LOW-RANK AND JOINT-SPARSE SIGNAL RECOVERY FOR SPATIALLY AND TEMPORALLY CORRELATED DATA USING SPARSE BAYESIAN LEARNING

Yanbin Zhang, Yangqing Li, Changchuan Yin, and Kesen He

Beijing Laboratory of Advanced Information Networks,
Beijing Key Laboratory of Network System Architecture and Convergence,
Beijing University of Posts and Telecommunications, Beijing 100876, P. R. China
{iyanbin4, liyq, ccyin, kesenhe}@bupt.edu.cn

ABSTRACT

In order to meet the demands of data-intensive continuous monitoring in wireless body area network, we address a structured sparse signal recovery method to exploit both s-patial and temporal correlations in data using compressive sensing (CS). Using a simultaneously low-rank and joint-sparse (L&S) signal model, we employ a Bayesian learning treatment by incorporating an L&S-inducing prior over the data and the appropriate hyperpriors over all hyperparameters, resulting in effective reconstruction of the L&S data. Simulation results suggest that the proposed L&S-bSBL is superior to the state-of-the-art recovery methods in terms of computation burden and runtime cost.

Index Terms— Wireless Body Area Network, Bayesian Learning, Compressive Sensing, Low-rank and Joint-sparse

1. INTRODUCTION

Telemonitoring electroencephalogram (EEG) /electrocardiogram (ECG) data via wireless body area network (WBAN) [1, 2] is an evolving direction in e-Health. In WBAN, there are many sensors with limited computational power. They need to collect and encode data, then send the encoded data to a fusion center (FC) for decoding whose computational power is not at a premium. Therefore, new coding techniques are needed urgently. Compressive Sensing (CS) [1, 2] based techniques is a good solution to this problem since it requires a low-complexity encoder and sophisticated decoder.

In recent years, motivated by many source data which often contain some kinds of structures, people began to use structured signal model to improve the CS reconstruction performance and the reconstruction speed [2–5]. [1] assumes a block structure of the spatial correlated EEG signals in a transform domain (discrete cosine transform (DCT) or wavelet) and shows that a block Sparse Bayesian Learning (bSBL) algorithm yields good recovery results. [6] introduces a Bayesian message passing algorithm for solving the multiple measurement vector (MMV) problem when temporal correlation is present in the amplitudes of the non-zero signal coefficients. However, most of these works on energy-efficient data gathering only focus on exploiting either the spatial structure (i.e., the dependency relationship among different sources) or the temporal structure of data. The data in WBAN prevalently has both spatial and temporal correlations at the same time [7]. Instead of only using low-rankness [8] to recover the signal, [7] considers the structured spatial and temporal correlations jointly by assuming the spatio-temporal correlated data satisfies simultaneous low-rank and joint-sparse (L&S) structure, and obtains a superior performance.

In this paper, we propose a bSBL-based algorithm, L&S-bSBL, that incorporates the L&S-inducing prior over the data and the appropriate hyperpriors over all hyperparameters to recover compressed L&S data in WBAN. Specifically, we formulate our problem and transform it into a block single measurement vector (SMV) [6] problem. Then the structure of the covariance matrix of the L&S data is given. The inference problem is splited into two steps. Firstly, we get initial values of hyperparameters by assuming the covariance matrix of the data as a diagonal matrix. Secondly, we get the optimal reconstructed data by an expectation maximization (EM)-like algorithm. Simulation results show that L&S-bSBL has superior performances compared with existing algorithms with the same order of magnitude runtime consumption.

The rest of the paper is organized as follows: In Section 2, we introduce the problem statements. In Section 3, we provide the Bayesian learning inference for L&S-bSBL. In Section 4, we present the experimental results for both synthetic data and real-world applications. Finally the conclusions of this work are discussed in Section 5.

Notation: $p(\mathbf{A}) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ denotes the elements of $\mathbf{A}$ follows a Gaussian distribution with mean $\mathbf{0}$ and variance $\mathbf{\Sigma}$. $|\mathbf{A}|$ denotes the determinant of $\mathbf{A}$. $\|\mathbf{x}\|_2$ denotes the ℓ_2 norm of $\mathbf{x}$. $\mathbf{A} \otimes \mathbf{B}$ represents the Kronecker product of the two

This work was supported in part by the National Natural Science Foundation of China under Grant 61629101 and Grant 61671086, in part by the 111 Project under Grant B17007, and in part by the Director Funds of Beijing Key Laboratory of Network System Architecture and Convergence under Grant 2017BKL-NSAC-ZJ-04.

matrices $\mathbf{A}$ and $\mathbf{B}$. vec$[\mathbf{A}]$ denotes the vectorization of the matrix $\mathbf{A}$ formed by stacking its columns into a single column vector. $\mathbf{A}^\top$ denotes the transpose of $\mathbf{A}$. tr$(\mathbf{A})$ denotes the trace of $\mathbf{A}$.

2. PROBLEM FORMULATION AND SIGNAL MODEL

We consider a typical WBAN scenario in which there are m sensors to collect data $\mathbf{F} = [\mathbf{f}_1, \cdots, \mathbf{f}_m]^\top \in \mathbb{R}^{m\times n}$ in time synchronization, where $\mathbf{f}_i \in \mathbb{R}^{n\times 1}, i \in 1, 2, ..., m$ stands for the data collected by the ith sensor and $\mathbf{F}$ is the spatially and temporally correlated data matrix. Then, $\mathbf{F}$ is encoded by linearly mixing with $\mathbf{\Xi} \in \mathbb{R}^{p\times m}$ and transmitted to an FC, denoted as $\mathbf{Y} \in \mathbb{R}^{p\times n}$ after superimposed noise $\mathbf{V} \in \mathbb{R}^{p\times n}$. Finally, $\mathbf{F}$ is decoded in the FC using a CS algorithm. Fortunately, many spatially and temporally correlated WBAN data (e.g. EEG, ECG data) are highly correlated with the magnitude and locations of their non-zero elements in sparse domain. So we can get an approximately L&S matrix $\mathbf{X} \in \mathbb{R}^{m\times n}$ from $\mathbf{F} = \mathbf{\Psi}\mathbf{X}$, where $\mathbf{\Psi} \in \mathbb{R}^{m\times m}$ is a sparsifying basis (e.g. DCT matrix or wavelet matrix) [7]. And we have the formulation:

$$\mathbf{Y} = \mathbf{\Phi}\mathbf{X} + \mathbf{V}, \tag{1}$$

where $\mathbf{\Phi} = \mathbf{\Xi}\mathbf{\Psi}$. Here, this problem belongs to the multiple measurement vector (MMV) [6] problem.

Now, we consider a bSBL framework [1] by incorporating an L&S-inducing prior over the signals and the appropriate hyperpriors over all hyperparameters to recover $\mathbf{X}$. By transforming the MMV problem to the block single measurement vector (SMV) [6] problem, we have

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}, \tag{2}$$

where $\mathbf{y} = \text{vec}[\mathbf{Y}^\top] \in \mathbb{R}^{np\times 1}$, $\mathbf{A} = \mathbf{\Phi} \otimes \mathbf{I}_n \in \mathbb{R}^{np\times nm}$. Here $\mathbf{\Phi}$ is a known dictionary matrix and $\mathbf{I}_n \in \mathbb{R}^{n\times n}$ denotes an $n \times n$ identity matrix. $\mathbf{x} = \text{vec}[\mathbf{X}^\top] = [\mathbf{x}_1^\top, \cdots, \mathbf{x}_m^\top]^\top \in \mathbb{R}^{nm\times 1}$, $\mathbf{x}_i \in \mathbb{R}^{n\times 1}$ is the ith block in $\mathbf{x}$. K nonzero rows in $\mathbf{X}$ means K nonzero blocks in $\mathbf{x}$. Thus, $\mathbf{x}$ is block-sparse. $\mathbf{v} = \text{vec}[\mathbf{V}^\top] \in \mathbb{R}^{np\times 1}$.

We assume the elements $\mathbf{v}_i, i \in 1, \cdots, m$, of the noise vector $\mathbf{v}$ follow identically and independently distributed (i.i.d.) Gaussian variables with $p(\mathbf{v}_i) \sim \mathcal{N}(0, \lambda), \forall i$. For the problem in (2), we first define the Gaussian likelihood as

$$p(\mathbf{y}|\mathbf{x}; \mathbf{A}, \lambda) \sim \mathcal{N}_{y|x}(\mathbf{A}\mathbf{x}, \lambda\mathbf{I}) \propto \exp[-\frac{1}{2\lambda}\|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2], \tag{3}$$

then it produces the aggregate prior on $\mathbf{x}$ given by

$$p(\mathbf{x}; \gamma_i, \gamma_j, \mathbf{B}_{ij}, \forall i, j) \sim \mathcal{N}_x(\mathbf{0}, \mathbf{\Sigma}_0) \propto \exp[\mathbf{x}^\top \mathbf{\Sigma}_0^{-1}\mathbf{x}], \tag{4}$$

where $\mathbf{B}_{ij} \in \mathbb{R}^{n\times n}$ is a covariance matrix between $\mathbf{x}_i$ and $\mathbf{x}_j$, $i, j = 1, \cdots, m$. $\mathbf{\Sigma}_0 = \mathbf{\Gamma} \otimes \mathbf{B}_{ij}$, $\mathbf{\Gamma} = \mathbf{\Gamma}_0\mathbf{\Gamma}_0^\top$. $\mathbf{\Gamma}_0 = [\gamma_1, \cdots, \gamma_m]^\top$ is the sparsity pattern vector of $\mathbf{X}$ with the support indicates $\gamma_i \in \{0, 1\}$, $i = 1, \cdots, m$.

Specifically, we have

$$\mathbf{\Sigma}_0 = \begin{bmatrix} \gamma_1\gamma_1\mathbf{B}_{11} & \gamma_1\gamma_2\mathbf{B}_{12} & \cdots & \gamma_1\gamma_m\mathbf{B}_{1m} \\ \gamma_2\gamma_1\mathbf{B}_{21} & \gamma_2\gamma_2\mathbf{B}_{22} & \cdots & \gamma_2\gamma_m\mathbf{B}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_m\gamma_1\mathbf{B}_{m1} & \gamma_m\gamma_2\mathbf{B}_{m2} & \cdots & \gamma_m\gamma_m\mathbf{B}_{mm} \end{bmatrix}. \tag{5}$$

Fig.1 shows an example structure of the covariance matrix $\mathbf{\Sigma}_0$ of $\mathbf{x}$ with $m = 4, n = 6$. Since $\mathbf{x}_2 = \mathbf{0}$ and $\mathbf{x}_4 = \mathbf{0}$, the blocks associated with γ_2 and γ_4 become zeros. Thus, only four block matrices are non-zeros with the blocks of $\gamma_1\gamma_3\mathbf{B}_{13}$ and $\gamma_3\gamma_1\mathbf{B}_{31}$ are diagonal block matrices.

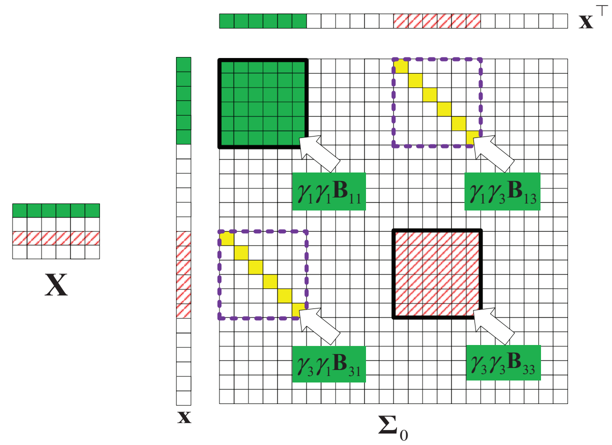


Fig. 1. An example structure of the covariance matrix $\mathbf{\Sigma}_0$ of $\mathbf{x}$, $\mathbf{x} = \text{vec}[\mathbf{X}^\top]$, $m = 4, n = 6$.

3. PROPOSED ALGORITHM

Using the Bayes rule, we have the posterior density of $\mathbf{x}$, which is also Gaussian,

$$p(\mathbf{x}|\mathbf{y}; \lambda, \gamma_i, \gamma_j, \mathbf{B}_{ij}, \forall i, j) \sim \mathcal{N}_x(\boldsymbol{\mu}_x, \mathbf{\Sigma}_x), \tag{6}$$

where the mean $\boldsymbol{\mu}_x$ and the covariance $\mathbf{\Sigma}_x$ are given by

$$\boldsymbol{\mu}_x = \frac{1}{\lambda}\mathbf{\Sigma}_x\mathbf{A}^\top\mathbf{y}, \tag{7}$$

$$\begin{aligned} \mathbf{\Sigma}_x &= (\mathbf{\Sigma}_0^{-1} + \frac{1}{\lambda}\mathbf{A}^\top\mathbf{A})^{-1} \\ &= \mathbf{\Sigma}_0 - \mathbf{\Sigma}_0\mathbf{A}^\top(\lambda\mathbf{I} + \mathbf{A}\mathbf{\Sigma}_0\mathbf{A}^\top)^{-1}\mathbf{A}\mathbf{\Sigma}_0. \end{aligned} \tag{8}$$

Given all the hyperparameters $\lambda, \gamma_i, \gamma_j, \mathbf{B}_{ij}, \forall i, j$, the maximum a posterior (MAP) estimate of $\mathbf{x}$ is given by

$$\begin{aligned} \hat{\mathbf{x}} = \text{vec}[\hat{\mathbf{X}}^\top] \triangleq \boldsymbol{\mu}_x &= (\lambda\mathbf{\Sigma}_0^{-1} + \mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}\mathbf{\Sigma}_0 \\ &= \mathbf{\Sigma}_0\mathbf{A}^\top(\lambda\mathbf{I} + \mathbf{A}\mathbf{\Sigma}_0\mathbf{A}^\top)^{-1}\mathbf{y}, \end{aligned} \tag{9}$$

where $\mathbf{\Sigma}_0$ is the block matrix given by (5) with most block matrices being zeros. Clearly, the sparsity of the blocks of $\hat{\mathbf{x}}$ is described by $\gamma_i\gamma_j, \forall i, j$. When $\gamma_k = 0$, the associated kth block in $\hat{\mathbf{x}}$ becomes zeros.

In order to obtain the estimation of $\mathbf{x}$ from (9), we need to obtain the hyperparameters firstly. Similar to the bSBL framework [1] which induces the temporal correlation in the prior density via the covariance matrices $\mathbf{B}_{ij}$, $i, j = 1, \cdots, m$, to avoid overfitting, we consider using a common positive definite matrix $\mathbf{B}$ to model all the covariance matrices $\mathbf{B}_{ij}$. Then, (5) turns into

$$\mathbf{\Sigma}_0 = \mathbf{\Gamma} \otimes \mathbf{B}, \tag{10}$$

where

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_1\gamma_1 & \gamma_1\gamma_2 & \cdots & \gamma_1\gamma_m \\ \gamma_2\gamma_1 & \gamma_2\gamma_2 & \cdots & \gamma_2\gamma_m \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_m\gamma_1 & \gamma_m\gamma_2 & \cdots & \gamma_m\gamma_m \end{bmatrix}. \tag{11}$$

In this regard, we use Bayesian strategy to marginalize over $\mathbf{x}$ and then maximize the resulting likelihood function with respect to $\mathbf{B}$ and $\mathbf{\Gamma}$, which are obtained from the following maximize problem

$$\max_{\mathbf{B}\in H^+, \mathbf{\Gamma}\geq 0} \int p(\mathbf{y}|\mathbf{x}; \mathbf{A}, \lambda) p(\mathbf{x}; \mathbf{\Gamma}, \mathbf{B}) d\mathbf{x}, \tag{12}$$

which is equivalent to minimizing the cost function

$$\mathcal{L}(\mathbf{\Gamma}, \mathbf{B}, \lambda) = \mathbf{y}^\top \mathbf{\Sigma}_y^{-1} \mathbf{y} + \log |\mathbf{\Sigma}_y|, \tag{13}$$

where H^+ denotes a set of $n \times n$ positive definite matrices.

$$\mathbf{\Sigma}_y = \mathbf{A}\mathbf{\Sigma}_0\mathbf{A}^\top + \lambda \mathbf{I}, \quad \mathbf{\Sigma}_0 = \mathbf{\Gamma} \otimes \mathbf{B}. \tag{14}$$

Here $\mathbf{\Sigma}_y$ can be interpreted as the covariance of $\mathbf{y}$ given $\mathbf{\Gamma}$ and $\mathbf{B}$.

Let $\mathbf{\Theta} = \{\mathbf{\Gamma}, \mathbf{B}, \lambda\}$, thus (13) becomes

$$\mathcal{L}(\mathbf{\Theta}) = \mathbf{y}^\top \mathbf{\Sigma}_y^{-1} \mathbf{y} + \log |\mathbf{\Sigma}_y|, \tag{15}$$

with $\mathbf{\Sigma}_y = \mathbf{A}\mathbf{\Sigma}_0\mathbf{A}^\top + \lambda \mathbf{I}$ and $\mathbf{\Sigma}_0 = \mathbf{\Gamma} \otimes \mathbf{B}$. We first treat $\mathbf{x}$ as hidden variables in the expectation maximization (EM) formulation proceeding and then maximize

$$\begin{aligned} \mathcal{Q}(\mathbf{\Theta}) =& E_{x|y;\mathbf{\Theta}^{(pre)}}[\log p(\mathbf{y}, \mathbf{x}; \mathbf{\Theta})] \\ =& E_{x|y;\mathbf{\Theta}^{(pre)}}[\log p(\mathbf{y}|\mathbf{x};\lambda)] \\ &+ E_{x|y;\mathbf{\Theta}^{(pre)}}[\log p(\mathbf{x}; \mathbf{\Gamma}, \mathbf{B})]. \end{aligned} \tag{16}$$

Here and in the sequel, notation $\mathbf{A}^{(pre)}$ denotes the estimated value of $\mathbf{A}$ in the last iteration.

To estimate $\mathbf{\Gamma}$ and $\mathbf{B}$, we assume $\mathbf{\Gamma} = \text{diag}(\gamma_1^2, \cdots, \gamma_m^2)$ where $\text{diag}(\cdot)$ denotes a diagonal matrix operator. Notice that the first term in (16) is unrelated to $\mathbf{\Gamma}$ and $\mathbf{B}$. So, we can simplify the $\mathcal{Q}$ function (16) to

$$\mathcal{Q}(\mathbf{\Gamma}, \mathbf{B}) = E_{x|y;\mathbf{\Theta}^{(pre)}}[\log p(\mathbf{x}; \mathbf{\Gamma}, \mathbf{B})], \tag{17}$$

then we have

$$\begin{aligned} \mathcal{Q}(\mathbf{\Gamma}, \mathbf{B}) \propto & -\frac{n}{2}\log(|\mathbf{\Gamma}|) - \frac{m}{2}\log(|\mathbf{B}|) \\ & - \frac{1}{2}\text{tr}[(\mathbf{\Gamma}^{-1} \otimes \mathbf{B}^{-1})(\mathbf{\Sigma}_x + \boldsymbol{\mu}_x\boldsymbol{\mu}_x^\top)]. \end{aligned} \tag{18}$$

Then, we plug $\boldsymbol{\mu}_x$ and $\mathbf{\Sigma}_x$ into (18). To estimate hyperparameters $\mathbf{\Theta}$, we get the gradients of (18) over γ_i^2 and $\mathbf{B}$, respectively, and then we obtain $\gamma_i^{(pre)}$, $i = 1, \cdots, m$, and $\mathbf{B}^{(pre)}$. Thus, we will get $\mathbf{\Gamma}^{(pre)}$. Using the same way, we can get $\lambda^{(pre)}$. Finally, we get $\mathbf{\Theta}^{(pre)}$.

In order to get an exact result of $\mathbf{\Theta}$, we employ standard upper bounds for solving (13) which known as a non-convex optimization problem leading to an EM-like algorithm. For the first and second terms of $\mathcal{L}(\mathbf{\Gamma}, \mathbf{B})$, we compute their bounds respectively.

Based on [9], for the first term in (13) we have

$$\mathbf{y}^\top \mathbf{\Sigma}_y^{-1} \mathbf{y} \leq \frac{1}{\lambda}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \mathbf{x}^\top \mathbf{\Sigma}_0^{-1} \mathbf{x}, \tag{19}$$

where equality is obtained when $\mathbf{x}$ satisfies (9).

For the second term,

$$\begin{aligned} \log|\mathbf{\Sigma}_y| &\equiv m\log|\mathbf{B}| + \log|\lambda \mathbf{A}^\top \mathbf{A} + \mathbf{\Sigma}_0^{-1}| \\ &\leq m\log|\mathbf{B}| + \text{tr}[\mathbf{B}^{-1}\nabla_{\mathbf{B}^{-1}}] + C, \end{aligned} \tag{20}$$

where for the second term $\log|\lambda \mathbf{A}^\top \mathbf{A} + \mathbf{\Sigma}_0^{-1}|$, we use a first-order approximation with a bias term C to approximate it with equality whenever the gradient satisfies

$$\nabla_{\mathbf{B}^{-1}} = \sum_{i=1}^{m} \mathbf{B} - \mathbf{B}\mathbf{A}_i^\top (\mathbf{A}\mathbf{\Sigma}_0\mathbf{A}^\top + \lambda \mathbf{I})^{-1} \mathbf{A}_i \mathbf{B}, \tag{21}$$

where $\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_m]$ and $\mathbf{A}_i \in \mathbb{R}^{p\times n}$. Finally using the upper bounds of (19), (20) and $\nabla_{\mathbf{B}^{-1}}$, we have the optimal $\mathbf{B}$ in closed form as

$$\begin{aligned} \mathbf{B}^{opt} &= \arg\min_{\mathbf{X}} \text{tr}[\mathbf{B}^{-1}(\mathbf{X}\mathbf{X}^\top + \nabla_{\mathbf{B}^{-1}})] + m\log|\mathbf{B}| \\ &= \frac{1}{m}[\hat{\mathbf{X}}\hat{\mathbf{X}}^\top + \nabla_{\mathbf{B}^{-1}}]. \end{aligned} \tag{22}$$

By starting with $\mathbf{B} = \mathbf{B}^{(pre)}$ and then iteratively computing (9), (21), and (22), we then have an estimate for $\mathbf{B}$, and a corresponding estimate for $\mathbf{x}$ given by (9). Here, we refer to this approach as L&S-bSBL algorithm which is outlined in Algorithm 1.

4. SIMULATION EXPERIMENTS

In this section, we present both synthetic data and real data results to compare the performance of the proposed L&S-bSBL algorithm with prior state-of-art TSBL [10], BARM [11] algorithms.

Algorithm 1 L&S-bSBL

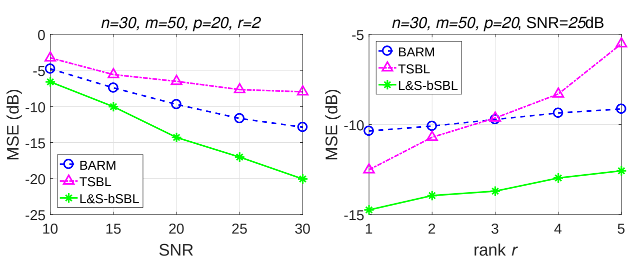
Input: $\mathbf{y}, \mathbf{A}$;
Output: $\mathbf{X}$;
procedure
Initialize
$iters = 0,\ \delta = 10^{-6}$, **max iteration number** $= 500$;
assume $\mathbf{\Gamma} = \mathrm{diag}(\gamma_1^2, \cdots, \gamma_m^2)$;
compute $\mathbf{\Gamma}, \mathbf{B}, \lambda$ from (18);
$\mathbf{\Sigma}_0 \leftarrow \mathbf{\Gamma} \otimes \mathbf{B}$;
while $\|\mathbf{X} - \hat{\mathbf{X}}\|_2^2 \geq \delta$ **do**
compute $\hat{\mathbf{X}}$ from (9);
compute $\nabla_{\mathbf{B}^{-1}}$ from (21);
compute $\mathbf{B}^{opt}$ from (22);
$iters = iters + 1$;
if $iters \geq 500$ **STOP**; **end if**
end while
Get the best $\mathbf{B}^{opt}$ and $\mathbf{X}$.
end procedure

4.1. Synthetic Experiments

Since L&S-bSBL is designed for spatially and temporally correlated data acquisition in WBAN, by assuming that the signal matrix satisfies the L&S structure, here we compare the performance of the algorithms for recovering the L&S matrices from their noisy linear measurements.

All the experiments consist of 100 independent trials. For the generation of $\mathbf{X}$, we first generate $\mathbf{X}_0 = \mathbf{M}_L\mathbf{M}_R$, with $\mathbf{M}_L \in \mathbb{R}^{m\times r}$ and $\mathbf{M}_R \in \mathbb{R}^{r\times n}$ ($n = 30$, $m = 50$, r is the rank of $\mathbf{X}$) [12]. Then we randomly generate the source matrix $\mathbf{X}$ with $K = r + 1$ nonzero rows. In each trial, the indexes of the sources are randomly chosen.

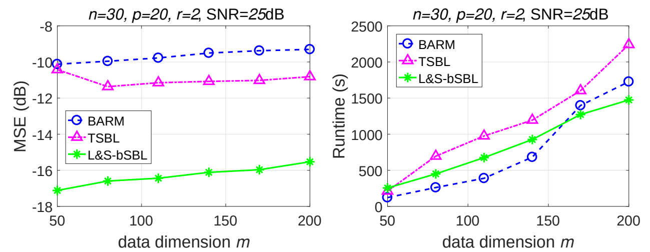
Fig.2(a) shows that with SNR increasing, the proposed L&S-bSBL algorithm has a better performance. And we see that the MSE gaps among these algorithms become bigger. Fig.2(b) shows that with rank r increasing, L&S-bSBL yields a better performance. The MSEs of both BARM and TSBL are almost the same when $r = 3$.



(a) (b)

Fig. 2. (a) MSE vs SNR. (b) MSE vs rank.

Next, we compare all the algorithms with different dimension m of data. Fig.3(a) shows that with m increasing, all of MSEs decrease when the compressive ratio becomes larger. In these cases, our proposed L&S-bSBL algorithm has a better performance. For instance, L&S-bSBL achieves at least 7dB reconstruction gain than the other algorithms and takes almost the same runtime when $m = 50$.



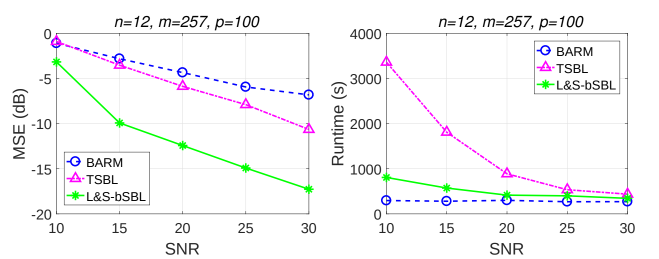
(a) (b)

Fig. 3. (a) MSE vs m. (b) Runtime vs m.

4.2. Experiments with Real Data

In real data experiments, we illustrate the potential of our algorithm by considering real-time ECG data compression in WBAN.

We assume the use of several body sensors to collect the 12-lead ECG data, compress and transmit the results to a FC, where the 12 signal vectors are recovered by using different algorithms above. Then, 100 continuous-time trials are run. In each trial, 12 length-257 contiguous sampled data vectors from the databases (I01[1]) are used. Thus, the data matrix is a 12×257 matrix. Fig.4 plots the performance of MSE versus SNR and runtime versus SNR, respectively. From Fig.4(a), we observe that L&S-bSBL outperforms all the other algorithms in terms of MSE, e.g., at SNR $= 25$dB, we note that L&S-bSBL achieves at least 7dB reconstruction gain than the other algorithms. Fig.4(b) shows that L&S-bSBL almost has the same runtime with BARM which is shorter than TSBL when SNR $= 20$dB.



(a) (b)

Fig. 4. MSE vs SNR and Runtime vs SNR for the recovery of real ECG data.

5. CONCLUSION

In this paper, we studied joint sparse reconstruction of spatially and temporally correlated data in WBAN, assuming that the signal matrix satisfies the L&S model. We proposed an algorithm based L&S structure to recover data using a bSBL-based algorithm. The proposed approach presented a better performance than other two methods through numerical results.

[1] Available at http://physionet.org/physiobank/database/incartdb.

6. REFERENCES

[1] Zhilin Zhang, Tzyy-Ping Jung, Scott Makeig, and Bhaskar D. Rao, "Compressed sensing of EEG for wireless telemonitoring with low energy consumption and inexpensive hardware," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 1, pp. 221–224, 2013.

[2] Zhilin Zhang, Tzyy-Ping Jung, Scott Makeig, and Bhaskar D. Rao, "Compressed sensing for energy-efficient wireless telemonitoring of noninvasive fetal ECG via block sparse Bayesian learning," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 2, pp. 300–309, 2013.

[3] Yonina C. Eldar and Moshe Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5302–5316, 2009.

[4] Namrata Vaswani, "Kalman filtered compressed sensing," in *Proceedings of 15th IEEE International Conference on Image Processing (ICIP)*, San Diego, CA, USA, Oct. 2008, pp. 893–896.

[5] Richard G. Baraniuk, Volkan Cevher, Marco F. Duarte, and Chinmay Hegde, "Model-based compressive sensing," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[6] Justin Ziniel and Philip Schniter, "Efficient high-dimensional inference in the multiple measurement vector problem," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 340–354, 2013.

[7] Yangqing Li, Wei Chen, Changchuan Yin, and Zhu Han, "Approximate message passing for sparse recovering of spatially and temporally correlated data," in *Proceedings of 2015 IEEE/CIC International Conference on Communications in China (ICCC)*, Shenzhen, China, Nov. 2015, pp. 1–6.

[8] Angshul Majumdar, Anupriya Gogna, and Rabab Ward, "A low-rank matrix recovery approach for energy efficient EEG acquisition for a wireless body area network," *Sensors*, vol. 14, no. 9, pp. 15729–15748, 2014.

[9] David P. Wipf, Bhaskar D. Rao, and Srikantan Nagarajan, "Latent variable Bayesian models for promoting sparsity," *IEEE Transactions on Information Theory*, vol. 57, no. 9, pp. 6236–6255, 2011.

[10] Zhilin Zhang and Bhaskar D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse bayesian learning," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 912–926, 2011.

[11] Bo Xin, Yizhou Wang, Wen Gao, and David Wipf, "Exploring algorithmic limits of matrix rank minimization under affine constraints," *IEEE Transactions on Signal Processing*, vol. 64, no. 19, pp. 4960–4974, 2016.

[12] Canyi Lu, Jinhui Tang, Shuicheng Yan, and Zhouchen Lin, "Generalized nonconvex nonsmooth low-rank minimization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 4130–4137.