gaussian matrix $\mathbf{\Omega} \in \mathbb{R}^{m\times n}$ is generated randomly for each trial as well. Let $\boldsymbol{\psi}$ denotes the wavelet transform basis, then compressed sensing measurement can be written as $\mathbf{x} = \Omega\,\boldsymbol{\psi}\,\theta = \mathbf{\Phi}\theta$, where θ is the coefficient of wavelet transform. The recovered sparse signal $\hat{\theta}$ and the corresponding reconstructed signal $\hat{\mathbf{x}}$ are illustrated in Fig. 10(d) and Fig. 10(c), respectively.

Fig. 11 displays the averaged ANMSE of different methods as a function of number of measurement. Note that the SP and BSL0 methods don't exist in this experiment, the reason is that the reconstruction property of SP and BSL0 methods is very poor. From Fig. 11, we find that the BgOMP method provides considerable re-

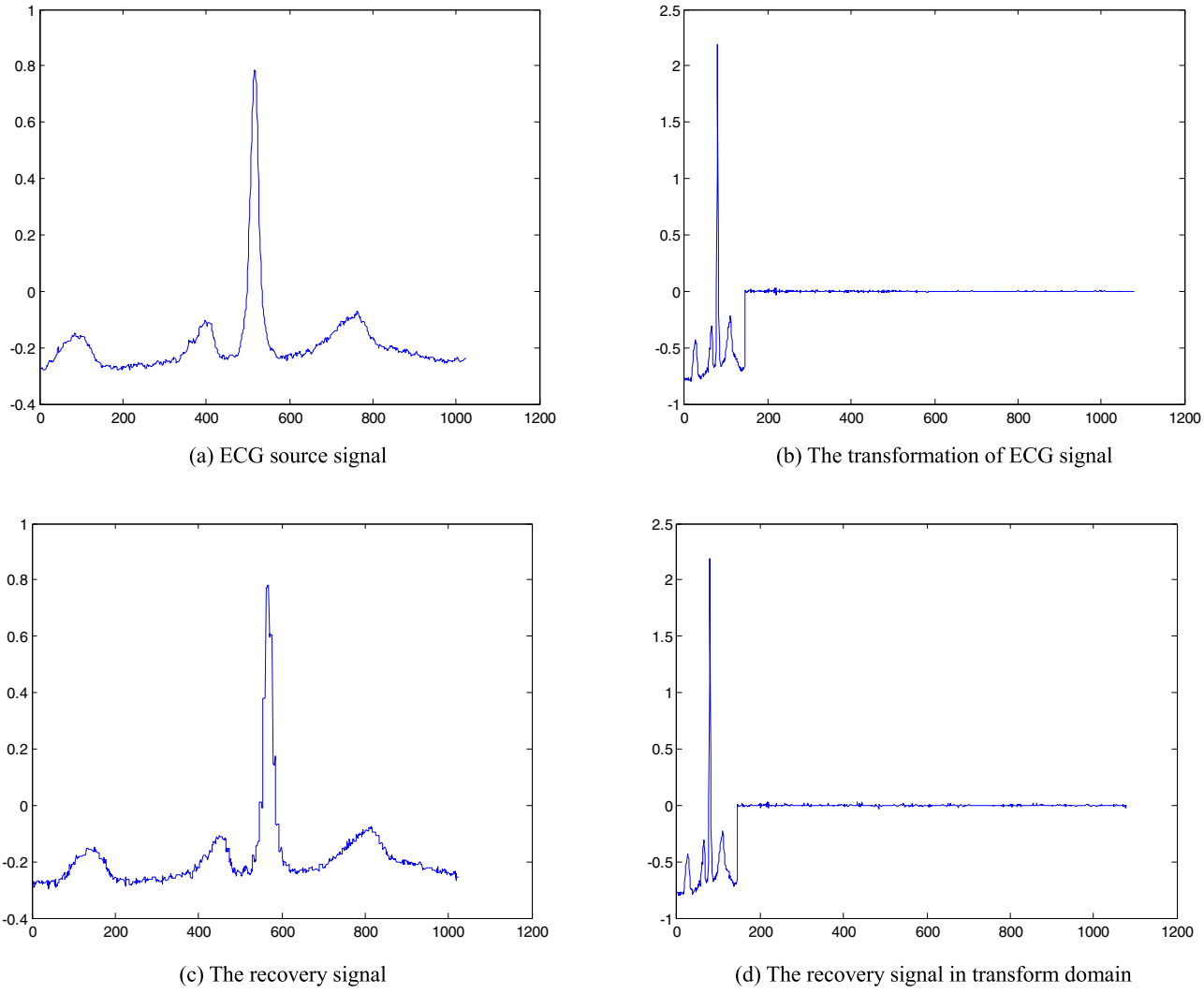


Fig. 10. The ECG signal selected randomly from the PTB Diagnostic ECG Database.

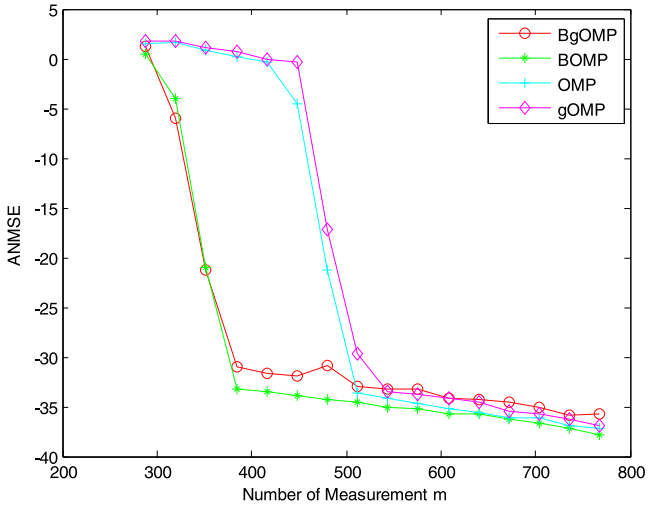


Fig. 11. The averaged ANMSE of different methods versus number of measurement.

covery performance, even if the original signal is not block sparse with even block size.

4.7. Image data recovery

In this subsection, we deal with image 'Lena' with size 256×256 to demonstrate the property of the proposed algorithm. To reduce the computational complexity, we reconstruct the image block by block. The image 'Lena' is first cut into 256 blocks with size 16×16, then the processing of whole image is divided into several smaller, and hence simple procedure. For each block, by using the 2D Haar Wavelet transform basis, the coefficients in the wavelet domain tend to be K-sparse, where for each block only K largest magnitude wavelet coefficients are kept. After each block has been recovered by using the CS procedure, the whole image will be obtained by concatenating all the blocks.

In order to evaluate the quality of estimation, Peak Signal-to-Noise Ratios (PSNR) is most commonly used as a measure, which is defined via the Mean Squared Error (MSE). Given two $m \times n$ monochrome images $\mathbf{F}$ and $\mathbf{G}$, MSE is defined as

$$\mathrm{MSE} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \|\mathbf{F}(i,j) - \mathbf{G}(i,j)\|. \tag{15}$$

The PSNR is defined as [39]

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{255^2}{\mathrm{MSE}}\right) = 20\log_{10}\left(\frac{255}{\sqrt{\mathrm{MSE}}}\right). \tag{16}$$

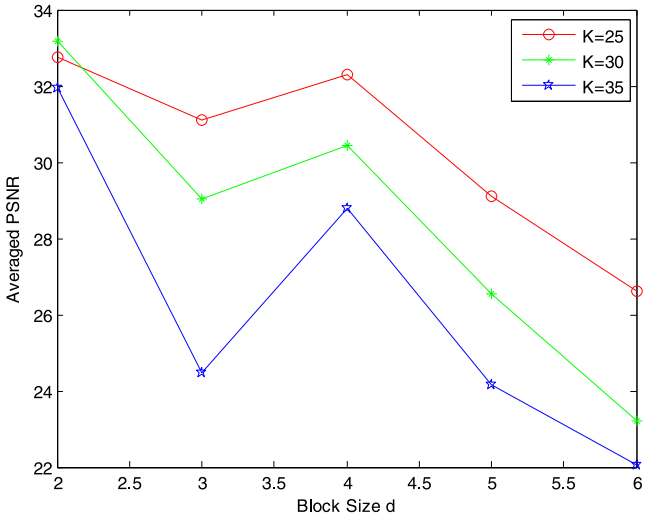


Fig. 12. PSNR versus K largest magnitude wavelet coefficients as a function of block size d.

(a) original image (b) recovery image

Fig. 13. The original image and recovery image by using BgOMP method.

To be mentioned, the block sparse signal by retaining the K largest magnitude wavelet coefficients is not of block sparse signal with even block size. Hence, we have two questions to be solved for our experiments. One is that how many largest magnitude wavelet coefficients should be retained, the other is that how to choose the block size d, under which the above block sparse signal can be regarded as block sparse signal with even size. Fig. 12 depicts the averaged PSNR versus K largest magnitude wavelet coefficients as a function of block size d. We compare the performance of BgOMP with BOMP, OMP and gOMP.

As shown in Fig. 12, we observe that for each K, the best block size should be chosen to be 2 or 4. On the other hand, the averaged PSNR is bigger when $K = 30$ for a fixed block size d. By using $K = 30$ and $d = 2$, the original and recovery images are shown in Fig. 13. In this case, the averaged PSNR of BgOMP algorithm is 32.7484. As shown in above, we find even if the image signal is block sparse signal with arbitrary block size, the BgOMP algorithm also gives considerable recovery property.

5. RIP based recovery condition analysis

In this section, we study the exact recovery condition under which the BgOMP algorithm can perfectly reconstruct the support set of block sparse signals. Firstly, we give several Lemmas used in our theoretical analysis, which were proposed in [32].

Lemma 1. [32] (*Monotonicity of δ_K*) *If the sensing matrix $\mathbf{\Phi}$ satisfies the block RIP of both order K_1 and K_2, then $\delta_{K_1} \le \delta_{K_2}$ for any $K_1 \le K_2$.*

Lemma 2. [32] *Let $\mathbf{\Phi}$ satisfy the block RIP of order K and Λ be a set with $|\Lambda| \le K$. Then for any $\mathbf{x} \in \mathbb{R}^m$,*

$$\left\| \mathbf{\Phi}_\Lambda^{\mathrm{T}} \mathbf{x} \right\|_2 \le (1+\delta_K) \|\mathbf{x}\|_2. \tag{17}$$

Lemma 3. [32] *Let Λ_1, Λ_2 be two subsets of Ω with $|\Lambda_2 - \Lambda_1| \ge 1$. If a matrix $\mathbf{\Phi}$ satisfies the block RIP of order $|\Lambda_1 \cup \Lambda_2|$, then for any vector $\mathbf{x} \in \mathbb{R}^{|\Lambda_2 - \Lambda_1| \times d}$,*

$$\left(1-\delta_{|\Lambda_1\cup\Lambda_2|}\right)\|\mathbf{x}\|_2^2 \le \left\| \mathbf{P}_{\Lambda_1}^{\perp} \mathbf{\Phi}_{\Lambda_2-\Lambda_1} \mathbf{x} \right\|_2^2 \le \left(1+\delta_{|\Lambda_1\cup\Lambda_2|}\right)\|\mathbf{x}\|_2^2. \tag{18}$$

We also give a useful lemma which plays a key role in our theoretical analysis.

Lemma 4. *Let set $\Lambda \subseteq \Omega$ satisfy $\Lambda = kN$ and $S \cap \Lambda = l$ for some integers N, k and l with $0 \le k \le l \le |S| - 1$ and $N(k+1) + |S| - k \le m/d$. Let $W \subseteq S^c$ satisfy $|W| = N$ and $W \cap \Lambda = \emptyset$. If $\mathbf{\Phi}$ satisfies the block RIP of order $N(k+1) + |S| - l \le m/d$, then*

$$\left\| \mathbf{\Phi}^{\mathrm{T}}{}_{S-\Lambda} \mathbf{P}_\Lambda^{\perp} \mathbf{\Phi}_{S-\Lambda} \mathbf{x}_{S-\Lambda} \right\|_{2,\infty} - \frac{1}{N} \sum_{j\in W} \left\| \mathbf{\Phi}^{\mathrm{T}}[j] \mathbf{P}_\Lambda^{\perp} \mathbf{\Phi}_{S-\Lambda} \mathbf{x}_{S-\Lambda} \right\|_2 \ge \frac{(1-\sqrt{(|S|-l)/N+1}\,\delta_{N(k+1)+|S|-l}) \|\mathbf{x}_{S-\Lambda}\|_2}{\sqrt{|S|-l}}. \tag{19}$$

Proof: See Appendix A

Remark 1. Lemma 4 can be regarded as an extension version of [25, Lemma 4] and [32, Lemma 4], the reason is that Lemma 4 extends [25, Lemma 4] for $d = 1$ to general d and [32, Lemma 4] for $N = 1$ to general N. To be mentioned, the proof of Lemma 4 is mainly based on the techniques for proving [25, Lemma 4] and [32, Lemma 4].

Remark 2. The condition $N(k+1) + |S| - k \le m/d$ is to ensure the assumption that $\mathbf{\Phi}$ satisfies block RIP of order $N(k+1) + |S| - k$ makes sense.

In the following, we state our main result as follows, whose proof is based on [25].

Theorem 1. *Suppose the noise vector $\mathbf{v}$ satisfies $\|\mathbf{v}\|_2 < \varepsilon$. Let $\mathbf{\Phi}$ satisfy block RIP with*

$$\delta_{N(k+1)+|S|-k} < 1/\sqrt{|S|/N+1}, \tag{20}$$

for some integer k and N with $0 \le k \le |S| - 1$ and $N(k+1) + |S| - k \le m/d$. Then the BgOMP identifies at least one index in S in each of the first $k+1$ iterations until all the indices in S are selected or BgOMP terminates provided that

$$\min_{i\in S} \|\mathbf{x}[i]\|_2 > \frac{2\varepsilon}{1-\sqrt{|S|/N+1}\,\delta_{N(k+1)+|S|-k}}. \tag{21}$$

Proof: See Appendix B.

Remark 3. Following from $N(k+1) + |S| - k \le m/d$, we conclude that $N \le (m-d)/Kd$ is required in our paper, which allows more choices of N for BgOMP algorithm. When $d = 1$, BgOMP algorithm degenerates to gOMP algorithm, where $N \le (m-1)/K$ is also required in paper [25].

Specializing for the case $k = |S| - 1$ and together with Lemma 1, we can easily obtain the following result.

Theorem 2. *Suppose $\mathbf{\Phi}$ satisfies block RIP of order $NK + 1$ with*

$$\delta_{NK+1} < 1/\sqrt{K/N+1}. \tag{22}$$

Then for an integer N with $1 \le N \le (m-d)/Kd$ and $\|\mathbf{v}\|_2 < \varepsilon$, the BgOMP identifies at least k_0 indices in S after performing k_0 iterations or recovers S in K iterations provided that

$$\min_{i\in S}\|\mathbf{x}[i]\|_2 > \frac{2\varepsilon}{1-\sqrt{K/N+1}\delta_{Nk+1}}. \tag{23}$$

When $N=1$, BgOMP reduces to BOMP, the following result can be directly obtained from Theorem 2.

Corollary 1. Suppose $\mathbf{\Phi}$ and $\mathbf{v}$ respectively satisfy block RIP of order $K+1$ with $\delta_{K+1} < 1/\sqrt{K+1}$ and $\|\mathbf{v}\|_2 < \varepsilon$. Then the BOMP identifies at least k_0 indices in S after performing k_0 iterations or recovers S in K iterations provided that

$$\min_{i\in S}\|\mathbf{x}[i]\|_2 > \frac{2\varepsilon}{1-\sqrt{K+1}\delta_{K+1}}. \tag{24}$$

Remark 4. The support recovery condition for BOMP [32, Theorem 3] is the same as Corollary 1 with respect to RIP and minimum norm of each $\mathbf{x}[i]$.

As mentioned in Theorem 2, the BgOMP algorithm may terminate after performing $0<k_0<K$ iterations, that is to say, the support set S cannot be guaranteed to be reconstructed under conditions (22) and (23). To ensure the exact reconstruction of support set S, we give the following Lemma.

Lemma 5. *Suppose that $\mathbf{v}=\mathbf{0}$, let $\mathbf{\Phi}$ satisfy the RIP with (20) for some integers N, k and $0\le k\le |S|-1$ and $1\le N\le (m-d)/Kd$. If there exists an integers k_0 with $0<k_0\le k$ and $|S\cap\Lambda^{k_0}|\ge k_0$ such that $\|\mathbf{r}^{k_0}\|_2=0$. Then $S\subseteq\Lambda^{k_0}$.*

Proof. The proof of procedure follows the same logic way of proof of [25, Lemma 5], thus we omit it here.

Since the BgOMP does not involve the backtracking step, which indicates the block indices selected in the previous stage will stay in the support set, but it may be wrong. If the BgOMP algorithm terminate after performing $0<k_0<|S|$ iterations, the final support set Λ^{k_0} may contains block indices not in S (i.e., $|\Lambda^{k_0}|>S$). Thankfully, even in this situation, the original signal will also be recovered successfully. The reason is that

$$\begin{aligned}\hat{\mathbf{x}}_{\Lambda^{k_0}} &= \mathbf{\Phi}^{+}_{\Lambda^{k_0}}\mathbf{y} = \left(\mathbf{\Phi}^{\mathbf{T}}_{\Lambda^{k_0}}\mathbf{\Phi}_{\Lambda^{k_0}}\right)^{-1}\mathbf{\Phi}^{\mathbf{T}}_{\Lambda^{k_0}}\mathbf{\Phi}_S\mathbf{x}_S\\ &= \left(\mathbf{\Phi}^{\mathbf{T}}_{\Lambda^{k_0}}\mathbf{\Phi}_{\Lambda^{k_0}}\right)^{-1}\mathbf{\Phi}^{\mathbf{T}}_{\Lambda^{k_0}}\mathbf{\Phi}_{\Lambda^{k_0}}\mathbf{x}_{\Lambda^t} + \left(\mathbf{\Phi}^{\mathbf{T}}_{\Lambda^{k_0}}\mathbf{\Phi}_{\Lambda^{k_0}}\right)^{-1}\\ &\quad \mathbf{\Phi}^{\mathbf{T}}_{\Lambda^{k_0}}\mathbf{\Phi}_{\Lambda^{k_0}-S}\mathbf{x}_{\Lambda^{k_0}-S}\\ &= \mathbf{x}_{\Lambda^{k_0}},\end{aligned} \tag{25}$$

where we have used the result that $\mathbf{x}_{\Lambda^{k_0}-S}=\mathbf{0}$. By induction, we find that as long as at least one correct block index is chosen in each iteration, the original block sparse signal will be successfully reconstructed within K iterations.

Combining Theorem 2 and the above statement, we have the following result.

Theorem 3. *Suppose that $\mathbf{v}=\mathbf{0}$, let $\mathbf{\Phi}$ satisfy the block RIP of order $NK+1$ with $\delta_{NK+1}<1/\sqrt{K/N+1}$ for an integer N with $1\le N\le (m-d)/Kd$. Then the BgOMP algorithm can successfully recover $\mathbf{x}$ within K iterations.*

Remark 5. Specializing the above results for the case $N=1$, we obtain the exact recovery condition of BOMP algorithm, which coincides with the result in [32, Theorem 2].

Remark 6. Specializing the above results for the case $d=1$, we can immediately obtain the corresponding exact recovery condition of gOMP algorithm, which coincides with the result in [25, Theorem 3].

6. Conclusions

In this paper, a new block sparse signal reconstruction algorithm, named BgOMP, is proposed. This algorithm can be viewed as the block version of gOMP algorithm. Due to the fact that multiple indices can be selected each time, the BgOMP algorithm requires less running time to reconstruct the original signal successfully. Simulation results show that BgOMP algorithm has excellent recovery performance comparable to many existing algorithms with respect to probability of exact reconstruction and running time. Actual datum including ECG signal and image signal are also considered in our paper, which demonstrates the effectiveness of the BgOMP algorithm.

In addition, the recovery performance is studied by using block RIP. It is shown that if the sensing matrix $\mathbf{\Phi}$ satisfies the block RIP of order $NK+1$ with a small isometry constant $\delta_{NK+1}<1/\sqrt{K/N+1}$, then under a condition on minimum norm of $\mathbf{x}[i]$, BgOMP algorithm identifies at least one index in the support set of $\mathbf{x}$ at each iteration. Moreover, it is a sufficient condition for exact recovering $\mathbf{x}$ within K iterations in noise-free case. However, the BgOMP algorithm needs the sparsity to be known as a prior. Future works include investigation of efficient block sparse algorithm dealing with block sparse signal with arbitrary block size.

Acknowledgment

The authors would like to thank the anonymous reviewers and the editors for their valuable comments and suggestions. This work is supported by Natural Science Foundation of China (no. 61601417), Natural Science Foundation of Hubei Province (no. 2016CFC717).

Appendix A. Proof of Lemma 4

In the following, we follow the proofs of [25, Lemma 4] and [32, Lemma 4] to prove this lemma.

We first prove the inequality

$$\left\|\mathbf{P}^{\perp}_{\Lambda}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2^2 \le \sqrt{|S|-l}\|\mathbf{x}_{S-\Lambda}\|_2\left\|\mathbf{\Phi}^{\mathrm{T}}{}_{S-\Lambda}\mathbf{P}^{\perp}_{\Lambda}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_{2,\infty}. \tag{26}$$

Since $|S\cap\Lambda|=l\le|S|-1$, then $|S-\Lambda|=S-l$, $\|\mathbf{x}_{S-\Lambda}\|_{2,1}\ne 0$. Thus, we have

$$\begin{aligned}&\sqrt{|S|-l}\|\mathbf{x}_{S-\Lambda}\|_2\left\|\mathbf{\Phi}^{\mathrm{T}}{}_{S-\Lambda}\mathbf{P}^{\perp}_{\Lambda}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_{2,\infty}\\ &\ge \|\mathbf{x}_{S-\Lambda}\|_{2,1}\left\|\mathbf{\Phi}^{\mathrm{T}}{}_{S-\Lambda}\mathbf{P}^{\perp}_{\Lambda}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_{2,\infty}\\ &= \Big(\sum_{i\in S-\Lambda}\|\mathbf{x}[i]\|_2\Big)\left\|\mathbf{\Phi}^{\mathrm{T}}{}_{S-\Lambda}\mathbf{P}^{\perp}_{\Lambda}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_{2,\infty}\\ &\ge \sum_{i\in S-\Lambda}\Big(\|\mathbf{x}[i]\|_2\left\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{P}^{\perp}_{\Lambda}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2\Big)\\ &\overset{a}{\ge} \sum_{i\in S-\Lambda}\Big(\left\|\mathbf{x}^{\mathrm{T}}[i]\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{P}^{\perp}_{\Lambda}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2\Big)\\ &\overset{b}{=} \sum_{i\in S-\Lambda}\left(\mathbf{x}^{\mathrm{T}}[i]\mathbf{\Phi}^{\mathrm{T}}[i](\mathbf{P}^{\perp}_{\Lambda})^{\mathrm{T}}\mathbf{P}^{\perp}_{\Lambda}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right)\\ &= \left\|\mathbf{P}^{\perp}_{\Lambda}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2^2,\end{aligned} \tag{27}$$

where we have used the Cauchy-Schwarz inequality in (a). The reason for (b) holding is due to the fact

$$\left(\mathbf{P}^{\perp}_{\Lambda}\right)^{\mathrm{T}}\mathbf{P}^{\perp}_{\Lambda} = \mathbf{P}^{\perp}_{\Lambda}\mathbf{P}^{\perp}_{\Lambda} = \mathbf{P}^{\perp}_{\Lambda}. \tag{28}$$

Therefore, (26) holds.

Let

$$\alpha = -\frac{\sqrt{(|S|-l)/N+1}-1}{\sqrt{(|S|-l)/N}}, \tag{29}$$

it is easy to obtain

$$\frac{2\alpha}{1-\alpha^2} = -\sqrt{\frac{|S|-l}{N}},\ \frac{1+\alpha^2}{1-\alpha^2} = -\sqrt{\frac{|S|-l}{N}+1}. \tag{30}$$

Next, let $W=\{\tau_1,\tau_2,\cdots,\tau_N\}\subseteq S^c$ and define the vector $\mathbf{e}\in\mathbb{R}^{Nd}$ with

$$\mathbf{e}=(e_{11},e_{21},\cdots e_{1d},e_{21},e_{21},\cdots,e_{2d},\cdots,e_{N1},e_{N2},\cdots e_{Nd})^{\mathrm{T}},\tag{31}$$

and

$$e_{ij}=\frac{\left(\mathbf{\Phi}^{\mathrm{T}}[i]\right)_j\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}}{\left\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2},\ 1\le i\le N,1\le j\le d.\tag{32}$$

It is obvious that

$$\mathbf{e}^{\mathrm{T}}[j]\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}=\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2,\tag{33}$$

which gives

$$\mathbf{e}^{\mathrm{T}}\mathbf{\Phi}_W^{\mathrm{T}}\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}=\sum_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2.\tag{34}$$

What's more, define

$$\mathbf{B}=\mathbf{P}_{\Lambda}^{\perp}[\mathbf{\Phi}_{S-\Lambda},\mathbf{\Phi}_W],\tag{35}$$

$$\mathbf{u}=\begin{bmatrix}\mathbf{x}_{S-\Lambda} & \mathbf{0}\end{bmatrix}^{\mathrm{T}}\in\mathbb{R}^{(|S-\Lambda|+N)d},\tag{36}$$

$$\omega=\begin{bmatrix}\mathbf{0} & \dfrac{\alpha\|\mathbf{x}_{S-\Lambda}\|_2}{\sqrt{N}}\mathbf{e}\end{bmatrix}^{\mathrm{T}}\in\mathbb{R}^{(|S-\Lambda|+N)d}.\tag{37}$$

Then,

$$\mathbf{B}\mathbf{u}=\mathbf{P}_{\Lambda}^{\perp}[\mathbf{\Phi}_{S-\Lambda},\mathbf{\Phi}_W]\mathbf{u}=\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda},\tag{38}$$

and

$$\|\mathbf{u}+\omega\|_2^2=\left(1+\alpha^2\right)\|\mathbf{x}_{S-\Lambda}\|_2^2,\tag{39}$$

$$\left\|\alpha^2\mathbf{u}-\omega\right\|_2^2=\alpha^2\left(1+\alpha^2\right)\|\mathbf{x}_{S-\Lambda}\|_2^2.\tag{40}$$

Thus

$$\begin{aligned}\omega^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{B}\mathbf{u}&\overset{a}{=}\frac{\alpha\|\mathbf{x}_{S-\Lambda}\|_2}{\sqrt{N}}\mathbf{e}^{\mathrm{T}}\mathbf{\Phi}_W^{\mathrm{T}}(\mathbf{P}_{\Lambda}^{\perp})^{\mathrm{T}}\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\\&\overset{b}{=}\frac{\alpha\|\mathbf{x}_{S-\Lambda}\|_2}{\sqrt{N}}\mathbf{e}^{\mathrm{T}}\mathbf{\Phi}_W^{\mathrm{T}}\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\\&\overset{c}{=}\frac{\alpha\|\mathbf{x}_{S-\Lambda}\|_2}{\sqrt{N}}\sum_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2,\end{aligned}\tag{41}$$

where (a) follows from (35) and (37) (38); (b) and (c) follow from (28) and (34), respectively. By induction, we have

$$\begin{aligned}&\|\mathbf{B}(\mathbf{u}+\omega)\|_2^2-\left\|\mathbf{B}(\alpha^2\mathbf{u}-\omega)\right\|_2^2\\&=(1-\alpha^4)\|\mathbf{B}\mathbf{u}\|_2^2+2(1+\alpha^2)\omega^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{B}\mathbf{u}\\&=(1-\alpha^4)\left(\|\mathbf{B}\mathbf{u}\|_2^2+\frac{2}{1-\alpha^2}\omega^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}\mathbf{B}\mathbf{u}\right)\\&=(1-\alpha^4)\left(\|\mathbf{B}\mathbf{u}\|_2^2+\frac{2\alpha}{1-\alpha^2}\frac{\|\mathbf{x}_{S-\Lambda}\|_2}{\sqrt{N}}\right.\\&\left.\sum_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2\right)\\&=(1-\alpha^4)\left(\|\mathbf{B}\mathbf{u}\|_2^2-\frac{\sqrt{|S|-l}\|\mathbf{x}_{S-\Lambda}\|_2}{N}\right.\\&\left.\sum_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2\right),\end{aligned}\tag{42}$$

on the other hand,

$$\begin{aligned}&\|\mathbf{B}(\mathbf{u}+\omega)\|_2^2-\left\|\mathbf{B}\left(\alpha^2\mathbf{u}-\omega\right)\right\|_2^2\\&\overset{a}{\ge}\left(1-\delta_{N(k+1)+|S|-l}\right)\|\mathbf{u}+\omega\|_2^2-\left(1+\delta_{N(k+1)+|S|-l}\right)\left\|\alpha^2\mathbf{u}-\omega\right\|_2^2\\&=\left(1-\delta_{N(k+1)+|S|-l}\right)\left(1+\alpha^2\right)\|\mathbf{x}_{S-\Lambda}\|_2^2\\&-\left(1+\delta_{N(k+1)+|S|-l}\right)\alpha^2(1+\alpha^2)\|\mathbf{x}_{S-\Lambda}\|_2^2\\&=(1+\alpha^2)\|\mathbf{x}_{S-\Lambda}\|_2^2\left[\left(1-\delta_{N(k+1)+|S|-l}\right)-\left(1+\delta_{N(k+1)+|S|-l}\right)\alpha^2\right]\\&=(1+\alpha^2)\|\mathbf{x}_{S-\Lambda}\|_2^2\left[(1-\alpha^2)-\delta_{N(k+1)+|S|-l}(1+\alpha^2)\right]\\&=(1-\alpha^4)\|\mathbf{x}_{S-\Lambda}\|_2^2\left(1-\frac{1+\alpha^2}{1-\alpha^2}\delta_{N(k+1)+|S|-l}\right)\\&\overset{b}{=}(1-\alpha^4)\|\mathbf{x}_{S-\Lambda}\|_2^2\left(1-\sqrt{(|S|-l)/N+1}\delta_{N(k+1)+|S|-l}\right),\end{aligned}\tag{43}$$

where (a) follows from (35) and Lemma 3 (note that $|S\cap\Lambda|=l,|W|=N$, and $\Lambda=kN$, thus $|\Lambda\cup((S-\Lambda)\cup W)|=N(k+1)+|S|-l$), (b) follows from the equality (30).

Combining (38) (42) (43), we obtain

$$\begin{aligned}&\left\|\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2^2-\frac{\sqrt{|S|-l}\|\mathbf{x}_{S-\Lambda}\|_2}{N}\sum_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2\\&\ge\|\mathbf{x}_{S-\Lambda}\|_2^2\left(1-\sqrt{(|S|-l)/N+1}\delta_{N(k+1)+|S|-l}\right).\end{aligned}\tag{44}$$

Using (26), we have

$$\begin{aligned}&\left\|\mathbf{\Phi}_{S-\Lambda}^{\mathrm{T}}\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_{2,\infty}-\frac{1}{N}\sum_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda}^{\perp}\mathbf{\Phi}_{S-\Lambda}\mathbf{x}_{S-\Lambda}\right\|_2\\&\ge\frac{(1-\sqrt{(|S|-l)/N+1}\delta_{N(k+1)+|S|-l})\|\mathbf{x}_{S-\Lambda}\|_2}{\sqrt{|S|-l}},\end{aligned}\tag{45}$$

which completes the proof.

Appendix B. Proof of Theorem 1

In the following, we follow the proofs of [25, Theorem 1] and [32, Theorem 1] to prove this theorem.

Suppose that BgOMP algorithm chooses at least one block index in the first k iterations, that is to say, $|S\cap\Lambda^k|=l\ge k$. We assume $S\not\subseteq\Lambda^k$ (i.e., $l\le|S|-1$) and the BgOMP algorithm performs at least $k+1$ iterations, otherwise, the result holds. Then, we need to show that $(\Lambda^{k+1}-\Lambda^k)\cap S\ne\emptyset$. Since $\Lambda^0=\emptyset$, the induction assumption $|S|>|\Lambda^k\cap S|\ge k$ holds with $k=0$. Thus, $k=0$ is also contained in this proof.

Let

$$W=\{\tau_1,\tau_2,\cdots,\tau_N\}\subseteq S^c,\tag{46}$$

such that

$$\begin{aligned}&\left\|\mathbf{\Phi}^{\mathrm{T}}[\tau_1]\mathbf{r}^k\right\|_2\ge\left\|\mathbf{\Phi}^{\mathrm{T}}[\tau_2]\mathbf{r}^k\right\|_2\ge\cdots\ge\left\|\mathbf{\Phi}^{\mathrm{T}}[\tau_N]\mathbf{r}^k\right\|_2\\&\ge\left\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{r}^k\right\|_2(i\in S^c-W),\end{aligned}\tag{47}$$

where τ_i denotes the ith block index.

Since $(\Lambda^{k+1}-\Lambda^k)\cap S\ne\emptyset$, we should show

$$\max_{i\in S}\left\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{r}^k\right\|_2>\left\|\mathbf{\Phi}^{\mathrm{T}}[\tau_N]\mathbf{r}^k\right\|_2,\tag{48}$$

by using (47), we can easily obtain

$$\left\|\mathbf{\Phi}^{\mathrm{T}}[\tau_N]\mathbf{r}^k\right\|_2<\frac{1}{N}\sum_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{r}^k\right\|_2.\tag{49}$$

Therefore, it suffices to show

$$\max_{i\in S}\left\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{r}^k\right\|_2 > \frac{1}{N}\sum_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{r}^k\right\|_2. \tag{50}$$

By Algorithm 1, we have

$$\begin{aligned}\mathbf{r}^k&=\mathbf{y}-\mathbf{\Phi}_{\Lambda^k}\hat{\mathbf{x}}_{\Lambda^k}=\left(\mathbf{I}-\mathbf{\Phi}_{\Lambda^k}(\mathbf{\Phi}_{\Lambda^k}^{\mathrm{T}}\mathbf{\Phi}_{\Lambda^k})^{-1}\mathbf{\Phi}_{\Lambda^k}^{\mathrm{T}}\right)\mathbf{y}\\&=\mathbf{P}_{\Lambda^k}^{\perp}(\mathbf{\Phi}\mathbf{x}+\mathbf{v})=\mathbf{P}_{\Lambda^k}^{\perp}(\mathbf{\Phi}_S\mathbf{x}_S+\mathbf{v})\\&=\mathbf{P}_{\Lambda^k}^{\perp}\left(\mathbf{\Phi}_{S\cap\Lambda^k}\mathbf{x}_{S\cap\Lambda^k}+\mathbf{\Phi}_{S-\Lambda^k}\mathbf{x}_{S-\Lambda^k}+\mathbf{v}\right)\\&\overset{a}{=}\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{\Phi}_{S-\Lambda^k}\mathbf{x}_{S-\Lambda^k}+\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{v},\end{aligned} \tag{51}$$

where (a) have used the result that $\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{\Phi}_{\Lambda^k}=0$.

In addition, for each block index $i\in\Lambda^k$, $\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{r}^k\|_2=0$. Thus, from (51) and the triangular inequality, we have

$$\begin{aligned}&\max_{i\in S}\left\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{r}^k\right\|_2=\max_{i\in S-\Lambda^k}\left\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{r}^k\right\|_2\\&>\max_{i\in S-\Lambda^k}\left(\left\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{\Phi}_{S-\Lambda^k}\mathbf{x}_{S-\Lambda^k}\right\|_2-\left\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{v}\right\|\right)_2,\end{aligned} \tag{52}$$

$$\begin{aligned}\frac{1}{N}\sum_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{r}^k\right\|_2&<\frac{1}{N}\sum_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{\Phi}_{S-\Lambda^k}\mathbf{x}_{S-\Lambda^k}\right\|_2\\&\quad+\max_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{v}\right\|_2.\end{aligned} \tag{53}$$

To show (50), we have

$$\beta_1>\beta_2, \tag{54}$$

where

$$\begin{aligned}\beta_1&=\max_{i\in S-\Lambda^k}\left\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{\Phi}_{S-\Lambda^k}\mathbf{x}_{S-\Lambda^k}\right\|_2\\&\quad-\frac{1}{N}\sum_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{\Phi}_{S-\Lambda^k}\mathbf{x}_{S-\Lambda^k}\right\|_2\\&=\left\|\mathbf{\Phi}_{S-\Lambda^k}^{\mathrm{T}}\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{\Phi}_{S-\Lambda^k}\mathbf{x}_{S-\Lambda^k}\right\|_{2,\infty}\\&\quad-\frac{1}{N}\sum_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{\Phi}_{S-\Lambda^k}\mathbf{x}_{S-\Lambda^k}\right\|_2,\end{aligned} \tag{55}$$

$$\beta_2=\max_{i\in S-\Lambda^k}\left\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{v}\right\|_2+\max_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{v}\right\|_2. \tag{56}$$

Clearly, there exist $i_0\in S-\Lambda^k$ and $j_0\in W$ such that

$$\max_{i\in S-\Lambda^k}\left\|\mathbf{\Phi}^{\mathrm{T}}[i]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{v}\right\|_2=\left\|\mathbf{\Phi}^{\mathrm{T}}[i_0]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{v}\right\|_2, \tag{57}$$

$$\max_{j\in W}\left\|\mathbf{\Phi}^{\mathrm{T}}[j]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{v}\right\|_2=\left\|\mathbf{\Phi}^{\mathrm{T}}[j_0]\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{v}\right\|_2. \tag{58}$$

Therefore,

$$\begin{aligned}\beta_2&=\left\|[\mathbf{\Phi}[i_0],\mathbf{\Phi}[j_0]]^{\mathrm{T}}\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{v}\right\|_{2,1}\le\sqrt{2}\left\|[\mathbf{\Phi}[i_0],\mathbf{\Phi}[j_0]]^{\mathrm{T}}\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{v}\right\|_2\\&\le\sqrt{2\left(1+\delta_{N(k+1)+|S|-k}\right)}\|\mathbf{v}\|_2,\end{aligned} \tag{59}$$

where we have used Lemma 2 and

$$\left\|\mathbf{P}_{\Lambda^k}^{\perp}\mathbf{v}\right\|_2\le\left\|\mathbf{P}_{\Lambda^k}^{\perp}\right\|_2\|\mathbf{v}\|_2\le\|\mathbf{v}\|_2\le\varepsilon. \tag{60}$$

In the following, we will give the lower bound on β_1. As shown in Algorithm 1, $|\Lambda^k|=kN$. It follows from the assumption of Lemma 4, we have

$$0\le k\le\left|S\cap\Lambda^k\right|=l\le|S|-1. \tag{61}$$

Note that $W\subset S^c$ and $|W|=N$. By using Lemma 4 and (55), we obtain

$$\begin{aligned}\beta_1&\ge\frac{\left(1-\sqrt{(|S|-l)/N+1}\delta_{N(k+1)+|S|-l}\right)\|\mathbf{x}_{S-\Lambda^k}\|_2}{\sqrt{|S|-l}}\\&\ge\frac{\left(1-\sqrt{|S|/N+1}\delta_{N(k+1)+|S|-k}\right)\|\mathbf{x}_{S-\Lambda^k}\|_2}{\sqrt{|S|-l}},\end{aligned} \tag{62}$$

where the second inequality holds owning to Lemma 1. By the fact that $|S\cap\Lambda^k|=l$, we have

$$\begin{aligned}\|\mathbf{x}_{S-\Lambda^k}\|_2&>\frac{\sum_{i\in S-\Lambda^k}\|\mathbf{x}[i]\|_2}{\sqrt{|S|-l}}>\sqrt{|S|-l}\min_{i\in S-\Lambda^k}\|\mathbf{x}[i]\|_2\\&>\sqrt{|S|-l}\min_{i\in S}\|\mathbf{x}[i]\|_2.\end{aligned} \tag{63}$$

By using (62) (63), we have

$$\beta_1\ge\left(1-\sqrt{|S|/N+1}\delta_{N(k+1)+|S|-k}\right)\min_{i\in S}\|\mathbf{x}[i]\|_2. \tag{64}$$

Combining (59) and (64) leads to the following result

$$\min_{i\in S}\|\mathbf{x}[i]\|_2>\frac{\sqrt{2\left(1+\delta_{N(k+1)+|S|-k}\right)}\varepsilon}{1-\sqrt{|S|/N+1}\delta_{N(k+1)+|S|-k}}. \tag{65}$$

It is obvious that (65) is guaranteed if

$$\min_{i\in S}\|\mathbf{x}[i]\|_2>\frac{2\varepsilon}{1-\sqrt{|S|/N+1}\delta_{N(k+1)+|S|-k}}, \tag{66}$$

which completes the proof.

References

[1] E.J. Candes, M.B. Wakin, An introduction to compressive sampling, IEEE Signal Process. Mag. 25 (2) (2008) 21–30.
[2] D.L. Donoho, Compressed sensing, IEEE Trans. Inf. Theory 52 (4) (2006) 1289–1306.
[3] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, SIAM J. Sci. Comput. 20 (1) (1999) 33–61.
[4] H. Mohimani, M. Babaie-Zadeh, C. Jutten, A fast approach for overcomplete sparse decomposition based on smoothed l_0 norm, IEEE Trans. Signal Process 57 (1) (2009) 289–301.
[5] J.A. Tropp, A.C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, IEEE Trans. Inf. Theory 53 (12) (2007) 4655–4666.
[6] D. Needell, R. Vershynin, Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit, Found. Comput. Math. 9 (3) (2009) 317–334.
[7] W. Dai, O. Milenkovic, Subspace pursuit for compressive sensing signal reconstruction, IEEE IEEE Trans. Inf. Theory 55 (5) (2009) 2230–2249.
[8] D.L. Donoho, Y. Tsaig, I. Drori, J.C. Starck, Sparse solutions of underdetermined linear equations by stagewise orthogonal matching pursuit, IEEE Trans. Inf. Theory 58 (2) (2012) 1094–1120.
[9] D. Needell, J.A. Tropp, CoSaMP: iterative signal recovery from incomplete and inaccurate samples, Appl. Comput. Harmon. Anal. 26 (3) (2009) 301–321.
[10] J. Wang, S. Kon, B. Shim, Generalized, orthogonal matching pursuit, IEEE Trans. Signal Process 60 (12) (2012) 6202–6216.
[11] M. Davenport, M. Wakin, Analysis of orthogonal matching pursuit using the restricted isometry property, IEEE Trans. Inf. Theory 56 (9) (2010) 4395–4401.
[12] E. Liu, V. Temlyakov, The orthogonal super greedy algorithm and applications in compressed sensing, IEEE Trans. Inf. Theory 58 (4) (2012) 2040–2047.
[13] Q. Mo, S. Yi, A remark on the restricted isometry property in orthogonal matching pursuit, IEEE Trans. Inf. Theory 58 (6) (2012) 3654–3656.
[14] J. Wang, B. Shim, On the recovery limit of sparse signals using orthogonal matching pursuit, IEEE Trans. Signal Process 60 (9) (2012) 4973–4976.
[15] L.H. Chang, J.Y. Wu, An improved RIP-based performance guarantee for sparse signal recovery via orthogonal matching pursuit, IEEE Trans. Inf. Theory 60 (9) (2014) 5702–5715.
[16] Q. Mo, A sharp restricted isometry constant bound of orthogonal matching pursuit, 2015. arXiv:1501.01708.
[17] J.M. Wen, X.M. Zhu, D.F. Li, Improved bounds on the restricted isometry constant for orthogonal matching pursuit, Electron. Lett. 49 (23) (2013) 1487–1489.
[18] Y. Shen, S. Li, Sparse signals recovery from noisy measurements by orthogonal matching pursuit, Inverse Probl. Imaging 9 (1) (2015) 231–238.
[19] R. Wu, W. Huang, D.R. Chen, The exact support recovery of sparse signals with noise via orthogonal matching pursuit, IEEE Signal Process. Lett. 20 (4) (2013) 403–406.

[20] J.M. Wen, Z.C. Zhou, J. Wang, X.H. Tang, Q. Mo, A sharp condition for exact support recovery with orthogonal matching pursuit, IEEE Trans. Signal Process 65 (6) (2017) 1370–1382.
[21] S. Satpathi, R.L. Das, M. Chakraborty, Improving the bound on the RIP constant in generalized orthogonal matching pursuit, IEEE Signal Process. Lett. 20 (11) (2013) 1074–1077.
[22] Y. Shen, B. Li, W.L. Pan, J. Li, Analysis of generalized Orthogonal Matching Pursuit using Restricted Isometry Constant, Electron. Lett. 50 (14) (2014) 1020–1022.
[23] B. Li, Y. Shen, S. Rajan, T. Kirubarajan, Theoretical results for sparse signal recovery with noises using generalized OMP algorithm, Signal Process 117 (2015) 270–278.
[24] B. Li, Y. Shen, Z.H. Wu, J. Li, Sufficient conditions for generalized Orthogonal Matching Pursuit in noisy case, Signal Process 108 (2015) 111–123.
[25] J.M. Wen, Z.C. Zhou, D.F. Li, X.H. Tang, A novel sufficient condition for generalized orthogonal matching pursuit, IEEE Commun. Lett. 21 (4) (2017) 805–808.
[26] F. Parvaresh, H. Vikalo, S. Misra, B. Hassibi, Recovering sparse signals using sparse measurement matrices in compressed DNA microarrays, IEEE J. Sel. Top. Signal Process 2 (3) (2008) 275–285.
[27] S. Cotter, B. Rao, Sparse channel estimation via matching pursuit with application to equalization, IEEE Trans. Commun. 50 (3) (2002) 374–377.
[28] M. Mishali, Y.C. Eldar, Blind multi-band signal reconstruction: compressed sensing for analog signals, IEEE Trans. Signal Process 57 (3) (2009) 993–1009.
[29] A. Majumdar, R. Ward, Compressed sensing of color images, Signal process 90 (12) (2010) 3122–3127.
[30] Y.C. Eldar, P. Kuppinger, Block-sparse signals: uncertainty relations and efficient recovery, IEEE Trans. Signal Process 58 (6) (2010) 3042–3054.
[31] J. Wang, G. Li, H. Zhang, X.Q. Wang, Analysis of block OMP using block RIP[J], 2011. arXiv:1104.1071.
[32] J.M. Wen, Z.C. Zhou, Z.L. Liu, M.J. Lai, X.H. Tang, Sharp sufficient conditions for stable recovery of block sparse signals by block orthogonal matching pursuit[J], 2016. arXiv:1605.02894.
[33] R.G. Baraniuk, V. Cevher, M.F. Duarte, C. Hegde, Model-based compressive sensing, IEEE Trans. Commun. 56 (4) (2010) 1982–2001.
[34] S. Hamidi-Ghalehjegh, M. Babaie-Zadeh, C. Jutten, Fast Block-Sparse Decomposition based on SL0, 9th international conference, 2010.
[35] A. Kamali, M.R. Aghabozorgi-Sahaf, A.M. Doost-Hooseini, A.A. Tadaion, Block subspace pursuit for block-sparse signal reconstruction, IJST Trans. Electr. Eng. 37 (1) (2013) 1–16.
[36] B.X. Huang, T. Zhou, Recovery of block sparse signals by a block version of StOMP, Signal Process. 106 (2015) 231–244.
[37] Y.C. Eldar, M. Mishali, Robust recovery of signals from a structured union of subspaces, IEEE Trans. Inf. Theory 55 (11) (2009) 5302–5316.
[38] A.L. Goldberger, L. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.K. Peng, H.E. Stanley, et al., Physiobank, physiotoolkit and physionet components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) e215–e215.
[39] W.S. Dong, L. Zhang, R. Lukac, Sparse representation based image interpolation with nonlocal autoregressive modeling, IEEE Trans. Image Process 22 (4) (2013) 1382–1394.