IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 6, 2021, accepted January 11, 2021, date of publication January 14, 2021, date of current version January 27, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3051644

Inverse Solution of Steady-State Responses Based on Sparse Bayesian Learning

MINGWEN QU[1], TINGTING CHEN[1,3], SHIQI LU[2], JIANLING HU[1], (Member, IEEE), JIAJUN WANG[1], AND NAN HU[1]

[1]School of Electronics and Information Engineering, Soochow University, Suzhou 215006, China
[2]Emergency Department, The First Affiliated Hospital of Soochow University, Suzhou 215006, China
[3]Unisoc Corporation, Nanjing 210018, China

Corresponding author: Nan Hu (hunan@suda.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61601316, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20171249, and in part by the Suzhou Science and Technology Planning Project under Grant SYS201521 and Grant SYS2019029.

ABSTRACT Steady-state responses (SSRs), evoked by various patterns of periodic stimuli, comprise an important category of evoked potentials. To explore the neural generators of SSRs, a unified framework solving the inverse problem for a single subject or integrating multiple subjects is indispensable. Inspired by the phenomenon that the oscillation frequency of an SSR follows that of the periodic stimulus, we consider the problem of source localization for SSRs using the Fourier components at the stimulation frequency instead of directly using the waveform in this paper. The multi-channel electroencephalogram (EEG) Fourier components at the stimulation frequency is shown to equal multiplying the lead field matrix (LFM) by a complex-valued vector that contains the amplitudes and phases of sources in the cortex, contaminated by spontaneous EEG and electrical noise. This complex-valued inverse problem is further solved in the framework of sparse Bayesian learning, where the non-stationarity of spontaneous EEG among epochs is considered, and the joint sparsity of complex-valued source component vectors is modeled and utilized to improve the source localization performance. Expectation-maximization (EM) is employed to give the ultimate SSR source localization algorithm. By the proposed method, not only a single subject's SSR source localization can be achieved, but also the common locations of a certain type of SSR integrating multiple subjects can be given, even when the electrode layout or number of electrodes varies among subjects. The validity and superior performance of the proposed method was verified by simulations compared with other methods. Real SSR stimulation/recording experiments were also performed, where the electric generators of 40-Hz auditory steady-state responses (ASSRs) by various stimulation patterns were investigated.

INDEX TERMS Electroencephalogram (EEG), joint sparsity, steady-state response (SSR), source localization, sparse Bayesian learning.

I. INTRODUCTION

Electroencephalogram (EEG) measurements collected on scalp can reflect the neuronal electric activities inside the brain and thus has numerous applications in clinical diagnosis and prognosis, psychological assessment, cognitive research, etc. Compared with spontaneous EEG signals, evoked responses yielded superior performances in many applications, e.g. they showed higher specificity and sensitivity in assessing cerebral function for comatose patients [1]. Evoked responses are aroused by a certain pattern of stimulation and can be categorized into transient evoked potentials and steady-state responses (SSRs). The presentation frequency of SSR stimulation is much higher than that of transient evoked potential stimulation. Being different from transient evoked potentials with a set of peaks and troughs in the waveform as latent components, the SSR has the form of sinusoidal signal, which has the same oscillation frequency as periodic stimulus [2]. It was conjectured that SSRs may be generated by the superimposition of transient evoked responses by each cycle of stimulation, while some inconsistent relationship is still shown to exist between transient evoked responses and SSRs [3].

It can be observed that SSRs focus their energy on their corresponding stimulus frequency, and they are believed to have higher signal-to-noise ratio (SNR) [4]. Hence, it is easy

The associate editor coordinating the review of this manuscript and approving it for publication was Hasan S. Mir.

to evaluate and analyze the SSRs in the frequency domain by Fourier analysis [2]. The SSRs commonly include auditory steady-state response (ASSR) [5], steady-state visual evoked potential (SSVEP) [6], and steady-state somatosensory evoked potential (SSSEP) [7], and they have been widely studied nowadays. E.g., ASSRs have been used in hearing test, anesthesia monitoring, neurological assessment and coma prognosis [8]. In brain computer interfaces (BCI), SSVEPs were used to build a ''speller'' device that transfers the neural activities to control signals [9], [10], and to solve the problem that prolonged visual attention is prone to induce visual fatigue, SSSEPs are also applied in BCI nowadays as they rely on the somatosensory nervous system instead [7]. Due to the wide usage of SSRs, there is a persistent interest in investigating the neural generators of them, and for patients it is also useful to study the relationship between the disease and location changes of neural generators of SSRs.

After decades of study, many methods for EEG or magnetoencephalogram (MEG) source localization, also known as solving the inverse problem, have been developed. According to the used source models, they can be classified into two categories: methods based on the equivalent current dipole (ECD) model [11] and methods based on the distributed source model [12]. The application of distributed source model has been prevalent in recent years, as it does not need any preliminary assumption or estimation of the number of the active sources like the ECD model and usually achieves better localization accuracy and resolution. In the distributed source model, the suspected dipoles was defined at the predetermined locations represented by dense groups of volumetric grids within the head volume, leading to a severe underdetermined problem [12]. To make the ill-posed problem solvable, regularization schemes were introduced in various linear distributed solutions. ℓ_2-norm was used in minimum norm estimation (MNE) [13], and the current density estimated by MNE was further used and standardized by its variance in standardized low resolution brain electromagnetic tomography (sLORETA) [14]. In [15], an inverse method named vector based spatial-temporal analysis using a ℓ_1-minimum-norm (VESTAL) was proposed. Linearly constrained minimum variance (LCMV) beamforming [16], though not specifically classified into distributed source model based methods, can still use a spatial filter in the distributed source model. In fact, solving the inverse problem, by minimizing the model matching error along with the regularization term, can be unified into the framework of Bayesian inference [17], where appropriate assumptions of prior distributions and practical parameter inference methods have been paid much attention. Friston *et al.* showed the general idea to apply the empirical Bayesian framework in processing neuroimaging data [18]. The advantage of Bayesian inference is that the regularization is easily modeled by a prior distribution and the unknown variables can be inferred from the data. Costa *et al.* developed a hierarchical Bayesian model whose robustness is enhanced by an introduced multivariate Bernoulli Laplacian structured sparsity prior [19]. Block sparse Bayesian learning considering the inner block structure was in introduced in [20]. Chen *et al.* applied the multi-trial priors and structural information in a Bayesian model to enhance the extraction of somatosensory evoked potential [21].

Some aforementioned source localization methods have also been applied in locating the neural generators of SSRs. We take the ASSR inverse methods for example. Herdman *et al.* treated the real and imaginary response components at the stimulation frequency as independent samples and used brain electric source analysis (BESA) for ASSR source localization [22]. Reyes *et al.* used PET-independent low-resolution electroencephalographic tomography (LORETA) as well as PET-weighted LORETA and MNE to study the source locations of 40-Hz ASSR [23]. Poulsen *et al.* used the BESA to study the age-related changes in neural generators of ASSR [24]. The ASSR source localization was also studied by applying LCMV method in MEG data in [25]. In [26], Farahani *et al.* developed a method employing independent component analysis (ICA) and fitting an ECD model to the projection weights of the independent components, to reconstruct the ASSR sources. Generally, these inverse methods applied in SSR inverse problem were originally designed for transient evoked potentials, and the input data are the EEG waveforms after narrowband filtering and epoch-wise averaging. However, directly applying these methods in SSR source localization may confront some performance degradation: the SSR sources have the sinusoidal form with unknown and probably distinct amplitudes and phases, hence the peaks in waveforms of sources may not synchronize, leading to probably imbalanced estimates of sources at a specified point in time.

Considering the property of SSRs that they have focused energy around the stimulus frequency, we can solve the inverse problem utilizing the multi-channel Fourier components at the stimulus frequency. The analysis on frequency domain has already been applied in inverse solutions of brain rhythms [27]–[29]. In [27], LORETA was employed to investigate the electric source distribution differences of the theta band, the alpha band, and the beta band in the cortex during psychometrically matched verbal and spatial cognitive tasks. LORETA was also used to study the correlation between Alzheimer's disease severity and electric generators of brain rhythms [28]. In [29], an improved version of sLORETA named swLORETA, was used to explore the neural generators of brain rhythms related to motor imagery. These methods may have the drawbacks that follow. Firstly, the existing inverse methods were only roughly applied in the multi-channel Fourier components, while a rigorous signal model was not established. Secondly, the cross-spectral matrices at a certain frequency estimated using multi-epoch data were essential in these methods, while the non-stationarity of spontaneous EEG among epochs would deteriorate the cross-spectral matrix estimation. Lastly, the comprehensive inverse solution integrating multiple

subjects was only obtained by simple averaging, without the consideration of individual difference.

In this paper, we consider the problem of source localization for SSRs using the Fourier components at the stimulation frequency instead of directly using the waveform. The signal model of multi-channel EEG Fourier components is given, based on which a Bayesian model for the inverse problem is established. In this established Bayesian model, the non-stationarity of spontaneous EEG among epochs is considered, and a sparse support vector expressing the joint sparsity of complex-valued source component vectors among epochs is also involved. The parameters in this Bayesian model is estimated by a data-driven iterative procedure realized by expectation maximization (EM), leading to a sparse Bayesian learning (SBL) algorithm. In the outputs of the proposed method, the sparse support vector gives the final ASSR source locations integrating multiple epochs, and the automatic derivation of variances of spontaneous EEG plus electrical noise in all epochs eliminates the adverse effect of EEG non-stationarity. Furthermore, the proposed method can easily give the comprehensive SSR inverse solution integrating multiple subjects, even when the electrode layout or number of electrodes varies among subjects. Both simulation comparisons and real SSR stimulation/recording experiments are carried out to illustrate the efficiency of the proposed method.

The reminder of this paper is organized as follows: Section II describes the signal model of the Fourier components of data recorded by multiple electrodes. Section III gives the hierarchical Bayesian model corresponding to the inverse problem to be solved. Section IV provides the sparse Bayesian learning method in SSR source localization. Section V describes the experiments carried out for performance verification. Section VI gives the conclusion.

II. SIGNAL MODEL

When a certain pattern of periodic stimulation is exerted on a subject, several sinusoidal dipole sources would be evoked at particular positions in the subject's cortex. These sinusoidal source signals have the identical frequency, which equals the stimulus frequency f_0, while they may have distinct amplitudes and phases. If K sinusoidal source signals $s_1(t)$, $s_2(t)$, $\ldots$, $s_K(t)$ are simultaneously generated and keep stationary in a certain period of time, they can be represented as

$$s_k(t) = a_k \sin(2\pi f_0 t + \varphi_k), k = 1, 2, \ldots, K, \quad (1)$$

where a_k and φ_k represent the amplitude and the phase of $s_k(t)$, respectively. They can be further represented as

$$s_k(t) = \frac{a_k}{2} e^{j\varphi_k} e^{j2\pi f_0 t} - \frac{a_k}{2} e^{-j\varphi_k} e^{-j2\pi f_0 t}, k = 1, 2, \ldots, K. \quad (2)$$

For long stimulation and EEG recording, we can assume that the amplitude of $s_k(t)$ remain constant in a short while defined as an epoch, and may be variant or invariant among epochs. Under this assumption, if the total recording is divided into L epochs, the kth source signal embedded in the lth epoch can be represented by

$$s_{k,l}(t) = \tilde{s}_{k,l} e^{j2\pi f_0 t} + \hat{s}_{k,l} e^{-j2\pi f_0 t}, \quad (3)$$

where $\tilde{s}_{k,l}$ and $\hat{s}_{k,l}$ represent the complex weights corresponding to the components $e^{j2\pi f_0 t}$ and $e^{-j2\pi f_0 t}$, respectively. It can be found that $\tilde{s}_{k,l}$ contains the information of amplitude and phase of the kth sinusoidal source signal in the lth epoch.

EEG measurements are generated when neuronal activities are transmitted through a real head model to the surface of scalp, where the head model can usually be divided into several layers with different electrical conductivities, e.g. the classical four layers including the cortex, cerebrospinal fluid (CSF), skull and scalp. Such a head model, mapping the neuronal activities in the cortex to scalp electrodes, commonly leads to a lead field matrix (LFM) $\mathbf{L}_0$, constructed by finite element method (FEM) or boundary element method (BEM). If M-electrode layout is used in EEG recording, when a certain type of periodic stimulation is given to evoke the corresponding SSR, the EEG measurements at these electrodes in the lth epoch can be arranged in a vector form and represented as

$$\mathbf{x}_l(t) = \sum_{k=1}^{K} \mathbf{L}_0(\mathbf{r}_k) \boldsymbol{\psi}_k s_{k,l}(t) + \boldsymbol{\xi}_l(t) + \mathbf{n}_l(t). \quad (4)$$

where $\mathbf{L}_0(\mathbf{r}_k)$ is an $M \times 3$ LFM for the kth source located at $\mathbf{r}_k$, the 3-element vector $\boldsymbol{\psi}_k$ represents the current orientation of the kth dipole source, and $\boldsymbol{\xi}_l(t)$ and $\mathbf{n}_l(t)$ represent the simultaneously recorded spontaneous EEG and electrical noise in the lth epoch, respectively. In fact, $\boldsymbol{\xi}_l(t)$ can also be denoted by mapping the neuronal activities to scalp surface through LFM. However, we aim at localizing the SSR related neuronal activities in this study, hence the spontaneous EEG will be deemed as interference.

When only the neuronal activities in the cortex are concerned, the model of EEG measurements in (4) can be further simplified, as the sources are mainly generated by the coherent activation of thousands of pyramidal neurons which are vertical to the cortex surface [12]. Now that $\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \cdots, \boldsymbol{\psi}_K$ are given by $\boldsymbol{\psi}(\mathbf{r}_1), \boldsymbol{\psi}(\mathbf{r}_2), \cdots, \boldsymbol{\psi}(\mathbf{r}_K)$, respectively, (4) can be transformed to

$$\mathbf{x}_l(t) = \mathbf{L}\mathbf{s}_l(t) + \boldsymbol{\xi}_l(t) + \mathbf{n}_l(t). \quad (5)$$

where the column of newly constructed LFM $\mathbf{L}$ corresponding to the kth source is given by $\mathbf{L}[:, k] = \mathbf{L}(\mathbf{r}_k) \triangleq \mathbf{L}_0(\mathbf{r}_k)\boldsymbol{\psi}(\mathbf{r}_k)$, the signal waveform vector is formed by $\mathbf{s}_l(t) \triangleq [s_{1,l}(t), s_{2,l}(t), \cdots, s_{K,l}(t)]^{\mathrm{T}}$, and $(\cdot)^{\mathrm{T}}$ denotes the transpose. Substituting (3) into (5), we have

$$\mathbf{x}_l(t) = \mathbf{L}\tilde{\mathbf{s}}_l e^{j2\pi f_0 t} + \mathbf{L}\hat{\mathbf{s}}_l e^{-j2\pi f_0 t} + \boldsymbol{\xi}_l(t) + \mathbf{n}_l(t), \quad (6)$$

where $\tilde{\mathbf{s}}_l = [\tilde{s}_{1,l}, \tilde{s}_{2,l}, \cdots, \tilde{s}_{K,l}]^{\mathrm{T}}$ and $\hat{\mathbf{s}}_l = [\hat{s}_{1,l}, \hat{s}_{2,l}, \cdots, \hat{s}_{K,l}]^{\mathrm{T}}$.

If some time-frequency analysis, e.g. fast Fourier transform (FFT), is applied to the EEG records in the lth epoch, the Fourier components at f_0 can be derived, represented by

$$\tilde{\mathbf{x}}_l = \mathbf{L}\tilde{\mathbf{s}}_l + \tilde{\boldsymbol{\xi}}_l + \tilde{\mathbf{n}}_l, l = 1,\ 2,\ \cdots,\ L, \tag{7}$$

where $\tilde{\boldsymbol{\xi}}_l$ and $\tilde{\mathbf{n}}_l$ denote the Fourier components of $\boldsymbol{\xi}_l(t)$ and $\mathbf{n}_l(t)$ at f_0, respectively. In this study, we assume that, $\tilde{\mathbf{s}}_l$ can be variant or invariant among epochs, and $\tilde{\mathbf{n}}_l$ always obeys a complex Gaussian distribution with a fixed variance σ_n^2, i.e. $\tilde{\mathbf{n}}_l \sim \mathcal{CN}\left(\mathbf{0},\ \sigma_n^2\mathbf{I}\right)$, where $\mathbf{I}$ is identity matrix. The spontaneous EEG component vector $\tilde{\boldsymbol{\xi}}_l$ is also assumed to be complex Gaussian distributed. According to the reality that the spontaneous EEG is non-stationary, the corresponding variance σ_l^2 may be variant among epochs, i.e.

$$\tilde{\boldsymbol{\xi}}_l \sim \mathcal{CN}\left(\mathbf{0},\ \sigma_l^2\mathbf{I}\right), l = 1,\ 2,\ \cdots,\ L, \tag{8}$$

where σ_l^2 is a positive random variable. For example, a Gamma distributed σ_l^2 makes $\tilde{\boldsymbol{\xi}}_l$ K-distributed.

III. HIERARCHICAL BAYESIAN MODEL

In this paper, we aim at solving the inverse problem for SSRs modeled in (7) in the framework of SBL. Recall that in (7), $\tilde{\boldsymbol{\xi}}_l$ and $\tilde{\mathbf{n}}_l$ are assumed to be complex Gaussian distributed with random variance and fixed variance, respectively. It is easy to find that $\mathbf{e}_l \triangleq \tilde{\boldsymbol{\xi}}_l + \tilde{\mathbf{n}}_l \sim \mathcal{CN}\left(\mathbf{0},\ \left(\sigma_l^2 + \sigma_n^2\right)\mathbf{I}\right)$, as $\tilde{\boldsymbol{\xi}}_l$ and $\tilde{\mathbf{n}}_l$ are independent of each other. If $\gamma_l \triangleq \sigma_l^2 + \sigma_n^2$ and $\tilde{\mathbf{s}}_l$ are given, the conditional distribution of $\tilde{\mathbf{x}}_l$ can be represented as

$$\begin{aligned} p\left(\tilde{\mathbf{x}}_l \left| \mathbf{L}\tilde{\mathbf{s}}_l, \gamma_l\right.\right) &= \mathcal{CN}\left(\tilde{\mathbf{x}}_l \left| \mathbf{L}\tilde{\mathbf{s}}_l, \gamma_l\mathbf{I}\right.\right) \\ &= \frac{1}{\pi^M \gamma_l^M} \exp\left[-\frac{\left(\tilde{\mathbf{x}}_l - \mathbf{L}\tilde{\mathbf{s}}_l\right)^H \left(\tilde{\mathbf{x}}_l - \mathbf{L}\tilde{\mathbf{s}}_l\right)}{\gamma_l}\right], \end{aligned} \tag{9}$$

where $(\cdot)^{\mathrm{H}}$ denotes the conjugate transpose.

In real applications, uncertainty in the LFM $\mathbf{L}$ usually could not be neglected. Sometimes this uncertainty comes from imprecise solution of the forward problem, while at other times it may be caused by individual differences among subjects' head models. To consider this issue, we assume that there exists some perturbations in $\mathbf{L}$, defined as $\Delta\mathbf{L}$, and then we have

$$\begin{aligned} \tilde{\mathbf{x}}_l &= (\mathbf{L} + \Delta\mathbf{L})\,\tilde{\mathbf{s}}_l + \mathbf{e}_l \\ &= \mathbf{L}\tilde{\mathbf{s}}_l + \left(\Delta\mathbf{L}\tilde{\mathbf{s}}_l + \mathbf{e}_l\right) = \mathbf{L}\tilde{\mathbf{s}}_l + \mathbf{e}'_l, l = 1,\ 2,\ \cdots,\ L, \end{aligned} \tag{10}$$

where $\mathbf{e}'_l \triangleq \Delta\mathbf{L}\tilde{\mathbf{s}}_l + \mathbf{e}_l$. If the entries of the perturbation matrix $\Delta\mathbf{L}$ are independent of each other and have identical Gaussian distributions: $\Delta\mathbf{L}[m,k] \sim \mathcal{N}\left(0,\ \sigma_{\mathrm{LP}}^2\right)$, it is easily to derive that $\Delta\mathbf{L}\tilde{\mathbf{s}}_l \sim \mathcal{CN}\left(\mathbf{0},\ \sum_{k=1}^{K}\left|\tilde{s}_{k,\,l}\right|^2\sigma_{\mathrm{LP}}^2\mathbf{I}\right)$ and hence $\mathbf{e}'_l \sim \mathcal{CN}\left(\mathbf{0},\ \left(\sum_{k=1}^{K}\left|\tilde{s}_{k,\,l}\right|^2\sigma_{\mathrm{LP}}^2 + \sigma_l^2 + \sigma_n^2\right)\mathbf{I}\right)$. Recalling the conditional distribution of $\tilde{\mathbf{x}}_l$ in (9), we find that the uncertainty in the LFM can be absorbed into γ_l, which will be estimated in our inverse solution method to be proposed.

The data derived from L epochs can be arranged as $\mathbf{X} = \left[\tilde{\mathbf{x}}_1,\ \tilde{\mathbf{x}}_2,\ \cdots,\ \tilde{\mathbf{x}}_L\right]$. As $\tilde{\mathbf{x}}_m \left|\tilde{\mathbf{s}}_m, \gamma_m\right.$ is independent of $\tilde{\mathbf{x}}_n \left|\tilde{\mathbf{s}}_n, \gamma_n\right.$ as long as $m \neq n$, the conditional distribution of $\mathbf{X}$ can be given by

$$\begin{aligned} &p\left(\mathbf{X}\left|\mathbf{LS}, \boldsymbol{\gamma}\right.\right) \\ &= \prod_{l=1}^{L} p\left(\tilde{\mathbf{x}}_l \left|\mathbf{L}\tilde{\mathbf{s}}_l, \gamma_l\right.\right) = \prod_{l=1}^{L}\mathcal{CN}\left(\tilde{\mathbf{x}}_l \left|\mathbf{L}\tilde{\mathbf{s}}_l, \gamma_l\mathbf{I}\right.\right) \\ &= \prod_{l=1}^{L}\frac{1}{\pi^M\gamma_l^M}\exp\left[-\frac{\left(\tilde{\mathbf{x}}_l - \mathbf{L}\tilde{\mathbf{s}}_l\right)^H\left(\tilde{\mathbf{x}}_l - \mathbf{L}\tilde{\mathbf{s}}_l\right)}{\gamma_l}\right], \end{aligned} \tag{11}$$

where $\mathbf{S} = \left[\tilde{\mathbf{s}}_1,\ \tilde{\mathbf{s}}_2,\ \cdots,\ \tilde{\mathbf{s}}_L\right]$ and $\boldsymbol{\gamma} = [\gamma_1,\ \gamma_2,\ \cdots,\ \gamma_L]^{\mathrm{T}}$.

As the locations of SSR sources are unknown at present, the whole range of cortex should be inspected. If the whole cortex is divided into N tessellation elements located at $\tilde{\mathbf{r}}_1$, $\tilde{\mathbf{r}}_2,\ \cdots,\ \tilde{\mathbf{r}}_N$ spanning the whole cortex, an extended LFM $\tilde{\mathbf{L}} = \left[\mathbf{L}_0\left(\tilde{\mathbf{r}}_1\right)\psi\left(\tilde{\mathbf{r}}_1\right),\ \mathbf{L}_0\left(\tilde{\mathbf{r}}_2\right)\psi\left(\tilde{\mathbf{r}}_2\right),\ \cdots,\ \mathbf{L}_0\left(\tilde{\mathbf{r}}_N\right)\psi\left(\tilde{\mathbf{r}}_N\right)\right]$ can be built. Generally, only a small portion of cortex would be activated during the cyclic stimulation for evoking SSRs. In this context, we have $\mathbf{LS} = \tilde{\mathbf{L}}\mathbf{W}$, where an $N \times L$ row-sparse matrix $\mathbf{W}$ is defined. $\mathbf{W}[n,:] = \mathbf{S}[k,:]$ if $\tilde{\mathbf{r}}_n = \mathbf{r}_k$ where $\mathbf{r}_k \in \{\mathbf{r}_1,\ \mathbf{r}_2,\ \cdots,\ \mathbf{r}_K\}$, otherwise $\mathbf{W}[n,:]$ is filled with all zeros. As long as $\tilde{\mathbf{r}}_1,\ \tilde{\mathbf{r}}_2,\ \cdots,\ \tilde{\mathbf{r}}_N$ are predefined, $\tilde{\mathbf{L}}$ can be determined beforehand. Now the conditional distribution of $\mathbf{X}$ in (11) can be transformed as

$$\begin{aligned} &p\left(\mathbf{X}\left|\mathbf{W}, \boldsymbol{\gamma}\right.\right) \\ &= \prod_{l=1}^{L} p\left(\tilde{\mathbf{x}}_l \left|\mathbf{w}_l, \gamma_l\right.\right) = \prod_{l=1}^{L}\mathcal{CN}\left(\tilde{\mathbf{x}}_l \left|\tilde{\mathbf{L}}\mathbf{w}_l, \gamma_l\mathbf{I}\right.\right) \\ &= \prod_{l=1}^{L}\frac{1}{\pi^M\gamma_l^M}\exp\left[-\frac{\left(\tilde{\mathbf{x}}_l - \tilde{\mathbf{L}}\mathbf{w}_l\right)^H\left(\tilde{\mathbf{x}}_l - \tilde{\mathbf{L}}\mathbf{w}_l\right)}{\gamma_l}\right], \end{aligned} \tag{12}$$

where $\mathbf{w}_l$ is the lth column of $\mathbf{W}$.

The row-sparsity of $\mathbf{W}$ is due to the fact that the locations of sources for a certain type of SSR would not change greatly among epochs. In this paper, the joint sparsity among columns of $\mathbf{W}$ is manifested by modeling that the columns of $\mathbf{W}$ obey the same complex Gaussian distribution, i.e.

$$p\left(\mathbf{w}_l\left|\boldsymbol{\alpha}\right.\right) = \mathcal{CN}\left(\mathbf{w}_l\left|\mathbf{0},\ diag\left(\boldsymbol{\alpha}\right)\right.\right), l = 1,\ 2,\ \cdots,\ L, \tag{13}$$

where the $N \times 1$ vector $\boldsymbol{\alpha}$ is the common variance vector. A larger entry of $\boldsymbol{\alpha}$ implies that some source is more probable to locate at the corresponding position. As $\mathbf{w}_l\left|\boldsymbol{\alpha}\right.$ is independent of $\mathbf{w}_{l'}\left|\boldsymbol{\alpha}\right.$ as long as $l \neq l'$, the conditional distribution of $\mathbf{W}$ is represented as

$$\begin{aligned} p\left(\mathbf{W}\left|\boldsymbol{\alpha}\right.\right) &= \prod_{l=1}^{L} p\left(\mathbf{w}_l\left|\boldsymbol{\alpha}\right.\right) = \prod_{l=1}^{L}\mathcal{CN}\left(\mathbf{w}_l\left|\mathbf{0},\ \mathrm{diag}\left(\boldsymbol{\alpha}\right)\right.\right) \\ &= \prod_{l=1}^{L}\prod_{n=1}^{N}\frac{1}{\pi\alpha_n}\exp\left(-\frac{\left|w_{n,l}\right|^2}{\alpha_n}\right), \end{aligned} \tag{14}$$

where $w_{n,l}$ is the (n,l)th entry of $\mathbf{W}$. In reality, the subjects' states may vary with time, so the identical distribution shared

by all columns of $\mathbf{W}$ seems not conforming to realistic scenarios. However, the distribution provided by (13) is only the prior distribution of $\mathbf{w}_l$, while a more reliable statistical modeling of $\mathbf{w}_l$ is given by its posterior distribution, which will be discussed below. It can be noticed that, in fact $\boldsymbol{\alpha}$ plays the role of controlling the row sparsity of $\mathbf{W}$, and hence if $\boldsymbol{\alpha}$ is correctly obtained, the locations of sources can also be given.

In (12) and (14), $\boldsymbol{\gamma}$ and $\boldsymbol{\alpha}$ are deemed as random variable vectors. In the following analyses, they will be treated as parameters instead, and then $p(\mathbf{X}|\mathbf{W}, \boldsymbol{\gamma})$ in (12) and $p(\mathbf{W}|\boldsymbol{\alpha})$ in (14) are replaced by $p(\mathbf{X}|\mathbf{W};\ \boldsymbol{\gamma})$ and $p(\mathbf{W};\ \boldsymbol{\alpha})$, respectively. Derivation of these parameters will be the objectives of Bayesian inference. Here we choose to estimate the sum of σ_l^2 and σ_n^2, i.e. γ_l, rather than each one of them.

A. BAYESIAN NETWORK FOR A SINGLE SUBJECT

For a single subject, the joint distribution $p(\mathbf{X}, \mathbf{W};\ \boldsymbol{\alpha}, \boldsymbol{\gamma})$ can be represented by the product of the conditional distributions $p(\mathbf{X}|\mathbf{W};\ \boldsymbol{\gamma})$ and $p(\mathbf{W};\ \boldsymbol{\alpha})$, i.e.

$$p(\mathbf{X}, \mathbf{W};\ \boldsymbol{\alpha}, \boldsymbol{\gamma}) = p(\mathbf{X}|\mathbf{W};\ \boldsymbol{\gamma})\, p(\mathbf{W};\ \boldsymbol{\alpha}), \tag{15}$$

where $\mathbf{W}$ would be deemed as a latent variable matrix in the Bayesian inference. The directed acyclic graph (DAG) for this hierarchical Bayesian model is shown in Fig. 1. This simple single-subject model was considered in our previous work [30]. However, the method proposed in [30] cannot handle any multi-subject data, and had unsatisfactory performance in many scenarios as the multithreaded iterative parameter update procedure described next in this paper was not employed.

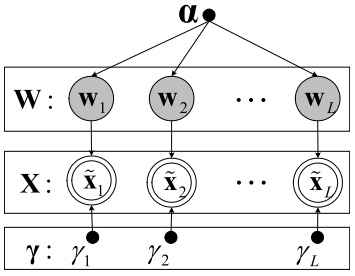

FIGURE 1. DAG of the Bayesian model for a single subject.

B. BAYESIAN NETWORK FOR MULTIPLE SUBJECTS

To exploit the common neural generators of a certain type of SSR, stimulation experiments were usually carried out on multiple subjects. The SBL provides us a convenient way to recover such common source locations by combining data of multiple subjects. If H subjects participated in the SSR experiments, their data matrices derived at f_0 are $\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H$, where $\mathbf{X}_h$ is the data matrix of the hth subject and it consists of L_h epochs. Similarly, $\mathbf{W} \triangleq [\mathbf{W}_1,\ \mathbf{W}_2,\ \cdots,\ \mathbf{W}_H]$ and $\boldsymbol{\gamma} \triangleq \left[\boldsymbol{\gamma}_1^{\mathrm{T}},\ \boldsymbol{\gamma}_2^{\mathrm{T}},\ \cdots,\ \boldsymbol{\gamma}_H^{\mathrm{T}}\right]^{\mathrm{T}}$ are also defined. The common location information for SSR among subjects can be modeled by the row-sparsity of $\mathbf{W}$, which is shared among $\mathbf{W}_1, \mathbf{W}_2, \cdots$, and $\mathbf{W}_H$, i.e.

$$p(\mathbf{W};\ \boldsymbol{\alpha}) = \prod_{h=1}^{H} p(\mathbf{W}_h;\ \boldsymbol{\alpha}) = \prod_{h=1}^{H}\prod_{l_h=1}^{L_h} p\left(\mathbf{w}_{h,l_h};\ \boldsymbol{\alpha}\right)$$

$$= \prod_{h=1}^{H}\prod_{l_h=1}^{L_h} \mathcal{CN}\left(\mathbf{w}_{h,l_h}|\mathbf{0},\ diag(\boldsymbol{\alpha})\right) \tag{16}$$

where $\boldsymbol{\alpha}$ is used to control the same row-sparse support for $\mathbf{W}_1, \mathbf{W}_2, \cdots$, and $\mathbf{W}_H$. Besides, it is easy to find that $p(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H|\mathbf{W};\ \boldsymbol{\gamma}) = \prod_{h=1}^{H} p\left(\mathbf{X}_h|\mathbf{W}_h;\ \boldsymbol{\gamma}_h\right)$, as $\mathbf{X}_{h_1}|\mathbf{W}_{h_1};\ \boldsymbol{\gamma}_{h_1}$ is independent of $\mathbf{X}_{h_2}|\mathbf{W}_{h_2};\ \boldsymbol{\gamma}_{h_2}$ if $h_1 \neq h_2$. We can easily find $p(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H, \mathbf{W};\ \boldsymbol{\alpha}, \boldsymbol{\gamma}) = p(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H|\mathbf{W};\ \boldsymbol{\gamma})\, p(\mathbf{W};\ \boldsymbol{\alpha})$, which has the similar formulation as (16) for the single-subject model. In fact, the single-subject model is the special case of multi-subject model when $H = 1$. This viewpoint can also be discovered form DAG for the hierarchical Bayesian model of multiple subjects shown in Fig. 2. This model can be further extended to the scenario that various electrode layouts or even various number of electrodes are probably applied for different subjects. When the same mesh is used to sample the cortex, the subjects can build their own extended LFMs: $\tilde{\mathbf{L}}_1, \tilde{\mathbf{L}}_2, \cdots$, and $\tilde{\mathbf{L}}_H$, where the dimension of $\tilde{\mathbf{L}}_h$ is $M_h \times N$. Now we have

$$p(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H|\mathbf{W};\boldsymbol{\gamma})$$
$$= \prod_{h=1}^{H}\prod_{l_h=1}^{L_h} p\left(\tilde{\mathbf{x}}_{h,\,l_h}\left|\mathbf{w}_{h,\,l_h};\,\gamma_{h,\,l_h}\right.\right)$$

$$= \prod_{h=1}^{H}\prod_{l_h=1}^{L_h} \mathcal{CN}\left(\tilde{\mathbf{x}}_{h,\,l_h}\left|\tilde{\mathbf{L}}_h\mathbf{w}_{h,\,l_h}, \gamma_{h,\,l_h}\mathbf{I}\right.\right), \tag{17}$$

where $\tilde{\mathbf{x}}_{h,\,l_h}$ and $\mathbf{w}_{h,\,l_h}$ are the l_hth column of $\mathbf{X}_h$ and $\mathbf{W}_h$, respectively, and $\gamma_{h,\,l_h}$ is the l_hth entry of $\boldsymbol{\gamma}_h$.

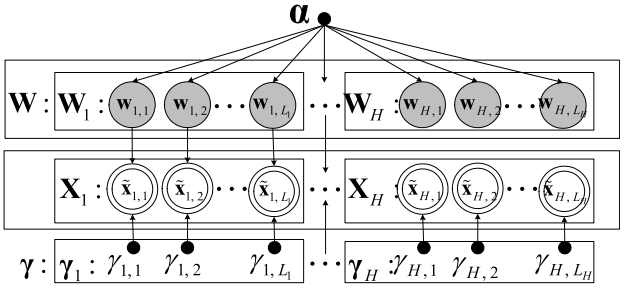

FIGURE 2. DAG of the Bayesian model for multiple subjects.

IV. SPARSE BAYESIAN LEARNING

Based on the Bayesian model established above, expectation-maximization (EM) is used to derive the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ in this paper. To give a unified solution, we consider the scenario involving multiple subjects (H subjects), and for single-subject scenario we only need to let $H = 1$.

The posterior distribution of latent variable matrices $p(\mathbf{W}|\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_H; \boldsymbol{\alpha}, \boldsymbol{\gamma})$ is the premise of applying EM. Based on Bayes' rule, when the SSR recordings collected from the hth subject have L_h epochs, we have

$$p\left(\mathbf{w}_{h,l_h}\left|\tilde{\mathbf{x}}_{h,l_h}; \boldsymbol{\alpha}, \boldsymbol{\gamma}\right.\right) = \mathcal{CN}\left(\mathbf{w}_{h,l_h}\left|\boldsymbol{\mu}_{h,l_h}, \boldsymbol{\Sigma}_{h,l_h}\right.\right), \quad h = 1, 2, \cdots, H, l_h = 1, 2, \cdots, L_h, \tag{18}$$

where $\boldsymbol{\mu}_{h,l_h} = \gamma_{h,l_h}^{-1}\boldsymbol{\Sigma}_{h,l_h}\tilde{\mathbf{L}}_h^{\mathrm{H}}\tilde{\mathbf{x}}_{h,l_h}$, $\boldsymbol{\Sigma}_{h,l_h} = (\boldsymbol{\Lambda}^{-1} + \gamma_{h,l_h}^{-1}\tilde{\mathbf{L}}_h^{\mathrm{H}}\tilde{\mathbf{L}}_h)^{-1}$, and $\boldsymbol{\Lambda} = diag(\boldsymbol{\alpha})$. As $\mathbf{w}_{h,l_h}\left|\tilde{\mathbf{x}}_{h,l_h}; \boldsymbol{\alpha}, \boldsymbol{\gamma}\right.$ is independent of $\mathbf{w}_{h',l_h'}\left|\tilde{\mathbf{x}}_{h',l_h'}; \boldsymbol{\alpha}, \boldsymbol{\gamma}\right.$ when $h \neq h'$ or $l_h \neq l_h'$, we further have

$$p(\mathbf{W}|\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_H; \boldsymbol{\alpha}, \boldsymbol{\gamma}) = \prod_{h=1}^{H}\prod_{l_h=1}^{L_h} p\left(\mathbf{w}_{h,l_h}\left|\tilde{\mathbf{x}}_{h,l_h}; \boldsymbol{\alpha}, \boldsymbol{\gamma}\right.\right). \tag{19}$$

According to the established Bayesian model, estimation of the parameter vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ should be given by

$$(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\gamma}}) = \underset{\boldsymbol{\alpha},\boldsymbol{\gamma}}{\operatorname{argmax}} \log p(\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_H; \boldsymbol{\alpha}, \boldsymbol{\gamma}). \tag{20}$$

Instead of such a direct optimization, in the procedure of EM, $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ are learned iteratively by maximizing a tight lower bound of $\log p(\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_H; \boldsymbol{\alpha}, \boldsymbol{\gamma})$. If terms irrelevant to $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ are ignored, this tight lower bound corresponds to a function $Q(\boldsymbol{\alpha}, \boldsymbol{\gamma})$, given by the expectation of $\log p(\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_H, \mathbf{W}; \boldsymbol{\alpha}, \boldsymbol{\gamma})$ under $p(\mathbf{W}|\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_H; \boldsymbol{\alpha}, \boldsymbol{\gamma})$, shown by the proof in Appendix A. The objective function $Q(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ is further expressed as

$$\begin{aligned} Q(\boldsymbol{\alpha}, \boldsymbol{\gamma}) &= E_{p(\mathbf{W}|\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_H; \boldsymbol{\alpha}, \boldsymbol{\gamma})} \left[\log p(\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_H, \mathbf{W}; \boldsymbol{\alpha}, \boldsymbol{\gamma})\right] \\ &= Q(\boldsymbol{\alpha}) + Q(\boldsymbol{\gamma}), \end{aligned} \tag{21}$$

where we have (22) and (23), as shown at the bottom of the next page, and $(\cdot)^*$, $Re(\cdot)$ and $tr(\cdot)$ denote the conjugate operation, the real part and the trace of a matrix, respectively. Now the task turns to maximizing $Q(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ with respect to the corresponding parameter vector.

It can be found that $Q(\boldsymbol{\alpha})$ and $Q(\boldsymbol{\gamma})$ are independent of $\boldsymbol{\gamma}$ and $\boldsymbol{\alpha}$, respectively. By making the derivative of $Q(\boldsymbol{\alpha})$ with respect to $\boldsymbol{\alpha}$ equal 0, we have the solution of $\boldsymbol{\alpha}$ given by

$$\alpha_n = \frac{\sum_{h=1}^{H}\sum_{l_h=1}^{L_h} \boldsymbol{\Sigma}_{h,l_h}[n,n] + \boldsymbol{\mu}_{h,l_h}[n]\,\boldsymbol{\mu}_{h,l_h}[n]^*}{\sum_{h=1}^{H} L_h}, \quad n = 1, 2, \cdots, N. \tag{24}$$

Similarly, let the derivative of $Q(\boldsymbol{\gamma})$ with respect to $\boldsymbol{\gamma}$ equal 0, and then we have (25), as shown at the bottom of the next page.

In the proposed SBL algorithm, $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ are updated iteratively. When some initialization is set, the convergence of iteration is guaranteed, as proved by Appendix A. Though $Q(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ can grow larger and converge ultimately, its corresponding lower bound of $\log p(\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_H; \boldsymbol{\alpha}, \boldsymbol{\gamma})$ cannot be guaranteed to be close enough to the maximum of $\log p(\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_H; \boldsymbol{\alpha}, \boldsymbol{\gamma})$ using an arbitrary initialization. To address this issue, we make the iteration procedure multithreaded, where each thread has different parameter initialization. The ultimate estimation of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ are given by the thread that achieves the largest $Q(\boldsymbol{\alpha}, \boldsymbol{\gamma})$. In this paper, the initializations of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ are given by

$$\boldsymbol{\alpha}_{\text{init}} = \sum_{h=1}^{H}\sum_{l_h=1}^{L_h} \boldsymbol{\rho}_{h,l_h}^* \odot \boldsymbol{\rho}_{h,l_h} / \sum_{h=1}^{H} L_h, \tag{26}$$

$$\gamma_{\text{init},\,h,l_h} = \left(\tilde{\mathbf{x}}_{h,l_h} - \tilde{\mathbf{L}}_h\boldsymbol{\rho}_{h,l_h}\right)^{\mathrm{H}}\left(\tilde{\mathbf{x}}_{h,l_h} - \tilde{\mathbf{L}}_h\boldsymbol{\rho}_{h,l_h}\right)/M_h, \tag{27}$$

where $\odot$ denotes Hadamard product, $\|\cdot\|_{\mathrm{F}}$ denotes Frobenius norm, and $\boldsymbol{\rho}_{h,l_h}$ is given by one of $D = 4$ choices:

$$\begin{cases} \boldsymbol{\rho}_{h,l_h}^1 = \sqrt{N}\tilde{\mathbf{L}}_h^{\mathrm{H}}\tilde{\mathbf{x}}_{h,l_h} / \left\|\tilde{\mathbf{L}}_h^{\mathrm{H}}\tilde{\mathbf{L}}_h\right\|_{\mathrm{F}}, & \text{or} \\ \boldsymbol{\rho}_{h,l_h}^2 = \tilde{\mathbf{L}}_h^{\mathrm{H}}\tilde{\mathbf{x}}_{h,l_h} / \left\|\tilde{\mathbf{L}}_h^{\mathrm{H}}\right\|_{\mathrm{F}}, & \text{or} \\ \boldsymbol{\rho}_{h,l_h}^3 = \tilde{\mathbf{L}}_h^{\mathrm{H}}\tilde{\mathbf{x}}_{h,l_h}/M_h, & \text{or} \\ \boldsymbol{\rho}_{h,l_h}^4 = \sqrt{N}\tilde{\mathbf{L}}_h^{\mathrm{H}}\tilde{\mathbf{x}}_{h,l_h}\tilde{\mathbf{x}}_{h,l_h}^{\mathrm{H}}\tilde{\mathbf{L}}_{\mathrm{h}}\tilde{\mathbf{L}}_h^{\mathrm{H}}\tilde{\mathbf{x}}_{h,l_h} / \left\|\tilde{\mathbf{L}}_h^{\mathrm{H}}\tilde{\mathbf{L}}_h\right\|_{\mathrm{F}}. \end{cases} \tag{28}$$

$\boldsymbol{\rho}_{h,l_h}^d$ is used for initialization in dth thread, and after convergence, $Q(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ corresponding to the initialization using $\boldsymbol{\rho}_{h,l_h}^d$ is represented by $Q^d(\boldsymbol{\alpha}, \boldsymbol{\gamma})$. The ultimate estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ are obtained from the thread corresponding to $\max_d Q^d(\boldsymbol{\alpha}, \boldsymbol{\gamma})$.

In each iteration, the computational complexity is dominated by the matrix inverse in calculating $\boldsymbol{\Sigma}_{h,l_h}$. As $N > M_h$, $\boldsymbol{\Sigma}_{h,l_h}$ is calculated by $\boldsymbol{\Sigma}_{h,l_h} = \boldsymbol{\Lambda} - \boldsymbol{\Lambda}\tilde{\mathbf{L}}_h^{\mathrm{H}}(\gamma_{h,l_h}\mathbf{I}_{M_h} + \tilde{\mathbf{L}}_h\boldsymbol{\Lambda}\tilde{\mathbf{L}}_h^{\mathrm{H}})^{-1}\tilde{\mathbf{L}}_h\boldsymbol{\Lambda}$ instead. The workflow of the proposed algorithm is listed in Table 1.

It was mentioned that the methods on frequency domain were applied in inverse solutions of brain rhythms. Our proposed method can also be employed for localizing cortical sources related to brain rhythms, if some adjustments are made. In our SSR source localization method, to alleviate the influence by the spontaneous brain rhythm components at the same frequency of SSRs, the joint sparsity among columns of $\mathbf{W}$ was modeled and utilized, and the variances of spontaneous EEG plus electric noise at all epochs were estimated. To adjust our method for inverse solutions of brain rhythms, $\mathbf{W}$ will be instead used to model the distributed sources of brain rhythms on the cortex, and electric noise is to be estimated, whose variance is assumed to be invariant among epochs. Though this adjustment seems simple, applying the adjusted method in brain rhythms should still be

careful, as brain rhythms cannot be assumed to keep their cortical locations for many epochs like SSRs, due to their non-stationarity property.

V. EXPERIMENTS FOR PERFORMANCE VERIFICATION

The performance of the proposed SBL-based SSR inverse solution is verified through experiments in two stages. In the first stage, simulation experiments are carried out, where the ground truth of simulated source locations is harnessed for judging the accuracy of inverse solution. Comparison between the proposed method and several existing inverse methods we adapted for SSR source localization is presented, and the outperformance of the proposed method is illustrated in various scenarios. After evaluation by simulations, in the second stage, the performance of the proposed method is further judged involving EEG data recorded during real SSR stimulation experiments. Here one typical category of SSRs, ASSR, is considered. The inverse solutions for ASSRs in various paradigms can be validated by conclusions on this issue in some existing literatures. Furthermore, referring to the advantages achieved by the proposed method in localization accuracy and spatial resolution revealed by simulation results, we expect that the proposed method may provide supplementary results for ASSR inverse solutions.

In this paper, the LFM $\tilde{\mathbf{L}}$ is calculated using OpenMEEG [31]. A common T1-weighted MRI anatomical template, the Montreal Neurological Institute/International Consortium for Brain Mapping (MNI/ICBM) 152 standard, is utilized for constructing $\tilde{\mathbf{L}}$. With the aid of OpenMEEG, a 3-shell realistic head model consisting of scalp, inner-skull, and outer-skull, was obtained using the symmetric boundary element method (BEM) as the distributed forward solver. The source space was manifested by deploying dense vertices on the surface of the cortex, and hence tessellating the cortex surface with a dense grid consisting of 1500 triangles corresponding to dipole positions. Given the electrode layout information, the LFM $\tilde{\mathbf{L}}$ mapping the source space to scalp electrodes is further constructed according to the distributed model. The generated lead field matrix is output by the Brainstorm software [32], to be reused by our method as well as the compared ones.

$$
\begin{aligned}
Q(\boldsymbol{\alpha}) &\triangleq E_{p(\mathbf{W}|\mathbf{X}_1,\,\mathbf{X}_2,\,\cdots,\,\mathbf{X}_H;\,\boldsymbol{\alpha},\boldsymbol{\gamma})}\left[\log p\left(\mathbf{W};\,\boldsymbol{\alpha}\right)\right] \\
&= \text{const} - \left(\sum_{h=1}^{H} L_h\right)\sum_{n=1}^{N}\log\alpha_n \\
&\quad - \sum_{h=1}^{H}\sum_{l_h=1}^{L_h}\sum_{n=1}^{N}\frac{E_{p(\mathbf{w}_{h,l_h}|\tilde{\mathbf{x}}_{h,l_h};\,\boldsymbol{\alpha},\boldsymbol{\gamma})}\left(\left|\mathbf{w}_{h,l_h}[n]\right|^2\right)}{\alpha_n} \\
&= \text{const} - \left(\sum_{h=1}^{H} L_h\right)\sum_{n=1}^{N}\log\alpha_n \\
&\quad - \sum_{h=1}^{H}\sum_{l_h=1}^{L_h}\sum_{n=1}^{N}\frac{\boldsymbol{\Sigma}_{h,l_h}[n,n] + \boldsymbol{\mu}_{h,l_h}[n]\,\boldsymbol{\mu}_{h,l_h}[n]^*}{\alpha_n}\,,
\end{aligned}
\tag{22}
$$

$$
\begin{aligned}
Q(\boldsymbol{\gamma}) &\triangleq E_{p(\mathbf{W}|\mathbf{X}_1,\,\mathbf{X}_2,\,\cdots,\,\mathbf{X}_H;\,\boldsymbol{\alpha},\boldsymbol{\gamma})}\left[\log p\left(\mathbf{X}_1,\;\mathbf{X}_2,\;\cdots,\;\mathbf{X}_H\,|\mathbf{W};\;\boldsymbol{\gamma}\right)\right] \\
&= \text{const} - \sum_{h=1}^{H}\left(M_h\sum_{l_h=1}^{L_h}\log\gamma_{h,l_h}\right) \\
&\quad - \sum_{h=1}^{H}\sum_{l_h=1}^{L_h}\frac{E_{p(\mathbf{w}_{h,l_h}|\tilde{\mathbf{x}}_{h,l_h};\,\boldsymbol{\alpha},\boldsymbol{\gamma})}\left[\left(\tilde{\mathbf{x}}_{h,l_h} - \tilde{\mathbf{L}}_h\mathbf{w}_{h,l_h}\right)^{\mathrm{H}}\left(\tilde{\mathbf{x}}_{h,l_h} - \tilde{\mathbf{L}}_h\mathbf{w}_{h,l_h}\right)\right]}{\gamma_{h,l_h}} \\
&= \text{const} - \sum_{h=1}^{H}\left(M_h\sum_{l_h=1}^{L_h}\log\gamma_{h,l_h}\right) - \sum_{h=1}^{H}\sum_{l_h=1}^{L_h}\left\{\tilde{\mathbf{x}}_{h,l_h}^{\mathrm{H}}\tilde{\mathbf{x}}_{h,l_h}\right. \\
&\quad \left. -2Re\left(\tilde{\mathbf{x}}_{h,l_h}^{\mathrm{H}}\tilde{\mathbf{L}}_h\boldsymbol{\mu}_{h,l_h}\right) + \mathrm{tr}\left[\tilde{\mathbf{L}}_h^{\mathrm{H}}\tilde{\mathbf{L}}_h\left(\boldsymbol{\Sigma}_{h,l_h} + \boldsymbol{\mu}_{h,l_h}\boldsymbol{\mu}_{h,l_h}^{\mathrm{H}}\right)\right]\right\}/\gamma_{h,l_h},
\end{aligned}
\tag{23}
$$

$$
\begin{aligned}
\gamma_{h,l_h} &= \frac{\tilde{\mathbf{x}}_{h,l_h}^{\mathrm{H}}\tilde{\mathbf{x}}_{h,l_h} - 2Re\left(\tilde{\mathbf{x}}_{h,l_h}^{\mathrm{H}}\tilde{\mathbf{L}}_h\boldsymbol{\mu}_{h,l_h}\right) + \mathrm{tr}\left[\tilde{\mathbf{L}}_h^{\mathrm{H}}\tilde{\mathbf{L}}_h\left(\boldsymbol{\Sigma}_{h,l_h} + \boldsymbol{\mu}_{h,l_h}\boldsymbol{\mu}_{h,l_h}^{\mathrm{H}}\right)\right]}{M_h}, \\
&\quad h = 1, 2, \cdots,\; H,\, l_h = 1,\; 2,\; \cdots,\; L_h\,.
\end{aligned}
\tag{25}
$$

TABLE 1. The proposed SSR inverse method.

Step 1	Input the preprocessed data matrices $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_H$ and LFMs $\tilde{\mathbf{L}}_1, \tilde{\mathbf{L}}_2, \cdots, \tilde{\mathbf{L}}_H$ for H subjects;
Step 2	Start D threads to calculating $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ iteratively, and assign the initializations to them, denoted as $\boldsymbol{\alpha}_{\text{old}}$ and $\boldsymbol{\gamma}_{\text{old}}$;
Step 3-1 (dth thread)	Calculate $\boldsymbol{\Sigma}_{h,J_h}$ and $\boldsymbol{\mu}_{h,J_h}$ using $\boldsymbol{\alpha}_{\text{old}}$ and $\boldsymbol{\gamma}_{\text{old}}$;
Step 3-2 (dth thread)	Update $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ using $\boldsymbol{\Sigma}_{h,J_h}$ and $\boldsymbol{\mu}_{h,J_h}$ according to Eqn(24) and Eqn(25);
Step 3-3 (dth thread)	If $\left\|\boldsymbol{\alpha}-\boldsymbol{\alpha}_{\text{old}}\right\|_2 < \varepsilon$ or the number of iterations reaches N_{iter}, stop iteration and calculate $Q(\boldsymbol{\alpha},\boldsymbol{\gamma})$, else make $\boldsymbol{\alpha}_{\text{old}} = \boldsymbol{\alpha}$, $\boldsymbol{\gamma}_{\text{old}} = \boldsymbol{\gamma}$, and go back to Step 3-1;
Step 4	Retain the $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ obtained from the thread that achieved the largest $Q(\boldsymbol{\alpha},\boldsymbol{\gamma})$; The inverse solution of SSR integrating H subjects is given by $\boldsymbol{\alpha}$.

ε denotes the iterative threshold and N_{iter} is the maximum number of iterations.

Electrode montages corresponding to two multi-electrode caps, developed by Easycap GmbH (Herrsching, Germany), were employed in the experiment verification. Their electrode layouts are displayed in Fig. 3. In simulation experiments, the generated simulation EEG recordings were assumed to be collected by EEG electrodes in the "64Ch-BrainCap-FEBA" or "32Ch-BrainCap-FEBA" cap. In the real SSR data verification stage, only the "64Ch-BrainCap-FEBA" cap was utilized for EEG recording.

A. SIMULATION EXPERIMENTS

In the simulation verification stage, the SSRs generated at specified source locations were synthesized, mapped to scalp by the LFM $\tilde{\mathbf{L}}$, and contaminated by electric noise and spontaneous EEG, leading to the EEG recordings at M scalp electrodes. Some pattern of stimulation was assumed to be given for arousing SSR, with the stimulus frequency f_0. The origins of K dipole sources were assumed located at K predetermined positions on the cortex, and hence mapping them to the scalp electrodes correspond to specified columns of $\tilde{\mathbf{L}}$. For each source signal, L epochs were generated. In each epoch, the source signal had the sinusoidal formulation, with random amplitude (uniformly distributed in $[0.8\sigma_s,\ 1.2\sigma_s]$) and random initial phase (uniformly distributed in $[-\pi, \pi]$), lasting for T_0. This implies that the amplitude and phase of SSR can vary among sources or even epochs, which will test the reasonability of setting the prior distributions of all columns of $\mathbf{W}$ to be an identical complex Gaussian distribution in (13). The spontaneous EEG plus electric noise in each epoch has temporally and spatially uncorrelated zero-mean Gaussian distribution, with variance $\sigma_l^2 + \sigma_n^2$, where σ_n^2 is fixed while $\sigma_l^2,\ l = 1, 2, \ldots, L$ obey Gamma distribution among epochs.

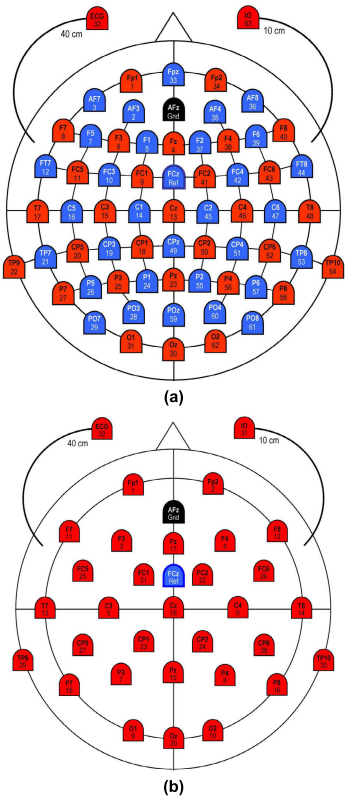

FIGURE 3. The montages of two electrode caps (Easycap GmbH, Herrsching, Germany) used in experiments. (a) Electrode layout of the employed "32Ch-BrainCap-FEBA" cap. (b) Electrode layout of the employed "64Ch-BrainCap-FEBA" cap.

The stimulus frequencies in the simulation experiments were all set as f_0 =39.1Hz, and the EEG recordings we simulated at scalp electrodes had the sampling frequency f_s =1000 Hz in this paper. The number of samples of EEG recordings in each epoch was fixed as 10230. For preprocessing the generated L-epoch multi-channel EEG simulation data, each epoch was divided into 10 nonoverlapped and successive segments and they were aligned and averaged to produce an "enhanced epoch" with length $T = 1023$, by a bit misusage of the concept. We believe that such a superposition averaging may bring two advantages: on one hand it reduced the number of FFT points, and on the other hand the SSR component can be "enhanced" compared with the spontaneous EEG and electric noise as the initial phases of sinusoidal sources in all the segments in the same epoch were guaranteed to be identical. After performing FFT on the lth "enhanced epoch", the Fourier components at f_0 = 39.1 Hz are derived to form the vector $\tilde{\mathbf{x}}_l$. For a simulated "subject", $\mathbf{X} = \left[\tilde{\mathbf{x}}_1,\ \tilde{\mathbf{x}}_2,\ \cdots,\ \tilde{\mathbf{x}}_L\right]$ obtained by preprocessing is used as the input to our proposed method. Change of stimulation frequency f_0 in simulation experiments will not lead to performance degradation, though the only thing to note is that the number of samples in data segments to be aligned and averaged should be adjusted, to guarantee that the initial phases of sinusoidal sources of f_0 in all the segments in the same epoch are identical.

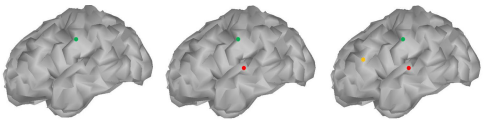

FIGURE 4. Locations of simulated sources in three cases, where the numbers of sources are 1, 2, and 3, respectively.

In the first simulation experiment, the one-trial examples of inverse solutions of simulated SSRs with preset cortical locations by the proposed method were given. We considered three cases, where sources were located at the left hemisphere of cortex, as displayed in Fig. 4. In the first case, the SSR had only one source originated at the postcentral gyrus. In the second case, another simulated source at the superior temporal gyrus was added. In the third case, three sources were generated, where a source located at the middle frontal gyrus was supplemented compared to the second case. The sources were mapped to be recorded by the EEG electrodes in the employed "64Ch-BrainCap-FEBA" cap with FCz as the reference, and hence the 62-channel EEG recording simulation data was generated for usage. For each case, we explored three signal-to-noise (SNR) scenarios: high SNR (SNR=12dB), medium SNR (SNR=0dB), and low SNR (SNR=-12dB). The EEG simulation data was assumed to be recorded from one subject, and the number of epochs was fixed as $L = 10$. In order to illustrate the superior performance of the proposed method, 3 existing inverse methods including MNE [13], sLORETA [14], and LCMV beamformer [16] were also simultaneously performed for comparison. For fair comparison, we adapted these three methods to handling the inverse problem through frequency domain analysis. The one-trial inverse solution results, projected on the surface of cortex in the form of color map, are displayed in Fig. 5, where "SBL" was used to label our method for distinguishing the proposed method from the compared ones, and the ground-truth simulated source location is marked by the small filled circle. Fig.5 shows that, apart from the performances of LCMV locating only one source, the compared methods yielded much smoother location results in various scenarios, and sometimes even merged or missed some sources. In contrast, the proposed method gave focal source localization, and the ground-truth sources lay right in the scope of its inverse solution in these trials. A trend that may be suspected from Fig. 5 is that, the proposed method may be much less affected by the SNR reduction, compared with the other methods.

The accuracy and resolution of source localization performance was further evaluated by Monte Carlo simulations. Statistical calculation of two indicators, root mean square error (RMSE) and spatial dispersion (SD) [33] was performed, involving the estimated locations and ground-truth source locations in Monte Carlo trials. The localization accuracy was evaluated by RMSE, which is given by

$$\mathrm{RMSE} = \sqrt{\sum_{k=1}^{K}\sum_{i=1}^{I} \left\|\tilde{\mathbf{r}}_{i,k} - \mathbf{r}_k\right\|_2^2 / (K \cdot I)}, \tag{29}$$

where $\mathbf{r}_k$ is the spatial location vector of the kth ground-truth source, $\tilde{\mathbf{r}}_{i,k}$ corresponds to the cluster center among K clusters of the localization results closest to the kth ground-truth source, and I is the total number of Monte Carlo trials. The SD was used to reflect the degree of localization ambiguity, represented as

$$\mathrm{SD} = \frac{1}{I}\sum_{i=1}^{I} \sqrt{\sum_{k=1}^{K}\sum_{j_k \in \mathcal{L}_k} \alpha_{j_k} \left\|\tilde{\mathbf{r}}_{i,j_k} - \mathbf{r}_k\right\|_2^2 / \sum_{n=1}^{N} \alpha_n}, \tag{30}$$

where $\mathcal{L}_k$ consists of the location indexes belonging to the cluster whose center is closest to the kth ground-truth source. In this paper, we used K-Nearest Neighbor (KNN) for clustering the localization result at each trial.

In Fig. 6, we plot the curves of RMSE and SD of source localization by the proposed method and the compared ones versus SNR, considering three cases of simulated sources given by Fig. 4. In all the cases, 500 Monte Carlo trials were carried out in each condition, and the number of epochs used in each trial was fixed at $L = 10$. The proposed method achieved the lowest RMSE and SD in low SNR range in all three cases. Specially, in the case of three sources, the proposed method gave substantially lower RMSE and SD curves compared with other methods in all SNR ranges. Some compared methods, which as mentioned were their frequency-domain-processing version, also revealed encouraging performances in localizing small number of SSR sources. E.g., LCMV yielded dominant performances in mediate and high SNR ranges when only one source was considered, and MNE gave lowest RMSEs in mediate and high SNR ranges at two sources' scenario. However, in real applications, as we never preliminarily know the true number of sources in the distributed source model, such "outperformance" of the compared methods cannot be confirmed. As observed from Fig. 6(c) that the lowest RMSE and SD curves versus SNR were achieved by the proposed method for localizing three sources, we are also interested in inspecting how these two indicators will vary with the increment of number of epochs used. Fig. 7 shows the RMSEs and SDs in the case of localizing three sources as displayed in Fig. 4, with variation of the number of epochs used and fixed SNR = 3dB. It can be found that all the performed methods, processing data in frequency domain, had the similar trend of very slow decline in RMSE and SD curves with respect to the increment of number of epochs used. Hence, we may suspect that the proposed method can already achieve appropriate source localization performance even when using a small number of epochs.

As we proved the convergence of EM procedure in our proposed method theoretically, we are interested in inspecting the convergence performance of our method in experiments. Still the case of localizing three sources displayed in Fig. 4 was considered, and the "64Ch-BrainCap-FEBA" cap was employed. We calculated the RMSEs and SDs of inverse solutions using the updated parameters at each round

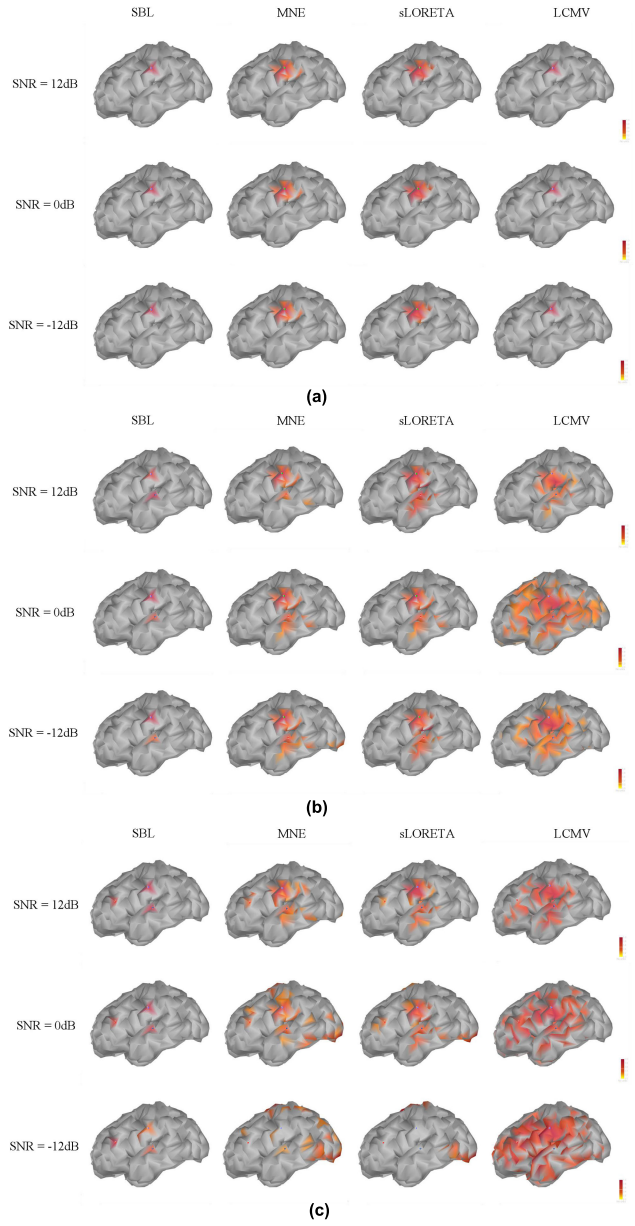

FIGURE 5. Examples of SSR source localization results using various localization methods in three cases. (a) One simulated source. (b) Two simulated sources. (c) Three simulated sources. The number of epochs used were all $L = 10$.

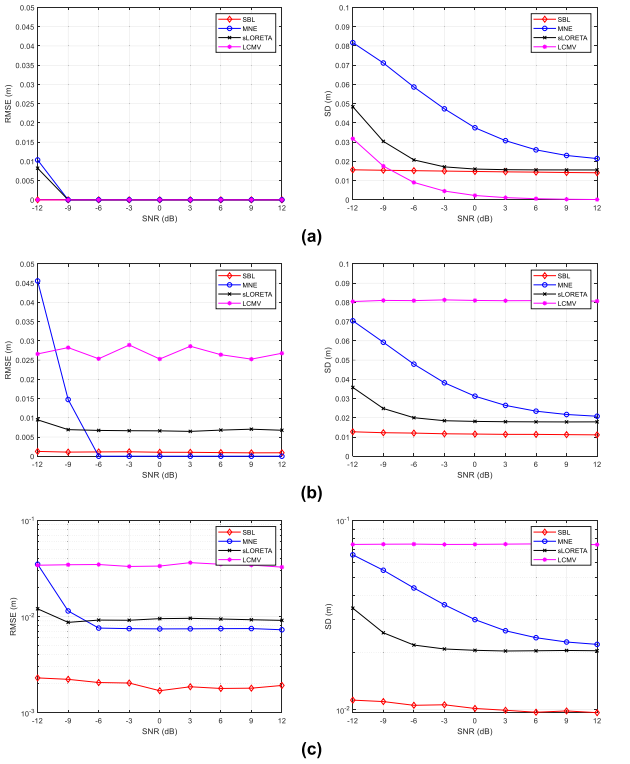

FIGURE 6. RMSE and SD of source localization results versus SNR in three cases, where the left column shows the RMSE curves and the right one shows the SD curves. (a) One simulated source. (b) Two simulated sources. (c) Three simulated sources. The number of epochs used were all fixed at $L = 10$.

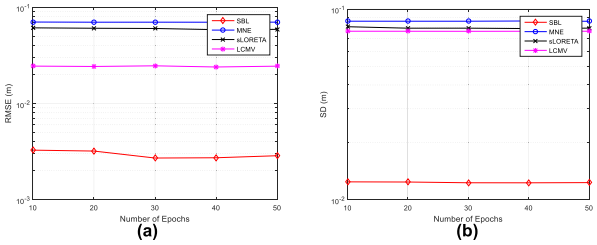

FIGURE 7. RMSE and SD of source localization results versus the number of epochs used in the case of three sources, where SNR was fixed at 3dB. (a) RMSE versus the number of epochs. (b) SD versus the number of epochs.

of iteration, and 500 Monte Carlo trials were carried out. Four scenarios were considered: (1) SNR=-20dB and $L = 5$; (2) SNR=-20dB and $L = 10$; (3) SNR=0dB and $L = 10$; (4) SNR=20dB and $L = 10$. Curves of RMSE and SD of inverse solutions versus time of iteration at these 4 scenarios are displayed in Fig. 8. It can be found that convergences were achieved around 100 rounds of iterations for these 4 simulation scenarios, and the stable values of SDs appeared much earlier than those of RMSEs. From Fig. 8, we can also speculate that, with fixed number of epochs our method will converge faster and achieve lower RMSE and SD of source localization in higher SNR scenario.

We examined our method's resistance to uncertainty in LFM $\tilde{\mathbf{L}}$ through simulation experiments. A random perturbation matrix $\Delta\tilde{\mathbf{L}}$ was added to $\tilde{\mathbf{L}}$ which was assumed given by precise forward solution. The entries of $\Delta\tilde{\mathbf{L}}$ are independent of each other and identically Gaussian distributed: $\Delta\tilde{\mathbf{L}}[m,n] \sim \mathcal{N}\left(0,\ p \times \left\|\tilde{\mathbf{L}}\right\|_{\mathrm{F}}^2 / \left(M^2N^2\right)\right)$, where p is a tunable parameter. The task was to localize three sources as displayed in Fig. 4 employing the "64Ch-BrainCap-FEBA" cap, under the conditions of $L = 10$ and SNR varied. 500 Monte Carlo trials were carried out in each condition, and we assumed that the perturbation in $\tilde{\mathbf{L}}$ was unknown in each trial for inverse solution. Four scenarios have been considered: (1) no perturbation existing in $\tilde{\mathbf{L}}$; (2) perturbation

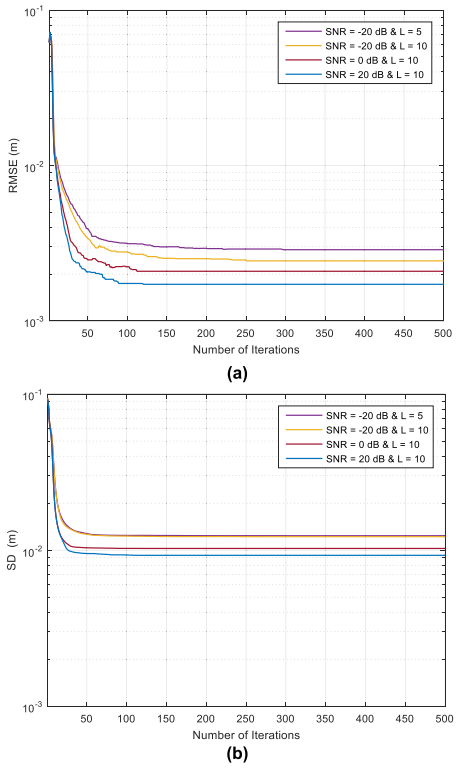

FIGURE 8. Convergence performances of the proposed method in various scenarios. (a) RMSE versus the number of iterations. (b) SD versus the number of iterations.

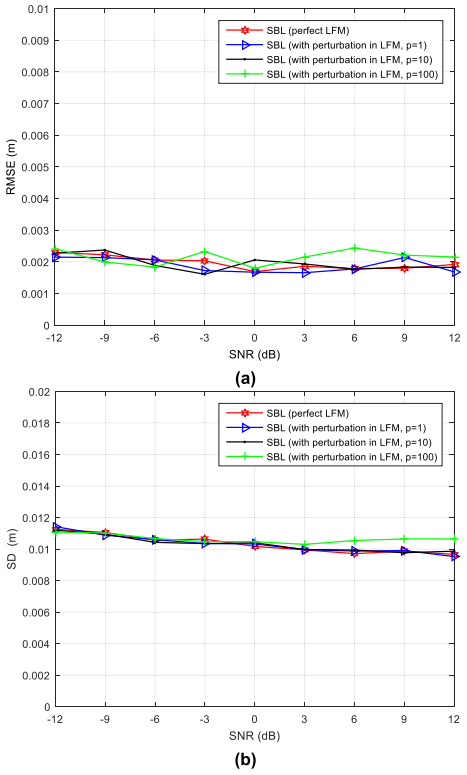

FIGURE 9. Performances of source localization by the proposed method confronting uncertainty in forward solution, where p controls the ratio of perturbation in LFM. (a) RMSE versus SNR. (b) SD versus SNR.

existing in $\tilde{\mathbf{L}}$ ($p = 1$); (3) perturbation existing in $\tilde{\mathbf{L}}$ ($p = 10$); (4) perturbation existing in $\tilde{\mathbf{L}}$ ($p = 100$). Fig. 9 depicts the RMSE curves and the SD curves with respect to SNR for these 4 scenarios. From these results, it can be noticed that the proposed method was almost immune to a small perturbation in $\tilde{\mathbf{L}}$ ($p = 1, 10$), and even for a larger perturbation ($p = 100$) the performance degradation was also very limited. This phenomenon coincides well with our conjecture from (10) that the perturbation in the LFM can be absorbed into γ_l which is estimated in the proposed method, hence endowing our method with some robustness to uncertainty in forward solution.

Compared with the existing method, our proposed SSR inverse solution method can easily fuse EEG recordings from multiple subjects to give an integrated source location result, even when they employ different electrode montages. We considered the scenario of combining 10-epoch simulation data by two subjects, who were assumed to use same or different electrode montages. These two subjects were assumed to have the same three-source SSR origins according to Fig. 4. For fair comparison, three cases were simulated, including: both of the subjects employing ''64Ch-BrainCap-FEBA'' caps, both of them employing ''32Ch-BrainCap-FEBA'' caps, and one employing ''64Ch-BrainCap-FEBA'' cap and the other one employing ''32Ch-BrainCap-FEBA'' cap. Fig. 10 shows the RMSE and SD curves versus SNR by our method, where ''SBL-64Ch'', ''SBL-32Ch'', and ''SBL-64Ch&32Ch'' mark the three cases mentioned above, respectively. As we expected, the source localization performance of ''SBL-64Ch'' outperformed that of ''SBL-32Ch'', due to the additional electrode recordings. ''SBL-64Ch&32Ch'' had the performance between ''SBL-64Ch'' and ''SBL-32Ch'' in low SNR range, while in high SNR range it had the smallest RMSE but largest SD among all three cases. The results displayed in Fig. 10 imply that the proposed method achieved a neat style of information integration of multiple subjects with probable different electrode montages in SSR inverse solution.

To show how the proposed method will perform when more SSR sources are to be localized, we considered the scenario where eight sources were simulated covering the left and right hemisphere of cortex. The ground-truth locations of simulated sources include: left middle frontal gyrus, left postcentral gyrus, left inferior occipital gyrus, left middle temporal gyrus, right middle frontal gyrus, right parieto-occipital fissure, right middle occipital gyrus, and right middle temporal gyrus. Fig. 11 shows one trial example of source localization by the proposed method in this scenario, where SNR = 12dB, the montage of ''64Ch-BrainCap-FEBA'' cap was used, and 10 epochs of simulation data were generated and used. The ground-truth simulation source locations are marked by colored solid points on the cortex. It can be observed that in this example, the SSR source localization results of our method were sparse, focusing right around the ground-truth simulation source locations. As four lobes of the cerebral cortex

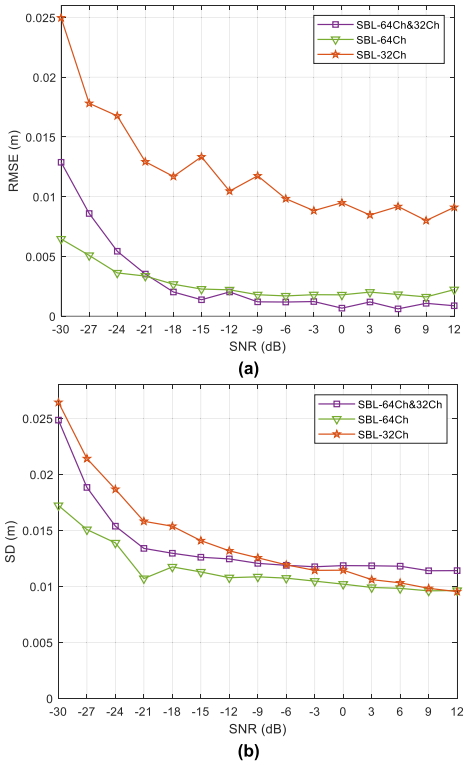

FIGURE 10. SSR Source localization results of the proposed method combining simulation data of two subjects using same or different recording electrode montages, where 10 epochs of data were simulated for each subject. (a) RMSE versus SNR. (b) SD versus SNR.

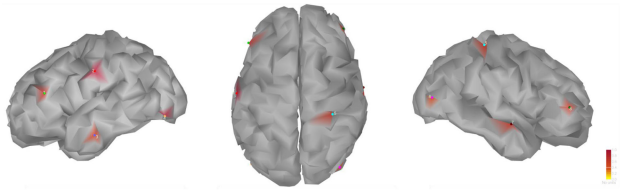

FIGURE 11. One trial example of SSR Source localization results of the proposed method in locating eight simulated sources.

were all involved when presetting the locations of simulated SSR sources in this example, the proposed method showed its potential in inspecting the locations of SSR sources from the whole cerebral cortex with fine resolution.

B. REAL SSR EXPERIMENTS: 40-Hz ASSR EXAMPLES

Encouraged by the performances of our method in simulation experiments, we are interested in exploring how well it works in solving inverse problems for real SSR data. In this paper, we considered 40-Hz ASSRs, stimulated by monaural or binaural periodic auditory stimulus at around 40 Hz, whose clinical applications were widely studied in hearing detection, anesthesia monitoring, consciousness state assessment, etc. It is of great significance to figure out the origins of 40-Hz ASSRs to support its clinical applications from the perspective of source generation. As aforementioned, the source locations of 40-Hz ASSR were explored in literatures with some conventional methods such as BESA [22], [24], LORETA [23], LCMV [25], and ECD fitting [26]. Hence, these existing studies may provide some proof for 40-Hz ASSR source localization results. Furthermore, according to the outperformance of our method shown in simulation experiments, we also expect that it may give supplementary or calibrated conclusions of 40-Hz ASSR inversion solution in comparison to the existing studies.

The collection of 40-Hz ASSR recordings comes from our previous study on developing a 40-Hz ASSR automatic detector and combining multi-paradigm 40-Hz ASSR automatic detection results for prognosis of comatose patients [8]. The 40-Hz ASSR automatic detector in [8] was developed by using EEG recordings of 26 healthy right-handed young volunteers experiencing multi-paradigm 40-Hz auditory stimulation. The involved volunteers all gave their signed informed consents before participating in the experiments. The study in [8] was carried out in accordance with the Declaration of Helsinki and approved by the Medical Ethics Committee of the Taicang Affiliated Hospital of Soochow University. Among the 26 healthy volunteers, 6 subjects (3 males and 3 females, aged 21$\pm$ 1.89 years) admitted to have their whole-scalp EEG recording, using the ''64Ch-BrainCap-FEBA'' cap. In this paper, we applied our SSR inverse method in the EEG recordings of these 6 subjects undergoing multi-paradigm 40-Hz auditory stimuli.

The auditory stimuli were in the form of sinusoidal amplitude modulated tones at $f_0 = 39.1$Hz, generated by MATLAB2015b and output by customized in-ear headphones, with sound intensity level of 70 dB SPL for healthy subjects. 4 paradigms of stimulation were presented: binaural stimulation with carrier frequency $f_c = 500$ Hz (denoted as ''39.1-500-both''), monaural stimulation at left ear with $f_c = 500$ Hz (denoted as ''39.1-500-left''), binaural stimulation with $f_c = 1000$ Hz (denoted as ''39.1-1000-both''), and monaural stimulation at left ear with $f_c = 1000$ Hz (denoted as ''39.1-1000-left''). When stimulation was only given at the left ear, the right ear of the subject was masked by white noise at the same time. For each stimulation paradigm, the auditory stimulation lasted for 150 seconds. Between every two adjacent stimulation paradigms, the subjects rest for 1 minute. The subjects were asked to keep their eyes closed during the whole experiment for excluding electrooculogram (EOG) artifacts.

The EEG recording was achieved by using BrainAmp (S/N AMP 13122084MRplus) and BrainVision Recorder (Version 1.20.0601) produced by Brain Products GmbH (Munich, Germany). Conductive gel was used to keep the resistance between each electrode and scalp under 5kΩ. The sampling rate was $f_s = 1000$ Hz, and a bandpass filter with 1$\sim$300 Hz passband (12dB/octave) and a 50-Hz notch filter were used for noise suppression during recording. For the recorded 150s data related to each paradigm of stimulation, preprocessing was applied to form a data matrix $\mathbf{X}$ consisting of $L = 10$ epochs, which was the input to our proposed method. Firstly, the EEG recording was re-referenced from FCz to the average of TP9 and TP10, and hence transformed from 63-channel data to 62-channel data. Secondly, the first 5s

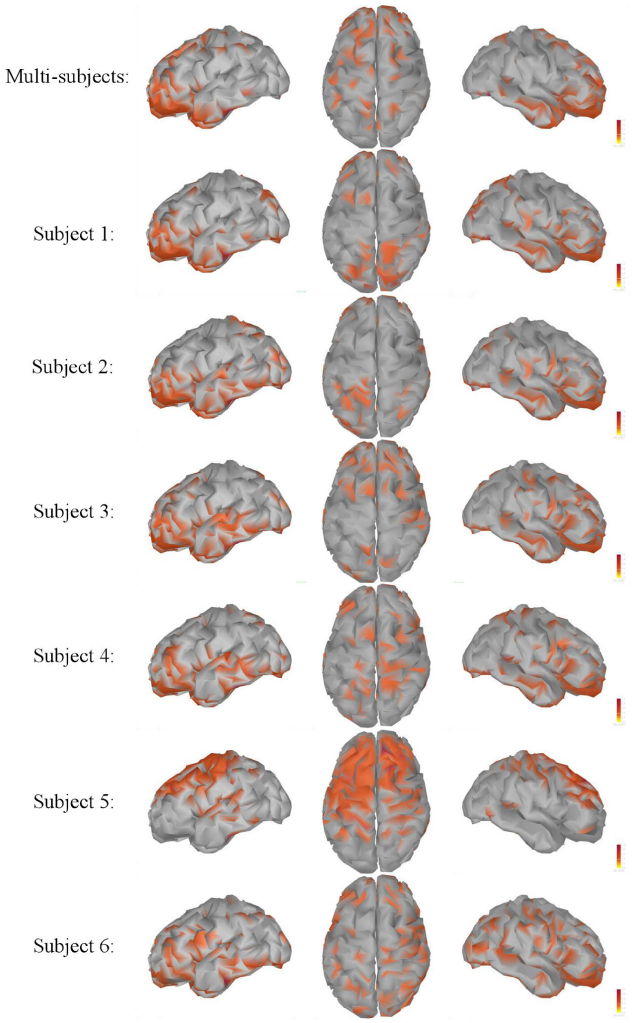

FIGURE 12. Cortical source localization results for ASSR evoked by stimulation paradigm "39.1-500-both". The top row displays the left, top, and right views of the ASSR inverse solution on cerebral cortex provided by the proposed method integrating all 6 subjects' data, while each subsequent row shows the solution using the corresponding one subject's data alone.

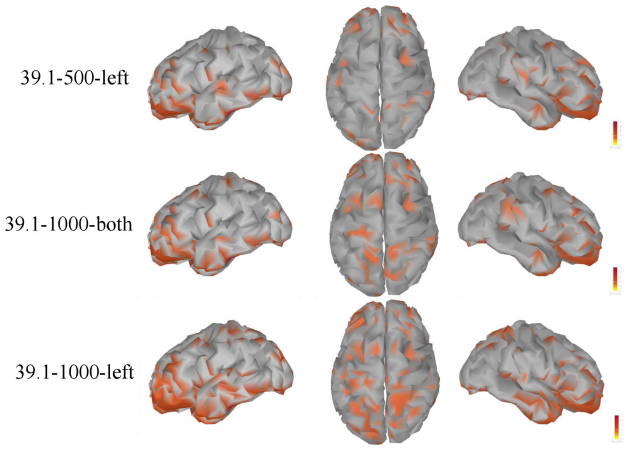

FIGURE 13. Cortical source localization results for ASSR evoked by stimulation paradigm "39.1-500-left", "39.1-1000-both", and "39.1-1000-left", where each row shows the left, top, and right views of the ASSR inverse solution on cerebral cortex for the corresponding stimulation paradigm. All solutions were given by our method combining all subjects' data under a certain stimulation paradigm.

data was used for baseline calibration, and the rest data was divided into successive and nonoverlapped data segments, with equivalent length of 1023 samples. Thirdly, artifacts rejection was performed at each data segment: one data segment was abandoned if the amplitude of any sample in it exceeded ± 40 μV; otherwise it was reserved. Fourthly, the reserved data segments were passed through a bandpass digital filter in 35$\sim$45 Hz, and then every 10 successive data segments were averaged to form a 1023-sample "epoch", where epochs were guaranteed to be nonoverlapped and only the first 10 epochs were reserved for subsequent usage. Lastly, the lth column of $\mathbf{X}$ ($l = 1, 2, \ldots, L$) was given by the Fourier components of the lth epoch data at $f_0 = 39.1$ Hz.

For each subject under one paradigm of stimulation, we obtained a 62×10 data matrix $\mathbf{X}$ as the input to our proposed method. Our method can also easily utilize all subjects' data to exploit the common source locations of ASSR evoked by a certain paradigm of stimulation. Fig. 12 shows the source location results on the cerebral cortex using each subject's data alone or combining all 6 subjects' data, where the ASSR evoked by "39.1-500-both" was taken for example. Comparing the integrated result with the one obtained from each subject only in Fig. 12, we can observe that the source location integration was not just a simple multi-subject-averaging version, but gave a collection of dominant located sources by our proposed method applied in all subjects' data. If more subjects are involved, we can expect to derive a more robust conclusion of cortical origins for the SSR evoked by some pattern of simulation.

The integrated source location results for the ASSRs evoked by the rest of paradigms of stimulation are displayed in Fig. 13. It seems that when the carrier frequency $f_c = 500$ Hz, the binaural stimulation may arouse 40-Hz ASSR with a little wider cortical source distribution than that aroused by the monaural one. However, this assumption did not hold in the case of $f_c = 1000$ Hz, as displayed in Fig. 13. In this paper, we tried to figure out the common source locations of 40-Hz ASSRs aroused by 4 paradigms of auditory stimulation that we carried out. From Fig. 12 and Fig. 13, our proposed method revealed that the locations of 40-Hz ASSR's cortical origins include: superior temporal gyrus, transverse temporal gyrus, middle temporal gyrus, inferior temporal gyrus, postcentral gyrus, angular gyrus, lingual gyrus, middle frontal gyrus, and lateral orbital gyrus. Most of the above derived source locations can find their counterparts in the literatures: in [23], temporal lobe, parietal lobe, and frontal lobe were claimed to have 40-Hz ASSR origins; temporal gyri coincided with the conclusions in [22] and [25]; postcentral gyrus, lingual gyrus, and middle frontal gyrus were also confirmed as locations of 40-Hz ASSR sources in [26]. Hence, our proposed method provided a more comprehensive conclusion of 40-Hz ASSR's origins compared with

the inverse solution given by any existing method alone. Furthermore, as illustrated by simulation experiments, our method achieved SSR source localization with much higher resolution covering the whole cortex. From the experiments involving 4 paradigms of auditory stimulation, the results of our method showed that, sources of 40-Hz ASSR may also appear in angular gyrus and lateral orbital gyrus, rather than vague conjectures on parietal lobe and frontal lobe.

VI. CONCLUSION

In this paper, the inverse problem for SSRs was considered in perspective of multi-channel EEG signal processing at frequency domain, where the framework of SBL was applied. We showed that the complex Fourier components of multi-channel SSR data at the stimulation frequency have a sparse representation counterpart on the cerebral cortex according to the distributed model. Based on the sparse representation of Fourier components, a hierarchical Bayesian model was established for SSR source localization using a single subject's data as well as combining multiple subjects' data. In designing the Bayesian model, the joint sparsity of complex-valued source component vectors among epochs was modeled and the non-stationarity of spontaneous EEG was considered. EM was employed to derive the iterative Bayesian inference procedure for solving the SSR inverse problem. Specially, the proposed method is also applicable in SSR source localization integrating the EEG recordings of multiple subjects even employing different electrode montages. The simulation experiments verified the advantages in source localization accuracy and resolution achieved by the proposed method, compared with the frequency-domain-analysis version of the conventional inverse methods. Real 40-Hz ASSR stimulation experiments were carried out and the recorded data was utilized for further performance verification. The yielded results coincided well with the existing literatures and provided SSR source location results with higher resolution on some lobes. Encouraged by the SSR source localization performances of the proposed method, we expect it to make more contributions to SSR related brain research and applications in the future.

APPENDIX A. PROOF OF CONVERGENCE

If $q(\mathbf{W})$ is an arbitrary distribution over $\mathbf{W}$, $\log p(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha},\boldsymbol{\gamma})$ can be reformed as

$$\begin{aligned}
&\log p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha},\boldsymbol{\gamma}\right)\\
&\quad=\int q\left(\mathbf{W}\right)\log\frac{p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H,\mathbf{W};\ \boldsymbol{\alpha},\boldsymbol{\gamma}\right)}{q\left(\mathbf{W}\right)}\\
&\qquad+KL\left(q\left(\mathbf{W}\right)\|p\left(\mathbf{W}\,|\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha},\boldsymbol{\gamma}\right)\right)\\
&\quad\geq\int q\left(\mathbf{W}\right)\log\frac{p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H,\mathbf{W};\ \boldsymbol{\alpha},\boldsymbol{\gamma}\right)}{q\left(\mathbf{W}\right)}\ ,
\end{aligned}\tag{31}$$

where $KL\left(q\,\|p\right)$ denotes the Kullback-Leibler divergence, which is nonnegative. (31) provides a tight lower bound of $\log p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha},\boldsymbol{\gamma}\right)$, i.e.

$$\begin{aligned}
&\mathcal{LB}\left[\log p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha},\boldsymbol{\gamma}\right)\right]\\
&\quad=\int q\left(\mathbf{W}\right)\log\frac{p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H,\mathbf{W};\ \boldsymbol{\alpha},\boldsymbol{\gamma}\right)}{q\left(\mathbf{W}\right)}\ ,
\end{aligned}\tag{32}$$

and the equality in (31) will hold when $q(\mathbf{W}) = p(\mathbf{W}\,|\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha},\boldsymbol{\gamma})$. In each iteration of EM, the E-step in (31) gives a lower bound of $\log p(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha},\boldsymbol{\gamma})$, while the M-step maximizes this lower bound to make it get closer to $\max_{\boldsymbol{\alpha},\boldsymbol{\gamma}}\log p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha},\boldsymbol{\gamma}\right)$. In the $(t+1)$th iteration, if $q^{(t)}(\mathbf{W}) = p\left(\mathbf{W}\,|\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha}^{(t)},\boldsymbol{\gamma}^{(t)}\right)$ built by $\boldsymbol{\alpha}^{(t)}$ and $\boldsymbol{\gamma}^{(t)}$ obtained from the t th iteration, the M-step is given by

$$\begin{aligned}
&\left(\boldsymbol{\alpha}^{(t+1)},\boldsymbol{\gamma}^{(t+1)}\right)\\
&=\underset{\boldsymbol{\alpha},\boldsymbol{\gamma}}{\operatorname{argmax}}\,\mathcal{LB}\left[\log p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha},\boldsymbol{\gamma}\right)\right]\Big|_{q^{(t)}(\mathbf{W})}\\
&=\underset{\boldsymbol{\alpha},\boldsymbol{\gamma}}{\operatorname{argmax}}\int q^{(t)}\left(\mathbf{W}\right)\log\frac{p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H,\mathbf{W};\ \boldsymbol{\alpha},\boldsymbol{\gamma}\right)}{q^{(t)}\left(\mathbf{W}\right)}\\
&=\underset{\boldsymbol{\alpha},\boldsymbol{\gamma}}{\operatorname{argmax}}\int q^{(t)}\left(\mathbf{W}\right)\log p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H,\mathbf{W};\ \boldsymbol{\alpha},\boldsymbol{\gamma}\right)\\
&=\underset{\boldsymbol{\alpha},\boldsymbol{\gamma}}{\operatorname{argmax}}Q\left(\boldsymbol{\alpha},\boldsymbol{\gamma}\right)\Big|_{q^{(t)}(\mathbf{W})=p\left(\mathbf{W}|\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha}^{(t)},\boldsymbol{\gamma}^{(t)}\right)}
\end{aligned}\tag{33}$$

Now combing (31) and (33), we further have

$$\begin{aligned}
&\log p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha}^{(t+1)},\boldsymbol{\gamma}^{(t+1)}\right)\\
&=\int q^{(t+1)}(\mathbf{W})\log\frac{p\left(\mathbf{X}_1,\ \mathbf{X}_2,\cdots,\mathbf{X}_H,\mathbf{W};\ \boldsymbol{\alpha}^{(t+1)},\boldsymbol{\gamma}^{(t+1)}\right)}{q^{(t+1)}\left(\mathbf{W}\right)}\\
&\geq\int q^{(t)}\left(\mathbf{W}\right)\log\frac{p\left(\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_H,\mathbf{W};\ \boldsymbol{\alpha}^{(t+1)},\boldsymbol{\gamma}^{(t+1)}\right)}{q^{(t)}\left(\mathbf{W}\right)}\\
&\geq\int q^{(t)}\left(\mathbf{W}\right)\log\frac{p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H,\mathbf{W};\ \boldsymbol{\alpha}^{(t)},\boldsymbol{\gamma}^{(t)}\right)}{q^{(t)}\left(\mathbf{W}\right)}\\
&=\log p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha}^{(t)},\boldsymbol{\gamma}^{(t)}\right)\ .
\end{aligned}\tag{34}$$

From (34), it can be observed that the sequence $\log p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha}^{(t)},\boldsymbol{\gamma}^{(t)}\right)$ is non-decreasing. As $\log p\left(\mathbf{X}_1,\ \mathbf{X}_2,\ \cdots,\ \mathbf{X}_H;\ \boldsymbol{\alpha}^{(t)},\boldsymbol{\gamma}^{(t)}\right)\leq 0$, this sequence is also upper bounded, so its convergence is guaranteed.

REFERENCES

[1] F. Logi, C. Fischer, L. Murri, and F. Mauguière, "The prognostic value of evoked responses from primary somatosensory and auditory cortex in comatose patients," *Clin. Neurophysiol.*, vol. 114, no. 9, pp. 1615–1627, Sep. 2003.

[2] T. Meigen and M. Bach, "On the statistical significance of electrophysiological steady-state responses," *Documenta Ophthalmol.*, vol. 98, no. 3, pp. 207–232, Jul. 1999.

[3] B. Lütkenhöner and R. D. Patterson, "Disruption of the auditory response to a regular click train by a single, extra click," *Exp. Brain Res.*, vol. 233, no. 6, pp. 1875–1892, Jun. 2015.

[4] F.-C. Lin, J. K. Zao, K.-C. Tu, Y. Wang, Y.-P. Huang, C.-W. Chuang, H.-Y. Kuo, Y.-Y. Chien, C.-C. Chou, and T.-P. Jung, "SNR analysis of high-frequency steady-state visual evoked potentials from the foveal and extrafoveal regions of human retina," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, San Diego, CA, USA, Aug. 2012, pp. 1810–1814.

[5] R. Gransier, A. Wieringen, and J. Wouters, "Binaural interaction effects of 30–50 Hz auditory steady state responses," *Ear Hear.*, vol. 38, no. 5, pp. 305–315, Sep. 2017.

[6] H. Cecotti, "Adaptive time segment analysis for steady-state visual evoked potential based brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 3, pp. 552–560, Mar. 2020.

[7] K.-T. Kim, J. Lee, H. Kim, C. H. Kim, and S. J. Lee, "Classification of selective attention within steady-state somatosensory evoked potentials from dry electrodes using mutual information-based spatio-spectral feature selection," *IEEE Access*, vol. 8, pp. 85464–85472, 2020.

[8] T. Chen, S. Lu, P. Qian, G. Chen, and N. Hu, "An automatic detection method for 40-Hz auditory steady state response and its application in prognosis of comatose patients," *Clin. Neurophysiol.*, vol. 131, no. 3, pp. 703–715, Mar. 2020.

[9] M. Middendorf, G. Mcmillan, G. Calhoun, and K. S. Jones, "Brain-computer interfaces based on the steady-state visual-evoked response," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 211–214, Jun. 2000.

[10] J. J. Podmore, T. P. Breckon, N. K. N. Aznan, and J. D. Connolly, "On the relative contribution of deep convolutional neural networks for SSVEP-based bio-signal decoding in BCI speller applications," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 611–618, Apr. 2019.

[11] B. He, T. Musha, Y. Okamoto, S. Homma, Y. Nakajima, and T. Sato, "Electric dipole tracing in the brain by means of the boundary element method and its accuracy," *IEEE Trans. Biomed. Eng.*, vol. BME-34, no. 6, pp. 406–414, Jun. 1987.

[12] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *IEEE Signal Process. Mag.*, vol. 18, no. 6, pp. 14–30, Nov. 2001.

[13] M. S. Hämäläinen and R. J. Ilmoniemi, "Interpreting magnetic fields of the brain: Minimum norm estimates," *Med. Biol. Eng. Comput.*, vol. 32, no. 1, pp. 35–42, Jan. 1994.

[14] R. D. Pascual-Marqui, "Standardized low resolution brain electromagnetic tomography (sLORETA): Technical details," *Methods Findings Experim. Clin. Pharmacol.*, vol. 24, pp. 5–12, Jan. 2002.

[15] M.-X. Huang, A. M. Dale, T. Song, E. Halgren, D. L. Harrington, I. Podgorny, J. M. Canive, S. Lewis, and R. R. Lee, "Vector-based spatial–temporal minimum L1-norm solution for MEG," *NeuroImage*, vol. 31, no. 3, pp. 1025–1037, Jul. 2006.

[16] B. D. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki, "Localization of brain electrical activity via linearly constrained minimum variance spatial filtering," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 867–880, Sep. 1997.

[17] S. Liu, Y. Huang, H. Wu, C. Tan, and J. Jia, "Efficient multitask structure-aware sparse Bayesian learning for frequency-difference electrical impedance tomography," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 463–472, Jan. 2021.

[18] K. J. Friston, W. Penny, C. Phillips, S. Kiebel, G. Hinton, and J. Ashburner, "Classical and Bayesian inference in neuroimaging: Theory," *NeuroImage*, vol. 16, no. 2, pp. 465–483, Jun. 2002.

[19] F. Costa, H. Batatia, T. Oberlin, C. D'Giano, and J.-Y. Tourneret, "Bayesian EEG source localization using a structured sparsity prior," *NeuroImage*, vol. 144, pp. 142–152, Jan. 2017.

[20] S. Saha, R. Rana, Y. Nesterets, M. Tahtali, F. de Hoog, and T. Gureyev, "Evaluating the performance of BSBL methodology for EEG source localization on a realistic head model," *Int. J. Imag. Syst. Technol.*, vol. 27, no. 1, pp. 46–56, Mar. 2017.

[21] W. Chen, C. Chang, and Y. Hu, "Single-trial extraction of pure somatosensory evoked potential based on expectation maximization approach," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 1, pp. 10–19, Jan. 2016.

[22] A. T. Herdman, O. Lins, P. Van Roon, D. R. Stapells, M. Scherg, and T. W. Picton, "Intracerebral sources of human auditory steady-state responses," *Brain Topogr.*, vol. 15, no. 2, pp. 69–86, Dec. 2002.

[23] S. A. Reyes, A. H. Lockwood, R. J. Salvi, M. L. Coad, D. S. Wack, and R. F. Burkard, "Mapping the 40-hz auditory steady-state response using current density reconstructions," *Hearing Res.*, vol. 204, nos. 1–2, pp. 1–15, Jun. 2005.

[24] C. Poulsen, T. W. Picton, and T. Paus, "Age-related changes in transient and oscillatory brain responses to auditory stimulation in healthy adults 19-45 years old," *Cerebral Cortex*, vol. 17, no. 6, pp. 1454–1467, Jun. 2007.

[25] M. Popescu, E.-A. Popescu, T. Chan, S. D. Blunt, and J. D. Lewine, "Spatio–temporal reconstruction of bilateral auditory steady-state responses using MEG beamformers," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 3, pp. 1092–1102, Mar. 2008.

[26] E. D. Farahani, T. Goossens, J. Wouters, and A. van Wieringen, "Spatiotemporal reconstruction of auditory steady-state responses to acoustic amplitude modulations: Potential sources beyond the auditory pathway," *NeuroImage*, vol. 148, pp. 240–253, Mar. 2017.

[27] Z. J. Koles, P. Flor-Henry, and J. C. Lind, "Low-resolution electrical tomography of the brain during psychometrically matched verbal and spatial cognitive tasks," *Hum. Brain Mapping*, vol. 12, no. 3, pp. 144–156, Mar. 2001.

[28] L. R. R. Gianotti, G. Künig, D. Lehmann, P. L. Faber, R. D. Pascual-Marqui, K. Kochi, and U. Schreiter-Gasser, "Correlation between disease severity and brain electric LORETA tomography in Alzheimer's disease," *Clin. Neurophysiol.*, vol. 118, no. 1, pp. 186–196, Jan. 2007.

[29] A.-M. Cebolla, E. Palmero-Soler, A. Leroy, and G. Cheron, "EEG spectral generators involved in motor imagery: A swLORETA study," *Frontiers Psychol.*, vol. 8, Dec. 2017, Art. no. 2133.

[30] T. Chen, J. Wang, B. Sun, and N. Hu, "Solving the inverse problem for auditory steady-state response via sparse Bayesian learning," in *Proc. IEEE 4th Inf. Technol., Netw., Electron. Autom. Control Conf. (ITNEC)*, Jun. 2020, pp. 1263–1267.

[31] A. Gramfort, T. Papadopoulo, E. Olivi, and M. Clerc, "OpenMEEG: Opensource software for quasistatic bioelectromagnetics," *Biomed. Eng. OnLine*, vol. 9, no. 1, 2010, Art. no. 45.

[32] F. Tadel, S. Baillet, J. C. Mosher, D. Pantazis, and R. M. Leahy, "Brainstorm: A user-friendly application for MEG/EEG analysis," *Comput. Intell. Neurosci.*, vol. 2011, Apr. 2011, Art. no. 879716.

[33] W.-T. Chang, A. Nummenmaa, J.-C. Hsieh, and F.-H. Lin, "Spatially sparse source cluster modeling by compressive neuromagnetic tomography," *NeuroImage*, vol. 53, no. 1, pp. 146–160, Oct. 2010.

MINGWEN QU received the B.E. degree in internet of things engineering from the Jiangsu University of Science and Technology, Zhenjiang, China, in 2019, and the master's degree from the School of Electronics and Information Engineering, Soochow University, Suzhou, China, in 2020. Her research interests include EEG signal processing and application.

TINGTING CHEN received the B.E. degree in biomedical engineering from Anhui Medical University, Hefei, China, in 2017, and the M.E. degree in signal and information processing from Soochow University, Suzhou, China, in 2020. She is currently an Engineer with Unisoc Corporation, Nanjing, China. Her research interests include EEG signal processing and application.

SHIQI LU received the B.M. degree from Zunyi Medical University, Zunyi, China, in 1982, and the M.M. and M.D. degrees in neurosurgery from Soochow University, Suzhou, China, in 2002 and 2005, respectively. He is currently an Associate Professor and the Director Physician with the Department of Emergency, The First Affiliated Hospital of Soochow University.

JIANLING HU (Member, IEEE) received the Ph.D. degree in communication and information systems from Shanghai Jiao Tong University, Shanghai, China, in 2000. From 2001 to 2008, he worked as an Associate Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University. Since 2008, he has been a Professor with the School of Electronics and Information Engineering, Soochow University, Suzhou, China. His research interests include wireless communications, embedded system design, and multimedia signal processing.

JIAJUN WANG received the B.Sc. and M.Sc. degrees in physics from Soochow University, China, in 1992 and 1995, respectively, and the Ph.D. degree in biomedical engineering from Zhejiang University, in 1999. He is currently a Professor with the School of Electronics and Information Engineering, Soochow University, China. He has published more than 50 scientific journal or conference papers. His research interests include medical imaging, image processing, pattern recognition, and bioinformatics.

NAN HU received the B.E. degree in electronic information engineering and the Ph.D. degree in signal and information processing from the University of Science and Technology of China, Hefei, China, in 2008 and 2013, respectively. From 2013 to 2014, he was an Engineer with 14th Research Institute of China Electronic Technology Group Corporation, Nanjing, China. He is currently an Associate Professor with the School of Electronics and Information Engineering, Soochow University, Suzhou, China. His research interests include multi-channel signal processing, biomedical signal processing, and machine learning.