Research Article

A New Weighted Correlation Coefficient Method to Evaluate Reconstructed Brain Electrical Sources

Jong-Ho Choi,¹ Min-Hyuk Kim,¹ Luan Feng,¹ Chany Lee,² and Hyun-Kyo Jung¹

¹ School of Electrical Engineering and Computer Science, College of Engineering, Seoul National University, Seoul 151744, Republic of Korea

² College of Medicine, Korea University, Seoul 136705, Republic of Korea

Correspondence should be addressed to Jong-Ho Choi, cjhaha@gmail.com

Received 28 October 2011; Accepted 18 January 2012

Academic Editor: Venky Krishnan

Copyright © 2012 Jong-Ho Choi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Various inverse algorithms have been proposed to estimate brain electrical activities with magnetoencephalography (MEG) and electroencephalography (EEG). To validate and compare the performances of inverse algorithms, many researchers have used artificially constructed EEG and MEG datasets. When the artificial sources are reconstructed on the cortical surface, accuracy of the source estimates has been difficult to evaluate. In this paper, we suggest a new measure to evaluate the reconstructed EEG/MEG cortical sources more accurately. To validate the usefulness of the proposed method, comparison between conventional and proposed evaluation metrics was conducted using artificial cortical sources simulated under different noise conditions. The simulation results demonstrated that only the proposed method could reflect the source space geometry regardless of the number of source peaks.

1. Introduction

Noninvasive measurements of brain electrical activities with electroencephalography (EEG) and magnetoencephalography (MEG) enabled us to estimate the underlying cortical activities, thereby contributing to the rapid development of clinical and cognitive neuroscience. To estimate the cortical electrical activities from EEG and MEG, of which the process is often called EEG/MEG source imaging, highly underdetermined inverse problems have to be solved using linear or nonlinear inverse algorithms since the source estimation from EEG and MEG signals is an ill-posed problem, which generally produces blurry or inaccurately positioned source estimates [1]. Many mathematical approaches and techniques have been proposed to estimate accurate source locations and strengths. Among them, minimum-norm estimate (MNE) has been the most widely studied inverse algorithm as MNE is simple and has linearity [2]. MNE chooses a source distribution where the *l*2 norm of the current

distribution is minimized. On the contrary, minimum current estimate (MCE) selects a source where the *l*1 norm of the current is minimized [3]. Other than those two representative algorithms, there have been several modifications of norm minimization, for example, low-resolution electrical tomography (LORETA) algorithm [4] and the focal underdetermined system solution (FOCUSS) algorithm [5].

When a new source imaging algorithm is proposed, the performance of the inverse algorithm needs to be verified and compared with those of the existing ones. Since the reconstructed source distributions are hard to be verified using in vivo experiments, many researchers have used artificial EEG/MEG human skull phantoms [6] or realistically simulated EEG/MEG datasets. Since the use of simulated EEG/MEG data allows us to readily adjust and control noise levels and source configurations, that is, the number and size of source patches, most inverse algorithms are generally verified using simulated EEG/MEG data [2–5]. In recent simulation studies, the source spaces are generally constrained only on the interface between white and gray matter of the cerebral cortex, generally called cortical surface, considering neurophysiology. The orientations of the cortical sources are also assumed to be perpendicular to the cortical surface [7]. In such simulations, both the original source patches and the reconstructed sources are commonly distributed on the cortical surface generally tessellated with surface triangular elements.

For the evaluation of the reconstructed sources, evaluation metrics or error metrics need to be introduced to probe the similarity between the simulated and reconstructed sources. The well-known evaluation metrics are root mean square error (RMSE), shift of the maximum (S_{max}), shift of the center of mass (S_{cm}), and the correlation coefficient (CC) [6]. Each metric has its own pros and cons. In contrast to the conventional geometric error metrics such as S_{max} and S_{cm} , RMSE and CC do not reflect the geometry of the cortical surface. However, compared to S_{max} and S_{cm} , RMSE and CC are reliable specifically when the source distributions are not concentrated to a single peak. For more accurate and robust estimation of the accuracy of reconstructed EEG/MEG sources, we modified CC by giving the geodesic distance weights to the reconstructed sources to reflect the geometric information of cortical surface. To validate the new evaluation metric, named weighted correlation coefficient (WCC), some representative examples were used.

2. Methods

2.1. EEG/MEG Inverse Problem

When a set of *n* possible source locations and *m* sensor positions is given, thanks to the linearity of Maxwell's equations, an EEG/MEG forward model can be described as $\mathbf{b} = \mathbf{Kj} + \mathbf{s}$, where **K** is an *m* by *n* EEG/MEG lead field matrix, **j** is an *n* by 1 unknown source vector, **b** is an *m* by 1 recorded EEG/MEG data, and **s** is the additional sensor noises. The inverse problem for estimating **j** from **b** has no unique solution. To estimate the possible solutions, MNE adopts the following minimization problem:

$$\min_{\mathbf{i}} \|\mathbf{j}\|_2 \quad \text{subject to } \mathbf{b} = \mathbf{K}\mathbf{j} + \mathbf{s}. \tag{2.1}$$

Then, the estimated solution \tilde{j} can be written as

$$\widetilde{\mathbf{j}} = \mathbf{K}^T \left[\mathbf{K} \mathbf{K}^T + \lambda \mathbf{I} \right]^{-1} \mathbf{b},$$
(2.2)

where λ is a regularization parameter, which was determined using the generalized cross-validation method [8].

2.2. Conventional Evaluation Metrics

We assume that both the simulated true sources \mathbf{j} and the estimated sources \mathbf{j} are distributed on the 3D cortical surface. The dimension of both vectors is n by 1, where n is the number of nodes on the source space. We firstly summarize four conventional evaluation metrics, having been frequently used for assessing the accuracy of the source estimates.

2.2.1. Root Mean Square Error

The root mean square error (RMSE) is the most well-known and convenient way to measure the error between the actual source and the estimated source. RMSE is formulated as

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(j_i - \tilde{j}_i \right)^2},$$
(2.3)

where j_i and \tilde{j}_i are the *i*th elements of **j** and \tilde{j} respectively.

This metric is easy to implement and can be used regardless of the shapes of the source distributions. However, RMSE does not reflect the geometry of the cortical surface since RMSE is computed with just vectored values.

2.2.2. Shift of the Maximum of the Estimate

The shift of the maximum of the estimate (S_{max}) is the simplest measure which reflects the geometry of the source space. S_{max} indicates the distance between the locations where the maximum intensities of sources are generated. The maximum intensities of the actual and reconstructed source are assumed to be located at \mathbf{r}_{max} and $\tilde{\mathbf{r}}_{max}$; respectively,

$$\mathbf{r}_{\max} = \max_{\mathbf{r}} j_i, \qquad \widetilde{\mathbf{r}}_{\max} = \max_{\mathbf{r}} j_i, \qquad (2.4)$$

where \mathbf{r}_i is the coordinate of *i*th node, then S_{\max} is defined as

$$S_{\max} = \sqrt{\left(\mathbf{r}_{\max} - \widetilde{\mathbf{r}}_{\max}\right)^2}$$
(2.5)

and ranged from 0 to d_{max} , the maximum distance within the brain.

This measure is reliable only when the actual source is concentrated around the location of the maximum source intensity because it does not consider the distributions of the cortical sources. When S_{max} is adopted as a measure, the merit of distributed source modeling disappears. For example, even when the extents of the true source and the reconstructed sources are largely different, identical maximum location makes the S_{max} value be 0.

2.2.3. Shift of the Center of Mass

The center of mass has been widely used for evaluating various algorithms adopted not only in EEG and MEG but also in other functional brain imaging techniques such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET). The center of mass of the actual source \mathbf{r}_{cm} and the center of mass of the reconstructed source $\tilde{\mathbf{r}}_{cm}$ are computed as

$$\mathbf{r}_{\mathbf{cm}} = \frac{\sum_{i=1}^{n} j_{i} \mathbf{r}_{i}}{\sum_{i=1}^{n} j_{i}}, \qquad \widetilde{\mathbf{r}}_{\mathbf{cm}} = \frac{\sum_{i=1}^{n} \widetilde{j}_{i} \mathbf{r}_{i}}{\sum_{i=1}^{n} \widetilde{j}_{i}}.$$
(2.6)

As assuming the distributed source to be a dipole source placed on the center of mass of the source, the shift of center of mass (S_{cm}) is defined as the distance between \mathbf{r}_{cm} and $\tilde{\mathbf{r}}_{cm}$:

$$S_{\rm cm} = \sqrt{\left(\mathbf{r}_{\rm cm} - \tilde{\mathbf{r}}_{\rm cm}\right)^2}.$$
 (2.7)

 $S_{\rm cm}$ is similar to $S_{\rm max}$ in that the distributed source is considered as a point source placed at a single location. Therefore, $S_{\rm cm}$ is also reliable only when the simulated source is concentrated around $\mathbf{r_{cm}}$. If the distribution of the source has a radial symmetry, $S_{\rm cm}$ becomes equivalent to $S_{\rm max}$.

2.2.4. Correlation Coefficient

The correlation coefficient (CC), a concept adopted from statistics, is a measure of linear dependency between two variables, and the value ranges between -1 and 1. It has been widely employed as a standard measure in various fields of engineering and sciences. The conventional CC is defined as the covariance of **j** and \tilde{j} divided by the product of their standard deviations:

$$CC = \frac{\operatorname{cov}(j, \tilde{j})}{\sqrt{\operatorname{cov}(j, j) \operatorname{cov}(\tilde{j}, \tilde{j})}},$$
(2.8)

where the covariance is defined as

$$\operatorname{cov}\left(\mathbf{j},\widetilde{\mathbf{j}}\right) = \frac{1}{n} \sum_{i=1}^{n} (j_i - j^*) \left(\widetilde{j}_i - \widetilde{j}^*\right), \tag{2.9}$$

and *j*^{*} represents the mean value of the source **j**:

$$j^* = \frac{1}{n} \sum_{i=1}^n j_i, \qquad \tilde{j}^* = \frac{1}{n} \sum_{i=1}^n \tilde{j}_i.$$
(2.10)

If the distribution of the reconstructed sources is similar to that of the actual sources, the value of CC is close to 1; if the distribution of the reconstructed sources is different from

that of the actual sources, CC is close to -1. CC is reliable even when the source distribution is not concentrated to a single location or when the true source has many distinct peaks. However, similar to RMSE, CC cannot reflect the real geometry of the cortical surface.

2.3. New Algorithm: Weighted Correlation Coefficient

When we categorize four conventional measures mentioned in Section 2.2 in terms of geometric consideration, contrast to S_{max} and S_{cm} , RMSE and CC do not reflect the geometry of the cortical surface. However, RMSE and CC are reliable compared to S_{max} and S_{cm} when the source distribution is not concentrated to a single peak. To combine the advantages of both types of conventional measures, we modified CC by giving the source vector a weight reflecting geometrical information of cortical surface. The new evaluation measure, named weighted correlation coefficient (WCC), is defined as

WCC =
$$\frac{\operatorname{cov}(Wj, W\tilde{j})}{\sqrt{\operatorname{cov}(Wj, Wj) \operatorname{cov}(W\tilde{j}, W\tilde{j})}},$$
(2.11)

and **W** is an *n* by *n* weighting matrix that can be computed as

$$\mathbf{W} = \frac{d_{\max}\mathbf{I_n} - \mathbf{D}}{d_{\max}},\tag{2.12}$$

where I_n is an *n* by *n* identity matrix. **D** is an *n* by *n* distance matrix whose element is given as

$$D_{ij} = \left\| \mathbf{r}_i - \mathbf{r}_j \right\|_k \tag{2.13}$$

and d_{max} is the maximum value in **D**. If k = 2, the Euclidean distance is employed, and if k = geo, then the geodesic distance is employed to obtain the distance matrix. The geodesic distance was computed by solving the Eikonal equation on the tessellated cortical surface [9]. The main diagonal of the weight matrix **W** was filled with 1, and the off-diagonal elements were filled with values between 0 and 1. By multiplying weight matrix **W** to the source vector **j**, the geometric information of cortical surface is considered.

Additionally, Euclidean or geodesic distance can be employed in the definition of the distance matrix **D**. Since the cortical surface of a human brain is folded, the geodesic distance is more suitable to reflect the geometric information of the cortical surface than the Euclidean distance. The Euclidean distance is computed by the Cartesian coordinates regardless of the geometrical feature of the cortical surface. However, as the geodesic distance implies the minimum distance along the surface, the geodesic distance between the two adjacent gyri should be greater than the Euclidean distance. Figure 1 is an example of the Euclidean and geodesic distance between each cortical surface vertex and a reference point located at right dorsolateral prefrontal cortex, corresponding to a column of the distance matrix.

Once the distance matrix is computed, the weighting matrix is determined from (2.13). Figure 2 shows the Euclidean and geodesic weights corresponding to the Euclidean and

Measures	Reflection of geometry	Multi-peaks	Bound	Unit
RMSE	no	Yes	$0 \sim \infty$	no unit
S_{\max}	yes	No	$0 \sim d_{\max}$	mm
$S_{\rm cm}$	yes	No	$0 \sim d_{\max}$	mm
CC	no	Yes	-1 ~ 1	no unit
WCC	ves	Yes	-1 ~ 1	no unit

Table 1: The merits and demerits of measures.

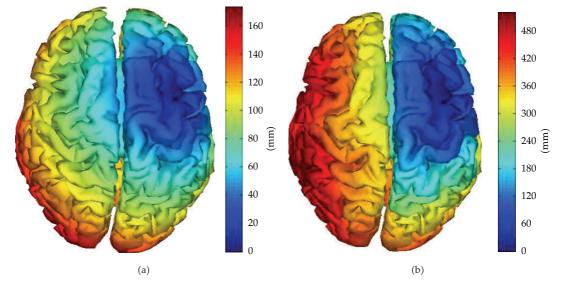


Figure 1: One column of (a) Euclidean and (b) geodesic distance matrix visualized on the cortical surface.

geodesic distance matrices exampled in Figure 1. In contrast to Figure 1, as getting farther from the reference point, the weight is getting smaller.

The characteristics of the conventional and proposed measures are summarized in Table 1. Low values of RMSE, S_{max} , or S_{cm} and high values of CC and WCC indicate the accurate reconstruction. Only WCC is applicable to the case of multipeak and can consider the geometry of source space.

3. Results

To compare and verify the conventional and proposed measures, a simple two-dimensional example was simulated as shown in Figure 3. The source space was defined as a two-dimensional rectangle. The actual source distribution x is given in Figure 3(a), and five reconstructed sources are given in Figures 3(b)-3(f), each of which was denoted as y_1 , y_2 , y_3 , y_4 , and y_5 . The source current intensities are indicated with different colors. If we evaluate the reconstructed sources based on visual inspection, anyone would agree that y_1 is the most accurate reconstruction and y_2 is the second best one. y_5 seem to be the worst reconstruction as the peak location is farthest from the actual one and no reconstructed source is overlapped with the actual one. y_3 and y_4 seems to be better matched than y_5 , but it is difficult to judge which result is better. The result y_3 has no commonly activated region with the actual

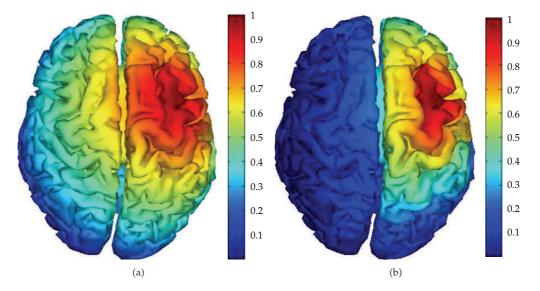


Figure 2: One column of (a) Euclidean and (b) geodesic weighted matrix visualized on the cortical surface.

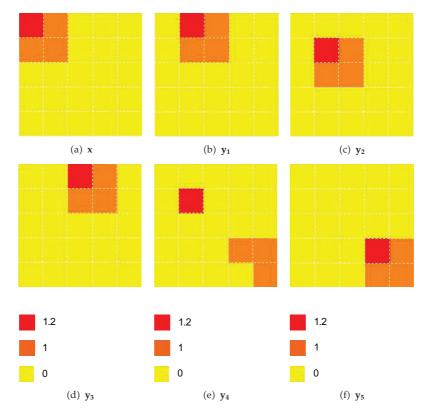


Figure 3: Example of a simulated two-dimensional source space: actual source distribution (a) and reconstructed sources (b)-(f).

	y 1	y 2	y 3	y 4	y 5
RMSE	2.12	2.53	2.98	2.73	2.98
$S_{\rm max}$	1.00	1.41	2.00	1.41	4.24
$S_{\rm cm}$	1.00	1.41	2.00	2.87	4.24
CC	0.40	0.13	-0.19	0.25	-0.19
WCC	0.89	0.80	0.48	-0.31	-0.79

Table 2: Evaluation of reconstructions depicted in Figure 1.

	V 1	y 2	¥3	V 4	Vr
RMSE	9.76	9.95	10.05	10.06	<u>y</u> 5 43.58
$S_{\max,E}$	10.07	21.94	31.39	59.54	20.07
$S_{\max,G}$	9.13	28.12	71.10	122.11	51.71
$S_{\mathrm{cm},E}$	4.16	18.04	31.60	54.95	13.35
$S_{\mathrm{cm},G}$	8.21	21.67	54.32	75.15	24.11
CC	0.32	0.09	0.00	0.00	0.00
WCC_E	0.92	0.89	0.82	0.60	0.58
WCC _G	0.95	0.86	0.43	0.36	0.35

Table 3: Evaluation of reconstructions depicted in Figure 4.

source, but the distribution is close to the actual source distribution, whereas y_4 has slightly overlapped region, but other regions are located far from the actual source location. If we assume visual inspection (VI) as a qualitative measure, the rank of the reconstructed sources can be expressed as $VI(x, y_1) > VI(x, y_2) > VI(x, y_3) \ge VI(x, y_4) > VI(x, y_5)$.

We then employed the conventional and proposed quantitative measures for the evaluation of the reconstructions depicted in Figure 1 and summarized the result in Table 2. All measures commonly indicated that y_1 is the best reconstruction and y_5 is the worst reconstruction. However, the different metrics showed different evaluation results for y_2 , y_3 , and y_4 . In the case of RMSE, y_3 was evaluated as the worst reconstruction, and y_4 and y_2 had an identical RMSE value, which was because RMSE was affected by the commonly activated regions regardless of the source geometry. In the case of S_{max} , which considers only the maximum location of the source, the results of y_2 and y_4 were equivalent. Similar to RMSE, CC classified y_3 as the worst reconstruction and y_4 and y_2 had an identical CC value. Both S_{cm} and WCC evaluated the reconstruction results identically to the visual inspection results. However, if the actual source has multiple peaks, S_{cm} cannot be accurately evaluated.

We then applied the conventional and proposed evaluation metrics to the evaluation of distributed sources on the cortical surface. The cortical surface was extracted from a structural MRI of a standard brain atlas provided by the Montreal Neurological Institute (MNI). To extract and tessellate the cortical surface, we used CURRY6 for Windows (Compumedics, Inc., El Paso, TX). The actual source **x** defined on the cortical surface is given in Figure 4(a), and the reconstructed sources are given in Figures 4(b)–4(f), each of which was denoted as **y**₁, **y**₂, **y**₃, **y**₄ and **y**₅. If we verify the reconstructed sources by visual inspection, **y**₁ seems to be the most accurate reconstruction and **y**₂ seems to be the second best one. The two distributions **y**₄ and **y**₅ are very different from the actual source distribution. We can roughly estimate the rank of the reconstructed sources as VI(**x**, **y**₁) > VI(**x**, **y**₂) > VI(**x**, **y**₄) \geq VI(**x**, **y**₅).

The quantitative evaluation results with conventional and proposed metrics are shown in Table 3. In the case of RMSE, the RMSE values corresponding to $y_1 \sim y_4$ were increasing,

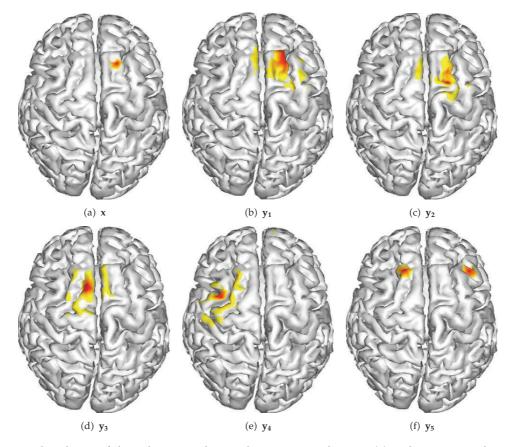


Figure 4: Simulation of three-dimensional cortical sources: actual source (a) and reconstructed sources (b)–(f).

which coincided with the visual inspection results, but the increment was very small compared to the absolute values of RMSE. The results of $S_{\max,E}$ and $S_{\max,G}$ showed that \mathbf{y}_5 is the better reconstruction than \mathbf{y}_3 . Moreover, since the center of mass of \mathbf{y}_5 was located near the actual source, $S_{cm,E}$ and $S_{cm,G}$ of \mathbf{y}_5 were less than those of \mathbf{y}_2 . Subscripted *E* and *G* indicate that the Euclidean and geodesic distance matrices were adopted, respectively. CC also could not distinguish the difference between \mathbf{y}_3 and \mathbf{y}_4 . WCC_{*E*} and WCC_{*G*} are the WCC results when the Euclidean and geodesic distance matrices were adopted, respectively, to construct the weighting matrix. Both WCC_{*E*} and WCC_{*G*} were proved to be reliable since both results coincided well with the visual inspection results. However, compared to WCC_{*G*}, WCC_{*E*} could not reflect the large difference between \mathbf{y}_2 and \mathbf{y}_3 , which are located even in different hemispheres.

We performed extensive computer simulations to quantitatively compare the performance of the proposed measure with that of conventional measures. 2,000 locations on the cortical surface are selected randomly, and on each location a single constant source patch is generated. Consequently, our computer simulations used 2,000 cortical patches whose averaged radius is 6 mm (\pm 1.2 mm). After solving the MEG forward problem with each cortical patch, different-level white Gaussian noise are added to the simulated MEG signal data. We set the signal-to-noise ratio (SNR) values from -10 dB to 30 dB. The reconstructed source

dB	30	20	10	0	-10
RMSE	11.21	11.37	11.33	32.22	53.81
$S_{\max,E}$	12.71	12.54	21.17	19.29	20.63
$S_{\max,G}$	14.32	18.90	30.82	29.68	23.53
$S_{\mathrm{cm},E}$	8.52	11.41	11.21	15.26	13.17
$S_{\mathrm{cm},G}$	15.11	21.67	32.43	31.52	34.21
CC	0.54	0.12	0.01	0.01	0.01
WCC_E	0.93	0.85	0.74	0.63	0.52
WCC_G	0.94	0.86	0.65	0.56	0.42

Table 4: Evaluation of reconstructions with 2,000 randomly selected source patches.

is computed by the minimum norm method with each MEG signal then evaluated with different measures. Table 4 shows the averaged accuracies with conventional and proposed measures with respect to SNR. We expect that results of reconstruction to become more accurate as SNR is getting higher. In the case of RMSE, though the RMSE is increasing as SNR becomes lower, the difference between 10, 20, and 30 dB cases is not clear as much the cases of 0 and -10 dB. In the cases of geodesic measures ($S_{max,E}$, $S_{cm,E}$, $S_{max,G}$, $S_{cm,G}$), the relation of accuracy and SNR is not consistent with the expected tendency. Moreover, the accuracy of low SNR is occasionally greater than that of high SNR in geodesic measures. In the case of CC with high SNR (10, 0, and -10 dB), there is only a marginal difference compared to that of low SNR. Only the accuracy measured by WCC is consistently decreasing as the SNR becomes lower.

4. Conclusion

The geometric measures (S_{max} , S_{cm}) could reflect the geometry of the source space, while the statistical measures (RMSE, CC) could be applied regardless of the distribution characteristics of the sources. In this paper, a new evaluation metric named WCC was proposed to combine the advantages of both types of evaluation metrics and showed enhanced performances compared to the conventional metrics. From the extensive simulation, we could conclude that the proposed measure is very promising to evaluate accuracy of reconstructed sources or EEG/MEG inverse algorithms.

References

- D. Cohen and E. Halgren, "Magnetoencephalography," in *Encyclopedia of Neuroscience*, L. R. Squire, Ed., pp. 615–622, Academic Press, Oxford, UK, 2009.
- [2] M. S. Hamalainen and R. J. Ilmoniemi, "Interpreting magnetic fields of the brain: minimum norm estimates," *Medical and Biological Engineering and Computing*, vol. 32, no. 1, pp. 35–42, 1994.
- [3] K. Uutela, M. Hämäläinen, and E. Somersalo, "Visualization of magnetoencephalographic data using minimum current estimates," *NeuroImage*, vol. 10, no. 2, pp. 173–180, 1999.
- [4] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, "Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain," *International Journal of Psychophysiology*, vol. 18, no. 1, pp. 49–65, 1994.
- [5] I. F. Gorodnitsky, J. S. George, and B. D. Rao, "Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm," *Electroencephalography and Clinical Neurophysiology*, vol. 95, no. 4, pp. 231–251, 1995.

Journal of Applied Mathematics

- [6] S. Baillet, J. J. Rira, G. Main, J. F. Magin, J. Aubert, and L. Ganero, "Evaluation of inverse methods and head models for EEG source localization using a human skull phantom," *Physics in Medicine and Biology*, vol. 46, no. 1, pp. 77–96, 2001.
- [7] A. M. Dale and M. I. Sereno, "Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach," *Journal of Cognitive Neuroscience*, vol. 5, no. 2, pp. 162–176, 1993.
- [8] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215–223, 1979.
- [9] A. Bartesaghi and G. Sapiro, "A system for the generation of curves on 3D brain images," *Human Brain Mapping*, vol. 14, no. 1, pp. 1–15, 2001.