COMPLEXITY REDUCTION TECHNIQUES IN MUSIC-BASED EEG SOURCE LOCALIZATION

Seyede Mahya Safavi[⋆]*, SeungJae Lee*[⋆]*, Beth Lopour*[†]*, Pai H. Chou*[⋆]

[⋆] Department of Electrical Engineering and Computer Science, University of California, Irvine
[†] Department of Biomedical Engineering, University of California, Irvine

ABSTRACT

Two techniques are proposed to alleviate the computational burden of MUltiple SIgnal Classification (MUSIC) algorithm applied to Electroencephalogram (EEG) source localization. A significant reduction was achieved by parsing the cortex surface into smaller regions and nominating only a few regions for the exhaustive search inherent in the MUSIC algorithm. The nomination procedure involves a dictionary learning phase in which each region is assigned an atom matrix. Moreover, a dimensionality reduction step provided by excluding some of the electrodes is designed such that the Cramer-Rao bound of localization is maintained. It is shown by simulation that computational complexity of the MUSIC-based localization can be reduced up to 80%.

Index Terms— EEG, MUSIC algorithm, complexity reduction, dictionary learning

1. INTRODUCTION

The electroencephalograph (EEG) source localization is a non-invasive brain imaging technique widely used in clinical and cognitive neuroscience applications [1]. In fact, EEG source localization is commonly utilized to pinpoint the epileptogenic areas in patients with intractable epilepsy [2]. Moreover, in cognitive neuroscience the localization analysis of EEG aimed to identify the areas activated during the sensorimotor tasks helps understand the brain functionality. Several methods have been proposed to solve the ill-posed source localization problem. In particular, state of the art techniques such as FOCUSS, LORETA, MUSIC, RAP-MUSIC, and LCMV are suffering from enormous iterations and extensively large dimensions due to high numbers of electrodes [3].

The LCMV, MUSIC, and RAP-MUSIC algorithms are among the equivalent current dipole techniques in which the cortical activity is modeled by a few current dipoles [3]. These techniques involve an exhaustive search over a grid of candidate cortical points extended across the entire cortex surface to find the the location of current dipoles. However, due to sparsity of the active dipoles, such exhaustive search is not efficient. A wiser approach is to identify the cortical areas with lower probabilities of having active dipoles and exclude them from the exhaustive search. In addition, although it was shown that increasing the number of electrodes will improve the localization accuracy [2], the computational complexity will grow quadratically with the number of electrodes. Thus, it is desirable to keep the number of electrodes minimum, while maintaining the localization accuracy.

In this paper, we address the problem of high computational burden, which imposes a need for bulky and power-hungry processing blocks. This paper proposes a method to smartly nominate the cortical areas with higher chances of activity for the exhaustive search, as well as an on-demand electrode selection to reduce the dimensionality of the problem. It is shown that by means of exploiting the proposed reduction techniques, the computational load of the localization can be alleviated by up to 80%.

2. PRELIMINARIES

The cortical activity can be modeled by a number of electric dipoles located at the surface of the cortex [1]. The i^{th} dipole is associated with a location $\boldsymbol{r}_i = [r_i^x, r_i^y, r_i^z]$ and a moment vector $\boldsymbol{s}_i = [s_i^x, s_i^y, s_i^z]$, where s_i^x, s_i^y, and s_i^z are the Cartesian coordinates of the dipole moments. Assuming the EEG setup has M electrodes, the response of each unit-dipole on the electrodes is denoted by an $M \times 3$ matrix $\boldsymbol{G}_i(\boldsymbol{r}_i) = [\boldsymbol{g}_i^x(\boldsymbol{r}_i), \boldsymbol{g}_i^y(\boldsymbol{r}_i), \boldsymbol{g}_i^z(\boldsymbol{r}_i)]$ and called the lead field matrix (LFM). The $M \times 1$ lead field vectors (LFV) $\boldsymbol{g}_i^x(\boldsymbol{r}_i)$, $\boldsymbol{g}_i^y(\boldsymbol{r}_i)$, and $\boldsymbol{g}_i^z(\boldsymbol{r}_i)$ are the response of a unit-magnitude dipole located at $\boldsymbol{r}_i$ and oriented towards x, y, and z axis, respectively. Assuming the dipole locations $\boldsymbol{r}_i$s are fixed, we shall omit the dependence on the location and denote it simply by $\boldsymbol{G}_i = [\boldsymbol{g}_i^x, \boldsymbol{g}_i^y, \boldsymbol{g}_i^z]$ for brevity.

The EEG signal measured on the M electrodes and denoted by an $M \times 1$ vector $\boldsymbol{y(t)} = [y_1(t), y_2(t), ..., y_M(t)]^T$ is the superposition of the contribution of individual active dipoles. Assuming N active dipoles, $N << M$, the measured signal can be represented as follows [1]:

$$\boldsymbol{y(t)} = \boldsymbol{Gs(t)} + \boldsymbol{n(t)}, \tag{1}$$

where $\boldsymbol{s}(t) = [s_1^x(t), s_1^y(t), s_1^z(t), ..., s_N^x(t), s_N^y(t), s_N^z(t)]^T$ is the $3N \times 1$ vector of active dipole moments at time t and

$\boldsymbol{G} = [\boldsymbol{G}_1, \boldsymbol{G}_2, ..., \boldsymbol{G}_N]$ is the $M \times 3N$ total LFM. $\boldsymbol{n}(t)$ is an $M \times 1$ vector of independent identically distributed noise with a Gaussian distribution $\mathcal{N}(0, {\sigma_n}^2)$.

The objective of source localization is to estimate the location of each dipole $\boldsymbol{r}_i, i = 1, 2, ..., N$. The techniques proposed to solve such an ill-posed problem define a grid of N_c candidate points across the cortex with locations $\{\tilde{\boldsymbol{r}}_i\}_{i=1}^{N_c}$ and among these N_c candidates search for the N-best points that minimize the following squared error as [4]:

$$\boldsymbol{r}, \boldsymbol{S} = \arg\min_{\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{S}}} ||\boldsymbol{Y} - \boldsymbol{G}(\tilde{\boldsymbol{r}})\tilde{\boldsymbol{S}}||^2, \quad (2)$$

where the $M \times T$ matrix Y and the $3N \times T$ matrix S are formed by stacking up all the time samples $\{\boldsymbol{y}(t)\}_{t=1}^{T}$ and $\{\boldsymbol{s}(t)\}_{t=1}^{T}$ as $\boldsymbol{Y} = [\boldsymbol{y}(1), ..., \boldsymbol{y}(T)]$ and $\boldsymbol{S} = [\boldsymbol{s}(1), ..., \boldsymbol{s}(T)]$. In this study, two new techniques are proposed to reduce the computational burden of the above exhaustive search. The first technique devises a reduction in the number of the candidate points while the second technique involves an on-demand electrode exclusion. In the following section, both methods are described in details.

3. PROPOSED REDUCTION TECHNIQUES

3.1. Candidate Grid Point Reduction

Considering the fact that the source localization problem described above has a sparse nature, ($N << N_c$), scanning the entire cortex surface to localize the sources is not efficient. On the other hand, the set of lead field vectors have some special characteristics making a smarter search possible as follow

- Any set of $P < M$ lead field vectors is linearly independent.
- The grid points in the spatial proximity have higher LFV cross correlations.

A wiser approach is to cluster the candidate grid points into groups of spatially adjacent points. Only clusters with a higher chance of finding optimal points are being searched for the localization purpose. First of all, based on their distance from the electrodes, the grid points are categorized into M groups. We shall denote the k^{th} electrode position in Cartesian coordinates by $\boldsymbol{r}_k^e = [r^{ex}{}_k, r^{ey}{}_k, r^{ez}{}_k]$. Define $\mathcal{C}_k, k = 1, 2, ..., M$ as the set of grid indices that are closest to k^{th} electrode as

$$\mathcal{C}_k = \{i | 1 < i < N_c, ||\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_k^e||^2 < ||\tilde{\boldsymbol{r}}_i - \boldsymbol{r}_j^e||^2, j \neq k\}, \quad (3)$$

where $\{\tilde{\boldsymbol{r}}_i\}_{i=1}^{N_c}$ are the grid points locations. After clustering the grid points, the best dictionary by which the LFVs can be sparsely represented is derived. Next, the measured EEG signal is projected on the atoms of the learned dictionary. The significance of the projection coefficients will determine which clusters to nominate and the points belonging to those elected clusters will be the candidate points for the exhaustive search. The procedure of dictionary learning and deriving their corresponding EEG signal projection coefficients is brought in the following.

3.1.1. *Dictionary learning*

The trend in learning is that for each cluster, three atom vectors are learned. Let the three $M \times 1$ atoms in k^{th} cluster be denoted as $\boldsymbol{d}_k^x$, $\boldsymbol{d}_k^y$, and $\boldsymbol{d}_k^z, k = 1, ..., M$. The three $M \times M$ matrices $\boldsymbol{D}^x$, $\boldsymbol{D}^y$, and $\boldsymbol{D}^z$ are comprised of these atoms as $\boldsymbol{D}^\ell = [\boldsymbol{d}_1^\ell, \boldsymbol{d}_2^\ell, ..., \boldsymbol{d}_M^\ell], \ell \in \{x, y, z\}$ The best possible dictionary should minimize

$$\hat{\boldsymbol{D}}, \hat{\boldsymbol{\varphi}}_i = \arg\min_{\boldsymbol{D}, \boldsymbol{\psi}_i} ||\boldsymbol{G}_i - \boldsymbol{D}\boldsymbol{\psi}_i||^2, \quad i = 1, ..., N_c \quad (4)$$

where $\hat{\boldsymbol{\psi}}_i$ is the $3M \times 3$ coefficients matrix derived by projecting the $\boldsymbol{G}_i$ on the learned dictionary atoms $\hat{\boldsymbol{D}}$ and $\boldsymbol{G}_i = [\boldsymbol{g}_i^x, \boldsymbol{g}_i^y, \boldsymbol{g}_i^z]$. The $M \times 3M$ dictionary matrix is defined as $\boldsymbol{D} = [\boldsymbol{D}^x, \boldsymbol{D}^y, \boldsymbol{D}^z]$. In addition to (4), the matrices $\boldsymbol{D}^x$, $\boldsymbol{D}^y$, and $\boldsymbol{D}^z$ should be designed orthonormal.

Algorithm 1 demonstrates the proposed procedure for learning the sparse dictionary $\boldsymbol{D}$. According to Algorithm 1, the learning procedure has two phases. In the initialization phase, the atoms of each cluster are initialized by the principle components of the LFVs inside that cluster. However, since the LFVs of spatially adjacent grid points are quite similar, the learned atoms from adjacent clusters are non-orthogonal. Phase two will make the columns of $\boldsymbol{D}^\ell, \ell \in \{x, y, z\}$ orthogonal. The orthogonality property will help when there are multiple active dipoles with different moment amplitudes. For instance, consider a case of two dipoles where one is strong and the other one is weak. The presence of the strong dipole will dominate the principle component of the signal. Thus, the algorithm nominating the clusters for exhaustive search might miss the cluster containing the weak dipole. However, by ensuring the orthogonality of the atoms, each cluster is offering a component that does not exist in other clusters. The presence of such a component in the measured signal is the sign to nominate its corresponding cluster. As will be shown in the simulation section, this will reduce the localization error dramatically.

3.1.2. *Nominating Clusters for Localization Search*

Once the dictionary is learned, one should assess the EEG signal projection on the atom vectors. Inspection of the potential source locations will be carried out among the points belonging to the clusters with highest coefficients. To formulate this in mathematical form, consider the covariance matrix of the EEG $\boldsymbol{y}$ and its eigenvalue decomposition as

$$\boldsymbol{R}_y = \frac{1}{T} \sum_{t=1}^{\mathrm{T}} \boldsymbol{y}(t)\boldsymbol{y}^H(t) = \boldsymbol{U}_y \Delta_y \boldsymbol{U}_y^H, \quad (5)$$

Algorithm 1 Dictionary learning

• **Phase 1, Initialization**:
for all clusters k=1,...,M **do**
-Find covariance Matrix of the Lead Field Vectors for Each Cluster, Factorize the Covariance Matrices Using Eigenvalue Decomposition:

$$\boldsymbol{R}^x_{\mathcal{C}_k} = \frac{1}{|\mathcal{C}_k|}\boldsymbol{G}^x_{\mathcal{C}_k}\boldsymbol{G}^{x\,H}_{\mathcal{C}_k} \qquad \boldsymbol{R}^x_{\mathcal{C}_k} = \boldsymbol{U}^x_{\mathcal{C}_k}\boldsymbol{\Delta}^x_{\mathcal{C}_k}\boldsymbol{U}^{x\,H}_{\mathcal{C}_k}$$
$$\boldsymbol{R}^y_{\mathcal{C}_k} = \frac{1}{|\mathcal{C}_k|}\boldsymbol{G}^y_{\mathcal{C}_k}\boldsymbol{G}^{y\,H}_{\mathcal{C}_k} \qquad \boldsymbol{R}^y_{\mathcal{C}_k} = \boldsymbol{U}^y_{\mathcal{C}_k}\boldsymbol{\Delta}^y_{\mathcal{C}_k}\boldsymbol{U}^{y\,H}_{\mathcal{C}_k}$$
$$\boldsymbol{R}^z_{\mathcal{C}_k} = \frac{1}{|\mathcal{C}_k|}\boldsymbol{G}^z_{\mathcal{C}_k}\boldsymbol{G}^{z\,H}_{\mathcal{C}_k} \qquad \boldsymbol{R}^z_{\mathcal{C}_k} = \boldsymbol{U}^z_{\mathcal{C}_k}\boldsymbol{\Delta}^z_{\mathcal{C}_k}\boldsymbol{U}^{z\,H}_{\mathcal{C}_k}$$

$\boldsymbol{G}^x_{\mathcal{C}_k}$ is the $M \times |\mathcal{C}_k|$ matrix containing the lead field vectors of dipoles inside k^{th} cluster oriented toward x-axis. The columns of $\boldsymbol{U}^y_{\mathcal{C}_k}$ are the eigenvectors of $\boldsymbol{R}^y_{\mathcal{C}_k}$.

-Assign the Principle Components of the Covariances to k^{th} atoms of dictionary matrices $\boldsymbol{D}^x$, $\boldsymbol{D}^y$, $\boldsymbol{D}^z$:

$$[\boldsymbol{d}^x]_k = [\boldsymbol{U}^x_{\mathcal{C}_k}]_1, \quad [\boldsymbol{d}^y]_k = [\boldsymbol{U}^y_{\mathcal{C}_k}]_1\,, \quad [\boldsymbol{d}^z]_k = [\boldsymbol{U}^z_{\mathcal{C}_k}]_1$$

$[\boldsymbol{U}^x_{\mathcal{C}_k}]_1$ is the first column of the matrix
end for
• **Phase 2, Update the Codebook**:
repeat
for all clusters k=1,...,M **do**
-Compute the Cluster Residual Error

$$\boldsymbol{E}^x_{\mathcal{C}_k} = \arg\min_{\psi_j} ||\boldsymbol{G}^x_{\mathcal{C}_k} - \textstyle\sum_{j\neq k}[\boldsymbol{d}^x]_j\psi_j||^2$$
$$\boldsymbol{E}^y_{\mathcal{C}_k} = \arg\min_{\psi_j} ||\boldsymbol{G}^y_{\mathcal{C}_k} - \textstyle\sum_{j\neq k}[\boldsymbol{d}^y]_j\psi_j||^2$$
$$\boldsymbol{E}^z_{\mathcal{C}_k} = \arg\min_{\psi_j} ||\boldsymbol{G}^z_{\mathcal{C}_k} - \textstyle\sum_{j\neq k}[\boldsymbol{d}^z]_j\psi_j||^2$$

-Deriving the Covariance Matrices and the Principle Components of the Residual Errors

$$\boldsymbol{Q}^x_{\mathcal{C}_k} = \frac{1}{|\mathcal{C}_k|}\boldsymbol{E}^x_{\mathcal{C}_k}\boldsymbol{E}^{x\,H}_{\mathcal{C}_k}, \quad \boldsymbol{Q}^x_{\mathcal{C}_k} = \boldsymbol{U}^x_{\mathcal{C}_k}\boldsymbol{\Delta}^x_{\mathcal{C}_k}\boldsymbol{U}^{x\,H}_{\mathcal{C}_k}$$
$$\boldsymbol{Q}^y_{\mathcal{C}_k} = \frac{1}{|\mathcal{C}_k|}\boldsymbol{E}^y_{\mathcal{C}_k}\boldsymbol{E}^{y\,H}_{\mathcal{C}_k}, \quad \boldsymbol{Q}^y_{\mathcal{C}_k} = \boldsymbol{U}^y_{\mathcal{C}_k}\boldsymbol{\Delta}^y_{\mathcal{C}_k}\boldsymbol{U}^{y\,H}_{\mathcal{C}_k}$$
$$\boldsymbol{Q}^z_{\mathcal{C}_k} = \frac{1}{|\mathcal{C}_k|}\boldsymbol{E}^z_{\mathcal{C}_k}\boldsymbol{E}^{z\,H}_{\mathcal{C}_k}, \quad \boldsymbol{Q}^z_{\mathcal{C}_k} = \boldsymbol{U}^z_{\mathcal{C}_k}\boldsymbol{\Delta}^z_{\mathcal{C}_k}\boldsymbol{U}^{z\,H}_{\mathcal{C}_k}$$

-Update the Codebook using the Residual Principle Component:

$$[\boldsymbol{d}^x]_k = [\boldsymbol{U}^x_{\mathcal{C}_k}]_1, \qquad [\boldsymbol{d}^y]_k = [\boldsymbol{U}^y_{\mathcal{C}_k}]_1, \qquad [\boldsymbol{d}^z]_k = [\boldsymbol{U}^z_{\mathcal{C}_k}]_1$$

end for
until Convergence

where T is the total number of time samples, columns of the $M \times M$ matrix $\boldsymbol{U}_y$ are the eigenvectors of $\boldsymbol{R}_y$ and $\boldsymbol{\Delta}_y$ is a diagonal matrix of its eigenvalues. Define the $M \times 3N$ matrix $\overline{\boldsymbol{U}}_y = [\boldsymbol{u}_1, \boldsymbol{u}_2, ..., \boldsymbol{u}_{3N}]$ to have the $3N$ eigenvectors corresponding to the $3N$ most significant eigenvalues. The subspace spanned by $\overline{\boldsymbol{U}}_y$ is called the signal subspace. Finding the correlation of the subspace spanned by the atom vectors and signal subspace is the solution to the generalized eigenvalue problem as follows[4]:

$$\hat{\boldsymbol{\phi}}_k = \arg\max_{\boldsymbol{\phi}_k} \quad \frac{\boldsymbol{\phi}_k^T\boldsymbol{D}_k^T\overline{\boldsymbol{U}}_y\overline{\boldsymbol{U}}_y^T\boldsymbol{D}_k\boldsymbol{\phi}_k}{\boldsymbol{\phi}_k^T\boldsymbol{D}_k^T\boldsymbol{D}_k\boldsymbol{\phi}_k}, \tag{6}$$

where $\boldsymbol{D}_k = [[\boldsymbol{d}^x]_k, [\boldsymbol{d}^y]_k, [\boldsymbol{d}^z]_k]$ is the $M \times 3$ matrix of the three atom vectors learnt from k^{th} cluster and $\boldsymbol{\phi}_k$ is the 3 coefficient vector for the best subspace projection. Equation (6) will be assessed for all clusters $k = 1, ..., M$. Next, the L clusters whose atom vectors mostly maximize the right hand side of (6) will be elected. Searching for the source locations will be confined only to the grid points belonging to these clusters. Algorithm 2 summarizes the reduction procedure.

3.2. Reducing the Number of Electrodes

In research or clinical applications of EEG, there can be as high as 256 evenly spaced electrodes around the scalp. However for localizing the cortical activity, these electrodes do not contribute equally in improving the accuracy of localization. The objective is to introduce an adaptive sensor selection scheme intended to reduce the computational burden. We shall consider the Cramer-Rao bound to quantify the performance for our source localization problem. The Fisher Information Matrix (FIM) is obtained as [5]

$$\boldsymbol{F}(t) = \mathbb{E}\left\{\left(\frac{\partial \ln p(\boldsymbol{y}(t)|\boldsymbol{r})}{\partial \boldsymbol{r}}\right)\left(\frac{\partial \ln p(\boldsymbol{y}(t)|\boldsymbol{r})}{\partial \boldsymbol{r}}\right)^H\right\}, \tag{7}$$

where $\boldsymbol{r} = [\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N]$ is a $1 \times 3N$ vector of N active dipole locations in Cartesian coordinates with $\boldsymbol{r}_i = [r_i^x, r_i^y, r_i^z], i = 1, ..., N$. Let the observations on different electrodes have independent Gaussian distributions $y_i \sim (\mu_i, \sigma_n), i = 1, ..., M$, where the mean μ_i is the expected value of the total responses from the N active dipoles on the i^{th} electrode derived as

$$\mu_i = \sum_{j=1}^{N}[\boldsymbol{g}_j^x]_i\mathbb{E}\{s_j^x(t)\} + [\boldsymbol{g}_j^y]_i\mathbb{E}\{s_j^y(t)\} + [\boldsymbol{g}_j^z]_i\mathbb{E}\{s_j^z(t)\} \tag{8}$$

$[\boldsymbol{g}_j^x]_i$ is the i^{th} element of the $M \times 1$ lead field vector $\boldsymbol{g}_j^x$. σ_n is the measurement noise presumed to be equal on all electrodes. Defining the $M \times 1$ mean vector $\boldsymbol{\mu} = [\mu_1, \mu_2, ..., \mu_M]$ and the $M \times M$ diagonal covariance matrix as $\Sigma = \sigma_n^2\boldsymbol{I}_{(M)}$, the probability distribution of the observation $\boldsymbol{y}$ given N active dipoles at location $\boldsymbol{r}$ is obtained by

$$p(\boldsymbol{y}(t)|\boldsymbol{r}) = \frac{1}{(2\pi\sigma_n^2)^{M/2}}exp\left(-\frac{1}{2}(\boldsymbol{y}(t)-\boldsymbol{\mu})^H\Sigma^{-1}(\boldsymbol{y}(t)-\boldsymbol{\mu})\right). \tag{9}$$

With a slight manipulation from (7) and (9), the FIM is (refer to [5])

$$\boldsymbol{F} = \frac{1}{\sigma_n^2}\mathbb{E}\{\widetilde{\boldsymbol{S}}(t)\boldsymbol{J}^H\boldsymbol{J}\widetilde{\boldsymbol{S}}(t)^H\}, \tag{10}$$

where $\widetilde{\boldsymbol{S}}$ is an $3N \times 9N^2$ block diagonal matrix defined as

$$\widetilde{\boldsymbol{S}}(t) \triangleq \begin{bmatrix} \boldsymbol{s}(t) & \boldsymbol{0} & ... & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{s}(t) & ... & \boldsymbol{0} \\ \vdots & & & \vdots \\ \boldsymbol{0} & \dots & \boldsymbol{0} & \boldsymbol{s}(t) \end{bmatrix}_{3N\times 9N^2}. \tag{11}$$

Algorithm 2 Candidate Point Reduction Procedure

for all clusters k=1,...,M **do**
Calculate $\hat{\boldsymbol{\phi}}_k = \arg\max_{\boldsymbol{\phi}_k} \frac{\boldsymbol{\phi}_k^T\boldsymbol{D}_k^T\overline{\boldsymbol{U}}_y\overline{\boldsymbol{U}}_y^T\boldsymbol{D}_k\boldsymbol{\phi}_k}{\boldsymbol{\phi}_k^T\boldsymbol{D}_k^T\boldsymbol{D}_k\boldsymbol{\phi}_k}$

$z_k = \frac{\hat{\boldsymbol{\phi}}_k^T\boldsymbol{D}_k^T\overline{\boldsymbol{U}}_y\overline{\boldsymbol{U}}_y^T\boldsymbol{D}_k\hat{\boldsymbol{\phi}}_k}{\hat{\boldsymbol{\phi}}_k^T\boldsymbol{D}_k^T\boldsymbol{D}_k\hat{\boldsymbol{\phi}}_k}$
end for
Define Λ as the set of indices k for which $|\boldsymbol{z}_k|$ are the L most significant .
Define $\mathcal{I} = \{\boldsymbol{r}_i | i \in \mathcal{C}_j, j \in \Lambda\}$ as the set of grid points belonging to clusters set Λ.
Restrict the exhaustive search for the localization problem to the points belonging to $\mathcal{I}$

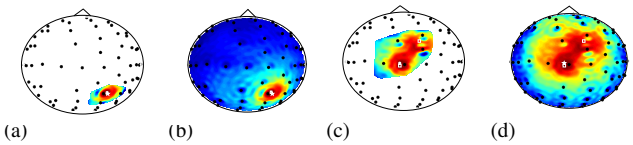

Fig. 1. (a) The topograph of the reduced single active dipole case. (b) The regular MUSIC for the single dipole case in a. (c) The reduced double active dipole case. (d) The regular MUSIC for the double case in c. The actual locations of the dipoles are marked by white squares.

$\boldsymbol{s}(t) = [s_1^x(t), s_1^y(t), s_1^z(t), ..., s_N^x(t), s_N^y(t), s_N^z(t)]$ is a $1 \times 3N$ vector of active dipole moments at time t, and $\mathbf{0}$ is a $1 \times 3N$ vector of all zeros. The Jacobian of $\boldsymbol{G}$, $\boldsymbol{J}$, is defined as the derivative of $\boldsymbol{G}$ with respect to entries of the location vector $\boldsymbol{r}$ as follows:

$$\boldsymbol{J}_{M\times 9N^2} \triangleq \nabla_r \boldsymbol{G} = \left[\nabla_{r_1^x}\boldsymbol{G}, \nabla_{r_1^y}\boldsymbol{G}, \nabla_{r_1^z}\boldsymbol{G}, \ldots, \nabla_{r_N^x}\boldsymbol{G}, \nabla_{r_N^y}\boldsymbol{G}, \nabla_{r_N^z}\boldsymbol{G}\right], \tag{12}$$

The sensor selection problem can be formulated as

$$\tilde{\boldsymbol{F}} = \frac{1}{\sigma_n^2}\mathbb{E}\{\widetilde{\boldsymbol{S}}(t)\boldsymbol{J}^H\boldsymbol{W}\boldsymbol{J}\widetilde{\boldsymbol{S}}(t)^H\}, \tag{13}$$

where $\boldsymbol{W}$ is an $M \times M$ diagonal matrix with diagonal element i specified as follows:

$$w_i = \begin{cases} 1 & if\ the\ i^{th}\ sensor\ is\ included \\ 0 & if\ the\ i^{th}\ sensor\ is\ excluded. \end{cases} \tag{14}$$

As a scalar metric to quantify the performance, the logarithm of the determinant of $\widetilde{\boldsymbol{F}}$ is maximized in a D-optimal design [5]. The objective to adjust the diagonal elements of $\boldsymbol{W}$ in

$$\boldsymbol{W} = \arg\max_{\boldsymbol{W}} \quad \ln\det\widetilde{\boldsymbol{F}} + \lambda||M\boldsymbol{I}_{(M)} - \boldsymbol{W}||^2, \tag{15}$$

where λ is the Lagrange multiplier. We shall resort to a gradient based method to solve the above optimization problem. However, entries of $\boldsymbol{W}$ can only take discrete values of 1s and 0s. In order to make the above expression a continuous function for derivation, diagonal elements of $\boldsymbol{W}$ are substituted by sigmoid functions. A new diagonal matrix $\widetilde{\boldsymbol{W}}$ is defined with diagonal elements as $\widetilde{w}_i = \frac{1}{1+e^{-\rho\upsilon_i}}, \quad \rho > 0$, where $\upsilon_i \in \mathbb{R}$ is a continuous variable. The sigmoid function experiences an abrupt change at the output from 0 to 1 when υ_i increments from negative values toward positive values. Being continuous, the sigmoid function makes the derivation with respect to υ_i possible while it mimics the discrete behavior of elements of $\boldsymbol{W}$.

4. SIMULATION RESULTS

In our simulation, we first synthesized an EEG signal based on equation (1) with 64 electrodes. The LFMs were calculated using a four shell volume conduction model as in [6]. The dipoles were randomly located on the cortex with

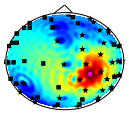

Fig. 2. The selected electrodes, the excluded electrodes, the actual and estimated locations of the active dipoles are depicted by black stars, black squares, a pink star and a pink square respectively.

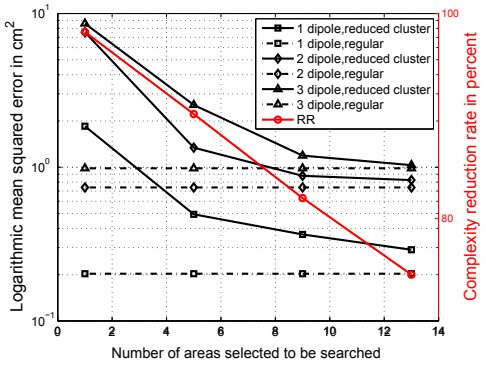

Fig. 3. The mean squared error for three scenarios of one, two , and three active dipoles. The SNR is kept 10 dB.

moments changing over time following independent normal probability distribution. A white Gaussian noise was added to all electrodes to have a signal to noise ratio (SNR) of 10 dB. In Fig. 1, 5 clusters are nominated and it is observed that clusters including the active sources are well among the nominated 5 clusters. We applied the proposed electrode exclusion technique to reduce the dimension of the problem. 28 electrodes were excluded. The optimization problem finds the best set of 28 electrodes to preserve the localization performance. As it is demonstrated in Fig. 2, those electrodes with the greatest distance from the active dipoles are the ones to be omitted by the algorithm. The mean squared error for three cases of one, two and three active dipoles is demonstrated in Fig. 3. In addition, the error values for the regular MUSIC algorithm is also depicted by straight dashed lines for comparison. The reduction rate (RR) illustrated on the right of the figure is

$$RR = \left(1 - \frac{\text{number of nominated grid points}}{\text{total number of grid point over the cortex}}\right) \times 100 \tag{16}$$

The RR is maximum when only one cluster is nominated for search. As the number of nominated clusters increases, both the RR and the mean squared error are decreased.

5. CONCLUSION

The two techniques proposed in this paper reduce the complexity burden of the MUSIC-based EEG source localization up to 80 %. Such reduction will make the real-time, low-power, compact hardware implementation of the EEG source localization feasible to be used in long-term monitoring of the brain activity.

References

[1] H. Becker, A. Laurent, C. Pierre, R. Gribonval, F. Wendling, I. Merlet. "Brain-Source Imaging: From sparse to tensor models." IEEE Signal Processing Magazine, no. 6, pp. 100-112, 2015.

[2] A. Sohrabpour,L. Yunfeng,K. Pongkiat, J. Blount, H. Kim, B. He. "Effect of EEG electrode number on epileptic source localization in pediatric patients." Clinical Neurophysiology 126, no. 3, pp. 472-480, 2015.

[3] M. A. Jatoi, N. Kamel, A.S. Malik, I. Faye, T. Begum, "A survey of methods used for source localization using EEG signals," Elsevier Journal of Biomedical Signal Processing and Control, 11, pp.42-52, 2014.

[4] S. C. Wu, A. L. Swindlehurst, P. T. Wang and Z. Zoran Nenadic, "Projection Versus Prewhitening for EEG Interference Suppression," in IEEE Transactions on Biomedical Engineering, vol. 59, no. 5, pp. 1329-1338, May 2012.

[5] S.P. Chepuri, G. Leus. "Sparsity-promoting sensor selection for non-linear measurement models," IEEE Transaction on Signal Processing, vol. 63, no. 3, pp. 684-698, 2015.

[6] M. Sun, "An Effieceint Algorithm for Computing Multishell Spherical Volume Conductor Models in EEG Dipole Source Localization," in IEEE Transactions on Biomedical Engineering, vol. 44, no. 12, pp. 1243-1252, Dec. 1997.