

Ultra-low Noise EEG at LSBB: Effective Connectivity Analysis

Nazanin Hamzei¹, John Steeves², John (Kip) Kramer², Matt Yedlin¹, and Guy A. Dumont^{1*}

¹Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

²International Collaboration On Repair Discoveries (ICORD), Blusson Spinal Cord Centre, 818 West 10th Avenue, Vancouver, BC, V5Z 1M9, Canada

Abstract. In this study, we further investigate electroencephalographic (EEG) data recorded during October 2014 in the ultra-shielded capsule at LSBB, with a focus on the study of task-specific Granger-causal effective connectivity patterns. In previous studies, we showed that noise-free EEG signals acquired in LSBB are suitable for analysis of activity patterns in high frequency bands, i.e. 30 Hz and above. We previously demonstrated that increases in task/rest gamma band (30-70 Hz) energy ratios during ankle and wrist movements are more prominent in the LSBB capsule than in an above-ground hospital environment. The present study extends previous analyses by examining gamma-band connectivity, i.e. the functional patterns of interaction between 64 channels of EEG within the gamma band during motor tasks. We use parameters from a MultiVariate Auto-Regressive (MVAR) model to estimate effective connectivity in 10-second batches of EEG and report the average patterns across all batches in which subjects repetitively move their ankle/wrist. We report the gamma-band connectivity results in a reduced form as strength of hemispheric and inter-regional connections. The analysis reveals that for some subjects, significant channel-wise connections in the LSBB capsule outnumber those in the hospital, suggesting that patterns of gamma-band connectivity are better reflected in low-noise environments. This study again demonstrates the potential of the ultra-shielded capsule and motivates further protocol enhancements and analysis methods for conducting future high-frequency EEG studies within LSBB. .

1 Introduction

In two previous studies [1], [2] we have demonstrated the potential of the ultra-low noise shielded capsule at the heart of LSBB is an ideal environment for acquiring ultra-high, wide-band electroencephalographic (EEG) signals without the need of filtering. The EEG has proven to be a useful information source in the analysis of brain activity, neural plasticity, and the diagnosis of various neurological disorders, as well as the development of intelligent brain-computer interfaces (BCIs). The EEG can be defined as a representation of neuronal activity (e.g. summated extracellular action potential frequencies) within different brain regions and is usually acquired non-invasively from the surface of the scalp. To reduce the

*Corresponding author: guyd@ece.ubc.ca

disruptive effects of electrical interference and movement (i.e. noise), depending on the neuronal activity of interest, the EEG is often band-pass filtered at different frequencies from 0.5 to 100 Hz and is then studied in terms of activity across different frequency bands; including δ (0-4 Hz), θ (4-7 Hz), α (8-12 Hz), β (12-30 Hz), and γ (30-100 Hz). Clinically, EEG frequencies are most often studied from 0.5 to 30 Hz with the gamma (γ) band being ignored. The gamma band is often ignored due to the use of filters to suppress unwanted electromagnetic noise from powered (50 or 60 Hz) electrical outlets. However, in recent years, the gamma band has received greater scrutiny within the research community. In [1], and in particular in [2] the use of time-frequency analysis revealed that the task-induced increase in gamma band (>30 Hz) energy relative to the resting state EEG is more prominent in signals acquired at LSBB than in a typical hospital environment, suggesting that task-specific changes in EEG are better reflected and more readily detected in signals acquired at LSBB. In this report we go further by analyzing the connectivity in the gamma band of EEG acquired during motor tasks.

1.1 Motivation for Connectivity Analysis

For the last 20 years or so, the study of *functional integration* of the brain; i.e. viewing the brain as an interconnected system in which the interplay among different parts acts as a crucial element in neural operation and formation of coherent cognitive and behavioral states [3] has been at the heart of neuroscience. The recognition that while the neuron is a separable entity in itself, its operation largely depends on the input gathered from other neurons has given birth to the term *Connectome*.

Brain connectivity patterns are dynamic in nature; links among neuronal assemblies may form or disappear in milliseconds [4], allowing for fast and transient information transfer among brain regions. Therefore, it is important to study connectivity in a framework that encompasses both time and space simultaneously. While fMRI has been the modality of choice for studying brain function, it suffers from relatively low temporal resolution. Moreover, recent research suggests that typical assumptions and statistical methods used in fMRI introduce a high number of false positives and lead to erroneous conclusions [5]. On the other hand, low-cost, non-invasive scalp EEG offers substantially higher temporal and spectral resolution. With the aid of proper computational data analysis tools and an acquisition system comprising of densely distributed electrodes, EEG is therefore an attractive candidate for studying rapidly changing spatiotemporal interactions among brain regions.

1.2 Different Categories of Connectivity

Connectivity patterns of the brain can be studied from several perspectives:

Anatomical or *structural* connectivity aims at describing the biophysical (i.e. axonal) communication links between neuronal assemblies. One example of the methodologies in this field is the neuroanatomical tract tracing [6], an invasive technique to provide information about direct axonal connections *in vivo*. Anatomical connectivity can also be studied non-invasively, though with lower spatial resolution, by means of diffusion weighted imaging techniques such as diffusion tensor MRI (DTI) [7].

The goal of *functional* connectivity, on the other hand, is not to understand the physical capability of axonal links. Rather, functional connectivity is concerned with finding evidences of statistical dependence among large-scale neuronal units, regardless of the presence of direct anatomical connections. The dependence among brain regions is typically examined by means of cross-correlation, spectral coherence, or mutual information, and therefore has

no notion of direction or causation. Primary tools for the analysis of functional connectivity, including fMRI and Positron Emission Tomography (PET), have shown that functional connectivity is related to behavior in a variety of different tasks [8].

In addition to functional connectivity, *effective* connectivity attempts to understand causation among neural units, i.e. the influence that one neural system has over another. In comparison with functional connectivity, effective connectivity provides additional information about the *direction* of interactions as well as their presence. Current techniques for determining effective connectivity include Granger-causal modeling, dynamic causal modeling, structural equation modeling, and transfer entropy, applied to neuroimaging data such as fMRI as well as EEG/MEG time series [9]. Methodologies implemented in this paper typically fall under the category of task-relevant effective connectivity. It is worth mentioning that there is not always a sharp distinction between the terminology of effective and functional connectivity, as sometimes directed connectivity relevant to a particular function might be called functional connectivity.

1.3 Modeling And Estimation of Connectivity

While most of methods for quantifying connectivity rely completely on the data to infer connectivity patterns (data-driven approaches), some techniques rely on prior assumptions on the connectivity structure (model-based approaches). Dynamic causal modeling (DCM) [10] is a popular example of a model-based technique which relies on comparing different models of connectivity based on the relative evidence for one model compared to another [11]. Structural Equation Modeling (SEM) is another approach based on explaining the observed covariance among several variables by a defined anatomical network [12]. On the other hand, correlation-based methods and information-theoretic techniques such as transfer entropy [13] and mutual information [14] exemplify methods that do not depend on prior assumptions on the model and try to find connectivity structure solely based on the data.

From a different perspective, the connectivity estimation metrics can be classified into either linear or nonlinear measures. Typically, most of the methods in the literature are based on linear assumptions; while information-theoretic approaches or methods such as the imaginary part of coherency [15], do not assume linear relations among variables. Linear methods are generally easy to implement and are sufficient for detecting interactions such as coupling of oscillations at similar frequencies, while nonlinear methods are useful if we are interested in nonlinear forms of coupling, such as cross-frequency couplings [16]. Moreover, while it may seem counter-intuitive to apply linear methods to problems of highly nonlinear nature such as EEG, it has been shown that many biomedical signals can be sufficiently characterized and analyzed by means of linear methods [17].

Connectivity metrics may also be classified into bivariate or multivariate measures. Bivariate measures find patterns of interaction among a multiplicity of signals by calculating pair-wise connectivity separately for each channel pair. On the other hand, for computing multivariate measures all channels are taken into account at once by fitting a full multivariate model. Most of the nonlinear connectivity measures such as mutual information and phase synchronization are bivariate [18]. In the case of densely inter-connected networks, multivariate measures are strongly preferred since bivariate measures may lead to misleading and spurious connectivity patterns [18].

Through the rest of this paper, we focus on a model-based, linear, and multivariate method based on Granger causality [19]. Originated from the field of economic time series, Granger causality describes a framework for quantifying the influence of signal $A(t)$ on another signal $B(t)$ by the ability of A to predict subsequent instances of B . Due to their simplicity, interpretability, and the minimal prior assumptions posed on the data, Granger-causal methods

have been extensively used in biomedical data analysis [20]. Their use has been extended to more than two signals by means of MultiVariate AutoRegressive (MVAR) modeling [21]. Moreover, many of these methods operate in the frequency domain and have proper adaptive variants to deal with nonstationarity, making them good candidates for analysis of connectivity within specific frequency bands of a nonstationary multivariate process such as EEG.

Until recently, analysis of connectivity was based on the implicit assumption that the pattern of interactions among brain regions remains fairly constant over the course of performing a task or the period of data collection [8]. This assumption has led to an extensive number of studies contributing to our understanding of large-scale interactions. However, results of these studies typically represent the aggregate or average connectivity patterns 'smeared' across time. In cases such as studying the formation and propagation of epileptic seizures, however, analysis of changes in dynamic connectivity over smaller, near instantaneous time scales will lead to greater insights into the fundamentals of brain networks [22]. In this paper, we analyzed EEG data both from a static and dynamic perspective, with the results of the former and latter approaches presented in sections 5.1 and 5.2, respectively.

This paper is organized as follows. In section 2 we describe the theory behind MVAR modeling and Granger-causal methods and introduce a number of connectivity measures based on Granger causality. Section 4 discusses the practical issues faced when these theories are to be implemented on real EEG data, as well as our methods and approaches in solving these issues. Once the theory and methods are introduced, the results of applying the static measures on our data are presented in section 5.1. In section 5.2, we extend our use of these measures to the time-varying case in order to incorporate time evolutions of the connectivity patterns.

2 Connectivity Analysis

2.1 Granger Causality

Granger's original definition of causality [19] is based on the fact that causes precede their effects in time, and that knowledge of the cause aids in predicting the effect. Assume that we are interested in predicting the value of a time series x_1 at time t based on a linear combination of its p previous values:

$$x_1(t) = \sum_{i=1}^p A_{11}(i)x_1(t-i) + \epsilon(t) \quad (1)$$

where $\epsilon(t)$ is the prediction error. If, in predicting the current value of $x_1(t)$, we incorporate also q previous values of another signal $x_2(t)$, we will attain a different prediction error $\epsilon'(t)$:

$$x_1(t) = \sum_{i=1}^p A'_{11}(i)x_1(t-i) + \sum_{j=0}^{q-1} A_{12}(j)x_2(t-j) + \epsilon'(t) \quad (2)$$

In this context, x_2 is said to Granger-cause x_1 if it can be shown that ϵ' is an improvement over ϵ . This improvement in the prediction error needs to be assessed in a statistical sense, e.g. by performing an F-test on the variances of $\epsilon(t)$ and $\epsilon'(t)$, given assumptions of covariance stationarity on $x_1(t)$ and $x_2(t)$.

2.2 MultiVariate AutoRegressive (MVAR) Models

Granger-causal relations between signals $x_1(t)$ and $x_2(t)$ in equation 2 can also be inferred through parameters $A'_{11}(i)$, $i = 1, \dots, p$ and $A_{12}(j)$, $j = 0, \dots, q-1$. These parameters, along

with a similar set of parameters relating $x_2(t)$ with past values of itself and $x_1(t)$, comprise an autoregressive (AR) model. Autoregressive modeling is a simple yet effective approach for characterization of time series and their spectra. The order of the model (p, q , etc.) is the number of preceding observations included in the model and depends on the dynamics of the signal. The coefficients A_{ij} are essentially features carrying information about the behavior of the time series. In the context of Granger causality, these features represent the amount by which past values of a signal aid in prediction of the current values of another signal (hence an implicit notion of causation).

Multivariate autoregressive (MVAR) models extend this approach to more than two time series by predicting each of the signals based on the previous values of all other signals. Specifically, let \mathbf{X} denote a K -dimensional stochastic multivariate process of length T . In our case, \mathbf{X} corresponds to the set of $K = 60$ channels of EEG recorded over T time points. A value of this process at time instant t is the K -dimensional data vector $X(t) = (X_1(t), X_2(t), \dots, X_k(t))'$. This vector can be estimated as a regression on its p previous values (a vector autoregressive process) as:

$$X(t) = \sum_{m=1}^p \mathbf{A}_m X(t-m) + E(t) \quad (3)$$

Here, \mathbf{A}_m 's are $K \times K$ model coefficient matrices where $\mathbf{A}_m(i, j)$ represents the effect (weight) of sub-process X_j on X_i at lag m ; and $E(t)$ is a K -dimensional zero-mean white noise process with a non-singular covariance matrix Σ . We have assumed, without loss of generality, that $X_1(t), X_2(t), \dots, X_k(t)$, $k = 1, \dots, K$ are zero-mean sub-processes and that the same model order p is required to regress on all signals.

Relating equations 2 and 3, extension of Granger causality to more than two variables thus involves fitting an MVAR model to the data. In this context, a time series $X_i(t)$ is called a Granger cause of the time series $X_j(t)$ if at least at one lag m , $m = 1, \dots, p$, the corresponding element of the coefficient matrix $\mathbf{A}_m(i, j)$ is significantly greater than zero in absolute value sense [23], [24].

2.3 Autoregressive Modeling of Nonstationary Data

In practice, autoregressive models are typically restricted to stationary time series so that an accurate model fit can be realized. However, many biomedical signals, specially the EEG, are highly nonstationary in nature. There are two approaches for modeling nonstationary EEG time series:

- I. **Segmentation:** We may assume that segments of EEG in small overlapping windows are at least quasi-stationary, so an MVAR model can be accurately fit to each of the segments. Adopting a similar concept to Short Time Fourier Transform (STFT), we therefore model local sections of a multivariate signal as it changes over time.

A problem with this approach is the concern of having sufficient data points falling within each segment. Fitting an MVAR model to a high-dimensional time series amounts to determining a large number of parameters relating each channel to lagged values of other channels. Therefore, a large number of data points are needed to have a well-posed fitting problem and this restricts our ability to choose a short window length for satisfying assumptions of quasi-stationarity. Nonetheless, fitting MVAR models to long segments of EEG has been suggested in the literature [25], [26], and can be justified as being useful in assessing the aggregate connectivity structure within each window, disregarding the transients. In section 5.1, we will fit models to 10-second epochs of

EEG and average the results over all epochs. This may lead us to an overview of the average connectivity structure inferred from a full-length recording during a specific task.

II. Adaptive models: Alternatively, we can model the nonstationary EEG using adaptive variants of the MVAR process, i.e. assuming that the model itself varies over time in accordance with the data. In this sense, the matrices \mathbf{A}_m (and hence the connectivity structure) are dependent on time and the equation 3 can be modified to represent instantaneous model parameters:

$$X(t) = \sum_{m=1}^p \mathbf{A}_m(t)X(t-m) + E(t) \quad (4)$$

This approach will capture the transient features of effective connectivity and result in *dynamic* task-specific connectivity analysis. Details regarding adaptive MVAR modeling are discussed in section 4.2.

Through the rest of this paper, we have used a constant parameter matrix notation \mathbf{A} . It goes without saying that in the adaptive case, this matrix (and its corresponding variants) can be implicitly substituted with $A(t)$, its instantaneous value at time t .

2.4 Representation in Frequency Domain

The formulation in section 2.2 can be transformed to the frequency domain to study couplings in different frequencies, as is common with EEG analysis. Specifically, rearranging 3 and assuming $\hat{\mathbf{A}}_0 = I$ and $\hat{\mathbf{A}}_m = -\mathbf{A}_m$,

$$E(t) = \sum_{m=0}^p \hat{\mathbf{A}}_m X(t-m). \quad (5)$$

Transforming 5 into the frequency domain, we get

$$\mathbf{E}(f) = \mathbf{A}(f)X(f) \quad (6)$$

$$\mathbf{A}(f) = \sum_{m=0}^p \hat{\mathbf{A}}_m \exp(-2\pi i m f / f_s) \quad (7)$$

where f_s is the sampling frequency. 6 can also be rearranged in the following form:

$$X(f) = \mathbf{A}^{-1}(f)\mathbf{E}(f) = \mathbf{H}(f)\mathbf{E}(f) \quad (8)$$

Equation 8 suggests that the AR approach models the process \mathbf{X} as a filter acting on the white noise process \mathbf{E} . Since the spectrum of white noise is flat over all frequencies, information about the spectral content of \mathbf{X} is contained in the matrix $\mathbf{H}(f) \triangleq \mathbf{A}^{-1}(f)$, also known as the transfer matrix. From the transfer matrix $\mathbf{H}(f)$ and the prediction error covariance matrix Σ , the spectral density matrix \mathbf{S} of the process can be calculated as

$$\mathbf{S}(f) = \mathbf{X}(f)\mathbf{X}^*(f) = \mathbf{H}(f)\mathbf{E}(f)\mathbf{E}^*(f)\mathbf{H}^*(f) = \mathbf{H}(f)\Sigma\mathbf{H}^*(f) \quad (9)$$

where $*$ denotes complex conjugation. As we shall see in the following section, matrices $\mathbf{S}(f)$, $\mathbf{H}(f)$, and $\mathbf{A}(f)$ derived from the EEG process carry information about directed connectivity, and several quantitative connectivity metrics have been defined in the literature based upon these matrices, each targeting a different aspect of information flow.

2.5 Frequency Domain Estimators of Directed Connectivity

Here we introduce and define a selection of quantitative metrics for effective connectivity in a coherence and/or Granger-causal sense, derived from the matrices defined in the previous sub-section. Our goal is not to provide a comprehensive review on these measures, so the list goes well beyond the few measures introduced herein.

- **Coherency (Coh):** Perhaps the simplest measure of coupling in the frequency domain is Coherency, defined in terms of the spectral matrix \mathbf{S} as

$$C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}} \quad (10)$$

Coherency measures the degree of synchrony among the subprocesses at different frequencies and is not a directional measure. The directional versions of coherency, such as Directed Coherence [27] are limited to bivariate models and do not fully consider the multivariate nature of the process [28].

- **partial Coherence (pCoh):** In a multivariate process, coherence between two subprocesses might be influenced by all other variables. The Partial Coherence [29], [30] attempts to find the portion of coherence between two subprocesses which cannot be explained by a linear combination of other common inputs. Partial Coherence is defined as:

$$P_{ij}(f) = \frac{M_{ij}(f)}{\sqrt{M_{ii}(f)M_{jj}(f)}} \quad (11)$$

where M_{ij} is a minor determinant of \mathbf{S} with the i -th row and j -th column removed. It can be shown [31] that (11) is equivalent to

$$P_{ij}(f) = \frac{\hat{S}_{ij}(f)}{\sqrt{\hat{S}_{ii}(f)\hat{S}_{jj}(f)}} \quad (12)$$

where $\hat{\mathbf{S}}(f) = \mathbf{S}^{-1}(f)$.

- **Partial Directed Coherence (PDC):** Another estimator based on the matrix $\mathbf{A}(f)$, the Partial Directed Coherence (PDC), has been proposed in [32]. PDC is defined in terms of the coefficient matrix \mathbf{A} as:

$$\pi_{ij}(f) = \frac{A_{ij}(f)}{\sqrt{\sum_{k=1}^K |A_{kj}|^2}} \quad (13)$$

The complex quantity $\pi_{ij}(f)$ can be interpreted as the causal flow from channel j to channel i normalized by all outflows from channel j . Since it is based on the values A_{ij} (the parameters of the MVAR model), PDC can be viewed as a frequency-domain equivalent of multivariate Granger causality [33].

- **Directed Transfer Function (DTF):** Similarly to PDC, the Directed Transfer Function (DTF) [21] is a multivariate directional measure defined based on the elements of the transfer matrix \mathbf{H} as

$$\gamma_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_{k=1}^K |H_{ik}(f)|^2} \quad (14)$$

γ_{ij}^2 represents the directional flow from channel j to channel i normalized by the sum of flow from all channels to channel i . The normalization in the original definition of DTF (14) is done in order to compare directed components in signals with different power spectra [34]. However, DTF can also be defined in a simpler, non-normalized format as:

$$\theta_{ij}^2(f) = |H_{ij}(f)|^2 \quad (15)$$

It is argued in [23] that DTF does not represent Granger Causality. Rather, DTF and Granger causal tools such as PDC focus on different and complementary aspects of the connectivity structure.

- **full-frequency Directed Transfer Function (ffDTF)**: Integrating the denominator in (14) over frequency leads to a variant of DTF, full-frequency DTF (ffDTF) [35]:

$$\lambda_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_f \sum_{k=1}^K |H_{ik}(f)|^2} \quad (16)$$

Compared to DTF, ffDTF allows for better interpretation of the estimator characteristics at different frequencies.

- **direct Directed Transfer Function (dDTF)**: DTF and its variants show not only direct but also indirect, mediated interactions [35]. For instance, if two non-interacting channels A and B are influencing channel C such that $A \rightarrow C$ and $B \rightarrow C$, then DTF will falsely detect the indirect interaction $A \rightarrow B$. A more robust variant of DTF that is able to distinguish between direct and indirect flows among channels is introduced in [35] as the product of full-frequency DTF and partial coherence:

$$\begin{aligned} \delta_{ij}^2(f) &= \lambda_{ij}^2(f) P_{ij}^2(f) \\ &= \frac{|H_{ij}(f)|^2}{\sum_f \sum_{k=1}^K |H_{ik}(f)|^2} \times \frac{(\hat{S}_{ij}(f))^2}{\hat{S}_{ii}(f) \hat{S}_{jj}(f)} \end{aligned} \quad (17)$$

Direct DTF (dDTF) is thus a multivariate directional measure which combines information from both DTF and Partial Coherence.

A comparison of some of the measures introduced above is presented in [36], where it is concluded that all measures perform nearly equivalently under reasonable recording conditions. In this paper, we proceed with dDTF as our measure of effective connectivity since it has the combined advantages of ffDTF and Partial Coherence.

3 Data Acquisition

The battery-operated NR SIGN EEG 5000Q 64-channel EEG system (NR SIGN Inc., Vancouver, BC, V6S2L9) was used for non-invasive EEG recording (Figure 1). Data was acquired at a sampling rate of 500 Hz with 16-bit resolution, and was transferred using a USB cable to a laptop computer (also running on battery power) through the NR SIGN EEG application software. The equipment was used to acquire raw, unfiltered continuous EEG recordings at LSBB from seven subjects. Similar recordings were then acquired in ICORD (International Collaboration On Repair Discoveries), a clinical research facility within Vancouver General Hospital, using the same equipment and protocols to allow for comparison between the two recording environments. Five subjects had volunteered for data acquisition at ICORD, four of whom were also present during acquisition in LSBB. For the purpose of the study, three subjects were selected out of the four subjects common to both environments, since the acquired data for one subject was found to be corrupted with large movement

and muscle artifacts, as well as EOG (high-amplitude electro-oculogram artifact caused by blinking of the eyes) to a high extent. In each of the environments, the subjects were asked to perform a variety of cognitive, sensory, and motor tasks while EEG was acquired in real time. First, EEG data was collected from the subjects in a darkened room during a relaxed (resting) state as a baseline condition (with either eyes open or eyes closed). The subjects were then asked to perform specific tasks over a 30-second period, while resting for 30 seconds in between the test intervals. The tasks included: 1) counting backwards by 7 from a large randomly selected number; 2) performing an increasingly challenging matching memory task on an iPad; 3) performing reciprocal dorsal and plantar flexion movements of the right ankle; and 4) performing repeated flexion-extension of the right wrist. All subjects were right-hand dominant. To assess the effects of pain on the EEG, the subjects were subjected to 5) brushing of the adductor pollicis region of the right thumb with a cotton swab as an innocuous tactile stimulation task; and 6) application of a hot pack (recorded skin temperature = 34-36 deg C) to the same adductor pollicis region as a noxious stimulus, while rating their pain perception using the (0-10, 10 being most painful) visual analog scale (VAS). For the latter task, only two out of the three subjects were included in the analysis since one subject did not continue the recording at the hospital environment, claiming the stimulus was too hot to tolerate.

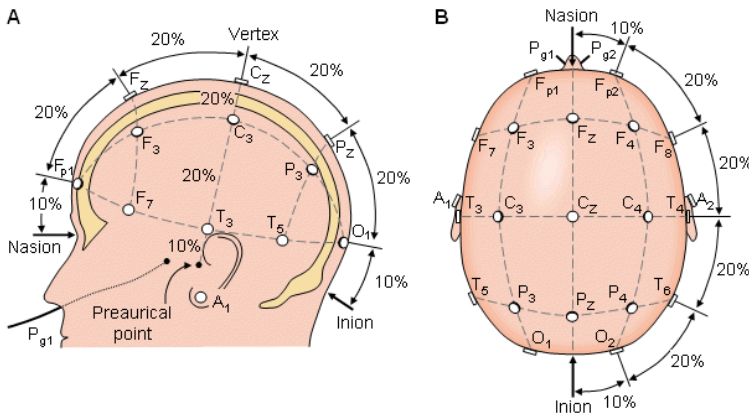


Figure 1: 10-10 electrode montage used for EEG acquisition. Reference and ground electrodes are not shown in the figure.

4 Workflow, Methods, And Practical Considerations

In section 2 we outlined the theory underlying Granger-causal analysis and introduced a number of connectivity estimators. We concluded that we can estimate effective connectivity in multi-channel EEG using an approach based on linear Multivariate Autoregressive models; and that this approach can be used adaptively to infer instantaneous connectivity patterns. The detailed procedure for obtaining an estimate of the connectivity structure from raw EEG data is depicted in Figure 2. In the following, we will discuss methods and practical issues related to each step of the procedure.

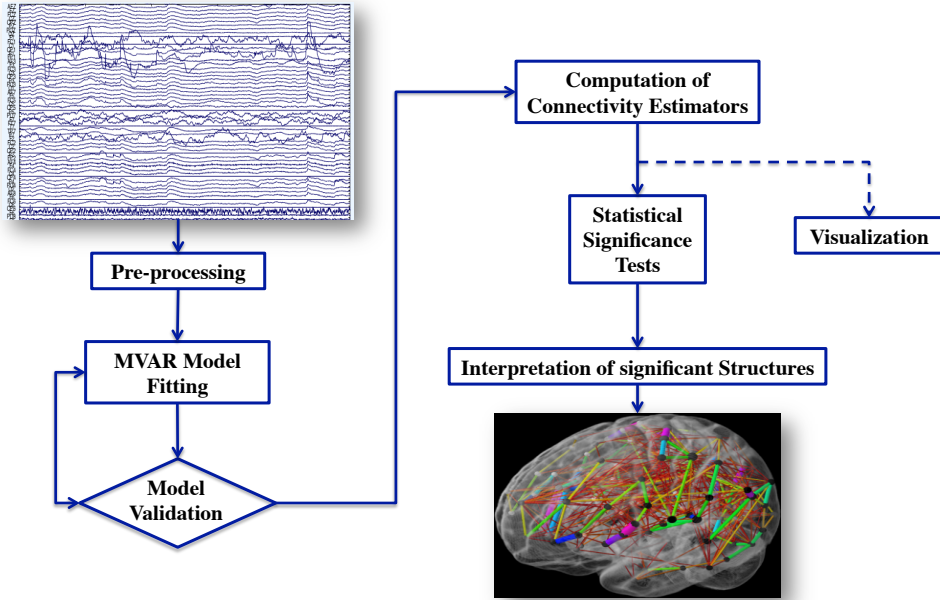


Figure 2: From raw time series to connectivity patterns: steps for estimating effective connectivity based on multivariate autoregressive models

4.1 Pre-processing And Artifact Removal

The first step is to remove artifacts from the data using stringent criteria. Indeed, contrary to our previous work, here we also reject, as much as possible, portions of data corresponding to blinks and low-frequency artifacts as they affect and distort the parameters of the MVAR model, especially in the adaptive, time-varying case.

Extra caution needs to be exercised in pre-processing of the EEG for estimation of the connectivity measures. Since these measures are dependent on the signal phase, any filtering with phase distortion would invalidate the final results. Re-referencing is another procedure that would distort the estimates, as the choice of reference is shown to have a significant impact on the derived network attributes [37], and average-rereferencing mixes the signals and introduces false correlations between them. We thereby chose to have as little pre-processing steps as possible, bypassing the conventional pre-filtering steps. We do, however, normalize every channel by removing the mean and dividing by the standard deviation according to the Equation (18) below prior to segmentation and MVAR modeling:

$$x_k^{norm}(t) = \frac{x_k(t) - \overline{x_k(t)}}{\sigma_k}, \quad k = 1, \dots, K \quad (18)$$

where $\overline{x_k(t)}$ and σ_k refer to the temporal mean and the temporal standard deviation of the signal at electrode k , respectively. Since many of the connectivity measures described in section 2.5 are highly dependent on scale and variances of the signals, this step is performed

to scale the variances among different signals to a comparable range in order to prevent model misspecifications.

4.2 Model Fitting And Validation

Having ensured that the data is free of artifacts, the next step would be to fit an MVAR model to the time series. Model fitting refers to implementation of an algorithm for computing the coefficient matrices \mathbf{A}_m , $m = 1, \dots, p$ and the error covariance matrix Σ (See Equation 3) given the process time series \mathbf{X} collected over T time points. Here we discuss our choice of the fitting algorithm and the criterion for choosing the model order p .

4.2.1 The Model Fitting Algorithm

As discussed in section 2.3, there are two approaches for MVAR modeling of nonstationary EEG time series: 1) Segmentation, which results in batch-averaged connectivity estimation, and 2) Adaptive MVAR modeling, resulting in instantaneous (dynamic) connectivity estimation.

Static MVAR models can be fitted to a batch of signals falling within a window using various methods including least-squares (Yule-Walker) approaches, Burg's method, and the Vieira-Morf algorithm. In this paper, we have used an efficient step-wise least squares method proposed in [38] and [39]. The only consideration here is the choice of the window length, which is dependent upon the minimum analysis frequency and the number of parameters of the MVAR model. We chose the window length to be 10 seconds as in [25] and [26]. Considering the estimation problem, there are $K^2 p$ free parameters to estimate in an MVAR model and as a rule of thumb, at least 10 times more data samples are needed for an accurate estimation ([40]). With a generic order of $p = 8$, we have $60^2 * 8 \approx 29000$ parameters to estimate and we need at least 290000 data points, which are provided in a 10-second window ($60(\text{channels}) * 500(\text{Hz}) * 10(\text{s}) = 300000$). Therefore, the selected window length of 10 s seems to be a reasonable choice. We have also included an overlap of 50% (5 seconds) between windows in order to have more batches of available data as well as a smoother distribution of connectivity parameters.

In addition to batch-based models, adaptive MVAR models are estimated using the Recursive Least Squares (RLS) algorithm with forgetting factor [41]. The RLS algorithm was preferred over methods such as Kalman filtering due to being better suited for high-dimensional data. Specifically, in Kalman-filter based approaches, the required matrix inversions cannot be avoided [41] and this is specially undesirable in cases when the dimension of the time series is large. Also, it is shown in [41] that the model dimension has no influence on the RLS algorithm's adaptation speed and its estimation properties. The only tuning parameter of the algorithm, the forgetting factor, represents the trade-off between adaption speed and the variance of the estimation (this parameter was empirically set to 0.002 and was fixed throughout all recordings).

4.2.2 Choice of the Model Order

In an autoregressive model, the order p is the number of lags used for regression on the previous process values. Since the order is not known a priori, we need a data-driven criterion that would determine an optimal value for p given the multivariate time series. The most popular methods to this end are information-theoretical approaches such as Akaike's Final Prediction Error (FPE) [42], and the Schwarz Bayesian Criterion (SBC) [43]. Generally, these methods attempt to minimize an entropy-based objective function comprising of a prediction

error term and a penalty term for including too many parameters (large model orders). By minimizing both terms over a range of model orders, they search for an order which is optimal in the sense that it is both parsimonious and that it predicts the data well.

The FPE and SBC methods function rather conservatively and impose too high a penalty for large model orders; that is, when used on a 10-second segment of our EEG data, FPE and SBC criteria yield optimal orders of 3 and 1 respectively, which are likely too low for accurate spectral identification. Therefore, we have used a rather heuristic approach for picking a proper model order. This approach is based on the fact that if the model is accurately representative of the data, the correlation structure of the data will be completely described by the model. Hence the residuals $\mathbf{U}(t) = \mathbf{X}(t) - \sum_{m=1}^p \hat{\mathbf{A}}_m \mathbf{X}(t-m)$ should not exhibit significant correlation, and validating the model amounts to checking the whiteness (uncorrelatedness) of its residuals. This can be assessed by computing cross-correlations of the residuals up to some maximum lag, and checking whether they will be sufficiently small with the current choice of the model order.

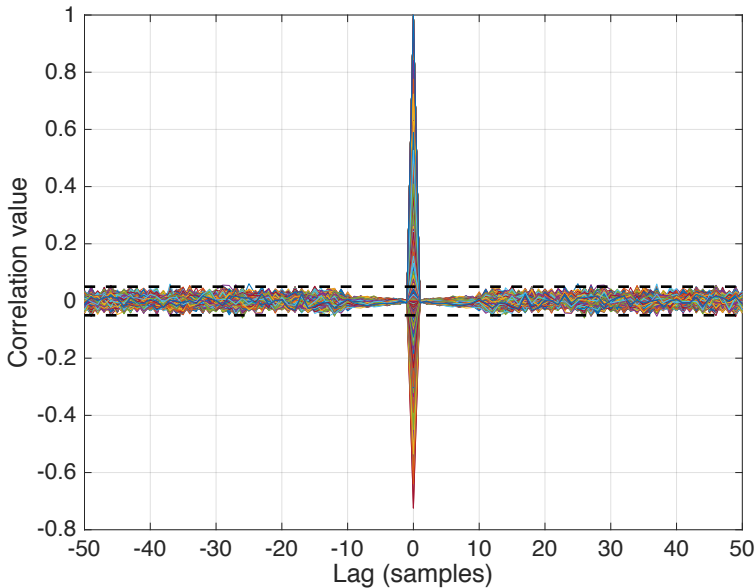


Figure 3: Normalized auto and cross-correlations among all 60 channels of the model residuals

We found that a model order $p = 10$ is enough to keep the normalized covariance of the residuals reasonably small. Figure 3 depicts the correlation structure of the residuals of a model with $p = 10$ fitted to a 10-second segment of the data (the 60^2 autocovariance and cross-covariance sequences for all combinations of the residual dimensions are overlaid on the same plot). We observe that with $p = 10$, residual correlations will be bounded by ± 0.05 for all lags other than zero. Hence, we may claim that with %95 confidence, the residuals are white and the model is valid. We have also empirically assumed that this choice of the optimal model order is sufficient to keep the residuals white among all of the recordings. Based on this observation, the order was also kept fixed at $p = 10$ for dynamic MVAR modeling in the RLS algorithm.

4.3 Computation of Connectivity Estimators

Once we have fitted a valid MVAR model to the data, we may proceed to computing the quantitative measures of effective connectivity described in section 2.5 from the parameters of the model. As shown in section 2.2, fitting MVAR models to the time series data would result in K -by- K parameter matrices A_m , $m = 1, \dots, p$. These parameter matrices are then transformed to the frequency domain (Equation 7) to yield arrays of the form $\mathbf{A}(f)$ ($K \times K \times N_t \times N_f$), where

- N_t is, for the static connectivity case, the number of overlapping windows (segments) extracted from the whole length of the recording; whereas in the dynamic connectivity case N_t is the number of discrete time samples;
- N_f is the number of discrete frequency values at which $\mathbf{A}(f)$ is evaluated.

These four-dimensional arrays encapsulate the model parameter information in terms of time (or epoch), frequency, and channel-wise interactions. Once computed, they are passed to functions calculating the connectivity estimators explained in section 2.5, which in turn yield connectivity arrays of a similar form \mathbf{C} ($K \times K \times N_t \times N_f$). As a case in point, assuming we have calculated the connectivity measure DTF (Equation 15), the element (i, j, t, f) in the connectivity array \mathbf{C} represents the causal interaction of channel j on channel i at time t and frequency f as measured by the metric DTF.

An issue arising in computation of dynamic (time-dependent) connectivity structures is that these arrays can easily become so large that they exceed the available computer memory and cause programs to be unresponsive. For instance, when calculating the time-varying dynamic connectivity of 10 seconds of data in 50 frequency points, the connectivity array will take up to $60 \times 60 \times 10 \times 500(F_s) \times 50 \times 8(\text{bytes per array element}) = 7200$ megabytes in memory. Considering the fact that these arrays are to be further manipulated through the rest of the program and compared among conditions and subjects, they are impractical to use unless somehow compressed and reduced. To this end, we have eliminated the third dimension (frequency) by integrating the measures over frequency in the bands of interest. Moreover, in this paper, we have examined the time-varying connectivity structure only during a short period of time after the beginning of performing a task (usually the first 5 seconds).

4.4 Tests for Statistical Significance

Proper interpretation of connectivity patterns is not achievable without a suitable statistical testing scheme which is able to distinguish significant interactions from insignificant ones. Specifically, we have seen that computation of any connectivity estimator on a segment of EEG at a particular time and frequency yields K^2 values, each corresponding to the directional flow from one channel to another. These values often need to be compared between two conditions and assessed among subjects to infer significant and consistent patterns. Similar to our approach in the previous paper, we have compared pair-wise connectivity values during each task to those of the resting state. In this sense, a task-specific connection from channel A to channel B is deemed significant if and only if its connectivity value is greater (in a statistical sense) than that of the resting state. Our methods for statistical testing are customized to the static and dynamic connectivity estimation as follows:

- In the case of static connectivity, we are comparing frequency-integrated arrays $C_{rest}(K \times K \times N_{rest})$ and $C_{task}(K \times K \times N_{task})$. Each directed pair of channels has N_{rest} realizations (hence a 'distribution' of connectivity values) in the resting state and N_{task} realizations during performance of the task (Figure 4a). For each directed pair of channels, we can therefore compare means of the two distributions for the rest and task conditions using a

Wilcoxon signed rank test. Once the means of all resting-state batches are compared with those of the task state, a $K \times K$ 'significance matrix' \mathbf{S} of zeros and ones is derived in which $\mathbf{S}(i, j) = 1$ implies that for the interaction from channel j to channel i , the mean connectivity value during task is significantly greater than the mean connectivity value during rest, and hence the connection from j to i is task-specific. By comparing the means of the two distributions, we are essentially examining the time-averaged (or more specifically, batch-averaged) connectivity over all epochs of the rest and task recordings.

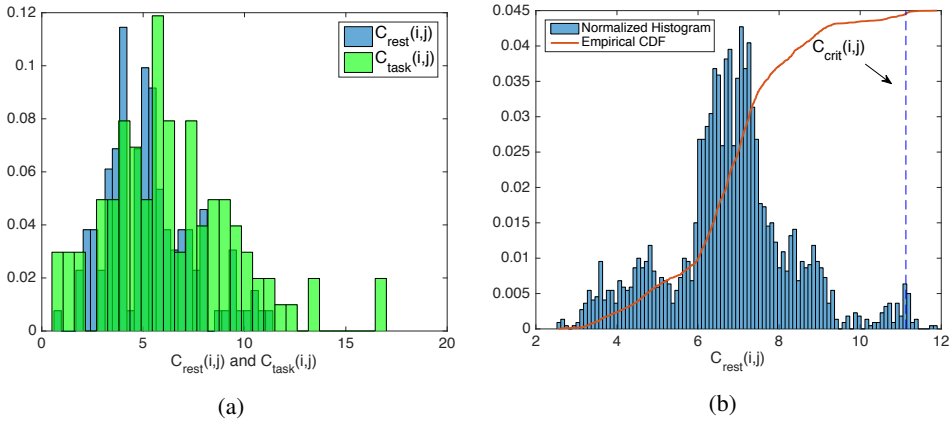


Figure 4: Our statistical significance testing method for determination of significant pair-wise interactions. The directed pair whose connectivity values are depicted above are $FC2$ and $F3$, with $C(i, j)$ signifying the directed interaction $FC2 \rightarrow F3$. (a) In static connectivity estimation, all values of pair-wise interactions at rest are compared with those during task in an offline manner; (b) In dynamic connectivity estimation, instantaneous values of interactions during task are compared with a critical value obtained from the cumulative distribution of all resting-state values.

- As for dynamic connectivity, significant interactions need to be identified at every instant throughout the length of the recording. Hence, pair-wise distributions of rest and task connectivity values cannot be assessed in an offline manner that leads to loss of temporal information. Rather, we may gather all resting-state connectivity realizations (at all time instants) into a 'baseline distribution', and examine at what point in time the connectivity value during a task is exceeding a 'critical' value (Figure 4b). In this context, the critical connectivity value $C_{crit}(i, j)$ for each directed pair of channels is defined as the value at which the cumulative distribution of resting-state connectivity values reaches $1 - \alpha$, where α is some significance level. Essentially, significant connectivity values during a task are values which are unlikely to occur during rest; and that is the rationale behind looking at the tail of the resting-state distribution. The significance matrix \mathbf{S} in this case would be of size $K \times K \times N_t$ and would comprise of the instantaneous significant interactions evolving over time.

It is worth mentioning that simultaneous comparison of a large number of interactions will raise the chances of occurrence of type 1 errors. Hence, the significance level needs to be corrected and set to a more conservative value to decrease the number of false-positive significant interactions. We have used the Bonferroni correction method for this purpose, i.e. lowered the significance level from α to α/K^2 .

4.5 Visualization and Further Data Reduction

So far, we have fitted MVAR models to batches or instants of the time series, computed the connectivity metrics, and assessed the statistical significance of these metrics. At this point, we need to be able to visualize and interpret a vast number (i.e. 60^2) of significant and insignificant interactions that may or may not differ among subjects and even different instances within the same recording. Visualization of a connectivity structure is not an easy task, especially when the number of electrodes is large and in scenarios like ours where the notion of direction needs to be preserved by the visualization method.

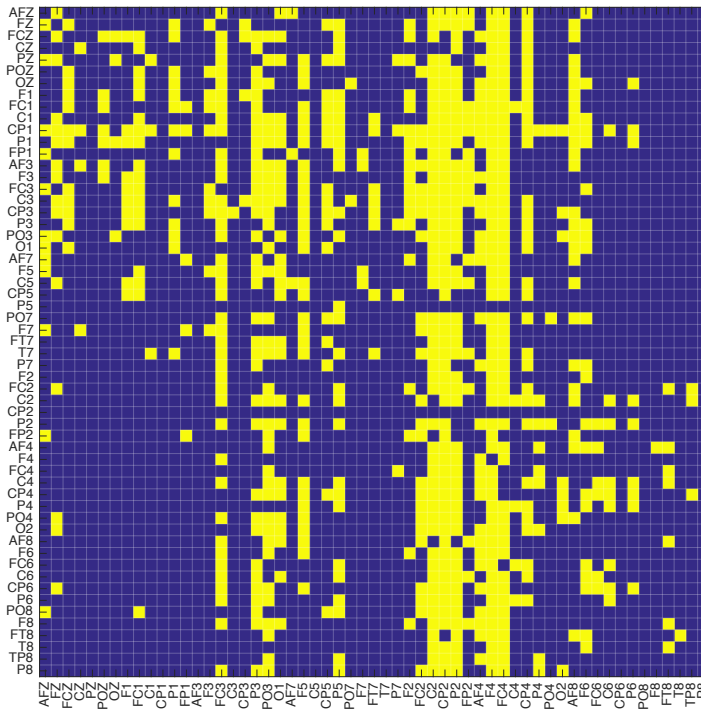


Figure 5: Visualization of the connectivity structure by plotting the significance matrix. Starting from top left, the ij -th element represents the connection from channel j (corresponding column below the pixel) to channel i (corresponding row to the left). Yellow pixels correspond to statistically significant interactions, while blue ones represent insignificant connections.

Perhaps the most readily available visualization scheme is to plot the significance matrix as an image, such that each pixel corresponds to a directed pair-wise interaction and its color indicates whether or not the interaction is significant. Figure 5 depicts such an image obtained from comparison of gamma band static connectivity between the ankle movement and rest conditions for subject 1. Even though a few high-level features can be identified from this figure (see section 5), it contains too much detail to be informative of the overall connectivity structure in the first glance. Moreover, the significance matrix changes instantly in the dynamic scenario, making it extremely difficult to follow the patterns. There is thus an inevitable need to reduce this information and represent it in a compact and more interpretable manner.

Brain Region	Electrodes
Prefrontal	FPZ, FP1, FP2, AFZ, AF3, AF4, AF7, AF8
Frontal	FZ, F1, F2, F3, F4, F5, F6, F7, F8, FCZ, FC1, FC2, FC3, FC4
Central	CZ, C1, C2, C3, C4, CPZ, CP1, CP2, CP3, CP4
Left Temporal	FC5, FT7, C5, T7, CP5, TP7
Right Temporal	FC6, FT8, C6, T8, CP6, TP8
Left Parietal	P3, P5, P7, PO3, PO7
Right Parietal	P4, P6, P8, PO4, PO8
Parietal	PZ, P1, P2, POZ
Occipital	OZ, O1, O2

Table 1: Grouping of the electrodes in a 10-10 montage based on their locations

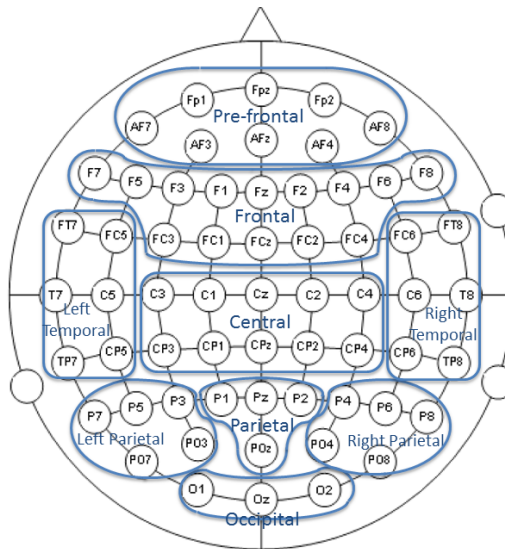


Figure 6: Topographical map of electrode groups based on the classification in Table 1

To this end, we further reduce the significance matrix, by grouping the neighbouring electrodes into nine specified brain regions. In this sense, the 60^2 channel-wise interactions will be reduced to 9^2 regional interactions, where the ij -th element in the 'regional interaction matrix' indicates the average number of significant interactions from the j -th to the i -th brain region, where the average is taken over all directed channel pairs between the two regions. Application of this averaging procedure on the connectivity structure in Figure 5 yields Figure 7, where brighter colors show stronger significant interactions. This figure represents information not directly perceivable from Figure 5, such as the strong frontal-to-parietal interaction. Nonetheless, a major shortcoming of the brain regioning approach is that regardless of the choice of boundaries, neighboring electrodes do not always fall into the same group (e.g. in Figure 6, channels FC4 and FC6 are immediate neighbors, but belong to two distinct groups). In other words, 'spatial discretization' of the electrodes does not take into account the interactions between adjacent electrodes near the boundaries, resulting in a relatively smeared representation of regional interactions.

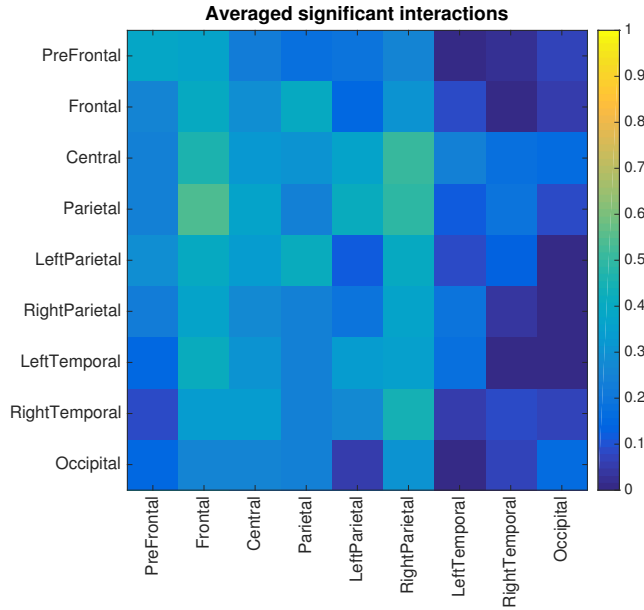


Figure 7: Reduction of Figure 5 by averaging channel interactions within brain regions, where the average is taken over all directed channel pairs (e.g. the regional interaction value from region i to region j is calculated by taking the average of all $N_i \times N_j$ pair-wise connectivity values between the two regions, where N_i and N_j are, respectively, the number of channels in regions i and j). Each colored block represents the average strength of directional connections from the region below the block to the region on its left. For illustrative and comparison purposes, interaction matrices are normalized to have unit Frobenius norm.

4.5.1 Graph-theoretical Measures

In order to overcome these limitations and further reduce the high-dimensional data for visualization (especially in the case of dynamic connectivity patterns), we may turn to graph theory. The main idea behind graph-theoretical measures is that large connectivity datasets have the same characteristics as 'networks' emerging in biology, economy, internet, and other fields; and their properties can be characterized as holistic, compact, meaningful, and easily computable network measures. Graph theory is a field of mathematics defined to study these networks and their properties and has been extensively used in the past decade for the analysis of brain networks [44].

In mathematics, a graph is a structure consisting of a set of 'nodes' having some sort of inter-connections represented as 'edges' (Figure 8). In the context of brain networks, nodes represent channels and edges represent the strength (or presence) of the connectivity measures. Once the computed connectivity structure is defined in terms of a graph, features of the network can be summarized in terms of quantitative graph-theoretical measures such as centrality, clustering coefficient, efficiency, path length etc. [44]. In this paper, we have used a few basic graph-theoretical measures that are outlined below; including inflow, outflow, and causal asymmetry ratio [45]. Denoting the computed connectivity measure from channel j to channel i by c_{ij} ,

- **Inflow** (I_i) is defined as the sum of causal information flowing from the rest of the system toward channel i : $I_i = \sum_{j=1}^K c_{ij}$,

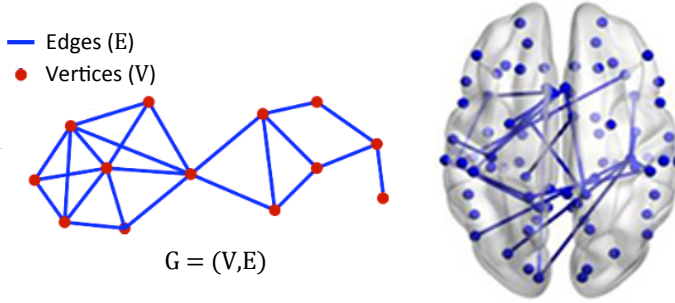


Figure 8: A graph G (shown on the left) is an inter-connected set of vertices (V) and edges (E). Brain networks can be represented as graphs. The non-causal connectivity network depicted on the right can be obtained by thresholding the connectivity values and discarding their direction.

- **Outflow** (O_j) is defined as the sum of causal information flowing from channel j toward the rest of the system: $O_j = \sum_{i=1}^K c_{ij}$,
- **Causal Asymmetry Ratio** (CAR_i) is a normalized value indicating asymmetry of information flowing in and out of channel i : $CAR_i = \frac{O_i - I_i}{O_i + I_i}$.

Based on definitions above, a high value of outflow indicates that the channel acts as a source (causally influencing the system), and channels having high inflow values acts as sinks (being causally influenced by the system). The causal asymmetry ratio has values ranging from -1 to 1 , with positive values close to 1 indicating source behavior and negative values close to -1 suggesting that the channel influences the system more as a sink. A CAR value around zero would indicate that the channel is relatively passive in that the amount of influence it imposes on the network is equalized by the amount of influence it receives from the network.

Since these measures are summing the pairwise connectivity values to obtain channel-wise values, they essentially reduce the dimensionality of the connectivity structure from K^2 to K . In addition to being compact, these measures are intuitive and neurologically insightful. They can easily be computed and visualized on a topographical plot, circumventing the enforced abstraction of location information (as in Figures 5 and 7) and hence simplifying visualization of the directed connectivity structure.

5 Results And Discussion

Having laid the foundations and a framework for estimation of effective connectivity, we now present the results of applying this framework to the data at hand. For simplicity and based on results from the previous paper, we only consider motor tasks for connectivity estimation and will analyze the rest of our data in future publications.

We begin by presenting the results of batch-averaged (static) connectivity in section 5.1. To clarify and sum up, the results presented in section 5.1 are obtained by:

1. cleaning the raw recordings by rejecting artifactual channels (yielding $K' < K$ clean channels per recording) and rejection of transient artifactual time segments,
2. extracting 10-second epochs with 50% overlap from clean rest and task recordings (usually, around 130 rest epochs and 30 task epochs are extracted per subject),

3. fitting an MVAR model ($p = 10$) to each epoch, calculating the epoch dDTF measure (Equation (17)), and integrating the dDTF values in the gamma band,
4. statistically comparing the motor dDTF values with those of the resting state (section 4.4) and obtaining the $K \times K$ pair-wise significance matrices per subject, where values corresponding to the rejected channels are left empty ¹,
5. averaging the $K \times K$ pair-wise significance matrices (ignoring empty values) across subjects to determine significant dDTF connections which are common to all subjects and hence survive averaging.

Steps 4 and 5 above outline our strategy for group-level (between-subjects) analysis: we first obtain subject-specific significance matrices by statistically comparing the two conditions for each individual subject. This procedure rules out, for each subject, any connection which is not task-specific. Subsequently, group-level results are obtained by averaging these significance matrices across subjects. Averaging promotes general trends among all subjects by preserving common connections and suppressing connections which are specific only to a single subject.

Next, results from dynamic connectivity analysis are presented in section 5.2. Similarly, these results have been obtained by:

1. cleaning the raw recordings by rejecting artifactual channels and transient artifactual time segments,
2. extracting the first 5 seconds from the rest and task recordings per subject,
3. fitting an instantaneous MVAR model ($p = 10$) using the RLS algorithm, calculating the dDTF measure at each instant, and integrating the instantaneous dDTF values in the gamma band,
4. statistically comparing the instantaneous motor dDTF values with the overall resting state values and obtaining the $K' \times K'$ pair-wise significance matrices for all time instants.

For reasons explained in section 5.2, dynamic connectivity is not analyzed on a group-level basis and is inspected for all subjects individually.

5.1 Static Connectivity Results

Figure 9 depicts the significance matrices of three subjects in both environments for one of the motor tasks (ankle movements). Although somewhat too detailed, a number of overall structural properties can be deduced from these figures.

First, in many cases, one or more well-defined columns of the significance matrix with connections to most of the channels can be seen. These columns represent prominent sources of information flow and can be useful in identification of 'critical nodes' in the network. However, we observed that these sources are not consistent as they differ among subjects, and even among different recordings of the same subject. They are thus not reported here in isolation, but re-examined shortly using Graph-theoretical measures.

Interestingly, we also observe that for a specific subject, more significant connections are present in recordings at LSBB than their counterparts at the hospital environment. This is a

¹In MATLAB, empty matrix values are implemented as NaN, the IEEE arithmetic representation for Not-a-Number.

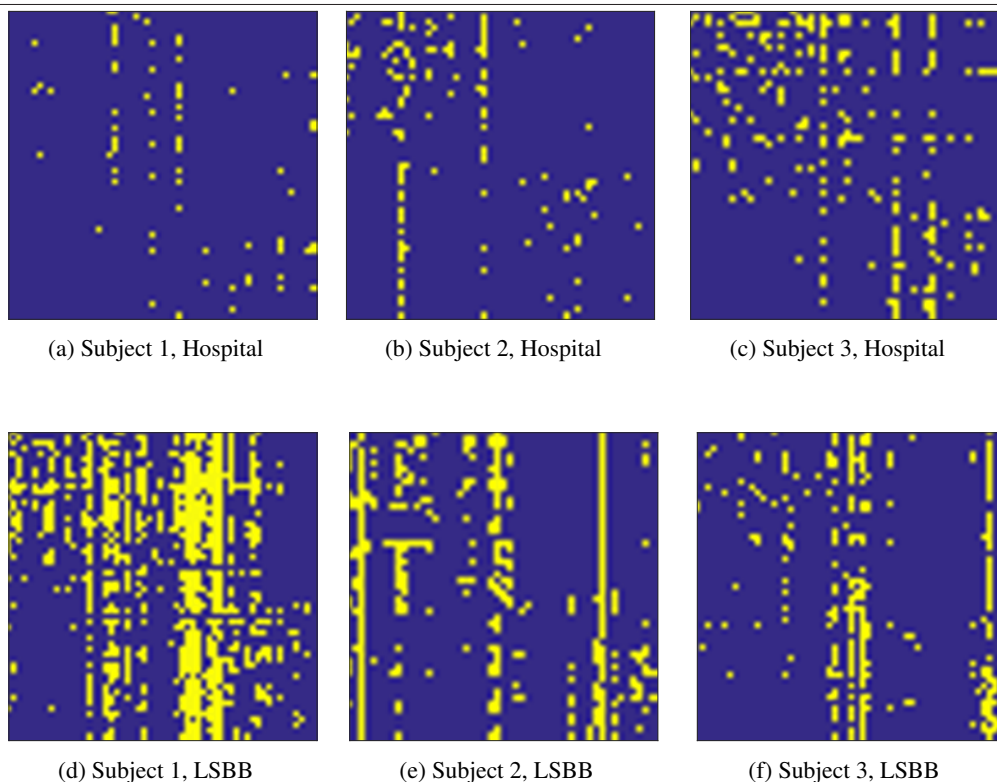


Figure 9: Significant connections during ankle movement from three subjects in the two environments (chosen arbitrarily from the four subjects common to both environments). Electrode labels are not shown on the axes, as the overall structures are the main point of notice. 'Global' sources of activity, i.e. channels with connections to the majority of other channels are evident in subfigures **(d)** and **(e)**; while more local sources, connected only to a subgroup of electrodes, can be spotted in other subfigures.

direct consequence of low-noise conditions, and is in accordance with our findings in [2], as task-specific effective connectivity is another gamma band correlate which is enhanced and better detected in the low-noise environment.

Another property deducible from Figure 9 is laterality and the amount of inter and intra-hemispheric connections. Based on the arrangement of electrodes depicted in detail in Figure 5, and the 10-10 montage shown in Figure 6, entries in a significance matrix can be classified according to the corresponding hemispheres they are connecting. Figure 10 below shows a mapping of all of the pixels to their corresponding inter and intra-hemispheric connections. This crude classification of electrodes leads to the observation that on average, there are more connections within the same hemisphere than between hemispheres, since the yellow pixels (significant connections) are more densely distributed in the top-left and bottom-right corners of the images in Figure 9.

In Figure 11 we have quantified the inter and intra-hemispheric connections across all of the significance matrices, as well as those corresponding to wrist movements. Each entry is the result of calculating the ratio of significant interactions to the total number of possible

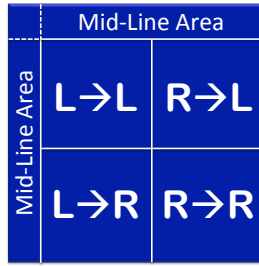


Figure 10: Segmentation of significance matrices into hemispheric connections. This particular arrangement stems from the ordered sequence of electrodes in Figure 5, having the mid-line electrodes first (names ending in 'Z'), electrodes in the left hemisphere second (names ending with an odd number), and lastly, electrodes in the right hemisphere (names ending with an even number). $L \rightarrow L$ denotes connections within the left hemisphere, $L \rightarrow R$ denotes connections from the left hemisphere to the right, and so on.

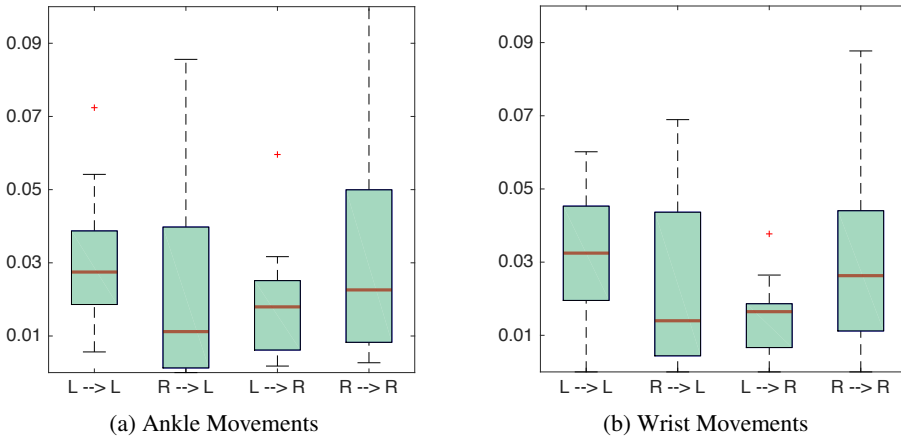


Figure 11: Overall strengths of inter and intra-hemispheric task-specific connectivity during motor tasks in eight subjects and two environments. Box plots represent the median and quartiles of the distribution of hemispheric strength values across all subjects in both environments. Each strength value is essentially the number of significant connections from one hemisphere to the other, divided by the total number of significant connections.

interactions within the regions designated in Figure 10, discarding the pixels falling in the mid-line areas. The figure confirms our previous observation that the number of significant intra-hemispheric connections is more than the number of inter-hemispheric connections. The overall number of connections is also shown to be slightly higher in the left (contralateral) hemisphere than in the right. This suggests that, in a Granger-causal sense, the contralateral hemisphere is more gamma-band-inter-connected during ankle and wrist movements than the ipsilateral hemisphere. Further, there is seemingly more left-to-right connectivity than from the right hemisphere to the left during both ankle and wrist movements. Left-to-right connections are also less variable (as measured by the inter-quartile range of their distribution) among different subjects than other inter and intra-hemispheric connections. They might

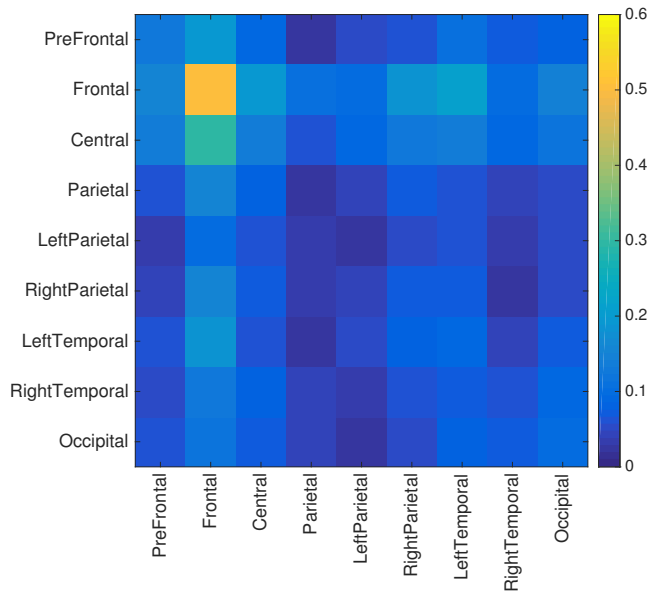
therefore be deemed, with relatively higher confidence, as subject-independent movement correlates. However, these observations may vary if more subjects are added to the analysis.

To further assess which local regions of the brain contribute the most to the task-specific connectivity structure, we can count the number of significant interactions within the brain regions specified in Figure 6. In Figure 12, we have illustrated the mean region-wise significant interactions averaged over all of the recordings (the process is explained in section 4.5). Figure 12a demonstrates that during ankle movements, frontal and central regions have the most inter-relations relative to other sites of the brain, with frontal-to-frontal and frontal-to-central activities being the most prominent region-wise interactions. In contrast, the mid- and left parietal regions are relatively less active in the connectivity process, while there is noticeable information flow from the right parietal and left temporal regions to the frontal region. Figure 12b shows the same characteristics for the wrist movements, although the overall connectivity pattern seems to be more structured and focused around the frontal and central areas compared to ankle movements (i.e. connections between other areas are sparser and connections between frontal and central areas are stronger).

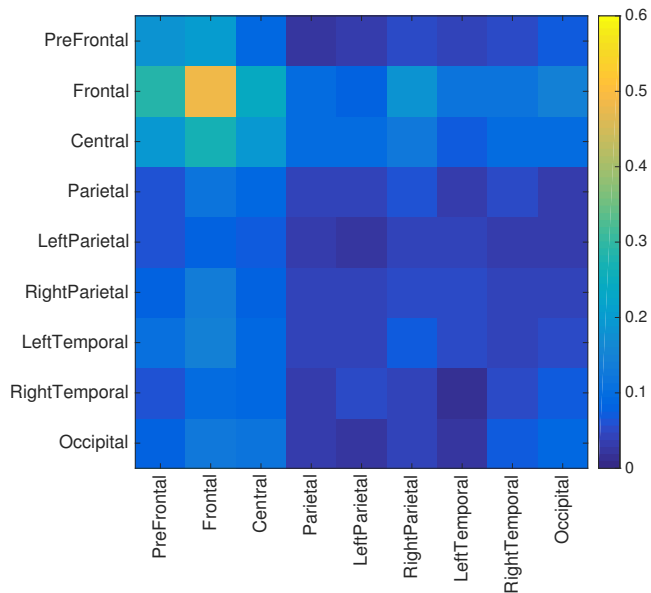
More accurate and interpretable results are obtained using the Graph-Theoretical (GT) measures described in section 4.5.1. We begin by computing the average of significance matrices across all subjects and environments. We will then compute the channel-wise GT measures of inflow, outflow, and Causal Asymmetry Ratio on the average significance matrix. The resulting GT values can then be shown on topographical scalp plots, allowing for a more intuitive representation of significant sources and sinks of information flow. Figure 13 illustrates the results for significant gamma band activity, averaged over recordings from eight subjects, for both of the motor tasks.

According to Figure 13a, on average, a cluster of frontal electrodes on the right hemisphere (F4, F6, FC2) as well as another cluster of parietal electrodes on the left hemisphere (P5, CP5) are identified as the most prominent sources of gamma band activity during ankle movement. These sources propagate information within the gamma band to the central and posterior parts of the brain on the left hemisphere, most notably to the fronto-central and parietal nodes such as FC1, C5, CP1, CZ, and PZ. Meanwhile, Figure 13b shows a similar structure for wrist movements, with dominant pre-frontal and parietal sources slightly shifted toward the midline in comparison with ankle movements. An interesting observation is that for both of the motor tasks, almost all channels with high values of inflow (sinks of information) are located on the contralateral hemisphere and clustered around fronto-central and parietal regions. On the other hand, sources of gamma band activity exist on both hemispheres, feeding information to the rest of the system while being organized in the form of several separate clusters on fronto-central (on the right hemisphere) and parietal (on the left hemisphere) regions of the brain.

It is argued in [46] that gamma band activity in the prefrontal cortex is linked to the maintenance of the behaviorally relevant items. This can explain the observed prefrontal sources of gamma band during the motor activity, as they could be responsible for planning of the next repetitive movement. This information then drives the specific part of the brain known to be responsible for motor function, namely, the central nodes within the contralateral hemisphere, in order to perform the act of movement. It is important to stress here that the confidence of these conclusions is confined by the limited number of subjects in the study, as the objective of this paper was not to reach rigorous neurological discoveries, but rather to report the results on available data as a pilot gamma band study.



(a) Ankle Movements



(b) Wrist Movements

Figure 12: Average regional interactions during motor tasks in eight subjects and both environments. We have recorded the number of significant interactions between the nine previously defined regions of the brain, in an attempt to represent the data in Figure 9 in a more compact manner, and extend the analysis of hemispheric connections.

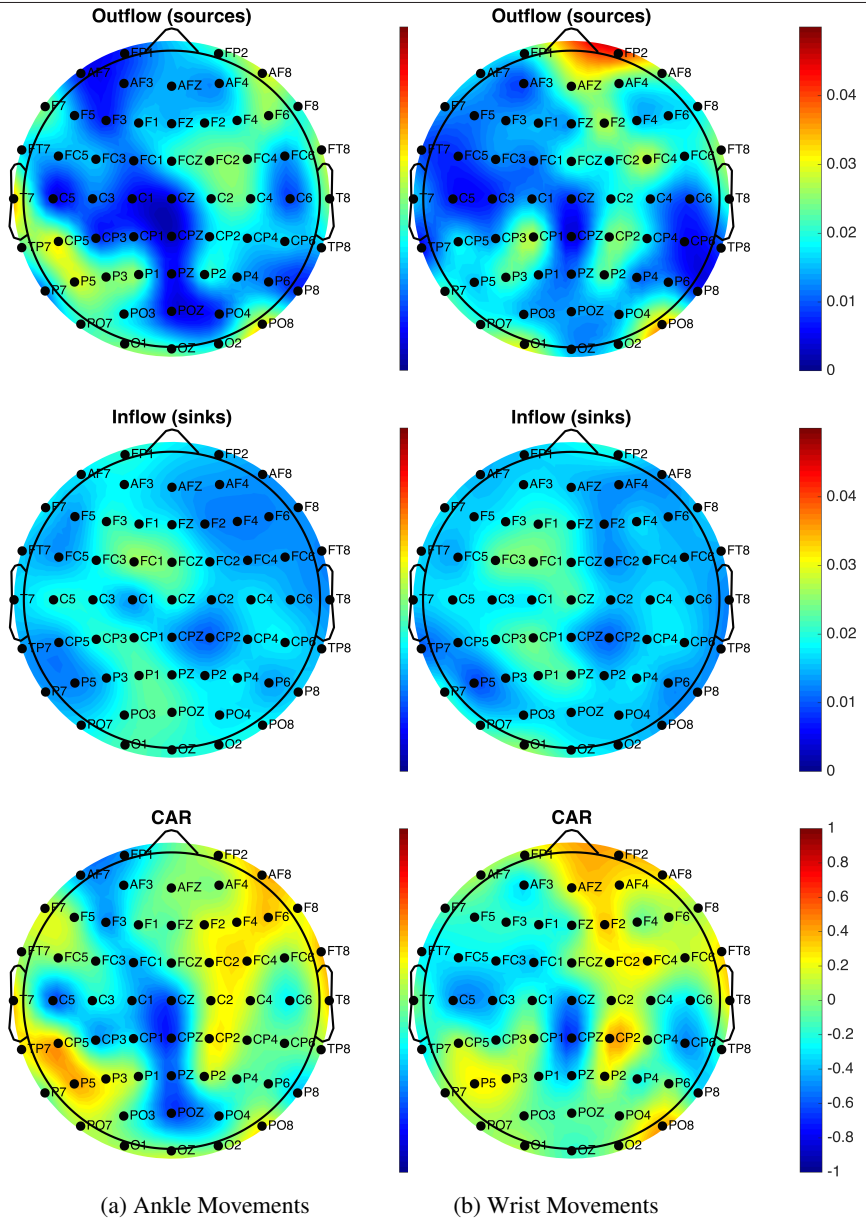


Figure 13: Topographical plots showing the static channel-wise graph-theoretical measures calculated on the mean significance matrix during motor tasks. Top: outflow; middle: inflow; bottom: Causal Asymmetry Ratio (CAR).

5.2 Dynamic Connectivity Results

As explained before, dynamic or instantaneous connectivity is the result of letting the parameters of the model (and hence the connectivity values) vary adaptively with time. Due to memory limitations, calculation of connectivity parameters from model parameters is done one at a time for small consecutive time windows of 5 seconds. Given the fact that a huge volume of data is generated in this manner, we have resorted to observing only the first 5 seconds of connectivity values in order to see which interactions dominate the connectivity structure at the initial stages of performing a task.

In terms of visualization of the results, all of the Figures in the previous section would turn into videos whose frames depict instantaneous connectivity structures. Naturally, the least detailed and most illustrative videos would be those corresponding to graph-theoretical topographical plots. In Figure 14, we have illustrated a few video snapshots from three subjects showing the evolution of the ankle movement connectivity structure throughout the first five-second interval in one-second steps. Shown in the plots is the instantaneous *CAR* value calculated on gamma band significance matrices obtained from recordings at LSBB.

Alternatively, if the directed interaction between a specific pair of channels $A \rightarrow B$ is of interest, we can plot the time course of the computed connectivity measure (integrated within the gamma band) flowing from A to B . For instance, having identified the major sources ($CP2$, $P5$, $F4$, $FC4$) and sinks ($C3$, $CP1$, CZ , PZ , $CP6$) of significant gamma band activity for a single subject from plots similar to those in Figure 13a, we might be curious to know the specific times at which connections between each of these channels are stronger during ankle movements. Figure 15 depicts such information obtained from the first five-second interval of ankle movements from subject 1 at LSBB. The figure illustrates the variability of interactions over time. Specifically, we observe that connections $F4 \rightarrow C3$, $F4 \rightarrow CP1$, and $P5 \rightarrow CZ$ are mainly inactive during this period. Moreover, while $FC4$ is propagating activity only during the initial stages of the task, connections such as $P5 \rightarrow C3$ are activated at later times, and connections $P5 \rightarrow CP1$ and $CP2 \rightarrow CP1$ are consistently active throughout the whole five-second period.

Unfortunately, interpretability of these results and performing a multi-subject analysis is severely limited by the protocol definition and design of experiments. It is apparent from Figures 14 and 15 that the connectivity structure changes rapidly over short instances of time. Hence, in order to be able to compare the time-dependent patterns across subjects and recordings to find consistent connectivity patterns, the movement tasks need to be performed at repeatable epochs, commenced at precisely known start times and executed at closely similar paces. On the contrary, task start times in our dataset were not properly annotated, and the movements were performed at varying self-selected paces. In other words, the subjects in Figure 14 are likely in different stages of ankle and wrist movements due to the high uncertainty in task start times, and hence their activities cannot be compared across time.

5.3 Conclusion

Brain connectivity structure comprises networks of different brain sites connected by anatomical, functional, or causal (effective) associations. In this paper, we extended the segregated analysis of [2] to search for the presence and direction of task-specific gamma band connectivity links in our EEG dataset during motor tasks.

Using the parameters from a data-driven auto-regressive model, we calculated a Granger-causal measure of connectivity, the direct Directed Transfer Function (dDTF), on 10-second segments of the continuous data. This gave rise to a batch-averaged (so-called static) manifestation of the connectivity structure, signifying the general patterns of interaction between

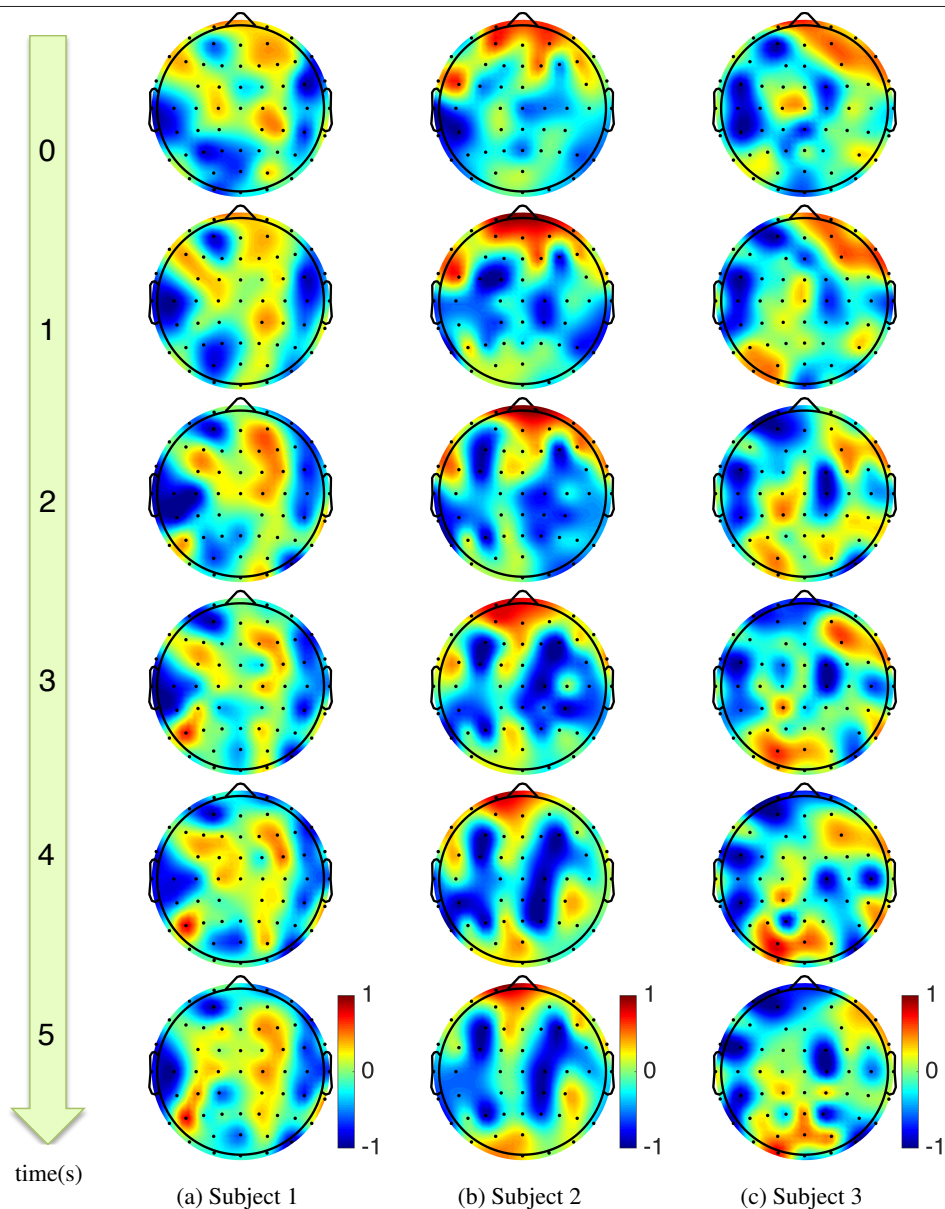


Figure 14: Topographical plots of instantaneous significant *CAR* values during the first five seconds of ankle movement for three subjects in LSBB.

channels within the whole batch period. Using the dDTF values obtained from different batches of the same recording, we devised a statistical significance selection procedure based on the distribution of dDTF values across channel pairs, conditions, and subjects in order to distinguish significant connections from insignificant ones. Moreover, to present the high-dimensional results in a compact manner, we introduced graph-theoretical measures of inflow, outflow, and Causal Asymmetry Ratio. The static connectivity results indicated high inter-connectivity across a wide range of brain regions within the gamma band. Hemispher-

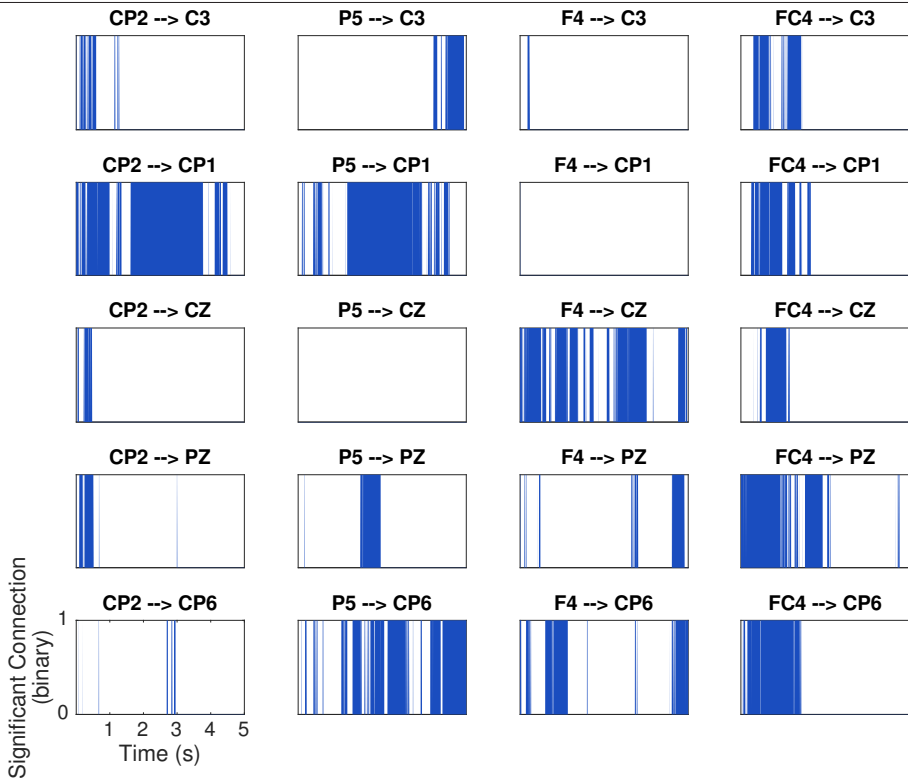


Figure 15: Time course of significant gamma band interactions between a selected array of sources and sinks throughout the first five seconds of ankle movement (subject 1, LSBB). Shaded areas represent the presence of significant gamma band interactions across time (horizontal axis).

ical analysis demonstrated more intra-hemispheric than inter-hemispheric connectivity, and more left-to-right connections than right-to-left (for tasks involving movements of the right hand/ foot). Frontal and central regions contained the most number of significant connections during motor tasks, while significant sources and sinks of information were also seen in other (e.g. parietal) regions.

Furthermore, we extended the auto-regressive model to the adaptive, time-varying case (so-called dynamic connectivity) to incorporate the dimension of time in the analysis. We observed that patterns of connectivity change very rapidly over time, limiting interpretability of the results given the uncertainty of task start times in the current dataset.

Apart from the timing of the experiments, statistical rigor of the results obtained from this dataset is also limited by the small number of participants. More so than any other biological phenomenon, task-specific EEG correlates are known to differ significantly in topology, time-frequency, and connectivity patterns across subjects. Substantial variability was observed even between recordings obtained from the same subject at different times. Hence, many subjects (more than 30, according to the Central Limit Theorem) are needed to obtain more consistent and reliable results, as well as many trials of the same task performed by the same subject. The thorough pipeline and methods introduced in this paper would be of value to the future studies addressing these limitations.

References

- [1] A. Zandi, G. Dumont, M. Yedlin, P. Lapeyrie, C. Sudre, S. Gaffet, *IEEE Transactions on Biomedical Engineering* **58**, 2407 (2011)
- [2] Hamzei, N, Bastany, Z, Jutzeler, CR, Yedlin, M, Kramer, JLK, Steeves, JD, Dumont, GA, *E3S Web of Conf.* **12**, 05003 (2016)
- [3] G. Tononi, O. Sporns, G. Edelman, *Proceedings of the National Academy of Sciences* **91**, 5033 (1994)
- [4] V. Calhoun, R. Miller, G. Pearlson, T. Adalı, *Neuron* **84**, 262 (2014)
- [5] A. Eklund, T. Nichols, H. Knutsson, *Proceedings of the National Academy of Sciences* p. 201602413 (2016)
- [6] F. Wouterlood, M. Vinkenoog, M. van den Oever, *Network: Computation in Neural Systems* **13**, 327 (2002)
- [7] G. Gong, Y. He, L. Concha, C. Lebel, D. Gross, A. Evans, C. Beaulieu, *Cerebral cortex* **19**, 524 (2009)
- [8] M. Hutchison, T. Womelsdorf, E. Allen, P. Bandettini, V. Calhoun, M. Corbetta, S. Della Penna, J. Duyn, G. Glover, J. Gonzalez-Castillo et al., *Neuroimage* **80**, 360 (2013)
- [9] K. Friston, *Brain connectivity* **1**, 13 (2011)
- [10] K. Friston, L. Harrison, W. Penny, *Neuroimage* **19**, 1273 (2003)
- [11] W. Penny, K. Stephan, A. Mechelli, K. Friston, *Neuroimage* **23**, S264 (2004)
- [12] A. Protzner, A. McIntosh, *Human brain mapping* **27**, 935 (2006)
- [13] R. Vicente, M. Wibral, M. Lindner, G. Pipa, *Journal of computational neuroscience* **30**, 45 (2011)
- [14] Z. Wang, P. Lee, M. McKeown, *Biomedical engineering online* **8**, 1 (2009)
- [15] G. Nolte, O. Bai, L. Wheaton, Z. Mari, S. Vorbach, M. Hallett, *Clinical neurophysiology* **115**, 2292 (2004)
- [16] A. Bastos, J. Schoffelen, *Frontiers in systems neuroscience* **9** (2015)
- [17] K. Blinowska, M. Malinowski, *Biological cybernetics* **66**, 159 (1991)
- [18] K. Blinowska, *Medical & biological engineering & computing* **49**, 521 (2011)
- [19] C. Granger, *Econometrica: Journal of the Econometric Society* pp. 424–438 (1969)
- [20] S. Bressler, A. Seth, *Neuroimage* **58**, 323 (2011)
- [21] M. Kaminski, K. Blinowska, *Biological cybernetics* **65**, 203 (1991)
- [22] M. Centeno, D. Carmichael, *Frontiers in neurology* **5**, 93 (2014)
- [23] M. Eichler, *Biological cybernetics* **94**, 469 (2006)
- [24] H. Lütkepohl, *New introduction to multiple time series analysis* (Springer Science & Business Media, 2005)
- [25] S. Chennu, P. Finoia, E. Kamau, J. Allanson, G. Williams, M. Monti, V. Noreika, A. Arnatkeviciute, A. Canales-Johnson, F. Olivares et al., *PLoS Comput Biol* **10**, e1003887 (2014)
- [26] M. Kamiński, K. Blinowska, W. Szelenberger, *Electroencephalography and clinical neurophysiology* **102**, 216 (1997)
- [27] Y. Saito, H. Harashima, Yamaguchi N, Fujisawa K (eds) *Recent advances in {EEG} and {EMG} data processing*. Elsevier pp. 133–146 (1981)
- [28] L. Astolfi, F. Cincotti, D. Mattia, C. Babiloni, F. Carducci, A. Basilisco, P. Rossini, S. Salinari, L. Ding, Y. Ni et al., *Clinical neurophysiology* **116**, 920 (2005)
- [29] J. Rosenberg, D. Halliday, P. Breeze, B. Conway, *Journal of neuroscience methods* **83**, 57 (1998)

- [30] D. Bringer, Expanded Edition, HoldenDay (1981)
- [31] V. Jirsa, A. McIntosh, *Handbook of brain connectivity*, Vol. 1 (Springer, 2007)
- [32] L. Baccalá, K. Sameshima, *Biological cybernetics* **84**, 463 (2001)
- [33] K. Sameshima, L. Baccalá, *Journal of neuroscience methods* **94**, 93 (1999)
- [34] L. Astolfi, F. Babiloni, *Synthesis Lectures on Biomedical Engineering* **2**, 1 (2007)
- [35] A. Korzeniewska, M. Mańczak, M. Kamiński, K. Blinowska, S. Kasicki, *Journal of neuroscience methods* **125**, 195 (2003)
- [36] L. Astolfi, F. Cincotti, D. Mattia, M. Marciani, L. Baccala, F. de Vico Fallani, S. Salinari, M. Ursino, M. Zavaglia, L. Ding et al., *Human brain mapping* **28**, 143 (2007)
- [37] F. Chella, V. Pizzella, F. Zappasodi, L. Marzetti, *Journal of neural engineering* **13**, 036016 (2016)
- [38] A. Neumaier, T. Schneider, *ACM Transactions on Mathematical Software (TOMS)* **27**, 27 (2001)
- [39] T. Schneider, A. Neumaier, *ACM Transactions on Mathematical Software (TOMS)* **27**, 58 (2001)
- [40] A. Schlögl, G. Supp, *Progress in brain research* **159**, 135 (2006)
- [41] E. Möller, B. Schack, M. Arnold, H. Witte, *Journal of neuroscience methods* **105**, 143 (2001)
- [42] H. Akaike, *Annals of the institute of Statistical Mathematics* **21**, 243 (1969)
- [43] G. Schwarz et al., *The annals of statistics* **6**, 461 (1978)
- [44] M. Rubinov, O. Sporns, *Neuroimage* **52**, 1059 (2010)
- [45] T. Mullen, Z. Acar, G. Worrell, S. Makeig, *Modeling cortical source dynamics and interactions during seizure*, in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE)*, 2011, pp. 1411–1414
- [46] F. Roux, M. Wibral, H. Mohr, W. Singer, P. Uhlhaas, *Journal of Neuroscience* **32**, 12411 (2012)