

# Ensemble multi-modal brain source localization using theory of evidence

Ashkan Olliaiee<sup>a</sup>, Sepideh Hajipour Sardouie<sup>b,\*</sup>, Mohammad Bagher Shamsollahi<sup>a</sup>

<sup>a</sup> Biomedical Signal and Image Processing Laboratory (BiSIPL), Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran

<sup>b</sup> Human Machine Interface Laboratory (HML), Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran

## ARTICLE INFO

### Keywords:

EEG  
MEG  
Integration  
Source localization  
Inverse problem  
Dempster-Shafer theory  
Theory of evidence  
Epilepsy

## ABSTRACT

The primary aim in pre-surgical evaluations in patients with neurological disorders such as epilepsy is determining the precise location of the cortical region responsible for the malfunctions which is called source localization. Different modalities unravel different views of brain activity. Combining these complementary aspects of the brain yields more accurate source localization.

In this paper, a method is proposed for combining localization methods in different modalities based on the theory of evidence, the result of some localization methods in modalities are integrated using weights in accordance to their relative performance and are combined using Dempster's rule of combination and is used for the case of EEG and MEG combinatory source localization.

The proposed method is evaluated on simulated realistic MEG and EEG data and in different noise and artifact levels and finds the zone of interest in a more accurate way. The AUC criterion is used as a metric for the evaluation. The proposed method results in better localization accuracy in terms of AUC showing the combination of modalities could lead to superior performance.

Combining two modalities needs an exact knowledge of the phenomena happening in each modality, making the combination difficult. Here rather of combining EEG and MEG information at the initial phase, the results of some source localization techniques on both modalities are combined. In spite of the simplicity of use, the experimental results of combination showed improvement in epileptic source localization accuracy even in cases that one method shows poor performance. Using the proposed method any number of modalities can be combined without complex consideration and a better brain source localization could be obtained.

## 1. Introduction

In any brain study, the final goal is finding out what is happening inside the brain and what parts are involved and what their mutual behaviors are; better saying, estimating the spatiotemporal dynamics of neural activities are of particular interest. Our knowledge about the brain's function has been flourished in the last decade due to the development of imaging techniques. In most of these studies, to get rid of complex reactions inside neurons of the brain, brain activity is imputed to the hypothetical dipoles as representatives for regional brain activity. The primary aim of pre-surgical evaluations in patients with neurological disorders such as epilepsy is determining the exact location of the cortical region responsible for the malfunctions. This process is called "brain source localization", which is not a well-posed problem. In most of the cases, especially those who are pharmacoresistant, the success of surgery depends on the precise determination of this responsible zone and finding the exact part of the brain to be resected. Epilepsy is a

chronic disorder, the hallmark of which is recurrent and unprovoked seizures. A person is diagnosed with epilepsy if he has unprovoked seizures that were not caused by some known and reversible medical conditions like alcohol withdrawal or extremely low blood sugar level. The seizures in epilepsy may be related to a brain injury or a family tendency, but often the cause is entirely unknown [1–3].

Nowadays, clinics are equipped with many medical devices and imaging systems. Using these systems, different modalities show different aspects of the brain's structure and activity; for example, EEG and MEG, respectively, record electrical and magnetic signals on the scalp, on the other hand, fMRI measures the hemodynamic responses related to neural activity, and Diffusion Tensor Imaging (DTI) can unravel the structural connectivity of brain network. These methods of imaging have their pros and cons; to be mentioned fMRI benefits from excellent spatial resolution while having inferior time resolution and in contrast, EEG and MEG both have good temporal resolutions while their spatial resolutions are lower compared to fMRI [4]. Even in EEG and

\* Corresponding author.

E-mail addresses: [olliaee.ashkan@ee.sharif.edu](mailto:olliaee.ashkan@ee.sharif.edu) (A. Olliaiee), [hajipour@sharif.edu](mailto:hajipour@sharif.edu) (S.H. Sardouie), [mbshams@sharif.edu](mailto:mbshams@sharif.edu) (M.B. Shamsollahi).

<https://doi.org/10.1016/j.bspc.2021.102668>

Received 15 January 2021; Received in revised form 10 April 2021; Accepted 24 April 2021

Available online 19 June 2021

1746-8094/© 2021 Published by Elsevier Ltd.

MEG, the orientation selectivity combined with different effects of skull and scalp make different views of brain activity; in a way that sources tangential to the cortex are better observed on MEG localization while the orthogonal ones are best localized through the use of EEG.

The complementary features of different modalities drive the idea on the combination of modalities, as an ongoing trend which has been posed these days regarding overcoming the shortcomings of modalities. The combination of EEG and MEG is one of the most common combinations of modalities which is popular since there is the feasibility of concurrent recording of these two modalities. The popularity of this multimodal recording in some disorder studies such as epilepsy is due to the need for high temporal resolution and also long term monitoring of patients available on EEG and MEG [5,6]. Previous studies have reported many Electromagnetic Source Imaging (ESI) methods to integrate EEG and MEG in terms of source localization [7–17]. Methods involve sequential steps to find radial and tangential source components separately [7,18], normalizing them according to the noise level, to create unit-free measures such as row normalization of lead field matrices, minimization of mutual information for channel selectivity, using the Bayesian framework and beamforming [13,14,19,20,12,21–23], but there is still a controversy on the best strategy [24].

In this paper, a new strategy is proposed for extracting the synergistic results of EEG and MEG source localization. As the Bayesian framework became so popular regarding combining different modalities, and since the theory of evidence is known as a more comprehensive case of Bayesian theory [25–28], the proposed method of combining modalities was developed based on Dempster-Shafer theory of evidence which has been well known in control and management society before [29,30]. Even though the proposed method is apt to be used for the combination of different modalities, here it was used for the case of combining EEG and MEG. In this paper, a new source localization method was not proposed, rather a method was developed to combine the results of other popular localization methods such as MNE [31], WMNE [32], LORETA [33], LAURA [34] and so on. The results of testing our method on realistic simulated epileptic EEG and MEG signals showed an improvement in determining the foci of epileptic zone. Better saying, the proposed method makes it possible to ensemble popular source localization methods in different modalities.

The organization of the remaining is as follows. Section 2 provides a comprehensive overview of the localization methods in its first part followed by a review of modalities integration techniques in the second part. In 2.3 the explanation of Dempster's Shafer theory and the preliminaries needed are introduced. In Section 3 the detailed explanation of the method is provided followed by evaluation criteria in Section 4. Section 5 explains the simulation procedures and presents the result of applying the method to the intended data. In the last section, the discussion and conclusion will be expressed.

## 2. Background

As mentioned earlier, the brain's activities are imputed to the activity of dipoles inside the brain. In any localization process, two types of problems exist. The first one, known as “*Forward Problem*”, is defined as finding the recorded signal, e.g., potentials on the scalp electrodes in case of EEG, by having spatial and temporal activities of dipole sources. This is done by considering the equations of propagation such as those of Maxwell's. This is a straightforward procedure requiring simpler calculations. The other significant problem, namely the “*Inverse Problem*”, is exactly the opposite. In this case, the unknowns are temporal and spatial activities of dipoles given the recorded signal and propagation relations. In inverse problems, sources are usually modeled in two ways: the *Equivalent Current Dipole (ECD)* and the *Distributed Model (DM)*. In the ECD methods, also called parametric methods, the activity of a large part of the brain is considered as one dipole moment, and the whole recorded activity is attributed to few dipoles (usually less than 6) [35]. On the other hand, in the DM or non-parametric methods, many dipoles are

considered in predefined places, and the problem is changed to find the orientations and moments of dipoles. In many cases, because dipoles are perpendicular to the surface of the cortex, the problem is reduced to find the magnitude of sources in time. The ECD methods usually convert the inverse problem to a non-linear optimization one which should be solved using the iterative algorithms. Accordingly, they are susceptible to the correct estimation of the number of dipoles. In the DM ones, the subject's brain anatomy, gained by high-resolution MRI, is used, and the propagation medium is determined. Considering these, the equations and relations are defined. Since the number of dipoles is more than the number of equations in hand, a highly underdetermined system of equations has to be solved. Therefore, this system has lots of possible solutions; the reasonable one must be chosen by adding adequate constraints. To this end, lots of priors based upon mathematical, physiological, and functional assumptions have been considered [36,37]. Each of these priors leads to a method with a unique answer, some of which are mentioned in Section 2.1.

### 2.1. Source localization methods

Considering the non-parametric model, by denoting the number of electrodes, the number of dipoles and the number of time samples with  $n$ ,  $p$  and  $t$ , respectively, the EEG/MEG signals and the activity of dipoles can be related through:

$$\mathbf{M}^{\text{E,M}} = \mathbf{G}^{\text{E,M}}\mathbf{D} + \boldsymbol{\eta}^{\text{E,M}} \quad (1)$$

where  $\mathbf{M}$  is an  $n$  by  $t$  data matrix, having the recorded EEG/MEG signals,  $\mathbf{G}$  is an  $n$  by  $p$  matrix, called *leadfield*, describing the propagation equations in the head,  $\mathbf{D}$  is a  $p$  by  $t$  matrix representing the activity of dipoles and  $\boldsymbol{\eta}$  is the noise matrix. The superscripts E and M correspond to EEG and MEG, respectively. In the inverse problem, the goal is estimating  $\mathbf{D}$  or activity of dipole sources having  $\mathbf{M}$  (the recorded signal) given  $\mathbf{G}$ . Among the popular methods of solving the mentioned problem, Bayesian-based methods are more prevalent. Generally, in Bayesian methods, an estimate of source activity  $\hat{\mathbf{D}}$  is of interest which maximizes the probability of  $\mathbf{D}$  given  $\mathbf{M}$ :

$$\hat{\mathbf{D}} = \operatorname{argmax}_{\mathbf{D}} [p(\mathbf{D}|\mathbf{M})] \quad (2)$$

Considering the Bayes relation, it could be written as:

$$p(\mathbf{D}|\mathbf{M}) = \frac{\exp[-F_{\alpha}(\mathbf{D})]/z}{p(\mathbf{M})} \quad (3)$$

where:

$$F_{\alpha}(\mathbf{D}) = U(\mathbf{D}) + \alpha L(\mathbf{D}) \quad (4)$$

in which  $U(\mathbf{D})$  and  $L(\mathbf{D})$  are energy functions related to  $p(\mathbf{M}|\mathbf{D})$  and  $p(\mathbf{D})$  respectively. In fact,  $U(\mathbf{D})$  is the consistency factor and  $L(\mathbf{D})$  corresponds to temporal and spatial constraints of dipoles. The  $z$  and  $\alpha$  factors are normalization and regularization factors, respectively. Different  $U(\mathbf{D})$  and  $L(\mathbf{D})$  lead to different methods of localization, which MNE could be mentioned amongst the simplest ones, in which the cost function is as:

$$\hat{\mathbf{D}} = \min_{\mathbf{D}} [\|\mathbf{GD} - \mathbf{M}\|_F^2 + \alpha\|\mathbf{D}\|_F^2] \quad (5)$$

in which  $\|\cdot\|_F$  shows the Frobenius norm and the closed form answer is achieved by:

$$\hat{\mathbf{D}}_{\text{MNE}} = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T + \alpha\mathbf{I}_p)^{-1}\mathbf{M} \quad (6)$$

where  $\mathbf{I}_p$  is the identity matrix of size  $p$  by  $p$ . This method is intended to find the superficial and weak sources. To overcome this issue, the WMNE method was proposed which by considering a weight matrix  $\mathbf{W}$ , deep sources are more involved. The estimate of  $\mathbf{D}$  in the WMNE method is as follows:

$$\hat{D} = \min_D [\|GD - M\|_F^2 + \alpha \|WD\|_F^2] \quad (7)$$

where the closed-form answer can be achieved by:

$$\hat{D}_{\text{WMNE}} = G^T (GG^T + \alpha W^T W)^{-1} M \quad (8)$$

One of the simplest weights to be applied is based on column normalization of matrix  $G$ . Even though this method leads to deeper and more accurate sources compared to MNE, still the accuracy is not acceptable in many cases. Since sources close together have somehow similar moments, or in other words have similar regional activity, the Laplacian operator is used instead of  $W$  in the LOERTA method. Therefore, the LOERTA estimate of  $D$  will be as follows:

$$\hat{D} = \min_D \left[ \|GD - M\|_F^2 + \alpha \|\Delta \bar{G} \cdot D\|_F^2 \right] \quad (9)$$

and the closed-form solution is obtained as:

$$\hat{D}_{\text{LOERTA}} = \left( GG^T + \alpha \bar{G} \Delta^T \Delta \bar{G} \right)^{-1} G^T M \quad (10)$$

where matrix  $\bar{G}$  is obtained from column-wise normalization of the  $G$  matrix, and  $\Delta$  is the Laplacian operator. As the name of sLOERTA may wrongly considered to be similar to the LOERTA, it is completely different and uses both the signal noise variance and biological variance impressive in making  $W$  [38,39] and Laplacian operator is not used. To mitigate some limitations of sLOERTA under realistic noisy conditions, swLOERTA, a modification of sLOERTA obtained by incorporating a singular value decomposition-based lead field weighting was suggested [40]. Another modification, called LAURA (Local AUtoRegressive Average) [41], incorporates the biophysical law that the strength of the source falls off with the inverse of the cubic distance for vector fields. LAURA integrates this law in terms of a local autoregressive average with coefficients depending on the distances between solution point. Among this type of methods, some others like eLOERTA [42], EPIFOCUS [41], VARETA [43], S-MAP [44], ST-MAP [45], BESA [46] can be named. In addition, among the ECD methods, MUSIC [47,11] and BeamForming [47,48] worth mentioning.

## 2.2. Combination of different modalities

As mentioned before, any neuro-imaging modality unravels a view of brain activity, hereupon, multimodal setups can take the advantage of complementary views of brain activity [49]. Since measurements are not of the same type in different modalities, combining them needs special considerations and techniques some of which will be mentioned in the following. Joint analysis and data integration of different modalities can be categorized from different points of view. The most accepted categories are:

- *Symmetric and Asymmetric methods,*
- *Supervised and Unsupervised methods,*
- *Model-based and Data-driven methods.*

These divisions are general and not independent of each other. One method could be model-based asymmetric and the other could be symmetric and model-based. These divisions are explained in the following:

- *Symmetric and asymmetric methods:* In asymmetric analysis, measurements of one modality are used as a base for data analysis of the other modality. Its better to say that, the data of one modality is suggested as primary conditions for analysis of the data in the other modality. It is even possible that, the result of analysis in the second modality is in contrast to the suggestion in the first modality. In other words, the main focus is the analysis of one modality but with the

help of another one. Contrarily, in the symmetric methods, both modalities are of the same importance and are analyzed concurrently [49–51]. Most of the asymmetric methods are regression in nature. In these cases, the stimuli are used for feature extraction in one modality and they are used as the regressor for the other modality. There are also some methods that can be used in both ways. But some methods could be named only in the symmetric group, from which the information theory combination can be named [51–53].

- *Supervised and unsupervised methods:* This kind of division is adapted from literature of classification methods. *Statistical Parametric Maps (SPMs)* and *Multivariate Pattern Analysis (MPAs)* are the most popular supervised integration methods. Both of them are based on regression and mostly the linear one, but the role of the target variable is different. In SPM methods, the time series of a voxel or a dipole is considered a target variable and is estimated through a regression with practical parameters. In contrast to SPM, in MPA the main goal is finding practical parameters from the observed data. Better saying, in MPAs finding patterns in data to best describe the stimuli is of the main concern [54–56]. Unsupervised methods, are well known in integration of modalities. Methods like ICA and PCA are of the most important methods of data analysis. Unsupervised methods are usually unimodal methods in first steps. Features are extracted individually and separately in each modality by means of methods like ICA and PCA and these features are integrated in the next step. Unlike unimodal unsupervised methods in recent years, special methods were adapted to jointly analyze the data in modalities [57]. They automatically reflect the best common underlying processes and there is no need for the concatenating step mentioned before. Amongst the leading ones Joint ICA [58], Linked ICA [59], Parallel ICA [58,60], mCCA [61], NPLS [58], CC-ICA [58] could be named.
- *Model-based and data-driven methods:* In most of the above mentioned methods, especially supervised ones, besides the relationship between parameters, such as orthogonality in PCA or linear combination of sources in ICA, no further assumption exists in signal generation. Better saying, they are purely data-driven methods and do not offer a real model for the underlying neural processes. Since there is a neural model causing the signal generation, it is better to assume a model for data generation; in other words, not every signal can be accepted as the neural response on the modalities. Hypothesis or prior knowledge, can be incorporated in models of physiological processes underlying the measurements. These physiological models are called forward models. Most of these model-based integration use a probabilistic frame called *state space model* or often *dynamic casual models* [62]. Bayesian integration methods are among the most popular ones. Bayesian source localization methods are general types of methods maximizing the posterior probability of dipole parameter by assuming a model for parameters, having measurements, and based on the Bayesian rule. This method is different as can change the underlying assumptions and facilitate the prior knowledge usage [63]. Mainly the Bayes rule is defined as:

$$p(\theta | \text{measurements}) \propto p(\text{measurements} | \theta) p(\theta) \quad (11)$$

where  $\theta$  shows the desired unknowns,  $p(\theta | \text{measurements})$  is the conditional probability of unknowns given the measurement information,  $p(\text{measurements} | \theta)$  is the conditional probability distribution of measurements given  $\theta$  and represents the physical relation between  $\theta$  and the measurements and finally  $p(\theta)$  is the prior distribution of unknowns. The main idea here is maximizing  $p(\theta | \text{measurements})$ . The parameters involved are covariances and noise models which have to be selected in a precise way.

Integration based on Kalman filtering is another model-based method benefiting from considering time series information. Considering the fact that the brain source state depends on the state of earlier times, and the measurements in the modalities like EEG/MEG seem to have a linear relation to these states, usage of Kalman

filtering seems to be rational [64].

### 2.3. Dempster-Shafer theory

In the mid-1960s, Arthur P. Dempster proposed a theory about upper and lower bounds of probability problems [25,26,65]; It soon became apparent that this theory could be used in describing uncertainty in systems. Ten years later, Glen Shafer, reviewed and modified the statements of Dempster [26] and published them under the name of “*Theory of Evidence*” also known as “*Dempster-Shafer Theory of Evidence (DST)*”. Afterward, the activity of others showed a greater difference between DST and probability theory. In this paper, we proposed a novel method to make the multimodal integration possible based on DST.

Basic concepts of DST comprises of three main parts: namely, i) *hypothesis*, ii) *pieces of evidence*, and iii) *data sources* [26]. Suppose  $\Theta$  is a finite non-empty set having all possible states of an event of interest. In this manner, the  $\Theta$  is called “*frame of discernment*”. To avoid misconception,  $\Theta$  is interpreted as a set of mutually exclusive and exhaustive propositions [66]. In this way, any subset of  $\Theta$  could be a hypothesis, a target, or a situation of a system. The power set of  $\Theta$  shown by  $2^\Theta$ , is a set comprising all subsets of  $\Theta$ :

$$2^\Theta = \{A|A \subseteq \Theta\} \quad (12)$$

Any subset of  $2^\Theta$  could be a hypothesis or overlap of hypotheses. In this theory, any observation or event accessible is called “*observation*”. Each of these observations will signify a hypothesis and the important property is that no two observations can imply one hypothesis. The relation of these observations and hypotheses is asserted by “*Data Sources*” and through a projection called “*mass function*”  $m$ :

$$\begin{cases} m : 2^\Theta \rightarrow [0, 1] \\ m(\emptyset) = 0 \\ \sum_{A \subseteq \Theta} m(A) = 1 \end{cases} \quad (13)$$

The mass function is also called “*Basic Probability Assignment (BPA)*” in this context [29].  $m(A)$  shows the share of set  $A$  from all other observations and supports the statement about the element of  $\Theta$  which only can be a member of  $A$  [29]. Despite similarities between DST and probability theory, major differences exist between these two. Some conditions like what follows in the probability may not hold in DST:

1.  $m(\Theta) = 1$
2. if  $A \subseteq B \Rightarrow m(A) \leq m(B)$
3.  $m(A)$  and  $m(\bar{A})$  has a relation.

Starting from  $m(A)$ , lots of other functions are defined in this context: functions such as “*belief*”, “*plausibility*”, “*commonality*” and so on which are out of the scope of this paper. The “*rule of combination*” is one of the essential rules in DST making a combination of two independent data sources with two different mass functions possible:

$$m(Z) = \frac{\sum_{A \cap B = Z \neq \emptyset} m(A) \cdot m(B)}{1 - \sum_{A \cap B = \emptyset} m(A) \cdot m(B)} \quad (14)$$

where the numerator shows the degree of agreement of two data sources on the  $Z$  hypothesis [29]. The second part of the denominator shows the disagreement and also completely ignoring any BPA associated with the null set [67]. This term is determined by summing the products of BPA of all sets where the intersection is null. The whole dominator makes the sum of all new mass functions equal to one in order to be consistent with (13). This rule is commutative, associative but not idempotent [68]. The rule above could be extended to many data sources [67,68]. Based on DST and specifically by means of combining rule, we proposed a method for multi-modal source localization.

### 3. Methodology

Regarding explaining the idea, suppose that there is a source localization problem and this has to be done using the integration of two modalities, without losing generality suppose these two to be EEG and MEG. In this case, the frame of discernment is any possible activity of dipoles in the brain. Data sources are these two modalities, presenting different mass functions on any hypothesis of dipoles activity. In other words, in each modality, any source localization method proposes an activation map of dipoles held as an exclusive hypothesis. Different methods have got different mass functions in each modality.

Suppose having  $modal^1$  and  $modal^2$  with  $P$  and  $Q$  methods of localization intended to be combined respectively; Weights of importance for these methods, which are their relative mass functions will be  $w_j^i$  ( $w$  is used instead of  $m$  for ease of use) where:

$$\sum_{j=1, \dots, P} w_j^1 = 1. \quad \sum_{j=1, \dots, Q} w_j^2 = 1. \quad (15)$$

The set of all possible dipoles is defined by  $S = \{s_1, s_2, \dots, s_D\}$  and the dipoles chosen by the  $j$ th localization method using each modality ( $modal^i$ ,  $i = 1, 2$ ) is shown by  $S_j^i \subseteq S$ .

By definitions, the rule of combination (14) can be applied on the dipoles extracted using  $modal^1$  and  $modal^1$  as follows:

$$\begin{aligned} \forall S' \subseteq S, \quad S' \neq \emptyset \\ \sum_{S_j^1 \cap S_k^2 = S'} (w_j^1 \cdot w_k^2) \\ m(S') = \frac{\sum_{j=1:P, k=1:Q} (w_j^1 \cdot w_k^2)}{1 - \sum_{S_j^1 \cap S_k^2 = \emptyset} (w_j^1 \cdot w_k^2)} \end{aligned} \quad (16)$$

where  $m(S')$  is the mass function of the dipole set  $S' \subseteq S$ . After calculating the mass values for all possible dipole sets, the dipole set with the highest mass function is chosen as the active dipole set.

To clarify more, the proposed method is explained through an example. Suppose that there are three methods for localization. Considering having EEG data in hand, the first method which is the weakest one and so has the least degree of trust, proposes sources  $\{s_1, s_2\}$ , the strongest one proposes sources  $\{s_2, s_3, s_4\}$ , and the mid-performance one proposes that source  $\{s_5\}$  is active. The degree of trust in the results of strongest method is twice the trust in the mid-performance method and six times as much as the weakest one. It worth mentioning that these values are just assumed as an example. As the sum of mass functions must be equal to one, here we have a mass function of EEG data source with these values:

$$m^{EEG} = \begin{cases} 0.1 & \{s_1, s_2\} \\ 0.6 & \{s_2, s_3, s_4\} \\ 0.3 & \{s_5\} \end{cases} \quad (17)$$

And again using MEG and by the same or different localization methods by supposing different degrees of trust in methods, we will have another mass function called  $m^{MEG}$  defined:

$$m^{MEG} = \begin{cases} 0.2 & \{s_1, s_2, s_3\} \\ 0.3 & \{s_1, s_3, s_5\} \\ 0.5 & \{s_4, s_5\} \end{cases} \quad (18)$$

where sources  $s_1, s_2, \dots$  in both cases indicate the same dipole sources. Using (16), the combinatory mass function will be obtained. To mention one, the mass function of set  $\{s_1\}$  calculated as  $\frac{0.1 \times 0.3}{1 - (0.1 \times 0.5 + 0.3 \times 0.2)} = \frac{3}{89}$ . Similarly, the combinatory mass function of all possible dipole sets will be obtained as:

$$m^{\text{DS}} = \begin{cases} \frac{3}{89} \{s_1\} \\ \frac{18}{89} \{s_3\} \\ \frac{24}{89} \{s_4\} \\ \frac{21}{89} \{s_5\} \\ \frac{3}{89} \{s_1, s_2\} \\ \frac{18}{89} \{s_2, s_3\} \end{cases} \quad (19)$$

So hypothesis  $\{s_4\}$ , having the largest mass value, will be chosen as the active dipole obtained by the integration of EEG and MEG. It must be noticed that, as supposed in the above example, the values (weights) of importance for different methods in two modalities are not usually the same. One method may have a better performance in one modality while having poor one on the other. The proposed idea is combining these mass functions using Dempster's rule of combination. It should be noted that, the problem is not as simple as the example explained above. First of all, in the DM model, any method suggests many active sources normally far from two or three active sources. Also, dipoles have got different values as activity magnitude, not a bi-state of being active or inactive as supposed in the example. These are difficulties somehow handled in the proposed method which will be further explained. The method could be expressed in a more detailed way as follows.

Using the data of each modality after performing the localization method, the activation map or activation set of each method in each modality will be determined (the activities of dipoles is normalized in zero to one interval).

$$\text{Activation Set}(i, j) : A_j^i = \{(s_1, \alpha_1^i), (s_2, \alpha_2^i), \dots, (s_D, \alpha_D^i)\} \quad (20)$$

where the  $(s_k, \alpha_k^i)$  represents the  $k$ th dipole and its related activity of modality  $i$  and localized using method  $j$  and  $D$  is the total number of dipoles considered in the cortex area.

The zero to one interval will be chunked into  $k$  equal intervals and the chunk limited activity will be determined;

$${}_m A_j^i = \{(s_d, \alpha_d^i) \mid (m-1)\frac{1}{k} \leq \alpha_d^i \leq m\frac{1}{k}\} \quad (21)$$

where the set  ${}_m A_j^i$ , is the set of dipoles having the activity limited to the  $m$ th chunk obtained using method  $j$  and by means of  $modal^i$ .

Using the Dempster's rule of combination, the limited activation sets of each chunk are combined using the related weights.

$${}_m A = \text{DSComb}_{ij} ({}_m A_j^i) \quad \text{using } w_j^i \quad (22)$$

where the  ${}_m A$  shows different sets of dipoles having the value equal to middle of the  $m$ th chunk.

From these candidates, the one with greatest value of related mass will be chosen;

$${}_m \bar{A} = \text{Max}_w ({}_m A) \quad (23)$$

$${}_m \bar{A} = \{(s_\alpha, c_m), (s_\beta, c_m), \dots, (s_\theta, c_m)\} \quad (24)$$

where  $c_m$  is the middle value of  $m$ th chunk.

$$c_m = \left(m - \frac{1}{2}\right) \left(\frac{1}{k}\right) \quad (25)$$

The final combinatory activation set, having all possible values of activity, will be obtained by aggregating all the  ${}_m \bar{A}$ s.

$$A = \bigcup_m ({}_m \bar{A}) \quad (26)$$

where the  $A$  represents the desired activation set or activation map. The whole process is shown in Fig. 1.

#### 4. Evaluation criteria

To evaluate the performance of the proposed method, AUC measure was used. The AUC (Area under Curve) shows the ability of the method to correctly determine the values of the dipole moments. AUC is the area under the ROC (Receiver Operating Curve) curve. The ROC curve is created by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold settings. The TPR is also known as sensitivity and the FPR is also known as fall-out and can be calculated as below:

$$\text{TPR} = \frac{\sum \text{number of active dipoles truly retrieved}}{\text{number of all active dipoles}} \quad (27)$$

$$\text{FPR} = \frac{\sum \text{number of active dipoles falsely retrieved}}{\text{number of all inactive dipoles}} \quad (28)$$

To calculate these values, at each threshold, dipoles having the moment greater than the threshold were supposed to be active and the others to be inactive. In addition to the AUC criterion, to get better insight of what is really detected as active dipoles, the visual representations of the brain real and detected patches were also indexed.

#### 5. Simulation and results

To evaluate the proposed method, epileptic realistic simulations of EEG and MEG were used. The simulated EEG data were generated using a realistic model developed in LTSI, University of Rennes 1 [69,70]. In these simulations, head structure and cortical mesh were gained using a 3D MRI T1 image of the subject consisting of a total of 19626 triangles of the mean surface of 5mm<sup>2</sup> each. A current dipole is placed at the barycenter of each triangle oriented orthogonally to the triangle surface.  $P$  distributed sources, called patches, generating interictal spikes were defined. Each patch is composed of 100 dipole sources to which we have assigned hyper-synchronous spike-like activities generated from a model of neural populations [69]. Having this setup, and considering 32 EEG electrodes and 80 MEG magnetometers in the cortical region, the forward problem is then solved using a realistic head model made of three nested homogeneous volumes shaping the brain, the skull and the scalp (Boundary Element Method). Then, the EEG and MEG signals are calculated using the forward model and the epileptic patches. Finally, real EEG and MEG data, having muscle activity, background signal, and instrument noise are added to the generated epileptic activity with specified Signal to Noise Ratio (SNR). We considered three scenarios accounting for some interesting source configurations common in partial epilepsy. For each of these three scenarios, 75 different dynamics were considered. Here, by dynamics, we mean different sources' activation and synchronization patterns over time. In the first scenario, as shown in Fig. 4(a), one single patch localized in the superior temporal gyrus in the left hemisphere was considered. In the second scenario, the patch was supposed to be in the inferior frontal, and finally in the third one, it was supposed to be in the rostral frontal as shown in Fig. 4(b) and (c), respectively.

From the aforementioned source localization methods, MNE, WMNE, LAURA, EPIFOCUS, sLORETA and swLORETA were used for both modalities. To do the combination, the activity of dipoles was limited between zero and one and this interval was chopped into discrete intervals of length 0.005. The combination was performed in each interval separately, and in the case that two values were obtained for a dipole in two different chunks, the greater one was accepted. An example of a synthetic EEG and MEG signal is shown in Figs. 2 and 3, respectively. As

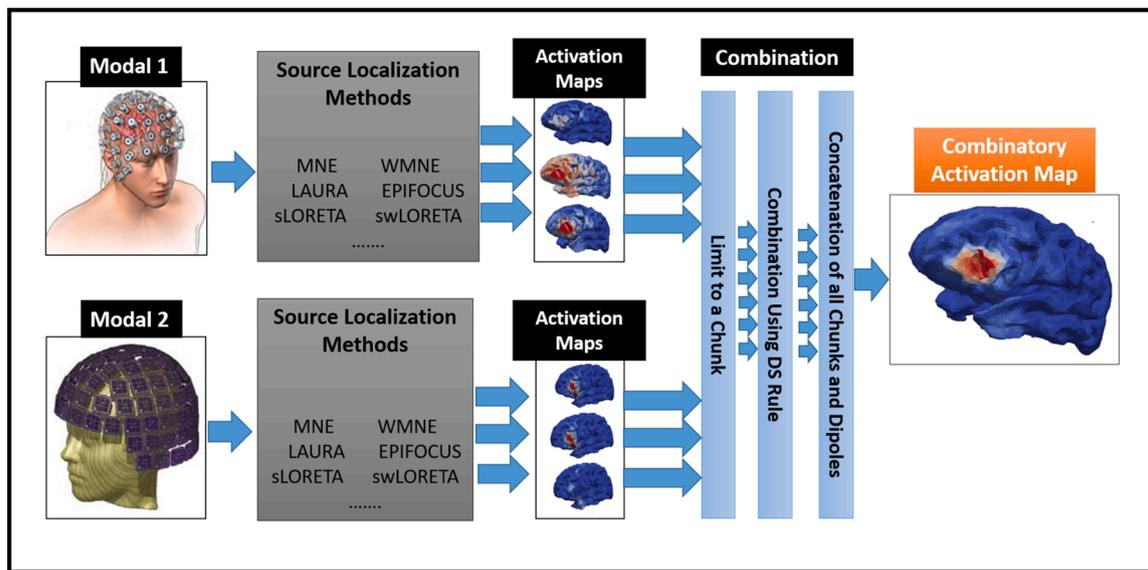


Fig. 1. Process of combination, starting from modalities' data fed into source localization methods yielding activation maps. These normalized activation maps are chunked into zero to one interval and in each interval combination through Dempster's rule of combination is executed and set of dipoles with activity value equal to middle of the chunk is determined. Concatenating these sets with their related values yields the final combinatory activation map.

shown in these figures, an epileptic spike shows up between seconds 5 and 10. As localization of epileptic foci is the main goal, the localization algorithms were applied in this segment.

As shown in Fig. 4, three patches were considered as they are common places of interictal activity (epileptic spikes). In each of these patches, the moments of dipoles are greater than the ones outside the patch and these moments are varying with time and in relation to each other depending on the chosen dynamic. Applying the source localization algorithms to synthetic signals resulted in an estimate of dipole activities, as shown in Figs. 5 and 6 for the case of activation as Fig. 4(C). Combining these results with pre-determined weights of the proposed methods resulted in estimated activities as shown in Fig. 7.

It is worth mentioning that these weights are obtained through the test and train procedure. In each modality and from each scenario, 20 randomly chosen dynamics were used to investigate the localization performance of each method. In accordance with these performances or better saying AUC ratios, the combinatory weights were obtained. The other remained dynamics were used as the test data. The AUCs of the source localization methods on the EEG and MEG train dynamics are also presented in Fig. 8. The first six boxplots are related to the average AUC of source localization methods using EEG and the next six ones shows the average AUC of these methods using MEG of train trials. These mean values of AUC were chosen as the weight of combination for each

method. The localization process was performed at different noise levels starting with clean EEG and MEG signals having no noise and artifact, ending at -30 db contaminated EEG and MEG signals. The aforementioned weights obtained from the comparison of different methods, were derived from the clean setup and these weight were used for combination through the proposed method at all other SNRs. The performance evaluation of the proposed method in comparison to the other localization methods at different noise levels are represented in Fig. 9 and Fig. 10 for EEG and MEG, respectively. The precise values AUCs' mean and standard deviation at is also presented in 1.

### 6. Discussion and conclusion

Considering the simulations, it is inferred that LAURA outperforms other methods in EEG and EPIFOCUS is slightly better than others in MEG in these simulation setups, and with no surprise MNE is the weakest method amongst them. In spite of having a great performance in EEG for LAURA, the method shows poor performance in MEG, so the best strategy is giving the LAURA the highest weight for EEG and a low weight for MEG. The lowest weight was given to MNE in both modalities as it was the case in the derived weights in the proposed method. Source Localization with these methods yields different results on EEG compared to MEG attributing to the lead-field properties of these two

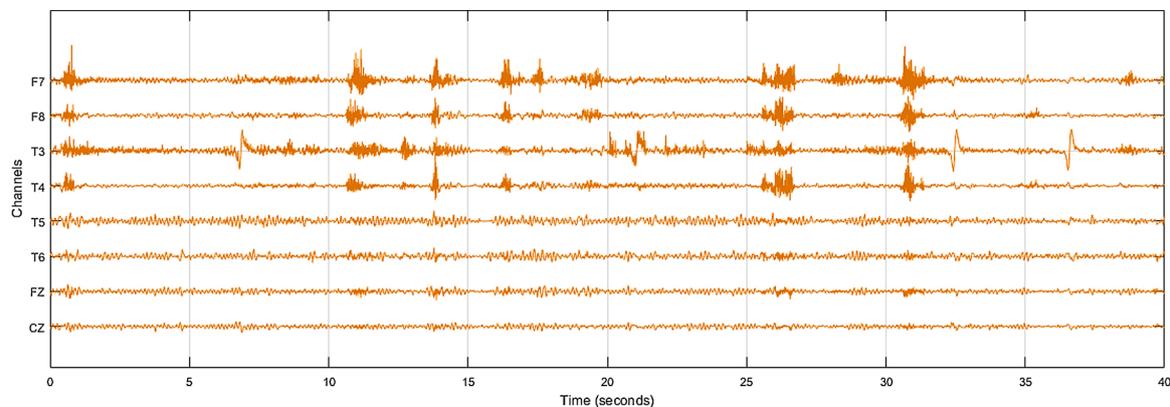


Fig. 2. Sample synthetic EEG corrupted with background noise and muscle artifact, the interictal spikes are best seen on channel T3 on seconds 7, 21, 35 and 37.

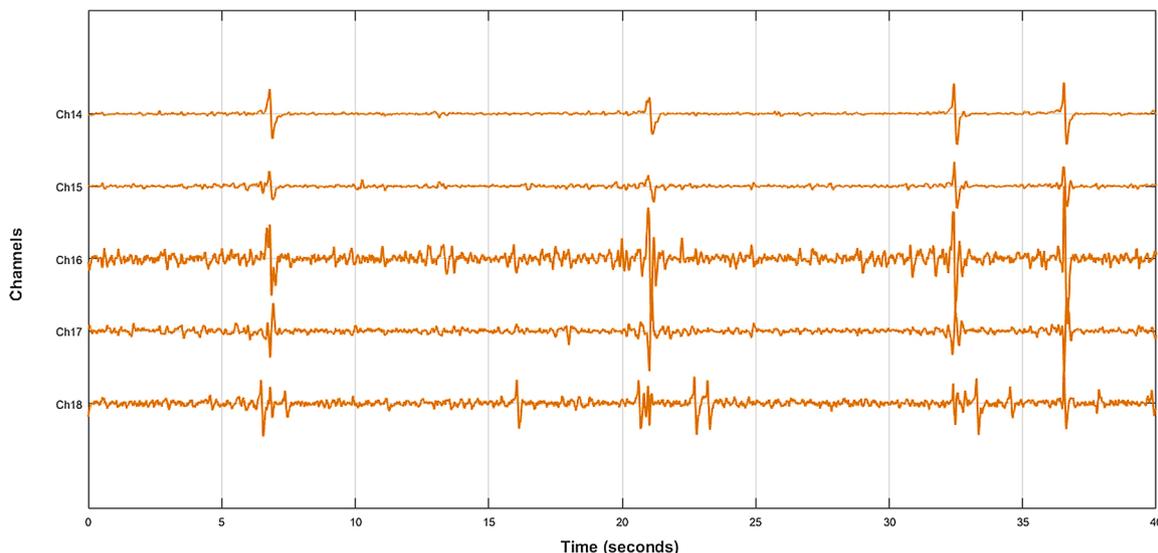


Fig. 3. Sample synthetic MEG corrupted with background noise, the interictal spikes are best seen on channel 14 on seconds 7, 21, 35 and 37.

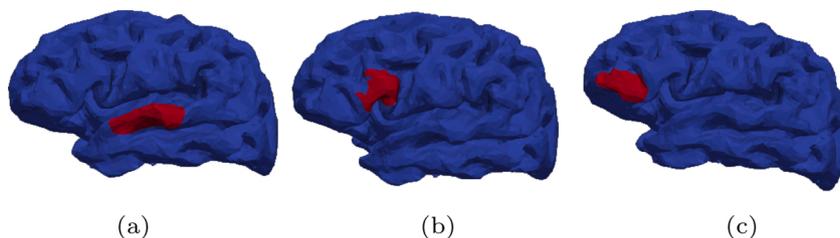


Fig. 4. Patch locations: (a) Temporal superior; (b) frontal inferior; (c) frontal rostral.

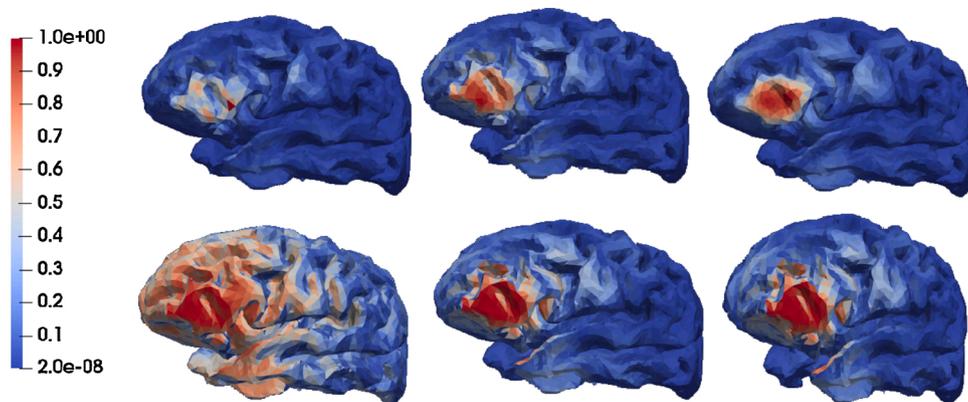


Fig. 5. Applying localization methods to EEG from upper left to lower right: MNE, WMNE, LAURA, EPIFOCUS, sLORETA, swLORETA.

modalities in which MEG’s lead-field as a matrix has more near-zero eigenvalues and more equations related to squid channels. In some cases, as it is in the case of LAURA, superior performance in EEG source localization does not necessarily ensue better performance in the MEG case. As expected, choosing different weights for the methods resulted in different final combination results complying with the nature of theory of evidence. But in this case as far as the weight of best method is not smaller than the others, the total performance will be comparable to the best method. As seen in Figs. 9 and 10 at low noise level source localization using MEG, gives better results but with increasing the noise level, the performance of the best methods on both modalities becomes comparable. By having the SNR less than  $-10$  db, the best method using EEG outperforms the best using MEG. The proposed method shows

better performance in comparison to the best method in any case no matter whether it is on EEG or MEG or what the method is. There is a remarkable point here that in some special cases a good method shows poor performance, as it was the case for LAURA in MEG. By ensemble the solutions of different methods and not relying on one method, more trustful results will be obtained. In these setups the performance of LAURA using EEG was already good enough which made further improvement difficult. Despite the fact that the AUC criterion showed slightly superior performance of the proposed method in comparison to the LAURA, from the visionary point of view, the region of activity has remarkable higher precision in the proposed method. To justify this, suppose the case in which all active values are found with value one and all inactive ones are found with value zero; in this scenario, the AUC is

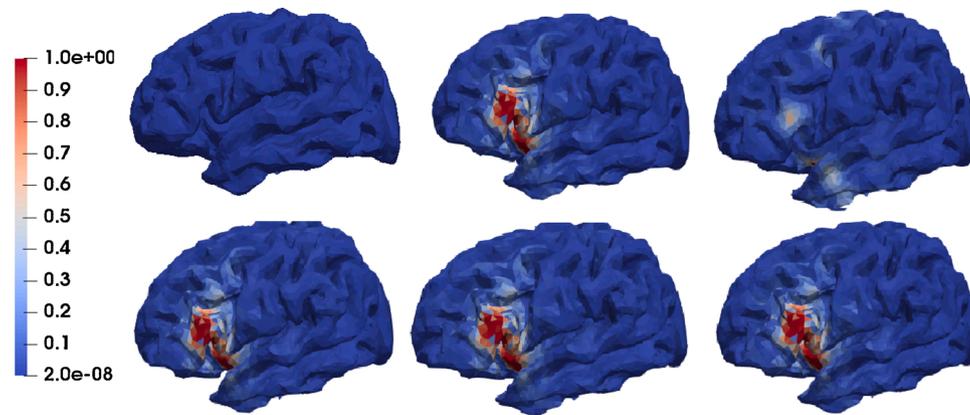


Fig. 6. Applying localization methods to MEG from upper left to lower right: MNE, WMNE, LAURA, EPIFOCUS, sLORETA, swLORETA.

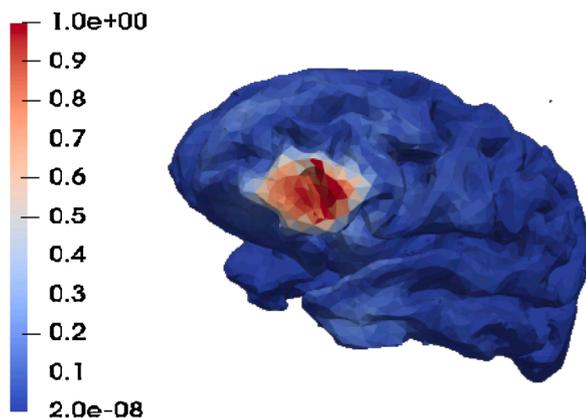


Fig. 7. Localization using the proposed method.

low but the region of activation is exactly determined, showing the necessity of using visionary metric in addition to AUC. Better saying, the values and magnitudes of dipoles' activities are not as accurate as LAURA but their labeling, being active or inactive, is done better in the proposed method.

One of the major benefits of the proposed method is that there is no need to deal with the mathematical formulation of the methods. No

matter how the method does it, the participation of its results in the proposed framework, is possible. There is no limit in the number of methods to be combined. The proposed method enables the further combination of modalities regardless of what their physiological relationships are, as opposed to most of the combination cases; so other multimodal combinations could be quickly performed by this method, such as fMRI and EEG combination. The proposed method does not impose a computational burden and the ponderous part is implementing the methods used. It is somehow expected that in cases that different and noncomplying regions are marked as active in different modalities, as it is the case for fMRI and EEG, more accurate results will be gained that will be the topic of another paper. Further investigations for distributed sources and usage for the case of other disorders will be fascinating. Developing a sophisticated mathematical framework is our main concern and hoped to be fulfilled.

**Financial disclosure**

There has been no significant financial support for this work that could have influenced its outcome.

**Authors' contribution**

Ashkan Oliaiee: Conceptualization, Data curation, Formal analysis, Visualization, writing – original draft. Sepideh Hajipour Sardouie: Supervision, Writing – Review & Editing, Resources. Mohammad Bagher

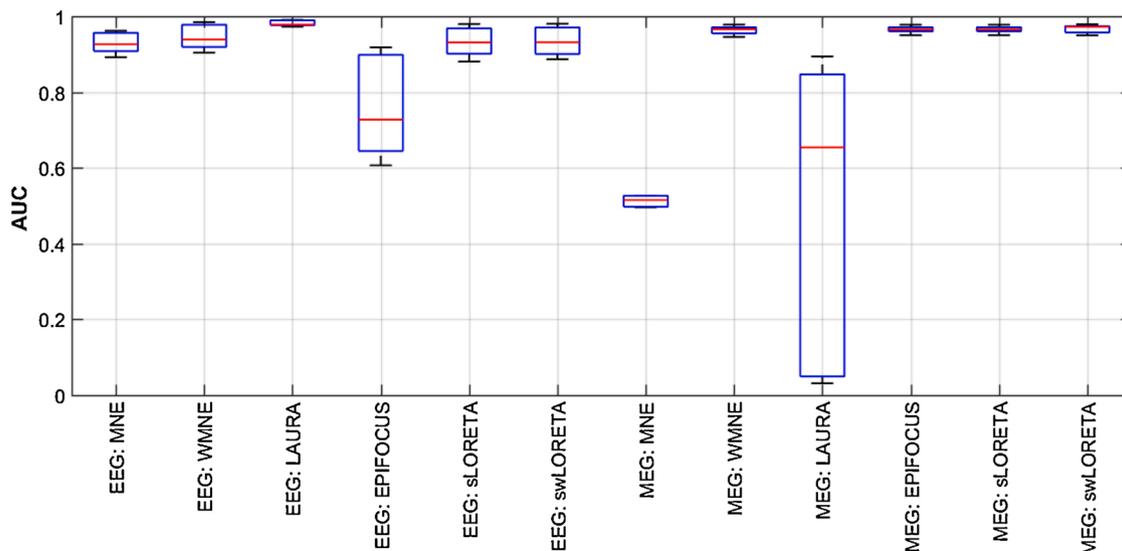


Fig. 8. The AUC localization results of different methods using EEG and MEG on clean train data, used to determine the weights for the proposed method.

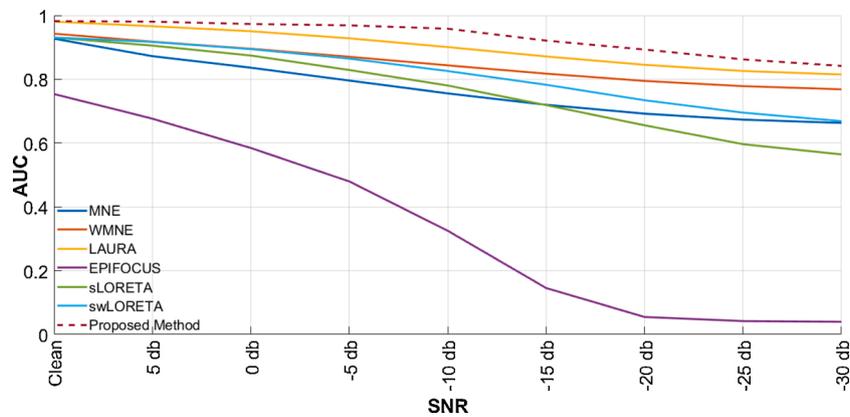


Fig. 9. The localization results obtained by the proposed method in comparison with the common localization methods using EEG data for the clean data and data with different noise levels.

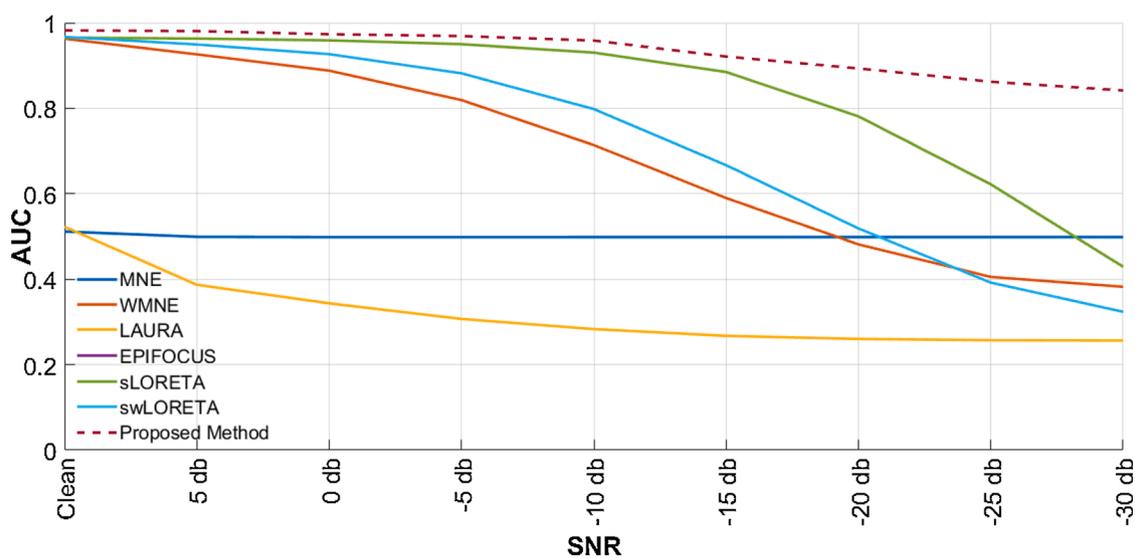


Fig. 10. The localization results obtained by the proposed method in comparison with the common localization methods using MEG data for the clean data and data with different noise levels.

Table 1

Methods performance in terms of AUC and for cases of different Signal to Noise Ratio (SNR), ranging from clean synthetic signal to synthetic data with SNR of -30 db. The table is in accordance to values of Figs. 9 and 10.

Methods	Modality	AUCs (mean ± std) at different SNRs								
		Clean	5 db	0 db	-5 db	-10 db	-15 db	-20 db	-25 db	-30 db
<b>MNE</b>	EEG	0.93 ± 0.02	0.87 ± 0.00	0.84 ± 0.07	0.79 ± 0.07	0.75 ± 0.06	0.72 ± 0.05	0.69 ± 0.03	0.67 ± 0.02	0.66 ± 0.02
	MEG	0.51 ± 0.01	0.50 ± 0.00	0.50 ± 0.00	0.49 ± 0.00	0.56 ± 0.00	0.50 ± 0.07	0.49 ± 0.00	0.50 ± 0.00	0.50 ± 0.00
<b>WMNE</b>	EEG	0.94 ± 0.03	0.92 ± 0.04	0.80 ± 0.05	0.87 ± 0.06	0.86 ± 0.06	0.82 ± 0.06	0.79 ± 0.04	0.78 ± 0.02	0.77 ± 0.01
	MEG	0.96 ± 0.01	0.93 ± 0.04	0.89 ± 0.06	0.89 ± 0.09	0.71 ± 0.11	0.50 ± 0.00	0.48 ± 0.00	0.41 ± 0.15	0.38 ± 0.14
<b>LAURA</b>	EEG	<b>0.98 ± 0.01</b>	0.97 ± 0.02	0.95 ± 0.04	0.93 ± 0.05	0.90 ± 0.50	0.87 ± 0.05	0.84 ± 0.03	0.83 ± 0.03	0.81 ± 0.03
	MEG	0.52 ± 0.35	0.39 ± 0.23	0.34 ± 0.02	0.30 ± 0.17	0.28 ± 0.14	0.27 ± 0.19	0.26 ± 0.18	0.26 ± 0.12	0.26 ± 0.11
<b>EPIFOCUS</b>	EEG	0.75 ± 0.12	0.68 ± 0.16	0.58 ± 0.27	0.48 ± 0.29	0.32 ± 0.29	0.14 ± 0.16	0.70 ± 0.07	0.04 ± 0.05	0.04 ± 0.04
	MEG	0.96 ± 0.01	0.96 ± 0.01	0.96 ± 0.02	0.95 ± 0.03	0.93 ± 0.05	0.88 ± 0.09	0.78 ± 0.19	0.62 ± 0.25	0.43 ± 0.26
<b>sLORETA</b>	EEG	0.93 ± 0.03	0.92 ± 0.05	0.87 ± 0.07	0.83 ± 0.1	0.78 ± 0.12	0.72 ± 0.13	0.66 ± 0.13	0.60 ± 0.12	0.56 ± 0.13
	MEG	0.96 ± 0.01	0.96 ± 0.01	0.96 ± 0.02	0.95 ± 0.03	0.93 ± 0.05	0.88 ± 0.01	0.78 ± 0.19	0.62 ± 0.25	0.43 ± 0.25
<b>swLORETA</b>	EEG	0.93 ± 0.03	0.92 ± 0.03	0.89 ± 0.05	0.86 ± 0.08	0.82 ± 0.06	0.78 ± 0.11	0.73 ± 0.12	0.69 ± 0.14	0.67 ± 0.17
	MEG	0.97 ± 0.01	0.95 ± 0.03	0.93 ± 0.04	0.88 ± 0.07	0.80 ± 0.11	0.67 ± 0.17	0.52 ± 0.19	0.39 ± 0.2	0.32 ± 0.21
<b>Proposed method</b>		<b>0.98 ± 0.01</b>	<b>0.98 ± 0.03</b>	<b>0.97 ± 0.04</b>	<b>0.97 ± 0.05</b>	<b>0.96 ± 0.06</b>	<b>0.92 ± 0.07</b>	<b>0.89 ± 0.05</b>	<b>0.86 ± 0.05</b>	<b>0.84 ± 0.06</b>

Shamsollahi: Supervision, Writing – Review & Editing, Project administration.

## Acknowledgement

The authors wish to thanks Isabelle Merlet, Laurent Albera, Amar Kachenoura and Gwénaél Birot from LTSI, University of Rennes 1, for generating the data and some Matlab codes.

## Declaration of Competing Interest

The authors report no declarations of interest.

## References

- [1] T.-H. Eom, J.-H. Shin, Y.-H. Kim, S.-Y. Chung, I.-G. Lee, J.-M. Kim, Distributed source localization of interictal spikes in benign childhood epilepsy with centrotemporal spikes: a standardized low-resolution brain electromagnetic tomography (sLORETA) study, *J. Clin. Neurosci.* 38 (2017) 49–54.
- [2] M. Hassan, I. Merlet, A. Mheich, A. Kabbara, A. Biraben, A. Nica, F. Wendling, Identification of interictal epileptic networks from dense-EEG, *Brain Topogr.* 30 (1) (2017) 60–76.
- [3] P. Nemtsas, G. Birot, F. Pittau, C.M. Michel, K. Schaller, S. Vulliemoz, V. K. Kimiskidis, M. Seeck, Source localization of ictal epileptic activity based on high-density scalp EEG data, *Epilepsia* 58 (6) (2017) 1027–1036.
- [4] A. Hashemi, S. Haufe, Improving EEG source localization through spatio-temporal sparse Bayesian learning, in: 2018 26th European Signal Processing Conference (EUSIPCO), IEEE, 2018, pp. 1935–1939.
- [5] F.E. Abd El-Samie, T.N. Alotaiby, M.I. Khalid, S.A. Alshebeili, S.A. Aldosari, A review of EEG and MEG epileptic spike detection algorithms, *IEEE Access* 6 (2018) 60673–60688.
- [6] C. Cai, K. Sekihara, S.S. Nagarajan, A novel scanning algorithm for MEG/EEG imaging using covariance partitioning and noise learning, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2019, pp. 4803–4806.
- [7] D. Cohen, B. Cuffin, A method for combining MEG and EEG to determine the sources, *Phys. Med. Biol.* 32 (1) (1987) 85.
- [8] F. Babiloni, C. Babiloni, F. Carducci, G.L. Romani, P.M. Rossini, L.M. Angelone, F. Cincotti, Multimodal integration of EEG and MEG data: a simulation study with variable signal-to-noise ratio and number of sensors, *Hum. Brain Mapp.* 22 (1) (2004) 52–62.
- [9] R. Leahy, J. Mosher, M. Spencer, M. Huang, J. Lewine, A study of dipole localization accuracy for MEG and EEG using a human skull phantom, *Electroencephalogr. Clin. Neurophysiol.* 107 (2) (1998) 159–173.
- [10] A.K. Liu, A.M. Dale, J.W. Belliveau, Monte Carlo simulation studies of EEG and MEG localization accuracy, *Hum. Brain Mapp.* 16 (1) (2002) 47–62.
- [11] J.C. Mosher, R.M. Leahy, P.S. Lewis, EEG and MEG: forward solutions for inverse methods, *IEEE Trans. Biomed. Eng.* 46 (3) (1999) 245–259.
- [12] H. Huizenga, T. Van Zuijlen, D.J. Heslenfeld, P. Molenaar, Simultaneous MEG and EEG source analysis, *Phys. Med. Biol.* 46 (7) (2001) 1737.
- [13] F. Babiloni, C. Babiloni, F. Carducci, D.C. Gratta, G. Romani, P. Rossini, F. Cincotti, Cortical source estimate of combined high resolution EEG and fMRI data related to voluntary movements, *Methods Inf. Med.* 41 (05) (2002) 443–450.
- [14] S. Baillet, L. Garnero, G. Marin, J.-P. Huguonin, Combined MEG and EEG source imaging by minimization of mutual information, *IEEE Trans. Biomed. Eng.* 46 (5) (1999) 522–534.
- [15] S.P. van den Broek, F. Reinders, M. Donderwinkel, M. Peters, Volume conduction effects in EEG and MEG, *Electroencephalogr. Clin. Neurophysiol.* 106 (6) (1998) 522–534.
- [16] F.S. Avarvand, S. Bartz, C. Andreou, W. Samek, G. Leicht, C. Mulert, A.K. Engel, G. Nolte, Localizing bicoherence from EEG and MEG, *Neuroimage* 174 (2018) 352–363.
- [17] N. Mäkelä, M. Stenroos, J. Sarvas, R.J. Ilmoniemi, Truncated rap-music (trap-music) for MEG and EEG source localization, *NeuroImage* 167 (2018) 73–83.
- [18] M.-X. Huang, T. Song, D.J. Hagler Jr., I. Podgorny, V. Jousmaki, L. Cui, K. Gaa, D. L. Harrington, A.M. Dale, R.R. Lee, et al., A novel integrated MEG and EEG analysis method for dipolar sources, *Neuroimage* 37 (3) (2007) 731–748.
- [19] S.S. Dalal, S. Baillet, C. Adam, A. Ducorps, D. Schwartz, K. Jerbi, O. Bertrand, L. Garnero, J. Martinerie, J.-P. Lachaux, Simultaneous MEG and intracranial EEG recordings during attentive reading, *Neuroimage* 45 (4) (2009) 1289–1304.
- [20] R.N. Henson, E. Mouchlianitis, K.J. Friston, MEG and EEG data fusion: simultaneous localisation of face-evoked responses, *Neuroimage* 47 (2) (2009) 581–589.
- [21] S.C. Jun, MEG and EEG fusion in Bayesian frame, *Proc. ICEIE 2* (2010) 295–299.
- [22] S. Ko, S.C. Jun, Beamformer for simultaneous magnetoencephalography and electroencephalography analysis, *J. Appl. Phys.* 107 (9) (2010) 09B315.
- [23] H. Liu, N. Tanaka, S. Stufflebeam, S. Ahlfors, M. Hämäläinen, Functional mapping with simultaneous MEG and EEG, *J. Vis. Exp.* (40) (2010) e1668.
- [24] E. Pirondini, B. Babadi, G. Obregon-Henao, C. Lamus, W.Q. Malik, M.S. Hämäläinen, P.L. Purdon, Computationally efficient algorithms for sparse, dynamic solutions to the EEG source localization problem, *IEEE Trans. Biomed. Eng.* 65 (6) (2017) 1359–1372.
- [25] A.P. Dempster, A generalization of Bayesian inference, *J. R. Stat. Soc. Ser. B: Methodol.* 30 (2) (1968) 205–232.
- [26] G. Shafer, *A Mathematical Theory of Evidence*, vol.42, Princeton university press, 1976.
- [27] P. Smets, Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *Classic Works of the Dempster-Shafer Theory of Belief Functions*, Springer, 2008, pp. 633–664.
- [28] C. Cai, K. Sekihara, S.S. Nagarajan, Hierarchical multiscale Bayesian algorithm for robust MEG/EEG source reconstruction, *NeuroImage* 183 (2018) 698–715.
- [29] R.U. Kay, Fundamentals of the Dempster-Shafer theory and its applications to system safety and reliability modelling, *Reliabil. Theory Appl.* 2 (3–4) (7).
- [30] K. Zaman, S. Rangavajhala, M.P. McDonald, S. Mahadevan, A probabilistic approach for representation of interval uncertainty, *Reliabil. Eng. Syst. Saf.* 96 (1) (2011) 117–130.
- [31] M.S. Hämäläinen, R.J. Ilmoniemi, Interpreting magnetic fields of the brain: minimum norm estimates, *Med. Biol. Eng. Comput.* 32 (1) (1994) 35–42.
- [32] R.D. Pascual-Marqui, Review of methods for solving the EEG inverse problem, *Int. J. Bioelectromagnet.* 1 (1) (1999) 75–86.
- [33] R.U. Kay, Pascual-Marqui, C.M. Michel, D. Lehmann, Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain, *Int. J. Psychophysiol.* 18 (1) (1994) 49–65.
- [34] M.A. Jatoi, N. Kamel, A.S. Malik, I. Faye, T. Begum, A survey of methods used for source localization using EEG signals, *Biomed. Signal Process. Control* 11 (2014) 42–52.
- [35] M.A. Jatoi, N. Kamel, *Brain Source Localization Using EEG Signal Analysis*, CRC Press, 2017.
- [36] R. Grech, T. Cassar, J. Muscat, K.P. Camilleri, S.G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, B. Vanrumste, Review on solving the inverse problem in EEG source analysis, *J. Neuroeng. Rehabil.* 5 (1) (2008) 25.
- [37] A. Al Hilli, L. Najafizadeh, A.P. Petropulu, A weighted approach for sparse signal support estimation with application to EEG source localization, *IEEE Trans. Signal Process.* 65 (24) (2017) 6551–6565.
- [38] R.D. Pascual-Marqui, et al., Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details, *Methods Find. Exp. Clin. Pharmacol.* 24 (Suppl D) (2002) 5–12.
- [39] M. Wagner, M. Fuchs, J. Kastner, Evaluation of sLORETA in the presence of noise and multiple sources, *Brain Topogr.* 16 (4) (2004) 277–280.
- [40] E. Palmero-Soler, K. Dolan, V. Hadamschek, P.A. Tass, sWLORETA: a novel approach to robust source localization and synchronization tomography, *Phys. Med. Biol.* 52 (7) (2007) 1783.
- [41] R.G. de Peralta Menendez, S.G. Andino, G. Lantz, C.M. Michel, T. Landis, Noninvasive localization of electromagnetic epileptic activity: I. Method descriptions and simulations, *Brain Topogr.* 14 (2) (2001) 131–137.
- [42] R.D. Pascual-Marqui, D. Lehmann, M. Koukkou, K. Kochi, P. Anderer, B. Saletu, H. Tanaka, K. Hirata, E.R. John, L. Prichep, et al., Assessing interactions in the brain with exact low-resolution electromagnetic tomography, *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* 369 (1952) (2011) 3768–3784.
- [43] P. Valdes-Sosa, F. Marti, F. Garcia, R. Casanova, Variable resolution electric-magnetic tomography, *Biomag* 96, Springer, 2000, pp. 373–376.
- [44] S. Baillet, Toward functional brain imaging of cortical electrophysiology: Markovian models for magneto and electroencephalogram source estimation and experimental assessments, Orsay, France 11.
- [45] L. Gavitt, S. Baillet, J.-F. Mangin, J. Pescatore, L. Garnero, A multiresolution framework to MEG/EEG source imaging, *IEEE Trans. Biomed. Eng.* 48 (10) (2001) 1080–1087.
- [46] W. Miltner, C. Braun, R. Johnson Jr., G. Simpson, D. Ruchkin, A test of brain electrical source analysis (BESA): a simulation study, *Electroencephalogr. Clin. Neurophysiol.* 91 (4) (1994) 295–310.
- [47] S. Baillet, J.C. Mosher, R.M. Leahy, Electromagnetic brain mapping, *IEEE Signal Process. Mag.* 18 (6) (2001) 14–30.
- [48] K. Sekihara, S. Nagarajan, D. Poeppel, Y. Miyashita, Reconstructing spatio-temporal activities of neural sources from magnetoencephalographic data using a vector beamformer, in: *icassp*, IEEE, 2001, pp. 2021–2024.
- [49] A. Salek-Haddadi, K. Friston, L. Lemieux, D. Fish, Studying spontaneous EEG activity with fMRI, *Brain Res. Rev.* 43 (1) (2003) 110–133.
- [50] C.M. Bishop, *Pattern recognition and machine learning*, 2006 60(1) (2012) 78–78.
- [51] F. Biessmann, S. Plis, F.C. Meinecke, T. Eichele, K.-R. Müller, Analysis of multimodal neuroimaging data, *IEEE Rev. Biomed. Eng.* 4 (2011) 26–58.
- [52] F. Biessmann, F.C. Meinecke, A. Gretton, A. Rauch, G. Rainer, N.K. Logothetis, K.-R. Müller, Temporal kernel CCA and its application in multimodal neuronal data analysis, *Mach. Learn.* 79 (1–2) (2010) 5–27.
- [53] Y. Murayama, F. Biessmann, F.C. Meinecke, K.-R. Müller, M. Augath, A. Oeltermann, N.K. Logothetis, Relationship between neural and hemodynamic signals during spontaneous activity studied with temporal kernel CCA, *Magn. Reson. Imaging* 28 (8) (2010) 1095–1103.
- [54] A.M. Dale, M.I. Sereno, Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach, *J. Cognit. Neurosci.* 5 (2) (1993) 162–176.
- [55] S. Fazli, J. Mehnert, J. Steinbrink, G. Curio, A. Villringer, K.-R. Müller, B. Blankertz, Enhanced performance by a hybrid NIRS-EEG brain computer interface, *Neuroimage* 59 (1) (2012) 519–529.
- [56] R.N. Henson, H. Abdulrahman, G. Flandin, V. Litvak, Multimodal integration of M/EEG and fMRI data in spm12, *Front. Neurosci.* 13 (2019) 300.
- [57] S. Dähne, F. Biessmann, W. Samek, S. Haufe, D. Goltz, C. Gundlach, A. Villringer, S. Fazli, K.-R. Müller, Multivariate machine learning methods for fusing multimodal functional neuroimaging data, *Proc. IEEE* 103 (9) (2015) 1507–1530.

- [58] J. Sui, T. Adali, Q. Yu, J. Chen, V.D. Calhoun, A review of multivariate methods for multimodal fusion of brain imaging data, *J. Neurosci. Methods* 204 (1) (2012) 68–81.
- [59] L. Landini, V. Positano, M. Santarelli, *Advanced Image Processing in Magnetic Resonance Imaging*, CRC Press, 2005.
- [60] J. Liu, V. Calhoun, Parallel independent component analysis for multimodal analysis: application to fMRI and EEG data, in: 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007, IEEE, 2007, pp. 1028–1031.
- [61] N.M. Correa, T. Eichele, T. Adali, Y.-O. Li, V.D. Calhoun, Multi-set canonical correlation analysis for the fusion of concurrent single trial ERP and functional mri, *Neuroimage* 50 (4) (2010) 1438–1445.
- [62] K.J. Friston, L. Harrison, W. Penny, Dynamic causal modelling, *Neuroimage* 19 (4) (2003) 1273–1302.
- [63] A. Ojeda, K. Kreutz-Delgado, T. Mullen, Fast and robust Block-Sparse Bayesian learning for EEG source imaging, *NeuroImage* 174 (2018) 449–462.
- [64] L. Hamid, Ü. Aydin, C. Wolters, U. Stephani, M. Siniatchkin, A. Galka, MEG-EEG fusion by Kalman filtering within a source analysis framework, in: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2013, pp. 4819–4822.
- [65] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping. *Classic Works of the Dempster-Shafer Theory of Belief Functions*, Springer, 2008, pp. 57–72.
- [66] N. Wilson, Algorithms for Dempster-Shafer theory. *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, Springer, 2000, pp. 421–475.
- [67] R.R. Yager, On the Dempster-Shafer framework and new combination rules, *Inf. Sci.* 41 (2) (1987) 93–137.
- [68] K. Sentz, S. Ferson, et al., *Combination of Evidence in Dempster-Shafer Theory*, vol. 4015, Citeseer, 2002.
- [69] F. Wendling, J.-J. Bellanger, F. Bartolomei, P. Chauvel, Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals, *Biol. Cybern.* 83 (4) (2000) 367–378.
- [70] D. Cosandier-Rimélie, I. Merlet, J.-M. Badier, P. Chauvel, F. Wendling, The neuronal sources of EEG: modeling of simultaneous scalp and intracerebral recordings in epilepsy, *NeuroImage* 42 (1) (2008) 135–146.