

# Graph Fractional-Order Total Variation EEG Source Reconstruction\*

Ying Li<sup>1</sup> Jing Qin<sup>2</sup> Stanley Osher<sup>3</sup> and Wentai Liu<sup>4</sup>

**Abstract**—EEG source imaging is able to reconstruct sources in the brain from scalp measurements with high temporal resolution. Due to the limited number of sensors, it is very challenging to locate the source accurately with high spatial resolution. Recently, several total variation (TV) based methods have been proposed to explore sparsity of the source spatial gradients, which is based on the assumption that the source is constant at each subregion. However, since the sources have more complex structures in practice, these methods have difficulty in recovering the current density variation and locating source peaks. To overcome this limitation, we propose a graph Fractional-Order Total Variation (gFOTV) based method, which provides the freedom to choose the smoothness order by imposing sparsity of the spatial fractional derivatives so that it locates source peaks accurately. The performance of gFOTV and various state-of-the-art methods is compared using a large amount of simulations and evaluated with several quantitative criteria. The results demonstrate the superior performance of gFOTV not only in spatial resolution but also in localization accuracy and total reconstruction accuracy.

**Index Terms**—EEG source localization, inverse problem, fractional-order total variation, alternating direction method of multipliers

## I. INTRODUCTION

Electroencephalography (EEG) is a valuable tool for studying the brain activity with excellent temporal resolution ( $\sim$ ms) as well as supporting free body movement. EEG source imaging aims at localizing brain sources from scalp measurements by solving the corresponding *inverse problem*. However, due to the limited number of measurements ( $\sim 10^2$ ) comparing with the large number of possible dipole source localizations ( $\sim 10^4$ ), the EEG inverse problem is highly under-determined and has infinite solutions. By imposing various types of regularizations, the uniqueness of the solution can be guaranteed. So far, a large number of EEG reconstruction methods have been proposed to improve reconstruction resolution and localization accuracy by employing various regularization techniques. One of the most widely used methods is based on the Tikhonov regularization using the  $\ell_2$ -norm, such as minimum norm estimate (MNE)[1] and standardized low resolution brain electromagnetic tomography (sLORETA)[2]. The major drawback of

this type of methods is its relatively low spatial resolution. To further improve the spatial resolution, the  $\ell_1$ -norm based methods are proposed with sparsity constraint in the original source domain, such as minimum current estimate (MCE)[3]. Although the spatial resolution is greatly improved, these methods only involve a few elements in each source region, thus failing to identify the spatial extent of distributed sources. Recently, the total variation (TV) based methods [4], [5] are developed by imposing sparsity of the first spatial derivative rather than the original source. It shows great potential in preserving sharpness of edges and estimating the spatial extent of distributed sources. However, relying on the assumption that the underlying source is constant in each subregion, the TV-based methods suffer from the so-called staircase artifacts for the sources with higher order smoothness, which results in loss of the gradual intensity change in the reconstructed source and poor localization of source peaks.

In this paper, we propose a novel EEG source reconstruction method based on the Fractional-Order Total Variation (FOTV) to improve the reconstruction accuracy of the brain image. FOTV has been recently proposed to solve image processing problems [6], [7]. Different from TV which imposes sparsity of first-order spatial derivatives, FOTV can choose a more elegant smoothness order for the underlying source by imposing sparsity of  $\alpha$ -order derivatives ( $\alpha > 0$ ). Therefore, the proposed method is capable of reconstructing the brain image with higher-order smoothness and preserving natural intensity changes of the brain image. As a consequence, the localization accuracy of peaks is greatly improved compared to TV-based methods. In order to extend the traditional FOTV defined on a 2D rectangular grid to an irregular cortex surface, we treat the cortex surface as a graph and define a novel graph FOTV (gFOTV) using shortest paths on the graph. In fact, TV-based methods can be considered as a special case in our proposed framework when  $\alpha = 1$ . In addition, we derive an efficient algorithm using the alternating direction method of multipliers (ADMM). Finally, a large variety of simulations are conducted under different source sizes with a realistic head model. The proposed method is compared with several state-of-the-art ones both qualitatively and quantitatively, which demonstrates superior performance in terms of total reconstruction accuracy, spatial resolution and localization accuracy.

## II. EEG INVERSE PROBLEM

Given acquired electrical recordings, the “EEG inverse problem” aims to identify the distributions of currents in a human brain. More precisely, after discretizing the cortex

\*This work was supported in part by the California Capital Equity LLC and the Keck foundation.

<sup>1</sup>Ying Li is with the Department of Bioengineering, University of California, Los Angeles, CA 90095, USA [yingli.ucla@gmail.com](mailto:yingli.ucla@gmail.com)

<sup>2</sup>Jing Qin is with the Department of Mathematics, University of California, Los Angeles, CA 90095, USA [jxq@ucla.edu](mailto:jxq@ucla.edu)

<sup>3</sup>Stanley Osher is with the Department of Mathematics, University of California, Los Angeles, CA 90095, USA [sjo@math.ucla.edu](mailto:sjo@math.ucla.edu)

<sup>4</sup>Wentai Liu is with the Department of Bioengineering, Electrical Engineering and California NanoSystem Institute, University of California, Los Angeles, CA 90095, USA [wentai@ucla.edu](mailto:wentai@ucla.edu)

surface into  $M$  triangles (called ‘‘voxels’’), the electrical potential  $b \in \mathbb{R}^N$  on the scalp measured by electrodes is related to the current density  $u \in \mathbb{R}^M$  via

$$b = Au + n, \quad (1)$$

where  $n$  denotes the noise, and the lead field matrix  $A$  can be calculated by constructing a head model using the geometric and electrical properties of the head. In this work we use the realistic head model and source model templates provided by Fieldtrip toolbox [8]. In general, since the number of voxels  $M$  ( $\sim 10^4$ ) is much larger than that of electrodes  $N$  ( $\sim 10^2$ ), the corresponding inverse problem is ill-posed. In order to obtain a unique solution for the problem, it is necessary to consider an appropriate regularization based on geometry of the cortex surface as well as the current density distribution of sources.

### III. GRAPH FRACTIONAL-ORDER TOTAL VARIATION

Fractional-order TV has been proposed recently in image processing to enhance smoothness with relatively low computational cost [6], [7], [9]. In this paper, we focus on the anisotropic version in which the objective function of the derived minimization problem is separable. Recall that the anisotropic fractional-order TV in a 2D rectangular mesh is defined as follows:

$$TV_\alpha(u) = \|\nabla^\alpha u\|_1 = \sum_{i,j=1}^M (|(D_x^\alpha u)_{i,j}| + |(D_y^\alpha u)_{i,j}|),$$

where  $\alpha \in (1, 2)$ . Here the fractional derivative is based on the Grwald-Letnikov derivative definition [10]

$$(D_x^\alpha u)_{i,j} = \sum_{k=0}^K w_\alpha(k) u(i-k, j), \quad (D_y^\alpha u)_{i,j} = \sum_{k=0}^K w_\alpha(k) u(i, j-k),$$

where the coefficients are

$$w_\alpha(k) = (-1)^k \frac{\Gamma(\alpha + 1)}{k! \Gamma(\alpha - k + 1)}.$$

It is easy to see that  $w_\alpha(0) = 1$  and  $w_\alpha(1) = -\alpha$ . Moreover, if  $\alpha = 1$ , then  $TV_\alpha$  is the traditional TV. If  $\alpha = 2$ ,  $D_x^\alpha/D_y^\alpha$  approximates the second partial derivative of  $u$  along the  $x/y$ -direction. Although the above definition is also valid for  $\alpha \in (0, 1) \cup (2, \infty)$ , it has experimentally shown that  $\alpha \in (1, 2)$  achieves the best performance in applications [7].

Since the cortex surface is an irregular 3D surface consisting of gyrus and sulcus, we define a graph  $\alpha$ -order TV with  $\alpha \in [1, 2]$  tailored to such surface. After discretizing the cortex surface as mentioned in Section II, we create a graph whose nodes correspond to the centroids of all triangles. For a specific node  $v_i$ , let  $d(v_i, v_j)$  be the number of nodes on the shortest path connecting the nodes  $v_i$  and  $v_j$ , which is in or close to a geodesic of the underlying surface passing through  $v_i$  and  $v_j$ . Given a path  $p = (v_{i=m_0}, v_{m_1}, \dots, v_{m_K})$  where the shortest distance between the nodes  $v_{m_0}$  and  $v_{m_j}$  is  $j$  nodes, the fractional-order derivative along the path  $p$  is defined as

$$(D_p^\alpha u)_i := D_p^\alpha u(v_i) = \sum_{v \in p} w_\alpha(d(v_i, v)) u(v) = \sum_{j=0}^K w_\alpha(j) u(v_{m_j}).$$

Then the discretized fractional-order TV of  $u$  is defined as follows:

$$TV_\alpha(u) = \|D_\alpha u\|_1 = \sum_{i=1}^M \sum_{p \in \mathcal{P}(i;K)} |(D_p^\alpha u)_i|,$$

where  $\mathcal{P}(i;K)$  is the set of all paths starting from the  $i$ th node with length of  $K$  nodes. The shortest path between each node pair can be calculated by the breadth-first search (BFS) algorithm. For a specific node  $v_i$ , the nodes at level  $k$ , i.e., the nodes have shortest distance  $k$  from  $v_i$ , are assigned the weight  $w_\alpha(k)$ . In particular, if  $\alpha = 1$ , then  $D_\alpha$  is exactly the finite difference operator used in the TV-based methods. If  $\alpha = 2$ , then  $w_\alpha(k) = 0$  for  $k > 2$  which implies that all nodes at level more than two are assigned zero weight. If  $\alpha \in (1, 2)$ , then the weights  $w_\alpha(k)$  will gradually decay as  $k$  goes to infinity. As the value of  $\alpha$  increases from 1 to 2, the decay rate of weights  $w_\alpha(k)$  becomes larger. Note that  $K$  specifies the maximal level of nodes to be used. Due to the sparse structure of the underlying  $u$ , it is sufficient to choose  $K \leq 4$  levels of neighboring nodes to achieve the desired accuracy in our experiments.

### IV. MODEL AND ALGORITHM

After defining FOTV on triangular mesh of cortex surface, we propose the following fractional-order TV regularized EEG source reconstruction model to improve high order smoothness of the brain image

$$\min_u \left\{ \frac{1}{2} \|Au - b\|_2^2 + \lambda TV_\alpha(u) \right\}. \quad (2)$$

where  $\lambda > 0$  is regularization parameter, which controls the tradeoff between the data fidelity term and the sparsity term. By change of variables, the above problem can be converted to a linear equality constrained minimization problem

$$\min_{u,v} \left\{ \frac{1}{2} \|Au - b\|_2^2 + \lambda \|v\|_1 \right\} \quad \text{subject to} \quad D_\alpha u = v. \quad (3)$$

Then the ADMM yields the following algorithm

$$\begin{cases} v = \text{shrink}(D_\alpha u + \tilde{v}, \lambda/\rho) \\ u = \underset{u}{\text{argmin}} \left\{ \frac{1}{2} \|Au - b\|_2^2 + \frac{\rho}{2} \|D_\alpha u - v + \tilde{v}\|_2^2 \right\} \\ = (A^T A + \rho D_\alpha^T D_\alpha)^{-1} (A^T b + \rho D_\alpha^T (v - \tilde{v})) \\ \tilde{v} \leftarrow \tilde{v} + \gamma (D_\alpha u - v) \end{cases} \quad (4)$$

Here the parameters  $\rho > 0$ ,  $\gamma \in (0, (\sqrt{5} + 1)/2)$  and the shrinkage operator is defined componentwise as  $\text{shrink}(u, \mu)_i = \text{sign}(u_i) \max\{|u_i| - \mu, 0\}$ .

### V. EXPERIMENTAL RESULTS

#### A. Simulation Protocol

In our simulation, Gaussian patches are used to simulate sources in the brain. To represent sources at different locations, we randomly select three sources located at different lobes of the cortex surface. In addition, to evaluate the performance of the proposed methods for different source sizes, we simulate sources containing 50, 100, 150, 200, 250

voxels, corresponding to 1.0cm, 1.4cm, 1.7cm, 1.9cm, 2.1cm in radius, respectively. After sources are generated, random independent and identically distributed (i.i.d.) Gaussian noise is added to each voxel as background neural noise. Additionally, electrode and electronic noise are added to each channel. For the signal-to-noise ratio (SNR), it is set to be 10dB by default. The number of electrode used is 346, and the number of voxels in the source model is 10240. Finally, the signal is normalized to between  $10\mu V$  to  $100\mu V$ , which is a typical range of amplitude for an adult EEG signal [12]. To reduce round-off errors, both the lead field matrix  $A$  and the electrical potential  $b$  are scaled by  $10^5$ .

### B. Quantitative Criteria

The following three quantitative criteria are employed to evaluate performance of source reconstruction methods.

- *Total reconstruction error (TRE)*, aiming to estimate the total reconstruction accuracy. It measures the relative error between the reconstructed source and the true one, and is defined as  $TRE = \|\hat{u} - u\|_2 / \|u\|_2$ , where  $u$  and  $\hat{u}$  are the true and reconstructed source respectively.
- *Localization error (LE)*, aiming at evaluating the localization accuracy. It measures the averaged localization error between the peaks of the true and reconstructed sources [13], [14], and can be expressed as  $LE = \sum_k LE_k / K$ ,  $LE_k = \{d_{ki} | i = \operatorname{argmin}_{i' \in I_k} \|u_{i'}\|_2\}$ , where  $d_{ki}$  is the distance between the  $i$ th voxel to the peak of the  $k$ th true source, and  $I_k$  is a set of the voxel indices that are closest to the peak of the  $k$ th source.
- *Degree of focalization (DF)* is used to evaluate spatial resolution by measuring how focal the reconstructed source is. It is defined as  $DF = \|\hat{u}_S\|_2^2 / \|u_S\|_2^2$  [14], where  $u_S$  is  $u$  restricted on the original patch area  $S$ .

### C. Parameter Selection

In the operator  $D_\alpha$ , the parameter  $\alpha$  specifies the order of the spatial derivative domain where we want to impose sparsity constraint on. As  $\alpha$  becomes larger, we impose sparsity of higher order derivatives, and therefore the reconstructed sources become smoother and decay faster. When  $\alpha = 1$ , the source decay speed is the slowest, which is the case of TV that the current density is piecewise constant. In our experience, it works well when  $\alpha = 2$ . By choosing  $\alpha$  to be a fraction between 1 and 2, all level nodes will be assigned a non-zero weight, which enhances reconstruction smoothness. Specifically,  $\alpha = 1.6$  is appropriate for all our experiments. In addition, the spatial resolution will be higher than that of  $\alpha = 2$ . In the proposed Algorithm (4), the regularization parameter  $\lambda$  controls the balance between the data fidelity term and the sparsity term. When the source size becomes larger,  $\lambda$  should be tuned to be a little larger. As the noise level becomes higher,  $\lambda$  needs to be a little smaller. According to our experience, it works pretty well by simply choosing  $\lambda$  around 1. The parameter  $\rho$  affects the convergence speed of Algorithm (4), and is set as  $10\lambda$  by default. Finally, the parameter  $\gamma$  is fixed as 1.

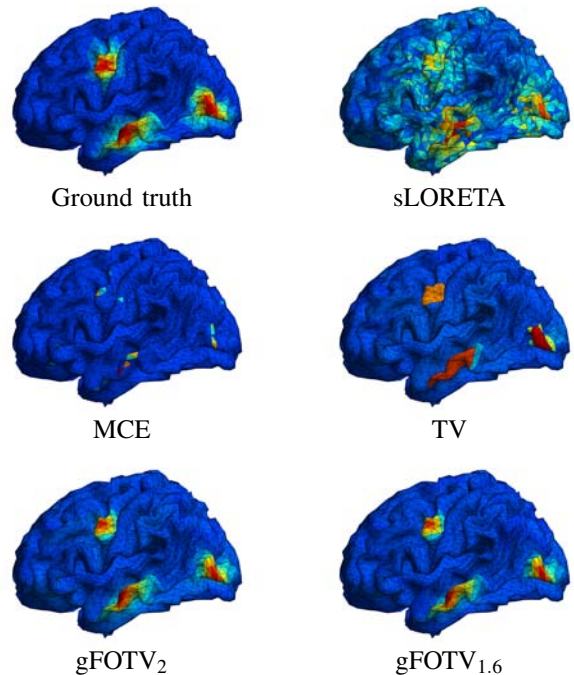


Fig. 1: Source localization results of various methods on synthetic data with three sources. The color scale gradation goes from blue being the minimum to red being the maximum.

### D. Simulation Results

First, we evaluate the performance of the proposed method using  $\alpha = 2$  and  $\alpha = 1.6$ , and compare with several state-of-the-art methods: sLORETA,  $\ell_1$ -norm based and TV-based methods. Fig.1 shows the reconstruction results of three sources at different locations of the cortex surface. One can see that the spatial resolution of sLORETA is very low, since it is based on the Tikhonov regularization using  $\ell_2$ -norm. The  $\ell_1$ -norm based method MCE greatly improves the spatial resolution because it imposes sparsity of the source in itself, but it yields over-focused reconstructed sources and fails to identify the spatial extent of the sources. The TV-based method shows good performance in identifying the spatial extent and preserving edges. However, it does not recover the varying intensity of the source, because of the assumption that the source intensity is piecewise constant. In contrast, the proposed method generates the reconstructed images closest to the ground truth. It not only provides high spatial resolution, but also successfully reconstructs the intensity variation and the spatial extent of the sources. It is worth noting that  $\alpha = 1.6$  is able to further enhance the spatial resolution compared to  $\alpha = 2$ .

We further evaluate the proposed method quantitatively by using various quantitative criteria defined above. To show the robustness and consistency, we conduct a large number of simulation tests with different source sizes, and average results of 50 realizations with different noise configurations, illustrated in Fig.2. In the TRE curve, one can see that the proposed method provides the smallest total reconstruction error. Notice that the TV-based method also

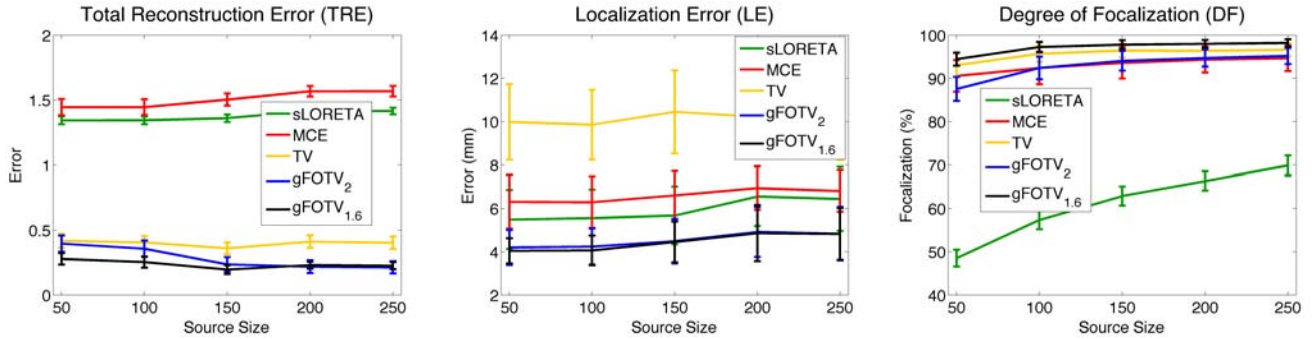


Fig. 2: Quantitative comparisons of various methods with different source sizes (average of 50 realizations of different noise configurations, where the error bar shows the sample standard deviation).

shows a relatively small total reconstruction error compared to other methods. In the LE curve, our proposed method shows the smallest localization error, since it makes the brain image smoother by imposing sparsity in higher order derivative instead of the first order derivative. As is well known that sLORETA can provide high localization accuracy, sLORETA also shows small localization error. TV-based method, however, produces relatively large localization error, which is because the reconstructed sources are constant in each subregion therefore it has difficulty pinpointing the peak of the source accurately. Finally, the DF curve shows that all of the methods with sparsity constraints show high focalization degree. It is interesting to notice that for gFOTV-based methods,  $\alpha = 1.6$  further enhances the spatial resolution. Compared to other methods, the spatial resolution of sLORETA is the lowest, as indicated by its name “standardized low resolution brain electromagnetic tomography”. To sum up, our proposed method demonstrates superior performance from all the criteria including total reconstruction accuracy, localization accuracy and spatial resolution.

## VI. CONCLUSIONS

In this paper, we propose an efficient and accurate EEG source reconstruction method by defining a novel graph fractional-order total variation (gFOTV) adapted for a triangular mesh of the cortical surface. This method imposes sparsity in  $\alpha$ -order spatial derivatives with  $\alpha \in [1, 2]$ , which includes the TV-based methods as a special case when  $\alpha = 1$ . By tuning the parameter  $\alpha$ , the proposed method provides the freedom to choose a more elegant order of smoothness for the underlying brain image. Therefore, it not only provides high spatial resolution, but also recovers the current density variation and localizes the source peaks with high accuracy. In addition, the proposed algorithm is parameter friendly in the sense that the parameter selection is not sensitive to source size and noise level. A variety of simulation experiments demonstrate that the proposed method consistently outperforms the state-of-the-art methods both qualitatively and quantitatively. In future work, our proposed framework could be extended by using spatially variant smoothness orders for different subregions in an automatic way to further

improve the reconstruction accuracy.

## ACKNOWLEDGMENT

The authors would like to thank Shuo Tan for help with implementation of the shortest path search algorithm.

## REFERENCES

- [1] M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa, “Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain,” *Reviews of modern Physics*, 65(2):413, 1993.
- [2] R. D. Pascual-Marqui, “Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details,” *Methods Find Exp Clin Pharmacol*, 24(Suppl D): 5–12, 2002.
- [3] K. Uutela, M. Hämäläinen, and E. Somersalo, “Visualization of magnetoencephalographic data using minimum current estimates,” *NeuroImage*, 10(2):173–180, 1999.
- [4] L. Ding, “Reconstructing cortical current density by exploring sparseness in the transform domain,” *Physics in Medicine and Biology*, 54(9):2683, 2009.
- [5] H. Becker, L. Albera, P. Comon, R. Gribonval, and I. Merlet. “Fast, variation-based methods for the analysis of extended brain sources,” In *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*, IEEE, 2014, pp. 41–45.
- [6] Z. Ren, C. He, and Q. Zhang, “Fractional order total variation regularization for image super-resolution,” *Signal Processing*, 93(9):2408–2421, 2013.
- [7] D. Chen, S. Sun, C. Zhang, Y. Chen, and D. Xue, “Fractional-order TVL2 model for image denoising,” *Central European Journal of Physics*, 11(10):1414–1422, 2013.
- [8] R. Oostenveld, D. F. Stegeman, P. Praamstra, and A. van Oosterom, “Brain symmetry and topographic analysis of lateralized event-related potentials,” *Clinical neurophysiology*, 114(7):1194–1202, 2003.
- [9] J. Bai and X. Feng, “Fractional-order anisotropic diffusion for image denoising,” *IEEE Trans. Image Process.*, 16(10):2492–2502, 2007.
- [10] K. B. Oldham and J. Spanier, “The Fractional Calculus,” Academic Press, New York, 1974.
- [11] O. R. Ryyanen, J. A. Hyttinen, P. H. Laarne, and J. A. Malmivuo, “Effect of electrode density and measurement noise on the spatial resolution of cortical potential distribution,” *Biomedical Engineering, IEEE Transactions on*, 51(9):1547–1554, Sept 2004.
- [12] H. Aurlien, I. O. Gjerde, J. H. Aarseth, G. Eldøen, B. Karlsen, H. Skeidsvoll, and N. E. Gilhus, “EEG background activity described by a large computerized database,” *Clinical Neurophysiology*, 115(3):665–673, 2004.
- [13] A. Molins, S. M. Stufflebeam, E. N. Brown, and M. S. Hämäläinen, “Quantification of the benefit from integrating MEG and EEG data in minimum  $\ell_2$ -norm estimation,” *Neuroimage*, 42(3):1069–1077, 2008.
- [14] C. H. Im, K. O. An, H. K. Jung, H. Kwon, and Y. H. Lee, “Assessment criteria for MEG/EEG cortical patch tests,” *Physics in medicine and biology*, 48(15):2561, 2003.