2020 IEEE/CIC International Conference on Communications
in China (ICCC)

Fast Recovery of Low-Rank and Joint-Sparse Signals in Wireless Body Area Networks

Yanbin Zhang*, Longting Huang†, Yangqing Li*, Kai Zhang*, and Changchuan Yin*

*Beijing Laboratory of Advanced Information Networks,
Beijing Key Laboratory of Network System Architecture and Convergence,
Beijing University of Posts and Telecommunications, Beijing 100876, China
†School of Information Engineering, Wuhan University of Technology, Wuhan 430205, China
{iyanbin4, liyq, kaizhang, ccyin}@bupt.edu.cn, huanglt08@whut.edu.cn

***Abstract*—E-health monitoring signals collected from wireless body area networks (WBANs) usually have some highly correlated structures in a certain transform domain (e.g., discrete cosine transform (DCT)). We exploit these structures and propose a fast recovery algorithm for low-rank and joint-sparse (L&S) structured WBAN signal in the framework of compressed sensing (CS). By using a simultaneously L&S signal model, we employ the number of the bigger singular values and Bayesian learning which incorporates an L&S-inducing prior over the signal and the appropriate hyperpriors over all hyperparameters to recover the signal. Experiments show that the proposed algorithm has a superior performance to state-of-the-art algorithms.**

***Index Terms*—Internet of things, wireless body area network, sparse Bayesian learning, compressed sensing, low-rank and joint-sparse, fast recovery**

I. Introduction

Wireless Body Area Networks (WBANs) are becoming promising facilities in building e-health monitoring systems in Healthcare Internet of Things (IoT) [1]–[4]. Generally, WBAN includes lots of sensor nodes which are capable of sampling, processing, and communicating vital signals (e.g., heart rate, blood pressure, oxygen saturation, and activity) or environmental parameters (e.g., location, temperature, humidity, and light). Typically, sensor nodes are placed properly on the human body as tiny patches or hidden in users' clothes (so called wearable devices) allowing ubiquitous health monitoring in their native environment for extended periods of time. As studied in [5], WBAN signals usually contain certain structure, such as low-rank and joint-sparse.

How we could reduce the amount of signals being transmitted via WBANs for e-health monitoring to make wearables have better user experience (e.g., longer battery life, fewer computing resource requirements, lower latency) is badly needed. Fortunately, compressed sensing (CS)-based methods are proposed for this kind of signals [6]–[9]. Assuming that electroencephalogram (EEG) signal has a spatial block structure in a transform domain (discrete cosine transform (DCT) or wavelet transform), the authors in [10] demonstrate a good recovery performance when a block Sparse Bayesian Learning (bSBL) algorithm is used. The authors in [11] solved the multiple measurement vector (MMV) problem by using a Bayesian message passing algorithm when the amplitude coefficients of non-zero signals are temporal correlated.

Usually, the e-health monitoring signals in WBAN are spatio-temporal correlated. Therefore, we can assume that they have low-rank and joint-sparse structures [12]. However, most of the existing works only use one of the two structures in their recovery algorithms. Instead of considering low-rank individually [13] to recover the signals, the authors in [12] use structured spatial and temporal correlation to recover compressed signals and obtain a superior performance by assuming that the signals in wireless sensor networks (WSNs) have simultaneous low-rank and joint-sparse (L&S) structure.

In this paper, we proposed a fast bSBL-based data recovery algorithm, called fast L&S-bSBL, to recover compressed e-health monitoring signals in WBAN. The proposed algorithm not only incorporates the L&S-inducing prior over the signal and the appropriate hyperpriors over all hyperparameters but also uses the information of the principal components. Firstly, we give the analysis of the EEG signal, its DCT domain and singular values. Instead of computing the sparse position, by employing a Principal Components Analysis (PCA)-like method to get the sparse information, a fast bSBL-based algorithm is proposed. Then, by assuming that the covariance matrix of the signal is a diagonal matrix, the initial value of the hyperparameters is obtained. Finally, we use an expectation maximization (EM)-like algorithm to get the optimal reconstruction signal. Simulation results show that our proposed algorithm has better performance than the state-of-art algorithms through extensive experiments.

The rest of this work is organized as follows. Section II introduces the problem formation and signal model. In Section III, we provide the Bayesian learning inference for L&S-bSBL. Section IV presents the experimental results using two kinds of real-world application signals. Finally, Section V draws the conclusions.

Notation: Boldface upper-case letter denotes a matrix, and boldface lower-case letter denotes a column vector. rank($\mathbf{S}$)

This work was supported in part by Beijing Natural Science Foundation-Haidian Original Innovation Foundation (L192003), the National Natural Science Foundation of China under Grants 61629101, 61671086, and 61871041, in part by Beijing Laboratory Funding under Grant 2019BJLAB01, and in part by the 111 Project under Grant B17007.

denotes the rank of $\mathbf{S}$. $|\mathbf{S}|$ denotes the determinant of $\mathbf{S}$. $\|\mathbf{S}\|_0$ denotes the ℓ_0 norm of $\mathbf{S}$. $\|\mathbf{s}\|_2$ denotes the ℓ_2 norm of $\mathbf{s}$. $\mathbf{I}_n \in \mathbb{R}^{n\times n}$ denotes an n-dimensional identity matrix. $\mathbf{S} \otimes \mathbf{H}$ denotes the Kronecker product of the two matrices $\mathbf{S}$ and $\mathbf{H}$. $\text{vec}(\mathbf{S})$ denotes the vector of $\mathbf{S}$ which stacks the columns of $\mathbf{S}$ into a single vector. $\mathbf{S}^\top$ denotes the transpose of $\mathbf{S}$. $\text{tr}(\mathbf{S})$ denotes the trace of $\mathbf{S}$. $\text{diag}(\mathbf{s})$ denotes the matrix with $\mathbf{s}$ as the main diagonal.

II. Problem Formulation and Signal Model

This section shows the structures of WBAN signals firstly. Then problem formulation and signal model are presented. Fig. 1(a), (b) shows a segment of EEG signal in its time domain and DCT domain, respectively. The singular values of the EEG signal in time domain and in DCT domain are showed in Fig. 1(c). From Fig. 1(b), we can see that in the DCT domain most of the signal components are concentrated in the low frequency part, and the amplitude in other parts are almost zeros. From Fig. 1(c), we can observe that there are similarity for the singular values and the DCT of EEG signals. Specifically, the distribution of the singular values of the EEG signals are the same with that of its power spectrum. In this paper, we will use this feature to find the position of sparsity.

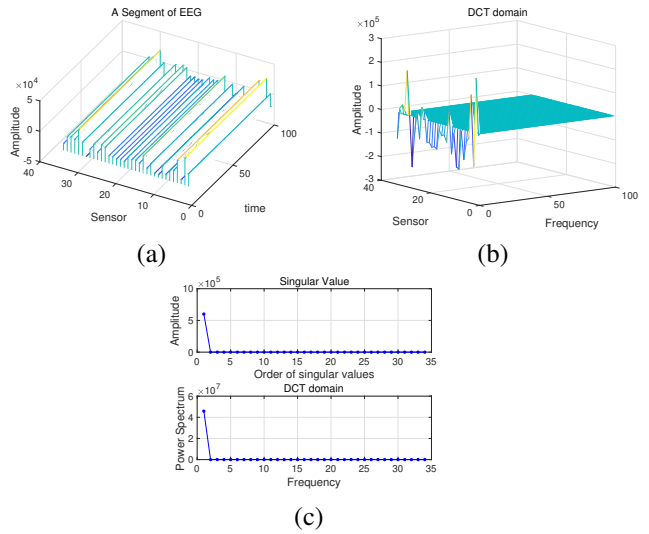


Fig. 1. (a) EEG. (b) its DCT. (c) its SVD vs. its DCT in 2D.

A. Problem Formulation

Due to the EEG signals have sparse and low-rank structures in some transform domain, we assume that it exhibits a simultaneously L&S structure. In this paper, we consider a typical e-health monitoring WBAN scenario and assume there are n time-varying signals being collected (we can also say n sensors which collect signals in synchronization), denoted as $\mathbf{F} = [\mathbf{f}_1, \cdots, \mathbf{f}_n] \in \mathbb{R}^{m\times n}$, where $\mathbf{f}_i \in \mathbb{R}^{m\times 1}, i \in \{1, 2, ..., n\}$ denotes the ith signal collected from the ith sensor with m samples. And then, signal $\mathbf{F}$ is encoded by $\mathbf{\Xi} \in \mathbb{R}^{p\times m}$ which is a linear mix matrix. After superimposed with noise $\mathbf{E} \in \mathbb{R}^{p\times n}$, the encoded $\mathbf{F}$ will be transmitted to a cloud, denoted as $\mathbf{X} \in \mathbb{R}^{p\times n}$. In addition, we assume that the signal $\mathbf{F}$ in DCT domain, denoted as $\mathbf{S} \in \mathbb{R}^{m\times n}$, is an approximately L&S matrix. Here, we can write it as $\mathbf{F} = \mathbf{\Psi}\mathbf{S}$, where $\mathbf{\Psi} \in \mathbb{R}^{m\times m}$ is the DCT matrix [12]. In the cloud, we use certain algorithm to recover the signals. As such, we have

$$\mathbf{X} = \mathbf{\Xi}\mathbf{F} + \mathbf{E} = \mathbf{\Phi}\mathbf{S} + \mathbf{E}, \tag{1}$$

where $\mathbf{\Phi} = \mathbf{\Xi}\mathbf{\Psi}$ is a dictionary matrix, $\mathbf{S}$ needs to be recovered in the cloud. We can obtain the e-health monitoring signals $\mathbf{F}$ by using the DCT matrix $\mathbf{\Psi}$.

B. Signal Model

By following the bSBL framework in [10], which recovers $\mathbf{S}$ by combining L&S-induced priors on the signal and appropriate priors on all hyperparameters, we can transform the problem into a block single measurement vector (SMV) [11] problem as follows

$$\mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{e}, \tag{2}$$

where $\mathbf{x} = \text{vec}(\mathbf{X}^\top) = [\mathbf{x}_1^\top, \cdots, \mathbf{x}_p^\top]^\top \in \mathbb{R}^{np\times 1}$, $\mathbf{H} = \mathbf{\Phi} \otimes \mathbf{I}_n \in \mathbb{R}^{np\times nm}$. $\mathbf{s} = \text{vec}(\mathbf{S}^\top) = [\mathbf{s}_1^\top, \cdots, \mathbf{s}_m^\top]^\top \in \mathbb{R}^{nm\times 1}$, $\mathbf{s}_i \in \mathbb{R}^{n\times 1}$ is the ith block of $\mathbf{s}$. l non-zero columns in $\mathbf{S}$ mean there are l non-zero blocks in $\mathbf{s}$. So, $\mathbf{s}$ is a block-sparse vector. $\mathbf{e} = \text{vec}(\mathbf{E}^\top) = [\mathbf{e}_1^\top, \cdots, \mathbf{e}_p^\top]^\top \in \mathbb{R}^{np\times 1}$. Here, we assume $\mathbf{e}_i, i \in \{1, \cdots, p\}$, is identically and independently distributed (i.i.d.) and follows Gaussian distribution, which is denoted by $p(\mathbf{e}_i) \sim \mathcal{N}(0, \lambda), \forall i$. For problem in (2), we define the Gaussian likelihood function as

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

and the Gaussian prior distribution of $\mathbf{s}$ as

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

where $\mathbf{B}_{ij} \in \mathbb{R}^{n\times n}$ is a covariance matrix that captures the correlation between $\mathbf{s}_i$ and $\mathbf{s}_j$ $(i, j = 1, \cdots, m)$. $\mathbf{\Sigma}_0 = \mathbf{\Gamma} \otimes \mathbf{B}_{ij}$, $\mathbf{\Gamma} = \boldsymbol{\gamma}\boldsymbol{\gamma}^\top$, where $\boldsymbol{\gamma} = [\gamma_1, \cdots, \gamma_m]^\top$ is a indicator vector with the support indicates $\gamma_i \in \{0, 1\}$, $(i = 1, \cdots, m)$.

Therefore, we can write $\mathbf{\Sigma}_0$ as

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

In this paper, we assume $\mathbf{S}$ is a low-rank matrix with joint-sparse in columns without losing generality. Fig. 2 shows an example of $\mathbf{S}$ and the structure of the covariance matrix $\mathbf{\Sigma}_0$ of $\mathbf{s}$ with $n = 4, m = 6$. In this example, we find that $\mathbf{s}_3 = \mathbf{0}$, $\mathbf{s}_4 = \mathbf{0}$, $\mathbf{s}_5 = \mathbf{0}$ and $\mathbf{s}_6 = \mathbf{0}$, so $\gamma_3 = \gamma_4 = \gamma_5 = \gamma_6 = 0$. Thus, only $\gamma_1\gamma_1\mathbf{B}_{11}, \gamma_1\gamma_2\mathbf{B}_{12}, \gamma_2\gamma_1\mathbf{B}_{21}$ and $\gamma_3\gamma_3\mathbf{B}_{22}$ are non-zeros with $\gamma_1\gamma_2\mathbf{B}_{12}$ and $\gamma_2\gamma_1\mathbf{B}_{21}$ are almost diagonal matrices.

III. Proposed Algorithm

In this section, the proposed algorithm is given, namely L&S-bSBL, to recover the e-health monitoring signals $\mathbf{S}$ from the measurement matrix $\mathbf{X}$ at the cloud, based on the model in Section II.

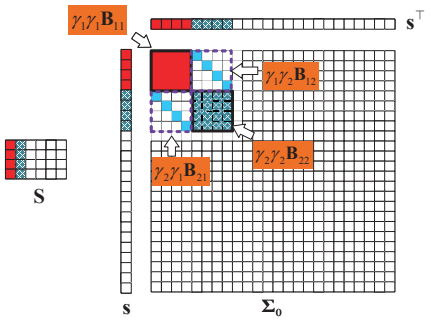


Fig. 2. An example of $\mathbf{S}$ and the structure of the covariance matrix $\mathbf{\Sigma}_0$ of $\mathbf{s}$.

A. Estimation of $\mathbf{S}$

We can obtain the Gaussian posterior density of $\mathbf{s}$ by using Bayesian rule

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

where $\boldsymbol{\mu}_{\mathbf{s}}$ and $\mathbf{\Sigma}_{\mathbf{s}}$ are given by [14], [15]

$$\boldsymbol{\mu}_{\mathbf{s}} = \frac{1}{\lambda}\mathbf{\Sigma}_{\mathbf{s}}\mathbf{H}^\top\mathbf{x}, \tag{7}$$

$$\mathbf{\Sigma}_{\mathbf{s}} = \mathbf{\Sigma}_0 - \mathbf{\Sigma}_0\mathbf{H}^\top(\lambda\mathbf{I}+\mathbf{H}\mathbf{\Sigma}_0\mathbf{H}^\top)^{-1}\mathbf{H}\mathbf{\Sigma}_0. \tag{8}$$

Therefore, we can obtain the maximum a posterior (MAP) estimation of $\mathbf{s}$ as

$$\hat{\mathbf{s}} = \mathrm{vec}(\hat{\mathbf{S}}^\top) \triangleq \boldsymbol{\mu}_{\mathbf{s}} = \mathbf{\Sigma}_0\mathbf{H}^\top(\lambda\mathbf{I}+\mathbf{H}\mathbf{\Sigma}_0\mathbf{H}^\top)^{-1}\mathbf{x}, \tag{9}$$

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

B. Hyperparameters Updating

We need to obtain the hyperparameters $\lambda,\gamma_i,\gamma_j,\mathbf{B}_{ij},\forall i,j$ before we estimate $\mathbf{s}$ by using (9). Here, we use a common positive definite matrix $\mathbf{B}$ to avoid overfitting, which is used to model the matrices $\mathbf{B}_{ij}$ $(i;j=1,\cdots,m)$. This is similar to [10] in that induces $\mathbf{B}$ to instead of all the source covariance $\mathbf{B}_{ij}$. Then, (5) turns into

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

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

$$\max_{\mathbf{B}\in\mathbf{M}^+,\mathbf{\Gamma}\geq 0}\int p(\mathbf{x}|\mathbf{s};\mathbf{H},\lambda)p(\mathbf{s};\mathbf{\Gamma},\mathbf{B})d\mathbf{s}. \tag{11}$$

As we know, $\max p(\mathbf{x};\lambda,\mathbf{\Gamma},\mathbf{B})$ is equal to minimizing $-2\log p(\mathbf{x};\lambda,\mathbf{\Gamma},\mathbf{B})$, where the cost function is

$$\mathcal{L}(\lambda,\mathbf{\Gamma},\mathbf{B}) = \mathbf{x}^\top\mathbf{\Sigma}_{\mathbf{x}}^{-1}\mathbf{x} + \log|\mathbf{\Sigma}_{\mathbf{x}}|, \tag{12}$$

where $\mathbf{\Sigma}_{\mathbf{x}} = \mathbf{H}\mathbf{\Sigma}_0\mathbf{H}^\top + \lambda\mathbf{I}$, and $\mathbf{\Sigma}_{\mathbf{x}}$ denotes the covariance of $\mathbf{x}$.

For simplicity, we define $\mathbf{\Theta} = \{\lambda,\mathbf{\Gamma},\mathbf{B}\}$, rewrite (12) as

$$\mathcal{L}(\mathbf{\Theta}) = \mathbf{x}^\top\mathbf{\Sigma}_{\mathbf{x}}^{-1}\mathbf{x} + \log|\mathbf{\Sigma}_{\mathbf{x}}|. \tag{13}$$

Here, $\mathbf{s}$ is treated as hidden variables in the EM formulation proceeding and then to maximize

$$\begin{aligned}\mathcal{Q}(\mathbf{\Theta}) = &E_{\mathbf{s}|\mathbf{x};\mathbf{\Theta}^{(pre)}}\left[\log p(\mathbf{x}|\mathbf{s};\lambda)\right]\\ &+ E_{\mathbf{s}|\mathbf{x};\mathbf{\Theta}^{(pre)}}\left[\log p(\mathbf{s};\mathbf{\Gamma},\mathbf{B})\right],\end{aligned} \tag{14}$$

where $\mathbf{\Theta}^{(pre)}$ denotes the hyperparameters which is estimated in the previous iteration.

To estimate λ, we simplify the $\mathcal{Q}$ function in (14) as

$$\begin{aligned}\mathcal{Q}(\lambda) =&E_{\mathbf{s}|\mathbf{x};\mathbf{\Theta}^{(pre)}}\left[\log p(\mathbf{x}|\mathbf{s};\lambda)\right]\\ \propto&-\frac{pn}{2}\log\lambda\\ &-\frac{1}{2\lambda}\left[\|\mathbf{x}-\mathbf{H}\boldsymbol{\mu}_{\mathbf{s}}\|_2^2+\lambda^{(pre)}[mn-\mathrm{tr}(\mathbf{\Sigma}_{\mathbf{s}}\mathbf{\Sigma}_0^{-1})]\right],\end{aligned} \tag{15}$$

where $\lambda^{(pre)}$ denotes the estimation of λ in the previous iteration. Using the learning rule, we can obtain λ when we calculate the derivative of (15) over λ and set it equal to zero

$$\lambda \leftarrow \frac{\|\mathbf{x}-\mathbf{H}\boldsymbol{\mu}_{\mathbf{s}}\|_2^2+\lambda^{(pre)}\left[mn-\mathrm{tr}(\mathbf{\Sigma}_{\mathbf{s}}\mathbf{\Sigma}_0^{-1})\right]}{pn}. \tag{16}$$

To estimate $\mathbf{\Gamma}$ and $\mathbf{B}$, we assume $\mathbf{\Gamma}_0 = \mathrm{diag}(\gamma_1^2,\cdots,\gamma_m^2)$ and will find that only the second term in (14) is related with $\mathbf{\Gamma}$ and $\mathbf{B}$. So, we can simplify the $\mathcal{Q}$ function in (14) as

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

We have

$$\begin{aligned}\mathcal{Q}(\mathbf{\Gamma}_0,\mathbf{B}) \propto &-\frac{n}{2}\log(|\mathbf{\Gamma}_0|)-\frac{m}{2}\log(|\mathbf{B}|)\\ &-\frac{1}{2}\mathrm{tr}\left[(\mathbf{\Gamma}_0^{-1}\otimes\mathbf{B}^{-1})(\mathbf{\Sigma}_{\mathbf{s}}+\boldsymbol{\mu}_{\mathbf{s}}\boldsymbol{\mu}_{\mathbf{s}}^\top)\right].\end{aligned} \tag{18}$$

Then, we plug $\boldsymbol{\mu}_{\mathbf{s}}$ and $\mathbf{\Sigma}_{\mathbf{s}}$ into (18). To estimate hyperparameters $\mathbf{\Theta}$, we need to obtain the gradients of (18) with respect of $\mathbf{B}$, and then we obtain $\mathbf{B}^{(pre)}$.

Here, we use the feature mentioned at the beginning of Section II to obtain γ_i $(i=1,\cdots,m)$. Specifically, we obtain m singular values of $\mathbf{X}$ by Singular Value Decomposition (SVD) and rank the singular values in descending order. According to the ranked singular values, we can obtain γ_i $(i=1,\cdots,m)$. The γ_i corresponding to the larger singular values[1] are 1, otherwise 0. In particular, for a certain type of signals, we only need to analyze some parts of the signals in advance to obtain the larger singular values of the signals, e.g., EEG signals, Temperature signals in WBAN. Thus, we will get $\mathbf{\Gamma}$. Finally, we get $\mathbf{\Theta}^{(pre)}$.

We use the low-rank property of matrix to get an accurate B. As we know, (12) is a non-convex optimization problem, for solving (12), we use an EM-like algorithm to obtain standard upper bounds [16]. For the two terms in the right hand side of (12), we are able to calculate their bounds separately.

[1]The larger singular values are defined as the singular values which are greater than 1% of the largest singular value.

For the first term of (12), based on [17],

$$\mathbf{x}^\top \boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\mathbf{x} \leq \frac{1}{\lambda}\|\mathbf{x}-\mathbf{H}\mathbf{s}\|_2^2 + \mathbf{s}^\top \boldsymbol{\Sigma}_0^{-1}\mathbf{s}, \tag{19}$$

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

For the second term of (12), we have

$$\begin{aligned}\log|\boldsymbol{\Sigma}_{\mathbf{x}}| &\equiv m\log|\mathbf{B}| + \log|\lambda\mathbf{H}^\top\mathbf{H} + \boldsymbol{\Sigma}_0^{-1}| \\ &\leq m\log|\mathbf{B}| + \mathrm{tr}\left(\mathbf{B}^{-1}\nabla_{\mathbf{B}^{-1}}\right) + C,\end{aligned} \tag{20}$$

where C is a bias term. When the gradient satisfies

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

the equality of (20) will be obtained. From (19), (20) and (21), we get a closed form of $\mathbf{B}$ as follows

$$\begin{aligned}\mathbf{B} = \arg\min_{\mathbf{B}} \mathrm{tr}\left[(\mathbf{B}^{(pre)})^{-1}\left(\mathbf{S}\mathbf{S}^\top + \nabla_{(\mathbf{B}^{(pre)})^{-1}}\right)\right] \\ + m\log|\mathbf{B}^{(pre)}| = \frac{1}{m}\left(\hat{\mathbf{S}}\hat{\mathbf{S}}^\top + \nabla_{(\mathbf{B}^{(pre)})^{-1}}\right).\end{aligned} \tag{22}$$

Then, we obtain $\boldsymbol{\Theta}^{(pre)} = \{\boldsymbol{\Gamma}, \mathbf{B}, \lambda\}$.

C. Calculation of $\mathbf{S}^{opt}$

By starting with $\mathbf{B} = \mathbf{B}^{(pre)}$ and following the learning rules to compute (9), (21) and (22) iteratively, we will get an estimation for $\mathbf{B}$, and an estimation for $\mathbf{s}$ through (9).

$$\mathbf{S}^{opt} \leftarrow \mathbf{s}. \tag{23}$$

In summary, we outline the flowchart of the proposed fast L&S-bSBL in Algorithm 1.

Algorithm 1 Fast L&S-bSBL Recovery Algorithm

Source data analyzation
calculate singular values of $\mathbf{S}$ by SVD;
obtain γ_i, $i = 1, \cdots, m$ by using singular values;
Initialize
$\mathbf{x}, \mathbf{H}, iters = 0, \delta = 10^{-6}$, **max iters** $= 500$;
assume $\boldsymbol{\Gamma}_0 = \mathrm{diag}(\gamma_1^2, \cdots, \gamma_m^2)$;
Set $\lambda, \mathbf{B}$ by $\lambda = 10^{-10}$, $\mathbf{B} = ones(n, n)$;
$\boldsymbol{\Sigma}_0 \leftarrow \boldsymbol{\Gamma}_0 \otimes \mathbf{B}$;
while $\|\mathbf{S} - \hat{\mathbf{S}}\|_2^2 \geq \delta$ **do**
  $\hat{\mathbf{S}} \leftarrow$ (9), (23): compute (9), (23) to obtain $\hat{\mathbf{S}}$;
  $\lambda \leftarrow$ (16): compute (16) to obtain λ;
  $\nabla_{\mathbf{B}^{-1}} \leftarrow$ (21): compute (21) to obtain $\nabla_{\mathbf{B}^{-1}}$;
  $\mathbf{B} \leftarrow$ (22): compute(22) to obtain $\mathbf{B}$;
  $iters = iters + 1$;
  if $iters \geq 500$
  STOP while;
  end if
end while
Get the optimal $\mathbf{S}$ from (23).

IV. Numerical Results

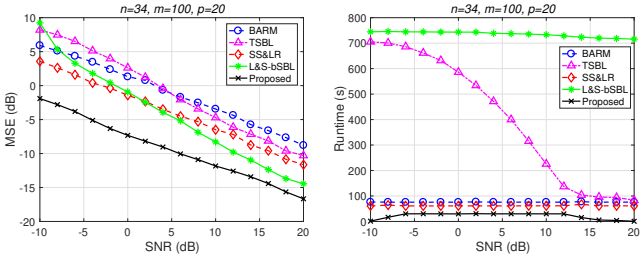
In this section, we conduct simulation experiments by using two real-world signal datasets to evaluate the performance of the proposed fast L&S-bSBL algorithm. For comparison purpose, we consider four state-of-art algorithms, BARM [15], TSBL [14], SS&LR [18] and our recently proposed L&S-bSBL [5].

A. Experiment with EEG signals

In this part, EEG datasets (subject1.zip[2]) are used. We do 100 independent continuous-time trials. In each trial, we use 34 length-100 contiguous sampled data vectors from the EEG datasets. Thus, the source matrix is a 100×34 matrix.

As mentioned earlier, we use random Gaussian matrix as the dictionary matrix $\mathbf{H}$ and DCT matrix as the matrix $\boldsymbol{\Psi}$.

Fig. 3 shows the performance of mean-squared error (MSE) and runtime with different signal-to-noise ratios (SNRs), respectively. Here $p = 20, n = 34, m = 100$, so the ratio of compression is $(p \times n)/(n \times m) = 680/3400$. The MSE is defined as $10 * \log(\|\hat{\mathbf{S}} - \mathbf{S}\|_2^2/\|\mathbf{S}\|_2^2)$, where $\hat{\mathbf{S}}$ is an estimation of $\mathbf{S}$. From Fig. 3(a), we observe that the proposed algorithm outperforms all the other algorithms in terms of MSE. When SNR $= 20$ dB, the proposed algorithm achieves at least 7 dB reconstruction gain compared to BARM and TSBL algorithms, 5 dB to SS&LR and 2 dB to L&S-bSBL. Fig. 3(b) shows that the proposed algorithm takes less runtime than the other four algorithms. From the results in Fig. 3, we can see that the proposed algorithm has a better performance not only in recovery MSE but also in computing time compared to the state-of-art algorithms.



(a) (b)

Fig. 3. (a) MSE vs. SNR. (b) Runtime vs. SNR for the recovery of EEG signals.

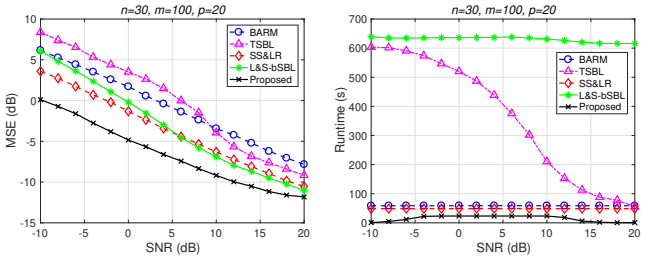
B. Experiment with temperature signal

In this part, we use temperature signals of Lausanne Urban Canopy Experiment (LUCE) datasets[3]. We do 100 independent continuous-time trials. In each trial, we use 30 length-100 contiguous sampled data vectors from the LUCE datasets. Thus, the source matrix is a 100×30 matrix.

Fig. 4 shows the performance of MSE and runtime with different SNRs, respectively. Here $p = 20, n = 30, m = 100$, so the ratio of compression is $(p \times n)/(n \times m) = 600/3000$. From Fig. 4(a), we observe that the proposed algorithm

[2]Available at https://mmspg.epfl.ch/BCI_datasets.
[3]Available at http://lcav.epfl.ch/cms/lang/en/1122pid/86035.

outperforms all the other algorithms in terms of MSE. When at SNR $= 0$ dB, the proposed algorithm achieves at least 7 dB reconstruction gain compared to BARM and TSBL algorithms, 5 dB to L&S-bSBL and 4 dB to SS&LR. Fig. 4(b) shows that the proposed algorithm take less runtime than the other four algorithms. From the results in Fig. 4, we can see that the proposed algorithm has a better performance not only in recovery MSE but also in computing time than the other four algorithms.



(a) (b)

Fig. 4. (a) MSE vs. SNR. (b) Runtime vs. SNR for the recovery of EEG data.

V. Conclusion

In this work, by taking advantage of the low-rank and joint-sparse structures of the spatially and temporally correlated WBAN signals, we proposed a fast L&S structure based algorithm to recover WBAN signals. The numerical results show that the proposed algorithm performs better than existed algorithms with synthetic and real EEG signal and temperature signal.

For future work, we may be able to take advantage of the properties of the WBAN signal structure in higher dimensions (more than two) to design an effective Bayesian-learning-based recovery algorithm; for example, tensor models could be considered to compress and recover multiple heterogeneous signals.

References

[1] Mingzhe Chen, Ursula Challita, Walid Saad, Changchuan Yin, and Mérouane Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3039–3071, 2019.

[2] Alireza Ghasempour, "Internet of things in smart grid: Architecture, applications, services, key technologies, and challenges," *Inventions*, vol. 4, no. 1, pp. 22, 2019.

[3] Alireza Ghasempour, "Optimum number of aggregators based on power consumption, cost, and network lifetime in advanced metering infrastructure architecture for smart grid internet of things," in *2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 2016, pp. 295–296.

[4] Alireza Ghasempour and Todd K Moon, "Optimizing the number of collectors in machine-to-machine advanced metering infrastructure architecture for internet of things-based smart grid," in *2016 IEEE Green Technologies Conference (GreenTech)*. IEEE, 2016, pp. 51–55.

[5] Yanbin Zhang, Yangqing Li, Changchuan Yin, and Kesen He, "Low-rank and joint-sparse signal recovery for spatially and temporally correlated data using sparse bayesian learning," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 4719–4723.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] Geoffrey McLachlan and Thriyambakam Krishnan, *The EM algorithm and extensions*, vol. 382, John Wiley & Sons, 2007.

[17] 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.

[18] Wei Chen, "Simultaneously sparse and low-rank matrix reconstruction via nonconvex and nonseparable regularization," *IEEE Transactions on Signal Processing*, vol. 66, no. 20, pp. 5313–5323, 2018.