Received December 14, 2018, accepted January 1, 2019, date of publication January 16, 2019, date of current version January 29, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2891568*

# Sparse Recovery With Block Multiple Measurement Vectors Algorithm

**YANLI SHI[1], LIBO WANG[2], AND RONG LUO[3]**
[1]School of Science, Jilin Institute of Chemical Technology, Jilin 132022, China
[2]College of Information Science and Technology, Jinan University, Guangzhou 510632, China
[3]College of Mathematics, Southwest Jiaotong University, Chengdu 610031, China

Corresponding author: Libo Wang (wanglibo12b@mails.ucas.edu.cn)

The work of Y. Shi was supported in part by the Science and Technology Innovation Development Program of Jilin City under Grant 201831769, and in part by the "Thirteen Five" Science and Technology Program, Department of Education, Jilin, under Grant JJKH20180552KJ. The work of L. Wang was supported in part by the Fundamental Research Funds for the Central Universities under Grant 21618331, in part by the National Natural Science Foundation of China under Grant 11871248, and in part by the National Key R&D Plan of China under Grant 2018YFB1003701. The work of R. Luo was supported by the Sichuan Science and Technology Program under Grant 2018JY0261.

**ABSTRACT** This paper investigates the performance of the block multiple measurement vectors (BMMV) algorithm in reconstructing block joint sparse matrices. We prove that if $\Phi$ obeys block restricted isometry property with $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$, then BMMV perfectly reconstructs any block $K$-joint sparse matrix $\boldsymbol{X}$ from observations $\boldsymbol{Y} = \Phi\boldsymbol{X}$ in $K$ iterations. We also show that BMMV may not reconstruct block $K$-joint sparse matrices in $K$ iterations under the condition $\delta_{K+1} \geq \frac{1}{\sqrt{K+1}}$. That is to say, the condition $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$ is optimal for the BMMV algorithm.

**INDEX TERMS** Sparse recovery, block restricted isometry property, block multiple measurement vectors (BMMV) algorithm.

## I. INTRODUCTION

In many application domains, such as multivariate regression [1], face recognition [2], direction of arrival estimation of multiple narrowband signals [3], [4], we need to reconstruct sparse matrix $\boldsymbol{X} \in \mathbb{R}^{N\times P}$ from the model

$$\boldsymbol{Y} = \Phi\boldsymbol{X}, \tag{1}$$

where $\boldsymbol{Y} \in \mathbb{R}^{M\times P}$ is an observation matrix, $\Phi \in \mathbb{R}^{M\times N}$ is a sensing matrix with $M \ll N$.

If $P = 1$, then the model (1) degenerates to

$$\boldsymbol{y} = \Phi\boldsymbol{x}. \tag{2}$$

The model (2) has a close relationship with a lot of applications, for more details, see [5]–[8] and references therein. Many effective and efficient greedy algorithms were proposed to reconstruct $\boldsymbol{x}$ in (2), for example, orthogonal matching pursuit [9], [10], generalized orthogonal matching pursuit [11] and subspace pursuit [13]. Various sufficient conditions were proposed for perfect reconstructing in model (2) with the above algorithms [12]–[21].

A number of effective and efficient algorithms have also been proposed to reconstruct $\boldsymbol{X}$ in (1), such as, MMV orthogonal matching pursuit and MMV order recursive matching pursuit [22]. There are also some other reconstructing algorithms and theoretical results, see, e.g., [23]–[25].

In many applications area including reconstructing multi-band signals [26], face recognition [27], the nonzero rows of matrix $\boldsymbol{X}$ appear in a few blocks, that is to say, the matrix is block joint sparse. To define block joint sparsity, we can view matrix $\boldsymbol{X}$ as concatenation of blocks of rows. Like in [28], we assume the lengths of all the blocks are $d$. Thus, we can rewrite $\boldsymbol{X}$ as:

$$\boldsymbol{X} = [\boldsymbol{X}[1]^T, \boldsymbol{X}[2]^T, \ldots, \boldsymbol{X}[L-1]^T, \boldsymbol{X}[L]^T]^T,$$

where, for $1 \leq i \leq L$,

$$\boldsymbol{X}[i] = [\boldsymbol{X}^T_{d(i-1)+1}, \boldsymbol{X}^T_{d(i-1)+2}, \ldots, \boldsymbol{X}^T_{di-1}, \boldsymbol{X}^T_{di}]^T$$

with $\boldsymbol{X}_j$ being the $j$-th row of $\boldsymbol{X}$. The block joint sparsity of a matrix $\boldsymbol{X} \in \mathbb{R}^{N\times P}$ is $K$ means that there are at most $K$ blocks $\boldsymbol{X}[i]$ are different from $d \times P$ zero matrix. Clearly, if $P = 1$, then $\boldsymbol{X}$ becomes a vector, and a block $K$-joint sparse matrix turns to a block $K$-sparse vector. The sensing matrix $\Phi$ can be rewritten as:

$$\Phi = [\Phi[1], \Phi[2], \ldots, \Phi[L-1], \Phi[L]],$$

where, for $1 \le i \le L$,

$$\Phi[i] = [\Phi_{d(i-1)+1}, \Phi_{d(i-1)+2}, \ldots, \Phi_{di-1}, \Phi_{di}]$$

with $\Phi_j$ being the $j$-th column of $\Phi$.

To analyze the reconstruction performance of block algorithms, the restricted isometry property (RIP) [5] was extended to block RIP in [29]. Specifically, $\Phi$ is said to satisfy the block RIP with parameter $\delta_{BK}$ if

$$(1 - \delta_{BK})\|\boldsymbol{x}\|_2^2 \le \|\Phi \boldsymbol{x}\|_2^2 \le (1 + \delta_{BK})\|\boldsymbol{x}\|_2^2 \quad (3)$$

for all block $K$-sparse vectors $\boldsymbol{x}$. The smallest constant $\delta_{BK}$ satisfying (3) is called block restricted isometry constant (RIC) of $\Phi$ with order $K$. By abuse of notation, we simply denote it by $\delta_K$.

Based on the multiple measurement vectors algorithm [2], a block multiple measurement vectors algorithm (BMMV) was proposed in [28], to reconstruct block joint sparse matrices by taking the block joint sparsity into account. For $\Gamma \subset \{1, 2, \ldots, L\}$, we denote its cardinality by $|\Gamma|$. Let $\Phi[\Gamma] \in \mathbb{R}^{M \times |\Gamma|}$ be the submatrix of $\Phi$ that only consist of the blocks of columns indexed by $\Gamma$ and $\boldsymbol{X}[\Gamma] \in \mathbb{R}^{|\Gamma| \times P}$ be the submatrix of $\boldsymbol{X}$ that only consist of the blocks of rows indexed by $\Gamma$, respectively. Then formally the BMMV algorithm [28] can be described as the following Algorithm 1.

**Algorithm 1** BMMV [28]

Input: $\boldsymbol{Y}$, $\Phi$, and sparsity $K$.
Initialize: $k = 0, \boldsymbol{R}^0 = \boldsymbol{Y}, \Lambda_0 = \emptyset$.
While $k < K$ do
1: $k = k + 1$,
2: $\lambda^k = \arg \max_{1 \le i \le L} \|\Phi[i]^T \boldsymbol{R}^{k-1}\|_F$,
3: $\Lambda_k = \Lambda_{k-1} \bigcup \{\lambda^k\}$,
4: $\hat{\boldsymbol{X}}[\Lambda_k] = \arg \min_{\boldsymbol{X}:\mathrm{supp}(\boldsymbol{X})=\Lambda_k} \|\boldsymbol{Y} - \Phi[\Lambda_k]\boldsymbol{X}\|_F$,
5: $\boldsymbol{R}^k = \boldsymbol{Y} - \Phi[\Lambda_k]\hat{\boldsymbol{X}}[\Lambda_k]$.
End
Output: $\hat{\boldsymbol{X}} = \arg \min_{\boldsymbol{X}:\mathrm{supp}(\boldsymbol{X})=\Lambda_K} \|\boldsymbol{Y} - \Phi \boldsymbol{X}\|_2$.

Like other reconstructing algorithms, sufficient conditions of reconstructing block joint sparse matrices with BMMV are very useful. It was shown in [28] that if $\Phi$ obeys $\delta_{K+1} < \frac{1}{\sqrt{K}+1}$, then BMMV perfectly reconstructs block $K$-joint sparse matrices in $K$ iterations. Thus, a natural question is: whether this condition can be further improved? We will answer the question in this paper. Specifically, we will firstly prove that if $\Phi$ obeys the condition $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$, then BMMV perfectly reconstructs block $K$-joint sparse matrices in $K$ iterations. Then, we will also show that BMMV may be failure in reconstructing block $K$-joint sparse matrices in $K$ iterations under the condition $\delta_{K+1} \ge \frac{1}{\sqrt{K+1}}$. Clearly, our sufficient condition is better than that in [28], which is $\delta_{K+1} < \frac{1}{\sqrt{K}+1}$. Moreover, our sufficient condition is sharp. Note that, when $P = 1$ (matrix $\boldsymbol{X}$ turns to vector $\boldsymbol{x}$) and $d = 1$, the BMMV algorithm reduces to orthogonal matching pursuit [9], and the above sufficient condition reduces to the condition in [15].

The rest of this paper is organized as follows. In Section II, we introduce three useful lemmas which are prepared for proving our main results which will be presented in Section III. Finally, we summarize this paper in Section IV. In the following, we introduce some notations.

*Notations:* Let $\|\boldsymbol{x}\|_2$ and $\|\Phi\|_F$ denote the $\ell_2$ norm of the vector $\boldsymbol{x}$ and Frobenius norm of the matrix $\Phi$, respectively. Let $\boldsymbol{x}_i$ be the $i$-th entry of vector $\boldsymbol{x}$. Let $\boldsymbol{I}$ be the identity matrix, and $\boldsymbol{0}$ be zero matrix or zero column vector. Let $\Lambda = \mathrm{supp}(\boldsymbol{X}) = \{i : \boldsymbol{X}^T[i] \neq \boldsymbol{0}\}$, then $|\Lambda| \le K$ for any block $K$-joint sparse matrix $\boldsymbol{X}$, where $\boldsymbol{X}^T[i]$ is the transpose of the $i$-th row of $\boldsymbol{X}$. Let $\Lambda \setminus \Gamma = \{i | i \in \Lambda \text{ and } i \notin \Gamma\}$ for any set $\Gamma \subset \{1, 2, \ldots, L\}$. Let $\Lambda^c = \{1, 2, \ldots, L\} \setminus \Lambda$ and $\Gamma^c = \{1, 2, \ldots, L\} \setminus \Gamma$, where $L$ is the number of blocks of $\boldsymbol{X}$. If $\Phi[\Gamma]$ has full column rank, then the pseudoinverse of $\Phi[\Gamma]$ is $\Phi[\Gamma]^\dagger = (\Phi[\Gamma]^T \Phi[\Gamma])^{-1}\Phi[\Gamma]^T$. Therefore, $\mathcal{P}[\Gamma] = \Phi[\Gamma]\Phi[\Gamma]^\dagger$ and $\mathcal{P}^\perp[\Gamma] = \boldsymbol{I} - \mathcal{P}[\Gamma]$ denote the projector and orthogonal complement projector on the column space of $\Phi[\Gamma]$, respectively.

## II. SOME USEFUL LEMMAS

We recall three lemmas, which are respectively [30, Lemma 1], [30, Lemma 2] and [4, Lemma 1], for proving our main results.

*Lemma 1:* If $\Phi$ satisfies the block RIP of orders $K_1$ and $K_2$ with $K_1 < K_2$, then $\delta_{K_1} \le \delta_{K_2}$.

*Lemma 2:* Let $S_1$ and $S_2$ satisfy $|S_2 \setminus S_1| \ge 1$. Let $\Phi$ obey $|S_1 \cup S_2|$-order block RIP, then for any vector $\boldsymbol{x} \in \mathbb{R}^{|S_2 \setminus S_1|}$,

$$\begin{aligned}(1 - \delta_{|S_1 \cup S_2|})\|\boldsymbol{x}\|_2^2 &\le \|\mathcal{P}^\perp[S_1]\Phi[S_2 \setminus S_1]\boldsymbol{x}\|_2^2 \\ &\le (1 + \delta_{|S_1 \cup S_2|})\|\boldsymbol{x}\|_2^2.\end{aligned}$$

In the following, we introduce [4, Lemma 1], which is useful for proving Lemma 4 in Section III.

*Lemma 3:* Let $\boldsymbol{B} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{D} \in \mathbb{R}^{n \times p}$. Then

$$\|\boldsymbol{BD}\|_F^2 \le \|\boldsymbol{D}\|_F \, \|\boldsymbol{B}^T\boldsymbol{BD}\|_F. \quad (4)$$

For the sake of reading, we recall the proof of [4, Lemma 1] as follows.

*Proof:* Define vectors $\boldsymbol{u}, \boldsymbol{w} \in \mathbb{R}^p$ as

$$u_i = \|\boldsymbol{D}_i\|_2, \quad w_i = \|\boldsymbol{B}^T\boldsymbol{BD}_i\|_2, \ 1 \le i \le p,$$

where $\boldsymbol{D}_i$ is the $i$-th column of $\boldsymbol{D}$.

Then, we have

$$\|\boldsymbol{D}\|_F = \|\boldsymbol{u}\|_2, \quad \|\boldsymbol{B}^T\boldsymbol{BD}\|_F = \|\boldsymbol{w}\|_2.$$

Moreover,

$$\begin{aligned}\|\boldsymbol{BD}\|_F^2 &= \sum_{i=1}^p \|\boldsymbol{BD}_i\|_2^2 = \sum_{i=1}^p \left(\boldsymbol{D}_i \cdot \boldsymbol{B}^T\boldsymbol{BD}_i\right) \\ &\le \sum_{i=1}^p \left(\|\boldsymbol{D}_i\|_2 \cdot \|\boldsymbol{B}^T\boldsymbol{BD}_i\|_2\right) = \boldsymbol{u}^T\boldsymbol{w} \\ &\le \|\boldsymbol{u}\|_2\|\boldsymbol{w}\|_2 = \|\boldsymbol{D}\|_F \, \|\boldsymbol{B}^T\boldsymbol{BD}\|_F,\end{aligned}$$

both of the above inequalities are from the Cauchy-Schwarz inequality. Therefore (4) holds. □

## III. MAIN RESULTS

We will firstly prove that if $\Phi$ obeys $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$, then BMMV perfectly reconstructs any block $K$-joint sparse matrices in $K$ iterations. Then, we will also show that BMMV may be failure in reconstructing block $K$-joint sparse matrices in $K$ iterations under the condition $\delta_{K+1} \geq \frac{1}{\sqrt{K+1}}$.

We start with the following Lemma 4, which provides a lower bound on BMMV decision-metric for the columns belonging to $\Lambda \setminus \Gamma$.

*Lemma 4:* Let $\Gamma \subseteq \Lambda$ with $|\Gamma| < |\Lambda|$ (recall that $\Lambda = \text{supp}(\boldsymbol{X})$), then

$$\max_{i\in\Lambda\setminus\Gamma} \|\Phi^T[i]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F \geq \frac{\|\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F^2}{\sqrt{|\Lambda\setminus\Gamma|}\|\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F}. \quad (5)$$

*Proof:* Since $\Gamma \subseteq \Lambda$ with $|\Gamma| < |\Lambda|$, $\|\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F \neq 0$. Thus,

$$\begin{aligned}
&\max_{i\in\Lambda\setminus\Gamma} \|\Phi^T[i]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F \\
&\geq \frac{1}{\sqrt{|\Lambda\setminus\Gamma|}}\|\Phi^T[\Lambda\setminus\Gamma]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F \\
&\overset{(a)}{=} \frac{1}{\sqrt{|\Lambda\setminus\Gamma|}}\|(\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma])^T\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F \\
&\overset{(b)}{\geq} \frac{1}{\sqrt{|\Lambda\setminus\Gamma|}\|\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F}\|\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F^2,
\end{aligned}$$

where (a) is due to

$$(\mathcal{P}^{\perp}[\Gamma])^T\mathcal{P}^{\perp}[\Gamma] = \mathcal{P}^{\perp}[\Gamma]\mathcal{P}^{\perp}[\Gamma] = \mathcal{P}^{\perp}[\Gamma], \quad (6)$$

and (b) is from (4) with $\boldsymbol{B} = \mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]$ and $\boldsymbol{D} = \boldsymbol{X}[\Lambda\setminus\Gamma]$. Thus, (5) holds. □

The following lemma provides an upper bound on the BMMV decision-metric for the columns belonging to $\Lambda^c$.

*Lemma 5:* Let $\Phi$ in (1) obey the $K+1$ order block RIP with

$$\delta_{K+1} < \frac{1}{\sqrt{K+1}} \quad (7)$$

and $\Gamma \subseteq \Lambda$ with $|\Gamma| < |\Lambda|$. Then

$$\max_{j\in\Lambda^c} \|\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F < \frac{\|\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F^2}{\sqrt{|\Lambda\setminus\Gamma|}\|\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F}.$$

For convenience of reading, we postponed the proof of Lemma 5 to Appendix.

From Lemmas 4 and 5, one can immediately get the following corollary, which shows the robustness of the BMMV algorithm.

*Corollary 1:* Let $\Phi$ in (1) satisfy (7) and $\Gamma \subseteq \Lambda$ with $|\Gamma| < |\Lambda|$, then

$$\max_{i\in\Lambda\setminus\Gamma} \|\Phi^T[i]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F > \max_{j\in\Lambda^c} \|\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F \quad (8)$$

Corollary 1 is one of our main results, and it will play an important role in proving Theorem 1, which is one of our main theorems.

*Remark 1:* If $\Gamma = \emptyset$ in Corollary 1, then (8) becomes to

$$\max_{i\in\Lambda} \|\Phi^T[i]\Phi\boldsymbol{X}\|_F > \max_{j\in\Lambda^c} \|\Phi^T[j]\Phi\boldsymbol{X}\|_F. \quad (9)$$

Moreover, if $d = 1$, then (9) reduces to

$$\max_{i\in\Lambda} \|\Phi_i^T\Phi\boldsymbol{X}\|_2 > \max_{j\in\Lambda^c} \|\Phi_j^T\Phi\boldsymbol{X}\|_2,$$

which is actually the main inequality presented by [4, Lemma 2]. Thus, Corollary 1 is a generalized version of [4, Lemma 2].

Then we give the main result of this paper.

*Theorem 1:* Suppose (7) holds in model (1). The BMMV algorithm can perfectly reconstruct any block $K$-joint sparse matrices in $K$ iterations.

*Proof:* By Algorithm 1, we only need to prove that BMMV selects an index in $\Lambda$ in each iteration. We prove it by mathematical induction. Suppose that BMMV selects an index belonging to $\Lambda$ in each of the first $k-1$ iterations, which means that $\Lambda_{k-1} \subseteq \Lambda$, where $1 \leq k < |\Lambda|$. This assumption obviously holds for $k = 1$ since $\Lambda_0 = \emptyset$. Then, we have to show that BMMV selects an index in $\Lambda$ at the $k$-th iteration, by Algorithm 1, to show that $\lambda^k \in \Lambda$.

By steps 4 and 5 of Algorithm 1, we can see that

$$\|\Phi^T[\Lambda_{k-1}\boldsymbol{R}^{k-1}]\|_F = 0.$$

To show $\lambda^k \in \Lambda$, by step 2 of Algorithm 1, it suffices to prove

$$\max_{i\in\Lambda\setminus\Lambda_{k-1}} \|\Phi^T[i]\boldsymbol{R}^{k-1}]\|_F > \max_{j\in\Lambda^c} \|\Phi^T[j]\boldsymbol{R}^{k-1}]\|_F. \quad (10)$$

By step 4 of Algorithm 1,

$$\hat{\boldsymbol{X}}[\Lambda_{k-1}] = (\Phi^T[\Lambda_{k-1}]\Phi[\Lambda_{k-1}])^{-1}\Phi^T[\Lambda_{k-1}]\boldsymbol{Y}. \quad (11)$$

Further, by step 5 of Algorithm 1 and (11),

$$\begin{aligned}
\boldsymbol{R}^{k-1} &= \boldsymbol{Y} - \Phi[\Lambda_{k-1}]\hat{\boldsymbol{X}}[\Lambda_{k-1}] \\
&= \big(\boldsymbol{I} - \Phi[\Lambda_{k-1}](\Phi^T[\Lambda_{k-1}]\Phi[\Lambda_{k-1}])^{-1}\Phi^T[\Lambda_{k-1}]\big)\boldsymbol{Y} \\
&\overset{(a)}{=} \mathcal{P}^{\perp}[\Lambda_{k-1}]\Phi\boldsymbol{X} \\
&\overset{(b)}{=} \mathcal{P}^{\perp}[\Lambda_{k-1}]\Phi[\Lambda]\boldsymbol{X}[\Lambda] \\
&\overset{(c)}{=} \mathcal{P}^{\perp}[\Lambda_{k-1}] \\
&\quad \times(\Phi[\Lambda_{k-1}]\boldsymbol{X}[\Lambda_{k-1}] + \Phi[\Lambda\setminus\Lambda_{k-1}]\boldsymbol{X}[\Lambda\setminus\Lambda_{k-1}]) \\
&\overset{(d)}{=} \mathcal{P}^{\perp}[\Lambda_{k-1}]\Phi[\Lambda\setminus\Lambda_{k-1}]\boldsymbol{X}[\Lambda\setminus\Lambda_{k-1}],
\end{aligned} \quad (12)$$

where (a) is due to $\boldsymbol{Y} = \Phi\boldsymbol{X}$ and the definition of $\mathcal{P}^{\perp}[\Lambda_{k-1}]$, (b) is because $\Lambda = \text{supp}(\boldsymbol{X})$, (c) is because $\Lambda_{k-1} \subseteq \Lambda$ (induction assumption), (d) is because $\mathcal{P}^{\perp}[\Lambda_{k-1}]\Phi[\Lambda_{k-1}] = \boldsymbol{0}$.

By (12), for $i \in \Lambda \setminus \Lambda_{k-1}$ and $j \in \Lambda^c$, we get

$$\|\Phi^T[i]\boldsymbol{R}^{k-1}]\|_F = \|\Phi^T[i]\mathcal{P}^{\perp}[\Lambda_{k-1}]\Phi[\Lambda \setminus \Lambda_{k-1}]\boldsymbol{X}[\Lambda \setminus \Lambda_{k-1}]\|_F,$$
$$\|\Phi^T[j]\boldsymbol{R}^{k-1}]\|_F = \|\Phi^T[j]\mathcal{P}^{\perp}[\Lambda_{k-1}]\Phi[\Lambda \setminus \Lambda_{k-1}]\boldsymbol{X}[\Lambda \setminus \Lambda_{k-1}]\|_F.$$

Thus, in order to show (10), we need to show

$$\max_{i\in\Lambda\setminus\Lambda_{k-1}} \|\Phi^T[i]\mathcal{P}^{\perp}[\Lambda_{k-1}]\Phi[\Lambda \setminus \Lambda_{k-1}]\boldsymbol{X}[\Lambda \setminus \Lambda_{k-1}]\|_F > \max_{j\in\Lambda^c} \|\Phi^T[j]\mathcal{P}^{\perp}[\Lambda_{k-1}]\Phi[\Lambda \setminus \Lambda_{k-1}]\boldsymbol{X}[\Lambda \setminus \Lambda_{k-1}]\|_F.$$

Applying Corollary 1 with $\Gamma = \Lambda_{k-1}$, one can see that the aforementioned inequality holds, and so does (10). Completing the proof. □

Then, we investigate the necessary condition for perfect reconstructing of block joint sparse matrices with BMMV.

*Theorem 2:* Let $d, K \geq 1$ be arbitrary positive integers. We can construct a block $K$-joint sparse matrix $\boldsymbol{X}$, and a matrix $\Phi$ obeys

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

such that BMMV can not reconstruct $\boldsymbol{X}$ in $K$ iterations.

In order to prove Theorem 2, we firstly introduce Lemma 6, which is obtained from the proof of [30, Theorem 2].

*Lemma 6:* Let $d, K \geq 1$ be arbitrary positive integers and

$$\Psi = \begin{bmatrix} \frac{K}{K+1}\boldsymbol{I}_{dK} & \frac{\boldsymbol{F}_{(dK)\times d}}{K+1} \\ \frac{\boldsymbol{F}^T_{(dK)\times d}}{K+1} & \frac{K+2}{K+1}\boldsymbol{I}_d \end{bmatrix},$$

where

$$\boldsymbol{F}_{(dK)\times d} = (\boldsymbol{I}_d, \ldots, \boldsymbol{I}_d)^T \in \mathbb{R}^{(dK)\times d}. \quad (13)$$

Then, $\Psi$ can be expressed as $\Psi = \Phi^T\Phi$, where $\Phi \in \mathbb{R}^{d(K+1)\times d(K+1)}$ and $\Phi$ obeys the condition $\delta_{K+1} = \frac{1}{\sqrt{K+1}}$.

*Proof of Theorem 2:* Let $d, K \geq 1$ be two arbitrary positive integers. Let $\Phi$ be defined in Lemma 6, and

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{E}_{(dK)\times P} \\ \boldsymbol{0}_{d\times P} \end{bmatrix} \in \mathbb{R}^{d(K+1)\times P},$$

where $\boldsymbol{E}_{(dK)\times P} \in \mathbb{R}^{dK\times P}$ with all entries being 1. Then, we will show that BMMV fails to reconstruct $\boldsymbol{X}$ in $K$ iterations from $\boldsymbol{Y} = \Phi\boldsymbol{X}$ .

Note that $\Phi^T\Phi = \Psi$, we have

$$\|\Phi^T[K+1]\boldsymbol{Y}\|_F = \|(\boldsymbol{0}_{d\times dK}, \boldsymbol{I}_d)\Phi^T\Phi\boldsymbol{X}\|_F = \|(\boldsymbol{0}_{d\times dK}, \boldsymbol{I}_d)\Psi\boldsymbol{X}\|_F = \frac{K\sqrt{dP}}{K+1}.$$

For $1 \leq i \leq K$, it is easy to verify that

$$\begin{aligned}\|\Phi^T[i]\boldsymbol{Y}\|_F &= \|\Phi^T[i]\Phi\boldsymbol{X}\|_F \\ &= \|(\boldsymbol{0}_{d\times(i-1)d}, \boldsymbol{I}_d, \boldsymbol{0}_{d\times(K+1-i)d})\Phi^T\Phi\boldsymbol{X}\|_F \\ &= \|(\boldsymbol{0}_{d\times(i-1)d}, \boldsymbol{I}_d, \boldsymbol{0}_{d\times(K+1-i)d})\Psi\boldsymbol{X}\|_F \\ &= \frac{K\sqrt{dP}}{K+1}.\end{aligned}$$

That is to say,

$$\max_{i\in\Lambda}\|\Phi^T[i]\boldsymbol{Y}\|_F = \max_{j\in\Lambda^c}\|\Phi^T[j]\boldsymbol{Y}\|_F.$$

Therefore, the BMMV algorithm fails in the first iteration, i.e., BMMV can not exactly reconstruct the matrix $\boldsymbol{X}$ in $K$ iterations. □

## IV. CONCLUSION

In this article, we studied block RIP based sufficient conditions of perfect reconstructing block joint sparse matrices with BMMV. If $\Phi$ satisfies the condition $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$, then we have shown that BMMV can perfectly reconstruct all block $K$-joint sparse matrices in $K$ iterations. Besides, we have also shown that BMMV may fail to reconstruct block $K$-joint sparse matrices in $K$ iterations under the condition $\delta_{K+1} = \frac{1}{\sqrt{K+1}}$. This sufficient condition is sharp in the sense that BMMV may not reconstruct a block $K$-joint sparse matrix in $K$ iterations under the condition $\delta_{K+1} \geq \frac{1}{\sqrt{K+1}}$.

## APPENDIX

## PROOF OF LEMMA 5

In the following, we follow the proof of [10, Lemma 1] and [15, Lemma II.2] to prove Lemma 5.

*Proof:* To prove the lemma, we show that for each given $j \in \Lambda^c$,

$$\|\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F < \frac{\|\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F^2}{\sqrt{|\Lambda\setminus\Gamma|}\|\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F}. \quad (14)$$

Let

$$\nu = -\frac{\sqrt{|\Lambda\setminus\Gamma|+1}-1}{\sqrt{|\Lambda\setminus\Gamma|}}, \quad (15)$$

then it is easy to obtain

$$\frac{2\nu}{1-\nu^2} = -\sqrt{|\Lambda\setminus\Gamma|}, \quad \frac{1+\nu^2}{1-\nu^2} = \sqrt{|\Lambda\setminus\Gamma|+1}. \quad (16)$$

To simplify notation, we introduce a new matrix $\boldsymbol{Z} \in \mathbb{R}^{d\times P}$, and the $p$-th column of $\boldsymbol{Z}$ is defined as

$$\boldsymbol{Z}_p = \frac{\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}_p[\Lambda\setminus\Gamma]}{\|\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F}, \quad 1 \leq p \leq P, \quad (17)$$

where $\boldsymbol{X}_p[\Lambda\setminus\Gamma]$ is the $p$-th column of $\boldsymbol{X}[\Lambda\setminus\Gamma] \in \mathbb{R}^{|\Lambda\setminus\Gamma|\times P}$. Then, $\|\boldsymbol{Z}\|_F = 1$ and

$$\sum_{p=1}^{P}\boldsymbol{Z}_p^T\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{Z}_p[\Lambda\setminus\Gamma] = \|\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F. \quad (18)$$

Furthermore, we define

$$\boldsymbol{U} = \begin{bmatrix}\boldsymbol{X}[\Lambda\setminus\Gamma] \\ \boldsymbol{0}_{d\times P}\end{bmatrix},$$
$$\boldsymbol{W} = \nu\|\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F\begin{bmatrix}\boldsymbol{0}_{|\Lambda\setminus\Gamma|d\times P} \\ \boldsymbol{Z}\end{bmatrix}.$$

For $j \in \Lambda^c$, denote

$$\boldsymbol{C} = \mathcal{P}^{\perp}[\Gamma]\left[\Phi[\Lambda \setminus \Gamma] \ \Phi[j]\right]. \tag{19}$$

Then

$$\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda \setminus \Gamma]\boldsymbol{X}[\Lambda \setminus \Gamma] = \boldsymbol{CU} \tag{20}$$

and

$$\|\boldsymbol{U} + \boldsymbol{W}\|_F^2 = (1+\nu^2)\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F^2, \tag{21}$$

$$\|\nu^2\boldsymbol{U} - \boldsymbol{W}\|_F^2 = \nu^2(1+\nu^2)\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F^2. \tag{22}$$

Moreover,

$$\begin{aligned}
&\sum_{p=1}^{P} \boldsymbol{W}_p^T\boldsymbol{C}^T\boldsymbol{C}\boldsymbol{U}_p \\
&\overset{(a)}{=} \nu\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F \sum_{p=1}^{P} \boldsymbol{Z}_p^T\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\mathcal{P}^{\perp}[\Gamma] \\
&\quad \times \Phi[\Lambda \setminus \Gamma]\boldsymbol{X}_p[\Lambda \setminus \Gamma] \\
&\overset{(b)}{=} \nu\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F \sum_{p=1}^{P} \boldsymbol{Z}_p^T\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda \setminus \Gamma]\boldsymbol{X}_p[\Lambda \setminus \Gamma] \\
&\overset{(c)}{=} \nu\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F \|\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda \setminus \Gamma]\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F,
\end{aligned} \tag{23}$$

where (a) follows from (19) and (20); (b) is due to (6) and (c) is from (18).

Applying (23) yields

$$\begin{aligned}
&\|\boldsymbol{C}(\boldsymbol{U}+\boldsymbol{W})\|_F^2 \\
&= \sum_{p=1}^{P} \|\boldsymbol{C}(\boldsymbol{U}_p + \boldsymbol{W}_p)\|_2^2 \\
&= \sum_{p=1}^{P} (\|\boldsymbol{C}\boldsymbol{U}_p\|_2^2 + \|\boldsymbol{C}\boldsymbol{W}_p\|_2^2 + \boldsymbol{W}_p^T\boldsymbol{C}^T\boldsymbol{C}\boldsymbol{U}_p) \\
&= \|\boldsymbol{CU}\|_F^2 + \|\boldsymbol{CW}\|_F^2 + 2\sum_{p=1}^{P} \boldsymbol{W}_p^T\boldsymbol{C}^T\boldsymbol{C}\boldsymbol{U}_p \\
&= \|\boldsymbol{CU}\|_F^2 + \|\boldsymbol{CW}\|_F^2 \\
&\quad + 2\nu\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F \|\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda \setminus \Gamma]\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F.
\end{aligned}$$

Similarly, by (23), we obtain

$$\begin{aligned}
&\|\boldsymbol{C}(\nu^2\boldsymbol{U}-\boldsymbol{W})\|_F^2 \\
&= \nu^4\|\boldsymbol{CU}\|_F^2 + \|\boldsymbol{CW}\|_F^2 \\
&\quad - 2\nu^3\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F \|\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda \setminus \Gamma]\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F.
\end{aligned}$$

Applying the aforementioned equations yields

$$\begin{aligned}
&\|\boldsymbol{C}(\boldsymbol{U}+\boldsymbol{W})\|_F^2 - \|\boldsymbol{C}(\nu^2\boldsymbol{U}-\boldsymbol{W})\|_F^2 \\
&= (1-\nu^4)\|\boldsymbol{CU}\|_F^2 + 2\nu(1+\nu^2) \\
&\quad \times \|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F \|\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda \setminus \Gamma]\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F \\
&= (1-\nu^4)(\|\boldsymbol{CU}\|_F^2 + \frac{2\nu}{1-\nu^2} \\
&\quad \times \|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F \|\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda \setminus \Gamma]\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F) \\
&= (1-\nu^4)(\|\boldsymbol{CU}\|_F^2 - \sqrt{|\Lambda \setminus \Gamma|} \\
&\quad \times \|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F \|\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda \setminus \Gamma]\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F),
\end{aligned} \tag{24}$$

where the last equality is because of the first equality in (16).

By (15), one can check that $1 - \nu^4 > 0$. Thus, if

$$\|\boldsymbol{C}(\boldsymbol{U}+\boldsymbol{W})\|_F^2 > \|\boldsymbol{C}(\nu^2\boldsymbol{U}-\boldsymbol{W})\|_F^2, \tag{25}$$

then by (24), we have

$$\frac{\|\boldsymbol{CU}\|_2^2}{\sqrt{|\Lambda\setminus\Gamma|}\|\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F} > \|\Phi^T[j]\mathcal{P}^{\perp}[\Gamma]\Phi[\Lambda\setminus\Gamma]\boldsymbol{X}[\Lambda\setminus\Gamma]\|_F.$$

By combing with (20), one can see that (14) holds and which implies the lemma holds. Therefore, what remains to show is (25). One can check that

$$\begin{aligned}
&\|\boldsymbol{C}(\boldsymbol{U}+\boldsymbol{W})\|_F^2 - \|\boldsymbol{C}(\nu^2\boldsymbol{U}-\boldsymbol{W})\|_F^2 \\
&= \sum_{p=1}^{P}(\|\boldsymbol{C}(\boldsymbol{U}_p+\boldsymbol{W}_p)\|_2^2 - \|\boldsymbol{C}(\nu^2\boldsymbol{U}_p-\boldsymbol{W}_p)\|_2^2) \\
&\overset{(a)}{\geq} (1-\delta_{|\Lambda|+1})\sum_{p=1}^{P}\|(\boldsymbol{U}_p+\boldsymbol{W}_p)\|_2^2 \\
&\quad - (1+\delta_{|\Lambda|+1})\sum_{p=1}^{P}\|(\nu^2\boldsymbol{U}_p-\boldsymbol{W}_p)\|_2^2 \\
&= (1-\delta_{|\Lambda|+1})\|(\boldsymbol{U}+\boldsymbol{W})\|_F^2 - (1+\delta_{|\Lambda|+1})\|(\nu^2\boldsymbol{U}-\boldsymbol{W})\|_F^2 \\
&\overset{(b)}{=} (1-\delta_{|\Lambda|+1})(1+\nu^2)\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F^2 \\
&\quad - (1+\delta_{|\Lambda|+1})\nu^2(1+\nu^2)\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F^2 \\
&= (1+\nu^2)\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F^2\big((1-\delta_{|\Lambda|+1}) - (1+\delta_{|\Lambda|+1})\nu^2\big) \\
&= (1+\nu^2)\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F^2\big((1-\nu^2) - \delta_{|\Lambda|+1}(1+\nu^2)\big) \\
&= (1-\nu^4)\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F^2\big(1 - \frac{1+\nu^2}{1-\nu^2}\delta_{|\Lambda|+1}\big) \\
&\overset{(c)}{=} (1-\nu^4)\|\boldsymbol{X}[\Lambda \setminus \Gamma]\|_F^2\big(1 - \sqrt{|\Lambda \setminus \Gamma| + 1}\delta_{|\Lambda|+1}\big) \overset{(d)}{>} 0.
\end{aligned}$$

where (a) follows from Lemma 2 and (19), (b) is due to (21) and (22), (c) follows from the second equality in (16), and (d) is from (7) and Lemma 1. □

## REFERENCES

[1] G. Obozinski, B. Taskar, and M. I. Jordan, "Joint covariate selection and joint subspace selection for multiple classification problems," *Statist. Comput.*, vol. 20, no. 2, pp. 231–252, Apr. 2010.

[2] A. Majumdar and R. K. Ward, "Face recognition from video: An MMV recovery approach," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 2221–2224.

[3] X. Wei, Y. Yuan, and Q. Ling, "DOA estimation using a greedy block coordinate descent algorithm," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6382–6394, Dec. 2012.

[4] J. Wen, J. Tang, and F. Zhu, "Greedy block coordinate descent under restricted isometry property," *Mobile Netw. Appl.*, vol. 22, no. 3, pp. 371–376, 2017.

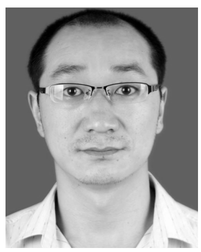
[5] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[6] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[7] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best $k$-term approximation," *J. Amer. Math. Soc.*, vol. 22, no. 1, pp. 211–231, 2009.
[8] J. Wen, D. Li, and F. Zhu, "Stable recovery of sparse signals via $l_p$-minimization," *Appl. Comput. Harmon. Anal.*, vol. 38, no. 1, pp. 161–176, 2015.
[9] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[10] J. Wen, Z. Zhou, J. Wang, X. Tang, and Q. Mo, "A sharp condition for exact support recovery with orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 65, no. 6, pp. 1370–1382, Mar. 2017.
[11] J. Wang, S. Kwon, and B. Shim, "Generalized orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6202–6216, Dec. 2012.
[12] J. Wen, Z. Zhou, and H. Chen, "An optimal condition for the block orthogonal matching pursuit algorithm," *IEEE Access*, vol. 6, pp. 38179–38185, 2018.
[13] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.
[14] J. Wang and B. Shim, "On the recovery limit of sparse signals using orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4973–4976, Sep. 2012.
[15] Q. Mo. (2015). "A sharp restricted isometry constant bound of orthogonal matching pursuit." [Online]. Available: https://arxiv.org/abs/1501.01708
[16] B. Li, Y. Shen, S. Rajan, and T. Kirubarajan, "Theoretical results for sparse signal recovery with noises using generalized OMP algorithm," *Signal Process.*, vol. 117, pp. 270–278, Dec. 2015.
[17] B. Li, Y. Shen, Z. Wu, and J. Li, "Sufficient conditions for generalized orthogonal matching pursuit in noisy case," *Signal Process.*, vol. 108, pp. 111–123, Mar. 2015.
[18] J. Zhao, R. Song, J. Zhao, and W.-P. Zhu, "New conditions for uniformly recovering sparse signals via orthogonal matching pursuit," *Signal Process.*, vol. 106, pp. 106–113, Jan. 2015.
[19] J. Wen, Z. Zhou, D. Li, and X. Tang, "A novel sufficient condition for generalized orthogonal matching pursuit," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 805–808, Apr. 2017.
[20] J. Wen, J. Wang, and Q. Zhang, "Nearly optimal bounds for orthogonal least squares," *IEEE Trans. Signal Process.*, vol. 65, no. 20, pp. 5347–5356, Oct. 2017.
[21] X. Cai, Z. Zhou, Y. Yang, and Y. Wang, "Improved sufficient conditions for support recovery of sparse signals via orthogonal matching pursuit," *IEEE Access*, vol. 6, pp. 30437–30443, 2018.
[22] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.
[23] E. van den Berg and M. P. Friedlander, "Theoretical and empirical results for recovery from multiple measurements," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2516–2527, May 2010.
[24] M.-J. Lai and Y. Liu, "The null space property for sparse recovery from multiple measurement vectors," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 3, pp. 402–406, 2011.
[25] J. A. Zhang, Z. Chen, P. Cheng, and X. Huang, "Multiple-measurement vector based implementation for single-measurement vector sparse Bayesian learning with reduced complexity," *Signal Process.*, vol. 118, pp. 153–158, Jan. 2016.
[26] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.
[27] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
[28] Y. Fu, H. Li, Q. Zhang, and J. Zou, "Block-sparse recovery via redundant block OMP," *Signal Process.*, vol. 97, pp. 162–171, Apr. 2014.
[29] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.
[30] J. Wen, Z. Zhou, Z. Liu, M.-J. Lai, and X. Tang, "Sharp sufficient conditions for stable recovery of block sparse signals by block orthogonal matching pursuit," *Appl. Comput. Harmon. Anal.*, to be published, doi: 10.1016/j.acha.2018.02.002.

**YANLI SHI** received the M.S. degree in basic mathematics from Beihua University, Jilin, China, in 2005, and the Ph.D. degree in basic mathematics from the Dalian University of Technology, Dalian, China, in 2017. Since 2005, she has been with the Jilin Institute of Chemical Technology. In 2014, she became an Associate Professor with the School of Science. Her research interests include fuzzy mathematics and data mining.

**LIBO WANG** received the B.S. degree in mathematics from Xiangtan University, Hunan, China, in 2012, and the Ph.D. degree in applied mathematics from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, in 2017. He is currently a Lecturer of the College of Information Science and Technology, Jinan University. His research interests include cryptography functions and sparse recovery.



**RONG LUO** received the Ph.D. degree in representation of algebra from Nanjing University, Nanjing, China, in 2008. From 2014 to 2015, he was a Visiting Scholar with the Department of Computer Science, The University of Melbourne, Australia. He is currently an Associate Professor with the School of Mathematics, Southwest Jiaotong University.