CrossMark

ORIGINAL PAPER

Integrated Analysis of EEG and fMRI Using Sparsity of Spatial Maps

S. Samadi[1,2,3] · H. Soltanian-Zadeh[1,4,5] iD · C. Jutten[2,6]

Abstract Integration of electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) is an open problem, which has motivated many researches. The most important challenge in EEG-fMRI integration is the unknown relationship between these two modalities. In this paper, we extract the same features (spatial map of neural activity) from both modality. Therefore, the proposed integration method does not need any assumption about the relationship of EEG and fMRI. We present a source localization method from scalp EEG signal using jointly fMRI analysis results as prior spatial information and source separation for providing temporal courses of sources of interest. The performance of the proposed method is evaluated quantitatively along with multiple sparse priors method and sparse Bayesian learning with the fMRI results as prior information. Localization bias and source distribution index are used to measure the performance of different localization approaches with or without a variety of fMRI-EEG mismatches on simulated realistic data. The method is also applied to experimental data of face perception of 16 subjects. Simulation results show that the proposed method is significantly stable against the noise with low localization bias. Although the existence of an extra region in the fMRI data enlarges localization bias, the proposed method outperforms the other methods. Conversely, a missed region in the fMRI data does not affect the localization bias of the common sources in the EEG-fMRI data. Results on experimental data are congruent with previous studies and produce clusters in the fusiform and occipital face areas (FFA and OFA, respectively). Moreover, it shows high stability in source localization against variations in different subjects.

Keywords Integration analysis · Referenced-based signal processing · Weighted sparse decomposition · Elastic net · Pareto optimization

✉ H. Soltanian-Zadeh
hszadeh@ut.ac.ir

C. Jutten
christian.jutten@gipsa-lab.grenoble-inp.fr

1   CIPCE, Electrical and Computer Engineering Department, University of Tehran, Tehran, Iran

2   GIPSA-LAB, University of Grenoble Alpes, Domaine universitaire- BP 46, 38402 Grenoble Cedex, France

3   Department of Electrical Engineering, Hashtgerd Branch, Islamic Azad University, Hashtgerd, Iran

4   Radiology Image Analysis Laboratory, Henry Ford Health System, Detroit, MI 48202, USA

5   School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

6   Institut Universitaire de France, Paris, France

Introduction

EEG is a non-invasive technique that records scalp electrical activity generated by brain structures and it greatly helps to figure out the aspects of cognitive processes. The EEG data is analyzed by solving its inverse problem, which looks for the localization and dynamics of cerebral activity sources. EEG reflects neural activity directly with high temporal resolution (around millisecond), but it lacks good spatial resolution. Recently, EEG-fMRI integration has been intensively addressed due to their complementary temporal and spatial resolutions. fMRI relies on the blood oxygen level dependent (BOLD) effect and measures hemodynamic signal. fMRI has high spatial resolution around millimeters. Therefore, it seems that the fMRI high

🙂 Springer

spatial resolution can assist solving the EEG inverse problem.

Different EEG and fMRI integration methods have been proposed recently which can be categorized into symmetric and asymmetric groups (Jorge et al 2014; Rosa et al 2010). Fusion or symmetric methods use the data of both modalities as the observation variables and consider their interactions to extract the maximum information about the neural activity (Henson 2010). These methods are complex and most of them require accurate physiological models of the interaction between neuronal activity and BOLD signals. In asymmetric methods the data from one modality is used to constrain the inference [fMRI-informed EEG (-Murta et al 2015)] or as a predictor [EEG-informed fMRI (Dong et al 2015)] of the other modality.

Here, we focus on the fMRI-informed EEG group. The fMRI-informed EEG methods apply the fMRI result in the inverse problem of EEG to improve the spatial resolution of source localization. Two main challenges are related to EEG-informed fMRI method: (1) How to solve the EEG inverse problem? (2) How to apply fMRI information?

Two main approaches are used for solving the inverse problem: *non-parametric* (imaging) and *parametric* (scanning) methods (Baillet et al 2001). A brief review of these methods can be found in Bolstad et al (2009). Parametric or scanning methods use a small number of dipoles (Michel et al 2004; Bai and He 2005; Shun Chi Wu et al 2012), multi-poles (Mosher et al 1999; Bandyopadhyay et al 2004), or cortical patches (Limpiti et al 2006; Wagner et al 2000) and scan over any possible locations to find the best set of sources to represent the data. The examples of scanning methods include MUSIC (Mosher and Leahy 1996), beamforming (Van Veen 1991; Fuchs et al 2011), and maximum likelihood estimation (Inan et al 2001). Non-parametric methods are also referred to as imaging methods. In these models several dipoles with fixed locations and possibly fixed orientations are distributed in the whole brain volume or cortical surface. In imaging method, the EEG forward problem is estimated with a linear system equations. Then, using head model, estimating the sources from scalp EEG is an underdetermined and ill-posed problem, which requires prior information and constraints for converging to a unique solution.

Minimum norm estimation (MNE) is one of the earliest methods to solve the EEG inverse problem. MNE searches a solution with the minimum ℓ_2 norm, obtained from a closed-form equation involving a simple matrix product (Tikhonov and Arsenin 1977). This makes the estimation extremely fast. However, using the ℓ_2 norm regularization method leads to smooth solution. Therefore, it fails to recover solutions with high spatial frequencies, because the extent of active regions is often over-estimated.

Recently, sparsity has been used as an effective tool to find sparse (unique) solutions for an underdetermined system of linear equations. Sparsity means that the solution should only have a small number of non-zero elements. The ℓ_0 pseudo-norm[1] is the mathematical way to measure sparsity of a vector **x**, which is nothing but the number of non-zero components of **x**. In the last few years, many algorithms have been developed to solve the ill-posed underdetermined linear system when considering sparse priors (Gorodnitsky et al 1995; Daubechies et al 2008; Tibshirani 1996; Efron et al 2004; Tibshirani et al 2005; Beck and Teboulle 2009; Friedman et al 2010; Bruce et al 1998).

It has been proved that ℓ_1 regularization leads to a sparse solution (Tibshirani 1996; Chen et al 1999) and under some conditions, produces the sparsest solution (Donoho 2004; Donoho and Elad 2003; Donoho 2006; Candès and Tao 2005). The ℓ_p norm is a real-valued function which is convex when $p \geq 1$. For $p = 1$, ℓ_1 norm is marginally convex and is the best convex estimation of the ℓ_0 norm (Wipf and Nagarajan 2009). Recently, some theoretical and practical studies show that the ℓ_0 regularization criterion is superior to that of ℓ_1 regularization in a sense of accuracy (Mancera and Portilla 2006; Mohimani et al 2009; Babaie-Zadeh and Jutten 2010). Since ℓ_0 norm regularization leads to NP-hard problems, (Mohimani et al 2009) proposed an algorithm to minimize the ℓ_0 norm by approximating ℓ_0 norm by a continuous and differentiable function. The algorithm, called smooth ℓ_0 norm (SL0), searches the optimal solution based on gradient ascent method. SL0 brings us two new opportunities. First, It is faster than the interior-point linear programming (LP) solvers, while at the same time having the same or better accuracy. Second, contrary to the previous approaches, ℓ_0 norm minimization algorithms can provide a stable estimation from noisy data, with a weaker sparsity requirement on the data than using ℓ_1 norm[2] (Babaie-Zadeh and Jutten 2010).

Some studies have solved the EEG inverse problem for a sparse solution such as the self-coherence enhancement algorithm (Dezhong Yao 2001), the ℓ_1 norm solution (i.e., LASSO) (Tibshirani 1996), and the ℓ_p norm iterative sparse source localization (LPISS) (Xu et al 2007). Also, the combination of the ℓ_2 and ℓ_1 norms has been proposed (Gramfort et al 2012, 2013). Different strategies are used to combine the ℓ_2 and ℓ_1 norms. Mixed-Norm Estimates (MxNE) algorithm minimizes the ℓ_1 norm of the ℓ_2 norm of the time series of the sources ($MxNE = \sum_{i=1}^{M} \sqrt[2]{\sum_{j=1}^{T} |\mathbf{X}_{ij}|^2}$,

---

[1] The ℓ_0 pseudo-norm does not satisfy the mathematical definition of a norm, However, in the following, we simply say ℓ_0 norm.

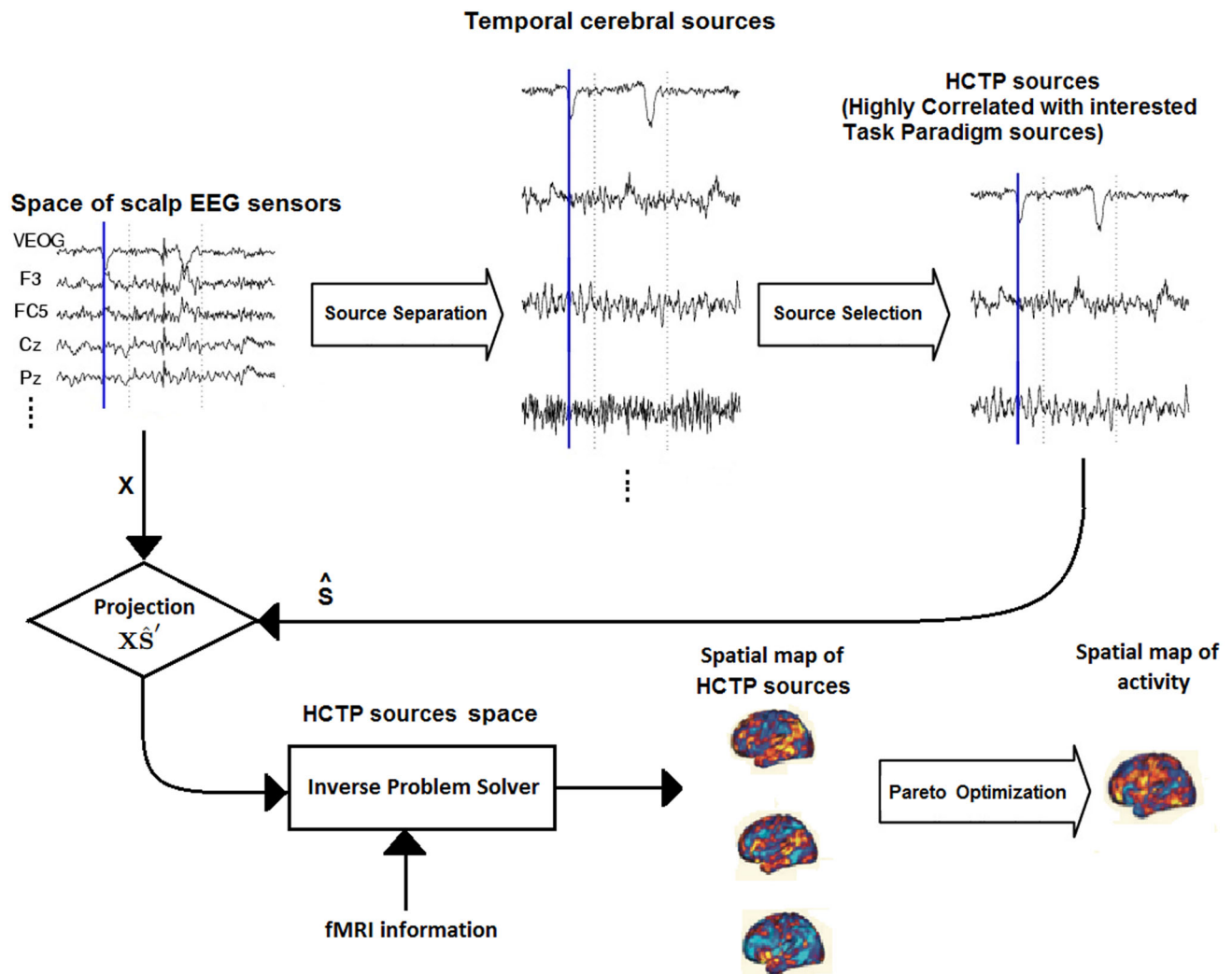[2] i.e. with data less sparse than required with algorithm based on ℓ_1 norm.

Fig. 1 The different spaces and their relationships

where $\mathbf{X}$ is an unknown matrix ($M \times T$)). It estimates the time series of the sources and their spatial maps simultaneously, but with a high computational cost. Therefore, it is used in a region of interest but cannot be applied for the whole brain. Group LASSO is used when we can partition the sparse variable into different groups (Yuan and Lin 2006; Cassidy et al 2012). Elastic Net uses the weighted sum of the ℓ_2 and ℓ_1 norms ($\lambda_1 \|\mathbf{x}\|_1 + \lambda_2 \|\mathbf{x}\|_2$, where $\mathbf{x}$ is an unknown vector) in the regularization term to estimate the connected sparse regions.

Here, we are interested in the brain neural activities, which are sparsely localized, thus a relevant constraint is to use the sparse characteristics of the spatial distribution of the sources (Gramfort et al 2012; Ding et al 2011; Ding 2009; Friston et al 2008). We choose the ℓ_0 norm to represent sparsity and benefit its accuracy and stability against the noise using the smooth ℓ_0 norm (SL0) algorithm (Mohimani et al 2009; Babaie-Zadeh and Jutten

2010). For using sparsity, since EEG is not potentially sparse, we have to project the EEG data to another space in which the data are represented by sparse unknown variables (see Fig. 1). This space is made up of temporal sources which we are interested in, i.e. sources which are highly correlated with the paradigm of the task of interest. We call them highly correlated with interested task paradigm sources (HCTP sources). The HCTP source space is found by applying the reference-based source separation (R-SS) method proposed in (Samadi et al 2013). R-SS is a semi-blind source separation method which extracts the discriminating sources of two data groups, one group being related to the task of interest. By projection to the HCTP source space, we reformulate the EEG inverse problem to achieve a sparse unknown variable which is the spatial map of HCTP sources. For a given sparsity level (i.e. ℓ_0 norm), we would like to promote neural activity concentrated in neighbour pixels

of a small number of cerebral areas, instead to neural activity distributed on isolated voxels. For enforcing this sparsity, usually called block-sparsity, we will combine the ℓ_0 norm and the ℓ_2 norm like what was done in elastic net. We also use the spatial map extracted from the thresholded F-map of the fMRI data as a weight for the spatial map of the EEG source activation. It means that EEG source activity is assumed to be located in areas of significant fMRI activation with high probability. Here, we use the same feature, extracted from EEG and fMRI, which is the spatial map of neural activity. Therefore, we did not need to use any assumption about the relationship between EEG and fMRI, which may not be generally congruent with reality.

The different spaces used in our method are illustrated in Figure 1, and defined as follows:

– *Spatial map* means the map of activation inside the brain.
– *Temporal cerebral sources* are the signals related to the cerebral sources extracted from the scalp EEG data using the source separation method.
– *HCTP sources* are the temporal sources (signals) highly correlated with the activity of interest: they are the outputs of the source selection algorithm.
– *HCTP sources space* is the space spanned by the HCTP sources, i.e. whose HCTP sources are the basis.

The main contributions of the current work are:

1. Using the same feature from both modalities which is the spatial map of the neural activity.
2. By projecting the forward problem of EEG onto the HCTP source space, we can reach enough sparsity to use sparse decomposition methods.
3. Promoting solutions characterized by the concentration of active voxels in a few regions (block-sparsity) using the ℓ_0 norm combined with the ℓ_2 norm.
4. Using fMRI spatial map as a weighting matrix in the penalty term to solve the inverse problem of EEG.
5. Exploiting Pareto optimization to find the final results.

The rest of the paper is organized as follows. The proposed method is explained in "Method" section. The data protocol and the results are presented in "Data" and "Results" sections, respectively. Discussion and concluding remarks are brought in "Discussion" and "Conclusion" sections, respectively.

Method

In the imaging method a grid of thousands of dipoles is used to model the forward problem of EEG as:

$$\mathbf{X} = \mathbf{GJ} + \mathbf{n} \tag{1}$$

where $\mathbf{X}$ is a $N \times T$ matrix that contains scalp recordings, N is the number of channels and T is the number of time points, $\mathbf{G}_{N \times M}$ is the gain matrix, $\mathbf{J}_{M \times T}$ contains dipole magnitudes at different time instants, $\mathbf{n}$ represents the noise, and M is the number of mesh vertices with, practically, $N \ll M$. Under this notation, the inverse problem consists in finding an estimate of the dipole temporal courses, which are the rows of the matrix $\mathbf{J}$, given the electrode positions and scalp recordings $\mathbf{X}$ and using the gain matrix $\mathbf{G}$, which can be calculated with boundary element method (BEM).

As denoted before, the EEG inverse problem is underdetermined (due to the fewer observations than the unknown variables (since $N \ll M$)) and ill-posed so, it lacks a unique solution. Therefore, some additional priors as constraints are needed to find a unique solution. The constraints are applied by adding a regularization term to the usual term of data fit, leading to the solution:

$$\hat{\mathbf{J}} = \arg \min_{\mathbf{J}} \|\mathbf{X} - \mathbf{GJ}\| + \lambda h(\mathbf{J}) \tag{2}$$

Here, we are interested in localized brain neural activities, i.e. only a few regions of the brain that are responsible for the specific brain activity. Therefore, using the spatial sparsity as a constraint in solving the EEG inverse problem is consistent with this prior (Simoncelli and Olshausen 2001; Op de Beeck et al 2008). The most important question is: How can we find the spatial map of neural activity to apply sparsity constraint?

We find the source spatial map in four steps. First by using a referenced based source separation (R-SS) method (Samadi et al 2013) the time courses of the HCTP sources are estimated, secondly the forward model is projected on the HCTP source space and thirdly, the spatial map of each HCTP source is estimated using a sparse decomposition method. Finally, Pareto optimization method is used to find the active regions. From now on, as the proposed method is based on Spatial Sparsity in the Source Space, we call it "S4" method. The architecture of the proposed method is shown in Fig. 2, each block of which will be explained in details in next subsections.

Source Space Estimation

Here, we use the R-SS method introduced in Samadi et al (2013) to extract the uncorrelated temporal courses which have high similarity with a reference state (HCTP sources). The R-SS method is used when we can segment the signal into non-overlapping time intervals each exclusively associated to one state among two different states. Defining states are dependent on the application. In this study, state refers to different tasks in the experiment. In the experiment we may

Fig. 2 Diagram of the proposed method (called S4). WEN0 is the regularization method, and $h(SM) = (1 - \alpha) + \alpha SM$
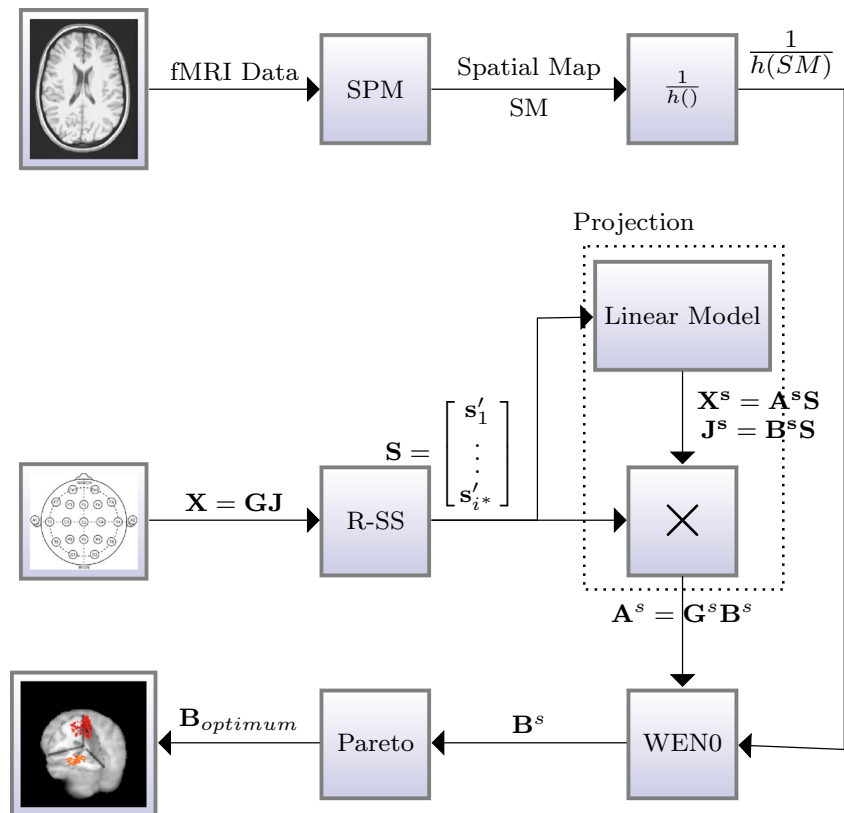
have task, control or rest states. We choose the state of interest as the reference state (denoted C^1) and the other state as the non-reference state (denoted C^2). The reference interval or state is related to the state of interest, here is the task state, and the non-reference interval or state can be considered as background, control or rest state. R-SS uses the concept of periodic independent component analysis or common spatial pattern which can be solved using the generalized eigenvalue decomposition (GEVD) (Sameni et al 2008; Koles et al 1990). Moreover, it uses the eigenvalues to estimate the number of HCTP sources (denoted i^*). The HCTP sources are the sources with the highest correlation with the reference state, which is the paradigm of the task of interest. A brief review of R-SS is explained in Appendix 1. The EEG recordings can be assumed as a linear combination of uncorrelated sources:

$$X = AS \tag{3}$$

where, A is a $N \times N$ mixing matrix, a_j is its j-th column, S is a $N \times T$ matrix, containing uncorrelated sources, and s'_i is its i-th row. Without the loss of generality the temporal sources are normalized, which means that $SS' = I$, where I is an $N \times N$ identity matrix, and S' denotes the transpose of S. We use R-SS to extract the i^* HCTP sources. Therefore, Eq. (3) can be rewritten as:

$$X = [a_1 \ldots a_{i^*}][s_1 \ldots s_{i^*}]' + [a_{i^*+1} \ldots a_N][s_{i^*+1} \ldots s_N]'$$

We rename the HCTP part of A as $A^s = [a_1 \ldots a_{i^*}]$ and the rest is renamed as $A^{\bar{s}} = [a_{i^*+1} \ldots a_N]$. So Eq. (3) can be rewritten as:

$$X = X^s + X^{\bar{s}} = A^s[s_1 \ldots s_{i^*}]' + A^{\bar{s}}[s_{i^*+1} \ldots s_N]' \tag{4}$$

where the dimensions of A^s and $A^{\bar{s}}$ are $N \times i^*$ and $N \times (N - i^*)$, respectively.

As X and J are linearly related (if one neglects the additive noise) according to Eq. (1), J can also be estimated from S as:

$$J = BS \tag{5}$$

where B is an $M \times N$ unknown mixing matrix which predicts J from S.

We can partition J in two parts estimated by HCTP sources and other sources, as done for X:

$$J = J^s + J^{\bar{s}} = B^s[s_1 \ldots s_{i^*}]' + B^{\bar{s}}[s_{i^*+1} \ldots s_N]' \tag{6}$$

where the dimensions of B^s and $B^{\bar{s}}$ are $M \times i^*$ and $M \times (N - i^*)$, respectively.

We put the HCTP sources, estimated with R-SS, in the rows of the matrix $\hat{S} = [\hat{s}_1 \cdots \hat{s}_{i^*}]'$ and call it from now on the HCTP source space.

Spatial Localization

Source Space Projection

In this step, we project the EEG forward problem to the HCTP source space by multiplying the forward model with the matrix $\hat{\mathbf{S}}$. Therefore, the projected version of the forward problem would be as:

$$\mathbf{X}\hat{\mathbf{S}}' = \mathbf{GJ}\hat{\mathbf{S}}' + \mathbf{n}\hat{\mathbf{S}}' \tag{7}$$

Using Eqs. (4) and (6) and the fact that sources are uncorrelated and normalized ($\mathbf{SS}' \sim \mathbf{I}$) and sources are uncorrelated with noise ($\mathbf{n}\hat{\mathbf{S}}' \sim \mathbf{0}$), we would have:

$$\mathbf{A}^s\mathbf{I} + \mathbf{A}^{\bar{s}}\mathbf{0} = \mathbf{GB}^s\mathbf{I} + \mathbf{GB}^{\bar{s}}\mathbf{0} + \mathbf{0} \tag{8}$$

where $\mathbf{I}$ is an $i^* \times i^*$ identity matrix and $\mathbf{0}$ is a zero matrix. Therefore, $\mathbf{A}^s$ and $\mathbf{B}^s$ are related by the following equation:

$$\mathbf{A}^s = \mathbf{GB}^s \tag{9}$$

The i-th column of $\mathbf{B}^s$ contains the projection of the current dipole signal of all vertices on the i-th HCTP source. Therefore, the active vertices related to the i-th HCTP source have high value in the i-th column of $\mathbf{B}^s$. In other words, each column of $\mathbf{B}^s$ represents the corresponding spatial map of the HCTP sources. Equation (9) shows $\mathbf{A}^s$, the projection of the observations on the HCTP source space, is a linear combination of the spatial maps. As the brain activity is localized, most of the vertices are not active and have a small value (near zero) in $\mathbf{B}^s$, so $\mathbf{B}^s$ is a sparse matrix.

Weighted Elastic Net

A newly developed variable selection method (Zou and Hastie 2005), called elastic net (EN), can produce a sparse model with good prediction accuracy, while encouraging a grouping effect. Experimental results and simulations have demonstrated the superiority of the elastic net over LASSO (Zou and Hastie 2005).

Weighted elastic net (WEN) has also been developed in (Hong and Zhang 2010; Jun-Tao and Ying-Min 2010). For a linear model $\mathbf{a}_k^s = \mathbf{Gb}_k^s$ ($\mathbf{b}_k^s$ -an $M \times 1$ vector-is the k-th column of $\mathbf{B}^s$ and $\mathbf{a}_k^s$-an $N \times 1$ vector-is the k-th column of $\mathbf{A}^s$), weighted elastic net estimates are computed as:

$$\hat{\mathbf{b}}_k^s = \arg\min_{\mathbf{b}_k^s} \|\mathbf{a}_k^s - \mathbf{Gb}_k^s\| + \lambda_2\|\mathbf{W}_2\mathbf{b}_k^s\| + \lambda_1\|\mathbf{W}_1\mathbf{b}_k^s\|_1$$

where $\|.\|$ and $\|.\|_1$ represent the ℓ_2 and ℓ_1 norms, respectively, and $\mathbf{W}_1$ and $\mathbf{W}_2$ are $M \times M$ weighting matrices. This optimization problem can be written as:

$$\begin{aligned} &\underset{\mathbf{b}_k^s}{\text{minimize}} &&\|\mathbf{W}_1\mathbf{b}_k^s\|_1 \\ &\text{subject to} &&\|\mathbf{a}_k^s - \mathbf{Gb}_k^s\| + \lambda_2\|\mathbf{W}_2\mathbf{b}_k^s\| < \delta \end{aligned}$$

Here, we replace ℓ_1 with ℓ_0 (denoted $\|.\|_0$) in the weighted elastic net and call it WEN0:

$$\begin{aligned} &\underset{\mathbf{b}_k^s}{\text{minimize}} &&\|\mathbf{W}_1\mathbf{b}_k^s\|_0 \\ &\text{subject to} &&\|\mathbf{a}_k^s - \mathbf{Gb}_k^s\| + \lambda_2\|\mathbf{W}_2\mathbf{b}_k^s\| < \delta \end{aligned}$$

Using the following change of variables (Hong and Zhang 2010):

$$\mathbf{G}^* = (1 + \lambda_2)^{-\frac{1}{2}}\begin{pmatrix} \mathbf{G} \\ \sqrt{\lambda_2}\mathbf{W}_2 \end{pmatrix}$$

$$\mathbf{a}_k^* = \begin{pmatrix} \mathbf{a}_k^s \\ \mathbf{0} \end{pmatrix}$$

$$\mathbf{b}_k^* = \sqrt{1 + \lambda_2}\mathbf{b}_k^s$$

WEN0 can be written in the simplified optimization problem:

$$\begin{aligned} &\underset{\mathbf{b}_k^*}{\text{minimize}} &&\|\mathbf{W}_1\mathbf{b}_k^*\|_0 \\ &\text{subject to} &&\|\mathbf{a}_k^* - \mathbf{G}^*\mathbf{b}_k^*\| < \delta \end{aligned} \tag{10}$$

Using the fast algorithm proposed by Mohimani et al (2009) and Babaie-Zadeh et al (2012) to solve Eq. (10), we estimate $\mathbf{b}_1^s, \ldots, \mathbf{b}_{i^*}^s$ which are the columns of $\mathbf{B}^s$. Stability against the noise and convergence of the algorithm are discussed in details in Babaie-Zadeh and Jutten (2010). In our problem, $\mathbf{a}_k^s$ (an $N \times 1$ vector) is the k-th column of $\mathbf{A}^s$ which is the projection of the EEG recordings onto the k-th HCTP source, and the weighting matrices $\mathbf{W}_2$ and $\mathbf{W}_1$ are $N \times N$ diagonal matrices initialized by the inverse of the spatial map extracted from the fMRI data. The fMRI spatial map is the thresholded F-map which is normalized by its maximum value and it is registered on the mesh vertices using SPM8, which uses the value of the nearest neighbor. It is defined as SM, which is a $M \times 1$ vector. Denoting SM_i the i-th component of SM, the diagonal elements of $\mathbf{W}_1$ and $\mathbf{W}_2$ are initialized as:

$$diag(\mathbf{W}_1)_i = diag(\mathbf{W}_2)_i = \frac{1}{(1-\alpha) + \alpha SM_i} \tag{11}$$

SM provides *maximum* values while the EEG localization is a *minimization* problem [Eq. (10)] so the *inverse* of fMRI spatial map is used in the weighting matrix. In other words, a low value for a given vertex implies that it is not crucial in the sparsity of the variable and the algorithm will focus on decreasing the values of the other vertices. Equation (11) can be replaced by any other well-behaved function in which the weights have the inverse relationship with the fMRI spatial map. The optimum function is not known and it will be investigated in future studies.

Although, we do not know exactly the relationship between EEG and fMRI, we extract the same feature from both i.e. spatial map of the HCTP sources. $\mathbf{b}_i^s$ is the spatial

map of the i-th HCTP temporal source, and it is weighted with the same feature (spatial map) from fMRI. In fact, this is very important because we are sure that we did not use any unreal assumption about the relationships between EEG and fMRI.

α can be estimated experimentally. $\alpha = 0$ removes the effect of the spatial map of the fMRI data and $\alpha = 1$ constrains the results strictly to the spatial map of the fMRI. Experimentally (Liu et al 1998), the best value of the parameter α [of Eq. (11)] is equal to 0.9 and this value has been used by others in the fMRI-informed EEG methods. However, it is not proved theoretically (Dale et al 2000; Babiloni and Cincotti 2005). Here, we do some experiments (see "Effect of parameter α" section) to check if the proposed value is also optimum for our experiment.

For i^* times (number of HCTP sources), we solve Eq. (10) for $\mathbf{b} = \mathbf{b}_k^s$ and $\mathbf{a} = \mathbf{a}_k^s$, $k = 1 \cdots i^*$, which are the columns of $\mathbf{B}^s$ and $\mathbf{A}^s$, respectively. Therefore, we estimate i^* spatial maps related to the i^* HCTP sources derived from R-SS. Now, we should investigate the HCTP spatial maps and extract a unique spatial map. To this goal, for finding the activation site or the place of optimum coefficients in $\mathbf{B}^s$, we use Pareto optimization (see Appendix 2).

Method Evaluation

Evaluation of the method is done in two steps. First, the method is applied to simulation data. Secondly, it is applied to experimental data (see detail in the next section).

Simulation data. Three simulation data sets are created as follows:

– The first data set includes 50 EEG-fMRI data. In this set, two sources are activated both in EEG and fMRI in the same area. This data set is produced for 6 different SNR values of the EEG data.
– The second data set includes 50 EEG-fMRI data. In this set, three sources are activated. Two of them are activated both in EEG and fMRI in the same area, and the third one, which is called fMRI extra source, is activated just in the fMRI data. In this data set, the SNR value of the EEG data is equal to 0 dB.
– The third data set includes 50 EEG-fMRI data. In this set, three sources are activated. Two of them are activated both in EEG and fMRI in the same area, and the third one, which is called EEG extra source, is activated just in the EEG data. In this data set, the SNR value of the EEG data is equal to 0 dB. We also use this data set to show the effect of the parameter α in (11).

To compare the performance of the methods, we use two measures. The first measure is the localization bias (LB), which is defined as Euclidean distance between the mean

site of the estimated activity area ($\hat{\boldsymbol{\mu}}$) and the actual mean site of the simulated activity area, $\boldsymbol{\mu}$.

$$LB = \|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\| \tag{12}$$

with

$$\hat{\boldsymbol{\mu}} = \frac{\sum_p \|\mathbf{b}_p\| \mathbf{r}_p}{\sum_p \|\mathbf{b}_p\|} \tag{13}$$

where, $\mathbf{r}_p$ is the MNI coordinate of a vertex and $\mathbf{b}_p$ is its related activation level, which is the p-th row of $\mathbf{B}^s$ ($\mathbf{b}_p = (\mathbf{b}_p^s)'$).

The second measure is the source distribution index (SDI), which is defined as the localization spread:

$$SDI = \frac{\sum_p \|\mathbf{b}_p\| \|\mathbf{r}_p - \hat{\boldsymbol{\mu}}\|}{\sum_p \|\mathbf{b}_p\|} \tag{14}$$

The details of the simulated data are explained in "3.1" section.

Experimental data. S4 has been also validated on experimental data and its group analysis is compared with the previous studies. The stability of the method against the value of α is investigated, using the idea of split-half resampling (Strother 2006). In this method, the subjects are divided in two equal groups. One group is used as a train data set, while the other one is used as test data set. The results, obtained from the train group is considered as ground truth and the results of the test group is evaluated using this ground truth.
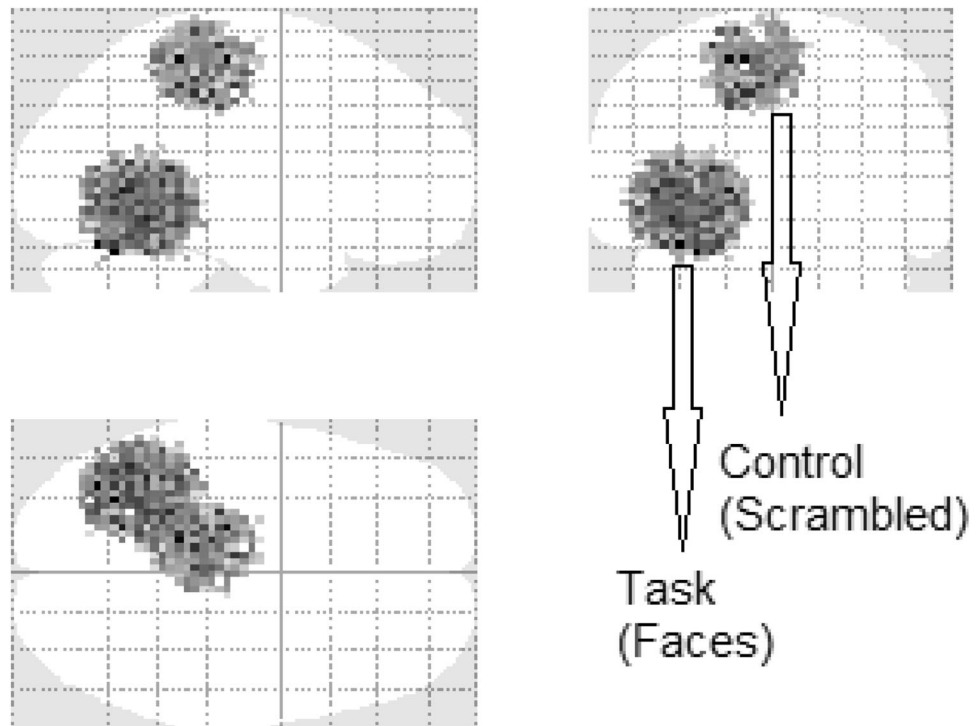
Data

Simulated Data

Monte Carlo modeling (sampling over randomized 50 source locations) is used to compare different methods with the proposed method. For each activation site two sets of EEG and fMRI data are created (details are given in the two next subsections). 50 random vertices are selected as the center of their active regions. The sources are assumed to be spread, so a smoothing Gaussian filter is used to model the spread sources.

Here, we simulate the EEG and fMRI data after the preprocessing step. It means that for example in fMRI, slice timing is not simulated and in EEG, the mean of the trials after preprocessing is simulated.

Sources are located in the grey matter. Grey matter is extracted from the tessellation of the real structural MRI of the first subject of our experimental data set. To be more realistic, we performed the source localization on the experimental EEG data and the spatial map is considered as an ROI (region of interest) in the simulated data set. We have put the sources in the ROI.

Fig. 3 Sample activation sites in simulated data (the figure is generated with xjView toolbox (http://www.alivelearn.net/xjview))

The experiment has two states: task and control. Onsets of task and control states are set as the onsets of the experimental data for face and scrambled states, respectively. All the parameters of imaging methods are chosen as the experimental data, like the electrode positions and sampling rate in EEG and TR and image dimension in fMRI.

fMRI

We use the package neuRosim in R (R Core Team 2012) for fMRI simulation, because the data generation in neuRosim is accurate and fairly fast and the parameters can be adjusted by the user (Welvaert et al 2011). A double-gamma function which is a linear combination of two gamma functions is used to model HRF (Friston et al 1998; Glover 1999). Fifty fMRI simulations are created with noisy measurements with a fixed SNR = 5 dB. In fact, the main known effect of low SNR in the result of our method is due to the number of false positive or false negative regions. Therefore, we evaluate the efficacy of our method in the presence of fMRI invisible or extra sources, at a fixed SNR, defined as:

$$SNR = 20 \log \left(\frac{\bar{S}}{\sigma_N} \right) \tag{15}$$

where $\bar{S}$ represents the average magnitude of the signal, and σ_N stands for the standard deviation of the noise (Krüger and Glover 2001). The noise is a mixture of

different random signals: (1) Rician system noise; (2) temporal noise of order 1; (3) low-frequency drift; (4) physiological noise; and (5) task-related noise (Krüger and Glover 2001). The background is set to the mean image of the experimental data.

In each fMRI data two sources are estimated: one is related to the task state and the other to the control state (see Fig. 3). The center of the spatial sources related to the task state is chosen randomly and the sources are extended with a radius of 10 mm. The control source is located at $(-10, -29, 55)$ mm in the Montreal Neurological Institute (MNI) coordinates and is extended as a sphere with a radius of 7 mm and a fading ratio of 0.01. The task sources are located at least 6 mm far from the control source.

Extra source in fMRI. For evaluating the effect of extra fMRI sources which cannot be detected in the EEG data, another data set is produced by appending another active region in the fMRI data but not in the EEG data. The extra region is supposed to be located in a fixed area and the other source which is active both in EEG and fMRI in the same area (common source) is chosen randomly in the ROI at least 8mm far from the extra source in fMRI.

EEG

For each fMRI data, the related EEG data is generated using the linear forward model [Eq. (1)]. The gain matrix is calculated using the structural MRI of a real subject and the

position of the electrodes in experimental data acquisition. The time courses of the sources are generated as an additive source problem in which two main sources are distinguished, namely (1) the activation caused by an experimental design, and (2) the background. The experiment is defined by two states, task and control trials. The details are as follows.

- *Activation site*. We assume 50 random sites to put a dipole related to the activation of interest (task). These random sites are the same for EEG and fMRI. We also assume a fixed site for a dipole related to the control state (like in fMRI it is located at $(-10, -29, 55)$ mm in the MNI coordinates).
- *Activation time courses*. For simulating the activation time courses, we extract EEG source time courses from the experimental data. To extract the task and control time courses, the R-SS method is used on the experimental data. To do this, alternatively, the task interval (faces) or control state (scrambled faces) is assumed to be the reference state, for estimating the task temporal course or the control temporal courses, respectively.
- *Background time course*. It is extracted from experimental data. It can be found by extracting the common sources between the task and control trials, $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively. Common source extraction is done by canonical correlation analysis (CCA) (Hotelling 1936). We use the first common source as an estimation of the background EEG signal. By back reconstruction, we compute the EEG background in the electrode space.
- *Time courses of the electrodes*. To compute the electrode time courses, we use the forward linear model [Eq. (1)]. To estimate the gain matrix, we use the structural MRI of a real subject with its fiducials and electrode positions. For this purpose, we use BEM with 8196 nodes developed in SPM8 (http://www.fil.ion.ucl.ac.uk/spm). As EEG is sensitive to the signal of the Pyramidal neurons of the cortex, which are perpendicular to the cortical surface, we assume that the dipoles are on the cortical grey matter, where each dipole is perpendicular to the cortical sheet. We put the task and control time courses in the related rows of the matrix $\mathbf{J}$. We put the estimated activation sources in the predefined sites and use a Gaussian filter with full width at half maximum (FWHM) of 8 mm to estimate the activation of the neighboring mesh vertices. After computing the effects of activations on the electrode space, we add the background signal to the electrodes with different SNR values in the range of $[-20, 5]$ dB with a step size of 5 dB.

Extra source in EEG. For evaluating the effect of extra EEG sources which cannot be detected in the fMRI data, another active region in the EEG data is added with a weak signal power which is not detected in the fMRI data.

Experimental Data

Experimental data is taken from ftp://ftp.mrc-cbu.cam.ac.uk/personal/rik.henson/wakemandg_hensonrn/. A brief summary of the acquisition characteristics, and preprocessing steps are brought in this section. More details can be found in (Henson et al 2009, 2011).

Participants

The participants are sixteen healthy young adults (eight female). The study protocol is approved by the local ethics review board (CPREC reference 2005.08). Each subject performs two days experiments; once for concurrent MEG + EEG and once for fMRI + sMRI. The protocol is the same each day, but the fMRI and EEG data are not recorded simultaneously.

Stimuli and Tasks

The paradigm is similar to that used previously under EEG, MEG, and fMRI (Henson et al 2009, 2011). A central fixation cross (presented for a random duration of 400–600 ms) is followed by a face or scrambled face (presented for a random duration of 800–1000 ms), followed by a central circle for 1700 ms.

EEG Acquisition

The EEG data is recorded simultaneously with MEG with a VectorView system (Elekta Neuromag, Helsinki, Finland). EEG data is recorded with 70 Ag–AgCl electrodes with the sampling rate of 1.1 kHz. The low-pass filter with the cut-off frequency of 350 Hz is applied on EEG data.

An elastic cap (EASYCAP GmbH, Herrsching- Breitbrunn, Germany) is used according to the extended 10–10 and a nose electrode is used as the recording reference. Vertical and horizontal EOG (and ECG) are also recorded.

fMRI + sMRI Acquisition

Structural MRI (sMRI) is a T1-weighted MPRAGE-MRI scan, which was acquired for each subject with voxel size $1 \times 1 \times 1$ mm, on a 3T Trio (Siemens, Erlangen, Germany) with TR = 2250 ms, TE = 2.99 ms, flip-angle = 9° and acceleration factor = 2.

The fMRI volumes comprise 33 T2-weighted transverse echoplanar images (EPI) (64×64, 3 mm $\times$ 3 mm pixels, TE = 30 ms) per volume, with blood oxygenation level dependent (BOLD) contrast. EPIs comprise 3 mm thick axial slices taken every 3.75 mm (3 mm thick with a 0.75 mm gap), acquire sequentially in a descending direction. 195 volumes were collected continuously with a repetition

time (TR) of 2.25s. The first six volumes were discarded to allow for equilibration effects.

EEG Preprocessing

The steps of preprocessing are as follows:

- Epoching from -500 to $+1000$ ms poststimulus onset.
- Down-sample to 250 Hz (using an antialiasing low-pass filter with a cutoff frequency of approximately 100 Hz).
- Reject the epochs with the EOG exceeded 150 µV (number of rejected epochs ranged from 0 to 47 across participants, median = 5).
- Re-reference the EEG data to the average over non-bad channels. The number of bad channels, which are rejected, varies between 0 to 7, according to the participants.

sMRI Processing

Structural MRI is used to model the forward problem. sMRI images of each participant are segmented and spatially normalized to an MNI T1 template brain in Talairach space. The inverse of the normalization transformation is then used to warp a cortical mesh from a template brain in MNI space to each participant's MRI space. The mesh has 8196 vertices (4098 per hemisphere) with a mean inter-vertex spacing of 5 mm. The lead-field matrix is then calculated for the dipoles at all points in the cortical mesh, oriented normal to the mesh. All is done in SPM.

fMRI Processing

The processing steps are performed with SPM8 (http://www.fil.ion.ucl.ac.uk/spm). General linear model (GLM) is used to model BOLD response and maximum likelihood estimation is used to estimate the parameters. A statistical parametric map (SM) of the F-statistic that compare faces against scrambled faces was thresholded for $p < 0.05$ (family-wise error corrected across the whole-brain) and the regions of at least 10 contiguous voxels are considered as active regions.

This analysis produces clusters in the fusiform and occipital face areas (FFA and OFA, respectively).

Results

The efficiency of S4 is compared with two other methods. The first method is the multiple sparse priors (MSP) (Friston et al 2008) with or without the fMRI spatial map as prior information. MSP is implemented in SPM8 (http://www.fil.ion.ucl.ac.uk/spm). We also compare S4 with

sparse Bayesian learning (SBL) with or without the fMRI spatial map as prior information (Wipf and Nagarajan 2009). Among the integration methods MSP (Friston et al 2008) and SBL (Tipping 2001) are the most similar methods to the proposed one, as they use the spatial sparsity of the activated regions. In these methods the active sources can be correlated and any extra information which gives information about the covariance matrix of the sources can be used e.g. the fMRI results. In these methods the information about the experiment is not used, while we have used it in the proposed method.

Simulated Data

Spatial Accuracy

To compare the spatial accuracy of different methods, mean values and standard deviations of LB (Eq. 12) are calculated for all methods (see Fig. 4). Also, mean values and standard deviations of SDI (Eq. 14) are represented in Fig. 5.

Figure 4 shows the mean localization bias of different methods in different SNR values. Comparing the results of MSP and SBL when fMRI information is used or not, shows that using fMRI information results in better resolution especially in lower SNR values. The figure also shows that SBL and S4 have the same accuracy. It is the result of using fMRI information directly to solve the EEG inverse problem and using the ℓ_0 norm. In addition, S4 is more robust to noise. This is because the reference-based source separation method removes the common sources between the two states and thus, it removes the background effect as much as possible.

Figure 5 shows the mean source distribution index (SDI) of different methods in different SNR values. The mean SDI of the actual source is 15 mm. Therefore, the results of S4 and SBL are more accurate. In contrary, MSP estimates sources with a too small SDI.

Effect of Parameter α

Extra Source in fMRI. For $SNR_{EEG} = 0$ dB, repeating the Monte Carlo simulation, using the second simulated data set, shows that the results would not change even for $\alpha = 0.9$ when there exists an extra source in fMRI, which is not detectable in EEG. In this case the mean of LB is equal to 18 mm, and no extra region has been found. It means that our method will miss the regions which are in fMRI but not in EEG, while the accuracy of the estimated activity site does not change much. It changes from 10 mm (See Fig. 4) to 18 mm (LB of the common source in the presence of the extra source in fMRI). In contrary, MSP and SBL fail to estimate even the common source site.

Fig. 4 Mean of localization bias over 50 random activation sites compared for five methods (MSP, MSP + fMRI, SBL, SBL + fMRI, S4)
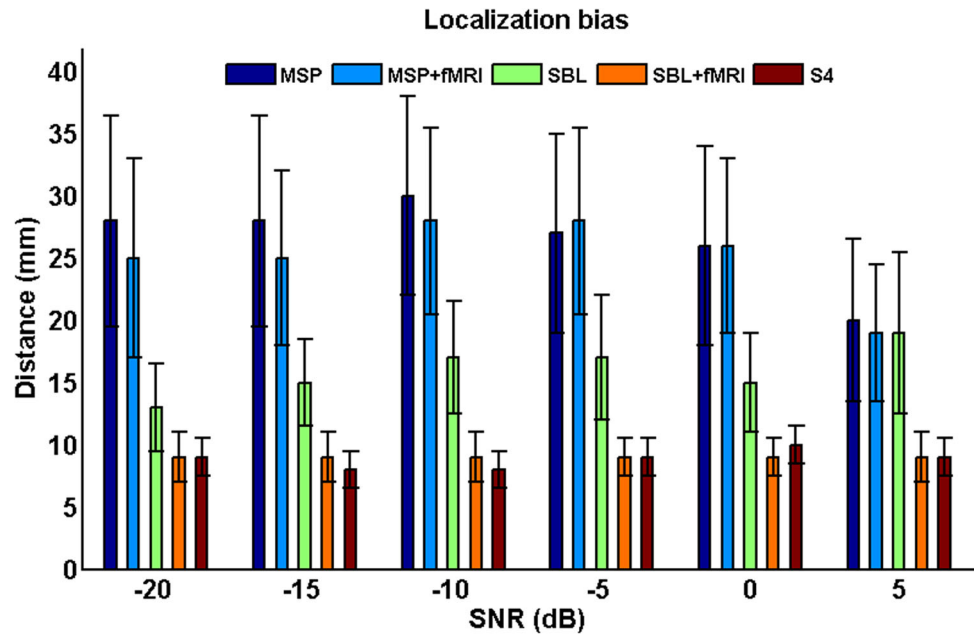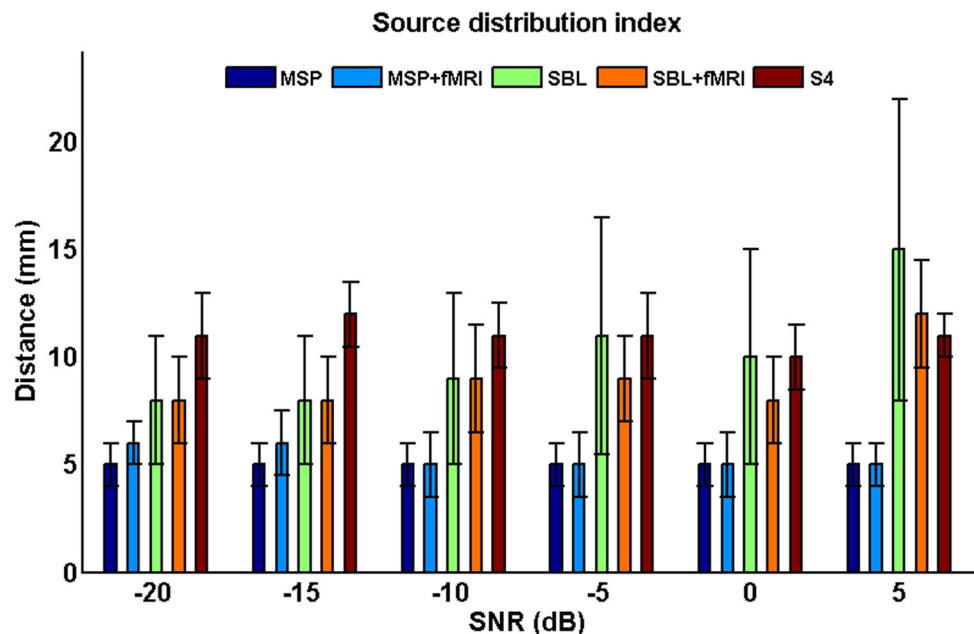

Fig. 5 Source distribution index over 50 random activation sites compared for five methods (MSP, MSP + fMRI, SBL, SBL + fMRI, S4)

Extra Source in EEG. In the case of the existence of an extra region, which is in EEG but not in fMRI, repeating the Monte Carlo simulation for 50 random common source localization and fixed extra source in EEG shows that for $SNR_{EEG} = 0$ dB and $\alpha = 0.9$, mean of LB is 13 mm for the common source between EEG and fMRI, and no extra region has been found in 45 experiments. α can balance the role of fMRI information in EEG inverse problem [see Eq. (11)]. To see the effect of α, we have repeated the experiment for $\alpha = 0.1, \dots \alpha = 0.9$. Figure 6 shows LB for three different values of α calculated for the common

source. Table 1 shows the percentage of the EEG extra source detection in the 50 EEG-fMRI data sets, and Table 2 shows the percentage of the common source detection in the 50 EEG-fMRI data sets. Detection happens when S4 finds a region less than 20 mm far from the EEG extra source.

Although $\alpha = 0.9$ is the optimum value, which is suggested in (Liu et al, 1998), in our experiments $\alpha = 0.7$ is the optimum value. Tables 1 and 2 show that for $\alpha = 0.7$, we have highest detection percentage of the extra source while the common source is estimated correctly in all experiments.
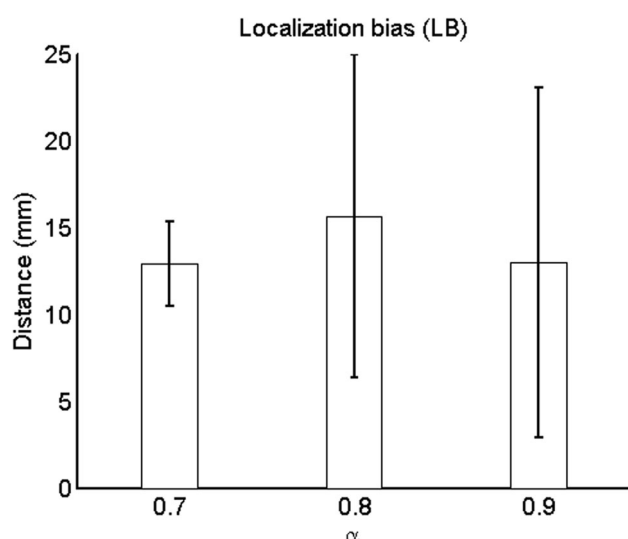
Fig. 6 Localization bias in the presence of an extra EEG source as a function of α

Table 1 Detection percentage of the extra EEG source

α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Detection (%)	58	58	52	56	52	46	54	20	10

Table 2 Detection percentage of the common EEG source

α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Detection (%)	36	46	46	54	68	100	100	100	100

We have applied MSP and SBL on the same data set, in which there exists an extra EEG source. These methods did not converge to the real solution and they fail to estimate even the common source site.

Experimental Data

Spatial Sources

The maps of spatial sources for group analysis of 16 subjects, using MSP and S4, are shown in Figs. 7 and 8, respectively. A zoom of the spatial map extracted by S4 algorithm is shown in Fig. 9. For group analysis, the EEG data were analyzed for each subject and SPM is used to do group analysis. One-sample t-test is used and the t-statistics is thresholded for $p < 0.05$ and the regions of at least 10 contiguous voxels are considered as active regions.

The other methods, MSP and SBL, with fMRI priors could not find a common region between different subjects as their solutions are too sparse, and they did not consider any grouping constraint.

Effect of Parameter α

To evaluate the stability of S4, we count the false positive and false negative regions using the idea of split-half resampling (Strother 2006). For this purpose, we partition the EEG data into two non-overlapping groups each with 8 subjects. For each group, the spatial map is extracted using S4. One group is assumed as the ground truth and the other as the test one. To evaluate false positives and false negatives, the localization bias of the regions in the test image is compared with the localization bias of the regions in the ground truth one. If any region in the test data is at least 20 mm far from all regions in the ground truth, it would be false positive. In reverse, if any region in the ground truth is at least 20 mm far from all regions in the test data, it would be false negative. 20 mm is chosen because the best resolution of MSP in the simulation was around 20 mm (see Fig. 4) and it is assumed that the regions that are less than 20 mm far from each other are acceptable as
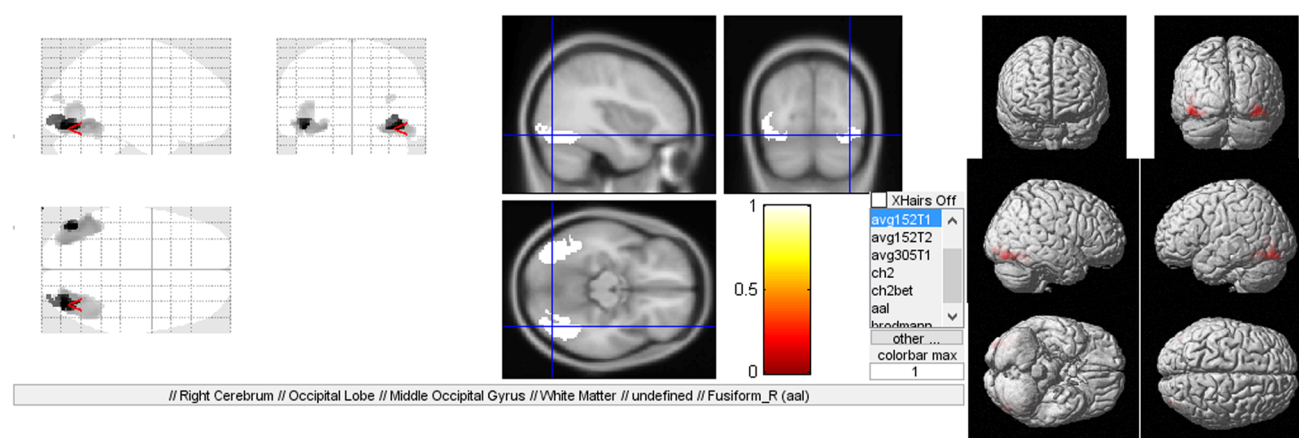

Fig. 7 Activation sites in the experimental data of 16 subjects by MSP (the figure is generated with xjView toolbox, http://www.alivelearn.net/xjview, and scale unit = 1)

// Right Cerebrum // Occipital Lobe // Middle Occipital Gyrus // White Matter // undefined // Fusiform_R (aal)

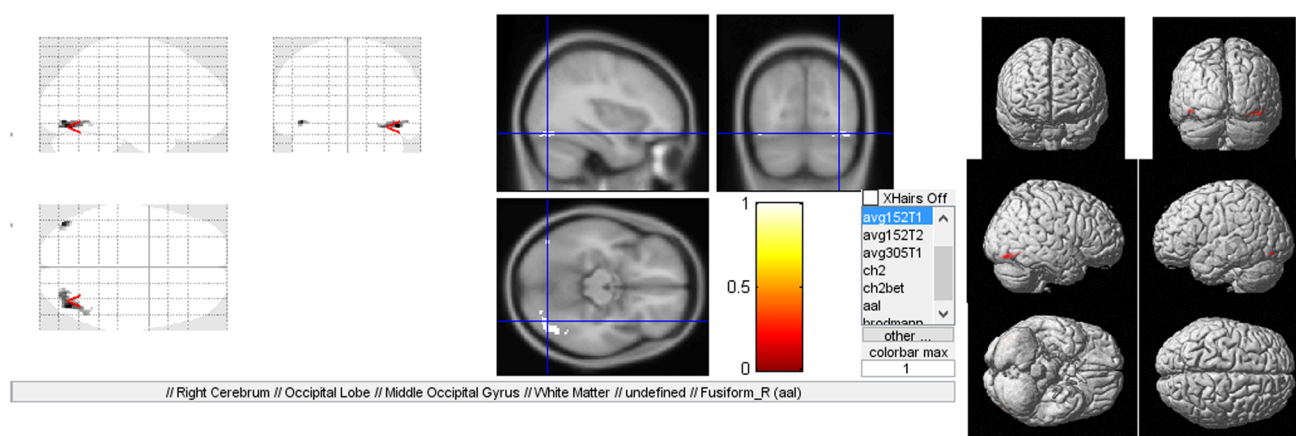**Fig. 8** Activation sites in the experimental data of 16 subjects by the proposed method (the figure is generated with xjView toolbox, http://www.alivelearn.net/xjview, and scale unit = 1)

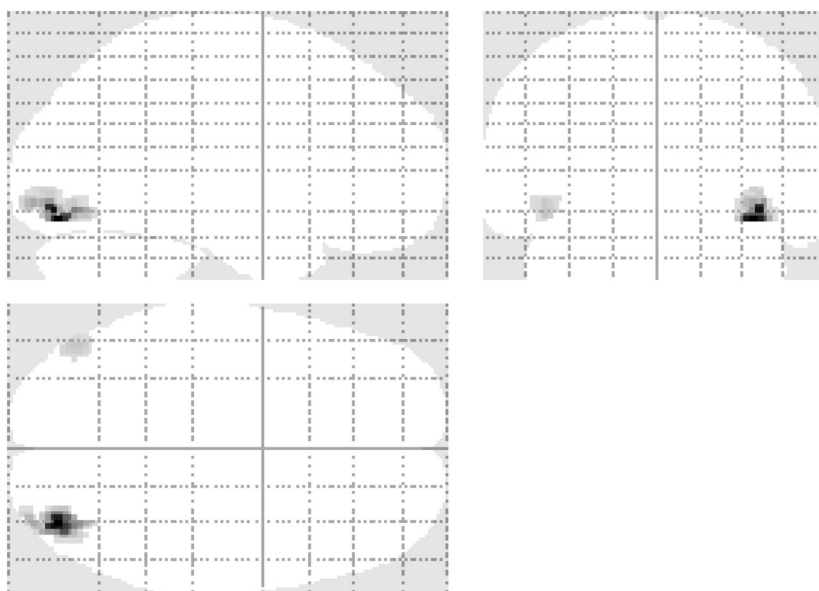**Fig. 9** Zoom of the activation sites in the experimental data of 16 subjects by S4 (the figure is generated with xjView toolbox, http://www.alivelearn.net/xjview)

Table 3 Effect of α in group analysis

α For the train group	0.7	0.9
α For the test group	0.7	0.9
Number of false positive regions	2	0
Number of false negative regions	8	2

close regions. This test is repeated 20 times for 20 different partitioning. The results for two different values of α (0.9 and 0.7) are shown in Table 3. For $\alpha = 0.9$, no false positive region is detected in all experiments and only two false negative regions occurred in all experiments. For $\alpha = 0.7$, two false positive regions are detected in all experiments and eight false negative regions occurred in all experiments. The number of regions detected in each experiment is between 1 and 5 (mean

= 3). The results show that $\alpha = 0.9$ is more stable against subject changes which was expected, because we have used the fMRI group analysis as the constraint, and the results of the EEG group analysis will be more restricted to the fMRI results, with higher value of α. Therefore the search space in different subjects is limited to the same space (fMRI group analysis spatial map) and the variability between subjects would be ignored in higher value of α.

Discussion

Although EEG-fMRI integration is intensively investigated in recent years, some challenges still exist as open problems. The most important challenge is the unknown relationship

between EEG and fMRI data. In this study, we investigate the EEG inverse problem from a new point of view. By projecting the EEG inverse problem to HCTP source space, we accurately associate the temporal space to the spatial space.

The projection basis used in S4 (HCTP sources) directly results in the spatial maps of interest, which are assumed to be sparse. Therefore, we can use sparse decomposition methods: in this work, the ℓ_0 norm regularization leads to high accuracy and stability against the noise.

After projection, the unknown variables are HCTP spatial maps, so we could use the spatial map of activation, extracted from fMRI, in the weighting matrix. Therefore, we solve the integration inverse problem using the same feature extracted from fMRI and EEG. This strategy was a remedy for overcoming (without assumption or estimation) the unknown relationships between EEG and fMRI.

However, if we have a more accurate information about the relationship between EEG and fMRI data, it may improve the results. Especially, we could optimize the weighting matrices, instead of choosing an arbitrary function to use the fMRI spatial map in the weighting matrix.

Another challenge in EEG-fMRI integration is the mismatch between EEG and fMRI sources. The proposed method has a good stability against an extra region in fMRI. Although it was not capable of detecting EEG extra regions, it achieves an acceptable detection rate of common sources between EEG and fMRI in the presence of mismatches.

Without definitely solving the open questions of EEG-fMRI integration, experiments show the S4 method achieves a step forward to use fMRI spatial maps in the EEG inverse problem with more feasibility.

The results are discussed in details as follows.

Simulated Data

The results of the method on the simulated data and its comparison with two methods, MSP and SBL, show that the spatial maps estimated by S4 and SBL, which minimize the ℓ_0 norm, have smaller mean and variance of localization bias (LB) for all SNR values, compared to results obtained with MSP (Fig. 4). SBL and S4 also estimate the mean of source distribution index (SDI) more accurately, while S4 has less variance of SDI (Fig. 4). Therefore, SBL and S4 are more accurate and stable against the noise.

Invisible sources either in fMRI or EEG result in an inconsistent constraint (here, it is a weighting matrix that contains fMRI information) in the EEG inverse problem. In this case, MSP and SBL fails even to estimate the common sources between EEG and fMRI, while S4 correctly estimates the common sources.

The activation regions which are missing in fMRI data may be related to neural activities which have a short

activation time or involve a few numbers of neurons that cannot produce a detectable increase of cerebral blood flow (Babiloni et al 2004). Missed regions can also be due to the analysis method. The result shows that the effect of missed regions in fMRI is dependent on the balancing variable α in Eq. (11). With α = 0.9 fMRI invisible sources are also missed in the EEG-fMRI result, but with α = 0.7, the EEG extra source could be detected in 54 percent of our experiments. Figure 6 and Tables 1 and 2 show that, by decreasing the α value, the probability of the extra source detection would increase while the mean of LB (localization bias) would not change much. The α parameter can be considered as a user-defined parameter for the method. User can decrease the α value to see if there exists any extra region in EEG. However, the optimum value of α is an open question by now.

Some brain cells (stellate cells) and cortical regions (thalamus) produce closed electromagnetic field. This field cannot be detected on the scalp electrodes, although their activation requires high blood flow. This would also happen when "sources of activation cancel each other according to the orientation disparity of the cortex" (Irimia et al 2012). Therefore, these neuronal populations present high metabolism requirements that can be detected by the fMRI technique, while at the same time they are invisible in the EEG modality (Babiloni and Cincotti 2005). We have evaluated the algorithm S4 when it exists an extra source in fMRI that is invisible in EEG. In this case, S4 increases the mean of the localization bias from 10 mm to 18 mm, but again, it outperforms the other methods, which fail to estimate the common source in the existence of an extra source.

Experimental Data

S4 algorithm is applied on experimental EEG and fMRI data of a face perception experiment. Spatial sources are shown in Figure 8. These activation regions are located in the fusiform face area (FFA) and the occipital face area (OFA). These clusters are in general agreement with previous studies reporting a similar contrast of faces versus scrambled faces (Henson et al 2009, 2011). Previous studies investigated the role of these two regions in face perception and they proposed that the OFA is the first stage in a hierarchical face perception network in which the OFA represents facial components (e.g., eyes, nose, and mouth) (Pitcher et al 2011), while the FFA encodes both the facial components and the configural information of faces (i.e., spatial relation among features) (Zhang et al 2012). The most important factor here is that this activation area should be similar in all participants. In other words, the activation site should have little spatial variation in different participants. The results show that S4 algorithm

can find a common region between subjects, while MSP and SBL are not able to estimate a common region between different subjects because their solutions are sparse but not block-sparse. In fact, we have counted the false positive regions and false negative regions using the split-half resampling algorithm proposed in (Strother 2006). The results show high stability or low spatial variation over different participant groups with $\alpha = 0.9$ in which the results are strictly biased with fMRI spatial map. For $\alpha = 0.7$, more false positives and false negatives are detected, but again it is acceptable.

Conclusion

We propose a new method, called S4, to integrate scalp EEG recording and fMRI data for the identification of the active regions of the brain with high spatial resolution. Using R-SS, a semi-blind source separation method, we estimate the spatial filter that maximizes the power of reference over non-reference states and provides temporal sources with high correlation with the paradigm of the task of interest (called HCTP sources). The EEG signal was projected to the HCTP source space. The inverse problem of EEG in the HCTP source space can be considered as a sparse decomposition problem. The unknown variables are the spatial maps of the HCTP sources. As we focus on the localized sources, their spatial maps are sparse. For promoting block-sparsity, we extend the elastic net idea by combining ℓ_0 and ℓ_2 norms. This algorithm, called WEN0, takes into account the fMRI information through weighting matrices related to the F-map of fMRI extracted with GLM.

In more details advantages of the proposed method are as follows.

–   The simulation results show that the localization quality of S4 and SBL, measure with localization bias (LB) and source distribution index (SDI), is 50 percent better than MSP. The results of S4 is significantly stable against the noise.
–   The stability of S4 against false positive and false negative regions in the prior information is investigated. S4 can estimate the common source between EEG and fMRI in the existence of an extra region and missed region in the fMRI data, while both SBL and MSP fail to estimate the sources.
–   Using Pareto optimization helps us in avoiding additional statistical methods and thresholding.

The drawback of S4 is that the final result is dependent on the first source separation step and any mistake in the first step will propagate to the rest. However, we did not obtain any false discovery in the simulation experiments, even for the lowest SNR value.

In future works, S4 may be applied to simultaneous EEG and fMRI data. Defining different weighting matrices for the ℓ_2 and ℓ_0 norms may be investigated and also, the optimized value of α, which controls the contribution of fMRI spatial map in the weighting matrix, can be defined using the data.

Appendix 1: Referenced-Based Source Separation

The objective is to extract the sources related to a reference activation model. We consider two states, denoted C^1 and C^2, which correspond to the reference and non-reference activations, respectively. Denoting $\mathcal{T}^\ell$, $\ell = 1, 2$, the set of time samples related to each state, we can build the corresponding segment matrix, $\mathbf{X}^\ell \in \mathbb{R}^{N \times M^\ell}$. The correlation matrix of data for each state can be estimated as:

$$\widehat{\mathbf{R}}^\ell = \frac{1}{M^\ell} \mathbf{X}^\ell \mathbf{X}^{\ell \prime}. \tag{16}$$

The spatial filters, $\mathbf{W}$ (whose columns are generically denoted $\mathbf{w}$), for which the temporal sources, $\mathbf{S} = \mathbf{W}'\mathbf{X}$ have maximum similarity with the reference activation state, i.e., maximum variance in the reference state compared to the other state, is computed as:

$$\max_{\mathbf{w}} \frac{\mathbf{w}'\widehat{\mathbf{R}}^1\mathbf{w}}{\mathbf{w}'\widehat{\mathbf{R}}^2\mathbf{w}} \tag{17}$$

Solving (17) leads to generalized eigenvalue decomposition (GEVD) of $(\widehat{\mathbf{R}}^1, \widehat{\mathbf{R}}^2)$:

$$\widehat{\mathbf{R}}^1\mathbf{W} = \widehat{\mathbf{R}}^2\mathbf{W}\boldsymbol{\Lambda} \tag{18}$$

Using $\mathbf{W}$, the spatial patterns, $\mathbf{A} = (\mathbf{W}')^{-1}$, and the temporal sources, $\mathbf{S} = \mathbf{W}'\mathbf{X}$, are extracted. The maximum eigenvalue in (18) is related to the maximum power ratio in (17). We rank the eigenvalues in decreasing order. This implies ranking of the estimated temporal sources, according to their resemblance to the reference activation state. Here, we propose a simple method to define the number of sources. Remember that the simplicity of this method is due to two facts; (1) the usage of experiment information in the source separation step, and (2) a special interpretation of the eigenvalues in the source selection step, which will be explained below.

After obtaining the discriminative sources ($\mathbf{s}_i$) between the reference and non-reference states, sources are ranked

according to their similarity to the reference state. Now, we need to select the sources, which are similar enough to the reference, for being considered as belonging to the reference class. To this end, we propose the following procedure.

The probability of the reference class (ω_1) membership is calculated as follows:

$$p(\mathbf{s}'_i \in \omega_1) = \frac{\lambda_i}{max(\lambda_{j=1,\ldots,N})} \qquad (19)$$

where λ_i, $i = 1, \ldots, N$ indicate the eigenvalues or the diagonal elements of Λ in (18). For the classification of the sources, the error probability using Bayes rule is defined as:

$$p_{error} = \sum_{j=1}^{N} \Big(p(\mathbf{s}'_j \in \omega_2 | \omega_1) p(\omega_1) \Big)$$
$$+ \sum_{j=1}^{N} \Big(p(\mathbf{s}'_j \in \omega_1 | \omega_2) p(\omega_2) \Big)$$

where $p(\mathbf{s}_j \in \omega_2 | \omega_1) = 1 - p(\mathbf{s}_j \in \omega_1)$ and $p(\mathbf{s}_j \in \omega_1 | \omega_2) = 1 - p(\mathbf{s}_j \in \omega_2)$ as ω_1 and ω_2 constitute a partition. $p(\omega_1)$ and $p(\omega_2)$ are the prior probabilities of the reference and non-reference classes, respectively. We remind that GEVD sorts the separated sources in the decreasing order of similarity with respect to the reference. Therefore, if we assume that only the first i sources belong to the reference class (and consequently the $N - i$ others belong to the non-reference class), then the total error probability (false positive plus false negative errors) can be written as follows:

$$p_{error}(i) = \sum_{j=1}^{i} p(\mathbf{s}'_j \in \omega_2 | \omega_1) p(\omega_1)$$
$$+ \sum_{j=i+1}^{N} p(\mathbf{s}'_j \in \omega_1 | \omega_2) p(\omega_2) \qquad (20)$$

where $p(\omega_1) = \frac{i}{N}$ and $p(\omega_2) = \frac{N-i}{N}$. Thus, the minimum of $p_{error}(i)$ provides the number i^* of the sources in the reference class, i.e., $i^* = \arg \min_i p_{error}(i)$.

Appendix 2: Pareto Optimization

$\mathbf{B}^s$ is a $M \times i^*$ matrix whose columns present the contributions of the related sources in each mesh vertex. All the i^* sources are involved in the neural activation. Therefore, all of them are important, and if we use the sum or the square sum of the contributions the weak sources would be hidden in the shadow of the sources with higher power. The remedy of this problem is to use multi-objective optimization method, called Pareto method (Deb 1999). Therefore, we are able to take into account the individual effect of each source. Pareto optimization finds the

optimum without any threshold, that obviates the need for arbitrary user-defined threshold.

A multi-objective optimization problem, in Pareto sense, has the following form:

$$\begin{aligned} \text{maximize} \quad & (\mathbf{b}^s_i)' & \text{for i} = 1, \ldots, \mathrm{M} \\ \text{subject to} \quad & (\mathbf{b}^s_i)' \in P \subset \mathfrak{R}^{\mathrm{i}^*} \end{aligned} \qquad (21)$$

It then consists of i^* objective functions that are aimed to be maximized simultaneously. $(\mathbf{b}^s_i)'$ is the i-th row of the matrix $\mathbf{b}^s$ which shows the contributions of the sources at the ith vertex. From the geometrical point of view, each $(\mathbf{b}^s_i)'$ can be considered as a point in a i^*-dimensional space. In Pareto optimization the non-dominated points should be chosen as the optimum points (Deb 1999): a point is non-dominated if either it dominates the others, or there is no other point dominating it. Point $(\mathbf{b}^s_i)'$ dominates point $(\mathbf{b}^s_k)'$, if $\forall l$, $\mathbf{b}^s_i(l) \geq \mathbf{b}^s_k(l)$, and $\exists l^*$, $\mathbf{b}^s_i(l) > \mathbf{b}^s_k(l^*)$,[3] where $\mathbf{b}^s_i(l)$ is the l-th component of the vector b^s_i. The set of all non-dominated points is called non-dominated layer. Let us consider M i^*-dimensional decision vectors, $(\mathbf{b}^s_i)'$, as M points in the search space P. The non-dominated layer, denoted by $D(P)$, is obtained using the following Pareto optimization algorithm (Deb 1999):

1. Initialize $D(P)$ with the first point ($i = 1$) with the value of $(\mathbf{b}_1)'$. This can be any point.
2. Choose a new point ($i = i + 1$):

   (a) If any node in $D(P)$ dominates point i go to Step 3.
   (b) Else add point i to $D(P)$ and remove any points of $D(P)$ that point i dominates.

3. If i is not equal to M go to Step 2.

References

Babaie-Zadeh M, Jutten C (2010) On the stable recovery of the sparsest overcomplete representations in presence of noise. IEEE Trans Signal Process 58(10):5396–5400. doi:10.1109/TSP.2010.2052357

Babaie-Zadeh M, Mehrdad B, Giannakis GB (2012) Weighted sparse signal decomposition. In: Proceedings of ICASSP2012, Kyoto, Japan, pp 3425–3428. doi:10.1109/ICASSP.2012.6288652

Babiloni F, Mattia D, Babiloni C, Astolfi L, Salinari S, Basilisco A, Rossini PM, Marciani MG, Cincotti F (2004) Multimodal integration of EEG, MEG and fMRI data for the solution of the neuroimage puzzle. Magn Reson Imaging 22:1471–1476

Babiloni F, Cincotti F (2005) Multimodal Imaging from neuroelectromagnetic and functional magnetic resonance recordings. In: He B (ed) Modeling and imaging of bioelectrical activity,

---

[3] Please note that in the two inequalities, one of them is a strict inequality.

bioelectric engineering. Springer, New York. doi:10.1007/978-0-387-49963-5-8

Bai X, He B (2005) On the estimation of the number of dipole sources in EEG source localization. Clin Neurophysiol 116:2037–2043. doi:10.1016/j.clinph.2005.06.001

Baillet S, Mosher JC, Leahy RM (2001) Electromagnetic brain mapping. Sig Process Mag IEEE 18(6):14–30. doi:10.1109/79.962275

Bandyopadhyay A, Tomassoni T, Omar A (2004) A numerical approach for automatic detection of the multipoles responsible for ill conditioning in generalized multipole technique (2004) IEEE MTT-S International Microwave Symposium Digest (IEEE Cat No04CH37535). doi:10.1109/MWSYM.2004.1338826

Beck R, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imaging Sci 2:183–202

Bolstad AK, Veen BDV, Nowak RD (2009) Space-time event sparse penalization for magneto-electroencephalography. NeuroImage 46:1066–1081

Bruce AG, Sardy S, Tseng P (1998) Block coordinate relaxation methods for nonparamatric signal denoising. doi:10.1117/12.304915, URL 10.1117/12.304915

Candès EJ, Tao T (2005) The dantzig selector: statistical estimation when p is much larger than n. Ann Stat 35:2313–2351

Cassidy B, Solo V, Seneviratne A (2012) Grouped l0 least squares penalised magnetoencephalography. In: 2012 9th IEEE international symposium on biomedical imaging (ISBI), pp 868–871. doi:10.1109/ISBI.2012.6235686

Chen SS, Donoho DL, Michael SA (1999) Atomic decomposition by basis pursuit. SIAM J Sci Comput 20:33–61

Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, Halgren E (2000) Dynamic statistical parametric mapping: combining fmri and meg for high-resolution imaging of cortical activity. Neuron 26(1):55–67

Daubechies I, DeVore R, Fornasier M, Gunturk S (2008) Iteratively re-weighted least squares minimization: proof of faster than linear rate for sparse recovery. In: 42nd Annual conference on information sciences and systems (CISS)

Deb K (1999) Multi-Objective Evolutionary Algorithms: Introducing Bias Among Pareto-Optimal Solutions. Tech. rep, Kanpur Genetic Algorithms Lab (KanGal), Technical report 99002

Dezhong Yao BH (2001) A self-coherence enhancement algorithm and its application to enhancing three-dimensional source estimation from EEGs. Ann Biomed Eng 29:1019–1027

Ding L (2009) Reconstructing cortical current density by exploring sparseness in the transform domain. Phys Med Biol 54:2683–2697

Ding L, Ni Y, Sweeney J, He B (2011) Sparse cortical current density imaging in motor potentials induced by finger movement. J Neural Eng 8(3):36008. URL:http://stacks.iop.org/1741-2552/8/i=3/a=036008

Dong L, Wang P, Bin Y, Deng J, Li Y, Chen L, Luo C, Yao D (2015) Local multimodal serial analysis for fusing EEG-fMRI: a new method to study familial cortical myoclonic tremor and epilepsy. IEEE Trans Auton Mental Dev 7(4):311–319. doi:10.1109/TAMD.2015.2411740

Donoho DL, Elad M (2003) Maximal sparsity representation via ℓ_1 minimization. Proc Natl Acad Sci 100:1297–2202

Donoho DL (2004) For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution. Comm Pure Appl Math 59:797–829

Donoho DL (2006) Compressed sensing. IEEE Trans Inf Theory 52:1289–1306

Efron B, Hastie T, Johnstone I, Tibshirani R (2004) Least angle regression. Ann Stat 32:407–451

Friedman J, Hastie T, Tibshirani R (2010) Regularization paths for generalized linear models via coordinate descent. J Stat Softw 33:1–22

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998) Event-related fMRI: characterizing differential responses. NeuroImage 7:30–40

Friston KJ, Harrison L, Daunizeau J, Kiebel SJ, Phillips C, Trujillo-Bareto N, Henson RNA, Flandin G, Mattout J (2008) Multiple sparse priors for the m/eeg inverse problem. NeuroImage 39:1104–1120

Fuchs A, Kelso JAS, Murzin V (2011) Anatomically constrained minimum variance beamforming applied to EEG. Exp Brain Res 214(4):515–528. doi:10.1007/s00221-011-2850-5

Glover GH (1999) Deconvolution of impulse response in event-related BOLD fMRI. NeuroImage 9:416–429

Gorodnitsky I, George J, Rao B (1995) Neuromagnetic source imaging with focuss: a recursive weighted minimum norm algorithm. Electroencephalogr Clin Neurophysiol 58:267–288

Gramfort A, Kowalski M, Hämäläinen M (2012) Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods. Phys Med Biol 57:1937–1961

Gramfort A, Kowalski M, Hämäläinen M (2013) Time-frequency mixed-norm estimates: sparse M/EEG imaging with non-stationary source activation. NeuroImage 70:410–422

Henson R (2010) Multimodal integration: constraining meg localization with EEG and fMRI. In: Supek S, Suac A (eds) IFMBE Proceedings of the 17th international conference on biomagnetism advances in biomagnetism biomag, 2010, vol 28. Springer, Berlin, pp 97–100. doi:10.1007/978-3-642-12197-5_18

Henson RN, Mouchlianitis E, Friston KJ (2009) MEG and EEG data fusion: simultaneous localisation of face-evoked responses. Neuroimage 47:581–589

Henson RN, Wakeman DG, Litvak V, Friston KJ (2011) A parametric empirical bayesian framework for the EEG/MEG inverse problem: generative models for multisubject and multimodal integration. Front Hum Neurosci 5(76):1–16

Hong D, Zhang F (2010) Weighted elastic net model for mass spectrometry imaging processing. Math Model Nat Phenom 5:115–133

Hotelling H (1936) Relations between two sets of variates. Biometrika 28(3–4):321–377. doi:10.1093/biomet/28.3-4.321. URL:http://biomet.oxfordjournals.org/content/28/3-4/321.short; http://biomet.oxfordjournals.org/content/28/3-4/321.full.pdf+html

Inan G, Kiymik M, Akin M, Alkan A (2001) AR spectral analysis of EEG signals by using maximum likelihood estimation. Comput Biol Med 31(6):441–450

Irimia A, Horn JDV, Halgren E (2012) Source cancellation profiles of electroencephalography and magnetoencephalography. NeuroImage 59(3):2464–2474. doi:10.1016/j.neuroimage.2011.08.104. http://www.sciencedirect.com/science/article/pii/S1053811911010378

Jorge J, van der Zwaag W, Figueiredo P (2014) EEG-fMRI integration for the study of human brain function. NeuroImage 102(1):24–34. doi:10.1016/j.neuroimage.2013.05.114

Jun-Tao LI, Ying-Min JIA (2010) An improved elastic net for cancer classication and gene selection. Acta Automat Sin 36:976–981

Koles ZJ, Lazar MS, Zhou SZ (1990) Spatial patterns underlying population differences in the background EEG. Brain Topogr 2(4):275–284. doi:10.1007/BF01129656

Krüger G, Glover G (2001) Physiological noise in oxygenation-sensitive magnetic resonance imaging. Magn Reson Med 46:631–637

Limpiti T, Van Veen BD, Wakai RT (2006) Cortical patch basis model for spatially extended neural activity. IEEE Transn Biomed Eng 53:1740–1754. doi:10.1109/TBME.2006.873743

Liu AK, Belliveau JW, Dale AM (1998) Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. Proc Natl Acad Sci 95:8945–8950

Mancera L, Portilla J (2006) L0-norm-based sparse representation through alternate projections. In: ICIP

Michel CM, Murray MM, Lantz G, Gonzalez S, Spinelli L (2004) EEG source imaging. Clin Neurophysiol 115:2195–2222. doi:10.1016/j.clinph.2004.06.001

Mohimani H, Babaie-Zadeh M, Jutten C (2009) A fast approach for overcomplete sparse decomposition based on smoothed L0 norm. IEEE Trans Signal Process 57(1):289–301. doi:10.1109/TSP.2008.2007606

Mosher JC, Leahy RM, Lewis PS (1999) EEG and MEG: forward solutions for inverse methods. IEEE Trans Biomed Eng 46:245–259. doi:10.1109/10.748978

Mosher J, Leahy R (1996) EEG and MEG source localization using recursively applied (RAP) MUSIC. In:Conference record of the thirtieth asilomar conference on signals, systems and computers. doi:10.1109/ACSSC.1996.599135

Murta T, Leite M, Carmichael DW, Figueiredo L, Lemieux L (2015) Electrophysiological correlates of the bold signal for eeg-informed fmri. Hum Brain Mapp 36:391–414

Op de Beeck HP, Haushofer J, Kanwisher N (2008) Interpreting fMRI data: maps, modules, and dimensions. Nat Rev Neurosci 9:123–135

Pitcher D, Walsh V, Duchaine B (2011) The role of the occipital face area in the cortical face perception network. Exp Brain Res 209(4):481–493. doi:10.1007/s00221-011-2579-1

R Core Team (2012) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org/

Rosa MJ, Daunizeau J, Friston KJ (2010) EEG- fMRI integration: a critical review of biophysical modeling and data analysis approaches. Curr Opin Neurol 23:374–381

Samadi S, Amini L, Cosandier-Rimélé D, Soltanian-Zadeh H, Jutten C (2013) Reference-based source separation method for identification of brain regions involved in a reference state from intracerebral EEG. IEEE Trans Biomed Eng 60:1983–1992

Sameni R, Jutten C, Shamsollahi MB (2008) Multichannel electrocardiogram decomposition using periodic component analysis. IEEE Trans Biomed Eng 55(8):1935–1940. doi:10.1109/TBME.2008.919714

Shun Chi W, Wang PT, Swindlehurst AL, Nenadic Z (2012) Efficient dipole parameter estimation in EEG systems with near-ML performance. IEEE Trans Biomed Eng 59(5):1339–1348. doi:10.1109/TBME.2012.2187336

Simoncelli EP, Olshausen BA (2001) Natural images statistics and neural repeeresentation. Annu Rev Neurosci 24:1193–1216

Strother SC (2006) Evaluating fMRI preprocessing pipelines. IEEE Eng Med Biol Mag 25:27–41

Tibshirani R (1996) Regression shrinkage and selection via the lasso. J R Stat 58:267–288

Tibshirani R, Saunders M, Rosset S, Zhu J, Knight K (2005) Sparsity and smoothness via the fused lasso. J R Stat Soc 67:91–108

Tikhonov AN, Arsenin VIA (1977) Solutions of ill-posed problems. Scripta series in mathematics. Winston, Great Falls

Tipping ME (2001) Sparse bayesian learning and the relevance vector machine. J Mach Learn Res 1:211–244. doi:10.1162/15324430152748236. URL:http://www.crossref.org/deleted_DOI.html

Van Veen BD (1991) Minimum variance beamforming with soft response constraints. IEEE Trans Signal Process 39(9):1964–1972. doi:10.1109/78.134429

Wagner P, Röschke J, Mann K, Fell J, Hiller W, Frank C, Grözinger M (2000) Human sleep EEG under the influence of pulsed radio frequency electromagnetic fields. Results from polysomnographies using submaximal high power flux densities. Technical report 2000

Welvaert M, Durnez J, Moerkerke B, Verdoolaege G, Rosseel Y (2011) neuRosim: an R package for generating fMRI data. J Stat Softw 44(10):1–18. http://www.jstatsoft.org/v44/i10/

Wipf D, Nagarajan S (2009) A unified Bayesian framework for MEG/EEG source imaging. NeuroImage 44:947–966

Xu D, Yan S, Tao D, Lin S, Zhang HJ (2007) Marginal fisher analysis and its variants for human gait recognition and content- based image retrieval. IEEE Transactions on Image Processing 16(11):2811–2821. URL http://dblp.uni-trier.de/db/journals/tip/tip16.html#XuYTLZ07

Yuan M, Lin Y (2006) Model selection and estimation in regression with grouped variables. J R Stati Soc 68(1):49–67

Zhang J, Li X, Song Y, Liu J (2012) The fusiform face area is engaged in holistic, not parts-based, representation of faces. PLoS One 7:e40390. doi:10.1371/journal.pone.0040390

Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. J R Stat Soc 67:301–320