MODELLING BRAIN CORTICAL CONNECTIVITY USING DIFFUSION ADAPTATION

Konstantinos Eftaxias⋆*, Saeid Sanei*⋆ *and Ali H. Sayed*†

⋆ Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, UK
† Department of Electrical Engineering, University of California, Los Angeles, USA

ABSTRACT

This work examines the flow of information among electrodes attached to the brain and uses diffusion adaptation strategies to assess brain cortical connectivity. The method uses the directed transfer function (DTF) technique to estimate combination coefficients to drive the adaptation and learning process. The diffusion strategy is then applied to the problem of recognizing left and right hand movements and its superior performance is demonstrated relative to solutions that rely on stand-alone electrodes and do not exploit coordination among multiple electrodes.

***Index Terms*—** Diffusion adaptation, brain connectivity, directed transfer function.

1. INTRODUCTION

There is a growing interest in using network models to explore brain connectivity with the intent of elucidating the anatomical and functional organization of the brain during specific tasks. In this paper, we examine the flow of information among electrodes attached to the scalp and use diffusion adaptation strategies to establish a model for performing a movement-related task. We employ the DTF technique to estimate combination coefficients that drive the adaptation and learning process. The diffusion strategy exploits the space-time characteristics of the measured signals more fully than non-cooperative models and leads to an enhanced model for recognising hand movements [1].

It is well-known that particular connectivity patterns between neurons in the brain are reflective of mental, cognitive, and movement activities. The patterns vary both in time and space and originate from distributed synaptic current sources. Modelling such connectivity patterns is therefore of great significance. One of the earliest measures of brain connectivity is the Pearson product correlation measure, also called "co-modulation" or Lexicor spectral correlation coefficient [2]. The coefficient is used to estimate the degree of association between amplitudes or magnitudes of the electroencephalography (EEG) signals, acquired by electrodes attached to the scalp, over intervals of time and frequency.

In some applications, such as the detection and classification of finger movements, it is important to find out how the electric signals propagate within the neural network of the brain. In these cases, there is a consistent movement of the source signals from the occipital to temporal regions. References [3][4][5] proposed a useful multivariate auto-regressive (MVAR) model to explain how the directionality of the cortical signal patterns changes within the brain. It is also clear that during mental tasks, different regions within the brain communicate with each other. The interactions and cross-talks among the EEG channels offer valuable clues towards understanding the processing at the brain during various tasks. For this purpose, it is important to recognize the transient periods of synchrony between various regions in the brain. These phenomena are not easy to observe by visual inspection of the EEGs. In some approaches, the connectivity, coherency, and synchronization of the brain regions are evaluated by examining the spatial statistics of scalp EEG signals using coherence measures.

Spectral coherence [6] is one common method for determining synchrony in EEG activity. Coherency is a normalized form of cross-spectrum. However, it does not provide information on the directionality of the coupling between recording sites. Granger causality (also called Wiener-Granger causality) [7] is another useful measure that attempts to extract and quantify the directionality from EEGs. Granger causality is based on bivariate AR estimates from the data. Based on Granger causality, if the past samples of a time series $y(i)$ can be used to predict another series $x(i)$, then $y(i)$ is said to cause $x(i)$.

Nevertheless, application of the Granger causality measure to multivariate data arising from multichannel recording is not computationally efficient. The directed transfer function (DTF) concept [3][4][5] is an extension of Granger causality and it can be used to detect and quantify the coupling directions. The advantage of DTF over spectral coherence is that it can determine the directionality in the coupling when the frequency spectra of the two brain regions have overlapping spectra [8]. A time-varying DTF can also be generated to track source signals by calculating the DTF over short windows to achieve the short time DTF (SDTF).

The work of A. H. Sayed was supported in part by NSF grant CCF-1011918.

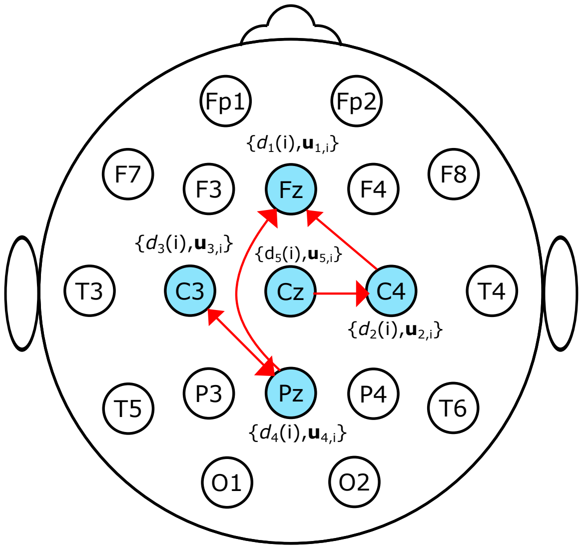

Fig. 1. An illustration of a brain connectivity pattern. EEG signals collected at the marked electrodes are used to train a cooperative network to estimate a model for right or left hand movements.

This SDTF technique plays an important role in the classification of left and right hand movements and in tracking the related sources of the brain signals. Using phase coherence, the delay between the onsets of similar frequency components can be estimated and it reveals the direction of propagation.

Most available techniques for evaluating connectivity patterns assume the underlying signals to be stationary. They also consider the sources separately. Network adaptation and diffusion strategies, on the other hand, allow us to exploit more fully the temporal and spatial characteristics of the brain signals. This approach models the sources from a number of neurons as agents that communicate with each other in space and evolve over time. One useful advantage of this approach is that it can cope with nonstationary data.

Notation: Bold uppercase letters denote matrices and bold lowercase letters denote column vectors.

2. DIFFUSION-BASED PROCESSING MODEL

In our proposed method towards studying brain connectivity, we couple the DTF technique with a diffusion adaptation strategy in order to enable a robust modelling of a motor task. The details of the approach are described in the sequel.

2.1. Directed Transfer Function

To define the DTF, we consider an $N-$channel multivariate process represented by the vector:

$$\mathbf{x}_i = \begin{bmatrix} u_1(i) & u_2(i) & u_3(i) & \cdots & u_N(i) \end{bmatrix} \tag{1}$$

The entries of this vector represent the EEG signals that are collected at N electrodes attached to the scalp. The vector signal is assumed to satisfy a multivariate auto-regressive model (MVAR) of the form:

$$\mathbf{x}_i = \sum_{m=1}^{L} \mathbf{A}_m \mathbf{x}_{i-m} + \mathbf{e}_i \tag{2}$$

where L denotes the model order, $\mathbf{e}_i$ is the error term, and $\mathbf{A}_m$ are $N \times N$ matrix coefficients. Let $\mathbf{H}(e^{j\omega})$ denote the frequency response of the system mapping $\mathbf{e}_i$ to $\mathbf{x}_i$, namely,

$$\mathbf{H}(e^{j\omega}) = \left[\mathbf{I}_N - \sum_{m=1}^{L} \mathbf{A}_m e^{-j\omega m} \right]^{-1} \tag{3}$$

This frequency response contains useful information about the relations between the electrode channels. We define the DTF coefficient of indices (ℓ, k) as the following measure of the causal influence of channel ℓ on channel k [9]:

$$\gamma_{\ell k}^2(e^{j\omega}) = \frac{|H_{\ell k}(e^{j\omega})|^2}{\sum_{q=1}^{N} |H_{qk}(e^{j\omega})|^2} \tag{4}$$

2.2. Diffusion Least-Mean-Squares

In diffusion adaptation, the goal is to estimate an $M \times 1$ unknown vector $\mathbf{w}^o$ in a distributed manner from measurements collected at N nodes spread over a network. Each node k has access to time realisations $\{d_k(i), \mathbf{u}_{k,i}\}$ of data that are assumed to be related via a linear regression model

$$d_k(i) = \mathbf{u}_{k,i} \mathbf{w}^o + v_k(i) \tag{5}$$

where $v_k(i)$ represents measurement noise, assumed to be temporally white and independent over space. There are several diffusion strategies that can be used for the estimation of $\mathbf{w}^o$ [10] [11]. The so-called Adapt-then-Combine (ATC) strategy takes the following form:

$$\boldsymbol{\psi}_{k,i} = \mathbf{w}_{k,i-1} + \mu_k \mathbf{u}_{k,i}^* (d_k(i) - \mathbf{u}_{k,i} \mathbf{w}_{k,i-1}) \tag{6}$$

$$\mathbf{w}_{k,i} = \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \boldsymbol{\psi}_{\ell,i} \tag{7}$$

where $\mathbf{w}_{k,i}$ denotes the estimate for $\mathbf{w}^o$ that is computed by node k at time i, and the $\{a_{\ell k}\}$ are nonnegative coefficients that satisfy

$$\sum_{\ell \in \mathcal{N}_k} a_{\ell k} = 1, \qquad a_{\ell k} = 0 \text{ if } \ell \notin \mathcal{N}_k \tag{8}$$

Moreover, $\mathcal{N}_k$ denotes the set of neighbors of node k. The ATC implementation involves two operations. In the first step, each node updates its intermediate estimate $\mathbf{w}_{k,i-1}$ to $\boldsymbol{\psi}_{k,i}$ by using its local data. And in the second step, the node aggregates the intermediate estimates of its neighbors to obtain $\mathbf{w}_{k,i}$.

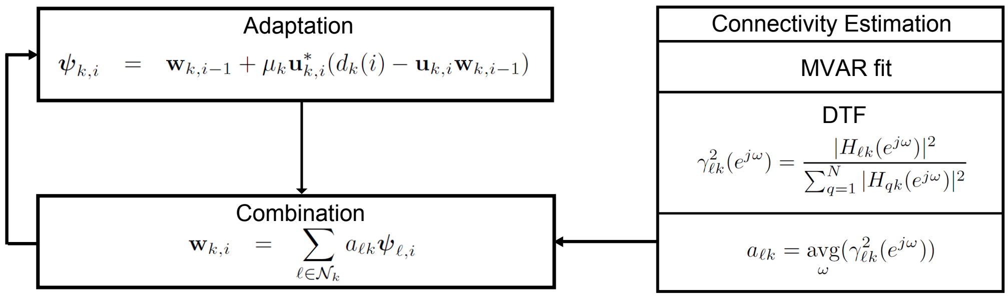

Fig. 2. Outline of the proposed method combining DTF and diffusion adaptation (avg refers to averaging operation).

2.3. Combining DTF and Diffusion

Suppose we acquire N EEG signals $\{u_k(i)\}$ over time $i \geq 0$ from N electrodes attached to the scalp. These signals are the result of a certain movement task by the individual, such as moving the right hand. For each electrode k, the EEG signal is used to construct the regression vector of size M as follows:

$$\mathbf{u}_{k,i} = \begin{bmatrix} u_k(i) & u_k(i-1) & \cdots & u_k(i-M+1) \end{bmatrix} \quad (9)$$

Each electrode location is treated as corresponding to the location of a node k in a diffusion network. All nodes are presented with reference signals, $d_k(i)$ for $k = 1, 2, \ldots, N$. These signals are constructed as follows. They are obtained by averaging different epochs of EEG signals that correspond to the same task. Therefore, the signals $d_k(i)$ are representative of the physiological characteristics that are specific to the motor task under examination. We can then apply diffusion adaptation on the data $\{d_k(i), \mathbf{u}_{k,i}\}$ in order to estimate the model $\mathbf{w}^o$ that relates the data. Once $\mathbf{w}^o$ is estimated, it can be subsequently used to classify whether a given EEG signal belongs to one class (right-hand movement) or another (left-hand movement). However, in order to apply the diffusion strategy, we need first to determine the connectivity (i.e.,the topology) pattern that is supposed to represent the interactions among the electrode locations.

For this purpose we distinguish the two steps of the diffusion LMS algorithm as the adaptation (6) and combination steps (7). Before performing the combination step, we use the EEG signals $\{u_k(i)\}$ to construct the vector process $\mathbf{x}_i$ defined by (1) and fit a MVAR model into the data to estimate the coefficients $\{\mathbf{A}_m\}$. This step can be accomplished, for example, by using the ARFit toolbox [12]; it can also be accomplished in an adaptive manner and we leave this extension for future consideration. Using the estimated $\{\mathbf{A}_m\}$, we can estimate the DTF coefficients $\{\gamma^2_{\ell k}(e^{j\omega})\}$ over a grid within some desired frequency band, especially since different motor actions tend to lead to more pronounced DTF coefficients over particular frequency bands. We subsequently average the DTF values $\gamma^2_{\ell k}(e^{j\omega})$ over the specific frequency band to obtain coefficient values that are representative of the cortical connectivity for the particular motor action. These averaged values are used as the combination coefficients $\{a_{\ell k}\}$ for the diffusion update. In general, the averaged DTF coefficients $\{\gamma^2_{\ell k}\}$ are sparse over the domain $\ell \times k$ and, therefore, the coefficients $\{a_{\ell k}\}$ will also be sparse. Specifically since the values of the DTF coefficients between most electrodes are close to zero, we set a threshold and discard values below the threshold. In this way, we end up defining the neighborhood $\mathcal{N}_k$ for each node k. After we normalise these values to ensure that they satisfy the first constraint of equation (8), we run the diffusion steps (6)-(7).

The model $\mathbf{w}^o$ that results from this procedure is based on incorporating information about brain connectivity and on exploiting the temporal and spatial features of the EEG signals more fully. Using this model to test other subjects or tasks can lead to useful insights about the connectivity patterns of the brain and the changes from task to task or between patient subjects.

3. EXPERIMENTS

We use a simulation scheme similar to the one used to validate brain connectivity measures in [9]. We use the BCI competition II dataset (Dataset III) [13], which consists of 140 training trials of 2 classes (imagery left and right hand movements). We isolate one random trial from each class (s_1 and s_2) and generate five signals from each by delaying them by five different delays (see Fig. 3). This way we can simulate the propagation of two signals from different starting points. We add Gaussian noise to introduce perturbation to the system. The outcome of this procedure is the generation of our experimental dataset. We proceed by using the diffusion LMS algorithm to estimate the filter coefficients $\mathbf{w}_{k,i}$. In this experiment, all nodes are presented with the same reference signal $d_k(i) = d(i)$ for $k = 1, 2, \ldots, N$. Therefore, we use the signal s_1 as the global reference signal $d(i)$ and use the measured signals generated by s_1 to construct the regression data $\{\mathbf{u}_{k,i}\}$ after adding some relatively low noise. Using the proposed method we are able to estimate the filter coefficients that describe our connectivity-enhanced model. The number of iterations and the parameter μ are fixed and the step-size is

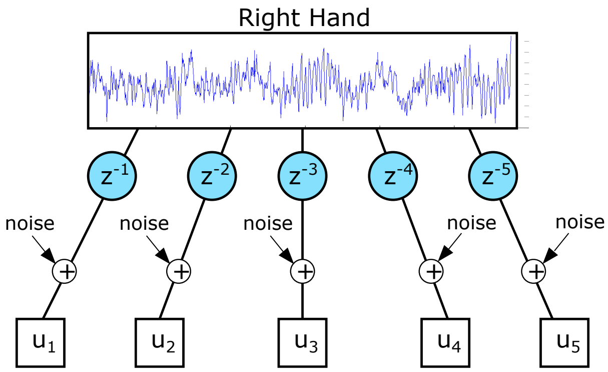

Fig. 3. Dataset generation; we delay and add noise to each signal in order to simulate the signal propagation across the nodes. We repeat the procedure for the left hand.

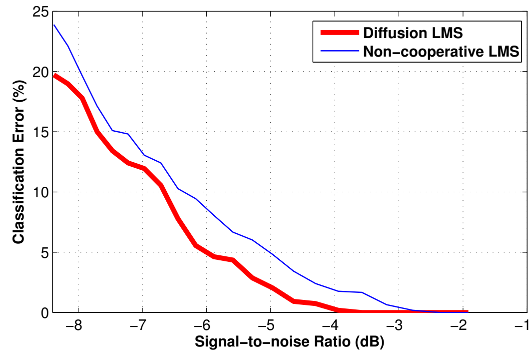

Fig. 4. Percentage of classification error versus signal-to-noise ratio for both diffusion and non-cooperation methods.

sufficiently small to ensure convergence of the filter.

Using the estimated model $\mathbf{w}^o$ that results from the diffusion step after training, we apply it to classify other data signals into right-hand and left-hand movements. We compare the performance of the diffusion implementation with a non-cooperative solution in Fig. 4. We simply classify the signals by only using the value of the mean square error and we calculate the percentage of the incorrectly classified signals. The results in the figure confirm the superior performance of the diffusion solution.

4. CONCLUSION

In this paper we introduced a modelling approach for describing a motor task using brain connectivity. The model is a realistic simulation of synaptic flow within the brain as the result of hand (or general body) movement. The approach benefits from the DTF connectivity measure to drive the learning process. The resulting model can be used for classification of left-right hand movement and therefore provides a useful direction in BCI and in places where movement-related connectivity changes occur as in Parkinson's patients.

5. REFERENCES

[1] S. Sanei, *Adaptive Processing of Brain Signals*, Wiley, NY, 2013.

[2] M. B. Sterman and D. Kaiser, "Comodulation: A new QEEG analysis metric for assessment of structural and functional disorders of the central nervous system," *Journal of Neurotherapy*, vol. 4, no. 3, pp. 73–83, 2001.

[3] M. J. Kaminski and K. J. Blinowska, "A new method of the description of the information flow in the brain structures," *Biological Cybernetics*, vol. 65, no. 3, pp. 203–210, 1991.

[4] M. Kaminski, K. J. Blinowska, and W. Szelenberger, "Topographic analysis of coherence and propagation of EEG activity during sleep and wakefulness," *Electroencephalography and Clinical Neurophysiology*, vol. 102, no. 3, pp. 216–227, 1997.

[5] A. Korzeniewska, S. Kasicki, M. Kamiński, and K. J. Blinowska, "Information flow between hippocampus and related structures during various types of rat's behavior," *Journal of Neuroscience Methods*, vol. 73, no. 1, pp. 49–60, 1997.

[6] C. Gerloff, J. Richard, J. Hadley, and A. E. Schulman, "Functional coupling and regional activation of human cortical motor areas during simple, internally paced and externally paced finger movements.," *Brain*, vol. 121, no. 8, pp. 1513–1531, 1998.

[7] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, vol. 37, no. 3, pp. pp. 424–438, 1969.

[8] J. Ginter Jr, K. J. Blinowska, M. Kamiński, and P. J. Durka, "Phase and amplitude analysis in time-frequency space–application to voluntary finger movement.," *Journal of Neuroscience Methods*, vol. 110, no. 1-2, pp. 113, 2001.

[9] K. J. Blinowska, "Review of the methods of determination of directed connectivity from multichannel data," *Medical and Biological Engineering and Computing*, vol. 49, no. 5, pp. 521–529, 2011.

[10] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE Trans. on Signal Processing*, vol. 56, no. 7, pp. 3122 –3136, July 2008.

[11] F. S. Cattivelli and A. H. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Trans. on Signal Processing*, vol. 58, no. 3, pp. 1035 –1048, March 2010.

[12] A. Neumaier and T. Schneider, "Estimation of parameters and eigenmodes of multivariate autoregressive models," *ACM Trans. on Mathematical Software (TOMS)*, vol. 27, no. 1, pp. 27–57, 2001.

[13] G. Pfurtscheller and A. Schlögl, "BCI Competition II dataset," Downloaded from http://www.bbci.de/competition/ii/.