Task-Related EEG Source Localization via Graph Regularized Low-Rank Representation Model

Feng Liu[a], Jay Rosenberger[a], Jing Qin[b], Yifei Lou[c], Shouyi Wang[a,*]

[a] *Department of Industrial and Manufacturing Systems Engineering*
University of Texas at Arlington
Arlington, TX, 76019, USA

[b] *Department of Mathematical Sciences*
Montana State University, Bozeman, MT, 59717, USA

[c] *Department of Mathematical Sciences*
University of Texas at Dallas, Richardson, TX, 75080, USA

Abstract

To infer brain source activation patterns under different cognitive tasks is an integral step to understand how our brain works. Traditional electroencephalogram (EEG) Source Imaging (ESI) methods usually do not distinguish task-related and spurious non-task-related sources that jointly generate EEG signals, which inevitably yield misleading reconstructed activation patterns. In this research, we argue that the task-related source signal intrinsically has a low-rank property, which is exploited to to infer the true task-related EEG sources location. Although the true task-related source signal is sparse and low-rank, the contribution of spurious sources scattering over the source space with intermittent activation patterns makes the actual source space lose the low-rank property. To reconstruct a low-rank true source, we propose a novel ESI model that involves a spatial low-rank representation and a temporal Laplacian graph regularization, the latter of which guarantees the temporal smoothness of the source signal and eliminate the spurious ones. To solve the proposed model, an augmented Lagrangian objective function is formulated and an algorithm in the framework of alternating direction method of multipliers is proposed. Numerical results illustrate the effectiveness of the proposed method in terms of reconstruction accuracy with high efficiency.

Keywords: EEG Source Imaging, Low-Rank Representation, Graph Regularization, Alternating Direction Method of Multiplier (ADMM)

*Corresponding author. Shouyi Wang, Tel.: 817-272-2921.

Email addresses: `feng.liu@mavs.uta.edu` (Feng Liu), `jrosenbe@uta.edu` (Jay Rosenberger), `shouyiw@uta.edu` (Shouyi Wang)

1. Introduction

As a direct measurement modality of neural electrical firing patterns, electroencephalogram (EEG) has a higher temporal resolution up to millisecond compared to positron emission topography (PET) and functional magnetic resonance imaging (fMRI) [1–3]. However, a limitation of EEG is its low spatial resolution since the measurement is on the scalp rather than inside the brain. Given the recorded scalp EEG data, to reconstruct of the brain sources signal inside the brain is known as EEG inverse problem or EEG source imaging (ESI) [4]. ESI technique has been widely used in the study of language mechanisms, cognition process, sensory function, as well as the localization of epileptic seizure [1, 5]. ESI is also an preliminary step for brain connectivity analysis [6, 7] on the source level [8–11].

Since the number of electrodes usually outnumbers that of brain sources, the ESI problem is highly ill-posed or mathematically underdetermined. A variety of methods have been proposed to address this challenging problem with different neurophysiological assumptions,formulated by various regularization techniques [4, 12]. There are two main types of inverse solvers for ESI: dipole fitting and distributed inverse solvers [13]. Dipole fitting empirically solves the MEG/EEG forward and inverse problems by characterizing the neural generators responsible for electrical potential detected on the scalp sensors [4, 14]. The algorithms proposed in recent years to solve the ESI problem within distributed dipoles paradigm can be further summarized into three categories in general, which are the (1) Bayesian framework [15–20], (2) state-space based algorithms [21–25], (3) models using sparse representation technique [13, 26–31].

In many sensory or cognitive studies, the underlying activated cortical responsible for the signal processing are relatively focal and thus sparse, which makes the spatial sparse constrain not only mathematically necessary but also neurologically reasonable [1]. One widely accepted assumption is that a sparse spatial structure is favored than a complicated source configuration to explain the same data [32]. An representative early pioneering work is the ℓ_2-norm based minimum norm estimate (MNE) inverse solver [33]. Based on MNE algorith, Pascual-Marqui et al. later proposed standardized low-resolution brain electromagnetic tomography (sLORETA) [34] that enforces spatial smoothness of the neighboring sources and normalizes the solution with respect to the estimated noise level. Some algorithms proposed to use combined multiple solvers, e.g., Weighted minimum norm-LORETA (WMN-LORETA) which combines the LORETA solver and a weighted minimum norm to compensate for deeper sources originate from the subcortical regions[35]. As the above mentioned algorithms are based on ℓ_2-norm to different extents, the estimated source area is over-diffuse. By replacing ℓ_2-norm by ℓ_1-norm, minimum current estimate (MCE) [36] is proposed to overcome overestimation of active area sizes incurred by ℓ_2-norm. Recent development of compressive sensing algorithm proved the ℓ_p ($p \leq 1$) regularization on the original source signal usually provides a set of discrete sources distributed across

the cortex due to high coherence of lead field matrix, in order to encourage reconstruction of extended source patches, it has been found that by enforcing sparsity in a transformed domain, e.g., total variation (TV) regularization [29, 37–39], focal source extents can be better estimated.

The aforementioned algorithms estimate source location at each time point independently, leading to discrepancy along the time direction. To encourage temporal smoothness, a number of regularization techniques based on spatiotemporal mixed norms have been developed, including the famous Mixed Norm Estimates (MxNE) which uses $\ell_{1,2}$-norm regularization [27], and time-frequency mixed-norm estimate (TF-MxNE) which uses structured sparse priors in time-frequency domain for better estimation of the non-stationary and transient source signal [26]. The disadvantage of MxNE is imposing ℓ_2 norm can not guarantee the temporal smoothness in neighboring time points when the source signal is corrupted, while the limitation of TF-MxNE is the lost of high temporal resolution by using of STFT Gabor matrix as a dictionary for time-frequency coefficients. Due to the existence strong spontaneous background source activations, discriminative source activation pattern corresponding to different cognitive tasks which provide more insights shall be reconstructed, Liu et al proposed to use label information of EEG data to find those discriminative source activation pattern [40, 41].

One common drawback of the existing ESI algorithms is that they only consider noises on the sensor level and ignore the spurious noise from the cortex. If perfectly reconstructed, the estimated source is aggregated by task-related source and spurious noise in the source space. The true task-related sources will be corrupted by spurious sources, which motivates us to develop new algorithms to find true task-related source. There are two commonly accepted assumptions (1) spatially sparse (2) temporally continuous for the task-related source activation pattern, which inevitably leads to the low-rank property of the source space. To better discover the task-related source, we impose the low-rank term in the goal function as we consider it is a more direct constraint for spatial sparse. Also, we use a more direct penalty term to temporal smoothness, which is to penalize dissimilarity of temporally neighboring samples based on manifold graph embedding. It is worth-noting that we used the graph regularization term in our previous paper, however the graph is defined to be fully connected for all the points within one class [40], which inevitably drive all the activate patterns at different time points having the same magnitude, thus making the previously defined graph regularization term rely on a strong assumption and limit its future application for realistic cases.

In this paper, we propose a novel EEG source imaging model based on temporal graph regularized low-rank representation. The model is solved based on the alternating direction method of multipliers (ADMM) [42]. We conducted extensive numerical experiments to verify the effectiveness on discovering task related low-rank sources. The reconstructed solution is temporally smooth and spatially sparse. The

contributions of our paper are summarized as follows:

1. We propose to consider the noise not only in the sensor level, but also in the source space.
2. A low-rank representation model (LRR) is proposed for the first time on EEG inverse problem inspired by the low-rank property of true task-related source configurations.
3. We redefined graph embedding regularization based on our previous paper that utilizes temporal vicinity information of samples to promote temporal smoothness.
4. A new algorithm based on ADMM is given which is capable of extracting the low-rank task-related source patterns.

2. Inverse Problem and Temporal Graph Structures

In this section, we briefly review the inverse problem and then discuss the design of temporal graph regularization.

2.1. The Inverse Problem

The cortex source activations propagate to EEG sensors through a linear mapping matrix called lead field matrix, and it can be described as the following linear model:

$$X = LS + E \tag{1}$$

where $X \in \mathbb{R}^{N_c \times N_t}$ is the EEG data measured at a set of N_c electrodes for N_t time points, $L \in \mathbb{R}^{N_c \times N_d}$ is the lead field matrix which maps the source signal to sensors on the scalp, each column of L represents the electrical field of a particular source to the EEG electrodes, $S \in \mathbb{R}^{N_d \times N_t}$ represents the corresponding driving potential in N_d sources locations for the N_t time instants. Since the number of sources is much larger than electrodes, solving S given X is ill-posed with infinite feasible solutions, which necessitates a regularization term to be imposed. Generally, an estimate of S can be found by minimizing the following cost function, which is composed of a quadratic error and a regularization term:

$$\arg\min_{S} \|E\|_F^2 + \gamma\Theta(S) \quad s.t. \quad X - LS = E, \tag{2}$$

where $\|\cdot\|_F$ is the Frobenius Norm. The penalty term $\Theta(S)$ is to encourage neurophysiologically plausible solutions. The regularization term take the form of ℓ_2, ℓ_1 or mixed norm. For example, spatially smooth formulation as in LORETA estimation or spatially sparse formulation with Least Absolute Shrinkage and Selection Operator (LASSO) estimate.

2.2. Temporal Graph Embedding

An important assumption on the source signal is that two temporal adjoint data points should have similar intrinsic activation pattern. In computer vision community, a lot of manifold learning methods have been proposed to find intrinsic similar structure on low-dimensional sub-manifolds embedded in a high dimensional ambient space, such as locally linear embedding [43], Locality Preserving Projection [44], Neighborhood Preserving Embedding [45]. A graph can be viewed as geometric neighborhood relationship between each vertex representing each data sample, the weight between vertex represents similarity between two points [46]. Inspired by the manifold theory [47] and work from Liu et al [48], we use a regularization term to penalize the difference of two neighboring source signal. In our previous work, we use a graph regularization term to promote intra-class consistency [40], but the assumption is too strong by requiring all the reconstructed sources at different time points has the same location as well as signal magnitude as long as they belong to the same class. Now define a temporal graph regularization as

$$R_t(S) = \sum_{i,j=1}^{N} \|s_i - s_j\|_2^2 W_{ij}, \tag{3}$$

where s_i is the i-th column of the matrix S, and a binary matrix W is designed as follows

$$W_{ij} = \begin{cases} 1, & \text{if } s_i \in N_k(s_j) \text{ or } s_j \in N_k(s_i) \\ 0, & \text{otherwise.} \end{cases}$$

The graph embedding matrix W contains temporal vicinity information. $N_k(s_i)$ is the set containing k temporally closest points to s_i. This formulation intends to force neighboring source signal having similar pattern. The benefits are twofold, one is for temporal smoothness, another advantage is to make the reconstructed source denoised for intermittent spurious source activates. By defining D as a diagonal matrix whose entries are row sums of the symmetric matrix W, i.e., $D_{ii} = \sum_j W_{ij}$, and denoting $G = D - W$, $R_t(S)$ can be rewritten as:

$$\begin{aligned} R_t(S) &= \sum_{i,j=1}^{N} ({s_i}^T s_i + {s_j}^T s_j - 2{s_i}^T s_j) W_{ij} = \sum_{i}^{N} s_i^T s_i w_{ii} - \sum_{i,j=1}^{N} s_i^T s_j w_{ij} \\ &= 2\,\mathrm{tr}(SDS^T) - 2\,\mathrm{tr}(SWS^T) = 2\,\mathrm{tr}(SGS^T), \end{aligned} \tag{4}$$

where $\mathrm{tr}(\cdot)$ is the trace operator of a matrix, i.e., adding up all diagonal entries of a matrix.

3. Proposed EEG Source Imaging Model

Before we present our low-rank model with temporal graph structures, we comment on the limitations of traditional model and come up with the decomposition of

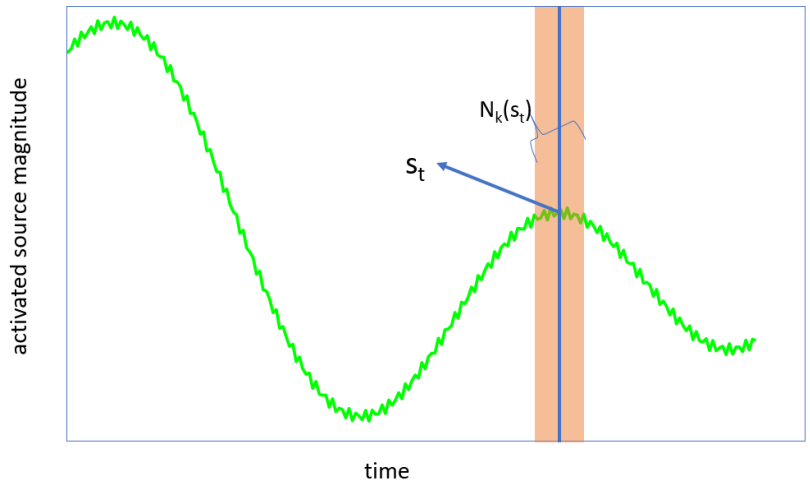


Figure 1: Illustration of temporal smoothness. By design the temporal graph matrix G, the reconstructed signal should have consistent pattern within the same neighborhood window.

task-related source with low-rank property and spontaneous non-task-related spurious sources that is sparsely distributed spatially and with transient patterns. We come up with a graph regularized low-rank representation model and detailed discussion on the purpose of each term in the goal function.

3.1. Decomposition of True and Spurious Sources

In general, two types of noises should be considered, one originates from inaccurate measurement of the sensors modeled by Guassian white noise, which is denoted as E in Eq.(1), the other type of noise is called biological noise that comes directly from the spontaneous activations in the source space, which are not task-related and termed as spurious source. The second types of noise (spurious source) also contributes to the EEG signal in the same way as ground truth source. A drawback of traditional models is that they didn't distinguish the spurious sources from the true sources, and the estimated source can be composed of both task-rated source and spurious sources. To address the above mentioned problem, we propose to use a decomposed source spaces, composed of a low-rank source space and spurious sources originates from spontaneous biological noises. The illustration of decomposition of source space as well as the whole procedure is given in Fig.2, where S_1 has a low-rank property and S_2 is sparse, and sum of S_1 and S_2 is no long low-rank, making X lose low-rank structure.

3.2. Basic Low-Rank Representation (LRR) Model

We argue that during a short period of task-related Evoked Repose Potential (EPR), the number of corresponding activated sources is sparse and remain activated

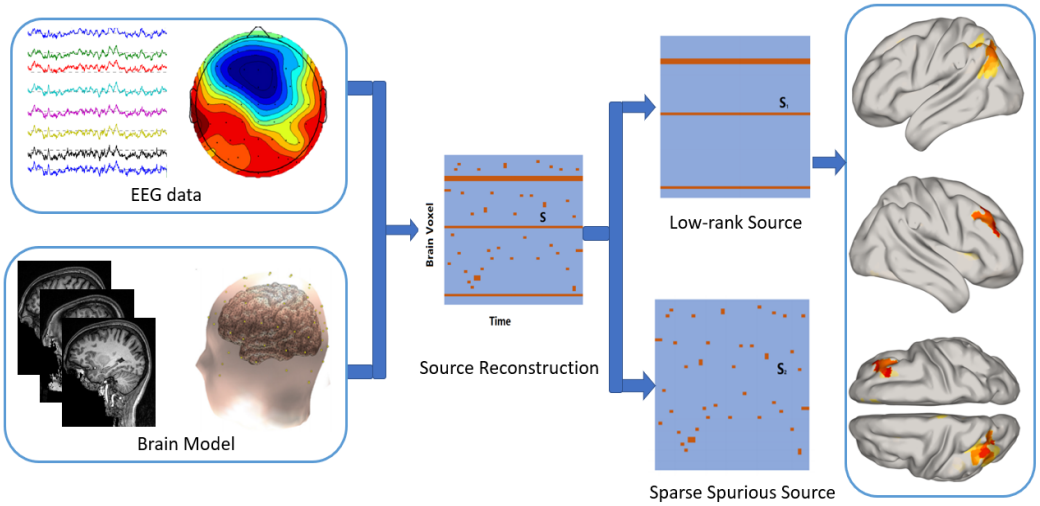


Figure 2: Extraction of low-rank true source from spurious source pipeline: After gathering the MRI scans of the head, tissue segmentation is conducted followed by mesh generation. By assigning conductivity values to different tissues and electrodes co-registered with the meshing model, boundary element method (BEM) was used to solve the forward model. Each triangle represents a brain source, the direction of the source is assumed to be perpendicular to the triangular surface. With EEG data and forward brain model, source reconstruction is calculated. The factual source signal S can be decomposed into two source matrix. The task related true sources S_1 have a low-rank property and the spurious sources S_2 are the sparse but not temporally consistent. The low-rank source solution is projected to cortex voxels to illustrate the activation pattern.

during this period of time, which makes the EPR source matrix low-rank with most rows being zero. However, the brain signal is known to be heavy noisy, spontaneous activations are also active from different source spots. The low-rank representation model for the EEG inverse problem is introduced as follows:

$$\begin{aligned} \min_{S,E} \ & \operatorname{rank}(S) + \beta\|S\|_{1,1} + \lambda\|E\|_{1,1} \\ \text{s.t.} \quad & X = LS + E, \end{aligned} \tag{5}$$

where β, λ and γ are positive scalars to balance the rank function, sparsity cost of sources and the reconstruction error. It is pointed out that ℓ_1 on the error term is more robust to outliers, we use ℓ_1 in the row-rank model [49]. The $\|E\|_{1,1}$ is defined as $\sum_j \sum_i |E_{ij}|$. Due to the discrete nature of the rank function, it is a common practice to use a surrogate nuclear norm $\|\cdot\|_*$ instead. The goal function is given as:

$$\begin{aligned} \min_{S,E} & \|S\|_* + \beta\|S\|_{1,1} + \lambda\|E\|_{1,1} \\ \text{s.t.} \quad & X = LS + E, \end{aligned} \tag{6}$$

The above formulation is a simple version trying to estimate the task-related source activation pattern by using low-rank constraint. To promote the temporal

smoothness, the Laplacian graph structure is included in the next section followed by the optimization algorithm.

3.3. LRR Model with Graph Regularization

Incorporating the previous temporal graph structures, we introduce our proposed model called low-rank representation with temporal graph structures ESI (LRR-TG-ESI). The model is composed of data fitting term to explain the EEG data, temporal graph embedding regularization term that promote temporal smooth, and a ℓ_1 norm for sparsity penalty and nuclear norm for the low-rank structure of ground true source. By combining the low-rank prior and the temporal graph regularization, we propose the following model for ESI:

$$\begin{aligned} &\min_{S,E} \|S\|_* + \lambda\|E\|_{1,1} + \beta\|S\|_{1,1} + \alpha\,\mathrm{tr}(SGS^T) \\ &\text{s.t.} \quad X = LS + E, \end{aligned} \tag{7}$$

where $\lambda, \beta, \alpha > 0$ are tuning parameters to balance the trade-off of different terms. Our proposed model is able to enforce row-sparse via low-rank prior and temporal smoothness via temporal graph regularization while fitting the EEG data X. Compared to earlier works on the ESI problem, both the low-rank prior and graph regularization is novel, although the graph regularization term has been discussed in our early paper [40], but it is not defined on the temporal manifold, and the previous definition in [40] make the magnitude of source signal to be equal intra-class, which is not realistic in real world. To further consider the spatial smoothness, a total variation term can be imposed as another penalty term, such as first order total variation (TV) regularization in Ref.[38, 50], fractional order TV in [37, 39], and similar algorithm can be derived under the framework of ADMM, however further investigation with constraints of TV is our future work.

4. Numerical Algorithm

To solve (7), an algorithm in the ADMM framework is developed. The augmented Lagrangian function of (7) is

$$\begin{aligned} L(S,M,E,T_1,T_2,\mu) &= \|S_1\|_* + \lambda\|E\|_{1,1} + \beta\|M\|_{1,1} + \alpha\,\mathrm{tr}(SGS^T) \\ &+ \langle T_1, X - LS - E\rangle + \langle T_2, M - S\rangle + \frac{\mu}{2} \times (\|X - LS - E\|_F^2 + \|M - S\|_F^2) \end{aligned} \tag{8}$$

By some simple algebra, (8) can be reformulated as

$$\begin{aligned} L(S,M,E,T_1,T_2,\mu) =& \|S\|_* + \lambda\|E\|_{1,1} + \beta\|M\|_{1,1} + \alpha\,\mathrm{tr}(SGS^T) + \frac{\mu}{2} \times (\|X - LS - E + \frac{T_1}{\mu}\|_F^2 \\ &+ \|M - S + \frac{T_2}{\mu}\|_F^2) - \frac{1}{2}\mu(\|T_1\|_F^2 + \|T_2\|_F^2), \end{aligned} \tag{9}$$

where T_1 and T_2 are Lagrangian multipliers and μ is a parameter for the augmented Lagragian term. The variables are updates alternately in a Gauss-Seidel manner by minimizing the augmented Lagrangian function, with other variables fixed. For symbolic simplicity, we rewrite Eq.(9) into the following form:

$$\begin{aligned} L(S, M, E, T_1, T_2\mu) &= \|S\|_* + \lambda\|E\|_{1,1} + \beta\|M\|_{1,1} \\ &+ h(S, E, M, T_1, T_2, \mu) - \frac{1}{2}\mu(\|T_1\|_F^2 + \|T_2\|_F^2), \end{aligned} \tag{10}$$

where

$$h(S, E, M, T_1, T_2, \mu) = \alpha\,\mathrm{tr}(SGS^T) + \frac{\mu}{2} \times (\|X - LS - E + \frac{T_1}{\mu}\|_F^2 + \|M - S + \frac{T_2}{\mu}\|_F^2).$$

If the augmented Lagrangian function is difficult to minimize with respect to a variable, a linearized approximate surrogate function can used, hence the algorithm bears the name Linearized Alternating Direction method [46, 51]. Updating S by minimizing $h(S, E^k, M^k, T_1^k, T_2^k, \mu^k)$ (suppose we are at iteration k) is equivalent to minimize the following goal function with the other variables fixed:

$$L_S = \|S\|_* + h(S, E^k, M^k, T_1^k, T_2^k, \mu^k), \tag{11}$$

which is approximated by optimizing its linearizion at S_1^k plus a quadratic proximal term, given as:

$$S = \underset{S}{\mathrm{argmin}}\|S\|_* + \left\langle \nabla_S h(S^k), S - S^k \right\rangle + \frac{\eta}{2}\left\|S - S^k\right\|_F^2. \tag{12}$$

Here η is a constant satisfying

$$\eta > 2\alpha\|G\|_2 + \mu(1 + \|L\|_2^2), \tag{13}$$

where $\|\cdot\|_2$ is the spectral norm of a matrix, i.e, the largest singular value. As long as (13) is satisfied, (12) is a good approximate to (11). The solution to (12) has a closed form using a singular value thresholding operator (SVT) [52] given as:

$$S^{k+1} = \Theta_{\eta^{-1}}(S^k - \nabla_S h(S^k)/\eta), \tag{14}$$

where $\Theta_\varepsilon(A) = US_\varepsilon(\Sigma)V^T$ is the SVT operator, in which $U\Sigma V^T$ is the singular value decomposition of A and $S_\varepsilon(s)$ is defined as $\sin(x)\max(|x| - \varepsilon, 0)$. $\nabla_{S_1} h(S_1^k)$ is calculated as

$$\nabla_S h(S^k) = \alpha(S^k G + S^k G^T) + \mu L^T(LS - X + E - \frac{T_1}{\mu}) + \mu(S - M - \frac{T_2}{\mu}) \tag{15}$$

To update M and E, it is equivalent to solve the following problem:

$$\underset{M}{\mathrm{argmin}}\,\frac{\mu}{2}\|M - S + \frac{T_2}{\mu}\|_F^2 + \beta\|M\|_{1,1} \tag{16}$$

Algorithm 1 Source Imaging Based on Spatial and Temporal Graph Structures

INPUT: Lead field matrix L, preprocessed EEG signal matrix X, graph matrix G, precalculated matrix D_α, parameters $\alpha, \zeta > 0$, and $\beta > 0$.
OUTPUT: Source matrix S.
Initialize: Set $S = S_0, J = 0, M = 0$.
while not converged **do**
  update S according to
  $S^{k+1} = \tilde{\Theta}_{\eta^{-1}}(S^k - \nabla_S h(S^k)/\eta)$,
  update M according to Equation (16),
  update E according to Equation (17),
  update T_1, T_2 according to Equation (19),
  update $\mu = \min(\rho\mu, \mu_{max})$,
end while

$$\underset{E}{\operatorname{argmin}} \frac{\mu}{2}\|X - LS - E + \frac{T_1}{\mu}\|_F^2 + \lambda\|E\|_{1,1} \tag{17}$$

The general form of (16)–(17) is a ℓ_1 norm proximal operator defined as

$$\operatorname{prox}_\mu(V) = \arg\min_X \mu\|X\|_{1,1} + \frac{1}{2}\|X - V\|_F^2, \tag{18}$$

with $\mu > 0$. The above problem (18) has a closed form solution, called soft thresholding, defined by a shrinkage function,

$$\operatorname{shrink}(V, \mu) = (|V| - \mu)_+ \operatorname{sgn}(V),$$

where $(x)_+$ is x when $x > 0$, otherwise 0. The shrinkage function is defined as element-wise operator. Problem (16)–(17) has a close form solution described with the shrinkage function. After updating all the variables, these Lagrange multipliers are updated by

$$\begin{aligned} T_1 &= T_1 + \mu(X - LS - E) \\ T_2 &= T_2 + \mu(M - S). \end{aligned} \tag{19}$$

The parameter μ is updated by $\mu = \min(\rho\mu, \mu_{max})$. A summarized algorithm is given as Algorithm 1. We initialize the S with the estimate S_0 from ℓ_1 solver.

It's worth noting that the data fitting term we use is $\ell_{1,1}$ norm of E in the model, and there are other options. Generally, if the Gaussian noise E is small, then the norm $\|E\|_F$, is an appropriate choice, but for random data corruption, $\ell_{1,1}$ should be used, and for sample specific data corruption, $\ell_{2,1}$ [53–55], should be used. It has been shown that $\ell_{2,1}$ is more robust to large outliers in some samples. Although the norm used in Algorithm 1 is $\ell_{1,1}$, it can be extended to $\ell_{2,1}$ norm of E, where $\|E\|_{2,1}$ is defined as

$$\|E\|_{2,1} = \sum_{i=1} \sqrt{\sum_{j=1} E_{ij}^2}.$$

In stead of solving (17) to update E, the following goal function (20) needs to be solved to update E.

$$\underset{E}{\operatorname{argmin}} \frac{\mu}{2}\|X - LS - E + \frac{T_1}{\mu}\|_F^2 + \lambda\|E\|_{2,1} \tag{20}$$

By substituting $K = X - LS + \frac{T_1}{\mu}$, if E^* is the optimal solution of

$$\min_E \frac{\lambda}{\mu}\|E\|_{2,1} + \frac{1}{2}\|E - K\|_F^2 \tag{21}$$

Based on the Lemmas in Ref[56], the solution to (21) is

$$e_i^* = \begin{cases} \frac{\|k_i\|_2 - \lambda/\mu}{\|k_i\|_2} k_i, & \text{if } \lambda/\mu \leqslant \|k_i\|_2, \\ 0, & \text{otherwise,} \end{cases}$$

where e_i^* and k_i is the i-th column of matrix E^* and K respectively.

Convergence: The convergence of Algorithm 1 can be easily derived from [51]. Even though the update of M and E is separated in Algorithm 1, they can be combined in one step to become a larger block step and simultaneously solving for (M, E) which is the same case described by LADMAP algorithm [51]. The convergence analysis in [51] can be applied to our case, thus the algorithm convergence is guaranteed [46].

5. Numerical Experiments

In this section, we conducted several experiments to illustrate the effectiveness of our proposed method. Since both the nuclear norm and the graph regularization is relative new for the ESI problem, we started from the simple model (6) to help the readers understand the property and impacts of low-rank prior along with data fidelity term and sparsity term for source reconstruction. In the first experiment, we did a comprehensive exploration for different parameter settings by varying the weights between low-rank term, data fitting term and sparsity term. In the second experiment, we illustrate the temporal smoothing functionality of the graph regularization term for uncorrupted smooth source and corrupted source with abrupt signal jumps. In the third experiment, we give comprehensive numerical results by testing our algorithm against the benchmark algorithms to showcase the effectiveness of the proposed method in reconstructing task-related source, where we show that our algorithm can not only find the activated locations, but also reconstruct the time-course of source activation with high precision.

5.1. Head Model

A realistic head model, referred to as New York Head model [57], is used in our numerical experiment. The New York Head model is based on highly detailed MRI

images derived ICBM152 anatomy, which is a nonlinear average of the T1-weighted structural MRI of 152 adults and calculated with state-of-the-art finite element electrical modeling. The New York Head model is considered to be highly accurate since it considers six tissue types when conducting segmentation, which is scalp, skull, CSF, gray matter, white matter, air cavities, with a native MRI resolution of 0.5 mm^3. The dimension of the lead field matrix used in the numerical experiment part is 2004 by 108, representing a linear mapping from 2004 sources to 108 electrodes.

5.2. Experiment 1: Test on Simple Low-rank Model

To understand the property of low-rankness, we start from a simple model to help the readers understand the property of low-rank in the EEG source imaging and the validity of low-rank when recovering the accurate location as well as time course of source signal. Like in [58], eight octants are divided as regions of interest (ROI) are considered, which are Right Anterior Inferior (RAI), Right Anterior Superior (RAS), Right Posterior Inferior (RPI), Right Posterior Superior (RPS), Left Anterior Inferior (LAI), Left Anterior Superior (LAS), Left Posterior Inferior (LPI), Left Posterior Superior (LPI). In the simple experiments, we selected 2 different ROIs and randomly select one activated voxel in each of these ROIs, and a $4th$ order moving average time series is generated, as illustrated in Fig.3.

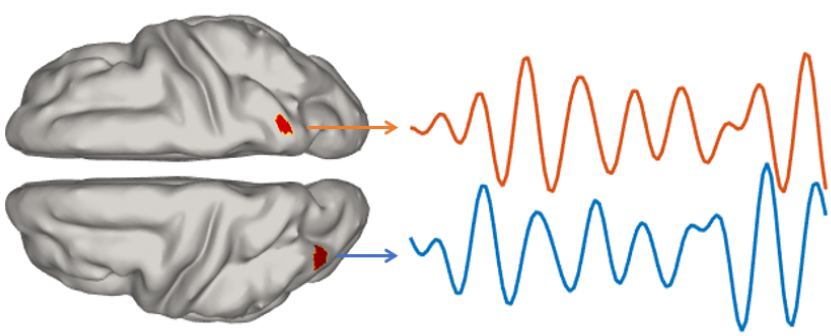


Figure 3: Illustration of two activated sources time series on two different ROIs.

At each location, a time series with length of 500 were generated to represent the source activation time-course. At each time point, two randomly picked sources are activated to simulate the non-task related spurious noise with mean of 0 and variance to be 1. The task-related activate pattern has low-rank property, however, the noise corrupted source space is no longer low-rank. We repeated our experiment 50 times for all the combinations of λ and β, where $\lambda = \{0.01, 0.02, 0.03, 0.05, 0.1, 0.2, 0.5\}$ and $\beta = \{0.005, 0.01, 0.015, 0.02, 0.05, 0.1\}$. The reconstructed error (RE) metric used here is

$$\mathrm{RE} = \|\hat{S} - S\|_2 / \|S\|_2, \tag{22}$$

where $\hat{S}$ represents the reconstructed source. We define the SNR $= 10\log_{10}\frac{P_s}{P_n}$, where P_s and P_n are the power of signal and noise respectively. The violin plot is used to visualize the distribution of reconstruction errors, in corresponding to β and λ respectively in Fig.4 and Fig.5. Increasing β will penalize the strength of signal and make the reconstruction error to be large when λ is small. However, When λ is set to be 0.5, the weight of data fidelity is high, thus driving the solution have a better data fidelity, which can balance off the increase of β.

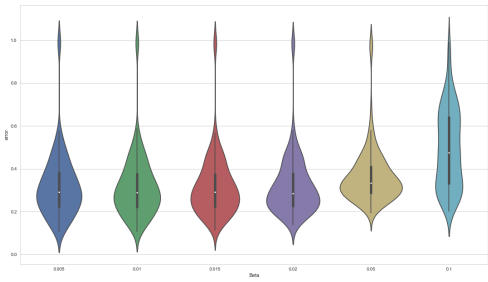


Figure 4: Reconstruction errors varying β.

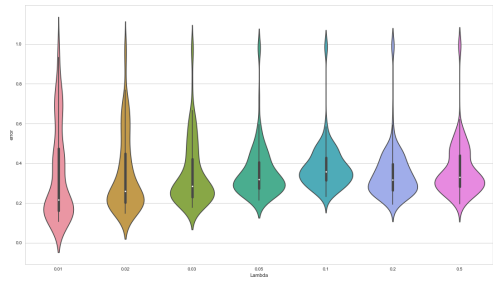


Figure 5: Reconstruction errors varying λ.

The averaged reconstruction error of 50 experiments for all of β and λ is given in Fig.6a and the averaged rank for the estimated source for all the combination of β and λ is given in Fig.6b. As we increase the value of λ, the rank is also increasing, which underlies the trade-off between explaining the data and finding the latent low-rank structure of the source space. Increasing λ means more weight on the data fitting term, and the spurious source can also be recovered, thus increasing the rank of the reconstructed source.

To empirically understand and explain the reconstruction error in Fig.6a, we visualize the curve fitting performance for truth source time series and the reconstructed source time-course corresponding to different level of reconstruction error. We picked the curve fitting cases when the reconstructed error is 0.2, 0.4, 0.8 and 1 respectively. For error equal to 0.2, we picked one experiment and plot the fitting of time series which demonstrated a very good curve fitting of the ground truth vs the reconstructed one as is shown in part (a) of Fig.7. Also we picked another experiment whose error equal to 0.4 and it is shown in part (b) of Fig.7. The curve fitting is slightly worse than the previous one but it is hard to notice the difference compared to the previous one, however according to our error metric, the error is 0.4. One thing to notice is that in both situations, the reconstructed rank equal to the ground truth rank, which is equal to 2, moreover, the estimated two active source location is exactly the ground true locations. We also examined the case when error is up to 0.8, and the curve fitting plot is given in part (c) of Fig.7. We examined this happened when we set $\lambda = 0.01$ and $\beta = 0.1$. In this case, the penalty for sparsity is 0.1, which means too much penalty for the sparse term and the signal magnitude is reduced by the shrinkage operation. We also notice than in the experiment, the error can be up to

1 no matter what parameter settings are given, as can be seen in the top region of Fig.4 and Fig.5. The curve fitting plot for the failed case is given in part (d) of Fig.7.

Although the situation is very rare when $RE = 1$, we want to visualize the activation pattern on the cortex to see how discrepant the reconstructed location compared to the true source location. For comparison, Fig. 8 is given when $RE = 0.4$ and $RE = 1$. The reconstructed source locations on two ROIs are exactly the same with the ground true location when $RE = 0.4$, when $RE = 1$, one source location is reconstructed perfectly while the other source location is not accurately located, however the neighboring sources are reconstructed.

To check how the rank of S evolve during the iterations, the boxplot of the rank at selected steps are given in Fig.9. Starting from an initial value with high rank, the rank of $\hat{S}$ is decreasing as the iteration proceeds. We set the maximum rank to be 20, during the iteration process, the rank of S is converged to very small number for most of the cases.

5.3. Experiments 2: Test LRR with Temoral Graph Prior

In this section, we solve the LRR-TG-ESI problem (7) with graph regularization term to test its impact on the reconstructed signal. Under the same setting with Experiment 1, we assign different values $\{0.01, 0.02, 0.05, 0.1, 0.5\}$ for the graph regularization parameter α. The original source signal was smooth, then it was corrupted with randomly number at some time points. There are also 2 randomly picked activated sources representing spurious sources with mean of 0 and variance to be 1. The "temporal smoothing" impact of the graph regularization is shown in Fig.10, where $\lambda = 0.02$ and $\beta = 0.01$. In Fig.10, the original signal is corrupted and not smooth at some time points, we set the neighbor size to be 2 (the closest signal before and after the one to be estimated) when calculating the Laplacian matrix. It is evident from the formulation (3) that the graph regularization term will decrease the dissimilarity of temporally neighbored reconstructed source. If α is set to be 0.5, the graph regularization term penalized heavily on the curvature of the reconstructed signal as is illustrated in Fig.10. We can see that with the temporal graph prior, the reconstructed source is more smooth. It is worth noticing that the main purpose of temporal graph prior is not to smooth the time course for the activated locations, the main purpose is to filter out the spurious activations that are short transients with abrupt jumps. Combined with the low-rank prior, the temporal graph prior can filter the spurious activations and reconstruct the task related activated source. The randomly planted spurious sources are filter out by penalizing the graph regularization and nuclear norm, and in most of the cases, the final rank is 2 can be achieved within a wide range of parameters. The time series plots of original EEG signal, corrupted EEG signal, and EEG signal recovered from the reconstructed sources from (7) by setting $\lambda = 0.02$ and $\beta = 0.01$, as well as the topoplots at 42 ms is given in Fig.11. To illustrate again the impact of the graph regularization, α is chosen from $\{0.01, 0.02, 0.05, 0.1, 0.5\}$. The first row is the EEG data generated from the

task-related sources (persistent and low-rank), the second row is the EEG data from the task-related sources and the spurious source, the SNR is -0.72 dB, which means the energy of spurious source is slightly larger than the task-related sources. The 3rd row is the EEG data calculated from forward model after the source is reconstructed from our proposed model with $\alpha = 0.01$. The 4th-7th row is the time series plots when $\alpha = 0.02, 0.05, 0.1, 0.5$ respectively. The topoplots on the right part of Fig.11 is are sampled from 42 ms of the EEG data on the left.

5.4. Experiments 3: Comprehensive Comparison with Benchmark Algorithms

The purpose of previous numerical experiments is to validate each term of the goal function and to understand their properties. The trade-off between low-rankness, data fidelity, sparsity, temporal smooth are fully discussed by varying different parameters. In this part, a comprehensive study is conducted to compare the proposed algorithm with the popular ESI algorithms such as MNE [33], sLORETA [34], MCE and well as the state-of-art algorithm mixed-norm estimate (MxNE) [27].

We generated independent sources in different ROIs for easy validation purpose, the number of independent sources varied from 2 to 4 corresponding to different rank of ground-truth source, and the number of spurious source are generated by randomly activating the sources on the cortex with a random scalar whose mean value to be 0, and the variance is 1. Moreover, the noise on sensor level is also added to the EEG data. Two of the MCE algorithms are selected, which are Homotopy and FISTA [59]. For MxNE algorithm, we choose $\ell_{2,1}$ to enforce ℓ_2 norm on the temporal time series of each voxel and ℓ_1 norm to impose spatial sparsity.

To measure the performance, we introduced 5 metrics, including 1) CPU time in seconds, 2) rank of the calculated source, 3) Sparsity, measuring the number of nonzero elements in the source space at each time point, 5) Reconstruction Error (RE) defined in Eq.(22), 5) Localization Error (LE), which is calculated using the shortest path algorithm over the irregular meshes from the reconstructed source location to the ground truth location. The LE metric is the most important one, since it measures the discrepancy in location, the other metrics give information of the property of the rendered solution. To calculate LE for each ROI with activated sources, we first locate source with the largest activation magnitude in this ROI, and calculate the shortest path distance from the located source to the ground truth location. We conduct the same procedure for all the activated ROI, and calculate the average value of all the distances at each time point. The final LE is the averaged distance value for all the 500 time points for each experiments.

All the algorithms are implemented in Matlab except MxNE which we call as a Python script (MultiTaskLasso.py) under the linear model of scikit-learn library[60] from Matlab. The formulation for MxNE is

$$||X - LS||_F^2 + \gamma ||S||_{2,1} \tag{23}$$

We found that parameter tuning process for MxNE algorithm is very time consuming, even though there is only one parameter for the unweighted version. If the parameter γ in Eq.(23) is tuned for one experiment with good performance, and we use the same setting of parameter and the same setting to generate random source activation patterns, the reconstructed source matrix can be a zero matrix. We tuned the parameter γ according to the best LE performance when the rank is 2 and the number of spurious activated source is 2. By conducting experiments for parameter tuning of MxNE, we set γ in Eq.(23). For our proposed algorithm, we set $\lambda = 0.01$ and $\beta = 0.01$, which were tested to have good performance for the same case when the rank is 2 and the number of spurious activated source is 2, and the graph parameter α is also set to be 0.01. 10 experiments were conducted under the same setting and the performance of all the algorithms are summarized in Table 1 to Table 3. The SNR is calculated after the noise signal is generated and it was averaged from 10 experiments under the same experiment setting. As can be seen from the tables, our algorithm is the most accurate to locate the task related activated source. The CPU time of our algorithm is between Homotopy and FISTA algorithm. The running time for Python version of MxNE is much faster than our proposed algorithm, but there are many cases that MxNE algorithm failed, thus making the overall accuracy drop significantly. Although our algorithm have more parameter, it is a more sophisticated model that allows better controllability to customize the weight of different terms in the goal function. Although MxNE has large RE and LE, it is still worth noting that the MxNE algorithm we used is a simple version with $\ell_{2,1}$ norm discussed in Ref.[27], and the algorithm to solve the goal function 23 is coordinate descent, other algorithms can be tested to solve the same problem. MCE model solve using ℓ_1 algorithms Homotopy and FISTA can render good LE accuracy, and Homotopy outperforms FISTA in all the experiment with better speed, which confirms the comparison discussed in [40][56].

Table 1: Source Reconstruction Performance Comparison (Rank=2)

Method	Rank =2; SNR= -0.356 dB					Rank =2; SNR= -1.12 dB					Rank =2; SNR= -1.67 dB				
	Time	Rank	Sparsity	RE	LE	Time	Rank	Sparsity	RE	LE	Time	Rank	Sparsity	RE	LE
Homotopy	0.41	449	65	0.80	2.02	0.41	449.8	84.4	0.95	3.31	0.50	459.3	91.15	0.97	4.57
FISTA	2.76	500	1812.4	0.62	3.93	2.76	500	1920.4	0.89	9.39	2.73	500	1888.0	0.80	11.25
sLORETA	0.055	500	2004	1.18	33.5	0.05	500	2004	1.26	34.13	0.056	500	2004.0	1.32	38.31
WMN	8.4e-5	107	2004	0.99	22.6	8.4e-5	107	2004	0.99	31.47	8.4e-5	107	2004.0	0.99	28.46
MxNE	0.026	7.9	7.9	0.95	42.1	0.023	10	10	0.97	45.79	0.028	11.8	11.8	1.04	68.80
Proposed	0.75	3.3	3.3	0.24	**0.145**	0.75	4.8	4.8	0.414	**1.94**	0.83	6.4	6.5	0.42	**2.27**

Table 2: Source Reconstruction Performance Comparison (Rank=3)

Method	Rank =3; SNR= 0.938 dB					Rank =3; SNR= -0.174 dB					Rank =3; SNR= -0.784 dB				
	Time	Rank	Sparsity	RE	LE	Time	Rank	Sparsity	RE	LE	Time	Rank	Sparsity	RE	LE
Homotopy	0.36	445.4	77.6	0.64	5.18	0.48	464	87.2	0.68	6.1	0.46	474.1	92.5	0.84	5.88
FISTA	2.64	500	1812.5	0.69	9.33	2.56	500	1969.0	1.23	21.5	2.67	500	1939.2	0.94	17.44
sLORETA	0.057	500	2004	1.90	41.93	0.06	500	2004	1.81	43.9	0.05	500	2004.0	1.25	46.60
WMN	8.4e-5	107	2004	6.79	29.55	7.6e-5	107	2004	6.82	35.63	7.3e-5	107	2004.0	0.99	34.41
MxNE	0.026	9.2	9.2	0.91	64.95	0.028	10	10	0.98	67.95	0.028	12.2	12.2	0.96	57.49
Proposed	0.72	5.2	5.2	0.43	**2.98**	0.80	6.1	6.1	0.488	**3.82**	0.75	7.3	7.3	0.45	**2.86**

Table 3: Source Reconstruction Performance Comparison (Rank=4)

Method	Rank = 4; SNR= 1.04 dB					Rank =4; SNR= 0.603 dB					Rank =4; SNR= −0.179 dB				
	Time	Rank	Sparsity	RE	LE	Time	Rank	Sparsity	RE	LE	Time	Rank	Sparsity	RE	LE
Homotopy	0.45	448	78.7	0.57	3.77	0.41	467.3	87.9	0.62	5.87	0.46	475.1	93.15	0.62	5.35
FISTA	2.97	500	1978.7	1.03	16.54	2.46	500	1964.5	1.01	16.64	2.45	500	1961.9	1.35	19.4
sLORETA	0.058	500	2004	1.99	45.3	0.058	500	2004	1.72	49.26	0.050	500	2004.0	1.71	49.38
WMN	7.8e-5	107	2004	7.43	31.0	7.8e-5	107	2004	6.63	33.49	7.8e-5	107	2004.0	6.66	38.00
MxNE	0.024	11.9	11.9	0.997	68.36	0.023	11.7	11.7	0.947	56.42	0.025	13.9	13.9	0.99	61.46
Proposed	0.79	5.9	6.8	0.38	**2.24**	6.9	7.8	7.8	0.43	**3.90**	0.76	10.1	14.3	0.44	**4.5**

6. Conclusion

In this paper, we propose to consider the noise not only on the sensor level, but also in the source space. we come up with an EEG source imaging model based on temporal graph structures and low-rank representation. The model is solved with our proposed algorithm based on ADMM. Numerical experiments are conducted to verify the effectiveness of the proposed work on discovering task related low-rank sources. We delineate the discussion on the properties and impacts for each term in the cost function to help better understand our proposed graph regularized low rank representation model. Compared the traditional model, our proposed one can find the intrinsic task related activation patterns and suppress the spurious source patterns.

7. Acknowledgment

This work been partially supported by the NSF funding under grant number CMMI-1537504 and DMS-1522786. The authors are also grateful to Dr. Stefan Haufe and Dr. Yu Huang for providing the head model.

References

[1] B. Babadi, G. Obregon-Henao, C. Lamus, M. S. Hämäläinen, E. N. Brown, and P. L. Purdon, "A subspace pursuit-based iterative greedy hierarchical solution to the neuromagnetic inverse problem," *NeuroImage*, vol. 87, pp. 427–443, 2014.

[2] A. Custo, S. Vulliemoz, F. Grouiller, D. Van De Ville, and C. Michel, "Eeg source imaging of brain states using spatiotemporal regression," *Neuroimage*, vol. 96, pp. 106–116, 2014.

[3] C. Li, D. Jacobs, T. Hilton, M. del Campo, Y. Chinvarun, P. L. Carlen, and B. L. Bardakjian, "Epileptogenic source imaging using cross-frequency coupled signals from scalp eeg," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 12, pp. 2607–2618, 2016.

[4] C. M. Michel, M. M. Murray, G. Lantz, S. Gonzalez, L. Spinelli, and R. G. de Peralta, "EEG source imaging," *Clinical neurophysiology*, vol. 115, no. 10, pp. 2195–2222, 2004.

[5] P. Mégevand, L. Spinelli, M. Genetti, V. Brodbeck, S. Momjian, K. Schaller, C. M. Michel, S. Vulliemoz, and M. Seeck, "Electric source imaging of interictal activity accurately localises the seizure onset zone," *J Neurol Neurosurg Psychiatry*, vol. 85, no. 1, pp. 38–43, 2014.

[6] C. J. Honey, E. L. Newman, and A. C. Schapiro, "Switching between internal and external modes: A multi-scale learning principle," *Network Neuroscience*, no. Early Access, pp. 1–18, 2017.

[7] K. Mahjoory, V. V. Nikulin, L. Botrel, K. Linkenkaer-Hansen, M. M. Fato, and S. Haufe, "Consistency of EEG source localization and connectivity estimates," *NeuroImage*, vol. 152, pp. 590–601, 2017.

[8] M. Hassan, O. Dufor, I. Merlet, C. Berrou, and F. Wendling, "EEG source connectivity analysis: from dense array recordings to brain networks," *PloS one*, vol. 9, no. 8, p. e105041, 2014.

[9] F. Liu, W. Xiang, S. Wang, and B. Lega, "Prediction of seizure spread network via sparse representations of overcomplete dictionaries," in *International Conference on Brain and Health Informatics*. Springer, 2016, pp. 262–273.

[10] A. Anwar, M. Muthalib, S. Perrey, A. Galka, O. Granert, S. Wolff, U. Heute, G. Deuschl, J. Raethjen, and M. Muthuraman, "Effective connectivity of cortical sensorimotor networks during finger movement tasks: a simultaneous fnirs, fmri, eeg study," *Brain topography*, vol. 29, no. 5, pp. 645–660, 2016.

[11] A. Mheich, M. Hassan, O. Dufor, M. Khalil, and F. Wendling, "Combining eeg source connectivity and network similarity: Application to object categorization in the human brain," *arXiv preprint arXiv:1606.01017*, 2016.

[12] R. Grech, T. Cassar, J. Muscat, K. P. Camilleri, S. G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, and B. Vanrumste, "Review on solving the inverse problem in EEG source analysis," *Journal of neuroengineering and rehabilitation*, vol. 5, no. 1, p. 1, 2008.

[13] W. Ou, M. S. Hämäläinen, and P. Golland, "A distributed spatio-temporal EEG/MEG inverse solver," *NeuroImage*, vol. 44, no. 3, pp. 932–946, 2009.

[14] R. Leahy, J. Mosher, M. Spencer, M. Huang, and J. Lewine, "A study of dipole localization accuracy for meg and eeg using a human skull phantom," *Electroencephalography and clinical neurophysiology*, vol. 107, no. 2, pp. 159–173, 1998.

[15] S. Baillet and L. Garnero, "A bayesian approach to introducing anatomo-functional priors in the EEG/meg inverse problem," *IEEE transactions on Biomedical Engineering*, vol. 44, no. 5, pp. 374–385, 1997.

[16] Y. Bekhti, F. Lucka, J. Salmon, and A. Gramfort, "A hierarchical bayesian perspective on majorization-minimization for non-convex sparse regression: application to M/EEG source imaging," *arXiv preprint arXiv:1710.08747*, 2017.

[17] D. Wipf and S. Nagarajan, "A unified bayesian framework for meg/eeg source imaging," *NeuroImage*, vol. 44, no. 3, pp. 947–966, 2009.

[18] F. Costa, H. Batatia, L. Chaari, and J.-Y. Tourneret, "Sparse EEG source localization using bernoulli laplacian priors," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 12, pp. 2888–2898, 2015.

[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, 2017.

[20] K. Liu, Z. L. Yu, W. Wu, Z. Gu, Y. Li, and S. Nagarajan, "Bayesian electromagnetic spatio-temporal imaging of extended sources with markov random field and temporal basis expansion," *NeuroImage*, vol. 139, pp. 385–404, 2016.

[21] M. J. Barton, P. A. Robinson, S. Kumar, A. Galka, H. F. Durrant-Whyte, J. Guivant, and T. Ozaki, "Evaluating the performance of kalman-filter-based eeg source localization," *IEEE transactions on biomedical engineering*, vol. 56, no. 1, pp. 122–136, 2009.

[22] Y. Yang, E. Aminoff, M. Tarr, and K. E. Robert, "A state-space model of cross-region dynamic connectivity in MEG/EEG," in *Advances in Neural Information Processing Systems*, 2016, pp. 1234–1242.

[23] C. Long, P. Purdon, S. Temereanca, N. Desai, M. Hamalainen, and E. Brown, "Large scale kalman filtering solutions to the electrophysiological source localization problem-a meg case study," in *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE.* IEEE, 2006, pp. 4532–4535.

[24] O. Yamashita, A. Galka, T. Ozaki, R. Biscay, and P. Valdes-Sosa, "Recursive penalized least squares solution for dynamical inverse problems of eeg generation," *Human brain mapping*, vol. 21, no. 4, pp. 221–235, 2004.

[25] C. Lamus, M. S. Hämäläinen, S. Temereanca, E. N. Brown, and P. L. Purdon, "A spatiotemporal dynamic distributed solution to the meg inverse problem," *NeuroImage*, vol. 63, no. 2, pp. 894–909, 2012.

[26] A. Gramfort, D. Strohmeier, J. Haueisen, M. S. Hämäläinen, and M. Kowalski, "Time-frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations," *NeuroImage*, vol. 70, pp. 410–422, 2013.

[27] A. Gramfort, M. Kowalski, and M. Hämäläinen, "Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods," *Physics in medicine and biology*, vol. 57, no. 7, p. 1937, 2012.

[28] D. Strohmeier, Y. Bekhti, J. Haueisen, and A. Gramfort, "The iterative reweighted mixed-norm estimate for spatio-temporal MEG/EEG source reconstruction," *IEEE transactions on medical imaging*, vol. 35, no. 10, pp. 2218–2228, 2016.

[29] A. Sohrabpour, Y. Lu, G. Worrell, and B. He, "Imaging brain source extent from EEG/MEG by means of an iteratively reweighted edge sparsity minimization (ires) strategy," *NeuroImage*, vol. 142, pp. 27–42, 2016.

[30] M. Zhu, W. Zhang, D. L. Dickens, and L. Ding, "Reconstructing spatially extended brain sources via enforcing multiple transform sparseness," *NeuroImage*, vol. 86, pp. 280–293, 2014.

[31] S. Haufe, V. V. Nikulin, A. Ziehe, K.-R. Müller, and G. Nolte, "Combining sparsity and rotational invariance in EEG/MEG source reconstruction," *NeuroImage*, vol. 42, no. 2, pp. 726–738, 2008.

[32] S. Castaño-Candamil, J. Höhne, J.-D. Martínez-Vargas, X.-W. An, G. Castellanos-Domínguez, and S. Haufe, "Solving the EEG inverse problem based on space–time–frequency structured sparsity constraints," *NeuroImage*, vol. 118, pp. 598–612, 2015.

[33] M. S. Hämäläinen and R. J. Ilmoniemi, "Interpreting magnetic fields of the brain: minimum norm estimates," *Medical & biological engineering & computing*, vol. 32, no. 1, pp. 35–42, 1994.

[34] R. D. Pascual-Marqui *et al.*, "Standardized low-resolution brain electromagnetic tomography (sloreta): technical details," *Methods Find Exp Clin Pharmacol*, vol. 24, no. Suppl D, pp. 5–12, 2002.

[35] C. Song, T. Zhuang, and Q. Wu, "Hybrid weighted minimum norm method a new method based loreta to solve EEG inverse problem," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. IEEE, 2006, pp. 1079–1082.

[36] K. Uutela, M. Hämäläinen, and E. Somersalo, "Visualization of magnetoencephalographic data using minimum current estimates," *NeuroImage*, vol. 10, no. 2, pp. 173–180, 1999.

[37] Y. Li, J. Qin, Y.-L. Hsin, S. Osher, and W. Liu, "s-SMOOTH: Sparsity and smoothness enhanced EEG brain tomography," *Frontiers in Neuroscience*, vol. 10, p. 543, 2016.

[38] L. Ding, "Reconstructing cortical current density by exploring sparseness in the transform domain," *Physics in Medicine and Biology*, vol. 54, no. 9, p. 2683, 2009.

[39] J. Qin, F. Liu, S. Wang, and J. Rosenberger, "EEG source imaging based on spatial and temporal graph structures," in *International Conference on Image Processing Theory, Tools and Applications*, 2017.

[40] F. Liu, J. Rosenberger, Y. Lou, R. Hosseini, J. Su, and S. Wang, "Graph regularized eeg source imaging with in-class consistency and out-class discrimination," *IEEE Transactions on Big Data*, vol. 3, no. 4, pp. 378–391, Dec 2017.

[41] F. Liu, S. Wang, J. Rosenberger, J. Su, and H. Liu, "A sparse dictionary learning framework to discover discriminative source activations in EEG brain mapping." in *AAAI*, 2017, pp. 1431–1437.

[42] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[43] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[44] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using laplacianfaces," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 3, pp. 328–340, 2005.

[45] X. He, D. Cai, S. Yan, and H.-J. Zhang, "Neighborhood preserving embedding," in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 2. IEEE, 2005, pp. 1208–1213.

[46] M. Yin, J. Gao, and Z. Lin, "Laplacian regularized low-rank representation and its applications," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 3, pp. 504–517, 2016.

[47] D. Cai, X. He, J. Han, and T. S. Huang, "Graph regularized nonnegative matrix factorization for data representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1548–1560, 2011.

[48] K. Liu, Z. L. Yu, W. Wu, Z. Gu, and Y. Li, "Straps: A fully data-driven spatio-temporally regularized algorithm for M/EEG patch source imaging," *International journal of neural systems*, vol. 25, no. 04, p. 1550016, 2015.

[49] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 171–184, 2013.

[50] V. Michel, A. Gramfort, G. Varoquaux, E. Eger, and B. Thirion, "Total variation regularization for fmri-based prediction of behavior," *IEEE transactions on medical imaging*, vol. 30, no. 7, pp. 1328–1340, 2011.

[51] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," in *Advances in neural information processing systems*, 2011, pp. 612–620.

[52] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[53] F. Nie, H. Huang, X. Cai, and C. H. Ding, "Efficient and robust feature selection via joint $\ell_{2,1}$-norms minimization," in *Advances in neural information processing systems*, 2010, pp. 1813–1821.

[54] S. Du, Y. Ma, and Y. Ma, "Graph regularized compact low rank representation for subspace clustering," *Knowledge-Based Systems*, vol. 118, pp. 56–69, 2017.

[55] M. Yin, J. Gao, Z. Lin, Q. Shi, and Y. Guo, "Dual graph regularized latent low-rank representation for subspace clustering," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 4918–4933, 2015.

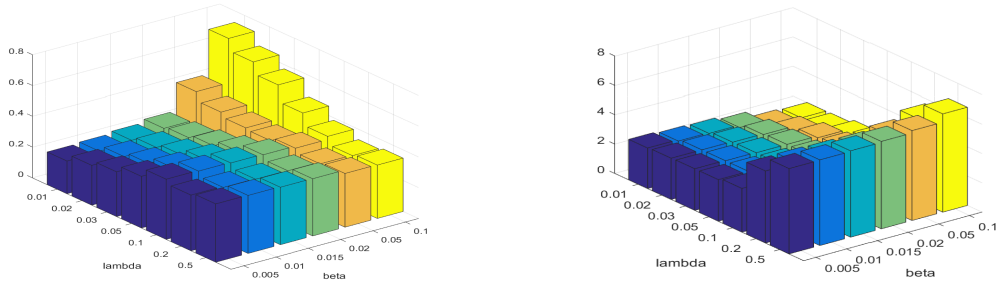
[56] J. Yang, W. Yin, Y. Zhang, and Y. Wang, "A fast algorithm for edge-preserving variational multichannel image restoration," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 569–592, 2009.

[57] Y. Huang, L. C. Parra, and S. Haufe, "The New York Head – a precise standardized volume conductor model for EEG source localization and tES targeting," *NeuroImage*, vol. 140, pp. 150 – 162, 2016, transcranial electric stimulation (tES) and Neuroimaging.

[58] S. Haufe and A. Ewald, "A simulation framework for benchmarking EEG-based brain connectivity estimation methodologies," *Brain topography*, pp. 1–18, 2016.

[59] A. Y. Yang, S. S. Sastry, A. Ganesh, and Y. Ma, "Fast ℓ 1-minimization algorithms and an application in robust face recognition: A review," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010, pp. 1849–1852.

[60] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, no. Oct, pp. 2825–2830, 2011.



(a) (b)

Figure 6: Averaged reconstruction error and rank varying λ and β over 50 experiments.(a) average of reconstruction error for different λ and β (b) average of rank for the source matrix S. Increasing λ means more weight on data fidelity and the rank becomes higher. Our model works well with a wide range of parameter setting.

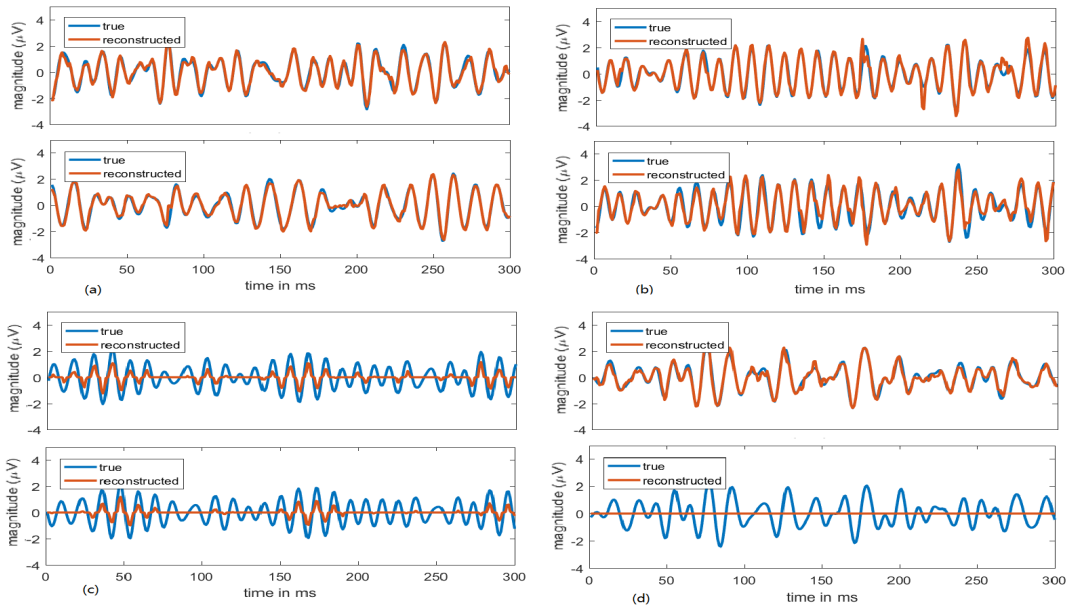
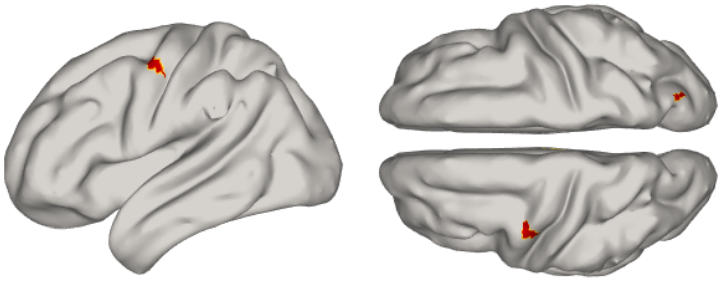
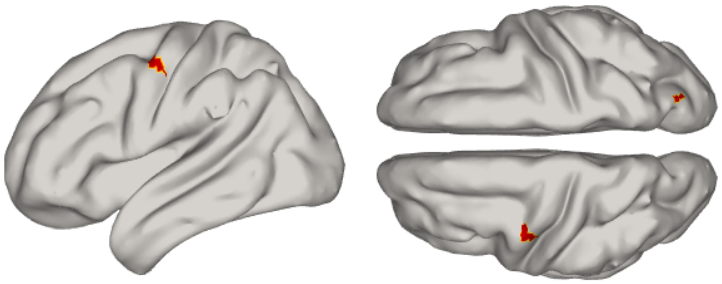


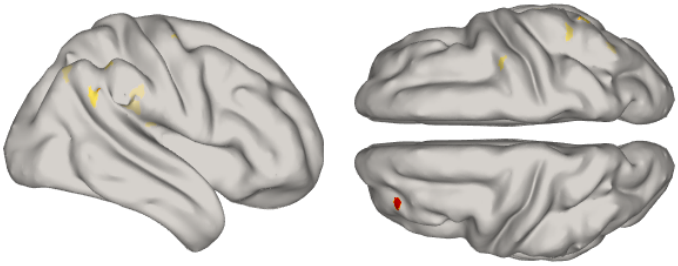
Figure 7: Source time course fitting illustration: (a): ground truth time course vs reconstructed for two activated source at different ROIs in one experiment when fitting error equal to 0.2. Here $\lambda = 0.02$ and $\beta = 0.005$.(b): ground truth time course vs reconstructed for two activated source when fitting error equal to 0.4. Here $\lambda = 0.5$, $\beta = 0.01$. (c): ground truth time course vs reconstructed for two activated source when fitting error equal to 0.8. Here $\lambda = 0.01$ and $\beta = 0.1$. (d): ground truth time course vs reconstructed for two activated source when fitting error equal to 1. Here $\lambda = 0.02$ and $\beta = 0.01$. The curve fitting of (b) is slightly worse than (a), corresponding to the RE= 0.2 and RE= 0.4. When the sparsity parameter is set too large, the reconstructed magnitude is smaller than the ground truth as is shown in (c). For some cases, only the time course in one source location is reconstructed shown in (d) with RE to be 1.0.



(a)



(b)



(c)

Figure 8: (a) Ground truth source activation pattern (b) Reconstructed source activation when 0.4. This plot illustrates the perfect localization of ground true sources. (c) Reconstructed source activation when RE is 1. This plot illustrates when our algorithm failed to recover one of the the exact locations of two activated sources, while the other one can be recovered perfectly. the pictures on the left is the reconstructed location, but still close to the ground truth.

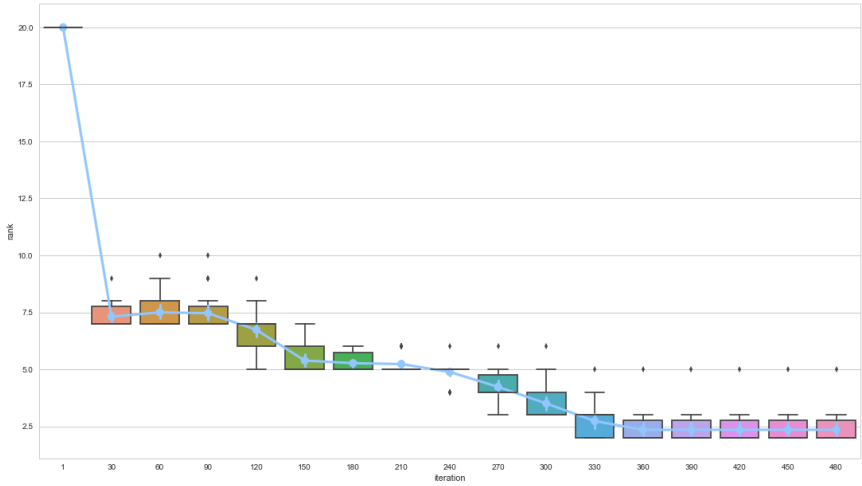


Figure 9: Convergence of rank of S during iteration procedure: In the first iteration, we set the maximum rank to be 20, and in most case the rank will converge to 2.

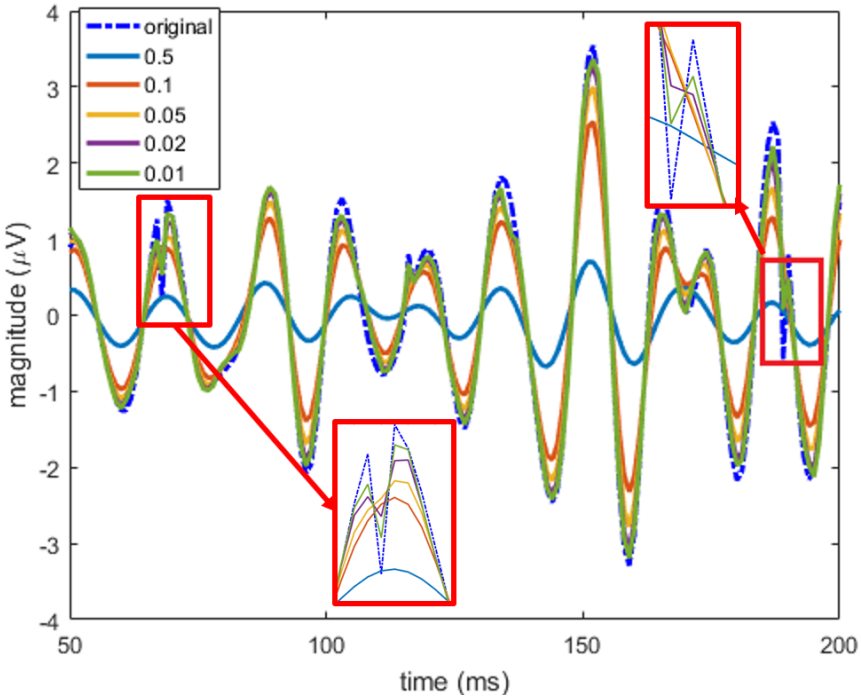


Figure 10: Illustration of the smoothing effect of temporal graph regularization: reconstructed time courses from varied graph regularization parameters.

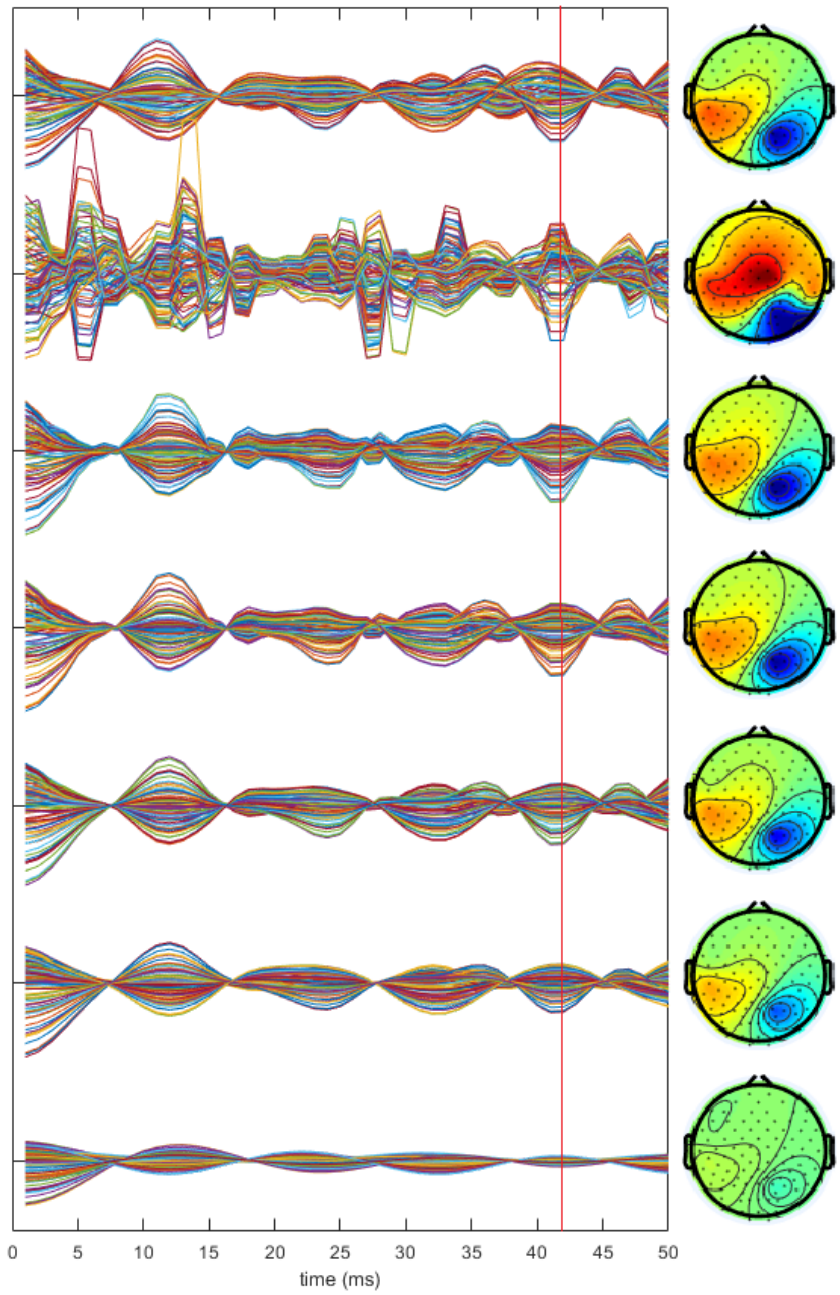


Figure 11: EEG time series plot of the uncorrupted EEG signal, corrupted EEG signal vs the reconstructed EEG signal using the proposed method and the corresponding topoplots at 42 ms: The 1st row is the time series plot of the original uncorrupted EEG data, the 2nd row is the plots for corrupted EEG data, the 3rd row is the EEG data reconstructed by applying our algorithm with $\alpha = 0.01$, the 4th row is the reconstructed EEG data with $\alpha = 0.02$, the 5th row is the reconstructed EEG data with $\alpha = 0.05$, the 6th row is the reconstructed EEG data with $\alpha = 0.10$, the 7th row is the reconstructed EEG data with $\alpha = 0.5$. The spurious source in the source space corrupted the task-related EEG data, and by using our graph regularized LRR model, the true EEG data can be recovered.