Source-Informed Segmentation: Towards Capturing the Dynamics of Brain Functional Networks through EEG

Ali Haddad and Laleh Najafizadeh
Department of Electrical and Computer Engineering, Rutgers University, NJ 08854, USA
Email: ali.haddad@rutgers.edu, laleh.najafizadeh@rutgers.edu

Abstract—In this paper, we propose a data-driven framework to objectively evaluate the dynamics of brain functional connectivity, during task execution, from electroencephalography (EEG) recordings. The proposed framework consists of two main steps: first, EEG recordings, in the sensor space, are segmented into intervals during which the spatial distribution of functionally connected networks, in the source space, stays quasi-stationary. This “source-informed segmentation” is achieved through statistically assessing the temporal variation in the dominant left singular subspace of the EEG recordings matrix. Second, recordings within each identified segment are transformed into the source space, where networks of functionally connected cortical points are localized based on their correlation to a common dominant right singular vector of the source data matrix. The proposed framework is capable of identifying multiple networks of functionally connected cortical points, per segment, with each network corresponding to an orthogonal time-course of estimated neuronal activity. Theoretical derivations are presented. The proposed framework is also used to examine the dynamics of networks of functionally connected cortical points during the execution of a visual oddball task, and results are discussed.

I. INTRODUCTION

Recent years have witnessed an increasing interest in studying the functionality of the brain within the context of “functional connectivity”, at multiple spatial and temporal scales, with the aim of advancing our understanding about the large scale functional organization of the brain [1]–[4]. In functional connectivity studies, statistical dependencies among signals recorded from remote regions of the brain are evaluated. Most of these studies, however, have been based on statistical measures that require the assumption of temporal stationarity (e.g. over the entire scan), and therefore, have ignored the temporal changes that occur in functional networks.

Electroencephalography (EEG) is a noninvasive brain imaging technique that measures the scalp electrical potentials related to neuronal activity. Compared to other non-invasive neuroimaging modalities, EEG offers excellent temporal resolution (in the range of milliseconds), and hence is a good candidate for examining the dynamics of brain function [5]. The majority of EEG-based dynamic functional connectivity studies have focused on the data that is obtained in the sensor space. Approaches such as using predefined segments based on having prior knowledge about the nature of the task [5], fixed-length sliding windows [6], frequency-dependent sliding windows [7], or based on graph metrics [8], [9] have been investigated to identify the dynamics of brain networks from EEG recordings. However, when evaluating the dynamics of brain connectivity using sensor-space data, volume conduction [10], [11] could increase the ambiguity of the correlation shown by the recordings obtained from different electrodes. To address this problem, spatial filtering is employed in [12] to enhance the spatial association of sensor-space data. In general, however, sensor-space dynamics represent a coarse picture of the underlying dynamics of functional networks that occur in the source space.

One solution to the problem of spatial ambiguity of EEG recordings is the use of source localization techniques to estimate the neuronal activity in the source space. However, due to the vastness of this space and the imposed computational complexity, most studies that are done in the source space have only considered a set of anchor points, e.g. corresponding to anatomically predefined cortical regions, for the analysis [13]. Such an approach can still suffer from the limitations imposed by making prior assumptions on the locations of functionally connected networks.

In this paper, we propose a novel data-driven framework to objectively examine the dynamic nature of neuronal activity, during the execution of tasks. The proposed framework starts in the temporal domain by segmenting EEG data (in the sensor space) into time intervals where the spatial distribution of networks of functionally connected cortical points in the source space stays quasi-stationary. For each segment, we then move to the spatial domain and determine the dominant time-courses in source-space activity to identify the locations of cortical points that are highly correlated to these time-courses.

The rest of the paper is organized as follows. The theoretical background of the proposed framework is presented in Section II. Experimental results are discussed in Section III, and the paper is concluded in Section IV.

II. THEORETICAL BACKGROUND

We tackle the problem of assessing the dynamics of functional connectivity in two main steps: the first step takes place in the temporal domain, where the recorded EEG data
in the sensor space is segmented into time intervals during which networks of functionally connected cortical points in the source space stay quasi-stationary, and, the second step takes place in the spatial domain, where for each identified segment, networks of functionally connected cortical points are localized. In what follows we present the theoretical background for each step, laying the foundation of our proposed framework.

A. Step 1: Source-Informed Segmentation of EEG Data

Here we tackle the problem of identifying time intervals in EEG data recorded in the “sensor space”, based on relevant events occurring in the underlying “source space”. We define a segment in the EEG data as an interval of time during which the spatial distribution of the underlying networks of functionally connected cortical points stays quasi-stationary [14], [15].

Taking a bottom-up approach, we begin in the sensor space, where the cortex is spatially discretized into $P$ cortical points. We model the temporal activity $s_k$ perceived at the cortical point with index $k$ ($k = 1, \ldots, P$), during the time interval $\{t_1, \ldots, t_T\}$ as $s_k = d_k \cdot \delta_{k,j} = d_k (1 + \varepsilon_{k,j})$, where $d_k$ represents the average activation intensity perceived at point $k$ during the interval, $\delta_{k,j}$ is a random variable denoting the scaled temporal activity at time $t_j$ ($j = 1, \ldots, T$) after factoring $d_k$ out, and $\varepsilon_{k,j}$ is the zero-mean component of the scaled temporal activity $\delta_{k,j}$. The activation intensities perceived at these points form the source matrix $S \in \mathbb{R}^{P \times T}$.

In the sensor space, during this interval, the EEG data matrix $Y \in \mathbb{R}^{C \times T}$, obtained through a set of $C$ electrodes positioned on the scalp, can be modeled as [16]

$$Y = G \cdot S + N,$$  \hspace{2cm} (1)

where, $G \in \mathbb{R}^{C \times P}$ represents the gain matrix, and $N \in \mathbb{R}^{C \times T}$ is the additive noise matrix.

In evaluating functionally connected networks, it should be noted that, here, we only deal with a discretized version of the source space. Therefore, a network of functionally connected cortical points can be defined as a set of $Q$ active cortical points such that the temporal activity of each point is highly correlated to a common time-course. We refer to $Q$ as the size of this network. Distinct cortical networks that are established during a given interval are theoretically characterized by their association to distinct time-courses. However, as it will be shown later, the information available in the sensor space can directly tie these networks to their locations in the cortex. The segmentation method we propose here use this information to detect the time points at which these networks are established/disestablished.

Assume that throughout the interval $\{t_1, \ldots, t_{T_{\text{seg}}}\}$ the temporal activity of each of the $Q_i$ cortical points $k^i$ ($k^i \in \{k^i_1, \ldots, k^i_{Q_i}\}$) stays highly correlated to the temporal activity of other points in this set. Then, these points form a network of functionally connected cortical points over this interval. Asymptotically, as the correlation throughout this interval increases, the temporal activity of each of these points can be expressed as

$$s_{k^i} \rightarrow d_{k^i} (1 + e_{k^i}^j) = d_{k^i} \cdot a_{k^i}^j,$$  \hspace{2cm} (2)

where $a_{k^i}^j$ is a random variable denoting the scaled temporal activity of all $Q_i$ functionally connected cortical points within the $i^{th}$ network at time $t_j$, for $j = 1, \ldots, T_{\text{seg}}$, and $e_{k^i}^j$ is the zero-mean component of the scaled temporal activity $a_{k^i}^j$.

Now, let $S_{\text{seg}} \in \mathbb{R}^{P \times T_{\text{seg}}}$ be the source matrix for the interval $\{t_1, \ldots, t_{T_{\text{seg}}}\}$ during which the spatial distribution of $R$ distinct networks of functionally connected cortical points, with sizes $Q_1, \ldots, Q_R \ll P$ points, stays quasi-stationary. $S_{\text{seg}}$ can be expressed as

$$S_{\text{seg}} = \begin{bmatrix} D_1 \cdot A_{1}^T \\ \vdots \\ D_R \cdot A_{R}^T \\ B_{\text{seg}} \end{bmatrix} = \begin{bmatrix} D_1 \cdot 1_1 \cdot A_{1}^T \\ \vdots \\ D_R \cdot 1_R \cdot A_{R}^T \\ B_{\text{seg}} \end{bmatrix},$$  \hspace{2cm} (3)

where $D_i \in \mathbb{R}^{Q_i \times Q_i}$ is a diagonal matrix with $d_{k^i}$ on its diagonal, $1_i$ is a $Q_i \times 1$ all-ones matrix, $B_{\text{seg}} \in \mathbb{R}^{(P-Q_1-\cdots-Q_R) \times T_{\text{seg}}}$ gives the background temporal activity for the cortical points outside any network of functional connectivity, and $A_i^T \in \mathbb{R}^{Q_i \times T_{\text{seg}}}$ has identical rows $a_{k^i}^j$ with the elements $a_{k^i}^j$ on each row, whereas $A_{\text{seg}}^T \in \mathbb{R}^{T_{\text{seg}} \times T_{\text{seg}}}$ is a diagonal matrix with $a_{k^i}^j$ on its diagonal.

The gain matrix $G$ can be then partitioned as

$$G = \begin{bmatrix} G_1 & \cdots & G_R \end{bmatrix} \begin{bmatrix} G_0 \end{bmatrix},$$  \hspace{2cm} (4)

where $G_i \in \mathbb{R}^{C \times Q_i}$ and $G_0 \in \mathbb{R}^{C \times (P-Q_1-\cdots-Q_R)}$. Thus, the sensor-space data matrix for this segment, $Y_{\text{seg}} \in \mathbb{R}^{C \times T_{\text{seg}}}$, can be expressed as

$$Y_{\text{seg}} = G \cdot S_{\text{seg}} + N_{\text{seg}},$$  \hspace{2cm} (5)

where $N_{\text{seg}} \in \mathbb{R}^{C \times T_{\text{seg}}}$ is the additive noise matrix for this segment. Substituting (3) into (5), $Y_{\text{seg}}$ can be written as

$$Y_{\text{seg}} = F_{\Sigma} \cdot A_{\Sigma}^T + F_{\text{bseg}} + N_{\text{seg}},$$  \hspace{2cm} (6)

where $A_{\Sigma}^T \in \mathbb{R}^{R \times T_{\text{seg}}}$ contains the distinct rows $a_{\Sigma}^j$; $F_{\text{bseg}} = (G_0 \cdot B_{\text{seg}}) \in \mathbb{R}^{C \times T_{\text{seg}}}$ corresponding to activities of cortical points that are outside any networks of functionally connected cortical points; and $F_{\Sigma} = [f_1 \cdots f_R] \in \mathbb{R}^{C \times R}$, with $f_i = (G_i \cdot D_i \cdot 1_i) \in \mathbb{R}^{C \times R}$ represents a linear combination of the gain vectors corresponding to the functionally connected cortical points within the $i^{th}$ network with $1_i$ being a $Q_i$-dimensional all-ones column vector.

Noting that during the considered segment interval the columns of $F_{\text{bseg}}$ span the space of gain vectors of all cortical points outside the networks of functionally connected cortical points, and can involve randomized linear combinations of these vectors, we can group $F_{\text{bseg}}$ with the additive noise term $N_{\text{seg}}$. The minimum mean-squared-error (MMSE) estimate of the term $F_{\Sigma} \cdot A_{\Sigma}^T$ can then be determined from the rank $R_{\Sigma}$ singular value decomposition (SVD) approximation of $Y_{\text{seg}}$, i.e.,

$$U_{\Sigma} \cdot \Delta_{\Sigma} \cdot V_{\Sigma}^T \rightarrow F_{\Sigma} \cdot A_{\Sigma}^T.$$  \hspace{2cm} (7)
The $i^{th}$ column of $F$, namely, $f_i$, can be viewed as a feature vector characterizing the $i^{th}$ network of functionally connected cortical points, as it is bound to the cortical points within the corresponding set, being a linear combination of their corresponding gain vectors. Additionally, this vector assigns higher weights to the gain vectors of the cortical points with higher average activation intensities.

A basis for the feature space spanned by the columns of $F$ can be obtained from the column space of $U_{R_C}$, the most dominant $R_C$-dimensional left singular subspace of $Y_{seg}$. A significant change in the span of this space, along the time axis, could be an indication of the change in the underlying spatial distribution of the established networks of functionally connected cortical points. As such, segment boundaries can be identified via statistical testing. An algorithm to accomplish this task has been previously introduced in [15]. Using this algorithm, we can locate segment boundaries, by statistically comparing the residual error resulting from projecting the block of the sensor-space data matrix under a reference window, and the block of the sensor-space data matrix under a sliding window, onto a feature subspace of the basis which is the dominant left singular vectors of the data under the reference window. Statistical testing is then performed using non-parametric Kolmogorov-Smirnov (K-S) test. To enhance the reliability of K-S test, the consecutive K-S decisions are aggregated under a given decision window.

B. Step 2: Functional Connectivity Analysis

Having addressed the temporal dimension of the dynamic functional connectivity problem, the next step is to address the spatial dimension, namely, to identify the spatial location of networks of functionally connected cortical points during each segment. We start by transforming the EEG data from the sensor space to the source space. An estimate, $S_{seg}$, of the source-space data matrix described in (3) will, then, become available. The component of the source data corresponding to the average intensity could dominate the next processing steps, forcing the algorithm to focus on localizing the highly activated cortical points rather than the functionally connected ones. From (2), we can see that the zero-mean temporal activity component can be utilized to search for functionally connected cortical points. Thus, first, the average activation intensity of each cortical point needs to be removed from the source matrix. An estimate of the zero-mean temporal activity of cortical points can be modeled as

$$S_{segAC} = \begin{bmatrix}
D_1 \cdot E_1^T + \Psi_1 \\
\vdots \\
D_R \cdot E_R^T + \Psi_R \\
B_{segAC}
\end{bmatrix}, \quad (8)$$

where $E_j^T \in \mathbb{R}^{Q_i \times T_{seg}}$ has identical rows $e_j^T$ with the elements $e_j^i$ on each row, and $\Psi_j \in \mathbb{R}^{Q_i \times T_{seg}}$ carries the additive estimation error for the temporal activity of the corresponding cortical points.

By examining (8), it appears that a large portion of the remaining cortical activation energy resides within the span of the time-courses $\{e_1^T, \ldots, e_R^T\}$. It is assumed, here, that these time-courses are minimally cross-correlated, as the existence of high correlation would otherwise mean that the networks associated with each are functionally connected. Low correlation gives the SVD less freedom to compress the bulk of the energy in these time-courses, into fewer dominant dimensions. Minimizing the energy of all other components in $S_{segAC}$, an estimate of these time-courses can be found from the $R$ most dominant right singular vectors of the estimated source matrix. A similar technique was suggested in [17]. An important difference between the two techniques is that, here, the temporal activity of cortical points is not normalized prior to performing the SVD, since we are not interested in emulating the calculation of Pearson correlation coefficients. Pre-normalization would assign equal weights to the activity of all cortical points, and thereby, boosting low background activity.

To estimate $R$, here, we search for the index $\hat{R}$ of the entry before the elbow on the curve of singular values. Only the right singular vectors corresponding to the most dominant singular values are considered.

Finally, the cortical points that their significant components of their activation energies are correlated to the estimated time-courses $\{\hat{e}_1^T, \ldots, \hat{e}_R^T\}$ are associated to the corresponding network of functionally connected cortical points. To build a significance test, we first transform the source-space data segment using fast Fourier transform (FFT), to examine the statistics of its different harmonics. A separate probability mass function (PMF) for the magnitude and phase content at each frequency can be estimated from the corresponding histograms. A large number of surrogate source-space segments are then generated in the frequency domain, with random components picked from the corresponding PMFs. The surrogate segments are transformed back using inverse fast Fourier transform (IFFT) and the most dominant $\hat{R}$ right singular vectors of each surrogate segment are then determined. A PMF for the energy components of the surrogate source-space temporal activity correlated to the $j^{th}$ right singular vector of the corresponding surrogate segments can then be estimated and used to determine an appropriate threshold for testing the significance of such components, provided some false positive ($\alpha$) rate.

It is important to note that due to the orthogonality of the estimated time-courses $\{\hat{e}_1^T, \ldots, \hat{e}_R^T\}$, the resulting identified functional networks would have minimum spatial overlap, since a cortical point will unlikely have significant components correlated to multiple time-courses.

III. EXPERIMENTAL RESULTS

A. EEG Data Acquisition and Preprocessing Steps

EEG data was collected from 2 healthy, right-handed male volunteers, using a 128-channel EEG system (Brain Products), at a rate of 500 sample/sec. Written informed consents approved by Rutgers IRB were obtained prior to experiments. The electrodes were positioned across the head based on the international 10–5 electrode placement system, with
the reference electrode positioned at location Cz. Volunteers performed a 3-stimuli modified visual odd-ball task [5], where one of the three served as the infrequent target stimulus (number of target stimuli = 100). The participants were asked to press a mouse button with their right index finger, when the infrequent target stimulus appeared on the screen.

The EEG data was preprocessed using EEGLAB [18]. First, the data was filtered using a band-pass finite impulse response (FIR) filter, with lower and higher cutoff frequencies of 1 and 50 Hz, respectively. Artifacts were then removed using independent component analysis (ICA). For each subject, bad epochs were removed, and the remaining target epochs were averaged to generate the event-related potential (ERP) signal. To transform data to the source space, the standardized low-resolution brain electromagnetic tomography (sLORETA) source localization algorithm provided by the Brainstorm toolbox [19] was used. The surface of the cortex was discretized into 15,002 uniformly distributed points.

B. Results

Figs. 1-a and 1-b show the 500 msec post-stimulus interval of ERP for subjects 1 and 2, respectively. The P300 component, a signature of the oddball task, can be easily identified in both ERPs. Segment boundaries, shown as dashed lines, are estimated using the source-informed segmentation algorithm. An initial reference window of length 20 samples (40 msec) and a decision window of length 10 samples (20 msec), were used for the segmentation algorithm [15]. As can be seen, the segmentation algorithm has identified 7 and 8 intervals on the ERPs of subject 1 and 2, respectively, possibly due to individual differences in processing the information.

Figs. 1-c and 1-d show the temporal evolution of functional networks in the source space, from the timing of stimulus presentation to 500 msec afterwards, for subjects 1, 2, and 2, respectively. The temporal orders for functional networks are the same as the segment orders shown in Figs. 1-a and 1-b. The false positive rate was chosen as $\alpha = 0.01$. For each segment, the functional connectivity analysis framework only detected one or, less frequently, two distinct functional networks.

Following the progression of cortical maps in Figs. 1-c and 1-d, it can be seen that, early after the onset of the stimulus, functional networks are being formed in the left inferior temporal region, and to a lesser degree in the right inferior temporal region. Soon after, during the interval $\sim [60, 200\) msec, functionally connected bilateral regions become identifiable over the occipital region indicating that the brain is processing the visual information. Examining brain functional networks during the segment coinciding with the P300 component, the sixth segment for both ERPs, shows that functional networks have moved to the parietal region. These results are in line with previous studies [5], [20], [21].

IV. Conclusions

In this paper, we proposed a novel framework for examining the dynamics of brain functional connectivity from EEG recordings. The framework starts by segmenting sensor-space EEG data into time intervals where the spatial distribution of networks of functionally connected cortical points in the source space stays quasi-stationary. We demonstrated that these intervals can be identified by dominant left singular subspace, and utilized this subspace to perform the source-informed segmentation. The data segments were then transformed into the source space to localize networks of functionally connected cortical points. It was shown that the dominant time-courses of cortical activity during a segment can be estimated from the dominant right singular vectors of source data. Hence, the cortical points that their temporal activity show high linear dependence to a common dominant time-course can be considered as functionally connected. The proposed framework was then used to identify the dynamics of functional connectivity during the execution of a modified odd-ball task. The results met the nature of the examined task, verifying the capabilities of the proposed framework in identifying the dynamics of brain function during task execution.

References

Fig. 1: Dynamic functional connectivity analysis. a) and b) ERPs of subjects 1 and 2; c) and d) brain views ordered from top to bottom rows as: superior, inferior, anterior, and posterior. Temporal sequencing of functional networks is shown from left to right. Cortical regions colored blue or green belong to distinct networks, as ordered by the significance of the corresponding time-courses in the SVD. Cortical regions colored red belong to both networks.


