

RESEARCH

Open Access

# SBL-based multi-task algorithms for recovering block-sparse signals with unknown partitions

Ying-Gui Wang<sup>1\*</sup>, Le Yang<sup>1,2</sup>, Zheng Liu<sup>1</sup> and Wen-Li Jiang<sup>1</sup>

## Abstract

We consider in this paper the problem of reconstructing block-sparse signals with unknown block partitions. In the first part of this work, we extend the block-sparse Bayesian learning (BSBL) originally developed for recovering a single block-sparse signal in a single compressive sensing (CS) task scenario to the case of multiple CS tasks. A new multi-task signal recovery algorithm, called the extended multi-task block-sparse Bayesian learning (EMBSBL), is proposed. EMBSBL exploits the statistical correlation among multiple signals as well as the intra-block correlation within individual signals to improve performance. Besides, it does not need *a priori* information on block partition. As the second part of this paper, we develop the EMBSBL-based synthesized multi-task signal recovery algorithm, namely SEMBSBL, to make it applicable to the single CS task case. The idea is to synthesize new CS tasks from the single CS task via circular-shifting operations and utilizes the minimum description length principle to determine the proper set of the synthesized CS tasks for signal reconstruction. SEMBSBL can achieve better signal reconstruction performance over other algorithms that recover block-sparse signals individually. Simulations corroborate the theoretical developments.

**Keywords:** Sparse Bayesian learning; Block-sparse; Multi-task; Circular shifting; Minimum description length

## 1 Introduction

Compressive sensing (CS) enables reconstructing a signal that is sparse in a certain domain from its measurements obtained at a rate significantly lower than the Nyquist frequency [1]. If in addition to sparsity, the signal representation is also structured in the form of clustered non-zeros, the signal would be referred to as being block-sparse. In practice, block-sparsity can be found in multi-band signals [2] or in the measurements of gene expression levels [3]. It has been shown that exploring the block-sparsity enables robust signal recovery from fewer compressive measurements [4]. We shall consider in this paper the efficient recovery of block-sparse signals.

Several block-sparse signal reconstruction algorithms have been developed in literature. Based on the compressive sampling matching pursuit (CoSaMP) [5], the block compressive sampling matching pursuit (BCoSaMP) was proposed in [4]. It utilizes the knowledge on the number of non-zero blocks to achieve signal recovery. On the basis of the orthogonal matching pursuit (OMP) [6], the block

orthogonal matching pursuit (BOMP) was developed in [7]. Zou et al. proposed a block fixed-point continuation algorithm in [8] for block-sparse signal recovery. Elhamifar and Vidal approached the problem via the application of convex relaxation and convex optimization [9]. The two methods developed in [8] and [9] require the availability of the information on the block size. In [10], the dictionary optimization for block-sparse signal representation was studied and the work assumed that the maximum block length was known. More recently, CluSS-MCMC [11] and BM-MAP-OMP [12] have been proposed, which require little *a priori* knowledge on the block partition. On the basis of Bayesian sparse learning for temporally correlated signals [13,14] proposed two block-sparse signal recovery algorithms, the block-sparse Bayesian learning (BSBL) and its extended version, the EBSBL algorithm. The BSBL algorithm utilizes the block partition information while EBSBL handles signals with unknown block partitions. Most techniques reviewed above fall under the category of the single-task CS, where the focus is on recovering a block-sparse signal from its compressive measurements.

The contribution of this paper is twofold. We shall first consider the block-sparse signal reconstruction in a

\*Correspondence: wyinggui@gmail.com

<sup>1</sup> College of Electronic Science and Engineering, National University of Defense Technology, Deya Road, Changsha 410073, People's Republic of China  
Full list of author information is available at the end of the article

multi-task scenario, where the signals in different CS tasks are statistically correlated. The multi-task compressive sensing (MCS) was originally developed in [15]. Mathematically, we have  $L$  CS tasks

$$\mathbf{y}_i = \Phi_i \boldsymbol{\theta}_i + \mathbf{n}_i, \quad i = 1, \dots, L, \quad (1)$$

where  $\mathbf{y}_i$  is the compressive measurement vector of the  $i$ th task and  $\Phi_i$  is the  $M_i \times N$  measurement matrix ( $M_i \ll N$ ).  $\boldsymbol{\theta}_i$  is the original signal in the  $i$ th task to be recovered and  $\mathbf{n}_i$  represents the measurement noise. In MCS, the correlation among  $\boldsymbol{\theta}_i$  is explored so that  $\boldsymbol{\theta}_i$  are reconstructed jointly. MCS outperforms the single-task CS algorithm in terms of the reduced number of compressive measurements needed for efficient signal recovery. However, existing MCS techniques do not take into account the structural information in signals, such as block-sparsity. We shall therefore propose in this paper an extended version of the EBSBL algorithm from [14]. The original EBSBL method does not assume the knowledge on the block partition information, and it exhibits better block-sparse signal reconstruction performance over methods such as CluSS-MCMC and BM-MAP-OMP. We shall generalize EBSBL to the MCS scenario and obtain a new technique, referred to as extended multi-task block-sparse Bayesian learning (EMBSBL). Besides using the statistical correlation among  $\boldsymbol{\theta}_i$  as in MCS, EMBSBL also utilizes the intra-block correlation within each signal to improve performance. Simulations show that the block-sparse signal recovery performance of EMBSBL is superior to that of the benchmark algorithms.

When there is only one CS task, the proposed EMBSBL algorithm would become inapplicable. To address this problem, in the second part of this work, we shall augment EMBSBL with the concept of the synthesized multi-task-based signal recovery. The new algorithm is referred to as SEMBSBL in the rest of the paper. SEMBSBL first synthesizes multiple CS tasks from the single-task CS and then applies EMBSBL to recover the block-sparse signal. The multiple CS tasks are produced via simply circular-shifting the columns of the measurement matrix of the original CS model, which corresponds to circular-shifting the elements in the original signal vector and creates signals that have overlapping clusters, or equivalently speaking, correlated signals. The number of synthesized tasks is determined by the minimum description length (MDL) principle. With increase in the computational complexity, the newly proposed SEMBSBL technique outperforms the previously developed block-sparse signal recovery methods in terms of significantly reduced reconstruction errors and the removal of the needs for detailed information on the sparsity structure. Computer simulations are provided to demonstrate the good performance of the proposed SEMBSBL method.

The remainder of this paper is organized as follows. Section 2 presents the new EMBSBL algorithm for recovering multiple correlated block-sparse signals jointly. Section 3 illustrates the idea of synthesizing multiple CS tasks from a single one and presents the proposed SEMBSBL algorithm. Simulation results are given in Section 4 and Section 5 concludes the paper.

## 2 EMBSBL algorithm

The development of EMBSBL starts with extending BSBL-BO in [14] to the case of multiple CS tasks. The resulting algorithm, called MBSBL, can jointly recover block-sparse signals when their non-zero blocks all have the same size. We next generalize MBSBL to obtain EMBSBL that does not need the information on the signal sparsity structure.

### 2.1 MBSBL

Let  $S$  be the block size and  $K$  be the number of blocks in every signal to be recovered. If the measurement noise  $\mathbf{n}_i$  in (1) follows an i.i.d. Gaussian distribution with zero mean and covariance matrix  $\beta^{-1}\mathbf{I}$ , the conditional likelihood function of  $\mathbf{y}_i$  is

$$p(\mathbf{y}_i | \boldsymbol{\theta}_i, \beta) = \mathcal{N}(\mathbf{y}_i | \Phi_i \boldsymbol{\theta}_i, \beta^{-1}\mathbf{I}), \quad (2)$$

where  $\mathcal{N}(\mathbf{y}_i | \Phi_i \boldsymbol{\theta}_i, \beta^{-1}\mathbf{I})$  represents a Gaussian distribution with mean  $\Phi_i \boldsymbol{\theta}_i$  and covariance matrix  $\beta^{-1}\mathbf{I}$ . In BSBL, each block  $\boldsymbol{\theta}_{i,j} \in \mathcal{R}^{S \times 1}$  is assumed to satisfy a zero-mean multivariate Gaussian distribution

$$p(\boldsymbol{\theta}_{i,j} | \gamma_j, \mathbf{B}_j) = \mathcal{N}(\boldsymbol{\theta}_{i,j} | \mathbf{0}, \gamma_j \mathbf{B}_j), \quad j = 1, \dots, K. \quad (3)$$

If we further assume that blocks are mutually uncorrelated, the prior for  $\boldsymbol{\theta}_i$  is given by  $p(\boldsymbol{\theta}_i | \boldsymbol{\gamma}, \mathbf{B}_0) = \mathcal{N}(\boldsymbol{\theta}_i | \mathbf{0}, \boldsymbol{\Sigma}_0)$ , where  $\boldsymbol{\gamma} = \{\gamma_j\}_{j=1, \dots, K}$ ,  $\mathbf{B}_0 = \{\mathbf{B}_j\}_{j=1, \dots, K}$ ,

$$\boldsymbol{\Sigma}_0 = \begin{bmatrix} \gamma_1 \mathbf{B}_1 & & & \\ & \ddots & & \\ & & \gamma_K \mathbf{B}_K & \\ & & & \end{bmatrix}. \quad (4)$$

Here,  $\mathbf{B}_j$  is a positive definite matrix, capturing the correlation structure within the  $j$ th block, and  $\gamma_j$  is a nonnegative parameter controlling the block-sparsity of  $\boldsymbol{\theta}_{i,j}$ . When  $\gamma_j = 0$ , the  $j$ th block becomes zero. During the learning process, most  $\gamma_j$  tend to be zero, due to the mechanism of automatic relevance determination [13].

To avoid overfitting, we set  $\mathbf{B}_j = \mathbf{B}$ ,  $j = 1, \dots, K$ . Thus,  $\boldsymbol{\Sigma}_0 = \Gamma \otimes \mathbf{B}$ , where  $\Gamma \triangleq \text{diag}(\gamma_1, \dots, \gamma_K)$  and  $\otimes$  denotes the Kronecker product. The posterior distribution of  $\boldsymbol{\theta}_i$  is then given by

$$\begin{aligned} p(\boldsymbol{\theta}_i | \mathbf{y}_i, \beta, \boldsymbol{\gamma}, \mathbf{B}) &= \frac{p(\mathbf{y}_i | \boldsymbol{\theta}_i, \beta) p(\boldsymbol{\theta}_i | \boldsymbol{\gamma}, \mathbf{B})}{\int p(\mathbf{y}_i | \boldsymbol{\theta}_i, \beta) p(\boldsymbol{\theta}_i | \boldsymbol{\gamma}, \mathbf{B}) d\boldsymbol{\theta}_i} \quad (5) \\ &= \mathcal{N}(\boldsymbol{\theta}_i | \boldsymbol{\mu}_{\boldsymbol{\theta}_i}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}_i}) \end{aligned}$$

where

$$\boldsymbol{\mu}_{\theta_i} = \boldsymbol{\Sigma}_0 \boldsymbol{\Phi}_i^T \left( \beta^{-1} \mathbf{I} + \boldsymbol{\Phi}_i \boldsymbol{\Sigma}_0 \boldsymbol{\Phi}_i^T \right)^{-1} \mathbf{y}_i \quad (6)$$

$$\boldsymbol{\Sigma}_{\theta_i} = \left( \boldsymbol{\Sigma}_0^{-1} + \beta \boldsymbol{\Phi}_i^T \boldsymbol{\Phi}_i \right)^{-1}. \quad (7)$$

From (5), we note that  $\beta$ ,  $\boldsymbol{\gamma}$ , and  $\mathbf{B}$  are the sharing parameters of all CS tasks. To estimate them, let  $\mathbf{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_L\}$  be the measurement set of the  $L$  CS tasks. We have

$$p(\mathbf{Y} | \beta, \boldsymbol{\gamma}, \mathbf{B}) = \prod_{i=1}^L p(\mathbf{y}_i | \beta, \boldsymbol{\gamma}, \mathbf{B}). \quad (8)$$

The logarithm of  $p(\mathbf{Y} | \beta, \boldsymbol{\gamma}, \mathbf{B})$  is

$$\begin{aligned} L(\beta, \boldsymbol{\gamma}, \mathbf{B}) &\triangleq \sum_{i=1}^L \log p(\mathbf{y}_i | \beta, \boldsymbol{\gamma}, \mathbf{B}) \\ &= \sum_{i=1}^L \log \int p(\mathbf{y}_i | \boldsymbol{\theta}_i, \beta) p(\boldsymbol{\theta}_i | \boldsymbol{\gamma}, \mathbf{B}) d\boldsymbol{\theta}_i \\ &= -\frac{1}{2} \sum_{i=1}^L \left[ M_i \log(2\pi) + \log(\det(\mathbf{C}_i)) + \mathbf{y}_i^T \mathbf{C}_i^{-1} \mathbf{y}_i \right], \end{aligned} \quad (9)$$

where  $\mathbf{C}_i = \beta^{-1} \mathbf{I} + \boldsymbol{\Phi}_i \boldsymbol{\Sigma}_0 \boldsymbol{\Phi}_i^T$ . Maximizing  $L(\beta, \boldsymbol{\gamma}, \mathbf{B})$  would yield the estimates of the sharing parameters  $\beta$ ,  $\boldsymbol{\gamma}$  and  $\mathbf{B}$ . We shall adopt the approach used in [14] to identify  $\boldsymbol{\gamma}$  via the bound-optimization method, and find  $\beta$  and  $\mathbf{B}$  via expectation maximization (EM).

### 2.1.1 Estimating $\boldsymbol{\gamma}$

Maximizing (9) is equivalent to the minimization of  $\sum_{i=1}^L \left[ \log(\det(\mathbf{C}_i)) + \mathbf{y}_i^T \mathbf{C}_i^{-1} \mathbf{y}_i \right]$ . For this purpose, we replace the term  $\log(\det(\mathbf{C}_i))$  with an upper bound and apply a surrogate function for the term  $\mathbf{y}_i^T \mathbf{C}_i^{-1} \mathbf{y}_i$  and then minimize their summation.

The upper bound of  $\log(\det(\mathbf{C}_i))$  depends on its supporting hyperplane. Let  $\boldsymbol{\gamma}^*$  be a given point in the  $\boldsymbol{\gamma}$ -space and we have

$$\begin{aligned} &\log(\det(\mathbf{C}_i)) \\ &= \log\left(\det\left(\beta^{-1} \mathbf{I} + \boldsymbol{\Phi}_i \boldsymbol{\Sigma}_0 \boldsymbol{\Phi}_i^T\right)\right) \\ &\leq \log\left(\det\left(\beta^{-1} \mathbf{I} + \boldsymbol{\Phi}_i \boldsymbol{\Sigma}_0^* \boldsymbol{\Phi}_i^T\right)\right) \\ &\quad + \sum_{j=1}^K \text{Tr}\left(\left(\boldsymbol{\Sigma}_{\mathbf{y}_i}^*\right)^{-1} \boldsymbol{\Phi}_i^j \mathbf{B} \left(\boldsymbol{\Phi}_i^j\right)^T\right) (\gamma_j - \gamma_j^*) \quad (10) \\ &= \log\left(\det\left(\boldsymbol{\Sigma}_{\mathbf{y}_i}^*\right)\right) \\ &\quad + \sum_{j=1}^K \text{Tr}\left(\left(\boldsymbol{\Sigma}_{\mathbf{y}_i}^*\right)^{-1} \boldsymbol{\Phi}_i^j \mathbf{B} \left(\boldsymbol{\Phi}_i^j\right)^T\right) (\gamma_j - \gamma_j^*), \end{aligned}$$

where  $\boldsymbol{\Sigma}_{\mathbf{y}_i}^* = \beta^{-1} \mathbf{I} + \boldsymbol{\Phi}_i \boldsymbol{\Sigma}_0^* \boldsymbol{\Phi}_i^T$  and  $\boldsymbol{\Sigma}_0^* \triangleq \boldsymbol{\Sigma}_0 |_{\boldsymbol{\gamma}=\boldsymbol{\gamma}^*}$ .  $\boldsymbol{\Phi}_i^j \in \mathcal{R}^{M_i \times S}$  is a submatrix of  $\boldsymbol{\Phi}_i = [\boldsymbol{\Phi}_i^1, \boldsymbol{\Phi}_i^2, \dots, \boldsymbol{\Phi}_i^K]$ , which corresponds to the  $j$ th block of  $\boldsymbol{\theta}_i$ . We next introduce the surrogate function for the term  $\mathbf{y}_i^T \mathbf{C}_i^{-1} \mathbf{y}_i$ . The purpose is to facilitate evaluating the partial derivatives of  $\mathbf{y}_i^T \mathbf{C}_i^{-1} \mathbf{y}_i$  with respect to the sharing parameters  $\beta$  and  $\boldsymbol{\gamma}$ . Originally, the sharing parameters appear in the inverse of the matrix  $\mathbf{C}_i$  (see the definition of  $\mathbf{C}_i$  under (9)). The surrogate function for  $\mathbf{y}_i^T \mathbf{C}_i^{-1} \mathbf{y}_i$  is

$$\begin{aligned} \mathbf{y}_i^T \mathbf{C}_i^{-1} \mathbf{y}_i &= \mathbf{y}_i^T \left( \beta^{-1} \mathbf{I} + \boldsymbol{\Phi}_i \boldsymbol{\Sigma}_0 \boldsymbol{\Phi}_i^T \right)^{-1} \\ \mathbf{y}_i &= \min_{\boldsymbol{\theta}_i} \left( \frac{1}{\beta} \|\mathbf{y}_i - \boldsymbol{\Phi}_i \boldsymbol{\theta}_i\|_2^2 + \boldsymbol{\theta}_i^T \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\theta}_i \right). \end{aligned} \quad (11)$$

It can be easily verified that the cost function to be minimized on the righthand of (11) is the logarithm of the numerator of (5),  $p(\mathbf{y}_i | \boldsymbol{\theta}_i, \beta) p(\boldsymbol{\theta}_i | \boldsymbol{\gamma}, \mathbf{B})$ , and the solution to the minimization problem is  $\boldsymbol{\mu}_{\boldsymbol{\theta}_i}$  defined in (6).

Putting (10) and (11) into (9), we have

$$\begin{aligned} &\sum_{i=1}^L \left[ \log(\det(\mathbf{C}_i)) + \mathbf{y}_i^T \mathbf{C}_i^{-1} \mathbf{y}_i \right] \\ &\leq \sum_{i=1}^L \left[ \log\left(\det\left(\boldsymbol{\Sigma}_{\mathbf{y}_i}^*\right)\right) \right. \\ &\quad \left. + \sum_{j=1}^K \text{Tr}\left(\left(\boldsymbol{\Sigma}_{\mathbf{y}_i}^*\right)^{-1} \boldsymbol{\Phi}_i^j \mathbf{B} \left(\boldsymbol{\Phi}_i^j\right)^T\right) (\gamma_j - \gamma_j^*) \right] \\ &\quad + \sum_{i=1}^L \min_{\boldsymbol{\theta}_i} \left( \frac{1}{\beta} \|\mathbf{y}_i - \boldsymbol{\Phi}_i \boldsymbol{\theta}_i\|_2^2 + \boldsymbol{\theta}_i^T \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\theta}_i \right). \end{aligned} \quad (12)$$

Let  $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_L\}$  be the set of original signals from the  $L$  CS tasks. We can express the upper bound of  $\sum_{i=1}^L \left[ \log(\det(\mathbf{C}_i)) + \mathbf{y}_i^T \mathbf{C}_i^{-1} \mathbf{y}_i \right]$  as

$$\begin{aligned} G(\boldsymbol{\gamma}, \boldsymbol{\Theta}) &\triangleq \sum_{i=1}^L \left[ \log\left(\det\left(\boldsymbol{\Sigma}_{\mathbf{y}_i}^*\right)\right) \right. \\ &\quad \left. + \sum_{j=1}^K \text{Tr}\left(\left(\boldsymbol{\Sigma}_{\mathbf{y}_i}^*\right)^{-1} \boldsymbol{\Phi}_i^j \mathbf{B} \left(\boldsymbol{\Phi}_i^j\right)^T\right) (\gamma_j - \gamma_j^*) \right] \\ &\quad + \sum_{i=1}^L \left[ \frac{1}{\beta} \|\mathbf{y}_i - \boldsymbol{\Phi}_i \boldsymbol{\theta}_i\|_2^2 + \boldsymbol{\theta}_i^T \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\theta}_i \right]. \end{aligned} \quad (13)$$

Taking the partial derivative of  $G(\boldsymbol{\gamma}, \boldsymbol{\Theta})$  with respect to  $\gamma_j$  and setting the result to zero yield the desired estimate of  $\gamma_j$ , the  $j$ th element in  $\boldsymbol{\gamma}$ , which is given by

$$\gamma_j = \sqrt{\frac{\sum_{i=1}^L [(\boldsymbol{\theta}_i^j)^T \mathbf{B}^{-1} \boldsymbol{\theta}_i^j]}{\sum_{i=1}^L \text{Tr}((\boldsymbol{\Phi}_i^j)^T (\boldsymbol{\Sigma}_{y_i}^*)^{-1} \boldsymbol{\Phi}_i^j \mathbf{B})}}, j=1, \dots, K. \quad (14)$$

### 2.1.2 Estimating $\mathbf{B}$ and $\beta$

The EM technique is used to find  $\Omega = \{\mathbf{B}, \beta\}$ . We proceed by first treating  $\boldsymbol{\theta}_i$  as hidden variables and then maximizing

$$\begin{aligned} W(\beta, \mathbf{B}) &= E_{\boldsymbol{\Theta}} \Big|_{\mathbf{Y}, \Omega^{(\text{old})}} [\log p(\mathbf{Y}, \boldsymbol{\Theta} | \Omega)] \\ &= E_{\boldsymbol{\Theta}} \Big|_{\mathbf{Y}, \Omega^{(\text{old})}} [\log p(\mathbf{Y} | \boldsymbol{\Theta}, \beta)] \\ &\quad + E_{\boldsymbol{\Theta}} \Big|_{\mathbf{Y}, \Omega^{(\text{old})}} [\log p(\boldsymbol{\Theta} | \mathbf{B})], \end{aligned} \quad (15)$$

where  $\Omega^{(\text{old})}$  denotes the evaluated parameters in the previous iteration,

$$\begin{aligned} \log p(\boldsymbol{\Theta} | \mathbf{B}) &= \sum_{i=1}^L \log p(\boldsymbol{\theta}_i | \mathbf{B}) \\ &= -\frac{1}{2} \sum_{i=1}^L [M_i \log(2\pi)] \\ &\quad + \log(\det(\boldsymbol{\Sigma}_0)) + \boldsymbol{\theta}_i^T \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\theta}_i]. \end{aligned} \quad (16)$$

Here, we only consider the terms relating to  $\mathbf{B}$  in  $W(\beta, \mathbf{B})$  and use the notation

$$\begin{aligned} W_1(\mathbf{B}) &\triangleq E_{\boldsymbol{\Theta}} \Big|_{\mathbf{Y}, \Omega^{(\text{old})}} [\log p(\boldsymbol{\Theta} | \mathbf{B})] \\ &= -\frac{1}{2} \sum_{i=1}^L [M_i \log(2\pi)] \\ &\quad + \log((\det(\boldsymbol{\Gamma}))^S (\det(\mathbf{B}))^K) \\ &\quad + E[\boldsymbol{\theta}_i^T (\boldsymbol{\Gamma}^{-1} \otimes \mathbf{B}^{-1}) \boldsymbol{\theta}_i] \\ &= -\frac{1}{2} \sum_{i=1}^L [M_i \log(2\pi) + S \log(\det(\boldsymbol{\Gamma}))] \\ &\quad + K \log(\det(\mathbf{B})) \\ &\quad + \text{Tr}[(\boldsymbol{\Gamma}^{-1} \otimes \mathbf{B}^{-1}) (\boldsymbol{\Sigma}_{\boldsymbol{\theta}_i} + \boldsymbol{\mu}_{\boldsymbol{\theta}_i} \boldsymbol{\mu}_{\boldsymbol{\theta}_i}^T)]. \end{aligned} \quad (17)$$

The partial derivative of (17) with respect to  $\mathbf{B}$ , which is symmetric and positive definite because it characterizes

the covariance matrix of every signal block (see the definition given above (5)), is

$$\begin{aligned} \frac{\partial W_1(\mathbf{B})}{\partial \mathbf{B}} &= -\frac{1}{2} \sum_{i=1}^L \left[ K \mathbf{B}^{-1} - \sum_{j=1}^K \left( \frac{1}{\gamma_j} \mathbf{B}^{-1} \left( \boldsymbol{\Sigma}_{\boldsymbol{\theta}_i}^j \right. \right. \right. \\ &\quad \left. \left. + \boldsymbol{\mu}_{\boldsymbol{\theta}_i}^j (\boldsymbol{\mu}_{\boldsymbol{\theta}_i}^j)^T \right) \mathbf{B}^{-1} \right) \Big] \\ &= -\frac{1}{2} \left[ K L \mathbf{B}^{-1} - \mathbf{B}^{-1} \left( \sum_{i=1}^L \sum_{j=1}^K \left( \frac{1}{\gamma_j} \left( \boldsymbol{\Sigma}_{\boldsymbol{\theta}_i}^j \right. \right. \right. \right. \\ &\quad \left. \left. + \boldsymbol{\mu}_{\boldsymbol{\theta}_i}^j (\boldsymbol{\mu}_{\boldsymbol{\theta}_i}^j)^T \right) \right) \mathbf{B}^{-1} \right], \end{aligned} \quad (18)$$

where  $\boldsymbol{\mu}_{\boldsymbol{\theta}_i}^j \triangleq \boldsymbol{\mu}_{\boldsymbol{\theta}_i}((j-1)S+1:jS)$ ,  $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_i}^j \triangleq \boldsymbol{\Sigma}_{\boldsymbol{\theta}_i}((j-1)S+1:jS, (j-1)S+1:jS)$ . Setting (18) to zero yields

$$\mathbf{B} = \frac{1}{KL} \sum_{i=1}^L \sum_{j=1}^K \left( \frac{1}{\gamma_j} \left( \boldsymbol{\Sigma}_{\boldsymbol{\theta}_i}^j + \boldsymbol{\mu}_{\boldsymbol{\theta}_i}^j (\boldsymbol{\mu}_{\boldsymbol{\theta}_i}^j)^T \right) \right). \quad (19)$$

Similar to [14], we improve the performance of the algorithm by restraining the matrix  $\mathbf{B}$ . Specifically, we attempt to find a positive definite and symmetric matrix  $\hat{\mathbf{B}}$  to approximate  $\mathbf{B}$ . Mathematically, we set  $\hat{\mathbf{B}}$  to be a Toeplitz matrix equal to

$$\hat{\mathbf{B}} = \text{Toeplitz}([1, r, \dots, r^{S-1}]) = \begin{bmatrix} 1 & r & \dots & r^{S-1} \\ \vdots & & & \vdots \\ r^{S-1} & r^{S-2} & \dots & 1 \end{bmatrix},$$

where  $r = \frac{m_1}{m_0}$ ,  $m_0$ , and  $m_1$  are obtained by averaging the elements along the main diagonal and the main sub-diagonal of  $\mathbf{B}$  in (19). As a result, the approximated version of  $\mathbf{B}$  is fully characterized by  $r$ . This method can also be applied with some modifications to the case where signal blocks have different sizes. In particular, in this case, we first compute  $\bar{r} = \frac{\bar{m}_1}{\bar{m}_0}$ , where  $\bar{m}_0 = \sum_{j=1}^K m_0^j$  and  $\bar{m}_1 = \sum_{j=1}^K m_1^j$ .  $m_0^j$  and  $m_1^j$  are obtained by averaging the elements along the main diagonal and the main sub-diagonal of  $\mathbf{B}_j$ , where it can be shown that  $\mathbf{B}_j = \frac{1}{L\gamma_j} \sum_{i=1}^L \left( \boldsymbol{\Sigma}_{\boldsymbol{\theta}_i}^j + \boldsymbol{\mu}_{\boldsymbol{\theta}_i}^j (\boldsymbol{\mu}_{\boldsymbol{\theta}_i}^j)^T \right)$ .  $\mathbf{B}_j$  are approximated with  $\hat{\mathbf{B}}_j = \text{Toeplitz}([1, \bar{r}, \dots, \bar{r}^{S_j-1}])$  such that again,  $\hat{\mathbf{B}}_j$  depend on the value of  $\bar{r}$  only. Here,  $S_j$  is the size of block  $j$ .

We next evaluate  $\beta$ . Consider the terms relating to  $\beta$  in  $W(\beta, \mathbf{B})$  and use the notation

$$\begin{aligned}
 W_2(\beta) &\triangleq E_{\Theta|Y, \Omega(\text{old})} [\log p(\mathbf{Y}|\Theta, \beta)] \\
 &= -\frac{1}{2} \sum_{i=1}^L \left[ M_i S \log(2\pi) - M_i S \log(\beta) \right. \\
 &\quad \left. + \beta E_{\theta_i|y_i, \Omega(\text{old})} [\|\mathbf{y}_i - \Phi_i \theta_i\|_2^2] \right] \\
 &= -\frac{1}{2} \sum_{i=1}^L \left[ M_i S \log(2\pi) - M_i S \log(\beta) \right. \\
 &\quad \left. + \beta \left[ \|\mathbf{y}_i - \Phi_i \boldsymbol{\mu}_{\theta_i}\|_2^2 + E \left[ \|\Phi_i (\theta_i - \boldsymbol{\mu}_{\theta_i})\|_2^2 \right] \right] \right] \\
 &= -\frac{1}{2} \sum_{i=1}^L \left[ M_i S \log(2\pi) - M_i S \log(\beta) \right. \\
 &\quad \left. + \beta \left[ \|\mathbf{y}_i - \Phi_i \boldsymbol{\mu}_{\theta_i}\|_2^2 + \text{Tr}(\boldsymbol{\Sigma}_{\theta_i} \Phi_i^T \Phi_i) \right] \right]. \tag{20}
 \end{aligned}$$

Differentiating  $W_2(\beta)$  with respect to  $\beta$  and then setting the result to zero, we obtain

$$\begin{aligned}
 \frac{\partial W_2}{\partial \beta} &= -\frac{1}{2} \left[ -\frac{1}{\beta} \sum_{i=1}^L M_i S + \beta \sum_{i=1}^L \left[ \|\mathbf{y}_i - \Phi_i \boldsymbol{\mu}_{\theta_i}\|_2^2 \right. \right. \\
 &\quad \left. \left. + \text{Tr}(\boldsymbol{\Sigma}_{\theta_i} \Phi_i^T \Phi_i) \right] \right] = 0. \tag{21}
 \end{aligned}$$

We have, after some manipulations,

$$\beta = \frac{\sum_{i=1}^L M_i S}{\sum_{i=1}^L \left[ \|\mathbf{y}_i - \Phi_i \boldsymbol{\mu}_{\theta_i}\|_2^2 + \text{Tr}(\boldsymbol{\Sigma}_{\theta_i} \Phi_i^T \Phi_i) \right]}. \tag{22}$$

The iterative process for estimating  $\beta, \boldsymbol{\gamma}$ , and  $\mathbf{B}$  starts with initial solution guesses of  $\boldsymbol{\mu}_{\theta_i}, \boldsymbol{\Sigma}_{\theta_i}, \gamma_j, \mathbf{B}$ , and  $\beta$ . We then evaluate sequentially (6), (7), and (14) to find  $\gamma_j$  and proceed to find the updated estimates of  $\mathbf{B}$  and  $\beta$  using (19) and (22). With the obtained estimates of the sharing parameters, the original signals  $\theta_i$  of the  $L$  CS tasks can be reconstructed by following the MCS technique [15]. This completes the development of MBSBL.

## 2.2 EMBSBL

We shall present the new EMBSBL algorithm that is based on the developed MBSBL technique. Similar to [14], we first assume that all the blocks are of equal size  $h$  and the non-zero blocks are arbitrarily located. We will show via simulations that EMBSBL is not sensitive to the choice of  $h$ . There are  $p \triangleq N - h + 1$  possible blocks in every signal  $\theta_i$ . The  $j$ th block starts at the  $j$ th element of  $\theta_i$  and continues until the  $(j + h - 1)$ th element. All the non-zero

elements of  $\theta_i$  lie within a subset of these blocks. From the analysis above, we can have the decomposition of  $\theta_i$

$$\theta_i = \sum_{j=1}^p \mathbf{E}_j \mathbf{z}_{i,j}, \tag{23}$$

where  $\mathbf{z}_{i,j} \in \mathcal{R}^{h \times 1}$ ;  $E(\mathbf{z}_{i,j} \mathbf{z}_{i,k}^T) = \delta_{j,k} \gamma_j \mathbf{B}$  ( $\delta_{j,k} = 1$  if  $j = k$ ; otherwise,  $\delta_{j,k} = 0$ ); and  $\mathbf{z}_i = [\mathbf{z}_{i,1}^T, \dots, \mathbf{z}_{i,p}^T]^T \sim \mathcal{N}_{\mathbf{z}_i}(\mathbf{0}, \tilde{\boldsymbol{\Sigma}}_0)$ ,  $\tilde{\boldsymbol{\Sigma}}_0 = \text{diag}(\gamma_1 \mathbf{B}, \dots, \gamma_p \mathbf{B}) \in \mathcal{R}^{ph \times ph}$ ;  $\mathbf{E}_j \in \mathcal{R}^{N \times h}$  is a zero matrix except that the submatrix composed of its  $j$ th row to  $(j + h - 1)$ th row is replaced by the identity matrix  $\mathbf{I}$ , and  $\mathbf{E}_j$  is the same for every  $\theta_i$ . The CS model (1) can then be re-expressed as

$$\mathbf{y}_i = \sum_{j=1}^p \Phi_i \mathbf{E}_j \mathbf{z}_{i,j} + \mathbf{n}_i = \mathbf{A}_i \mathbf{z}_i + \mathbf{n}_i, (i = 1, \dots, L) \tag{24}$$

where

$$\mathbf{A}_i = [\mathbf{A}_{i,1}, \dots, \mathbf{A}_{i,p}] \tag{25}$$

and

$$\mathbf{A}_{i,j} = \Phi_i \mathbf{E}_j. \tag{26}$$

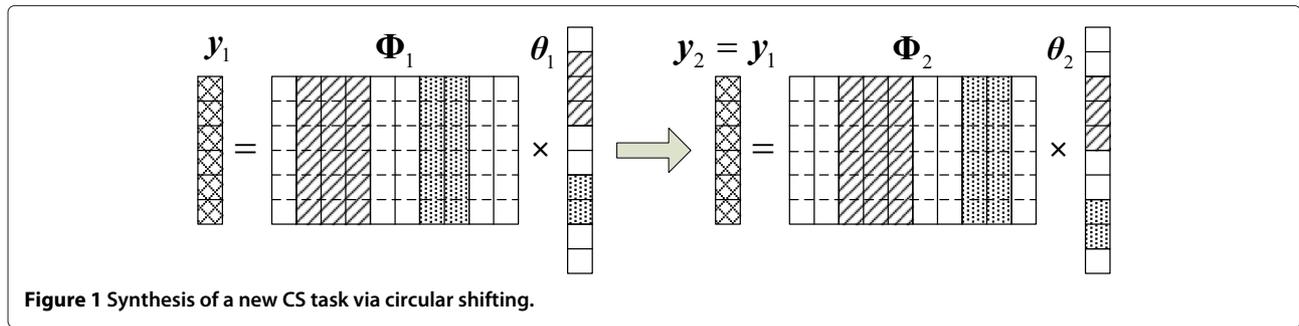
The new CS model (24) has its signals with the property of block-sparsity and the intra-block correlation is explicit.  $\mathbf{z}_i$  can be recovered using MBSBL and by utilizing (23), the original signals  $\theta_i$  of the CS tasks can then be found, which finishes the development of EMBSBL for recovering block-sparse signals under the MCS framework.

## 3 SEMBSBL algorithm

The EMBSBL cannot be directly applied to recover a single block-sparse signal in the single CS task scenario, due to its nature of being an MCS technique. We shall augment it with the idea of synthesized MCS to address this difficulty. CS task synthesis via circular-shifting operation is developed below. This section ends with the improved EMBSBL algorithm, namely synthesized EMBSBL (SEMBSBL), which utilizes the MDL principle to determine the optimal number of synthesized CS tasks to achieve satisfactory signal recovery performance.

### 3.1 Synthesis of multiple CS tasks

Figure 1 illustrates synthesizing multiple CS tasks from a single one. The absence of measurement noise is assumed here to improve clarity. The original CS task is  $\mathbf{y}_1 = \Phi_1 \theta_1$ , where  $\theta_1$  is the block-sparse signal to be recovered and it has two non-zero clusters (shadowed). The columns of the measurement matrix  $\Phi_1$  corresponding to the non-zero elements in  $\theta_1$  are also shadowed for illustration. Figure 1 indicates that a new CS task can be synthesized from the original one by circularly shifting the columns of



$\Phi_1$  to the right by one column. In this way, the new CS task has a new measurement matrix  $\Phi_2$  and a new signal  $\theta_2$  whose the elements are generated by circularly shifting  $\theta_1$  downward by one sample. The new CS task has the same compressive measurements as the original one. We assume that this observation holds also for the case where measurement noise is present. Comparing  $\theta_1$  with  $\theta_2$  reveals that the locations of their non-zero elements have overlaps. This implies that  $\theta_1$  and  $\theta_2$  are correlated. This forms the basis for utilizing EMBSBL in block-sparse signal recovery. Additional CS tasks can be synthesized by following a similar approach but with different directions and shifting amounts of the circular-shifting operations.

It can be expected that due to the block-sparsity of the signal to be recovered, the signals of some synthesized CS tasks may not be well correlated with others. In other words, they only have few overlapping non-zero elements. The utilization of these CS tasks in recovering the original signal via EMBSBL would lead to poor signal reconstruction performance. To address this problem, we propose to utilize the MDL principle to determine the number of synthesized CS tasks for the block-sparse signal reconstruction, as will be detailed in the following subsection.

### 3.2 SEMBSBL

This section presents the proposed SEMBSBL algorithm. We shall first provide a method for evaluating the signal recovery quality of EMBSBL for a given set of synthesized CS tasks. This is essential for selecting the optimal set of synthesized CS tasks for block-sparse signal recovery. For this purpose, we apply the MDL principle. Basically, it states that among a set of competing statistical models, the best model is the one having the minimum code length for the given data [16,17]. This is mathematically equivalent to solving  $\hat{Q} = \arg \min_{Q \in \mathfrak{M}} CL(\mathbf{y}, Q)$ , where  $\mathfrak{M}$

denotes the set of possible models and  $CL(\mathbf{y}, Q)$  is the code length function. We set  $CL(\mathbf{y}, Q)$  to be the Shannon code length [18], i.e.,  $CL(\mathbf{y}, Q) = -\log_2 p(\mathbf{y}, Q)$ , where  $p(\mathbf{y}, Q)$  is the probability density function of  $\mathbf{y}$  under the model  $Q$ .

For the problem of applying EMBSBL to recovering the block-sparse signal in a single CS task scenario (without loss of generality, we assume the task is  $\mathbf{y}_1 = \Phi_1 \mathbf{x} + \mathbf{n}_1$ ), we denote the estimates of the sharing parameters  $\beta, \gamma, \mathbf{B}$  as  $\hat{\beta}, \hat{\gamma}, \hat{\mathbf{B}}$ . They are output by the EMBSBL algorithm for a given set of synthesized CS tasks. The description length for  $\mathbf{y}_1$ ,  $CL(\mathbf{y}_1)$ , can be then expressed as, after using (9) and setting  $L = 1$ ,

$$\begin{aligned}
 CL(\mathbf{y}_1) &= CL(\mathbf{y}_1 | \hat{\beta}, \hat{\gamma}, \hat{\mathbf{B}}) + CL(\hat{\beta}, \hat{\gamma}, \hat{\mathbf{B}}) \\
 &= -\log_2 p(\mathbf{y}_1 | \hat{\beta}, \hat{\gamma}, \hat{\mathbf{B}}) \\
 &\quad - \log_2 p(\hat{\beta}) - \log_2 p(\hat{\gamma}) - \log_2 p(\hat{\mathbf{B}}) \\
 &= \frac{1}{2} (M_1 \log_2(2\pi) + \log_2(\det(\mathbf{C}_1)) \\
 &\quad + \mathbf{y}_1^T \mathbf{C}_1^{-1} \mathbf{y}_1 \log_2 e) + \text{const},
 \end{aligned} \tag{27}$$

where  $CL(\mathbf{y}_1 | \hat{\beta}, \hat{\gamma}, \hat{\mathbf{B}}) = -\log_2 p(\mathbf{y}_1 | \hat{\beta}, \hat{\gamma}, \hat{\mathbf{B}})$  measures the goodness of fit between the data and the current model,  $CL(\hat{\beta}, \hat{\gamma}, \hat{\mathbf{B}}) = -\log_2 p(\hat{\beta}) - \log_2 p(\hat{\gamma}) - \log_2 p(\hat{\mathbf{B}})$  represents the model complexity, and  $p(\hat{\beta})$ ,  $p(\hat{\gamma})$ , and  $p(\hat{\mathbf{B}})$  denote the prior distributions of  $\hat{\beta}$ ,  $\hat{\gamma}$ ,  $\hat{\mathbf{B}}$ ,  $e$  which are the base of the natural logarithm. Because we do not impose any specific distributions on  $\hat{\beta}, \hat{\gamma}, \hat{\mathbf{B}}$ , their prior probability distributions are thus set to be the uniform distributions. In other words,  $-\log_2 p(\hat{\beta}) - \log_2 p(\hat{\gamma}) - \log_2 p(\hat{\mathbf{B}})$  is a constant.  $\mathbf{C}_1 = \hat{\beta}^{-1} \mathbf{I} + \mathbf{A}_1 \hat{\Sigma}_0 \mathbf{A}_1^T$ ,  $\hat{\Sigma}_0 = \text{diag}(\hat{\gamma}_1 \hat{\mathbf{B}}, \dots, \hat{\gamma}_K \hat{\mathbf{B}})$ .

We are now ready to present the proposed SEMBSBL algorithm. It is an iterative method that improves the signal recovery quality gradually. In each iteration, a new CS tasks is synthesized using the circular-shifting operation as illustrated in Figure 1. The newly produced task is applied together with the previously synthesized CS tasks as well as the original CS task in EMBSBL for jointly reconstructing the block-sparse signal. The above process continues until the number of synthesized CS tasks reaches a pre-specified value or including the newly synthesized CS task does not lead to better signal reconstruction quality (or equivalently, reduced code length for describing the data, which is given in (27)).

The algorithm is summarized in Algorithm 1.  $l_{\max}$  is the user-specified maximum number of the synthesized CS tasks.  $\text{EMBSBL}^l(\mathbf{Y}, \mathbf{A})$  represents the application of EMBSBL for signal reconstruction in the  $l$ th iteration and it uses  $l$  CS tasks.  $\mathbf{Y}$  and  $\mathbf{A}$  collect the compressive measurements and their associated measurement matrices of the  $l$  CS tasks. The output of  $\text{EMBSBL}^l(\mathbf{Y}, \mathbf{A})$  is  $\hat{\beta}^l, \hat{\gamma}^l, \hat{\mathbf{B}}^l, \hat{\theta}_1^l$ , which are the estimates of the sharing parameters  $\beta, \gamma, \mathbf{B}$  and the original signal  $\theta_1$ . The operators Left( $\mathbf{A}_1, l$ ) and Right( $\mathbf{A}_1, l$ ) denote circular shifting the columns of  $\mathbf{A}_1$  to the left and to the right by  $l$  columns.

---

**Algorithm 1** SEMBSBL

---

- 1 Inputs:  $\mathbf{y}_1, \Phi_1, l_{\max}$ .
  - 2 Output:  $\hat{\theta}_1$ .
  - 3 Initialize  $\mathbf{Y} \leftarrow \{\mathbf{y}_1\}$ ; calculate  $\mathbf{A}_1$  using (25) and (26),  
 $\mathbf{A} \leftarrow \{\mathbf{A}_1\}; l \leftarrow 2$ ;  
 $\hat{\beta}^1, \hat{\gamma}^1, \hat{\mathbf{B}}^1, \hat{\theta}_1^1 \leftarrow \text{EMBSBL}^1(\mathbf{Y}, \mathbf{A})$ ; calculate  $CL_1(\mathbf{y}_1)$   
 using (27).
  - 4  $\mathbf{Y} \leftarrow \{\mathbf{Y}, \mathbf{y}_1\}; \mathbf{A} \leftarrow \{\mathbf{A}, \text{Left}(\mathbf{A}_1, l-1)\}$  if  $l$  is even, or  
 $\mathbf{A} \leftarrow \{\mathbf{A}, \text{Right}(\mathbf{A}_1, l-1)\}$   
 if  $l$  is odd;  $\hat{\beta}^l, \hat{\gamma}^l, \hat{\mathbf{B}}^l, \hat{\theta}_1^l \leftarrow \text{EMBSBL}^l(\mathbf{Y}, \mathbf{A})$ ; calculate  
 $CL_l(\mathbf{y}_1)$  using (27).
  - 5 If  $CL_l(\mathbf{y}_1) < CL_{l-1}(\mathbf{y}_1)$  or  $l = l_{\max}$ ,  $\hat{\theta}_1 \leftarrow \hat{\theta}_1^l$  and  
 terminate the algorithm; otherwise,  
 $l \leftarrow l + 1$  and goto step 4.
- 

#### 4 Simulations

We shall provide simulation results to demonstrate the performance of the EMBSBL algorithm proposed in Section 2 and the SEMBSBL algorithm developed in Section 3. The signal reconstruction error is quantified using  $\|\theta_i - \hat{\theta}_i\|_2 / \|\theta_i\|_2$ , where  $\theta_i$  and  $\hat{\theta}_i$  are the true and the estimated signals. The elements of the measurement matrix  $\Phi_i$  are initially drawn from the standard normal distribution  $\mathcal{N}(0, 1)$  and each row of  $\Phi_i$  is then normalized to have a unit norm.

In the first experiment, we simulate a two CS task scenario ( $L = 2$ ) where the original signals both have a length of  $N = 500$  and each contains 50 spikes with different amplitudes at random locations. The two signals also have six non-zero blocks with random sizes and they are at non-overlapping random locations. We consider two cases where 80% and 100% of the spikes of the two signals are at the same positions.

The signal-to-noise ratio (SNR) in log scale is defined as  $\text{SNR} \triangleq 20 \log_{10}(\sqrt{\beta} \|\Phi_i \theta_i\|_2)$  and zero-mean Gaussian noise is added to every CS measurement vector. We set  $\text{SNR} = 15$  dB and use MCS, BSBL-BO, and EBSBL-BO

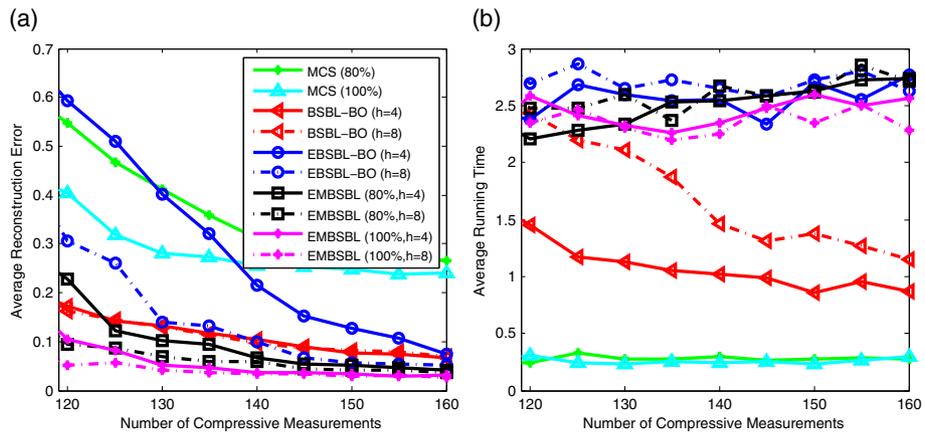
as benchmark techniques for comparison. When implementing BSBL-BO, EBSBL-BO, and EMBSBL, we set the block size parameter  $h$  to be  $h = 4$  and  $h = 8$  to illustrate the impact of different choice of  $h$  on their performance. The signal reconstruction error results shown are obtained via averaging over 50 ensemble runs.

Figure 2 compares MCS, BSBL-BO, EBSBL-BO, and EMBSBL in terms of their signal reconstruction errors as a function of the number of the compressive measurements. In this simulation, the intra-block correlation coefficient for each block is uniformly distributed between 0 and 0.1. Simulation results indicate that for BSBL-BO and EBSBL-BO, the performance curves when 80% and 100% of the spikes of the two original signals are at the same locations are very similar to each other. Therefore, to improve the clarity of the figures, we only provide in Figure 2 and the following Figures 3 and 4 the results when 80% of the spikes have the same locations.

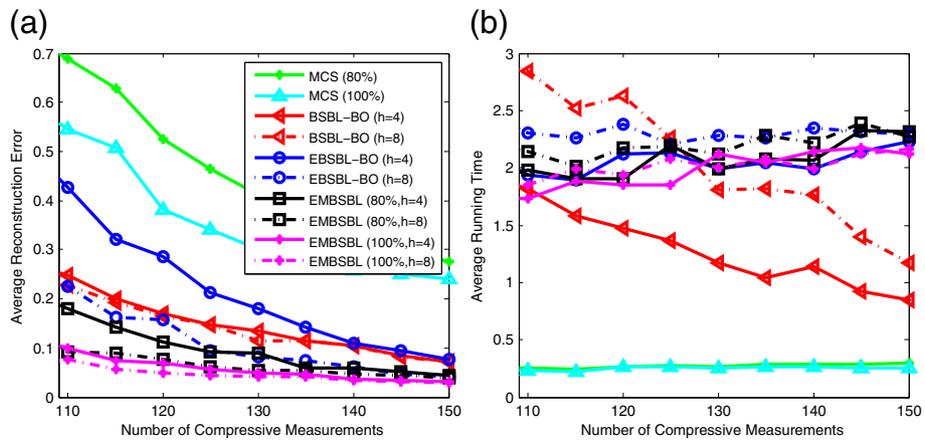
We can see from Figure 2a that the proposed EMBSBL has the least signal recovery error and its performance improves as more spikes of the original signals share the same locations. More importantly, EMBSBL is less sensitive to the choice of the block size parameter  $h$  than BSBL-BO and EBSBL-BO. This is because the new EMBSBL technique is an MCS algorithm that recovers the original signals of multiple CS tasks jointly. Compared with BSBL-BO and EBSBL-BO that are single-task CS algorithms, EMBSBL explores the inter-correlation among original signals to improve performance, besides the intra-block correlation. The use of this additional information improves the robustness of EMBSBL to the deviation of the presumed block size from the true value. Finally, as shown in Figure 2b, EMBSBL has the comparable complexity as EBSBL-BO, despite of being an MCS technique.

We repeat the simulation that produced Figure 2 but this time, we allow the intra-block correlation coefficient for each block to be uniformly distributed between 0.4 and 0.5. The obtained results are summarized in Figure 3. It can be seen that the proposed EMBSBL continues to offer the best signal recovery performance. Besides, the increase in the intra-block correlation improves the performance of the EBSBL-BO.

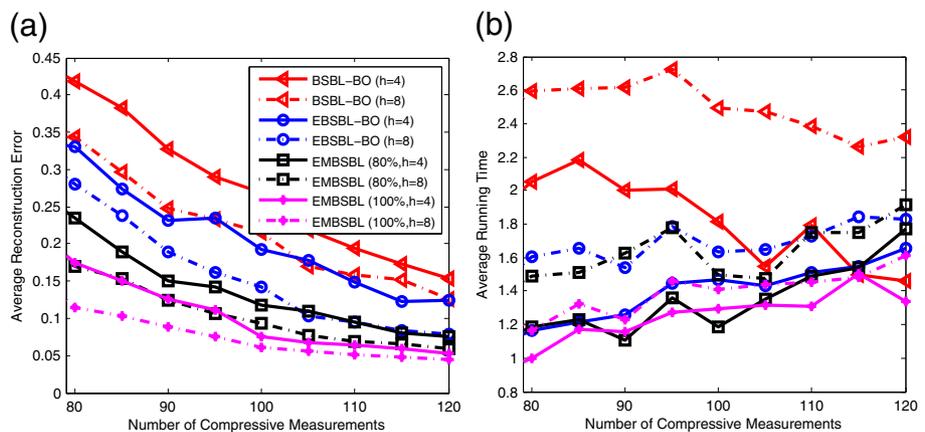
In producing Figure 4, we further increase the intra-block correlation coefficient for each block to be uniformly distributed from 0.8 to 0.9. The performance curve of MCS is excluded from the figure for sake of clarity, because in this case, the signal reconstruction error of MCS is large. We can find in Figure 4 that under significant intra-block correlation, the proposed EMBSBL still has the best signal recovery performance. Besides, comparing Figure 4 with Figures 2 and 3 reveals that higher intra-block correlation leads to improved performance of the proposed EMBSBL method. This is because



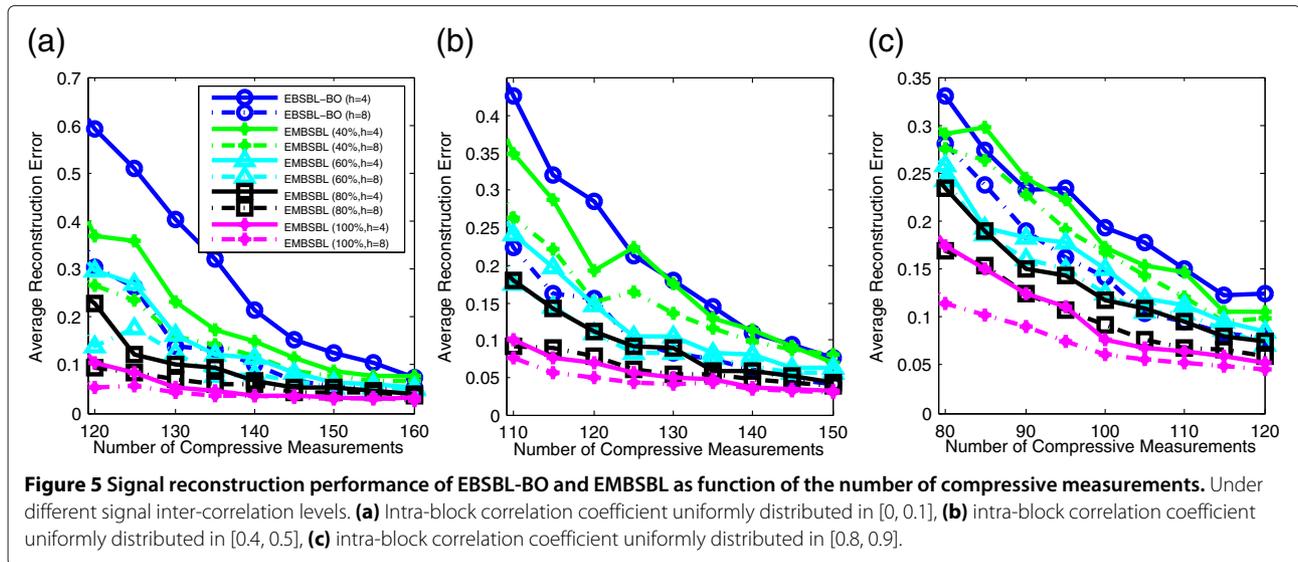
**Figure 2** Comparison of MCS, BSBL-BO, EBSBL-BO, and EMBSBL as a function of the number of compressive measurements. When the intra-block correlation coefficient for each block is uniformly distributed between 0 and 0.1. **(a)** Signal reconstruction error performance, **(b)** algorithm running time.



**Figure 3** Comparison of MCS, BSBL-BO, EBSBL-BO, and EMBSBL as function of the number of compressive measurements. When the intra-block correlation coefficient for each block is uniformly distributed between 0.4 and 0.5. **(a)** Signal reconstruction error performance, **(b)** algorithm running time.



**Figure 4** Comparison of MCS, BSBL-BO, EBSBL-BO, and EMBSBL as function of the number of compressive measurements. When the intra-block correlation coefficient for each block is uniformly distributed between 0.8 and 0.9. **(a)** Signal reconstruction error performance, **(b)** algorithm running time.

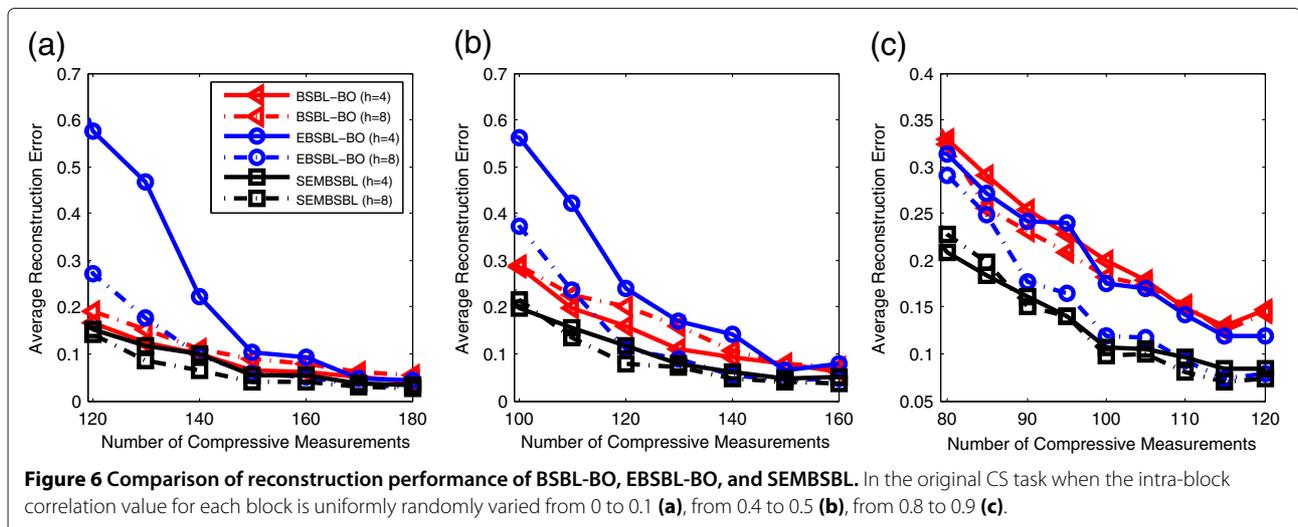


it explicitly utilizes the intra-block correlation for better signal recovery (see Section 2).

We next study the impact of different signal inter-correlation levels on the signal recovery performance of EMBSBL. The simulation setup is the same as that leading to Figure 2 except that the percentage of overlapping non-zeros of the two original signals is set to be 40%, 60%, 80%, and 100%. Three sets of simulation results are produced. They correspond to the intra-block correlation coefficient being uniformly distributed within  $[0, 0.1]$ ,  $[0.4, 0.5]$ , and  $[0.8, 0.9]$ . The obtained simulation results are summarized in Figure 5. We find that EMBSBL can recover the original signals with reduced signal reconstruction error as the inter-correlation among original signals increases. This observation holds for different intra-block correlation coefficients. The performance improvement is somewhat

expected, because the EMBSBL algorithm is an MCS technique and its signal recovery performance would benefit from increased inter-correlation among original signals.

To validate the development of SEMBSBL (see Section 3), we consider a single CS task scenario. Again, the original signal has a length of  $N = 500$  and it contains 50 spikes at random locations. Besides, the signal include six non-zero blocks with random sizes and random but non-overlapping locations. We set  $\text{SNR} = 15$  dB in the simulation and use BSBL-BO and EBSBL-BO as benchmarking techniques. For BSBL-BO, EBSBL-BO, and the proposed SEMBSBL, we generate signal recovery error performance curves with the block size parameter setting to be  $h = 4$  and  $h = 8$ . The results are averaged over 50 runs. For the proposed SEMBSBL, we set the pre-specified maximum number of the synthesized CS tasks



$l_{\max}$  to be six. We consider three cases, where the intra-block correlation coefficient for each block is uniformly distributed within  $[0, 0.1]$ ,  $[0.4, 0.5]$ , and  $[0.8, 0.9]$ . The obtained simulation results are shown in Figure 6. It can be observed that the proposed SEMBSBL outperforms benchmark algorithms and SEMBSBL is less sensitive to the choice of the parameter  $h$  than EBSBL-BO. But the running time of SEMBSBL is high, since it executes the EMBSBL algorithm a couple of times before it finds an optimal set of synthesized CS tasks for signal recovery.

## 5 Conclusion

In this paper, a novel algorithm for jointly recovering multiple block-sparse signals from their compressive measurements, termed as the EMBSBL algorithm, was developed. EMBSBL exploits both the statistical correlation among signals and signals' intra-block correlation to achieve superior signal recovery performance. Moreover, the new algorithm eliminates the requirement on the availability of *a priori* information on the sparsity structure of the original signal. We also developed in this paper SEMBSBL that applies EMBSBL to the single CS task case. It synthesizes new CS tasks from the single CS task via simple circular-shifting operations to make EMBSBL applicable. The MDL principle was adopted to determine the proper set of the synthesized CS tasks for reconstructing the block-sparse signal. Computer simulations were carried out and revealed that the proposed EMBSBL and SEMBSBL are able to outperform existing techniques in providing greatly enhanced block-sparse signal reconstruction performance at the cost of increased computational complexity.

### Competing interests

The authors declare that they have no competing interests.

### Acknowledgements

The authors wish to thank the associate editor and the anonymous reviewers for their constructive suggestions. The authors thank Zhilin Zhang, Bhaskar D. Rao, Shihao Ji, and David Dunson for sharing codes of their algorithms. This work was supported in part by Hunan Provincial Innovation Foundation for Postgraduates under Grant CX2012B019, Fund of Innovation, Graduate School of National University of Defense Technology under grant B120404, and National Natural Science Foundation of China (no. 61304264).

### Author details

<sup>1</sup>College of Electronic Science and Engineering, National University of Defense Technology, Deya Road, Changsha 410073, People's Republic of China.

<sup>2</sup>School of Internet of Things (IoT) Engineering, Jiangnan University, Lihu Road, Wuxi 214122, People's Republic of China.

Received: 13 June 2013 Accepted: 25 January 2014

Published: 1 February 2014

### References

1. Donoho DL, Compressed sensing. *IEEE Trans. Inf. Theory*. **52**(4), 1289–1306 (2006)
2. Mishali M, Eldar YC, Blind multi-band signal reconstruction: compressed sensing for analog signals. *IEEE Trans. Signal Process.* **57**(3), 993–1009 (2009)

3. Parvareh F, Vikalo H, Misra S, Hassibi B, Recovering sparse signals using sparse measurement matrices in compressed DNA microarrays. *IEEE J. Sel. Topics Signal Process.* **2**(3), 275–285 (2008)
4. Baraniuk RG, Cevher V, Duarte MF, Hegde C, Model-based compressive sensing. *IEEE Trans. Inf. Theory*. **56**(4), 1982–2001 (2010)
5. Needell D, Tropp JA, CoSaMP: iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmon. Anal.* **26**, 301–321 (2009)
6. Pati YC, Rezaifar R, Krishnaprasad PS, Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition, in *Proceedings of the 27th Ann. Asilomar Conf. Signals, Systems, and Computers* (Pacific Grove, CA, 1–3 Nov 1993)
7. Eldar YC, Kuppinger P, Bolcskei H, Block-sparse signals: uncertainty relations and efficient recovery. *IEEE Trans. Signal Process.* **58**(6), 3042–3054 (2010)
8. Zou J, Fu Y, Xie S, A block fixed point continuation algorithm for block-sparse reconstruction. *IEEE Trans. Signal Process. Lett.* **19**(6), 364–367 (2012)
9. Elhamifar E, Vidal R, Block-sparse recovery via convex optimization. *IEEE Trans. Signal Process.* **60**(8), 4094–4107 (2012)
10. Zelnik-Manor L, Rosenblum K, Eldar YC, Dictionary optimization for block-sparse representations. *IEEE Trans. Signal Process.* **60**(5), 2386–2395 (2012)
11. Yu L, Sun H, Barbot JP, Zheng G, Bayesian compressive sensing for cluster structured sparse signals. *Signal Process.* **92**(1), 259–269 (2012)
12. Peleg T, Eldar Y, Elad M, Exploiting statistical dependencies in sparse representations for signal recovery. *IEEE Trans. Signal Process.* **60**(5), 2286–2303 (2012)
13. Zhang Z, Rao BD, Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning. *IEEE J. Selected Topics Signal Process.* **5**(5), 912–926 (2011)
14. Zhang Z, Rao BD, Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation. *IEEE Trans. Signal Process.* **61**(8), 2009–2015 (2013)
15. Ji S, Dunson D, Carin L, Multi-task compressive sensing. *IEEE Trans. Signal Process.* **57**(1), 92–106 (2009)
16. Barron A, Rissanen J, Yu B, The minimum description length principle in coding and modeling. *IEEE Trans. Inf. Theory*. **44**(6), 2743–2760 (1998)
17. Ramirez I, Sapiro G, An MDL framework for sparse coding and dictionary learning. *IEEE Trans. Signal Process.* **60**(6), 2913–2927 (2012)
18. Cover T, Thomas J, *Elements of Information Theory*, 2nd ed (Wiley, New York, 2006)

doi:10.1186/1687-6180-2014-14

Cite this article as: Wang et al.: SBL-based multi-task algorithms for recovering block-sparse signals with unknown partitions. *EURASIP Journal on Advances in Signal Processing* 2014 **2014**:14.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)