



Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha



Sharp sufficient conditions for stable recovery of block sparse signals by block orthogonal matching pursuit

Jinming Wen^a, Zhengchun Zhou^b, Zilong Liu^c, Ming-Jun Lai^{d,*}, Xiaohu Tang^e^a College of Information Science and Technology, Jinan University, Guangzhou, 510632, China^b School of Mathematics, Southwest Jiaotong University, Chengdu 610031, China^c Institute for Communication Systems (ICS), 5G Innovation Centre (5GIC), University of Surrey, UK^d Department of Mathematics, University of Georgia, Athens, GA 30602, USA^e Information Security and National Computing Grid Laboratory, Southwest Jiaotong University, Chengdu 610031, China

ARTICLE INFO

Article history:

Received 7 July 2016

Received in revised form 8 January 2018

Accepted 2 February 2018

Available online xxxx

Communicated by Patrick Flandrin

Keywords:

Compressed sensing

Block restricted isometry property

Block orthogonal matching pursuit

Block sparse signals

 α -strongly-decaying

ABSTRACT

In this paper, we use the block orthogonal matching pursuit (BOMP) algorithm to recover block sparse signals \mathbf{x} from measurements $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}$, where \mathbf{v} is an ℓ_2 -bounded noise vector (i.e., $\|\mathbf{v}\|_2 \leq \epsilon$ for some constant ϵ). We investigate some sufficient conditions based on the block restricted isometry property (block-RIP) for exact (when $\mathbf{v} = \mathbf{0}$) and stable (when $\mathbf{v} \neq \mathbf{0}$) recovery of block sparse signals \mathbf{x} . First, on the one hand, we show that if \mathbf{A} satisfies the block-RIP with $\delta_{K+1} < 1/\sqrt{K+1}$, then every block K -sparse signal \mathbf{x} can be exactly or stably recovered by BOMP in K iterations. On the other hand, we show that, for any $K \geq 1$ and $1/\sqrt{K+1} \leq \delta < 1$, there exists a matrix \mathbf{A} satisfying the block-RIP with $\delta_{K+1} = \delta$ and a block K -sparse signal \mathbf{x} such that BOMP may fail to recover \mathbf{x} in K iterations. Then, we study some sufficient conditions for recovering block α -strongly-decaying K -sparse signals. We show that if \mathbf{A} satisfies the block-RIP with $\delta_{K+1} < \sqrt{2}/2$, then every α -strongly-decaying block K -sparse signal can be exactly or stably recovered by BOMP in K iterations under some conditions on α . Our newly found sufficient condition on the block-RIP of \mathbf{A} is less restrictive than that for ℓ_1 minimization for this special class of sparse signals. Furthermore, for any $K \geq 1$, $\alpha > 1$ and $\sqrt{2}/2 \leq \delta < 1$, the recovery of \mathbf{x} may fail in K iterations for a sensing matrix \mathbf{A} which satisfies the block-RIP with $\delta_{K+1} = \delta$. Finally, we study some sufficient conditions for partial recovery of block sparse signals. Specifically, if \mathbf{A} satisfies the block-RIP with $\delta_{K+1} < \sqrt{2}/2$, then BOMP is guaranteed to recover some blocks of \mathbf{x} if these blocks satisfy a sufficient condition. We further show that this condition is also sharp.

© 2018 Elsevier Inc. All rights reserved.

* Corresponding author.

E-mail addresses: jinming.wen@mail.mcgill.ca (J. Wen), zzc@home.swjtu.edu.cn (Z. Zhou), zilong.liu@surrey.ac.uk (Z. Liu), mjilai@uga.edu (M.-J. Lai), xhutang@swjtu.edu.cn (X. Tang).

1. Introduction

In many applications, such as computer vision, image reconstruction and blind sources separation (see, e.g., [9,21,14,37,25,63,33]), we need to recover a K -sparse signal $\mathbf{x} \in \mathbb{R}^n$ (i.e., \mathbf{x} has at most K nonzero entries) from the following linear model

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}, \quad (1.1)$$

where $\mathbf{y} \in \mathbb{R}^m$ is an observation signal, $\mathbf{A} \in \mathbb{R}^{m \times n}$ ($m \ll n$) is a known sensing matrix, and $\mathbf{v} \in \mathbb{R}^m$ is an ℓ_2 -bounded noise vector [34,23,7,59] (i.e., $\|\mathbf{v}\|_2 \leq \epsilon$ for some constant ϵ). In addition to the ℓ_2 -bounded noise, there are some other types of noises, such as the Gaussian noise [10] (i.e., $v \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$) and the ℓ_∞ -bounded noise [5] (i.e., $\|\mathbf{A}\mathbf{v}\|_\infty \leq \epsilon$ for some constant ϵ). In this paper, we focus only on the ℓ_2 -bounded noise as our results can be easily extended to the other two types of noises (see, e.g., [5], [60]).

A natural method of recovering \mathbf{x} in (1.1) is to solve the following ℓ_0 -minimization problem:

$$\min \|\mathbf{x}\|_0 : \text{subject to } \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon. \quad (1.2)$$

However, (1.2) is an NP-hard problem [1]. Fortunately, as demonstrated in the pioneering work [9] [22], under certain conditions, one can efficiently solve the following ℓ_1 -minimization problem, instead of solving (2), to stably reconstruct \mathbf{x} :

$$\min \|\mathbf{x}\|_1 : \text{subject to } \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon. \quad (1.3)$$

To analyze the recovery performance of sparse recovery algorithms, a commonly used concept is the so-called restricted isometry property (RIP) [9]. For a sensing matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and for any integer $1 \leq K \leq n$, the K -restricted isometry constant (RIC) $\delta_K \in (0, 1)$ of \mathbf{A} is the smallest constant such that

$$(1 - \delta_K)\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1 + \delta_K)\|\mathbf{x}\|_2^2 \quad (1.4)$$

for all K -sparse signals \mathbf{x} . Another commonly used concept is the mutual coherence [24] which is defined as

$$\mu = \max_{i \neq j} \frac{|\mathbf{A}_i^\top \mathbf{A}_j|}{\|\mathbf{A}_i\|_2 \|\mathbf{A}_j\|_2}, \quad (1.5)$$

where \mathbf{A}_i ($1 \leq i \leq n$) is the i -th column of \mathbf{A} .

There have been a variety of sufficient conditions for exactly (when $\|\mathbf{v}\|_2 = 0$) or stably recovering K -sparse signals \mathbf{x} (when $\|\mathbf{v}\|_2 \neq 0$) by solving (1.3). For example, $\delta_{2K} < \sqrt{2} - 1$ [7], $\delta_{2K} < 0.4531$ [32], $\delta_{2K} < 0.4652$ [30] and $\delta_{2K} < \sqrt{2}/2$ [6]. Moreover, it was shown in [6] and [19] that the exact recovery of \mathbf{x} may not be possible if $\delta_{2K} \geq \sqrt{2}/2$ which implies that $\delta_{2K} < \sqrt{2}/2$ [6] is a sharp sufficient condition for stably recovering K -sparse signals \mathbf{x} via solving (1.3).

In addition to the ℓ_1 minimization approach (1.3), the orthogonal matching pursuit (OMP) algorithm [52] is another widely-used algorithm to solve (1.2) due to its efficiency in computation. Generally speaking, solving (1.2) by the OMP algorithm is faster than solving (1.3) by the interior point method or simplex method or their variants. Thus, the OMP algorithm has attracted much attention. There are many sufficient conditions for exactly or stably recovering \mathbf{x} by using the OMP algorithm, see, e.g., [18,41,47,56,11]. Recently, it was shown in [60] [46] that if $\delta_{K+1} < 1/\sqrt{K+1}$, the exact recovery of K -sparse signals \mathbf{x} can be guaranteed by using the OMP algorithm in K iterations. Some necessary conditions on the recovery capability of the OMP algorithm can be found in [47] [56]. In particular, it was shown in [61] that if $\delta_{K+1} \geq 1/\sqrt{K+1}$, the OMP algorithm may fail to recover a K -sparse signal \mathbf{x} in K iterations.

It is easy to see that the sharp sufficient condition $\delta_{K+1} < 1/\sqrt{K+1}$ for the OMP algorithm is much stronger than the condition $\delta_{2K} < \sqrt{2}/2$ for (1.3). This leads to at least two interesting directions of research. One is to use more than K iterations in the OMP algorithm to recover a sparse signal and find a sufficient condition on $\delta_{c(K+1)}$ (for some constant c) to ensure the exact recovery as discussed in [66,42,31,43,15,57]. These results can be described as follows. When $\delta_{c(K+1)} < \delta^* < 1$ for some constant δ^* , any K -sparse signal can be exactly or stably recovered in cK iterations by the OMP algorithm. For example, in [66], $\delta^* = 1/3$ with $c = 30$. As explained in [15], c is dependent on δ^* . When $\delta^* < 1$ is bigger, so is c . When $c > 1$ is very large, the computational advantage of the OMP will lose. Another interesting direction is to characterize a class of K -sparse signals which can be exactly or stably recovered by the OMP algorithm in K iterations when $\delta_{K+1} < \sqrt{2}/2$ in order to maintain the computational advantage. It was shown in [18] that $\delta_{K+1} < 1/3$ is a sufficient condition for recovering α -strongly-decaying K -sparse signals (see below for the definition of α -strongly-decaying K -sparse signal) by using the OMP algorithm in K iterations under some conditions on α in noise-free case. This condition was extended to noise-corrupted models in [20]. Recently, Xu showed in [65] that under some conditions on α , $\delta_{K+1} < \sqrt{2} - 1$ is a sufficient condition. Recently, Herzet et al. showed [38] that under other sufficient condition on α , the mutual coherence $\mu < 1/K$ is a sharp condition to recover α -strongly-decaying signals.

Signals with decaying property are frequently encountered in practical applications, such as in speech communication [36] and audio source separation [54]. Without loss of generality, let us assume that all the entries of a K -sparse signal \mathbf{x} are ordered as

$$|x_1| \geq |x_2| \geq \dots \geq |x_K| \geq 0$$

and $|x_j| = 0$ for $K+1 \leq j \leq n$. Then, a K -sparse signal \mathbf{x} is called an α -strongly-decaying signal [18] ($\alpha > 1$) if

$$|x_i| \geq \alpha |x_{i+1}|, \quad 1 \leq i \leq K-1.$$

Thus, two questions are naturally raised: (1) what is the RIC based sharp condition of exactly or stably recovery of α -strongly-decaying K -sparse signals by using the OMP algorithm in K iterations under some conditions on α ? (2) Is the condition on α sharp? We shall answer these two questions by Theorems 3 and 4, and Corollary 4 in Section 3 of this paper.

In this paper, we focus on a generalized version of OMP, the Block OMP (BOMP) algorithm, which was independently proposed in [27] and [50], respectively. The main reason that we focus on the BOMP algorithm is that it has better reconstruction properties than the standard OMP algorithm as shown in [27]. Moreover, recovering block-sparse signals are frequently encountered, such as recovering signals that lie in unions of subspaces [4,28,26,35], reconstructing multiband signals [44,45], face recognition [64], recovering multiple measurement vectors [16,12,53,40] and clustering of data multiple subspaces [29].

To define block-sparsity, \mathbf{x} is viewed as a concatenation of blocks. Similar to [27], we assume \mathbf{x} consists of L blocks each having identical length of d (i.e., $n = Ld$). Then, \mathbf{x} can be expressed as

$$\mathbf{x} = [\mathbf{x}^\top[1], \mathbf{x}^\top[2], \dots, \mathbf{x}^\top[L]]^\top,$$

where

$$\mathbf{x}[i] = [x_{d(i-1)+1}, x_{d(i-1)+2}, \dots, x_{di}]^\top, \quad 1 \leq i \leq L.$$

A signal $\mathbf{x} \in \mathbb{R}^n$ is called block K -sparse if at most K blocks $\mathbf{x}[i], i = 1, \dots, K$, are not zero vectors [27]. Mathematically, a signal \mathbf{x} is block K -sparse if and only if

$$\sum_{i=1}^L I(\|\mathbf{x}[i]\|_2 > 0) \leq K,$$

where $I(\cdot)$ is the indicator function (i.e., it takes the value of 0 when its argument is zero and 1 otherwise). When the block size is 1, block-sparsity reduces to the conventional sparsity as defined in [8,21]. In the rest of the paper, the conventional sparsity will be simply called sparsity, in contrast to block-sparsity. Similarly, a sensing matrix \mathbf{A} can also be represented as a concatenation of column blocks, i.e.,

$$\mathbf{A} = [\mathbf{A}[1], \mathbf{A}[2], \dots, \mathbf{A}[L]],$$

where

$$\mathbf{A}[i] = [\mathbf{A}_{d(i-1)+1}, \mathbf{A}_{d(i-1)+2}, \dots, \mathbf{A}_{di}], \quad 1 \leq i \leq L.$$

For any ordered set $S \subset \{1, 2, \dots, L\}$ with ascending entries, let $\mathbf{A}[S]$ denote the submatrix of \mathbf{A} that contains only the blocks indexed by S and $\mathbf{x}[S]$ denote the subvector of \mathbf{x} that contains only the blocks indexed by S . For instance, if $S = \{1, 3, 4\}$, we have $\mathbf{A}[S] = [\mathbf{A}[1], \mathbf{A}[3], \mathbf{A}[4]]$ and $\mathbf{x}[S] = [\mathbf{x}^\top[1], \mathbf{x}^\top[3], \mathbf{x}^\top[4]]^\top$. Then, the BOMP algorithm can be formally described below in Algorithm 1.

Algorithm 1 The BOMP algorithm ([27], [50]).

Input: measurements \mathbf{y} , sensing matrix \mathbf{A} , sparsity K and the number of blocks L .

Initialize: $k = 0, \mathbf{r}^0 = \mathbf{y}, S_0 = \emptyset$.

Iterate the five steps below until the stopping criterion is met.

- 1: $k = k + 1$,
- 2: $s^k = \arg \max_{1 \leq i \leq L} \|\mathbf{A}[i]^\top \mathbf{r}^{k-1}\|_2$,
- 3: $S_k = S_{k-1} \cup \{s^k\}$,
- 4: $\hat{\mathbf{x}}[S_k] = \arg \min \|\mathbf{y} - \mathbf{A}[S_k]\mathbf{x}\|_2$,
- 5: $\mathbf{r}^k = \mathbf{y} - \mathbf{A}[S_k]\hat{\mathbf{x}}[S_k]$.

Output: $\hat{\mathbf{x}} = \arg \min_{\mathbf{x}: \text{supp}(\mathbf{x}) \subset S_K} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2$.

To analyze the performance of algorithms for block sparse signals, the classic RIP was extended to the block-RIP in [28]. Specifically, \mathbf{A} is said to satisfy the block-RIP with parameter $\delta_{BK} \in (0, 1)$ if

$$(1 - \delta_{BK})\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1 + \delta_{BK})\|\mathbf{x}\|_2^2 \tag{1.6}$$

for all block K sparse signals \mathbf{x} , where the smallest constant δ_{BK} is called as the block-RIC of \mathbf{A} . For convenience, we simply denote δ_{BK} by δ_K in the rest of this paper. It is worth mentioning that if the entries of \mathbf{A} independently and identically follow the Gaussian distribution $\mathcal{N}(0, 1/n)$ or the Bernoulli distribution namely $a_{ij} = \pm 1/\sqrt{n}$ with equal probability, then it is of overwhelming probability that \mathbf{A} satisfies the block-RIP for certain m and n , which respectively denote the row and column numbers of \mathbf{A} (see [28, Proposition 3] and the second paragraph above [28, Corollary 3]).

Similarly, we assume the block entries of $\mathbf{x} \in \mathbb{R}^n$ are ordered as

$$\|\mathbf{x}[1]\|_2 \geq \|\mathbf{x}[2]\|_2 \geq \dots \geq \|\mathbf{x}[K]\|_2 \geq 0 \tag{1.7}$$

with $\|\mathbf{x}[j]\|_2 = 0$ for $K < j \leq L$. In the following, we define the block α -strongly-decaying K -sparse signals, which is an extension of the α -strongly-decaying K -sparse explained above.

Definition 1. Let $\alpha > 1$, then a block K -sparse signal $\mathbf{x} \in \mathbb{R}^n$ is called a block α -strongly-decaying K -sparse signal if

$$\|\mathbf{x}[j]\|_2 \geq \alpha \|\mathbf{x}[j+1]\|_2, \quad 1 \leq j \leq K-1. \quad (1.8)$$

With the above explanation of the concepts and notations, we now summarize the main contributions of this paper as follows.

- If \mathbf{A} satisfies the block-RIP with $\delta_{K+1} < 1/\sqrt{K+1}$, then every block K -sparse signal can be exactly or stably recovered by the BOMP algorithm in K iterations, see Corollaries 1 and 2. Moreover, for any $K \geq 1$ and $1/\sqrt{K+1} \leq \delta < 1$, we show that there exist a matrix \mathbf{A} satisfying the block-RIP with $\delta_{K+1} = \delta$ and a block K -sparse signal \mathbf{x} such that the BOMP may fail to recover \mathbf{x} in K iterations, see Theorem 2. These results extend the existing results in [46,60,61] from $d = 1$ to any $d \geq 1$ in a non-trivial way.
- If \mathbf{A} satisfies the block-RIP with $\delta_{K+1} < \sqrt{2}/2$, then under some sufficient conditions on α , every block α -strongly-decaying K -sparse signal can be exactly or stably recovered by the BOMP Algorithm in K iterations, see Theorems 3 and Corollary 3. In addition, for any $K \geq 1$, $\alpha > 1$ and $\sqrt{2}/2 \leq \delta < 1$, we show that there exist a matrix \mathbf{A} satisfying the block-RIP with $\delta_{K+1} = \delta$ and an α -strongly-decaying block K -sparse signal \mathbf{x} such that the BOMP Algorithm may fail to recover \mathbf{x} in K iterations, see Theorem 4. Furthermore, the condition on α is also shown to be tight, see Corollary 4. Obviously, when the block size is 1, then the result for BOMP reduces to that for the standard OMP. Thus, it improves the result in [18] from $\delta_{K+1} < 1/3$ (for OMP) to $\delta_{K+1} \leq \sqrt{2}/2$ (for BOMP).
- Finally, we study partial recovery of block K -sparse signals. We show that if \mathbf{A} satisfies the block-RIP with $\delta_{K+1} < \sqrt{2}/2$ and $\|\mathbf{x}[j]\|_2/\|\mathbf{x}[j+1]\|_2, 1 \leq j \leq i$ satisfy a sufficient condition for some $1 \leq i \leq K$ (note that if $i = L$, this class of block K -sparse signals are block α -strongly-decaying K -sparse signals for certain α), then the BOMP algorithm chooses an index such that the corresponding subvector of \mathbf{x} is a nonzero vector in each of the first i iterations, see Theorem 5 and Corollary 5. Moreover, we show that this condition is tight, see Corollary 6. To the best of our knowledge, partial recovery of sparse signals with BOMP was not reported before.

As far as we know, there is not any reference which studies the recovery of α -strongly-decaying sparse vectors by solving the ℓ_1 -minimization problem (1.3). Thus, so far, the only known sharp sufficient condition for recovering α -strongly-decaying sparse vectors by solving the ℓ_1 -minimization problem (1.3) is still $\delta_{2K} < \sqrt{2}/2$. Thus, our sufficient condition $\delta_{K+1} < \sqrt{2}/2$ for using the BOMP algorithm to recover α -strongly-decaying sparse vectors is less restrictive. Also, using the BOMP algorithm, the recovery is done in K iterations which shows its efficiency.

It was shown in [38] that under a sufficient condition on α , $\mu < 1/K$ is a sharp condition for recovering α -strongly-decaying sparse vectors with OMP in K iterations. The condition requires $m \geq \mathcal{O}(K^2 \log n)$ observations (as pointed out in [38], this can be seen by invoking the Welch bound $\mu \geq \sqrt{\frac{n-m}{m(n-1)}}$ [33, Theorem 5.7]). Our sharp condition is $\delta_{K+1} < \sqrt{2}/2$ and only requires $m \geq \mathcal{O}(K \log(n/K))$ (see [3, Theorem 5.2] or [42, equation (1.3)]), and hence much less observations are needed. These further show the effectiveness of the BOMP algorithm. It is worth mentioning that although finding the mutual coherence of a given $m \times n$ matrix can be achieved in polynomial time of n according to its definition, it is an NP-hard problem to find the RIC of a given matrix [2] [51].

The rest of the paper is organized as follows. In section 2, we define some notations and give some useful lemmas that will be needed to prove our main results. We present our main results in Section 3 and prove them in Section 4. Finally, we summarize this paper and present some future research problems in Section 5.

2. Preliminaries

In this section, we first introduce some notations which will be used in the rest of this paper.

Notations: Let \mathbb{R} be the real field. Boldface lowercase letters denote column vectors, and boldface uppercase letters denote matrices, e.g., $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{m \times n}$. For a vector \mathbf{x} , x_i denotes the i -th entry of \mathbf{x} . Let \mathbf{e}_k denote the k -th column of an identity matrix \mathbf{I} and $\mathbf{0}$ denote a zero matrix or a zero vector. Let Ω be the set of all the indexes i such that $\mathbf{x}[i]$ are not zero vectors for block K -sparse signals \mathbf{x} , then $|\Omega| \leq K$, where $|\Omega|$ is the cardinality of Ω . Let $\Omega \setminus S = \{k | k \in \Omega, k \notin S\}$ for set S . Let Ω^c and S^c be the complement of Ω and S , i.e., $\Omega^c = \{1, 2, \dots, L\} \setminus \Omega$, and $S^c = \{1, 2, \dots, L\} \setminus S$. Let $\mathbf{A}[S]$ be the submatrix of \mathbf{A} that contains only the blocks of columns indexed by S , and $\mathbf{x}[S]$ be the subvector of \mathbf{x} that contains only the blocks indexed by S , and $\mathbf{A}^\top[S]$ be the transpose of $\mathbf{A}[S]$. For any full column rank matrix $\mathbf{A}[S]$, let $\mathbf{P}[S] = \mathbf{A}[S](\mathbf{A}^\top[S]\mathbf{A}[S])^{-1}\mathbf{A}^\top[S]$ and $\mathbf{P}^\perp[S] = \mathbf{I} - \mathbf{P}[S]$ denote the projector and orthogonal complement projector on the column space of $\mathbf{A}[S]$, respectively.

Next we introduce the ℓ_2/ℓ_p -norm for block vectors and some useful lemmas.

2.1. ℓ_2/ℓ_p -norm

Like in [27] and [40], for $\mathbf{x} \in \mathbb{R}^n$, we define a general mixed ℓ_2/ℓ_p -norm (where $p = 1, 2, \infty$) as

$$\|\mathbf{x}\|_{2,p} = \|\mathbf{w}\|_p, \quad (2.1)$$

where $\mathbf{w} \in \mathbb{R}^L$ with $w_\ell = \|\mathbf{x}[\ell]\|_2$ for $1 \leq \ell \leq L$. Note that $\|\mathbf{x}\|_{2,2} = \|\mathbf{x}\|_2$. Thus, we use $\|\mathbf{x}\|_2$ instead of $\|\mathbf{x}\|_{2,2}$ for short. Moreover, if the block size d is equal to 1, for $p = 1, 2, \infty$, we have $\|\mathbf{x}\|_{2,p} = \|\mathbf{x}\|_p$.

2.2. Some useful lemmas

We can easily get the following lemmas by extending [17, Lemma 1], [48, Proposition 3.1] and [49, Lemma 1] for $d = 1$ to general d (note that d is the length of the blocks).

Lemma 1. *If \mathbf{A} satisfies the block-RIP of orders K_1 and K_2 with $K_1 < K_2$, then $\delta_{K_1} \leq \delta_{K_2}$.*

Lemma 2. *Let \mathbf{A} satisfy the block-RIP of order K and S be a set with $|S| \leq K$. Then for any $\mathbf{x} \in \mathbb{R}^m$,*

$$\|\mathbf{A}^\top[S]\mathbf{x}\|_2^2 \leq (1 + \delta_K)\|\mathbf{x}\|_2^2.$$

Lemma 3. *Let sets S_1, S_2 satisfy $|S_2 \setminus S_1| \geq 1$ and matrix \mathbf{A} satisfy the block-RIP of order $|S_1 \cup S_2|$. Then for any vector $\mathbf{x} \in \mathbb{R}^{|S_2 \setminus S_1| \times d}$,*

$$(1 - \delta_{|S_1 \cup S_2|})\|\mathbf{x}\|_2^2 \leq \|\mathbf{P}^\perp[S_1]\mathbf{A}[S_2 \setminus S_1]\mathbf{x}\|_2^2 \leq (1 + \delta_{|S_1 \cup S_2|})\|\mathbf{x}\|_2^2.$$

3. Main results

3.1. An all-purpose sufficient condition

In this subsection, we give an all-purpose sufficient condition for the recovery of block K -sparse signals with the BOMP algorithm. We start with the following useful lemma.

Lemma 4. Let \mathbf{A} in (1.1) satisfy the block-RIP of order $K + 1$ and S be a proper subset of $\Omega \neq \emptyset$. If \mathbf{x} in (1.1) satisfies

$$\|\mathbf{x}[\Omega \setminus S]\|_{2,1}^2 \leq \nu \|\mathbf{x}[\Omega \setminus S]\|_2^2 \tag{3.1}$$

for some $\nu \geq 1$, then

$$\|\mathbf{A}^\top[\Omega \setminus S]\|_{2,\infty} - \|\mathbf{A}^\top[\Omega^c]\mathbf{q}[S]\|_{2,\infty} \geq \frac{(1 - \sqrt{\nu + 1}\delta_{|\Omega|+1})\|\mathbf{x}[\Omega \setminus S]\|_2}{\sqrt{\nu}}, \tag{3.2}$$

where

$$\mathbf{q}[S] = \mathbf{P}^\perp[S]\mathbf{A}[\Omega \setminus S]\mathbf{x}[\Omega \setminus S]. \tag{3.3}$$

Proof. See Section 4.1. \square

By using Lemma 4, we can prove the following all-purpose sufficient condition for the stable recovery of block K -sparse signals with the BOMP algorithm.

Theorem 1. Let $f(t)$ be a nondecreasing function of $t > 0$ with $0 < f(t) \leq t$. Suppose that block K -sparse \mathbf{x} satisfies

$$\|\mathbf{x}[\Omega \setminus S]\|_{2,1}^2 \leq f(|\Omega \setminus S|)\|\mathbf{x}[\Omega \setminus S]\|_2^2 \tag{3.4}$$

for any S which is a proper subset of $\Omega \neq \emptyset$. Suppose that \mathbf{v} satisfies $\|\mathbf{v}\|_2 \leq \epsilon$ and \mathbf{A} satisfies the block-RIP of order $K + 1$ with

$$\delta_{K+1} < \frac{1}{\sqrt{f(K) + 1}}. \tag{3.5}$$

Then, the BOMP algorithm with the stopping criterion $\|\mathbf{r}^k\| \leq \epsilon$ exactly recovers Ω in K iterations provided that

$$\min_{i \in \Omega} \|\mathbf{x}[i]\|_2 > \frac{2\epsilon}{1 - \sqrt{f(K) + 1}\delta_{K+1}}. \tag{3.6}$$

Moreover, the recovery error is bounded by

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 \leq \epsilon, \tag{3.7}$$

where $\hat{\mathbf{x}}$ is the block K -sparse signal returned by the BOMP algorithm.

Proof. See Section 4.2. \square

Remark 1. By (2.1) and the Cauchy–Schwarz inequality, it is easy to see that (3.4) always holds if $f(t) = t$. So it is not necessary to consider the case that $f(t) > t$ and this is the reason why we assume $f(t) \leq t$.

3.2. Recovery of general block sparse signals

In this subsection, we study the recovery of general block sparse signals. On the one hand, we will show that if \mathbf{A} satisfies the block-RIP of order $K + 1$ with $\delta_{K+1} < 1/\sqrt{K+1}$, then the BOMP Algorithm exactly or stably (under some conditions on $\min_{i \in \Omega} \|\mathbf{x}[i]\|_2$) recovers block K -sparse signals \mathbf{x} in K iterations. On the other hand, we will show that if $\delta_{K+1} \geq 1/\sqrt{K+1}$, then the BOMP algorithm may fail to recover a block K -sparse signal \mathbf{x} in K iterations.

By the Cauchy–Schwarz inequality and (2.1), it is easy to see that (3.4) holds if $f(t) = t$. Consequently, by Theorem 1, we can immediately obtain the following result which provides a sufficient condition for the exact support recovery of block K -sparse signal \mathbf{x} with BOMP in K iterations in noisy case.

Corollary 1. *Suppose that \mathbf{v} in (1.1) satisfies $\|\mathbf{v}\|_2 \leq \epsilon$, and \mathbf{A} in (1.1) satisfies the block-RIP of order $K + 1$ with*

$$\delta_{K+1} < \frac{1}{\sqrt{K+1}}. \quad (3.8)$$

Then the BOMP algorithm with the stopping criterion $\|\mathbf{r}^k\| \leq \epsilon$ exactly recovers Ω in K iterations provided that

$$\min_{i \in \Omega} \|\mathbf{x}[i]\|_2 > \frac{2\epsilon}{1 - \sqrt{K+1}\delta_{K+1}}.$$

Moreover, the recovery error can be bounded by (3.7).

From Corollary 1, we can immediately obtain the following sufficient condition for the exact recovery of block K -sparse signals with BOMP in K iterations in noise-free case.

Corollary 2. *Suppose that \mathbf{A} in (1.1) satisfies the block-RIP of order $K + 1$ with (3.8). Then, the BOMP algorithm exactly recovers block K -sparse signals \mathbf{x} in K iterations when $\mathbf{v} = \mathbf{0}$.*

When the block size $d = 1$, the block-RIP reduces to the classic RIP and the BOMP algorithm reduces to the OMP algorithm. Then Corollary 1 and Corollary 2 are [60, Theorem 1] and [46, Theorem III.1], respectively.

In the following, we give a necessary condition in terms of the RIC for the exact recovery of K -block sparse signals with the BOMP algorithm.

Theorem 2. *For any given positive integers $d, K \geq 1$ and any*

$$\frac{1}{\sqrt{K+1}} \leq \delta < 1, \quad (3.9)$$

there always exist a block K -sparse signal \mathbf{x} and a matrix \mathbf{A} satisfying the block-RIP of order $K + 1$ with $\delta_{K+1} = \delta$ such that the BOMP algorithm may fail to recover \mathbf{x} in K iterations.

Proof. See Section 4.3. \square

When the block size $d = 1$, Theorem 2 reduces to [61, Theorem 1]. Moreover, from Corollary 2 and Theorem 2, we can see that (3.8) is a sharp sufficient condition for recovering block K -sparse signals by the BOMP algorithm in K iterations.

Remark 2. Note that the sharp sufficient condition (3.8) has been independently obtained in [13] (see [13, Theorems 3.3, 3.4, 3.5]) which studies the recovery of block K -sparse signals with block orthogonal multi-matching pursuit, an improvement of BOMP.

3.3. Recovery of strongly-decaying block sparse signals

In this subsection, we will show that if \mathbf{A} satisfies the block-RIP of order $K+1$ with $\delta_{K+1} < \sqrt{2}/2$, then the BOMP algorithm exactly or stably (under some constraints on $\min_{i \in \Omega} \|\mathbf{x}[i]\|_2$) recovers all block α -strongly-decaying K -sparse signals in K iterations if α is larger than a value related to δ_{K+1} and K . Furthermore, we will show that if $\delta_{K+1} \geq \sqrt{2}/2$, the BOMP algorithm may fail to recover a block α -strongly-decaying K -sparse signal \mathbf{x} in K iterations, regardless how large is α . We will also show that the condition on α is sharp.

We first introduce the following lemma.

Lemma 5. For each $1 \leq i \leq K$, let

$$g_i(t) = \frac{(t^i - 1)(t + 1)}{(t^i + 1)(t - 1)}, \quad t > 1. \tag{3.10}$$

Then

$$1 = g_1(t) < g_2(t) < \dots < g_K(t). \tag{3.11}$$

Moreover, $g_i(t)$ is strictly decreasing with t and $1 < g_i(t) < i$ for $2 < i \leq K$.

Proof. See Section 4.4. \square

By Lemma 5, for $1 < s < K$, $g_K(t) = s$ has a unique solution t_s . To simplify the notation, we define

$$g_K^{-1}(s) = \begin{cases} t_s, & 1 < s < K \\ 1, & s \geq K \end{cases}. \tag{3.12}$$

Note that $g_K^{-1}(s)$ can be easily computed, for example, using Newton’s method.

By (3.12) and Theorem 1, we can obtain the following result which provides a sufficient condition for the exact support recovery of block α -strongly-decaying K -sparse signals with BOMP in K iterations in the noisy case.

Theorem 3. Suppose that \mathbf{v} in (1.1) satisfies $\|\mathbf{v}\|_2 \leq \epsilon$, and \mathbf{A} in (1.1) satisfies the block-RIP of order $K + 1$ with

$$\delta_{K+1} < \frac{\sqrt{2}}{2}. \tag{3.13}$$

If \mathbf{x} is a block α -strongly-decaying K -sparse signal with α satisfying

$$\alpha > g_K^{-1}(\delta_{K+1}^{-2} - 1). \tag{3.14}$$

Then BOMP with the stopping criterion $\|\mathbf{r}^k\| \leq \epsilon$ exactly recovers Ω in K iterations provided that

$$\min_{i \in \Omega} \|\mathbf{x}[i]\|_2 > \frac{2\epsilon}{1 - \sqrt{1 + \min\{s^{-2} - 1, K\}}\delta_{K+1}},$$

where s satisfies

$$\delta_{K+1} < s < \frac{\sqrt{2}}{2}, \quad g_K^{-1}(\delta_{K+1}^{-2} - 1) < g_K^{-1}(s^{-2} - 1) \leq \alpha. \tag{3.15}$$

Moreover, the recovery error can be bounded by (3.7).

Proof. See Section 4.5. \square

By Theorem 3, one can easily obtain the following result which provides a sufficient condition for the exact recovery of block α -strongly-decaying K -sparse signals with BOMP in K iterations in the noise-free case.

Corollary 3. Suppose that \mathbf{A} in (1.1) satisfies the block-RIP of order $K + 1$ with (3.13), and \mathbf{x} is a block α -strongly-decaying K -sparse signal with (3.14). Then the BOMP algorithm exactly recovers \mathbf{x} in K iterations when $\mathbf{v} = \mathbf{0}$.

Remark 3. By (3.12), $g_K^{-1}(\delta_{K+1}^{-2} - 1)$ is well defined for $\delta_{K+1} < \sqrt{2}/2$. By Lemma 5, $g_K(t)$ is strictly decreasing with t , so $g_K^{-1}(s)$ is strictly decreasing with s . Therefore, for each $\alpha > g_K^{-1}(\delta_{K+1}^{-2} - 1)$, there exists s such that (3.15) holds.

It was shown in [18, Theorem 4.1] that if \mathbf{A} in (1.1) satisfies the RIP of order $K + 1$ with $\delta_{K+1} < 1/3$, and \mathbf{x} is an α -strongly-decaying K -sparse signal with

$$\alpha > \frac{1 + 2 \frac{\delta_{K+1}}{1-\delta_{K+1}} \sqrt{K-1}}{1 - 2 \frac{\delta_{K+1}}{1-\delta_{K+1}}}. \tag{3.16}$$

Then, the OMP algorithm exactly recovers \mathbf{x} in K iterations when $\mathbf{v} = \mathbf{0}$.

If the block size $d = 1$, then the sufficient condition given by Theorem 3 is a condition for recovering α -strongly-decaying K -sparse signals by using the OMP algorithm. By (3.13), our condition on δ_{K+1} is much weaker than that given in [18, Theorem 4.1]. Moreover, our condition on α is also less restrictive than that given in [18, Theorem 4.1]. In the following, we show it. By (3.14) and (3.16), it is equivalent to show that

$$\frac{1 + 2 \frac{\delta_{K+1}}{1-\delta_{K+1}} \sqrt{K-1}}{1 - 2 \frac{\delta_{K+1}}{1-\delta_{K+1}}} > g_K^{-1}(\delta_{K+1}^{-2} - 1)$$

which is equivalent to

$$\frac{1 - \delta_{K+1} + 2\delta_{K+1}\sqrt{K-1}}{1 - 3\delta_{K+1}} > g_K^{-1}(\delta_{K+1}^{-2} - 1).$$

By the previous analysis, $g_K^{-1}(s)$ is decreasing with s . Thus, we only need to show

$$g_K \left(\frac{1 - \delta_{K+1} + 2\delta_{K+1}\sqrt{K-1}}{1 - 3\delta_{K+1}} \right) < \delta_{K+1}^{-2} - 1.$$

By (3.10), the inequality above holds if

$$\frac{\frac{1 - \delta_{K+1} + 2\delta_{K+1}\sqrt{K-1}}{1 - 3\delta_{K+1}} + 1}{\frac{1 - \delta_{K+1} + 2\delta_{K+1}\sqrt{K-1}}{1 - 3\delta_{K+1}} - 1} < \delta_{K+1}^{-2} - 1,$$

i.e.,

$$\frac{1 - 2\delta_{K+1} + \delta_{K+1}\sqrt{K-1}}{\delta_{K+1}(1 + \sqrt{K-1})} < \frac{1 - \delta_{K+1}^2}{\delta_{K+1}^2}.$$

To show the aforementioned inequality, it is equivalently to show

$$2\delta_{K+1}^2\sqrt{K-1} < (1 + \delta_{K+1}^2 - \delta_{K+1}) + \sqrt{K-1}.$$

It is clear to see that the above inequality holds when $\delta_{K+1} < \sqrt{2}/2$. Thus, our sufficient condition given by Theorem 3 is less restrictive than [18, Theorem 4.1] in terms of both δ_{K+1} and α .

In the following, we give two necessary conditions in terms of the RIC and α for the exact recovery of block α -strongly-decaying K -sparse signals by using the BOMP algorithm.

Theorem 4. *For any given positive integers $d, K \geq 1$, any given $\alpha > 1$ and any*

$$\frac{\sqrt{2}}{2} \leq \delta < 1, \tag{3.17}$$

there always exist a block α -strongly-decaying K -sparse signal \mathbf{x} and a matrix \mathbf{A} satisfying the block-RIP of order $K + 1$ with $\delta_{K+1} = \delta$ such that the BOMP algorithm may fail to recover \mathbf{x} in K iterations.

Proof. See Section 4.6. \square

Let

$$\delta_{K+1} = \frac{1}{\sqrt{K+1}}.$$

By (3.12), we have

$$g_K^{-1}(\delta_{K+1}^{-2} - 1) = g_K^{-1}(K) = 1.$$

By Definition 1, $\alpha > 1$, thus we cannot set $\alpha = g_K^{-1}(\delta_{K+1}^{-2} - 1)$ for all δ_{K+1} satisfying (3.13). Hence, we have the following result.

Corollary 4. *For any given positive integers $d, K \geq 1$, there always exist a matrix \mathbf{A} which satisfies the block-RIP of order $K + 1$ with δ_{K+1} satisfying (3.13) such that*

$$\alpha = g_K^{-1}(\delta_{K+1}^{-2} - 1) = 1,$$

where $g_K(t)$ is defined in (3.10).

From Theorems 3, 4 and Corollary 4, we can see that (3.13) is a sharp sufficient condition for recovering block α -strongly-decaying K -sparse signals \mathbf{x} by the BOMP algorithm in K iterations.

3.4. Partial recovery of block sparse signals

In this subsection, we will show that if \mathbf{A} satisfies the block-RIP of order $K + 1$ with $\delta_{K+1} < \sqrt{2}/2$ and $\|\mathbf{x}[j]\|_2/\|\mathbf{x}[j + 1]\|_2, 1 \leq j \leq i$ satisfy some conditions for some $1 \leq i \leq K$, then the BOMP algorithm is able to choose an index in Ω in each of the first i iterations. Moreover, we will show that this condition is sharp.

Before presenting our sufficient conditions, we introduce the following lemma.

Lemma 6. For each $1 \leq i \leq K - 1$, let

$$\hat{h}_i(t) = \frac{(t+i)^2}{t^2+i}, \quad t \geq 1. \tag{3.18}$$

Then, \hat{h}_i is strictly decreasing with t , and $1 < \hat{h}_i(t) \leq 1 + i$.

Moreover, for $1 < s \leq 1 + i$, $\hat{h}_i(t) = s$ if and only if $t = h_i(s)$, where for $1 \leq i \leq K - 1$,

$$h_i(s) = \frac{i + \sqrt{i^2s + (s - s^2)i}}{s - 1}, \quad 1 < s \leq 1 + i. \tag{3.19}$$

Note that for each $1 \leq i \leq K - 1$ and $1 < s \leq 1 + i$, we have $i^2s + (s - s^2)i = is[(i + 1) - s] \geq 0$. Thus, (3.19) is well-defined.

Proof. By some simple calculations and the assumption that $t \geq 1$, we have

$$\hat{h}_i(t)' = \frac{2(t+i)(1-t)i}{(t^2+i)^2} \leq 0.$$

Moreover, $\hat{h}_i(t)' = 0$ if and only if $t = 1$. Thus, \hat{h}_i is strictly decreasing with t . It is easy to see that

$$\hat{h}_i(1) = 1 + i, \quad \lim_{t \rightarrow +\infty} \hat{h}_i(t) = 1.$$

Thus, $1 < \hat{h}_i(t) \leq 1 + i$. For $1 < s \leq 1 + i$, it is straightforward to solve the quadratic equation $\hat{h}(t) = s$ which yields (3.19). \square

By (3.19) and Theorem 1, we can obtain the following sufficient condition for partial support recovery of block K -sparse signals with the BOMP algorithm in the noisy case.

Theorem 5. Suppose that in (1.1), \mathbf{A} satisfies the block-RIP of order $K + 1$ with δ_{K+1} satisfying (3.13) and \mathbf{v} satisfies $\|\mathbf{v}\|_2 \leq \epsilon$. If the block K -sparse \mathbf{x} in (1.1) satisfies

$$\|\mathbf{x}[j]\|_2 > h_{K-j}(\delta_{K+1}^{-2} - 1)\|\mathbf{x}[j+1]\|_2, \quad 1 \leq j \leq i. \tag{3.20}$$

Then the BOMP algorithm choose an index in Ω in each of the first i iterations provided that

$$\min_{i \in \Omega} \|\mathbf{x}[i]\|_2 > \frac{2\epsilon}{1 - \sqrt{1 + \min\{s^{-2} - 1, K\}}\delta_{K+1}},$$

where s satisfies $\delta_{K+1} < s < \frac{\sqrt{2}}{2}$ and

$$h_{K-1}(\delta_{K+1}^{-2} - 1) < h_{K-1}(s^{-2} - 1) \leq \min_{1 \leq j \leq i} \frac{\|\mathbf{x}[j]\|_2}{\|\mathbf{x}[j+1]\|_2}. \tag{3.21}$$

Proof. See Section 4.7. \square

By Theorem 5, we can get the following consequence which provides a sufficient condition for the partial recovery of block K -sparse signals with the BOMP algorithm in the noise-free case.

Corollary 5. Suppose that \mathbf{A} in (1.1) satisfies the block-RIP of order $K + 1$ with δ_{K+1} satisfying (3.13). If the block K -sparse \mathbf{x} in (1.1) satisfies (3.20). Then, the BOMP algorithm choose an index in Ω in each of the first i iterations when $\mathbf{v} = \mathbf{0}$.

Remark 4. By Lemma 6, $\hat{h}_1(t)$ is strictly decreasing with t , so $h_1(s)$ is strictly decreasing with s . Therefore, by (3.20), there exists an s such that (3.21) holds.

In the following, we give a necessary condition for partial recovery of block K -sparse signals by using the BOMP algorithm. Let

$$\delta_{K+1} = \frac{1}{\sqrt{K+1}}.$$

By (3.19), we have

$$h_{K-1}(\delta_{K+1}^{-2} - 1) = h_{K-1}(K) = 1.$$

Thus, from the proof of Theorem 2, we can easily get the following result.

Corollary 6. For any given positive integers $d, K \geq 1$, there always exist a matrix \mathbf{A} satisfying the block-RIP of order $K + 1$ with δ_{K+1} satisfying (3.13), and a block K -sparse signal with

$$\|\mathbf{x}[1]\|_2 = h_{K-1}(\delta_{K+1}^{-2} - 1)\|\mathbf{x}[2]\|_2 \tag{3.22}$$

such that the BOMP algorithm may fail to choose an index in the support Ω in the first iteration.

From Corollaries 5 and 6, and Theorem 5, we can see that Corollary 5 gives a sharp sufficient condition for partial recovery of block K -sparse signals \mathbf{x} with the BOMP algorithm in K iterations.

4. Proofs

In this section, we mainly prove the main results given in Section 3. We devote each subsection to one proof.

4.1. Proof of Lemma 4

The following proof follows a similar line as that for [60, Lemma 1], but with a key distinction. Specifically, we introduce a well-chosen vector \mathbf{h} (see (4.5)) to define the vector \mathbf{w} (see (4.7)) which incorporates the particular feature of block sparse.

Proof. The proof consists of two steps, in the first step, we prove

$$\sqrt{\nu}\|\mathbf{x}[\Omega \setminus S]\|_2\|\mathbf{A}^\top[\Omega \setminus S]\mathbf{q}[S]\|_{2,\infty} \geq \|\mathbf{q}[S]\|_2^2. \tag{4.1}$$

In the second step, we show for each $j \in \Omega^c$,

$$\|\mathbf{q}[S]\|_2^2 - \sqrt{\nu}\|\mathbf{x}[\Omega \setminus S]\|_2\|\mathbf{A}^\top[j]\mathbf{q}[S]\|_2 \geq (1 - \sqrt{\nu + 1}\delta_{|\Omega|+1})\|\mathbf{x}[\Omega \setminus S]\|_2^2. \tag{4.2}$$

We first show (4.1). It is easy to check that

$$\begin{aligned} & \sqrt{\nu}\|\mathbf{x}[\Omega \setminus S]\|_2\|\mathbf{A}^\top[\Omega \setminus S]\mathbf{q}[S]\|_{2,\infty} \stackrel{(a)}{\geq} \|\mathbf{x}[\Omega \setminus S]\|_{2,1}\|\mathbf{A}^\top[\Omega \setminus S]\mathbf{q}[S]\|_{2,\infty} \\ & \stackrel{(b)}{=} \left(\sum_{\ell \in \Omega \setminus S} \|\mathbf{x}[\ell]\|_2 \right) \|\mathbf{A}^\top[\Omega \setminus S]\mathbf{q}[S]\|_{2,\infty} \stackrel{(c)}{\geq} \sum_{\ell \in \Omega \setminus S} (\|\mathbf{x}[\ell]\|_2 \|\mathbf{A}^\top[\ell]\mathbf{q}[S]\|_2) \end{aligned}$$

$$\begin{aligned} &\stackrel{(d)}{\geq} \sum_{\ell \in \Omega \setminus S} (|\mathbf{x}^\top[\ell] \mathbf{A}^\top[\ell] \mathbf{q}[S]|) \geq \sum_{\ell \in \Omega \setminus S} (\mathbf{x}^\top[\ell] \mathbf{A}^\top[\ell]) \mathbf{q}[S] = (\mathbf{A}[\Omega \setminus S] \mathbf{x}[\Omega \setminus S])^\top \mathbf{q}[S] \\ &\stackrel{(e)}{=} (\mathbf{P}^\perp[S] \mathbf{A}[\Omega \setminus S] \mathbf{x}[\Omega \setminus S])^\top (\mathbf{P}^\perp[S] \mathbf{A}[\Omega \setminus S] \mathbf{x}[\Omega \setminus S]) = \|\mathbf{P}^\perp[S] \mathbf{A}[\Omega \setminus S] \mathbf{x}[\Omega \setminus S]\|_2^2 \stackrel{(f)}{=} \|\mathbf{q}[S]\|_2^2, \end{aligned}$$

where (a) and (b) respectively follow from (3.1) and (2.1), (c) is because for each $\ell \in \Omega \setminus S$,

$$\|\mathbf{A}^\top[\Omega \setminus S] \mathbf{q}[S]\|_{2,\infty} \geq \|\mathbf{A}[\ell]^\top \mathbf{q}[S]\|_2,$$

(d) is from the Cauchy–Schwarz inequality, and (e) and (f) are from (3.3) and the projection property:

$$(\mathbf{P}^\perp[S])^\top \mathbf{P}^\perp[S] = \mathbf{P}^\perp[S] \mathbf{P}^\perp[S] = \mathbf{P}^\perp[S]. \tag{4.3}$$

Thus, (4.1) holds.

In the following, we prove (4.2). Let

$$\mu = -\frac{\sqrt{\nu+1}-1}{\sqrt{\nu}}.$$

By a simple calculation, we have

$$\frac{2\mu}{1-\mu^2} = -\sqrt{\nu}, \quad \frac{1+\mu^2}{1-\mu^2} = \sqrt{\nu+1}. \tag{4.4}$$

To simplify notation, we define $\mathbf{h} \in \mathbb{R}^d$ by

$$h_i = \frac{(\mathbf{A}[j])_i^\top \mathbf{q}[S]}{\|\mathbf{A}^\top[j] \mathbf{q}[S]\|_2}, \quad 1 \leq i \leq d. \tag{4.5}$$

Then, it is easy to see that $\|\mathbf{h}\|_2 = 1$ and

$$\mathbf{h}^\top \mathbf{A}^\top[j] \mathbf{q}[S] = \|\mathbf{A}^\top[j] \mathbf{q}[S]\|_2. \tag{4.6}$$

Furthermore, we define two vectors

$$\begin{aligned} \mathbf{u} &= \begin{bmatrix} \mathbf{x}[\Omega \setminus S] \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^{(|\Omega \setminus S|+1)d}, \\ \mathbf{w} &= \begin{bmatrix} \mathbf{0} \\ \mu \|\mathbf{x}[\Omega \setminus S]\|_2 \mathbf{h} \end{bmatrix} \in \mathbb{R}^{(|\Omega \setminus S|+1)d}. \end{aligned} \tag{4.7}$$

For given $j \in \Omega^c$, we let

$$\mathbf{B} = \mathbf{P}^\perp[S] [\mathbf{A}[\Omega \setminus S] \quad \mathbf{A}[j]]. \tag{4.8}$$

Then

$$\mathbf{q}[S] = \mathbf{P}^\perp[S] \mathbf{A}[\Omega \setminus S] \mathbf{x}[\Omega \setminus S] = \mathbf{B} \mathbf{u} \tag{4.9}$$

and

$$\|\mathbf{u} + \mathbf{w}\|_2^2 = (1 + \mu^2) \|\mathbf{x}[\Omega \setminus S]\|_2^2, \tag{4.10}$$

$$\|\mu^2 \mathbf{u} - \mathbf{w}\|_2^2 = \mu^2 (1 + \mu^2) \|\mathbf{x}[\Omega \setminus S]\|_2^2. \tag{4.11}$$

Moreover, we have

$$\begin{aligned} \mathbf{w}^\top \mathbf{B}^\top \mathbf{B} \mathbf{u} &\stackrel{(a)}{=} \mu \|\mathbf{x}[\Omega \setminus S]\|_2 \mathbf{h}^\top \mathbf{A}^\top [j] (\mathbf{P}^\perp[S])^\top \mathbf{q}[S] \\ &\stackrel{(b)}{=} \mu \|\mathbf{x}[\Omega \setminus S]\|_2 \mathbf{h}^\top \mathbf{A}^\top [j] \mathbf{q}[S] \\ &\stackrel{(c)}{=} \mu \|\mathbf{x}[\Omega \setminus S]\|_2 \|\mathbf{A}^\top [j] \mathbf{q}[S]\|_2, \end{aligned}$$

where (a) follows from (4.8) and (4.9), (b) is from (3.3) and (4.3), and (c) follows from (4.6). Therefore,

$$\begin{aligned} &\|\mathbf{B}(\mathbf{u} + \mathbf{w})\|_2^2 - \|\mathbf{B}(\mu^2 \mathbf{u} - \mathbf{w})\|_2^2 \\ &= (1 - \mu^4) \|\mathbf{B} \mathbf{u}\|_2^2 + 2(1 + \mu^2) \mathbf{w}^\top \mathbf{B}^\top \mathbf{B} \mathbf{u} \\ &= (1 - \mu^4) \|\mathbf{B} \mathbf{u}\|_2^2 + 2\mu(1 + \mu^2) \|\mathbf{x}[\Omega \setminus S]\|_2 \|\mathbf{A}^\top [j] \mathbf{q}[S]\|_2 \\ &= (1 - \mu^4) (\|\mathbf{B} \mathbf{u}\|_2^2 + \frac{2\mu}{1 - \mu^2} \|\mathbf{x}[\Omega \setminus S]\|_2 \|\mathbf{A}^\top [j] \mathbf{q}[S]\|_2) \\ &= (1 - \mu^4) (\|\mathbf{B} \mathbf{u}\|_2^2 - \sqrt{\nu} \|\mathbf{x}[\Omega \setminus S]\|_2 \|\mathbf{A}^\top [j] \mathbf{q}[S]\|_2), \end{aligned} \tag{4.12}$$

where the last equality follows from the first equality in (4.4).

On the other hand, using (4.8) and Lemma 3, we have

$$\begin{aligned} &\|\mathbf{B}(\mathbf{u} + \mathbf{w})\|_2^2 - \|\mathbf{B}(\mu^2 \mathbf{u} - \mathbf{w})\|_2^2 \\ &\stackrel{(a)}{\geq} (1 - \delta_{|\Omega|+1}) \|\mathbf{u} + \mathbf{w}\|_2^2 - (1 + \delta_{|\Omega|+1}) \|\mu^2 \mathbf{u} - \mathbf{w}\|_2^2 \\ &\stackrel{(b)}{=} (1 - \delta_{|\Omega|+1}) (1 + \mu^2) \|\mathbf{x}[\Omega \setminus S]\|_2^2 - (1 + \delta_{|\Omega|+1}) \mu^2 (1 + \mu^2) \|\mathbf{x}[\Omega \setminus S]\|_2^2 \\ &= (1 + \mu^2) \|\mathbf{x}[\Omega \setminus S]\|_2^2 ((1 - \delta_{|\Omega|+1}) - (1 + \delta_{|\Omega|+1}) \mu^2) \\ &= (1 + \mu^2) \|\mathbf{x}[\Omega \setminus S]\|_2^2 ((1 - \mu^2) - \delta_{|\Omega|+1} (1 + \mu^2)) \\ &= (1 - \mu^4) \|\mathbf{x}[\Omega \setminus S]\|_2^2 (1 - \frac{1 + \mu^2}{1 - \mu^2} \delta_{|\Omega|+1}) \\ &\stackrel{(c)}{=} (1 - \mu^4) \|\mathbf{x}[\Omega \setminus S]\|_2^2 (1 - \sqrt{\nu + 1} \delta_{|\Omega|+1}), \end{aligned} \tag{4.13}$$

where (a) follows from Lemma 3 and (4.8), (b) is from (4.10) and (4.11), and (c) follows from the second equality in (4.4).

By (4.12), (4.13) and the fact that $1 - \mu^4 > 0$, we have

$$\|\mathbf{B} \mathbf{u}\|_2^2 - \sqrt{\nu} \|\mathbf{x}[\Omega \setminus S]\|_2 \|\mathbf{A}^\top [j] \mathbf{q}[S]\|_2 \geq \|\mathbf{x}[\Omega \setminus S]\|_2^2 (1 - \sqrt{\nu + 1} \delta_{|\Omega|+1}).$$

By (4.9), (4.2) holds, and this completes the proof. \square

4.2. Proof of Theorem 1

Proof. We first prove the first part of the results in Theorem 1. The proof consists of two steps. In the first step, we show that the BOMP algorithm selects correct indexes in all iterations. In the second step, we prove that it performs exactly $|\Omega|$ iterations.

We prove the first step by induction. Suppose that the BOMP algorithm selects correct indexes in the first k iterations, i.e., $S_k \subseteq \Omega$ for $0 \leq k \leq |\Omega| - 1$. Then, we need to show that the BOMP algorithm selects a correct index in the $(k + 1)$ -th iteration, i.e., by Algorithm 1, we show that $s^{k+1} \in \Omega$. Note that the proof

for the first selection is contained in the case that $k = 0$ because the induction assumption $S_k \subseteq \Omega$ holds in this case (not that $S_0 = \emptyset$).

By Algorithm 1, it is easy to see that

$$\|\mathbf{A}^\top [S_k] \mathbf{r}^k\|_2 = 0.$$

Thus, by line 2 of Algorithm 1, to show $s^{k+1} \in \Omega$, it suffices to show

$$\|\mathbf{A}^\top [\Omega \setminus S_k] \mathbf{r}^k\|_{2,\infty} > \|\mathbf{A}^\top [\Omega^c] \mathbf{r}^k\|_{2,\infty}. \tag{4.14}$$

By line 4 of Algorithm 1, we have

$$\hat{\mathbf{x}}[S_k] = (\mathbf{A}^\top [S_k] \mathbf{A}[S_k])^{-1} \mathbf{A}^\top [S_k] \mathbf{y}. \tag{4.15}$$

Then, by line 5 of Algorithm 1 and (4.15), we have

$$\begin{aligned} \mathbf{r}^k &= \mathbf{y} - \mathbf{A}[S_k] \hat{\mathbf{x}}[S_k] \\ &= (\mathbf{I} - \mathbf{A}[S_k] (\mathbf{A}^\top [S_k] \mathbf{A}[S_k])^{-1} \mathbf{A}^\top [S_k]) \mathbf{y} \\ &\stackrel{(a)}{=} \mathbf{P}^\perp [S_k] (\mathbf{A} \mathbf{x} + \mathbf{v}) \\ &\stackrel{(b)}{=} \mathbf{P}^\perp [S_k] (\mathbf{A}[\Omega] \mathbf{x}[\Omega] + \mathbf{v}) \\ &\stackrel{(c)}{=} \mathbf{P}^\perp [S_k] (\mathbf{A}[S_k] \mathbf{x}[S_k] + \mathbf{A}[\Omega \setminus S_k] \mathbf{x}[\Omega \setminus S_k] + \mathbf{v}) \\ &\stackrel{(d)}{=} \mathbf{P}^\perp [S_k] \mathbf{A}[\Omega \setminus S_k] \mathbf{x}[\Omega \setminus S_k] + \mathbf{P}^\perp [S_k] \mathbf{v}, \end{aligned} \tag{4.16}$$

where (a) is from the definition of $\mathbf{P}^\perp [S_k]$, (b) is from the definition of Ω , (c) follows from the induction assumption that $S_k \subseteq \Omega$ and (d) is from $\mathbf{P}^\perp [S_k] \mathbf{A}[S_k] = \mathbf{0}$.

Thus, by (4.16), we obtain

$$\|\mathbf{A}^\top [\Omega \setminus S_k] \mathbf{r}^k\|_{2,\infty} \geq \|\mathbf{A}^\top [\Omega \setminus S_k] \mathbf{P}^\perp [S_k] \mathbf{A}[\Omega \setminus S_k] \mathbf{x}[\Omega \setminus S_k]\|_{2,\infty} - \|\mathbf{A}^\top [\Omega \setminus S_k] \mathbf{P}^\perp [S_k] \mathbf{v}\|_{2,\infty}, \tag{4.17}$$

for the left-hand side of (4.14), and for the right-hand side of (4.14), we have

$$\|\mathbf{A}^\top [\Omega^c] \mathbf{r}^k\|_{2,\infty} \leq \|\mathbf{A}^\top [\Omega^c] \mathbf{P}^\perp [S_k] \mathbf{A}[\Omega \setminus S_k] \mathbf{x}[\Omega \setminus S_k]\|_{2,\infty} + \|\mathbf{A}^\top [\Omega^c] \mathbf{P}^\perp [S_k] \mathbf{v}\|_{2,\infty}. \tag{4.18}$$

Therefore, by (4.17) and (4.18), in order to show (4.14), we show

$$\begin{aligned} &\|\mathbf{A}^\top [\Omega \setminus S_k] \mathbf{P}^\perp [S_k] \mathbf{A}[\Omega \setminus S_k] \mathbf{x}[\Omega \setminus S_k]\|_{2,\infty} - \|\mathbf{A}^\top [\Omega^c] \mathbf{P}^\perp [S_k] \mathbf{A}[\Omega \setminus S_k] \mathbf{x}[\Omega \setminus S_k]\|_{2,\infty} \\ &> \|\mathbf{A}^\top [\Omega \setminus S_k] \mathbf{P}^\perp [S_k] \mathbf{v}\|_{2,\infty} + \|\mathbf{A}^\top [\Omega^c] \mathbf{P}^\perp [S_k] \mathbf{v}\|_{2,\infty}. \end{aligned} \tag{4.19}$$

By the induction assumption, $S_k \subseteq \Omega$, thus

$$\|\mathbf{x}[\Omega \setminus S_k]\|_2 \geq \sqrt{|\Omega| - k} \min_{i \in \Omega \setminus S_k} \|\mathbf{x}[i]\|_2 \geq \sqrt{|\Omega| - k} \min_{i \in \Omega} \|\mathbf{x}[i]\|_2 \geq \sqrt{f(|\Omega| - k)} \min_{i \in \Omega} \|\mathbf{x}[i]\|_2, \tag{4.20}$$

where the last inequality follows from the fact that $f(t) \leq t$ for $t > 0$.

In the following, we shall give a lower bound on the left-hand side of (4.19) and an upper bound on the right-hand side of (4.19), respectively. Since $S_k \subseteq \Omega$ and $|S_k| = k$, by Lemma 4 and (3.4), we have

$$\begin{aligned}
 & \|\mathbf{A}^\top[\Omega \setminus S_k] \mathbf{P}^\perp[S_k] \mathbf{A}[\Omega \setminus S_k] \mathbf{x}[\Omega \setminus S_k]\|_\infty - \|\mathbf{A}^\top[\Omega^c] \mathbf{P}^\perp[S_k] \mathbf{A}[\Omega \setminus S_k] \mathbf{x}[\Omega \setminus S_k]\|_\infty \\
 & \geq \frac{(1 - \sqrt{f(|\Omega| - k) + 1\delta_{|\Omega|+1}}) \|\mathbf{x}[\Omega \setminus S_k]\|_2}{\sqrt{f(|\Omega| - k)}} \stackrel{(a)}{\geq} \frac{(1 - \sqrt{f(K) + 1\delta_{|\Omega|+1}}) \|\mathbf{x}[\Omega \setminus S_k]\|_2}{\sqrt{f(|\Omega| - k)}} \\
 & \stackrel{(b)}{\geq} \frac{(1 - \sqrt{f(K) + 1\delta_{K+1}}) \|\mathbf{x}[\Omega \setminus S_k]\|_2}{\sqrt{f(|\Omega| - k)}} \stackrel{(c)}{\geq} (1 - \sqrt{f(K) + 1\delta_{K+1}}) \min_{i \in \Omega} \|\mathbf{x}[i]\|_2, \tag{4.21}
 \end{aligned}$$

where (a) is because $k \geq 1$, \mathbf{x} is block K -sparse (i.e., $|\Omega| \leq K$) and $f(t)$ is nondecreasing, (b) follows from Lemma 1, and (c) follows from (3.5) and (4.20).

To give an upper bound on the right-hand side of (4.19), we notice that there exist $i_0 \in \Omega \setminus S_k$ and $j_0 \in \Omega^c$ such that

$$\begin{aligned}
 \|\mathbf{A}^\top[\Omega \setminus S_k] \mathbf{P}^\perp[S_k] \mathbf{v}\|_{2,\infty} &= \|\mathbf{A}^\top[i_0] \mathbf{P}^\perp[S_k] \mathbf{v}\|_2, \\
 \|\mathbf{A}^\top[\Omega^c] \mathbf{P}^\perp[S_k] \mathbf{v}\|_{2,\infty} &= \|\mathbf{A}^\top[j_0] \mathbf{P}^\perp[S_k] \mathbf{v}\|_2.
 \end{aligned}$$

Therefore, we have

$$\begin{aligned}
 & \|\mathbf{A}^\top[\Omega \setminus S_k] \mathbf{P}^\perp[S_k] \mathbf{v}\|_{2,\infty} + \|\mathbf{A}^\top[\Omega^c] \mathbf{P}^\perp[S_k] \mathbf{v}\|_{2,\infty} \\
 &= \|\mathbf{A}^\top[i_0] \mathbf{P}^\perp[S_k] \mathbf{v}\|_2 + \|\mathbf{A}^\top[j_0] \mathbf{P}^\perp[S_k] \mathbf{v}\|_2 \\
 &= \|\mathbf{A}^\top[i_0 \cup j_0] \mathbf{P}^\perp[S_k] \mathbf{v}\|_{2,1} \\
 &\stackrel{(a)}{\leq} \sqrt{2} \|\mathbf{A}^\top[i_0 \cup j_0] \mathbf{P}^\perp[S_k] \mathbf{v}\|_2 \\
 &\stackrel{(b)}{\leq} \sqrt{2(1 + \delta_{K+1})} \|\mathbf{P}^\perp[S_k] \mathbf{v}\|_2 \\
 &\stackrel{(c)}{\leq} \sqrt{2(1 + \delta_{K+1})} \epsilon, \tag{4.22}
 \end{aligned}$$

where (a) is from (2.1) and the fact that $\mathbf{A}^\top[i_0 \cup j_0] \mathbf{P}^\perp[S_k] \mathbf{v}$ is a 2×1 block vector, (b) follows from Lemma 2 and (c) is because

$$\|\mathbf{P}^\perp[S_k] \mathbf{v}\|_2 \leq \|\mathbf{P}^\perp[S_k]\|_2 \|\mathbf{v}\|_2 \leq \|\mathbf{v}\|_2 \leq \epsilon. \tag{4.23}$$

From (4.21) and (4.22), we can see that (4.19) (or equivalently (4.14)) is guaranteed by

$$(1 - \sqrt{f(K) + 1\delta_{K+1}}) \min_{i \in \Omega} \|\mathbf{x}[i]\|_2 > \sqrt{2(1 + \delta_{K+1})} \epsilon,$$

i.e., by (3.5),

$$\min_{i \in \Omega} \|\mathbf{x}[i]\|_2 > \frac{\sqrt{2(1 + \delta_{K+1})} \epsilon}{1 - \sqrt{f(K) + 1\delta_{K+1}}}.$$

Thus, if (3.6) holds, then the BOMP algorithm selects a correct index in the k th iteration.

Next we need to show that the BOMP algorithm performs exactly $|\Omega|$ iterations, which is equivalent to show that $\|\mathbf{r}^k\|_2 > \epsilon$ for $1 \leq k < |\Omega|$ and $\|\mathbf{r}^{|\Omega|}\|_2 \leq \epsilon$. Since the BOMP algorithm selects a correct index in each iteration under (3.6), by (4.16), for $1 \leq k < |\Omega|$, we have

$$\begin{aligned}
 \|\mathbf{r}^k\|_2 &= \|\mathbf{P}^\perp[S_k] \mathbf{A}[\Omega \setminus S_k] \mathbf{x}[\Omega \setminus S_k] + \mathbf{P}^\perp[S_k] \mathbf{v}\|_2 \\
 &\geq \|\mathbf{P}^\perp[S_k] \mathbf{A}[\Omega \setminus S_k] \mathbf{x}[\Omega \setminus S_k]\|_2 - \|\mathbf{P}^\perp[S_k] \mathbf{v}\|_2
 \end{aligned}$$

$$\begin{aligned}
 &\stackrel{(a)}{\geq} \|\mathbf{P}^\perp[S_k]\mathbf{A}[\Omega \setminus S_k]\mathbf{x}[\Omega \setminus S_k]\|_2 - \epsilon \\
 &\stackrel{(b)}{\geq} \sqrt{1 - \delta_{|\Omega|}}\|\mathbf{x}[\Omega \setminus S_k]\|_2 - \epsilon \\
 &\stackrel{(c)}{\geq} \sqrt{1 - \delta_{K+1}} \min_{i \in \Omega} \|\mathbf{x}[i]\|_2 - \epsilon,
 \end{aligned} \tag{4.24}$$

where (a), (b) and (c) are respectively from (4.23), Lemma 3, and Lemma 1. Therefore, if

$$\min_{i \in \Omega} \|\mathbf{x}[i]\|_2 > \frac{2\epsilon}{\sqrt{1 - \delta_{K+1}}}, \tag{4.25}$$

then $\|\mathbf{r}^k\|_2 > \epsilon$ for $1 \leq k < \Omega$.

By a simple calculation, we can show that

$$\frac{2\epsilon}{1 - \sqrt{f(K) + 1}\delta_{K+1}} \geq \frac{2\epsilon}{\sqrt{1 - \delta_{K+1}}}. \tag{4.26}$$

Indeed, we have

$$1 - \sqrt{f(K) + 1}\delta_{K+1} \leq 1 - \delta_{K+1} \leq \sqrt{1 - \delta_{K+1}}.$$

Thus, (4.26) follows. Therefore, by (4.25) and (4.26), if (3.6) holds, then $\|\mathbf{r}^k\|_2 > \epsilon$ for $1 \leq k < \Omega$, i.e., the BOMP algorithm does not terminate before the $|\Omega|$ -th iteration.

Similarly, by (4.16), we have

$$\|\mathbf{r}^{|\Omega|}\|_2 = \|\mathbf{P}_{S_{|\Omega|}}^\perp \mathbf{A}_{\Omega \setminus S_{|\Omega|}} \mathbf{x}_{\Omega \setminus S_{|\Omega|}} + \mathbf{P}^\perp[S_{|\Omega|}]\mathbf{v}\|_2 \stackrel{(a)}{=} \|\mathbf{P}_{S_{|\Omega|}}^\perp \mathbf{v}\|_2 \stackrel{(b)}{\leq} \epsilon,$$

where (a) is because $S_{|\Omega|} = |\Omega|$ and (b) follows from (4.23). So, by the stopping condition, the BOMP algorithm terminates after performing the $|\Omega|$ -th iteration in total. That is, the BOMP algorithm performs $|\Omega|$ iterations.

Next let us prove the second part, i.e. (3.7) of Theorem 1. Since the support of the block K -sparse signal \mathbf{x} can be exactly recovered by the BOMP algorithm, we get

$$\hat{\mathbf{x}} = (\mathbf{A}^\top[\Omega]\mathbf{A}[\Omega])^{-1}\mathbf{A}^\top[\Omega]\mathbf{y} = (\mathbf{A}^\top[\Omega]\mathbf{A}[\Omega])^{-1}\mathbf{A}^\top[\Omega](\mathbf{A}\mathbf{x} + \mathbf{v}) = \mathbf{x} + \mathbf{P}[\Omega]\mathbf{v}.$$

Thus, we have

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 = \|\mathbf{P}[\Omega]\mathbf{v}\|_2 \leq \|\mathbf{v}\|_2 \leq \epsilon.$$

This completes the proof. \square

4.3. Proof of Theorem 2

Proof. For any given positive integers $d, K \geq 1$, let matrix function

$$\mathbf{B}(d) = \begin{bmatrix} \frac{K}{K+1}\mathbf{I}_{dK} & \frac{\mathbf{E}_{(dK) \times d}}{K+1} \\ \frac{\mathbf{E}_{(dK) \times d}^\top}{K+1} & \frac{K+2}{K+1}\mathbf{I}_d \end{bmatrix},$$

where

$$\mathbf{E}_{(dK) \times d} = (\mathbf{I}_d, \dots, \mathbf{I}_d)^\top \in \mathbb{R}^{(dK) \times d} \tag{4.27}$$

with \mathbf{I}_d being the $d \times d$ identity matrix.

Let

$$s = \delta - \frac{1}{\sqrt{K+1}}, \quad \mathbf{C}(d) = \mathbf{B}(d) - s\mathbf{I}_{d(K+1)} \tag{4.28}$$

and

$$\mathbf{x}(d) = (\mathbf{1}_d^\top, \dots, \mathbf{1}_d^\top, \mathbf{0}_d^\top) \in \mathbb{R}^{d(K+1)}, \tag{4.29}$$

where $\mathbf{1}_d \in \mathbb{R}^d$ with all of its entries being 1. Then $\mathbf{x}(d)$ is a block K -sparse signal.

In the following, we plan to show that $\mathbf{C}(d)$ is symmetric positive definite. Then there exists an upper triangular matrix $\mathbf{A}(d)$ such that $\mathbf{A}^\top(d)\mathbf{A}(d) = \mathbf{C}(d)$. We will continue to show that $\mathbf{A}(d)$ satisfies the block-RIP of order $K + 1$ with $\delta_{K+1} = \delta$, and the BOMP algorithm may fail to recover the block K -sparse signal $\mathbf{x}(d)$ from $\mathbf{y}(d) = \mathbf{A}(d)\mathbf{x}(d)$ in K iterations.

By some tedious calculations, it is not hard to show that the eigenvalues $\{\lambda_i\}_{i=1}^{K+1}$ of $\mathbf{B}(1)$ are

$$\lambda_1 = \dots = \lambda_{K-1} = \frac{K}{K+1}, \lambda_K = 1 - \frac{1}{\sqrt{K+1}}, \lambda_{K+1} = 1 + \frac{1}{\sqrt{K+1}}.$$

To show $\mathbf{C}(d)$ is positive definite, we claim that for each $\mathbf{x} \in \mathbb{R}^{(d+1)K}$,

$$\left(1 - \frac{1}{\sqrt{K+1}}\right) \|\mathbf{x}\|_2^2 \leq \mathbf{x}^\top \mathbf{B}(d)\mathbf{x} \leq \left(1 + \frac{1}{\sqrt{K+1}}\right) \|\mathbf{x}\|_2^2. \tag{4.30}$$

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{K+1}$ with $u_i = v_i = \|\mathbf{x}[i]\|_2$ for $1 \leq i \leq K$ and $u_{K+1} = -v_{K+1} = \|\mathbf{x}[K+1]\|_2$. Then

$$\begin{aligned} \mathbf{x}^\top \mathbf{B}(d)\mathbf{x} &\stackrel{(a)}{=} \frac{K}{K+1} \sum_{i=1}^K \mathbf{x}^\top [i]\mathbf{x}[i] + \frac{K+2}{K+1} \mathbf{x}^\top [K+1]\mathbf{x}[K+1] + \frac{2}{K+1} \sum_{i=1}^K (\mathbf{x}^\top [i]\mathbf{x}[K+1]) \\ &\stackrel{(b)}{\leq} \frac{K}{K+1} \sum_{i=1}^K \|\mathbf{x}^\top [i]\|^2 + \frac{K+2}{K+1} \|\mathbf{x}^\top [K+1]\|_2^2 + \frac{2}{K+1} \sum_{i=1}^K (\|\mathbf{x}^\top [i]\|_2 \|\mathbf{x}[K+1]\|_2) \\ &= \frac{K}{K+1} \sum_{i=1}^K u_i^2 + \frac{K+2}{K+1} u_{K+1}^2 + \frac{2}{K+1} \sum_{i=1}^K (u_i u_{K+1}) \\ &= \mathbf{u}^\top \mathbf{B}(1)\mathbf{u} \leq \left(1 + \frac{1}{\sqrt{K+1}}\right) \|\mathbf{u}\|_2^2 = \left(1 + \frac{1}{\sqrt{K+1}}\right) \|\mathbf{x}\|_2^2, \end{aligned}$$

where (a) and (b) respectively follow from (4.27) and the Cauchy-Schwarz inequality. Similarly,

$$\begin{aligned} \mathbf{x}^\top \mathbf{B}(d)\mathbf{x} &\geq \frac{K}{K+1} \sum_{i=1}^K \|\mathbf{x}^\top [i]\|^2 + \frac{K+2}{K+1} \|\mathbf{x}^\top [K+1]\|_2^2 - \frac{2}{K+1} \sum_{i=1}^K (\|\mathbf{x}^\top [i]\|_2 \|\mathbf{x}[K+1]\|_2) \\ &= \frac{K}{K+1} \sum_{i=1}^K v_i^2 + \frac{K+2}{K+1} v_{K+1}^2 + \frac{2}{K+1} \sum_{i=1}^K (v_i v_{K+1}) \\ &= \mathbf{v}^\top \mathbf{B}(1)\mathbf{v} \geq \left(1 - \frac{1}{\sqrt{K+1}}\right) \|\mathbf{v}\|_2^2 = \left(1 - \frac{1}{\sqrt{K+1}}\right) \|\mathbf{x}\|_2^2. \end{aligned}$$

Thus, (4.30) follows. By (4.28) and (4.30), we have

$$(1 - \delta)\|\mathbf{x}\|_2^2 \leq \mathbf{x}^\top \mathbf{C}(d)\mathbf{x} \leq \left[1 + \left(\frac{2}{\sqrt{K+1}} - \delta\right)\right] \|\mathbf{x}\|_2^2.$$

By (3.9), we can see that $\mathbf{C}(d)$ is a symmetric positive definite matrix. Therefore, there exists an upper triangular matrix $\mathbf{A}(d)$ such that $\mathbf{A}^\top(d)\mathbf{A}(d) = \mathbf{C}(d)$. Furthermore, $\mathbf{A}(d)$ satisfies the block-RIP of order $K + 1$ with $\delta_{K+1} \leq \delta$.

Next we need to show that $\mathbf{A}(d)$ satisfies the block-RIP of order $K + 1$ with $\delta_{K+1} \geq \delta$, leading to $\delta_{K+1} = \delta$. By (4.30), it suffices to show there exists a $\mathbf{x} \in \mathbb{R}^{(d+1)K}$ such that

$$\mathbf{x}^\top \mathbf{B}(d)\mathbf{x} = \left(1 - \frac{1}{\sqrt{K+1}}\right) \|\mathbf{x}\|_2^2.$$

Indeed, in this case, we will have

$$\|\mathbf{A}(d)\mathbf{x}\|_2^2 = \mathbf{x}^\top \mathbf{C}(d)\mathbf{x} = \mathbf{x}^\top (\mathbf{B}(d) - s\mathbf{I}(d))\mathbf{x} = \left(1 - \frac{1}{\sqrt{K+1}} - s\right) \|\mathbf{x}\|_2^2 = (1 - \delta)\|\mathbf{x}\|_2^2.$$

Let $\bar{\mathbf{u}} \in \mathbb{R}^{K+1}$ be the eigenvector of $\mathbf{B}(1)$ corresponding to the eigenvalue $1 - \frac{1}{\sqrt{K+1}}$. Then

$$\bar{\mathbf{u}}^\top \mathbf{B}(1)\bar{\mathbf{u}} = \left(1 - \frac{1}{\sqrt{K+1}}\right) \|\bar{\mathbf{u}}\|_2^2.$$

Let $\mathbf{x} \in \mathbb{R}^{d(K+1)}$ satisfying $\mathbf{x}[i] = \bar{u}_i \mathbf{e}_1$ (here $\mathbf{e}_1 \in \mathbb{R}^d$ is the first column of the $d \times d$ identity matrix) for $1 \leq i \leq K + 1$. Then

$$\mathbf{x}^\top \mathbf{B}(d)\mathbf{x} = \left(1 - \frac{1}{\sqrt{K+1}}\right) \|\bar{\mathbf{u}}\|_2^2 = \left(1 - \frac{1}{\sqrt{K+1}}\right) \|\mathbf{x}\|_2^2.$$

Thus, $\mathbf{A}(d)$ satisfies the block-RIP of order $K + 1$ with $\delta_{K+1} = \delta$.

Finally, we show that BOMP may fail to recover the block K -sparse signal $\mathbf{x}(d)$ given in (4.29) from $\mathbf{y}(d) = \mathbf{A}(d)\mathbf{x}(d)$ in K iterations. By the fact that $\mathbf{A}^\top(d)\mathbf{A}(d) = \mathbf{C}(d)$ and (4.29), we have

$$\begin{aligned} \max_{j \in \Omega^c} \|(\mathbf{A}(d)[j])^\top \mathbf{y}\|_2 &= \|(\mathbf{A}(d)[K+1])^\top \mathbf{y}\|_2 = \|(\mathbf{A}(d)[K+1])^\top \mathbf{A}(d)\mathbf{x}(d)\|_2 \\ &= \|(\mathbf{0}_{dK \times d}, \mathbf{I}_d)^\top \mathbf{C}(d)\mathbf{x}(d)\|_2 = \frac{K\sqrt{d}}{K+1}. \end{aligned}$$

Similarly, for $1 \leq i \leq K$, we have

$$\begin{aligned} \|(\mathbf{A}(d)[i])^\top \mathbf{y}\|_2 &= \|(\mathbf{A}(d)[i])^\top \mathbf{A}(d)\mathbf{x}(d)\|_2 = \|(\mathbf{0}_{(i-1)K \times d}, \mathbf{I}_d, \mathbf{0}_{(K+1-i)K \times d})^\top \mathbf{A}^\top(d)\mathbf{A}(d)\mathbf{x}(d)\|_2 \\ &= \left(\frac{K}{K+1} - s\right)\sqrt{d}. \end{aligned}$$

Since $s \geq 0$, it follows that

$$\max_{i \in \Omega} \|(\mathbf{A}(d)[i])^\top \mathbf{y}\|_2 \leq \max_{j \in \Omega^c} \|(\mathbf{A}(d)[j])^\top \mathbf{y}\|_2.$$

Therefore, the BOMP algorithm may fail in the first iteration. As a result, the BOMP algorithm may fail to recover the block K -sparse signal $\mathbf{x}(d)$ in K iterations. \square

4.4. Proof of Lemma 5

By (3.10), we have

$$g_i(t) = \frac{t+1}{t-1} \left(1 - \frac{2}{t^i+1}\right).$$

Since $t > 1$, it is easy to see that (3.11) holds.

In the following, let us show that $g_i(t)$ is strictly decreasing with t for $i \geq 2$. By (3.10) and a simple calculation, we have

$$g_i(t) = 1 + 2\phi(t) \text{ with } \phi(t) = \frac{(t^i - t)}{t^{i+1} - t^i + t - 1}.$$

Thus, we only need to show $\phi(t)$ is strictly decreasing with t for $i \geq 2$. By direct calculation, we have

$$\phi'(t) = \frac{\bar{\phi}(t)}{(t^{i+1} - t^i + t - 1)^2},$$

where

$$\bar{\phi}(t) = (1 + it^{i+1}) - (t^{2i} + it^{i-1}).$$

Thus, we only need to show $\bar{\phi}(t) < 0$ for $t > 1$ and $i \geq 2$. An elementary calculation yields

$$\bar{\phi}'(t) = i(i+1)t^i - (2it^{2i-1} + i(i-1)t^{i-2}) = t^{i-2}i[(i+1)t^2 - (2t^{i+1} + i - 1)] < 0$$

for $t > 1$. Thus, $\bar{\phi}(t) < \bar{\phi}(1) = 0$. Therefore, $g_i(t)$ is decreasing with t for $i \geq 2$. By some simple calculations, one can show that

$$\lim_{t \rightarrow 1} g_i(t) = i, \quad \lim_{t \rightarrow +\infty} g_i(t) = 1.$$

Thus, $1 < g_i(t) < i$ for $i \geq 2$. □

4.5. Proof of Theorem 3

Before proving Theorem 3, we need to introduce the following lemma which is a variant of [20, Lemma 1].

Lemma 7. Let $\alpha > 1$ and $\beta \geq \gamma \geq 1$ be given constants and let

$$\bar{g}_i(t_1, \dots, t_i, \beta, \gamma) = \frac{(\sum_{j=1}^{i-1} t_j + \beta t_i)^2}{\sum_{j=1}^{i-1} t_j^2 + \gamma t_i^2}, \quad 1 \leq i \leq K, \tag{4.31}$$

where

$$t_1 \geq \alpha t_2 \geq \dots \geq \alpha^{K-1} t_K > 0, \quad \sum_{j=1}^0 \cdot = 0. \tag{4.32}$$

Then, for $1 \leq i \leq K$, \bar{g}_i is increasing with t_i , and

$$\bar{g}_i(t_1, \dots, t_i, 1, 1) \leq g_i(\alpha), \tag{4.33}$$

where function g_i is defined in (3.10).

Proof. We first show that \bar{g}_i is increasing with t_i . By (4.31), we have

$$\frac{\partial \bar{g}_i}{\partial t_i} = 2 \left(\sum_{j=1}^{i-1} t_j + \beta t_i \right) \frac{\beta \left(\sum_{j=1}^{i-1} t_j^2 + \gamma t_i^2 \right) - \left(\sum_{j=1}^{i-1} t_j + \beta t_i \right) \gamma t_i}{\left(\sum_{j=1}^{i-1} t_j^2 + \gamma t_i^2 \right)^2}.$$

By (4.32), and the facts that $\alpha > 1$ and $\beta \geq \gamma > 1$, it is easy to see that $\frac{\partial \bar{g}_i}{\partial t_i} > 0$, so \bar{g}_i is increasing with t_i .
By (3.11), (4.33) holds with $i = 1$. For $2 \leq i \leq K$, we have

$$\begin{aligned} \bar{g}_i(t_1, \dots, t_i, 1, 1) &\leq \bar{g}_i(t_1, \dots, t_{i-1}, \alpha^{-1}t_{i-1}, 1, 1) = \bar{g}_{i-1}(t_1, \dots, t_{i-1}, (1 + \alpha^{-1}), (1 + \alpha^{-2})) \\ &\leq \dots \leq \bar{g}_1(t_1, \left(\sum_{j=1}^i \alpha^{1-j} \right), \left(\sum_{j=1}^i \alpha^{2-2j} \right)) = g_i(\alpha). \end{aligned}$$

Thus, (4.33) hold for $1 \leq i \leq K$. \square

Proof of Theorem 3. If $s^{-2} - 1 \geq K$, then $s \leq \frac{1}{\sqrt{1+K}}$. By (3.15), $\delta_{K+1} < \frac{1}{\sqrt{1+K}}$. Thus, by Corollary 1, Theorem 3 holds.

Next we consider the case that $s^{-2} - 1 < K$. By (3.12), we have

$$g_K(g_K^{-1}(s^{-2} - 1)) = s^{-2} - 1.$$

Let $f(t) = \min\{s^{-2} - 1, t\}$. By Theorem 1, it suffices to show (3.4) holds under (3.15) and (3.14).

By Definition 1, if \mathbf{x} is a block α -strongly-decaying K -sparse signal, so is $\mathbf{x}[\Omega \setminus S]$ for any subset S of Ω . By (4.31),

$$\begin{aligned} \frac{\|\mathbf{x}[\Omega \setminus S]\|_{2,1}^2}{\|\mathbf{x}[\Omega \setminus S]\|_2^2} &= \bar{g}_{|\Omega \setminus S|}(\|\mathbf{x}[\Omega \setminus S][1]\|_2, \dots, \|\mathbf{x}[\Omega \setminus S][|\Omega \setminus S]|\|_2, 1, 1) \\ &\stackrel{(a)}{\leq} g_{|\Omega \setminus S|}(\alpha) \stackrel{(b)}{<} g_{|\Omega \setminus S|}(g_K^{-1}(s^{-2} - 1)) \stackrel{(c)}{<} g_K(g_K^{-1}(s^{-2} - 1)) = s^{-2} - 1, \end{aligned}$$

where (a) follows from (4.33) and Definition 1, (b) is from (3.14) and Lemma 5, and (c) follows from (3.11).

By the Cauchy-Schwarz inequality, obviously,

$$\|\mathbf{x}[\Omega \setminus S]\|_{2,1}^2 \leq |\Omega \setminus S| \|\mathbf{x}[\Omega \setminus S]\|_2^2.$$

Thus, the above two inequalities imply that (3.4) holds with $f(t) = \min\{s^{-2} - 1, t\}$. Therefore, Theorem 3 follows from Theorem 1. \square

4.6. Proof of Theorem 4

Proof. For any given positive integers $d, K \geq 1$, δ satisfies (3.17), and $\alpha > 1$, let

$$\mathbf{B}(d) = \begin{bmatrix} \frac{1+\sqrt{2}-2\delta}{2} \mathbf{I}_d & \frac{1}{2} \mathbf{I}_d & \mathbf{0}_{d \times d(K-1)} \\ \frac{1}{2} \mathbf{I}_d & \frac{3+\sqrt{2}-2\delta}{2} \mathbf{I}_d & \mathbf{0}_{d \times d(K-1)} \\ \mathbf{0}_{d(K-1) \times d} & \mathbf{0}_{d(K-1) \times d} & \frac{1+\sqrt{2}-2\delta}{2} \mathbf{I}_{d(K-1)} \end{bmatrix}$$

and

$$\mathbf{x}(d) = (a_1 \mathbf{1}_d; \mathbf{0}_d; a_2 \mathbf{1}_d; \dots; a_K \mathbf{1}_d)^\top \in \mathbb{R}^{d(K+1)},$$

where all the entries of $\mathbf{1}_d \in \mathbb{R}^d$ are 1, and

$$a_1 \geq \alpha a_2 \geq \dots \geq \alpha^{K-1} a_K.$$

Thus, $\mathbf{x}(d)$ is a block α -strongly-decaying K -sparse signal.

Following the proof of Theorem 2, one can easily show that $\mathbf{B}(d)$ is symmetric positive definite. Thus, there exists an upper triangular matrix $\mathbf{A}(d)$ such that $\mathbf{A}^\top(d)\mathbf{A}(d) = \mathbf{B}(d)$. Moreover, one can similarly show that $\mathbf{A}(d)$ satisfies the block-RIP of order $K + 1$ with $\delta_{K+1} = \delta$.

In the following, we show that BOMP may fail to recover the block K -sparse signal $\mathbf{x}(d)$ from $\mathbf{y}(d) = \mathbf{A}(d)\mathbf{x}(d)$ in K iterations. By the fact that $\mathbf{A}^\top(d)\mathbf{A}(d) = \mathbf{B}(d)$, we have

$$\max_{j \in \Omega^c} \|\mathbf{A}^\top(d)[j]\mathbf{y}\|_2 = \|\mathbf{A}^\top(d)[2]\mathbf{y}\|_2 = \|(\mathbf{0}_{d \times d}, \mathbf{I}_d, \mathbf{0}_{d(K-1) \times d})\mathbf{A}^\top(d)\mathbf{A}(d)\mathbf{x}(d)\|_2 = \frac{1}{2}a_1\sqrt{d}.$$

Similarly, for $i \in \Omega$, we have

$$\|\mathbf{A}^\top(d)[i]\mathbf{y}\|_2 = \|\mathbf{A}^\top(d)[i]\mathbf{A}(d)\mathbf{x}(d)\|_2 = \frac{1 + \sqrt{2} - 2\delta}{2}a_i\sqrt{d}.$$

Thus, by (3.17), it holds that

$$\max_{i \in \Omega} \|\mathbf{A}^\top(d)[i]\mathbf{y}\|_2 = \|\mathbf{A}^\top(d)[1]\mathbf{y}\|_2 \leq \max_{j \in \Omega^c} \|\mathbf{A}^\top(d)[j]\mathbf{y}\|_2.$$

Therefore, BOMP may fail in the first iteration. As a result, the BOMP algorithm may fail to recover the block α -strongly-decaying K -sparse signal $\mathbf{x}(d)$ in K iterations. \square

4.7. Proof of Theorem 5

Before proving Theorem 5, we need to introduce the following lemma.

Lemma 8. Let $\beta \geq 1$ be a given constant and let

$$\bar{h}_i(t_1, \dots, t_i, \beta) = \frac{(\sum_{j=1}^{i-1} t_j + \beta t_i)^2}{\sum_{j=1}^{i-1} t_j^2 + \beta t_i^2}, \quad 1 \leq i \leq K, \tag{4.34}$$

where

$$t_1 \geq \alpha_1 t_2 \geq \dots \geq \alpha_{K-1} t_K > 0, \quad \sum_{j=1}^0 \cdot = 0 \tag{4.35}$$

for given $\alpha_1, \dots, \alpha_{K-1} \geq 1$. Then, \bar{h}_i is increasing with t_i . Moreover, for $2 \leq i \leq K$,

$$\bar{h}_i(t_1, \dots, t_i, 1) = \hat{h}_{i-1}(t_1/t_2) \leq \hat{h}_{i-1}(\alpha_1), \tag{4.36}$$

where \hat{h}_{i-1} is defined in (3.18).

Proof. By (4.31) and (4.34), we have

$$\bar{h}_i(t_1, \dots, t_i, \beta) = \bar{g}_i(t_1, \dots, t_i, \beta, \beta).$$

By Lemma 7, \bar{h}_i is increasing with t_i . Next we prove (4.36).

Since \bar{h}_i is increasing with t_i , by (4.35) and the fact that $\alpha_i \geq 1$ for $1 \leq i \leq K - 1$, we have

$$\begin{aligned} \bar{h}_i(t_1, \dots, t_i, 1) &\leq \bar{h}_i(t_1, \dots, t_{i-1}, t_{i-1}, 1) = \bar{h}_{i-1}(t_1, \dots, t_{i-1}, 2) \\ &\leq \dots \leq \bar{h}_2(t_1, t_2, i - 1) = \hat{h}_{i-1}(t_1/t_2) \leq \hat{h}_{i-1}(\alpha_1), \end{aligned}$$

where the last inequality follows from Lemma 6. \square

Proof of Theorem 5. We prove it by induction. We first show that the BOMP algorithm chooses an index in Ω in the first iteration.

If $s^{-2} - 1 \geq K$, then $s \leq \frac{1}{\sqrt{1+K}}$. By (3.21), $\delta_{K+1} < \frac{1}{\sqrt{1+K}}$. Thus, by Corollary 1, the BOMP algorithm chooses an index in Ω in the first iteration.

If $s^{-2} - 1 < K$, we let $\alpha_1 = \|\mathbf{x}[1]\|_2 / \|\mathbf{x}[2]\|_2$ and $f(t) = \min\{s^{-2} - 1, t\}$. Thus, by Theorem 3, to show the BOMP algorithm chooses an index in Ω in the first iteration, it suffices to show (3.4) holds with $S = \emptyset$ under (3.20), which is equivalent to show

$$\frac{\|\mathbf{x}\|_{2,1}^2}{\|\mathbf{x}\|_2^2} \leq s^{-2} - 1.$$

By Lemma 8 with $\alpha_2 = \dots = \alpha_{K-1} = 1$, Lemma 6, and (3.18)–(3.21), we have

$$\frac{\|\mathbf{x}\|_{2,1}^2}{\|\mathbf{x}\|_2^2} = \bar{h}_K(\|\mathbf{x}[1]\|_2, \dots, \|\mathbf{x}[K]\|_2, 1) \leq \hat{h}_{K-1}(\alpha_1) \leq \hat{h}_{K-1}(h_{K-1}(s^{-2} - 1)) = s^{-2} - 1.$$

Let us assume that the BOMP algorithm chooses an index in Ω in the first k iterations. Then, we show the BOMP chooses an index in Ω in the $(k + 1)$ -th iteration ($1 \leq k \leq i - 1$).

Suppose S is the set of all the indexes chosen by the BOMP algorithm in the first k iterations. By Theorem 1, it suffices to show (3.4) holds under (3.20), which is equivalent to show

$$\frac{\|\mathbf{x}[\Omega \setminus S]\|_{2,1}^2}{\|\mathbf{x}[\Omega \setminus S]\|_2^2} \leq f(|\Omega \setminus S|) = \min\{s^{-2} - 1, |\Omega \setminus S|\}. \tag{4.37}$$

By the Cauchy–Schwarz inequality and (2.1), we have

$$\|\mathbf{x}[\Omega \setminus S]\|_{2,1}^2 \leq (|\Omega \setminus S|)\|\mathbf{x}[\Omega \setminus S]\|_2^2.$$

Thus, if $s^{-2} - 1 \geq K - |S| \geq |\Omega \setminus S|$, then (4.37) holds. Hence, we can use Theorem 1.

It remains to consider the case $s^{-2} - 1 < K - |S|$. Let us prove

$$\frac{\|\mathbf{x}[\Omega \setminus S]\|_{2,1}^2}{\|\mathbf{x}[\Omega \setminus S]\|_2^2} \leq s^{-2} - 1.$$

To this end, let

$$\alpha_1 = \min_{1 \leq j \leq k+1} \frac{\|\mathbf{x}[j]\|_2}{\|\mathbf{x}[j+1]\|_2}.$$

Then, we have

$$\|(\mathbf{x}[\Omega \setminus S])[1]\|_2 \geq \alpha_1 \|(\mathbf{x}[\Omega \setminus S])[2]\|_2. \tag{4.38}$$

And by (3.19)–(3.21), we have

$$\alpha_1 \geq h_{K-1}(s^{-2} - 1) \geq h_{K-(k+1)}(s^{-2} - 1).$$

By Lemma 8 with $\alpha_2 = \dots = \alpha_{K-1} = 1$, Lemma 6, and (3.18)–(3.21), we have

$$\begin{aligned} \frac{\|\mathbf{x}[\Omega \setminus S]\|_{2,1}^2}{\|\mathbf{x}[\Omega \setminus S]\|_2^2} &= \bar{h}_{K-k}(\|(\mathbf{x}[\Omega \setminus S])[1]\|_2, \dots, \|(\mathbf{x}[\Omega \setminus S])[K-k]\|_2, 1) \\ &\leq \hat{h}_{K-(k+1)}(\alpha_1) \leq \hat{h}_{K-(k+1)}(h_{K-(k+1)}(s^{-2} - 1)) = s^{-2} - 1. \end{aligned}$$

Thus, with (4.38), we have (4.37) and hence, we can apply Theorem 1. These complete the proof. \square

5. Conclusions and future work

In this paper, we studied block-RIC based sufficient and necessary conditions for exact and stable recovery of block K -sparse signals \mathbf{x} from the measurements $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}$ by the BOMP algorithm in K iterations, where \mathbf{v} is an ℓ_2 -bounded noise vector. First, we showed that $\delta_{K+1} < 1/\sqrt{K+1}$ is a sharp sufficient condition. Second, under $\delta_{K+1} < \sqrt{2}/2$, we presented a class of block K -sparse signals which can be recovered exactly or stably by using the BOMP algorithm in K iterations. That is, we showed that under some conditions on α , the BOMP algorithm can exactly or stably recover any block α -strongly-decaying K -sparse signal in K iterations when $\delta_{K+1} < \sqrt{2}/2$. Moreover, we proved that the condition on α is sharp. Finally, we showed that $\delta_{K+1} < \sqrt{2}/2$ is a sufficient condition for partial recovery of block K -sparse signals \mathbf{x} by the BOMP algorithm under certain condition on the signals \mathbf{x} . The condition on the blocks of block K -sparse signals \mathbf{x} is also shown to be sharp for partial recovery.

In the following, we point out some research problems that will be investigated in the future

1. It is possible to use more than K iterations of OMP to recover a sparse signal as discussed in [66,42,43,15,57] or choose more than one indexes per iteration (see [39,41,55]). It can be expected that less restrictive sufficient conditions can be obtained, and we will investigate it in the future.
2. We are interested in extending our current work to this setting of study. For example, besides α -strongly-decaying sparse signals, what other signals can be recovered exactly or stably by using the BOMP in K iterations?
3. We assume that all the blocks have identical size of d in this paper. As a future work, we will extend our work to the case with non-uniform block sizes. Moreover, we are interested in investigating sufficient conditions of recovering α -strongly-decaying sparse signals by solving the ℓ_1 -minimization problem (see (1.3)).
4. An extension of the BOMP algorithm called as model-aware orthogonal matching pursuit (MA-OMP) has recently been proposed in [58] to recover the coherent MIMO radar imaging. Exact recovery of block K -sparse signals with structure by MA-OMP has been investigated in [62]. Whether the techniques developed in this paper can be used to develop a sharp condition for MA-OMP will be studied in the future.

Acknowledgments

This work was partially supported by the National Science Foundation of China under Grant 61661146003, the Joint fund of the Ministry of Education under Grant 6141A020223, Simons collaboration grant (No. 280646) and the National Science Foundation under the grant (No. DMS 1521537), and Major Frontier Project of Sichuan Province under Grant 2015JY0282.

Part of this work was carried out while Jinming Wen and Ming-Jun Lai were visiting Zhengchun Zhou and Xiaohu Tang at Southwest Jiaotong University. The hospitality received in this period is gratefully acknowledged.

References

- [1] E. Amaldi, V. Kann, On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems, *Theoret. Comput. Sci.* 209 (1998) 237–260.
- [2] A.S. Bandeira, E. Dobriban, D.G. Mixon, W.F. Sawin, Certifying the restricted isometry property is hard, *IEEE Trans. Inform. Theory* 59 (2013) 3448–3450.
- [3] B. Baraniuk, M. Davenport, R. DeVore, M. Wakin, A simple proof of the restricted isometry property for random matrices, *Constr. Approx.* 28 (2008) 253–263.
- [4] T. Blumensath, M.E. Davies, Sampling theorems for signals from the union of finite-dimensional linear subspaces, *IEEE Trans. Inform. Theory* 55 (2009) 1872–1882.
- [5] T. Cai, L. Wang, Orthogonal matching pursuit for sparse signal recovery with noise, *IEEE Trans. Inform. Theory* 57 (2011) 4680–4688.
- [6] T. Cai, A. Zhang, Sparse representation of a polytope and recovery of sparse signals and low-rank matrices, *IEEE Trans. Inform. Theory* 60 (2013) 122–132.
- [7] E.J. Candès, The restricted isometry property and its implications for compressed sensing, *C. R. Acad. Sci. Paris, Ser. I* 346 (2008) 589–592.
- [8] E.J. Candès, J. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, *Comm. Pure Appl. Math.* 59 (2006) 1207–1223.
- [9] E.J. Candès, T. Tao, Decoding by linear programming, *IEEE Trans. Inform. Theory* 51 (2005) 4203–4215.
- [10] E.J. Candès, T. Tao, The Dantzig selector: statistical estimation when p is much larger than n , *Ann. Statist.* 35 (2007) 2313–2351.
- [11] L.-H. Chang, J.-Y. Wu, An improved RIP-based performance guarantee for sparse signal recovery via orthogonal matching pursuit, *IEEE Trans. Inform. Theory* 60 (2014) 707–710.
- [12] J. Chen, X. Huo, Theoretical results on sparse representations of multiple-measurement vectors, *IEEE Trans. Signal Process.* 54 (2006) 4634–4643.
- [13] W. Chen, H. Ge, A sharp recovery condition for block sparse signals by block orthogonal multi-matching pursuit, *Sci. China Math.* 60 (2017) 1325–1340.
- [14] A. Cohen, W. Dahmen, R. DeVore, Compressed sensing and best k -term approximation, *J. Amer. Math. Soc.* 22 (2009) 211–231.
- [15] A. Cohen, W. Dahmen, R. DeVore, Orthogonal matching pursuit under the restricted isometry property, *Constr. Approx.* 45 (2017) 113–127.
- [16] S.F. Cotter, B.D. Rao, K. Engan, K. Kreutz-Delgado, Sparse solutions to linear inverse problems with multiple measurement vectors, *IEEE Trans. Signal Process.* 53 (2005) 2477–2488.
- [17] W. Dai, O. Milenkovic, Subspace pursuit for compressive sensing signal reconstruction, *IEEE Trans. Inform. Theory* 55 (2009) 2230–2249.
- [18] M. Davenport, M. Wakin, Analysis of orthogonal matching pursuit using the restricted isometry property, *IEEE Trans. Inform. Theory* 56 (2010) 4395–4401.
- [19] M.E. Davies, R. Gribonval, Restricted isometry constants where l^p sparse recovery can fail for $0 < p \leq 1$, *IEEE Trans. Inform. Theory* 55 (2009) 2203–2214.
- [20] J. Ding, L. Chen, Y. Gu, Perturbation analysis of orthogonal matching pursuit, *IEEE Trans. Signal Process.* 61 (2013) 398–410.
- [21] D.L. Donoho, Compressed sensing, *IEEE Trans. Inform. Theory* 52 (2006) 1289–1306.
- [22] D.L. Donoho, For most large underdetermined systems of linear equations the minimal l_1 -norm solution is also the sparsest solution, *Comm. Pure Appl. Math.* 59 (2006) 797–829.
- [23] D.L. Donoho, M. Elad, V.N. Temlyakov, Stable recovery of sparse overcomplete representations in the presence of noise, *IEEE Trans. Inform. Theory* 52 (2006) 6–18.
- [24] D.L. Donoho, X. Huo, Uncertainty principles and ideal atomic decomposition, *IEEE Trans. Inform. Theory* 47 (2001) 2845–2862.
- [25] M. Elad, M.A. Figueiredo, Y. Ma, On the role of sparse and redundant representations in image processing, *Proc. IEEE* 98 (2010) 972–982.
- [26] Y.C. Eldar, Compressed sensing of analog signals in shift-invariant spaces, *IEEE Trans. Signal Process.* 57 (2009) 2986–2997.
- [27] Y.C. Eldar, P. Kuppinger, H. Bölcskei, Block-sparse signals: uncertainty relations and efficient recovery, *IEEE Trans. Signal Process.* 58 (2010) 3042–3054.
- [28] Y.C. Eldar, M. Mishali, Robust recovery of signals from a structured union of subspaces, *IEEE Trans. Inform. Theory* 55 (2009) 5302–5316.
- [29] E. Elhamifar, R. Vidal, Sparse subspace clustering: algorithm, theory, and applications, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (2013) 2765–2781.
- [30] S. Foucart, A note on guaranteed sparse recovery via l_1 -minimization, *Appl. Comput. Harmon. Anal.* 29 (2010) 97–103.
- [31] S. Foucart, Stability and robustness of weak orthogonal matching pursuits, in: *Recent Advances in Harmonic Analysis and Applications*, Springer, 2012, pp. 395–405.

- [32] S. Foucart, M.-J. Lai, Sparsest solutions of underdetermined linear systems via l_q -minimization for $0 < q \leq 1$, *Appl. Comput. Harmon. Anal.* 26 (2009) 395–407.
- [33] S. Foucart, H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Birkhäuser Verlag, Basel, 2013.
- [34] J.J. Fuchs, Recovery of exact sparse representations in the presence of bounded noise, *IEEE Trans. Inform. Theory* 51 (2005) 3601–3608.
- [35] K. Gedalyahu, Y.C. Eldar, Time-delay estimation from low-rate samples: a union of subspaces approach, *IEEE Trans. Signal Process.* 58 (2010) 3017–3031.
- [36] E.A. Habets, S. Gannot, I. Cohen, Late reverberant spectral variance estimation based on a statistical model, *IEEE Signal Process. Lett.* 16 (2009) 770–773.
- [37] Z. He, S. Xie, S. Ding, A. Cichocki, Convolutional blind source separation in the frequency domain based on sparse representation, *IEEE/ACM Trans. Audio Speech Lang. Process.* 15 (2007) 1551–1563.
- [38] C. Herzet, A. Drémeau, C. Soussen, Relaxed recovery conditions for OMP/OLS by exploiting both coherence and decay, *IEEE Trans. Inform. Theory* 62 (2016) 459–470.
- [39] S. Huang, J. Zhu, Recovery of sparse signals using OMP and its variants: convergence analysis based on RIP, *Inverse Probl.* 27 (2011).
- [40] M.-J. Lai, Y. Liu, The null space property for sparse recovery from multiple measurement vectors, *Appl. Comput. Harmon. Anal.* 30 (2011) 402–406.
- [41] E. Liu, V. Temlyakov, The orthogonal super greedy algorithm and applications in compressed sensing, *IEEE Trans. Inform. Theory* 58 (2012) 2040–2047.
- [42] E.D. Livshits, On the efficiency of the orthogonal matching pursuit in compressive sensing, *Sb. Math.* 23 (2012) 183–195.
- [43] E.D. Livshitz, V.N. Temlyakov, Sparse approximation and recovery by greedy algorithms, *IEEE Trans. Inform. Theory* 60 (2014) 3989–4000.
- [44] M. Mishali, Y.C. Eldar, Blind multiband signal reconstruction: compressed sensing for analog signals, *IEEE Trans. Signal Process.* 57 (2009) 993–1009.
- [45] M. Mishali, Y.C. Eldar, From theory to practice: sub-Nyquist sampling of sparse wideband analog signals, *IEEE J. Sel. Top. Signal Process.* 4 (2010) 375–391.
- [46] Q. Mo, A sharp restricted isometry constant bound of orthogonal matching pursuit, arXiv:1501.01708, 2015.
- [47] Q. Mo, Y. Shen, A remark on the restricted isometry property in orthogonal matching pursuit, *IEEE Trans. Inform. Theory* 58 (2012) 3654–3656.
- [48] D. Needell, J.A. Tropp, CoSaMP: iterative signal recovery from incomplete and inaccurate samples, *Appl. Comput. Harmon. Anal.* 26 (2009) 301–321.
- [49] Y. Shen, B. Li, W. Pan, J. Li, Analysis of generalized orthogonal matching pursuit using restricted isometry constant, *Electron. Lett.* 50 (2014) 1020–1022.
- [50] G. Swirszcz, N. Abe, A.C. Lozano, Grouped orthogonal matching pursuit for variable selection and prediction, in: *Advances in Neural Information Processing Systems*, 2009, pp. 1150–1158.
- [51] A.M. Tillmann, M.E. Pfetsch, The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing, *IEEE Trans. Inform. Theory* 60 (2014) 1248–1259.
- [52] J.A. Tropp, A.C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, *IEEE Trans. Inform. Theory* 53 (2007) 4655–4666.
- [53] E. Van Den Berg, M.P. Friedlander, Theoretical and empirical results for recovery from multiple measurements, *IEEE Trans. Inform. Theory* 56 (2010) 2516–2527.
- [54] E. Vincent, N. Bertin, R. Gribonval, F. Bimbot, From blind to guided audio source separation: how models and side information can improve the separation of sound, *IEEE Signal Process. Mag.* 31 (2014) 107–115.
- [55] J. Wang, S. Kwon, B. Shim, Generalized orthogonal matching pursuit, *IEEE Trans. Signal Process.* 60 (2012) 6202–6216.
- [56] J. Wang, B. Shim, On the recovery limit of sparse signals using orthogonal matching pursuit, *IEEE Trans. Signal Process.* 60 (2012) 4973–4976.
- [57] J. Wang, B. Shim, Exact recovery of sparse signals via orthogonal matching pursuit: how many iterations do we need?, *IEEE Trans. Signal Process.* 64 (August 2016) 4194–4202.
- [58] L. Weiland, T. Wiese, W. Utschick, Coherent MIMO radar imaging with model-aware block sparse recovery, in: *2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, CAMSAP, Dec. 2015, pp. 425–428.
- [59] J. Wen, D. Li, F. Zhu, Stable recovery of sparse signals via l_p -minimization, *Appl. Comput. Harmon. Anal.* 38 (2015) 161–176.
- [60] J. Wen, Z. Zhou, J. Wang, X. Tang, Q. Mo, A sharp condition for exact support recovery with orthogonal matching pursuit, *IEEE Trans. Signal Process.* 65 (2017) 1370–1382.
- [61] J. Wen, X. Zhu, D. Li, Improved bounds on the restricted isometry constant for orthogonal matching pursuit, *Electron. Lett.* 49 (2013) 1487–1489.
- [62] T. Wiese, L. Weiland, W. Utschick, Exact recovery of structured block-sparse signals with model-aware orthogonal matching pursuit, in: *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications*, SPAWC, July 2016, pp. 1–5.
- [63] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T.S. Huang, S. Yan, Sparse representation for computer vision and pattern recognition, *Proc. IEEE* 98 (2010) 1031–1044.
- [64] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, Y. Ma, Robust face recognition via sparse representation, *IEEE Trans. Pattern Anal. Mach. Intell.* 31 (2009) 210–227.
- [65] Z. Xu, The performance of orthogonal multi-matching pursuit under RIP, *J. Comput. Math.* 33 (2015) 495–516.
- [66] T. Zhang, Sparse recovery with orthogonal matching pursuit in compressed sensing, *IEEE Trans. Inform. Theory* 57 (2011) 6215–6221.