Neural Computing and Applications (2021) 33:8513–8524
https://doi.org/10.1007/s00521-020-05603-1

ORIGINAL ARTICLE

# EEG extended source imaging with structured sparsity and $L_1$-norm residual

Furong Xu[1] · Ke Liu[1] · Zhuliang Yu[2,3] · Xin Deng[1] · Guoyin Wang[1]

Received: 10 June 2020 / Accepted: 11 December 2020 / Published online: 8 January 2021
© The Author(s), under exclusive licence to Springer-Verlag London Ltd. part of Springer Nature 2021

## Abstract

It is a long-standing challenge to reconstruct the locations and extents of cortical neural activities from electroencephalogram (EEG) recordings, especially when the EEG signals contain strong background activities and outlier artifacts. In this work, we propose a robust source imaging method called $L_1$R-SSSI. To alleviate the effect of outliers in EEG, $L_1$R-SSSI employs the $L_1$-loss to model the residual error. To obtain locally smooth and globally sparse estimations, $L_1$R-SSSI adopts the structured sparsity constraint, which incorporates the $L_1$-norm regularization in both the variation and original source domain. The estimations of $L_1$R-SSSI are efficiently obtained using the alternating direction method of multipliers (ADMM) algorithm. Results of simulated and experimental data analysis demonstrate that $L_1$R-SSSI effectively suppresses the effect of the outlier artifacts in EEG. $L_1$R-SSSI outperforms the traditional $L_2$-norm-based methods (e.g., wMNE, LORETA), and SISSY, which employs $L_2$-norm loss and structured sparsity, indicated by the larger AUC (average AUC $> 0.80$), smaller SD (average SD $<50$ mm), DLE (average DLE $<10$ mm) and RMSE (average RMSE $<1.75$) values under all the numerically simulated conditions. $L_1$R-SSSI also provides better estimations of extended sources than the method with $L_1$-loss and $L_p$-norm regularization term (e.g., LAPPS).

**Keywords** EEG source imaging · Outliers · Structured sparsity · ADMM

## 1 Introduction

Reconstructing brain activities from electroencephalography (EEG) plays an important role in neuroscience research and clinical treatment [2, 15, 34]. For example, for drug-resistant epilepsy, the epileptogenic zone can be removed through a surgical intervention. For this surgical planning, the precise delineation of the epileptic focuses is critical for surgeons [4].

✉ Ke Liu
liuke@cqupt.edu.cn

✉ Guoyin Wang
wanggy@cqupt.edu.cn

[1] Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

[2] School of Automation Science and Engineering, South China University of Technology, Guangzhou 510641, China

[3] Pazhou Lab, Guangzhou 510335, China

EEG source imaging (ESI) is to estimate cortical active areas from the scalp EEG signals, which is a highly ill-posed inverse problem with infinite solutions [15, 16]. To solve this inverse problem, there are mainly two classes of source models. One is the equivalent current dipole model (ECD), and the other is the distributed current density model (DCD). ECD uses a few dipoles to approximate cortical activities, which can estimate the locations of focal cortical activities [15, 20]. However, ECD provides little information of source extents.

DCD divides the cortical surface into several triangular grids, each of which represents a dipole, and a large number of fixed dipoles in the brain represent the continuous distribution of current activity [7, 15, 16]. The position of each dipole is fixed, so cortical activities can be estimated by solving a linear inverse problem. Since the dipoles largely outnumber the scalp sensors, the forward equation of DCD is underdetermined [2, 16, 34]. To obtain a unique solution, suitable constraints are needed to narrow the solution space.

One common DCD ESI solver is the $L_2$-norm-based methods, such as minimum norm estimate (MNE) [15, 16]. Since the source signals on the cortical surface are closer to the scalp sensors and easier to be detected by the sensors, MNE is biased toward superficial sources [15]. To overcome depth bias of MNE, the weighted minimum norm estimate (wMNE) weights the dipoles at different locations using the norm of the columns of lead-field matrix [23]. To consider the dependencies between adjacent sources, low resolution brain electromagnetic tomography (LORETA) penalizes the $L_2$-norm of the second-order spatial derivative of sources and obtains spatial coherent and smoothness solutions [15]. Although the $L_2$-norm-based methods are computationally efficient, they suffer from too diffused estimations, covering most areas of the cerebral cortex, beyond the actual cortical active area.

To improve the spatial resolution of reconstructed sources, the $L_p$ $(p \leq 1)$-norm regularization and sparse Bayesian learning are employed to obtain sparse solutions [10, 24, 33, 37]. Although the sparse constraint methods provide good estimations for focal cortical activities, it provides little information of the sizes for sources with large extents. Hence, some studies applied the sparsity constraint to transform domains instead of the original source domains to reconstruct the source extents [8, 22, 38]. Variation transform, which computes the difference between neighbored dipoles, was first reported to use transform sparseness [9]. By using variation sparseness, the variation-based sparse cortical current density (VB-SCCD) achieved better estimation results for extended sources [9].

The above ESI methods assume that the measurement noise satisfies Gaussian distribution and uses $L_2$-norm to fit the residual error. However, during EEG recordings, the EEG signals are inevitably contaminated by outliers (e.g., background noise of the measurement experiment and some artificial noise caused by the blink, head movement, etc.) [5, 29]. The $L_2$-norm is sensitive to these outliers. To tackle these outliers, least absolute $L_p$ penalized solution (LAPPS) [6] used the $L_1$-norm to fit residual to reduce the effect of outliers. Experiments show that LAPPS has better performance than methods based on $L_2$-norm residual. Nevertheless, due to the $L_p$-norm regularization term, LAPPS severely underestimates the extents for large size sources. Thus, our former work [36] proposed the $L_1$-norm loss and $L_1$-norm regularization of the variation sources, which provided accurate estimations for extended sources.

Nevertheless, the minimization of the variation transform (first-order derivative) does not constrain the global energies and often underestimates the source amplitudes. In this work, to reconstruct the locations, extents and amplitudes of cortical activities, especially when the EEG signals contain outlier artifacts, we propose a robust ESI method, named $L_1$-norm Residual and Structured Sparsity-based Source Image ($L_1$R-SSSI). Specifically, we employ $L_1$ loss to reduce the noise sensitivity. To estimate source extents, $L_1$R-SSSI applies $L_1$-norm regularization based on variation transform to obtain globally sparse and locally smooth solutions. Moreover, to alleviate the underestimated amplitudes due to variation sparseness, we further add $L_1$-norm regularization term of the original source. Alternating direction method of multipliers (ADMM) algorithm [18] is employed to efficiently estimate the sources. Both numerical and experimental data analyses validate the superior performance of the proposed algorithm compared to the conventional methods.

The remaining of the paper is organized as follows. In Sect. 2, we derive the $L_1$R-SSSI algorithm. In Sect. 3, we present the simulation design and evaluation metrics. In Sect. 4, we compare the performance of $L_1$R-SSSI method with the benchmark algorithms in simulated and real EEG data, followed by a brief discussion and summary in Sects. 5 and 6.

## 2 Methods

### 2.1 Background

Usually, the relationship between EEG signals and cortical sources can be described as [15]

$$\boldsymbol{b} = \boldsymbol{L}\boldsymbol{s} + \boldsymbol{\varepsilon} \tag{1}$$

where $\boldsymbol{b} \in \mathbb{R}^{d_b \times 1}$ is the EEG signal measured on $d_b$ sensors. $\boldsymbol{s} \in \mathbb{R}^{d_s \times 1}$ is the unknown source vector with $d_s$ candidate brain sources. $\boldsymbol{L} \in \mathbb{R}^{d_b \times d_s}$ is the lead-field matrix, which describes the relationship between EEG and cortical sources of the $d_s$ candidate locations. $\boldsymbol{\varepsilon}$ is the measurement noise.

Since the number of EEG electrodes is much less than the number of unknown cortical sources (i.e., $d_b \ll d_s$), Eq. (1) is seriously undetermined. And there exist infinite solutions for the inverse problem. To obtain unique solution, prior constraints or regularization terms are necessary to narrow the solution space. Typical constraints are the $L_2$-norm and $L_p$-norm $(p \leq 1)$ regularization terms. Evidences from other neuroimaging techniques, such as fMRI and ECoG [1], have revealed the compact nature of cortical activations, i.e., the sources are locally smooth and globally clustered. The conventional $L_2$-norm and $L_p$-norm-based methods provide little information of source extents. To reconstruct extended patches, several studies have employed the $L_1$-norm constraint in the transform domains, such as the variation transform [9, 22], which describes the

first-order differences of amplitudes between adjacent dipoles. For variation sparseness, a variation operator of cortical source $\boldsymbol{V} \in \mathbb{R}^{P\times d_s}$ is defined as [9, 22, 38]

$$\boldsymbol{V} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1d_s} \\ v_{21} & v_{22} & \cdots & v_{2d_s} \\ \vdots & \vdots & \ddots & \vdots \\ v_{P1} & v_{P2} & \cdots & v_{Pd_s} \end{bmatrix}$$
$$\begin{cases} v_{pi} = 1, v_{pj} = -1, i<j; \text{if source } i, j \text{ share edge} p \\ \qquad\qquad v_{pi} = 0; \quad \text{otherwise} \end{cases} \tag{2}$$

where $P$ is the number of edges of triangular elements. Each row of matrix $\boldsymbol{V}$ corresponds to a shared edge between two triangular grids. Each row of the variation source $\boldsymbol{u} = \boldsymbol{V}\boldsymbol{s} \in \mathbb{R}^{P\times 1}$ denotes the difference of amplitudes between the two adjacent dipoles. If each active clustered source has a uniform current density distribution, the sparseness largely occurs on the boundaries between active and inactive regions. Therefore, one can recover the sparseness in the variation domain to estimate the extents of cortical activations.

However, the minimization of variation (first-order derivative) sources does not constrain the global energies of inverse solutions and tends to overestimate the source extents [22]. Hence, additional constraints are needed to limit global energies [38]. Becker et al. [4] proposed SISSY to reconstruct the extended sources, which uses the $L_1$-norm regularization in both the variation and original source domains:

$$\hat{\boldsymbol{s}}_{\text{SISSY}} = \arg\min_{\boldsymbol{s}} \|\boldsymbol{L}\boldsymbol{s} - \boldsymbol{b}\|_2^2 + \lambda_1\|\boldsymbol{V}\boldsymbol{s}\|_1 + \lambda_2\|\boldsymbol{s}\|_1 \tag{3}$$

## 2.2 $L_1$R-SSSI algorithm

Typical ESI methods use $L_2$-norm to measure the residual error, which can exaggerate the effect of outliers caused by head movement or eye blinks during recordings [31, 35]. Prior studies have shown that $L_1$-norm loss for outliers is more robust and stable than the $L_2$-norm loss [12, 32]. Hence, we develop a robust ESI method to reconstruct extended sources, namely $L_1$R-SSSI, which uses $L_1$-norm to measure the residual error. To reconstruct the extents, locations and amplitudes of cortical activations, similar as [4], we employ the structured sparsity, which penalizes $L_1$-norm of both variation sources and original sources.

$$\hat{\boldsymbol{s}} = \arg\min_{\boldsymbol{s}} \|\boldsymbol{L}\boldsymbol{s} - \boldsymbol{b}\|_1 + \lambda_1\|\boldsymbol{V}\boldsymbol{s}\|_1 + \lambda_2\|\boldsymbol{s}\|_1 \tag{4}$$

where $\lambda_1 > 0$ and $\lambda_2 > 0$ are regularization parameters to adjust the balance between the residual term and the two regularization terms.

To solve Eq. (4), we rewrite it as

$$\begin{aligned} &\hat{\boldsymbol{s}} = \arg\min_{\boldsymbol{s}} \|\boldsymbol{e}\|_1 + \lambda_1\|\boldsymbol{u}\|_1 + \lambda_2\|\boldsymbol{w}\|_1 \\ &s.t., \boldsymbol{e} = \boldsymbol{L}\boldsymbol{s} - \boldsymbol{b}, \boldsymbol{u} = \boldsymbol{V}\boldsymbol{s}, \boldsymbol{w} = \boldsymbol{s} \end{aligned} \tag{5}$$

The augmented Lagrangian function associated with the optimization problem (5) is

$$\begin{aligned} \mathcal{L}(\boldsymbol{s},\boldsymbol{e},\boldsymbol{u},\boldsymbol{w},\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) = &\|\boldsymbol{e}\|_1 + \lambda_1\|\boldsymbol{u}\|_1 + \lambda_2\|\boldsymbol{w}\|_1 \\ &+ \boldsymbol{x}^\top(\boldsymbol{L}\boldsymbol{s} - \boldsymbol{b} - \boldsymbol{e}) + \frac{\rho_1}{2}\|\boldsymbol{L}\boldsymbol{s} - \boldsymbol{b} - \boldsymbol{e}\|_2^2 \\ &+ \boldsymbol{y}^\top(\boldsymbol{V}\boldsymbol{s} - \boldsymbol{u}) + \frac{\rho_2}{2}\|\boldsymbol{V}\boldsymbol{s} - \boldsymbol{u}\|_2^2 \\ &+ \boldsymbol{z}^\top(\boldsymbol{s} - \boldsymbol{w}) + \frac{\rho_3}{2}\|\boldsymbol{s} - \boldsymbol{w}\|_2^2 \end{aligned} \tag{6}$$

where $\boldsymbol{x} \in \mathbb{R}^{d_b\times 1}$, $\boldsymbol{y} \in \mathbb{R}^{P\times 1}$ and $\boldsymbol{z} \in \mathbb{R}^{d_s\times 1}$ are the Lagrangian multipliers, while $\rho_1 > 0$, $\rho_2 > 0$ and $\rho_3 > 0$ are penalty parameters. By minimizing the Lagrangian functions $\mathcal{L}$ with respect to $(\boldsymbol{s},\boldsymbol{e},\boldsymbol{u},\boldsymbol{w})$, these vectors can be updated alternately:

$$\begin{aligned} \boldsymbol{s}^{(k+1)} &= \left(\rho_1\boldsymbol{L}^\top\boldsymbol{L} + \rho_2\boldsymbol{V}^\top\boldsymbol{V} + \rho_3\boldsymbol{I}\right)^{-1}\boldsymbol{q}^{(k)} \\ \boldsymbol{e}^{(k+1)} &= \mathcal{S}_{\frac{1}{\rho_1}}\left(\boldsymbol{L}\boldsymbol{s}^{(k+1)} - \boldsymbol{b} + \frac{1}{\rho_1}\boldsymbol{x}^{(k)}\right) \\ \boldsymbol{u}^{(k+1)} &= \mathcal{S}_{\frac{\lambda_1}{\rho_2}}\left(\boldsymbol{V}\boldsymbol{s}^{(k+1)} + \frac{1}{\rho_2}\boldsymbol{y}^{(k)}\right) \\ \boldsymbol{w}^{(k+1)} &= \mathcal{S}_{\frac{\lambda_2}{\rho_3}}\left(\boldsymbol{s}^{(k+1)} + \frac{1}{\rho_3}\boldsymbol{z}^{(k)}\right) \end{aligned} \tag{7}$$

where $\boldsymbol{q}^{(k)} = \left[\rho_1\boldsymbol{L}^\top\left(\boldsymbol{b} + \boldsymbol{e}^{(k)}\right) + \rho_2\boldsymbol{V}^\top\boldsymbol{u}^{(k)} + \rho_3\boldsymbol{w}^{(k)} - \boldsymbol{L}^\top\boldsymbol{x}^{(k)} - \boldsymbol{V}^\top\boldsymbol{y}^{(k)} - \boldsymbol{z}^{(k)}\right]$, and $\mathcal{S}_\kappa(a)$ is

$$\mathcal{S}_\kappa(a) = \begin{cases} a - \kappa, & a > \kappa \\ 0, & |a| < \kappa \\ a + \kappa, & a < -\kappa \end{cases}$$

Using the dual ascent method, the corresponding update rules for the Lagrangian multipliers $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ can be deduced:

$$\begin{aligned} \boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)} + \rho_1\left(\boldsymbol{L}\boldsymbol{s}^{(k+1)} - \boldsymbol{b} - \boldsymbol{e}^{(k+1)}\right) \\ \boldsymbol{y}^{(k+1)} &= \boldsymbol{y}^{(k)} + \rho_2\left(\boldsymbol{V}\boldsymbol{s}^{(k+1)} - \boldsymbol{u}^{(k+1)}\right) \\ \boldsymbol{z}^{(k+1)} &= \boldsymbol{z}^{(k)} + \rho_3\left(\boldsymbol{s}^{(k+1)} - \boldsymbol{w}^{(k+1)}\right) \end{aligned} \tag{8}$$

where $x^{(k)}$ denotes the value of $x$ at the $k$th iteration.

By alternatively updating $(\boldsymbol{s},\boldsymbol{e},\boldsymbol{u},\boldsymbol{w})$ and $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$, we can get the solutions of $L_1$R-SSSI. When the maximum number of iteration is reached or the relative change of $\boldsymbol{s}$ reaches a specified tolerance, the iterative procedure stops.

### 2.3 Computational complexity and application details

To determine the computational complexity, we compute the number of floating-point operations (FLOPs, in terms of real-valued multiplications) needed for the completion of $L_1$R-SSSI. $L_1$R-SSSI is based on alternating updates of seven vectors, among which the update of the source vector $\boldsymbol{s}$ is the most complex. The update of $\boldsymbol{s}$ involves the inverse solution of a large matrix $(\rho_1\boldsymbol{L}^\top\boldsymbol{L}+\rho_2\boldsymbol{V}^\top\boldsymbol{V}+\rho_3\boldsymbol{I})^{-1}$, which can be solved by employing the Woodbury inversion lemma:

$$(\boldsymbol{L}^\top\boldsymbol{L}+\boldsymbol{M})^{-1}=\boldsymbol{M}^{-1}-\boldsymbol{M}^{-1}\boldsymbol{L}^\top(\boldsymbol{I}+\boldsymbol{L}\boldsymbol{M}^{-1}\boldsymbol{L}^\top)^{-1}\boldsymbol{L}\boldsymbol{M}^{-1} \tag{9}$$

where $\boldsymbol{M}=\frac{\rho_2}{\rho_1}\boldsymbol{V}^\top\boldsymbol{V}+\frac{\rho_3}{\rho_1}\boldsymbol{I}$, which is a sparse matrix and can be computed at a low computational cost of $O(\frac{3}{2}{d_s}^2)$ and inverted using only $O(4{d_s}^2)$ multiplications [3]. Compared to the multiplication $\boldsymbol{L}\boldsymbol{M}^{-1}\boldsymbol{L}^\top$, which requires $O(d_b d_s^2)$ FLOPs, the computation and inversion of $\boldsymbol{M}$ can be neglected. In practice, the computation of the inverse $(\boldsymbol{I}+\boldsymbol{L}\boldsymbol{M}^{-1}\boldsymbol{L}^\top)^{-1}$ is avoided by resorting to the Cholesky decomposition which requires $O(\frac{1}{6}{d_b}^3)$ real-valued multiplications [3, 4]. Once the hyperparameters $\rho_1$, $\rho_2$ and $\rho_3$ are determined, these operations are performed only once at the beginning of the algorithm. Additionally, at each iteration, the update of the signal vector $\boldsymbol{s}$ requires $O(2d_bd_s+2d_sP+d_s^2)$. Compared to computational cost of the vector $\boldsymbol{s}$, the computational complexity of the other six vectors is small and can be ignored. Hence, the computational complexity of $L_1$R-SSSI is around $O(d_bd_s^2+\frac{1}{6}d_b^3+(2d_bd_s+2d_sP+d_s^2)K)$, where $K$ is the number of iterations.

During our experimental simulations, we used a standard PC (Corei5-8500 CPU 3 GHz and 8 GB RAM) for numerical experiments. $\boldsymbol{e}$, $\boldsymbol{u}$, $\boldsymbol{w}$, $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ were initialized to 0. The regularization parameters $\lambda_1$, $\lambda_2$ were determined by cross-validation [32]. Detailed selection for $\lambda_1$ and $\lambda_2$ is shown in Appendix B. For EEG data with $d_b=62$ electrodes, a source space comprising $d_s=15{,}002$ dipoles, $P=44{,}986$ edges of triangular grids, and stop tolerance $\delta=10^{-4}$, $L_1$R-SSSI converges after about 1000 iterations, which takes less than 4 minutes.

## 3 Simulation design and evaluation metrics

To verify the performance of $L_1$R-SSSI, we carried out a series of Monte Carlo numerical simulations to compare $L_1$R-SSSI with SISSY and the followed algorithms.

(1) wMNE [23], which is a $L_2$-norm-based method to compensate depth bias of MNE.

$$\begin{aligned}\hat{\boldsymbol{s}}_{\text{wMNE}}&=\arg\min_{\boldsymbol{s}}\|\boldsymbol{L}\boldsymbol{s}-\boldsymbol{b}\|_2^2+\lambda\|\boldsymbol{W}\boldsymbol{s}\|_2^2\\&=\left(\boldsymbol{L}^\top\boldsymbol{L}+\lambda\boldsymbol{W}^\top\boldsymbol{W}\right)^{-1}\boldsymbol{L}^\top\boldsymbol{b}\end{aligned} \tag{10}$$

where $\boldsymbol{W}\in\mathbb{R}^{d_s\times d_s}$ is a diagonal matrix, and the $i$th diagonal element is $\boldsymbol{W}_{i,i}=\|\boldsymbol{l}_i\|_2^{-1}$, where $\boldsymbol{l}_i$ is the $i$th column of lead-field matrix $\boldsymbol{L}$.

(2) LORETA [25], which penalizes the second-order spatial derivative to obtain smooth and coherent sources.

$$\begin{aligned}\hat{\boldsymbol{s}}_{\text{LORETA}}&=\arg\min_{\boldsymbol{s}}\|\boldsymbol{L}\boldsymbol{s}-\boldsymbol{b}\|_2^2+\lambda\|\boldsymbol{D}\boldsymbol{s}\|_2^2\\&=\left(\boldsymbol{L}^\top\boldsymbol{L}+\lambda\boldsymbol{D}^\top\boldsymbol{D}\right)^{-1}\boldsymbol{L}^\top\boldsymbol{b}\end{aligned} \tag{11}$$

where $\boldsymbol{D}=\boldsymbol{I}-\boldsymbol{N}$ is the discrete Laplacian operators on the cortex, and $\boldsymbol{N}$ is defined as

$$N_j^i=\begin{cases}\dfrac{1}{|\mathcal{N}_i|} & \text{if} \quad j\in\mathcal{N}_i\\ 0 & \text{others}\end{cases}$$

where $\mathcal{N}_i$ is the source set adjacent to the $i$th source, $|\mathcal{N}_i|$ denotes the amount of elements in set $\mathcal{N}_i$.

(3) LAPPS [6], which employs $L_1$-norm for the residual and $L_p$-norm regularization term.

$$\hat{\boldsymbol{s}}_{\text{LAPPS}}=\arg\min_{\boldsymbol{s}}\|\boldsymbol{L}\boldsymbol{s}-\boldsymbol{b}\|_1+\lambda\|\boldsymbol{s}\|_p^p \tag{12}$$

ADMM was used to obtain the estimations of LAPPS in this work.

For the detailed application of these methods, in this work, we used the Bayesian Minimum-Norm method [7, 21] to determine the regularization parameter $\lambda$ for wMNE and LORETA. As for the LAPPS, we used cross-validation to determine the parameter $\lambda$. Furthermore, as suggested in [11], the nonconvex penalty function with $p=0.5$ yields good performance in sparse analysis. Hence, we set $p=0.5$ for LAPPS in our work. The regularization parameters $\lambda_1$, $\lambda_2$ for SISSY are also determined by cross-validation. To facilitate reproducibility, the codes of $L_1$R-SSSI and the above benchmark methods are available upon requests from the authors.

### 3.1 Simulation design

Due to the lack of ground truth for EEG source imaging, we first validate the performance of the proposed method with numerically simulated EEG data. For the Monte Carlo numerical simulations, we used Brainstorm software [27] to get a three-shell head model based on the default ICBM 152 anatomy. The high-resolution cortical surface was

downsampled to 15,002 triangular grids to obtain the source space for simulation. Each triangle grid indicated a dipole source, and the direction of each dipole was perpendicular to the cortical surface. The lead-field matrix $\boldsymbol{L}$ was calculated using Brainstorm with the sensor configuration of the 64-channel Neuroscan Quik-cap system. Since two channels are not EEG electrodes, $\boldsymbol{L} \in \mathbb{R}^{62\times 15002}$.

To construct an extended source, we randomly selected a seed triangle grid on the cortical surface and iteratively added adjacent triangle grids until the extent of cluster satisfied a specified value. The clean EEG recordings were obtained by multiplying the simulated sources with lead field $\boldsymbol{L}$. We then added Gaussian white noise and artifacts to the clean EEG data to simulate the actual EEG recordings. The noise level is controlled by the signal-to-noise ratio (SNR), which is defined as $\mathrm{SNR} = 10\log_{10}\left[\frac{\sigma^2(\boldsymbol{Ls})}{\sigma^2(\boldsymbol{\varepsilon})}\right]$. where $\sigma^2(\boldsymbol{Ls})$ is the variance of clean EEG data and $\sigma^2(\boldsymbol{\varepsilon})$ is the variance of mixed noise including Gaussian noise and artifacts. Similar to [6], the mixture noise $\boldsymbol{\varepsilon}$ was obtained as

$$\boldsymbol{\varepsilon} = \frac{\boldsymbol{\varepsilon}_1}{\sigma(\boldsymbol{\varepsilon}_1)}\left[10^{\left(-\frac{\mathrm{SNR}}{20}\right)}\right]\sigma(\boldsymbol{Ls}) \tag{13}$$

where $\boldsymbol{\varepsilon}_1 \in \mathbb{R}^{d_b\times 1}$ is the outliers and $\sigma$ is the standard deviation. Each element of $\boldsymbol{\varepsilon}_1$ obeys Gaussian distribution $\mathcal{N}(\mu + 10\sigma^2, \sigma^2)$, where $\mu$ and $\sigma^2$ are the mean and variance of the clean EEG data $\boldsymbol{Ls}$ across all channels. Then, for a given SNR, the mixture measurement noise was generated using Eq. (13) and added to the clean EEG data $\boldsymbol{Ls}$. During the numerical simulations, three scenarios were tested:

(1) One source with different extents (i.e., 0.5, 5, 10, 15, 30 $\mathrm{cm}^2$).
(2) Various number of patch sources (i.e., from 1 to 4), where the extents of each patch was around 10 $\mathrm{cm}^2$.
(3) EEG signals with different SNRs (i.e., − 10, − 5, 0, 5, 10 dB), where one cluster of around 10 $\mathrm{cm}^2$ was simulated for each SNR.

Except for the case of various SNRs, we set the SNR for the other two cases to be 5 dB. For each case, we carried out 100 Monte Carlo simulations, which guaranteed that the simulated patches covered most regions of the cortex.

Additionally, $L_1$R-SSSI was further applied to analyze real human data to test its practical use. The dataset includes the simultaneous MEG/EEG recordings of 16 subjects who performed a simple visual task on a large number of famous, unfamiliar and scrambled faces, which can be downloaded from the OpenNeuro website[1]. The MEG data consist of 102 magnetometers and 204 planar gradiometers from an Elekta VectorView system at 1100 Hz [30]. The same system was used to simultaneously record EEG data from 70 electrodes (using a nose reference). In this work, only EEG data were used for analysis.

### 3.2 Evaluation metrics

To quantitatively assess the performance of ESI methods, we employed four evaluation metrics.

(1) The area under the receiver operating characteristic (ROC) curve (AUC), which evaluates the detection sensitivity and specificity of reconstructed sources [14, 22].
(2) Spatial dispersion (SD), which measures the degree of spatial blurriness and dispersion of reconstructed sources [22, 38]. A low SD value indicates that the estimated source solution has less blurriness.
(3) Distance of localization error (DLE), which measures the location error of the recovered sources compared to the ground truth [8, 22].
(4) Relative mean square error (RMSE), which is the relative squared error between the estimated and simulated source activity: $\frac{\|\hat{\boldsymbol{s}}-\boldsymbol{s}\|_2^2}{\|\boldsymbol{s}\|_2^2}$ [14, 38].

A better ESI method is expected to yield a larger AUC, and a lower SD, DLE and RMSE values. Using the four metrics, we can comprehensively analyze the performance of ESI algorithms in various aspects, i.e., detection sensitivity, dispersion, locations, and amplitudes. Detailed computations of the four metrics are presented in the supplementary document of [22]. To assess the significance of results, we employed the Kruskal–Wallis test. If a test statistic from the Kruskal–Wallis test was significant, Wilcoxon rank sum tests were subsequently performed in $L_1$R-SSSI against each of the benchmark algorithms to determine whether $L_1$R-SSSI yielded significantly better estimations. For imaging visualization, we displayed the absolute value of estimated sources at specified time points. The imaging threshold was determined by the Otsu's method [22].

## 4 Results

### 4.1 Results of simulated data analysis

#### 4.1.1 Influence of source extents

We first tested the performance of $L_1$R-SSSI under one simulated cluster source with various extents. Figure 1 shows the performance metrics of wMNE, LORETA, LAPPS, SISSY, $L_1$R-SSSI. When the source extents increased, the AUC values of all algorithms except LORETA showed degraded values. The larger AUC values for

[1] https://openneuro.org/datasets/ds000117.

larger extents indicate that LORETA is suitable to localize extended sources, which is in line with the results in [14]. For wMNE, LORETA, SISSY and $L_1$R-SSSI, the SD, DLE and RMSE values gradually decreased with the increase in source extents. However, RMSE of LAPPS increased significantly, while SD and DLE values changed slightly. For focal source (e.g., 0.5 $cm^2$), LAPPS provided the smallest RMSE, SD, DLE values and large AUC ( $> 0.9$) values, showing that it was suitable for recovering focal sources. Nevertheless, LAPPS was less powerful in reconstructing extended sources, indicated by the largest RMSE and smallest AUC values for sources with large extents (e.g., 30 $cm^2$). For all source extents, $L_1$R-SSSI obtained the largest AUC ($p<0.05$), the smallest SD ($p<0.05$), DLE ($p<0.05$), and RMSE ($p<0.05$) values compared to wMNE, LORETA, and SISSY.

Table 1 shows the results of the Spearman's correlation analysis between the performance metrics of each algorithm and the source extents. The four performance metrics of wMNE, SISSY and $L_1$R-SSSI are significantly negatively correlated with the source extents. In contrast, LORETA's AUC and LAPPS' RMSE values are significantly positively correlated with the source extents. Therefore, as the source extents increase, the AUC values of LORETA and RMSE values of LAPPS will gradually increase. Furthermore, the SD and DLE values of LAPPS are weakly related to source extents ($|r|<0.3$), indicating that its SD and DLE values are only slightly sensitive to the extents.

Figure 2 shows an imaging example under different extents. The first column in Fig. 2 was the simulated sources, which was located in the left occipital lobe. The remaining columns were the recovered source maps by each ESI method, which were shown as the absolute value of sources. The threshold was determined by the Otsu's method. The results of wMNE and LORETA were overly diffused and covered multiple brain functional regions, which greatly exceeded the areas of the simulated source activities. Compared with wMNE, LORETA provided more smooth and coherent spatial solution. Due to the $L_p$-norm regularization term, LAPPS was not sensitive to the source extents and obtained several point sources around the ground truth. Using structured sparsity, SISSY obtained solutions with less blurriness and clearer boundaries of brain activities. Compared to the ground truth, $L_1$R-SSSI showed clearer and more accurate reconstructions than SISSY. As shown in Fig. 2, under spatial extents, the reconstructions of $L_1$R-SSSI matched the ground truth most accurately. According to the evaluation metrics and imaging results, $L_1$R-SSSI outperformed the other compared methods under different source extents.

#### 4.1.2 Influence of number of clusters

In this section, we further tested the performance of $L_1$R-SSSI in multiple patch sources. The extents of each patch were approximately 10 $cm^2$. Figure 3 depicts the evaluation metrics of all algorithms as the number of patches increases. As shown in Fig. 3, the AUC and SD values of all methods except LAPPS decreased with the increase in the number of sources, which is also validated by the results of Spearman's correlation analysis in Table 2. However, LAPPS's AUC and SD values were only slightly affected by the number of clusters (Spearman's correlation: AUC: $r = 0.214$, $p = 2.17e-6$; SD: $r = -0.216$, $p = 8.61e-4$). The DLE and RMSE values of all algorithms were not sensitive to the number of sources. In comparison, $L_1$R-SSSI provided higher AUC ($p<0.05$), the smaller DLE ($p<0.05$) and RMSE ($p<0.05$) values than other methods in estimating multiple patches. Due to the sparse constraint, LAPPS always produced several point sources in or around the ground truth and obtained the smallest SD values. However, LAPPS provided little information of source extents.

Figure 4 is an imaging example under various numbers of clusters. From the leftmost to the rightmost column, the number of simulated sources increased from 1 to 4. The four active sources were located in the left occipital lobe (Source A), left frontal lobe (Source B), right parietal lobe (Source C) and left central cortex (Source D) respectively. The estimations of wMNE and LORETA were too diffuse to obtain globally separated clusters. WMNE and LORETA were not able to separate the activations for adjacent sources (e.g., Source B and D). In contrast, LAPPS always produced focal and sparse solutions. Compared to LAPPS, SISSY and $L_1$R-SSSI were able to recover the extents of cortical activities. However, SISSY also produced some spurious activations and mis-localized some clusters (e.g., Source C). Compared with the other methods, $L_1$R-SSSI was more sensitive to source extents and showed no missing sources and visible false alarms.

#### 4.1.3 Influence of SNRs

Figure 5 depicts the performance metrics of all ESI methods with one active cluster under various SNR level. The noise levels greatly affected the accuracy of source estimation. Combined with the results of Spearman's correlation analysis between the performance metrics and SNRs in Table 3, when the noise pollution in EEG signals decreased, all algorithms show improved performance, which was indicated by the gradually increased AUC and decreased SD, DLE, RMSE values. Among all the imaging

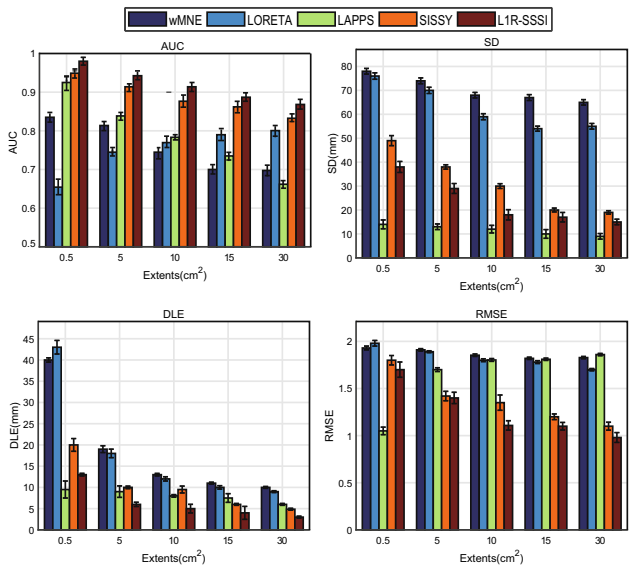


**Fig. 1** Evaluation metrics under various source extents. The figure is the results of 100 Monte Carlo simulations and is shown as Mean ± SEM (SEM: standard error of mean). The SNR is 5 dB

**Table 1** Correlations between the performance metrics and source extents

| Methods | Performance metrics | Correlation ($r$) | $p$ value |
|---|---|---|---|
| wMNE | AUC | − 0.778 | 2.86e–27 |
| | SD | − 0.593 | 2.15e–10 |
| | DLE | − 0.870 | 1.48e–30 |
| | RMSE | − 0.483 | 4.58e–14 |
| LORETA | AUC | 0.701 | 3.25e–35 |
| | SD | − 0.636 | 3.4e–21 |
| | DLE | − 0.851 | 1.38e–32 |
| | RMSE | − 0.536 | 3.23e–19 |
| LAPPS | AUC | − 0.805 | 4.21e–13 |
| | SD | − 0.203 | 3.45e–24 |
| | DLE | − 0.182 | 2.37e–32 |
| | RMSE | 0.750 | 1.35e–12 |
| SISSY | AUC | − 0.605 | 3.41e–14 |
| | SD | − 0.782 | 1.25e–13 |
| | DLE | − 0.752 | 3.57e–31 |
| | RMSE | − 0.714 | 3.42e–18 |
| $L_1$R-SSSI | AUC | − 0.683 | 5.61e–28 |
| | SD | − 0.713 | 1.89e–14 |
| | DLE | − 0.726 | 2.40e–23 |
| | RMSE | − 0.735 | 3.41e–13 |

methods, $L_1$R-SSSI always obtained the best performance with the largest AUC ($p<0.05$), the smallest DLE ($p<0.05$) and RMSE ($p<0.05$) values for all SNRs.

### 4.2 Results of real data analysis

Furthermore, $L_1$R-SSSI was applied to analyze the human face-processing data. The dataset includes the simultaneous MEG/EEG recordings of 16 subjects, which can be available at https://openneuro.org/datasets/ds000117. During the data recordings, the subject performed a simple visual task on a large number of famous, unfamiliar and scrambled faces. In this work, only the EEG signals were employed for analysis. We downsampled the EEG data to 275 Hz and averaged the 16 subjects' EEG recordings corresponding to face (famous and unfamiliar) stimulus for source estimation. The time window of EEG recordings was − 200 to 900 ms, where 0 ms denotes the stimulus onset time. The topography distribution of EEG recording at 170 ms and the averaged EEG time courses are shown in Fig. 6a, b, respectively. The lead-field matrix was calculated using OpenMEEG based on BEM model.

Figure 6c is the imaging results at 170 ms, including the ventral and lateral views of the cortex. As shown in Fig. 6, $L_1$R-SSSI and SISSY located bilateral fusiform and right temporal cortices, which was consistent with prior studies [17, 28]. However, SISSY also located some irrelevant areas. wMNE and LORETA can recognize the activities of bilateral fusiform area, but they were too widespread. In contrast, LAPPS only detected some point activities at the bilateral fusiform and right temporal areas. For the human face-processing EEG data analysis, the proposed method obtained more reasonable results that were also consistent with previous reports [17, 28, 30].

## 5 Discussions

In this work, we proposed a robust ESI method, $L_1$R-SSSI, to reconstruct extended sources, especially when strong background activity and outlier noise existed in the EEG recordings. $L_1$R-SSSI uses $L_1$-norm to fit the residual errors and implements the sparse constraint in the variation domain and the original source domain. The solution of $L_1$R-SSSI is efficiently solved by ADMM. Simulation studies reveal the superior performance of $L_1$R-SSSI compared with the benchmark algorithms (e.g., wMNE, LORETA, LAPPS and SISSY). $L_1$R-SSSI obtains more accurate estimation in terms of detection sensitivity, source extents, locations and amplitude errors. For real experimental data, $L_1$R-SSSI also provides more meaningful neurophysiological results.

Since the EEG inverse problem is highly ill-posed, suitable regularization constraints are necessary to obtain unique source solution [13, 34]. The traditional $L_2$-norm-based methods (e.g., wMNE and LORETA) always

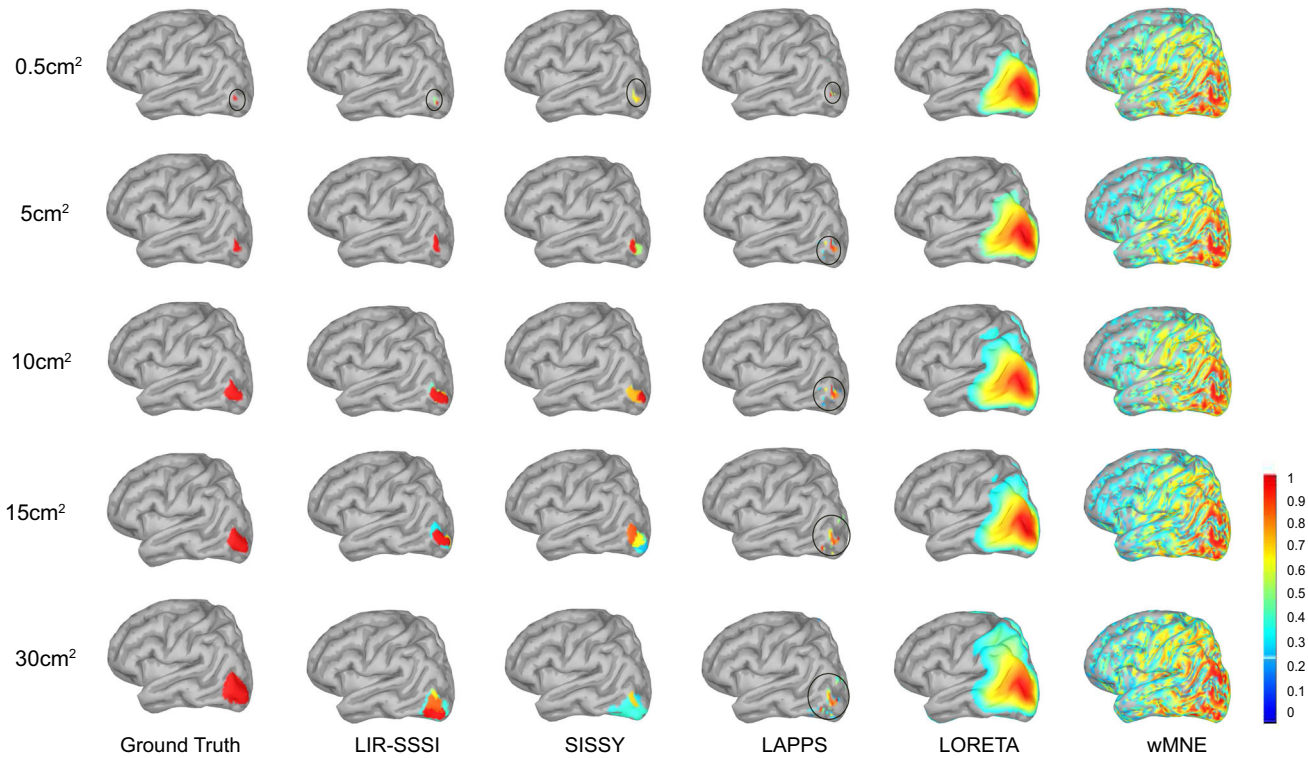


**Fig. 2** Imaging results for different extents. Source activity maps show the absolute value of the sources. The threshold is determined using Otsu's method. The SNR is 5 dB. Some sources are circled for illustration purpose

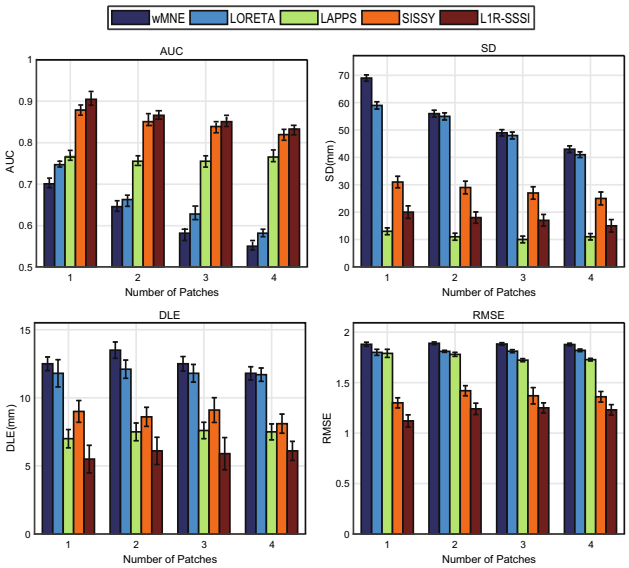


**Fig. 3** Evaluation metrics of different numbers of patches. The data are the results of 100 Monte Carlo simulations and is described as Mean ± SEM . The SNR is 5 dB and the extents of each cluster is about 10 cm$^2$

generate blurred and diffused source estimations, which is also indicated by the high SD and DLE values in Figs. 1, 3 and 5. To improve the spatial resolution of reconstructed sources, some studies have proposed sparse constraint

**Table 2** Correlations between the performance metrics and number of patches

| Methods | Performance metrics | Correlation ($r$) | $p$ value |
|---|---|---|---|
| wMNE | AUC | − 0.789 | 2.31e–15 |
| | SD | − 0.835 | 6.67e–36 |
| | DLE | − 0.730 | 0.193 |
| | RMSE | 0.026 | 0.719 |
| LORETA | AUC | − 0.801 | 3.18e–18 |
| | SD | − 0.792 | 2.42e–30 |
| | DLE | 0.043 | 0.564 |
| | RMSE | 0.016 | 0.826 |
| LAPPS | AUC | 0.214 | 2.17e–6 |
| | SD | − 0.216 | 8.61e–4 |
| | DLE | 0.179 | 0.122 |
| | RMSE | − 0.122 | 0.631 |
| SISSY | AUC | − 0.426 | 1.63e–10 |
| | SD | − 0.386 | 1.18e–15 |
| | DLE | 0.183 | 0.734 |
| | RMSE | − 0.135 | 0.794 |
| $L_1$R-SSSI | AUC | − 0.482 | 1.78e–14 |
| | SD | − 0.401 | 2.73e–13 |
| | DLE | 0.231 | 0.551 |
| | RMSE | 0.203 | 0.135 |

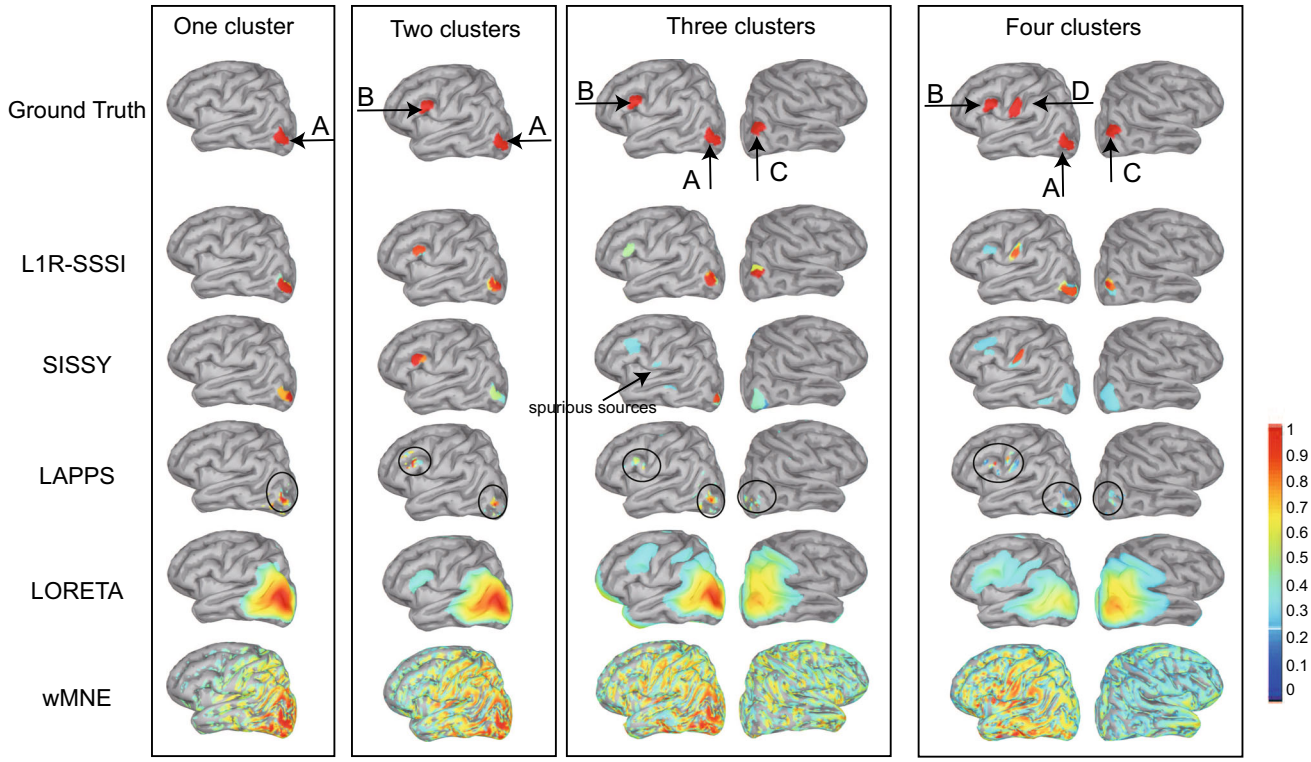


**Fig. 4** Imaging results for different numbers of clusters. Source activity maps show the absolute value of estimated sources. The threshold is determined using Otsu's method. The SNR is 5 dB and the extent of each patch is about 10 cm$^2$. The reconstructed sources of LAPPS are circled for illustration purpose

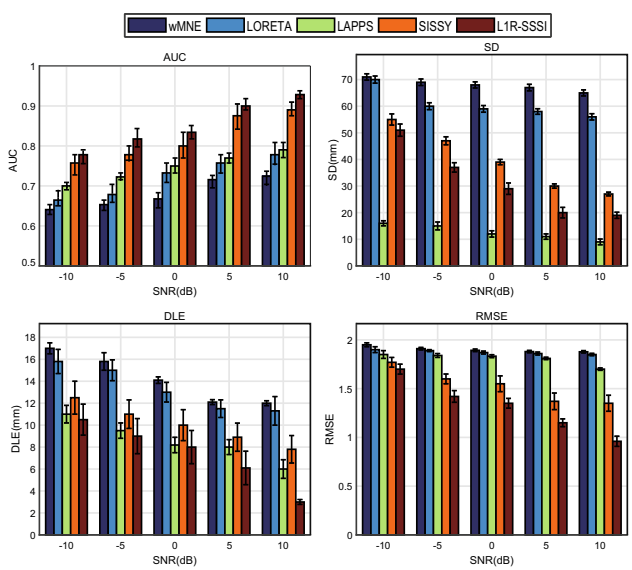


**Fig. 5** Evaluation metrics of various SNRs. The figure is the results of 100 Monte Carlo simulations and is described as Mean ± SEM. The extent of source is about 10 cm$^2$

methods, such as the $L_p$-norm ($p \leq 1$) and sparse Bayesian Learning-based methods. However, the sparse penalties in the original source space produce overly focal estimations [33]. These highly focal results are caused by enforced

**Table 3** Correlations between the performance metrics and SNRs

| Methods | Performance metrics | Correlation ($r$) | $p$ value |
|---|---|---|---|
| wMNE | AUC | 0.441 | 4.51e–7 |
| | SD | – 0.307 | 1.03e–18 |
| | DLE | – 0.798 | 2.79e–7 |
| | RMSE | – 0.331 | 5.21e–18 |
| LORETA | AUC | 0.495 | 1.37e–6 |
| | SD | – 0.363 | 1.25e–15 |
| | DLE | – 0.806 | 1.95e–10 |
| | RMSE | – 0.350 | 4.51e–21 |
| LAPPS | AUC | 0.631 | 3.78e–9 |
| | SD | – 0.314 | 3.43e–20 |
| | DLE | – 0.709 | 5.67e–9 |
| | RMSE | – 0.372 | 2.76e–18 |
| SISSY | AUC | 0.794 | 5.12e–5 |
| | SD | – 0.815 | 4.56e–13 |
| | DLE | – 0.738 | 3.21e–5 |
| | RMSE | – 0.626 | 2.84e–15 |
| $L_1$R-SSSI | AUC | 0.776 | 2.57e–4 |
| | SD | – 0.835 | 3.58e–19 |
| | DLE | – 0.755 | 2.18e–4 |
| | RMSE | – 0.687 | 3.12e–16 |

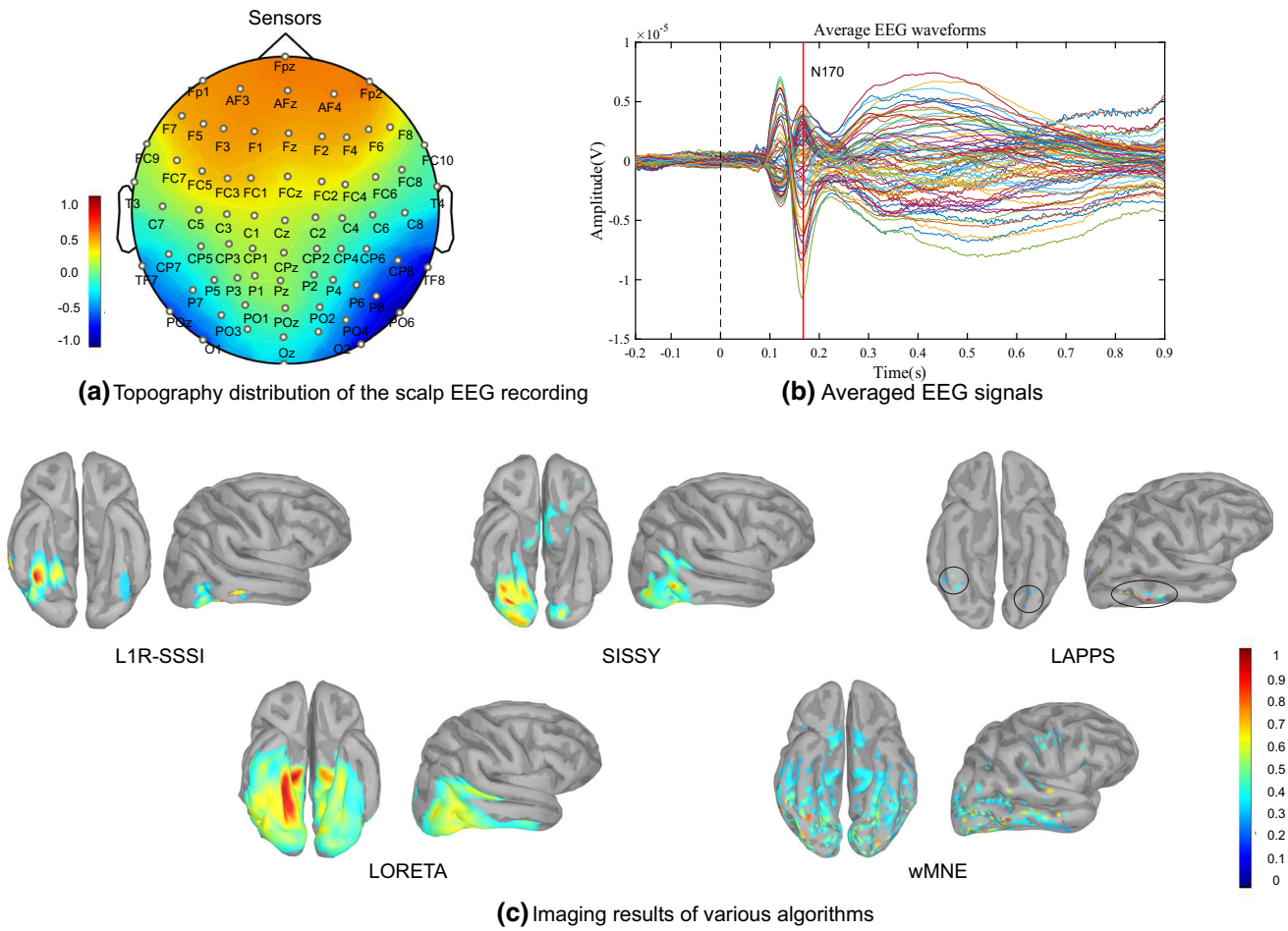


**Fig. 6** Imaging results for the EEG recordings of face recognition task, showing the absolute value of the source activities at 170 ms. The threshold is determined using Otsu's method. For each imaging algorithm, the ventral and lateral views of the brain are shown. The estimated source activities of LAPPS are circled for illustration purposes

sparseness in the original source domain. Therefore, to estimate both the locations and extents, some studies have proposed the sparsity penalties in the transform domains, such as the variation domain [9]. However, the sparsity in the variation domain produces amplitude-biased solutions. To remedy this issue, SISSY implements the $L_1$-norm constraint in the variation domain and original source domain simultaneously [4].

However, the above ESI methods usually employ $L_2$-norm to fit residual errors, which is based on the assumption that the EEG measurement noise satisfies Gaussian distribution. Nevertheless, EEG signals are inevitably contaminated by strong background activities and outliers caused by ocular or head movements during recordings. The $L_2$-norm loss may exaggerate the effect of these outliers. To remedy this issue, the study in [6] employed the $L_1$-loss for the residuals and $L_p$-penalty regularization, showing superior performance than the $L_2$-norm loss. As shown in our simulations, LAPPS can reconstruct the focal sources accurately (see Fig. 1 for 0.5 cm$^2$). However, for activities with large extents, LAPPS provides little information of source extents, as shown in Figs. 2 and 4.

To tackle the outliers in the EEG recordings and reconstruct extended sources, we proposed a robust ESI algorithm, i.e., $L_1$R-SSSI. The proposed method employs the $L_1$-norm to measure the fitting error, assuming that the EEG measurement noise satisfies the Laplace distribution. Our former work [36] has proposed the $L_1$-norm loss and $L_1$-norm regularization of variation sources to reconstruct the extents and locations of cortical activities. However, as suggested in [38], the minimization of variation sources usually severely underestimates the amplitudes of sources [36]. To more accurately reconstruct source locations, extents and amplitudes, $L_1$R-SSSI proposed the structured sparsity constraint (i.e., $L_1$-norm regularization in the variation and original source domain) as in [4]. The results of Monte Carlo simulations have verified that $L_1$R-SSSI is robust to the outlier noise in EEG recordings and provides more accurate reconstructions than the benchmark algorithms.

In this work, we develop $L_1$R-SSSI under the regularization framework. Our future work will model the inverse problem in the Bayesian probabilistic framework. Specifically, we can assume that both (1) the EEG measurement noise and (2) the prior of sources and variation sources satisfy the Laplace distribution. By solving the inverse problem under the probabilistic framework, we can obtain both the point estimations of sources and the corresponding uncertainty [22].

## 6 Conclusions

In summary, we propose a robust ESI method $L_1$R-SSSI to recover brain activities, which is efficiently solved by ADMM. $L_1$R-SSSI employs $L_1$-norm to fit residuals to alleviate the effect of outliers in EEG recordings and the structured sparsity to achieve globally sparse and locally smooth solution. For the Monte Carlo numerical simulations in this work, compared to the benchmark algorithms, $L_1$R-SSSI obtained larger AUC (average AUC $> 0.80$) and smaller SD (average SD $<50$ mm), DLE (average DLE $<10$ mm), RMSE (average RMSE $<1.75$). Comprehensively considering the four performance metrics, $L_1$R-SSSI is more powerful in estimating the source locations, extents and amplitudes. Human EEG data analysis demonstrates that $L_1$R-SSSI is a useful imaging method for estimating cortical activities, which will be helpful to the research of neuroscience and clinical cases. In this work, we only estimated the sources at a specified single time point. Due to high temporal resolution of EEG, the use of the temporal information is helpful to improve the accuracy of source reconstruction [21, 22, 28]. Our future work will also employ the temporal information under the framework of $L_1$R-SSSI to estimate the dynamic process of brain activities. We will also apply the proposed algorithm to the analysis of brain networks underlying various psychiatric disorders [26] and to the decoding of EEG signals for brain–computer interfaces [19].

**Acknowledgements** This work was supported in part by the National Natural Science Foundation of China under Grants 61703065 and 61836003, Chongqing Research Program of Application Foundation and Advanced Technology under Grant cstc2018jcyjAX0151, the Science and Technology Research Program of Chongqing Municipal Education Commission under Grant KJQN201800612 and Chongqing Graduate Research and Innovation Project under Grant CYS20257.

### Compliance with Ethical Standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## Appendix

The schema of $L_1$R-SSSI is summarized as follows.

**Algorithm 1** $L_1$R-SSSI algorithm

**Require:** EEG data $\boldsymbol{b}$, lead field matrix $\boldsymbol{L}$, variation operator $\boldsymbol{V}$, stop tolerance $\delta$ (e.g., $10^{-4}$) and parameters $\lambda_1$, $\lambda_2$, $\rho_1$, $\rho_2$, $\rho_3$.

**Ensure:** Current source solution $\boldsymbol{s}$.

1: Initialize vectors $\boldsymbol{e}$, $\boldsymbol{u}$, $\boldsymbol{w}$, $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$.
2: **repeat**
3: Update the values of source estimate $\boldsymbol{s}$ and latent vectors ($\boldsymbol{e}$, $\boldsymbol{u}$, $\boldsymbol{w}$) respectively by (7);
4: Update the values of Lagrangian multipliers ($\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$) by (8);
5: **until** $\frac{\|\boldsymbol{s}^{(k+1)}-\boldsymbol{s}^{(k)}\|^2}{\|\boldsymbol{s}^{(k+1)}\|^2} < \delta$.

## Appendix

Figure 7 shows the performance metrics under different values of $\lambda_1$ and $\lambda_2$. Clearly, the selection of regularization parameters has a great impact on algorithm performance. According to Fig. 7, $L_1R$-SSSI achieves better performance for $\lambda_1 \in [1.6, 2.1]$, $\lambda_2 \in [0.9, 1.3]$, indicated by the relatively larger AUC, smaller SD, DLE, and RMSE values. For simplicity, we employed a fixed value $\lambda_1 = 2.1$ and $\lambda_2 = 1.0$ for the experimental simulations.

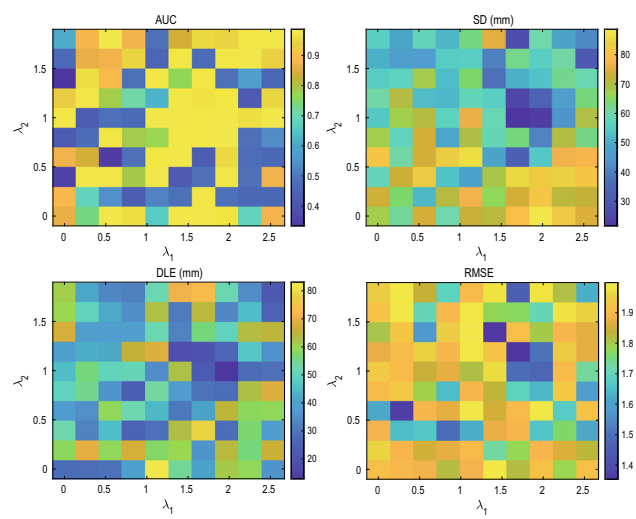


**Fig. 7** Performance metrics under different values $\lambda_1$ and $\lambda_2$

## References

1. Bai X, Towle VL, He EJ, He B (2007) Evaluation of cortical current density imaging methods using intracranial electrocorticograms and functional MRI. NeuroImage 35(2):598–608
2. Baillet S, Mosher JC, Leahy RM (2001) Electromagnetic brain mapping. IEEE Signal Process Mag 18(6):14–30

3. Becker H (2014) Denoising, separation and localization of EEG sources in the context of epilepsy. Ph.D. thesis
4. Becker H, Albera L, Comon P, Nunes JC, Gribonval R, Fleureau J, Guillotel P, Merlet I (2017) SISSY: an efficient and automatic algorithm for the analysis of EEG sources based on structured sparsity. NeuroImage 157:157–172
5. Blankertz B, Dornhege G, Krauledat M, Müller KR, Curio G (2007) The non-invasive Berlin brain-computer interface: fast acquisition of effective performance in untrained subjects. NeuroImage 37(2):539–550
6. Bore JC, Yi C, Li P, Li F, Harmah DJ, Si Y, Guo D, Yao D, Wan F, Xu P (2019) Sparse EEG source localization using LAPPS: least absolute l–p penalized solution. IEEE Trans Biomed Eng 66(7):1927–1939
7. Cai C, Sekihara K, Nagarajan SS (2018) Hierarchical multiscale Bayesian algorithm for robust MEG/EEG source reconstruction. NeuroImage
8. Chang W, Nummenmaa A, Hsieh J, Lin F (2010) Spatially sparse source cluster modeling by compressive neuromagnetic tomography. NeuroImage 53(1):146–160
9. Ding L (2009) Reconstructing cortical current density by exploring sparseness in the transform domain. Phys Med Biol 54(9):2683–2697
10. Ding L, He B (2008) Sparse source imaging in electroencephalography with accurate field modeling. Hum Brain Mapp 29(9):1053–1067
11. Fei W, Ling P, Yuan Y, Wenxian Y, Peilin L (2017) Efficient and robust recovery of sparse signal and image using generalized nonconvex regularization. IEEE Trans Comput Imaging
12. Foucart S, Lai M (2009) Sparsest solutions of underdetermined linear systems via lq-minimization for $0 < q < 1$. Appl Comput Harmon Anal 26(3):395–407
13. Friston KJ, Harrison LM, Daunizeau J, Kiebel SJ, Phillips C, Trujillobarreto NJ, Henson RN, Flandin G, Mattout J (2008) Multiple sparse priors for the M/EEG inverse problem. NeuroImage 39(3):1104–1120
14. Grova C, Daunizeau J, Lina J, Benar CG, Benali H, Gotman J (2006) Evaluation of EEG localization methods using realistic simulations of interictal spikes. NeuroImage 29(3):734–753
15. He B, Sohrabpour A, Brown EN, Liu Z (2018) Electrophysiological source imaging: a noninvasive window to brain dynamics. Annu Rev Biomed Eng 20(1):171–196
16. He B, Yang L, Wilke C, Yuan H (2011) Electrophysiological imaging of brain activity and connectivity-challenges and opportunities. IEEE Trans Biomed Eng 58(7):1918–1931
17. Henson RN, Goshengottstein Y, Ganel T, Otten LJ, Quayle AH, Rugg MD (2003) Electrophysiological and haemodynamic correlates of face perception, recognition and priming. Cereb Cortex 13(7):793–805
18. Hong M, Luo Z (2012) On the linear convergence of the alternating direction method of multipliers. arXiv: Optimization and Control **162**(1), 165–199
19. Hou Y, Zhou L, Jia S, Lun X (2020) A novel approach of decoding EEG four-class motor imagery tasks via scout ESI and CNN. J Neural Eng 17(1):016048
20. Kiebel SJ, Daunizeau J, Phillips C, Friston KJ (2008) Variational Bayesian inversion of the equivalent current dipole model in EEG/MEG. Neuroimage 39(2):728–741
21. Liu K, Yu ZL, Wu W, Gu Z, Li Y, Nagarajan SS (2016) Bayesian electromagnetic spatio-temporal imaging of extended sources with Markov random field and temporal basis expansion. NeuroImage 139:385–404
22. Liu K, Yu ZL, Wu W, Gu Z, Li Y, Nagarajan SS (2018) Variation sparse source imaging based on conditional mean for electromagnetic extended sources. Neurocomputing 313:96–110
23. Lucka F, Pursiainen S, Burger M, Wolters CH (2012) Hierarchical bayesian inference for the EEG inverse problem using realistic fe head models: Depth localization and source separation for focal primary currents. Neuroimage 61(4):1364–1382
24. Owen JP, Wipf D, Attias H, Sekihara K, Nagarajan SS (2012) Performance evaluation of the champagne source reconstruction algorithm on simulated and real M/EEG data. NeuroImage 60(1):305–323
25. Pascual-Marqui R, Esslen M, Kochi K, Lehmann D (2002) Functional imaging with low-resolution brain electromagnetic tomography (loreta): a review. Methods Find Exp Clin Pharmacol 24(Suppl C):91–95
26. Rolle CE, Fonzo GA, Wu W, Toll R, Jha MK, Cooper C, Chin-Fatt C, Pizzagalli DA, Trombello JM, Deckersbach T et al (2020) Cortical connectivity moderators of antidepressant versus placebo treatment response in major depressive disorder: secondary analysis of a randomized clinical trial. JAMA Psychiatry 77(4):397–408
27. Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM (2011) Brainstorm: a user-friendly application for MEG/EEG analysis. Comput Intell Neurosci 2011:879716–879716
28. Trujillo-Barreto NJ, Aubert-Vázquez E, Penny WD (2008) Bayesian M/EEG source reconstruction with spatio-temporal priors. NeuroImage 39:318–335
29. Wagner A, Wright J, Ganesh A, Zhou Z, Mobahi H, Ma Y (2012) Toward a practical face recognition system: robust alignment and illumination by sparse representation. IEEE Trans Pattern Anal Mach Intell 34(2):372–386
30. Wakeman DG, Henson RN (2015) A multi-subject, multi-modal human neuroimaging dataset. Sci Data 2(1):150001–150001
31. Wang Y, Yin W, Zeng J (2019) Global convergence of ADMM in nonconvex nonsmooth optimization. J Sci Comput 78(1):29–63
32. Wen F, Pei L, Yang Y, Yu W, Liu P (2017) Efficient and robust recovery of sparse signal and image using generalized nonconvex regularization. IEEE Trans Comput Imaging 3(4):566–579
33. Wipf D, Nagarajan SS (2009) A unified Bayesian framework for MEG/EEG source imaging. NeuroImage 44(3):947–966
34. Wu W, Nagarajan SS, Chen Z (2016) Bayesian machine learning: EEG/MEG signal processing measurements. IEEE Signal Process Mag 33(1):14–36
35. Xiao Y, Zhu H, Wu SY (2013) Primal and dual alternating direction algorithms for l1-l1-norm minimization problems in compressive sensing. Comput Optim Appl 54(2):441–459
36. Xu F, Liu K, Deng X, Wang G (2019) Imaging EEG extended sources based on variation sparsity with L1-norm residual. In: Liang P, Goel V, Shan C (eds) Brain informatics. Springer International Publishing, Cham, pp 95–104
37. Xu P, Tian Y, Chen H, Yao D (2007) Lp norm iterative sparse solution for EEG source localization. IEEE Trans Bio-Med Eng 54:400–409
38. Zhu M, Zhang W, Dickens D, Ding L (2014) Reconstructing spatially extended brain sources via enforcing multiple transform sparseness. NeuroImage 86:280–293


**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.