IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Single-trial effective brain connectivity patterns enhance discriminability of mental imagery tasks

This content has been downloaded from IOPscience. Please scroll down to see the full text.

Download details:

IP Address: 128.226.136.66
This content was downloaded on 11/06/2017 at 08:39

Manuscript version: Accepted Manuscript
Rathee et al

To cite this article before publication: Rathee et al, 2017, J. Neural Eng., at press: https://doi.org/10.1088/1741-2552/aa785c

This Accepted Manuscript is: © 2017 IOP Publishing Ltd

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.

As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by-nc-nd/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permission will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

When available, you can view the Version of Record for this article at: http://iopscience.iop.org/article/10.1088/1741-2552/aa785c

Single-trial Effective Brain Connectivity Patterns Enhance Discriminability of Mental Imagery Tasks

Dheeraj Rathee, Hubert Cecotti, Girijesh Prasad

Intelligent Systems Research Centre, School of Computing & Intelligent Systems, Ulster University, Derry ∼ Londonderry, N. Ireland, UK.

E-mail: rathee-d@email.ulster.ac.uk

October 2016

Abstract.
Objective: The majority of the current approaches of connectivity based BCI systems focus on distinguishing between different motor imagery (MI) tasks. Brain regions associated with MI are anatomically close to each other, hence these BCI systems suffer from low performances. Our objective is to introduce single-trial connectivity feature based BCI system for cognition imagery (CI) based tasks wherein the associated brain regions are located relatively far away as compared to those for MI.
Approach: We implemented time-domain partial granger causality (PGC) for the estimation of the connectivity features in a BCI setting. The proposed hypothesis has been verified with two publically available datasets involving MI and CI tasks.
Main results: The results support the conclusion that connectivity based features can provide a better performance than a classical signal processing framework based on bandpower features coupled with spatial filtering for CI tasks, including word generation, subtraction, and spatial navigation. These results show for the first time that connectivity features can provide a reliable performance for imagery-based BCI system.
Significance: We show that single-trial connectivity features for mixed imagery tasks (i.e. combination of CI and MI) can outperform the features obtained by current state-of-the-art method and hence can be successfully applied for BCI applications.

Keywords: Brain-Computer Interface, Mental Imagery, Effective Connectivity, Partial Granger Causality.
Submitted to: *J. Neural Eng.*

Accepted Manuscript

List of Abbreviations

AIC Akaike Information Criterion
BCI Brain-computer Interface
BLDA Bayesian Linear Discriminant Analysis
BOLD Blood-Oxygen-Level Dependent
CI Cognitive Imagery
CSD Current Source Density
CSP Common Spatial Pattern
DTF Directed Transfer Function
ECoG Electrocorticography
ERD Event Related Desynchronization
ERS Event Related Synchronization
FDR False Discovery Rate
fMRI Functional Magnetic Resonance Imaging
MEG Magnetoencephalography
MI Motor Imagery
MVAR Multivariate Autoregressive
PDC Partial Directed Coherence
PGC Partial Granger Causality
PMC Premotor Cortex
SBIC Schwarz Bayesian Information Criterion
SMA Supplementary Motor Area

1. Introduction

Electroencephalogram (EEG) based Brain-computer interface (BCI) technology can be implemented for a wide variety of communication and control purposes, such as controlling a cursor, wheelchair or prosthesis [1], virtual keyboards [2], and navigation through virtual environments [3]. These systems assert a strong positive impact on disabled users in terms of quality of life improvement and communication with their close environment. In addition, they offer alternative means of communication for healthy users in the form of EEG controlled entertainment systems such as computer games and music control [4, 5] as well as hybrid systems [6]. Moreover, recent studies have shown that the use of BCI can affect neural plasticity during the period of training [7], and possibly contribute to the enhancement of motor rehabilitation for stoke patients [8]. Thus, the development of this technology benefits several research domains, e.g., medical and healthcare, neuro-marketing, entertainment and games, smart environments, and security.

Most of the current BCI systems based on motor imagery (MI) detection rely on specific changes in the sensorimotor EEG rhythms (event-related desynchronization (ERD) and event-related synchronization (ERS)), which involve kinesthetic imagination of a particular motor action without its actual execution [9, 10, 11]. Although promising results and achievements have been reported in the literature [12], there remain many challenges and barriers to the use of this technology reliably and effectively for the intended beneficiaries [13]. One of the probable reasons for the limitations of MI-based BCI is the use of static channel derived features (e.g. band-power (BP), autoregressive (AR), common spatial patterns (CSP) and wavelets), which may not contain information about interactions among brain regions, while it is well known that multiple brain regions dynamically interact in accomplishing a mental task. Thus, it is reasonable to assume that the connectivity of spatially distributed regions could provide additional useful discriminant features for the classification of brain responses evoked during imagery tasks.

The human brain has been divided into several areas based on their anatomical and physiological characteristics [14]. These areas are connected to each other to form functional brain networks that are dynamically employed to perform various sensorimotor and cognitive tasks. Analyzing these network connectivities (connectome) and their dynamics during various brain states may provide a better understanding of physiological mechanisms related to them. However, functional connectivity evaluations are unable to provide exact information regarding the directionality of the interaction i.e. whether the information flow is from area A to area B, or vice-versa. Effective connectivity analysis can derive better relationships between two areas of interest by providing directed interactions. Effective connectivity is therefore a relevant measure for better assessing the induced physiological variations in the brain occurring during imagery tasks.

Over the last few years, several research groups attempted to implement connectivity-based BCIs [13, 15, 16, 17, 18] using different measures to map the interactions between distinct brain regions (e.g. transfer entropy [19], coherence, directed transfer function (DTF) [20], and partial directed coherence (PDC) [21]). A recent study implemented several vector autoregressive (VAR) model based methods for the realization of MI-based BCI [17]. The empirical findings provided low classification accuracies for most of the methods when compared to BP features. In this case, the connectivity between scalp electrodes failed to provide robust distinction among various MI tasks.

Investigations of the connectome associated with imagination as well as execution of motor tasks revealed the manifestation of induced activations and information flows at various cerebral structures, including, the primary motor areas (M1), the premotor

Accepted Manuscript

cortex (PMC), and the supplementary motor areas (SMA) [22]. These cortical areas are located in close vicinity to each other making estimation of the causal interaction a challenging task, in particular at scalp level analyses, due to the volume conduction (VC) effect. Furthermore, single-trial connectivity estimations become inherently more difficult due to the availability of small amount of EEG data. These confounding factors lead to low performance of connectivity-based BCI systems. In this study, we investigate the extent to which the performance of the imagery tasks detection using features of multivariate autoregressive (MVAR) model based connectivity methods depends on the imagery task type. In this context, we evaluate the performance of single-trial connectivity features for other imagery tasks, which are collectively known as cognition imagery (CI) tasks, e.g., mathematical, navigational and language-related tasks. The rationale behind our proposed work is that CI tasks involve activations of distinct regions of the brain that are relatively far from each other compared to that associated with MI tasks.

In the present work, in line with our previous study [23], we estimate the single-trial directed functional connectivity features to elucidate the interaction among the EEG signals during various MI tasks (left hand, right hand, feet, and tongue) and CI tasks (word generation, spatial navigation, and subtraction imagery task). In particular, we implemented time-domain partial granger causality (PGC) [24] along with PDC and DTF on the two publically available datasets. Thus, we included three different methods belonging to MVAR based modelling of time-series data. As the major aim of this study is to test the hypothesis that CI based connectivity features can provide better separability than MI based features, exhaustive discussion of different connectivity measures in the analysis is out of scope of this study. Nevertheless, several previous studies have compared various connectivity measures with both simulated and empirical datasets [17, 25, 26]. Furthermore, we compared the PGC results with state-of-the-art feature extraction method for BCI, i.e. log variance after spatial filtering using CSP.

The remainder of this paper proceeds as follows: Section 2 describes the connectivity methods, the BCI datasets, and the signal processing pipeline. Next, Section 3 presents the performance analysis. Finally, the results are discussed in Section 4, and Section 5 summarizes the findings of this study.

2. Methods

2.1. Multivariate Autoregressive (MVAR) Model

An MVAR model for a set of L observed time-sampled series $x(t) \in \mathbb{R}^L$, with $1 \leq t \leq N$, N is the total number of samples, can be defined as follows [27]:

$$x(t) = \sum_{p=1}^{r} \beta^{(p)} \begin{pmatrix} x_1(t-p) \\ \vdots \\ x_L(t-p) \end{pmatrix} + \begin{pmatrix} q_1(t) \\ \vdots \\ q_L(t) \end{pmatrix} \tag{1}$$

where r is the model order and $\mathbf{q} = [q_1, \ldots, q_L]^T$ is a zero-mean white noise vector with normally distributed real-values. The auto-regression coefficient matrices β^p are given by:

$$\beta^{(p)} = \begin{pmatrix} \beta^p_{1,1} & \cdots & \beta^p_{1,L} \\ \vdots & \ddots & \vdots \\ \beta^p_{L,1} & \cdots & \beta^p_{L,L} \end{pmatrix} \tag{2}$$

where $1 \leq p \leq r$. The matrix $\beta^p \in \mathbb{R}^{L \times L}$ reveals the linear interactions among multiple time-series at the time delay p. For a reliable estimation using MVAR modeling, the total number of available data points (LN) must be significantly higher than the total number of estimated parameters ($L^2 r$) [27].

The z-domain transformation of the model can be obtained as:

$$q(z) = \left(1 - \sum_{p=1}^{r} \beta^{(p)} z^{-p}\right) x(z) \tag{3}$$

$$A(z) = 1 - \sum_{p=1}^{r} \beta^{(p)} z^{-p} \tag{4}$$

$$x(z) = A(z)^{-1} q(z) = H(z) q(z) \tag{5}$$

where $H(z)$ and $A(z)$ are the transfer function matrix and its inverse matrix, respectively.

2.2. Time-domain Partial Granger Causality Analysis

Time-domain Partial Granger Causality (PGC) is a robust form of granger causality wherein causal interactions between multivariate data can be analyzed using MVAR modeling. Unlike bivariate GC and conditional GC, it provides better estimation of the true interactions by mitigating the effect of confounding variables [24].

Let's assume three time series data including $X(t)$, $Y(t)$ and $Z(t)$. Now, to analyze the effective connectivity between $X(t)$ and $Y(t)$ (conditioned on $Z(t)$) based on PGC rules, the reduced model (i.e. inclusion of past values of the sink variable conditioned on other variables) can be defined by:

$$X(t) = \sum_{p=1}^{k} (a_{1,p} X(t-p)) + \sum_{p=1}^{k} (c_{1,p} Z(t-p)) + \tag{6}$$

Accepted Manuscript

$$\epsilon_1(t) + \epsilon_1^E(t) + \beta_1(L)\epsilon_1^L(t)$$

$$Y(t) = \sum_{p=1}^{k}(b_{1,p}Y(t-p)) + \sum_{p=1}^{k}(d_{1,p}Z(t-p)) + \quad (7)$$

$$\epsilon_2(t) + \epsilon_2^E(t) + \beta_2(L)\epsilon_2^L(t)$$

where p is the model order, $\epsilon_i(t)$ is the prediction error, $\epsilon_i^E(t)$ and $\beta(L)\epsilon_i^L(t)$ are the residual errors corresponding to exogenous (E) and latent (L) inputs, respectively. Similarly, the full model (i.e. inclusion of past values of the sink variable along with past values of source variable conditioned on rest of variables) can be defined as:

$$X(t) = \sum_{p=1}^{k}(a_{2,p}X(t-p)) + \sum_{p=1}^{k}(b_{2,p}Y(t-p)) + \quad (8)$$

$$\sum_{p=1}^{k}(c_{2,p}Z(t-p)) + \epsilon_3(t) + \epsilon_3^E(t) + \beta_3(L)\epsilon_3^L(t)$$

$$Y(t) = \sum_{p=1}^{k}(d_{2,p}X(t-p)) + \sum_{p=1}^{k}(e_{2,p}Y(t-p)) + \quad (9)$$

$$\sum_{p=1}^{k}(f_{2,p}Z(t-p)) + \epsilon_4(t) + \epsilon_4^E(t) + \beta_4(L)\epsilon_4^L(t)$$

The collective prediction errors can be taken from the previous equations, and are represented as:

$$\mu_i = \epsilon_i(t) + \epsilon_i^E(t) + \beta_i(L)\epsilon_i^L(t) \quad (10)$$

with $1 \leq i \leq 4$.

The prediction error covariance matrix for the reduced model can be generated as:

$$R = \begin{bmatrix} \text{var}(\mu_1(t)) & \text{cov}(\mu_1(t), \mu_2(t)) \\ \text{cov}(\mu_2(t), \mu_1(t)) & \text{var}(\mu_2(t)) \end{bmatrix} \quad (11)$$

Likewise, the covariance matrix for the full model as:

$$L = \begin{bmatrix} \text{var}(\mu_3(t)) & \text{cov}(\mu_3(t), \mu_4(t)) \\ \text{cov}(\mu_4(t), \mu_3(t)) & \text{var}(\mu_4(t)) \end{bmatrix} \quad (12)$$

The PGC indices can be calculated by taking the log ratio between partial variance of prediction error generated by the reduced model and partial variance of prediction error generated by the full model. The following two equations provide the PGC indices for $Y(t)$ causing $X(t)$ and vice-versa, respectively:

$$\mathcal{G}_{Y\to X|Z} = ln(\frac{R_{1,1} - R_{1,2}R_{2,2}^{-1}R_{2,1}}{L_{1,1} - L_{1,2}L_{2,2}^{-1}L_{2,1}}) \quad (13)$$

$$\mathcal{G}_{X\to Y|Z} = ln(\frac{R_{2,2} - R_{2,1}R_{1,1}^{-1}R_{1,2}}{L_{2,2} - L_{2,1}L_{1,1}^{-1}L_{1,2}}) \quad (14)$$

2.3. PDC and DTF

The PDC and DTF are based on the concept of MVAR model, and their values can be obtained according to the following equations:

$$\mathbf{PDC}_{j\to i}(z) = \frac{|A_{ij}(z)|}{\sqrt{A_{:j}^T(z)A_{:j}(z)}} \quad (15)$$

$$\mathbf{DTF}_{j\to i}(z) = \frac{|H_{ij}(z)|}{\sqrt{H_{i:}(z)H_{i:}^T(z)}} \quad (16)$$

2.4. Common Spatial Pattern (CSP) Filtering

This method involves the extraction of log variance features in particular frequency bands after spatial filtering using common spatial patterns. The CSP method is highly successful in calculating spatial filters for detecting ERD/ERS during MI tasks. This two-step algorithm involves band-pass filtering of the scalp EEG signals followed by spatial filtering based on linear transformations that can be represented as:

$$Z_i = W^T E_i \quad (17)$$

where $E_i \in \mathbb{R}^{c\times N}$ denotes the EEG measurement from the band-pass filter of ith trial/segment; $Z_i \in \mathbb{R}^{c\times N}$ denotes the spatial filtered EEG data; $W^T \in \mathbb{R}^{c\times c}$ denotes the transpose of CSP projection matrix; c is the number of channels; and N is the number of time-samples in one trial/segment.

The CSP algorithm computes the transformation matrix W to yield features whose variances are optimal for discriminating two classes of EEG measurements. The m pair of CSP features (i.e. log variance of the spatial filtered EEG signal) are then given by

$$v_i = \log \frac{diag\left(\overline{Z}_i\overline{Z}_i^T\right)}{\text{Tr}\left(\overline{Z}_i\overline{Z}_i^T\right)} \quad (18)$$

where $v_i \in \mathbb{R}^{2m}$; $\overline{Z}_i$ represents the spatial filtered data with first m and the last m columns of Z_i; $diag\,(\cdot)$ and $\text{Tr}\,(\cdot)$ return the diagonal and the sum of the diagonal elements in the square matrix, respectively.

2.5. EEG Dataset-1

The BCI Competition IV dataset 2A has been analysed for investigating causal interaction during various MI tasks [28]. The dataset comprised of EEG signals acquired from nine healthy participants using a cue-based paradigm (see Figure 1.A) during two sessions on different days. The MI tasks include four different classes: left hand MI (class 1), right hand MI (class 2), both feet MI (class 3), and tongue MI (class 4). Each data acquisition session consists of 6 runs where each run comprised of 48 trials (12 trials for each class). Thus the complete study involved 576 trials from both sessions of the dataset. The total trial length is 7.5 s with variable inter-trial duration. The data were acquired from 25 channels (22 EEG channels

Accepted Manuscript

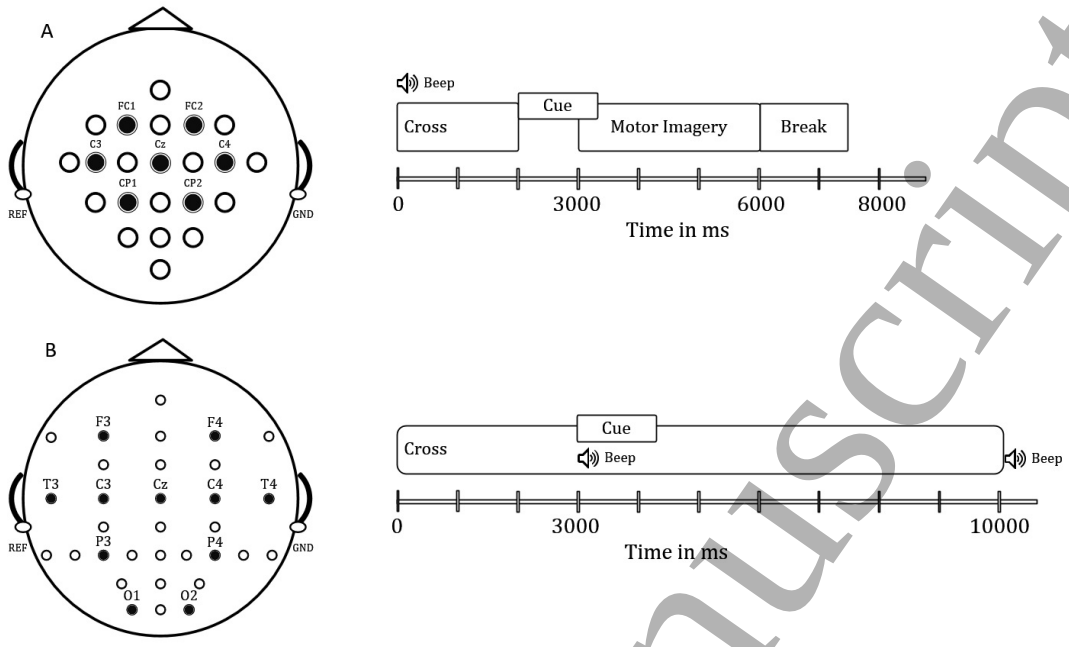


Figure 1. (A) EEG channel montage and timing of the experimental paradigm for dataset-1.Trial duration is 6 s with a break of 1.5 s between trials. Trial start is indicated with an acoustic beep and appearance of the fixation symbol. After 2 s, participants are cued for the MI task to be performed. **(B)** EEG channel montage and timing of the experimental paradigm for dataset-2. Trial duration is 10 s with a break of 2.5-3.5 s between trials. Trial start is indicated with the appearance of the fixation symbol, and the cue remains from 3 to 4.25 s. The appearance of the cue is indicated with acoustic beep.

along with three monopolar EOG channels) with a sampling frequency of 250 Hz and bandpass filtered between 0.5 Hz to 100 Hz (notch filter at 50 Hz). Reference and ground were placed at the left and right mastoid, respectively.

2.6. EEG Dataset-2

The mental imagery-based BCI dataset has been previously analysed [29] for pairwise comparison of BCI performances during various imagery tasks. The dataset comprised of EEG signals acquired from nine participants with severe motor disabilities using a cue-based paradigm (see Figure 1.B) during two sessions on different days. Details of the participants are summarized in Table A1 of Appendix A. For further exploration, readers are referred to the original study [29]. The mental tasks include five different classes: word generation imagery (class 1), mental subtraction imagery (class 2), spatial navigation imagery (class 3), right hand MI (class 4), and both feet MI (class 5). Each data acquisition session consists of 8 runs resulting in 200 trials (40 trials for each class). Thus the complete study involved 400 trials from both sessions of the dataset. The total trial length is 10 s with variable inter-trial duration. The data were acquired from 30 EEG channels with a sampling frequency of 256 Hz and bandpass filtered between 0.5 to 100 Hz (notch filter at 50 Hz). Reference and ground were placed at the left and right mastoid, respectively. This dataset is publically available at http://bnci-horizon-2020.eu/database/data-sets.

2.7. Data pre-processing and analysis

MVAR based connectivity methods are sensitive to volume conduction, thus the first step of data pre-processing involved the estimation of current source densities (CSD) for both datasets using spherical spline method [30]. A recent study showed that CSD estimation improves the interpretability of connectivity results by reducing the amount of spurious interactions [31]. Next, the channels that are noisy or severely contaminated with artefacts were interpolated for dataset-2 followed by exclusion of the bad trials. The information regarding the faulty EEG channels and noisy trials is provided in [29]. However, for dataset-1, no such information is present, so all the available trials were considered for the analysis. For the three connectivity methods, the features were generated for delta (δ) (1-4 Hz), theta (θ) (4-8 Hz), alpha (α) (8-12 Hz), beta (β) (13-25 Hz), lower gamma (γ) (25-40 Hz), and wide-band (1-40 Hz) frequency bands whereas for CSP we included a wideband approach (8-30 Hz) as incorporated in the original research article [29]. Bandpass filtering was applied by employing an 8th order, zero-phase forward and backward bandpass Butterworth filter. For further

Accepted Manuscript

analysis, seven channels (C3, FC1, CP1, Cz, FC2, CP2, and C4) related to the motor-cortical areas are selected for dataset-1. Likewise, for dataset-2, 11 channels (F3, F4, T3, C3, Cz, C4, T4, P3, P4, O1 and O2) were selected for considering the frontal, temporal, central, parietal, and occipital areas of the cortex. The motive behind this implementation is to select a common network for all imagery classes.

Furthermore, the data related to the imagery tasks were extracted from each trial and for both datasets (i.e. 3000 ms to 6000 ms for dataset-1 and 3000 ms to 10000 ms for dataset-2). The extracted data was segmented using a sliding window starting from 3000 ms to the end of the trial. For dataset-1, we extracted segments of length 1000 ms with 500 ms overlap while for dataset-2, the segment size was 2000 ms with an overlap of 500 ms. Thus, we obtained 5 and 11 segments from each trial for dataset-1 and dataset-2, respectively. The windowing strategy can effectively reduce the probability of spurious causal effects due to non-stationarity in EEG signal [32]. Longer segment may led to better model fitting, but they are frequently non-stationary. The implemented segment lengths provided a better trade-off between satisfactory model fitting and local stationarity. Moreover, to counter the issues related to inter-segment variations and non-stationarity, processes of detrending and demeaning of the data were performed wherein the average was subtracted from each segment separately along with division of each segment by the standard deviation.

The coefficients of the MVAR model for multi-trial data were estimated using the LWR algorithm [33]. The application of MVAR modeling for connectivity analysis requires estimation of the model order (MO) i.e., the number of time-lags. This process is highly crucial for the correct estimation of the connectivity networks. If the MO is too high, overfitting occurs and may introduce spurious links in the network. If the MO is too low, the fitted model may fail to capture the essential dynamics of the data resulting in exclusion of the existing links [34]. A previous study examined empirically the effect of incorrect MO selection on estimated functional connectivity networks with both simulated and real datasets [35]. Furthermore, the study compared the performances of various criteria available for estimating the correct model order and it suggested a cautious usage of these criteria. In this study, we implemented the Schwarz bayesian information criterion (SBIC) [36] and the Akaike information criterion (AIC) [37] for estimating the optimal value of the model order p. The expressions for these two methods are given as follows:

$$AIC(p) = \log[\det(\Sigma)] + \frac{2pL^2}{N} \tag{19}$$

$$SBIC(p) = \log[\det(\Sigma)] + \log(N)\frac{pL^2}{N} \tag{20}$$

Where Σ is the estimated noise covariance matrix, L is the number of EEG channels, and N is the number of data samples. Furthermore, the model order was selected for the measure provided minimum values. During the analysis of optimum model order, we found 8 as the best model order for most of the segments (more than 75%) for dataset-1, and 10 for majority of cases (more than 85%) for dataset-2. Hence to make the analysis unbiased, we set these model order values for the complete analysis. Furthermore, we implemented two different techniques to confirm the legitimacy of applied regression models. Durbin-Watson whiteness test [38] has been used for approximating whiteness of uncorrelated residuals. The model consistency assesses the proportion of the correlation structure that is shared by the real data and 'simulated' data generated from the MVAR model. A higher consistency value (i.e., ≥ 80) provide a confirmatory indication for the rejection of null hypothesis. Moreover, validation of the model was confirmed using the Ding method [32] by checking the consistency of the correlation structure. The Ding consistency test provided a higher value (nearly equal to 1), which shows that the selected MVAR model has effectively predicted the time series.

For CSP based analysis, 2 pairs of components (m=2) were selected for binary classification tasks. For computational assignments related to data processing, we used MATLAB (V8.1) and for estimation of single trial connectivity and CSP features, we implemented the algorithms with in-house scripts using functions from GCCA (V2.9) toolbox [39] and HERMES toolbox [40].

2.8. Single-trial connectivity classification

Single-trial connectivity feature based on 16 binary classifications were performed for both datasets. The binary classifications included left-right (L-R), left-feet (L-F), left-tongue (L-T), right-feet (R-F), right-tongue (R-T), and feet-tongue (F-T) for dataset-1 and word-subtraction (W-S), word-navigation (W-N), word-hand (W-H), word-feet (W-F), subtraction-navigation (S-N), subtraction-hand (S-H), subtraction-feet (S-F), navigation-hand (N-H), navigation-feet (N-F), and hand-feet (H-F) for dataset-2. A Bayesian linear discriminant analysis (BLDA) classifier was trained on the feature sets from the data of one session (training set) while the evaluation was performed on the feature set generated from the data of the second session (evaluation set). For the selection of the best time segment and the optimal number of features, cross validation was performed during the training phase. Consequently, the computational load

Accepted Manuscript

during the evaluation phase was efficiently reduced. We measure the classification performance with the area under the ROC curve (AUC) for each binary classification [41]. Furthermore, to evaluate the best connectivity measures for each pairwise binary classification task, Pearson's correlation coefficients were estimated using data from both sessions. The absolute values of class correlation coefficients $|r(i,j,s)|$ and their significance estimators ($p-values$: $p(i,j,s)$) of each connectivity index with their corresponding dummy class labels l were estimated for each pair of classes, all segments, all subjects, and both datasets. The dummy class label is a numeric value that indicates whether a trial contains class-1 ($l = 1$) or class-2 ($l = -1$) of the binary classification task. For instance, the dummy label l can be set to +1 for left-hand MI (class-1) and -1 for right-hand MI (class-1) or vice-versa as well in case of left-hand versus right-hand binary classification task. Thus, $|r(i,j,s)|$ indicates how strongly a connectivity measure can distinguish between both types of imagery classes for a channel pair (i,j) during a time segment s.

3. Results

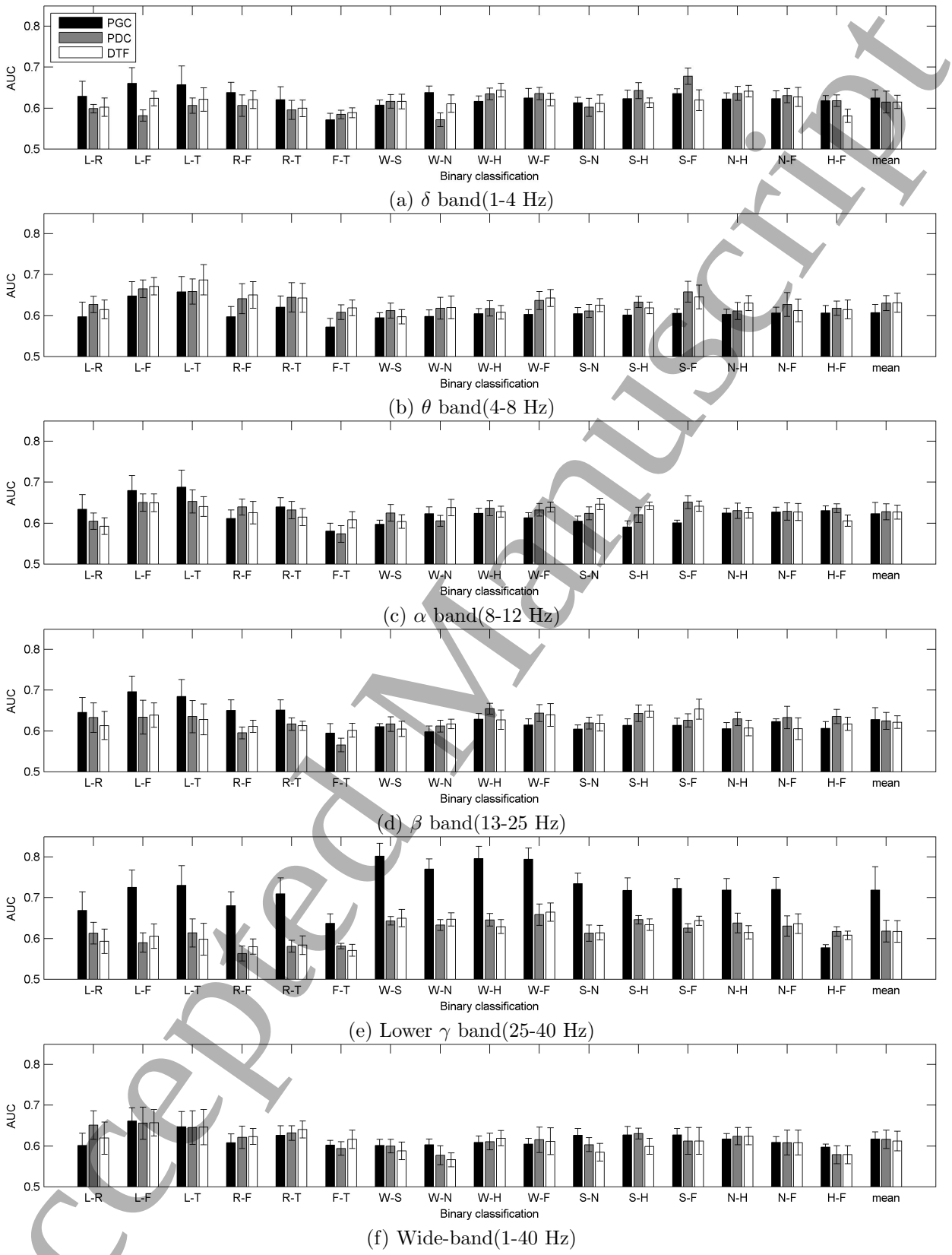
3.1. Performance comparison of PGC, PDC, and DTF for different frequency bands (δ, θ, α, β, lower γ, and wide-band).

Figure 2 presents the mean AUC values (across nine subjects) for PGC, PDC, and DTF for δ band (1-4 Hz), θ band (4-8 Hz), α band (8-12 Hz), β band (13-25 Hz), lower γ band (25-40 Hz), and wide-band (1-40 Hz) in 16 pairwise binary classification tasks involving both datasets. The first six comparisons (from left-hand side) belong to dataset-1 while the next ten comparisons belong to dataset-2. Furthermore, we compared the grand mean AUCs (across 16 binary classification tasks) for the three methods and found no statistically significant differences in δ, θ, α, β, and wide-band (1-40 Hz). Also, the grand mean AUC values are lower than 0.65 for all three methods for these frequency bands. Thus, the results clearly showed week discrimination power of connectivity based features for binary classification tasks. However, interestingly, PGC performed better than PDC and DTF for lower γ band. In this frequency band, the grand mean AUCs($\pm$ SD) for PGC, PDC, and DTF are 0.72($\pm$ 0.06), 0.62($\pm$ 0.03), and 0.62($\pm$ 0.3), respectively. In 15 out of the 16 binary classification tasks, PGC provided higher AUC values than PDC and DTF. Thus, the lower γ band with PGC provided the best classification results for both MI and CI binary tasks. The overall statistical significance analysis showed PGC in lower γ band performed significantly better than all the other combinations ($p < 0.025$, FDR corrected for multiple comparisons).

3.2. Performance comparison of PGC and CSP

The performances corresponding to PGC (lower γ band) and a state-of-the-art signal processing technique (i.e. CSP + log variance) are presented for both databases in Tables 1 and 2. For simplified presentation, we denote the PGC method by M_1 , and CSP algorithm by M_2 in the results. For dataset-1, the statistical analysis showed significantly better performance of CSP for all six binary classification tasks ($p < 0.05$). For dataset-2, the mean AUCs of eight out of ten binary classification tasks are better in case of PGC however we found statistically significant differences in two cases only i.e. W-S ($p = 0.039$) and W-N ($p = 0.01$), where improvements of 14% and 10% in mean AUC(%) are obtained, respectively. Furthermore, we compared the performance of different MI pairs for PGC results. The comparative analysis for dataset-1 (with FDR corrected p-value (p=0.026)) showed some interesting results wherein both L-F ($p = 0.004$) and L-T pair ($p = 0.023$) provided significantly better AUCs (%) as compared to F-T. Besides, the MI task pair F-T had the worst overall performance whereas L-F and L-T performed best among the six binary classifications. For dataset-2, group statistical analysis (with FDR corrected $p - value = 0.025$) showed significantly better performance of PGC during W-S, W-H, and W-F as compared to S-N (p=0.021, p=0.011, p=0.019), S-H (p=0.007, p=0.003, p=0.007), S-F (p=0.007, p=0.004, p=0.004), N-H (p=0.004, p=0.007, p=0.007), N-F (p=0.011, p=0.003, p=0.007), and H-F (p=0.004, p=0.007, p=0.004). Additionally, the outcome for W-H (p=0.011) is significantly better than W-N. In terms of average AUC(%) values, W-H and W-S performed the best, and H-F performed worse than all ten binary classification tasks. This analysis followed the most practical way of classification for a BCI setting i.e. training and evaluation with data from different sessions. However, to further explore the performances of the two competing methods and to evaluate the variations between two sessions, we estimated the 10-fold cross validation results for the data of session-1 and session-2 separately, and for the combined data of session-1 and session-2. The mean AUCs across subjects for session-1, session-2, and combined data are presented for both datasets in Figure 3(a) for CSP method and in Figure 3(b) for PGC methods. Although we found a slight increase in the grand mean AUC values for session-2 as compared to session-1, the difference is not statistically significant. Moreover, for all the three cases, CSP outperformed PGC for dataset-1 involving MI tasks while the latter outperformed CSP for dataset-2 involving both CI and MI tasks.

Accepted Manuscript



(a) δ band(1-4 Hz)

(b) θ band(4-8 Hz)

(c) α band(8-12 Hz)

(d) β band(13-25 Hz)

(e) Lower γ band(25-40 Hz)

(f) Wide-band(1-40 Hz)

Figure 2. Mean AUC measures for PGC, PDC, and DTF with 16 pairwise comparisons for: **(a)** δ band(1-4 Hz),**(b)** θ band(4-8 Hz), **(c)** α band(8-12 Hz), **(d)** β band(13-25 Hz), **(e)** lower γ band(25-40 Hz), and **(f)** wide-band(1-40 Hz). The first six comparisons from left side belong to dataset-1 while the rest belong to dataset-2. The error bars represent the standard error across subjects.

Accepted Manuscript

Furthermore, we estimated the outcomes by combining PGC and CSP features together. BLDA classifier was trained on the collective feature set of session-1 data and evaluated with combined feature-set of session-2 data. The mean AUC values (across all subjects) generated with CSP, PGC, and the combined features are presented for dataset-1 and dataset-2 in Figure 4 and Figure 5, respectively. Here, we compared the performance of the combined feature set with the best approach out of PGC and CSP for each dataset. For both datasets, we found no statistically significant difference during the comparisons. However, the difference between the grand mean AUC values with the combined feature set and PGC is higher for dataset-2 as compared to the difference between combined and CSP features.

3.3. Class-correlation analysis of the connectivity features

The absolute values of the class-correlation coefficients were averaged across all subjects and both sessions for each dataset separately. Next, for each pairwise comparison, the averaged values of these estimates were normalized between 0 and 1 to evaluate the significance of different connectivity indices within each binary classification task. Finally, the class-correlation matrices were generated using the normalized estimates. They are presented in Figure 6(a) and Figure 6(b) for six pairwise comparisons of dataset-1 and ten pairwise comparisons of dataset-2, respectively. Each matrix element represents directional connectivity from the ith column (source) to the jth row (sink). Table 3 illustrates the significant connectivity indices with $r > 0.5$ for various binary classification tasks for dataset-2. Furthermore, we scrutinize the indices with majority voting criterion i.e., selected only those features which maintain the r-value threshold (i.e. 0.5) for at least 5 subjects out of 9. We found no connectivity feature above the set threshold and the voting criterion for dataset-1. However, we can assess the most contributing connectivity features for dataset-1 with Figure 6(a), for instance, CP1→C4, C3→FC1, and C4→Cz for L-R, CP1→C4, C3→C4, and C3→CP1 for L-F, C3→C4, Cz→C4, and C3→CP1 for L-T, C3→Cz, and C3→FC1 for R-F, C3→FC1, and Cz→C3 for R-T, and finally CP1→FC1, Cz→CP2, Cz→C4, Cz→CP1,and Cz→FC1 for F-T. For dataset-2, we found significant association of frontal scalp area with W-S, W-N, W-H, and W-F binary classification tasks. Likewise, parietal area is more associated with W-N, W-H, and W-F while occipital area seems to be more activated during N-H, N-F, and S-F. Interestingly, central area electrodes contributed for almost all the binary classification tasks. Temporal lobe is majorly contributed in W-N, S-N, and W-F. However, we found no significant connectivity indices for H-F binary classification task.

4. Discussion

Most of the connectivity based single-trial BCI studies focus on MI related modulations in the motor cortex functional network. Investigations of the brain connectome related to imagination and execution of motor tasks revealed the exhibition of induced activations and information flows at various cerebral structures, incorporating the primary motor areas (M1), the premotor cortex (PMC), and the supplementary motor areas (SMA) [42]. These cortical areas are located in close vicinity to each other making discriminative estimation of the causal interaction a challenging task, in particular at scalp level analyses. Thus, the performances of the MI-based BCIs using connectivity features are typically low. In this study, we presented an alternative solution to overcome this confounding factor by employing cognition imagery (CI) features, including word generation, spatial navigation, and mathematical tasks. The rationale behind this hypothesis is that these imagery tasks activate distinct brain regions, which are relatively far from each other.

In this paper, we have considered PGC, DTF, and PDC to generate effective connectivity features in six frequency bands i.e., δ, θ, α, β, lower γ, and wide-band using two publically available imagery BCI datasets. Next, to obtain the best performing method and optimum frequency band, we compared the results for 16 binary classification tasks involving MI and CI tasks. Moreover, we compared the performances of PGC method and CSP algorithm by evaluating the results of binary classification tasks for both datasets. Our analysis yielded four major findings. First, we obtained higher discriminative connectivity features in lower γ band as compared to the rest of the canonical frequency bands (i.e. δ, θ, α, β) and wide-band. Second, the performance of PGC is significantly better than PDC and DTF in lower γ band. Third, and more importantly, the PGC and CSP comparative analysis showed significantly high improvement in classification performances with effective connectivity based features for mixed imagery tasks (i.e., CI versus MI) as compared to MI tasks. Finally, the post-hoc analysis showed higher values of class correlation coefficients for connectivity features in binary classification tasks involving word generation, spatial navigation and subtraction task as compared to MI tasks.

Our results illustrate that lower γ frequency range holds greater information regarding cortical

Accepted Manuscript

Table 1. Comparison of the AUC (%) between PGC in lower γ (M_1) and the log-variance of CSP features (M_2) approach for the dataset 1.

Subjects	1		2		3		4		5		6		7		8		9		Mean±SD	
	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2
L-R	73	92	51	67	82	99	56	79	63	66	54	73	59	79	74	99	91	97	67 ± 14	83 ± 14
L-F	88	99	61	78	81	93	66	83	60	63	61	68	68	95	70	85	96	98	73 ± 13	85 ± 13
L-T	94	99	52	72	81	96	67	85	64	74	60	74	75	92	70	97	95	99	73 ± 15	88 ± 11
R-F	88	99	69	83	81	96	64	92	62	67	59	69	68	98	61	91	60	66	68 ± 10	84 ± 14
R-T	93	100	60	72	88	99	65	89	64	72	61	71	71	95	72	92	65	86	71 ± 12	86 ± 12
F-T	61	67	60	78	76	74	58	75	58	66	58	69	65	77	63	86	74	82	64 ± 7	75 ± 7

Table 2. Comparison of the AUC (%) between PGC in lower γ (M_1) and the log-variance of CSP features (M_2) approach for the dataset 2.

Subjects	1		2		3		4		5		6		7		8		9		Mean±SD	
	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2	M_1	M_2
W-S	63	79	73	62	89	70	82	69	87	76	72	62	85	66	91	56	80	57	80 ± 9	66 ± 8
W-N	58	58	81	69	79	79	73	71	79	67	80	71	81	68	80	64	81	58	77 ± 8	67 ± 7
W-H	57	56	82	66	82	73	75	83	80	96	86	63	81	56	86	68	86	83	80 ± 9	71 ± 14
W-F	61	68	79	70	78	65	76	80	81	83	86	73	79	67	88	66	88	76	79 ± 8	72 ± 7
S-N	59	65	80	66	71	73	64	66	76	69	72	55	81	77	78	62	80	60	73 ± 8	66 ± 7
S-H	51	90	64	57	80	65	70	75	74	82	77	63	76	65	80	76	74	85	71 ± 9	73 ± 11
S-F	57	89	66	58	75	66	72	71	73	76	74	52	78	64	82	81	74	63	72 ± 7	69 ± 12
N-H	58	62	72	53	60	63	71	81	74	77	70	55	80	60	82	63	79	67	72 ± 8	65 ± 9
N-F	57	68	76	60	59	56	73	89	77	60	69	75	80	60	81	59	76	76	72 ± 9	67 ± 11
H-F	58	58	60	54	61	57	54	60	57	77	59	60	58	52	57	52	56	65	58 ± 2	60 ± 8

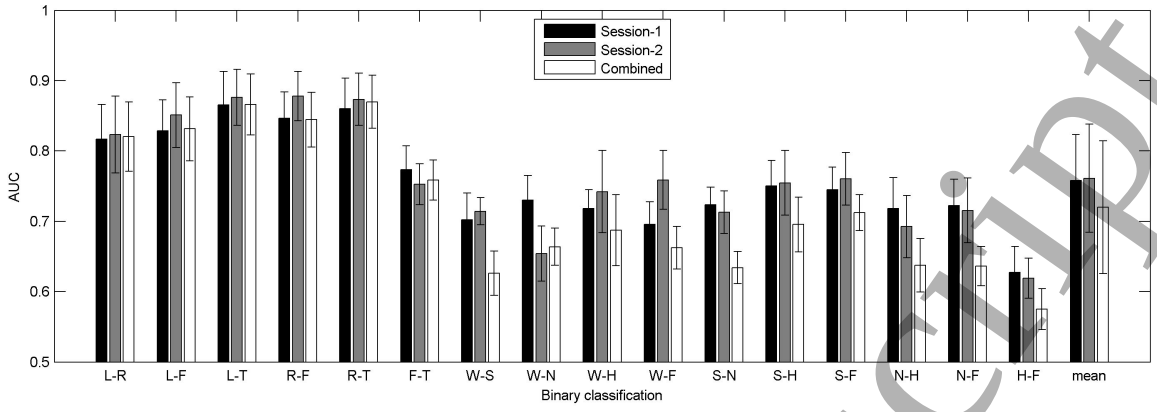
Table 3. Directed connectivities with statistically significant class-correlation coefficients ($r > 0.5$ and $p < 0.00045$, Bonferroni corrected) for each binary classification task on dataset-2.

	Relevant Connectivities
W-S	(P3→F4), (C4→C3), (C3→Cz), (C4↔Cz) (C3→C4), (F4→T4)
W-N	(C3→F4), (P3→F4), (T4→C4), (C4→T4) (P3→T4), (F4→P3), (C4→P3)
W-H	(Cz→F3), (Cz→F4), (P3→F4), (C3→Cz) (C4→Cz), (F3→C4), (T4↔C4), (C3→P3) (C3→P4), (C3→O1), (C3→O2)
W-F	(P3→F3), (F4→T3), (F4→C3), (P3→C3) (F3→C4), (P3→C4), (P3→T4), (P3→P4) (F4→O1)
S-N	(C4↔Cz), (T4↔C4), (P3↔T4)
S-H	(Cz→F3), (C4→C3), (T4→C3), (P4→C3) (F3→Cz), (C3→O1)
S-F	(Cz↔F3), (C4→C3), (T4→Cz), (C3→O1) (C3→O2), (Cz→O2)
N-H	(O2→F4), (P3→C3), (T4→Cz), (P3→Cz) (O2→C4), (O2→O1)
N-F	(O2→F4), (O2→C4), (T4→P3), (O2→P3) (O2↔O1)
H-F	Nil

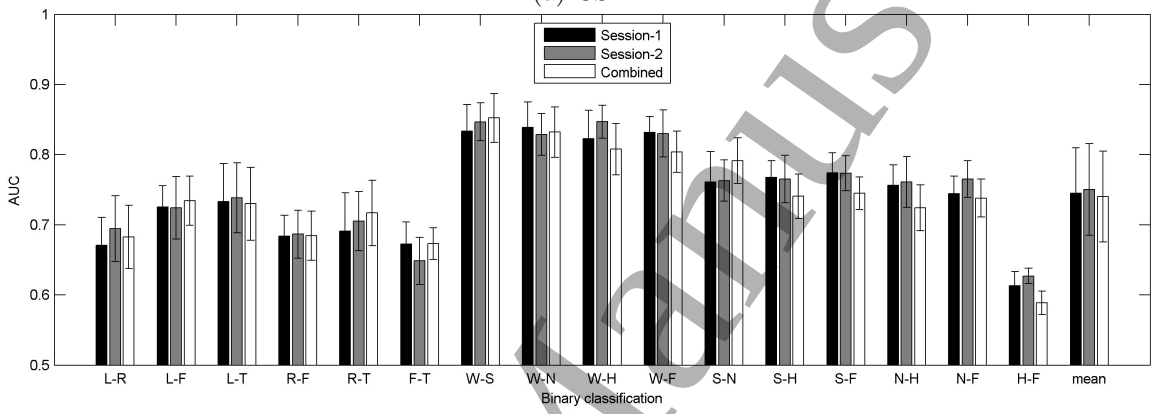
interactions during various brain imagery tasks than α and β frequency ranges. These findings are consistent with several prior studies. A transfer entropy based brain connectivity study suggested strongest modulation in γ band during MI task [16]. Similarly, detection of high frequency activations in motor cortex during motor execution and imagery was reported using invasive modalities including ECoG [43] and non-invasive recording methods including fMRI [44] during motor imagery. Furthermore, two recent studies presented high correlation between higher frequency range of ECoG signal with spatially focal BOLD peaks in primary sensorimotor areas [45] and in pre and post-central areas (i.e., covering motor area) [46]. However, it is evident that the signal to noise ratio is a significant issue when dealing with high frequencies in EEG signals. To mitigate the effect of muscular noise, we implemented CSD estimation as a pre-processing method [30]. A similar study involving various MVAR based connectivity measures reported low accuracies of PDC and DTF for MI-based BCI [17]. Moreover, the theoretical formulation of PDC and DTF is based on standard GC measure while the PGC method enhances the efficiency of conditional GC measure by mitigating the effect of confounding factors (i.e. latent variables and exogenous inputs) using a concept similar to partial correlation [24]. A recent study provided evidence supporting the higher accuracy and consistency with PGC as compared to conditional GC and PDC approaches in identifying causal connectivity of neural circuits with both simulated and empirical datasets [47].

We validated the performance improvement of single-trial connectivity-feature based BCI system by considering mental imagery of word generation,

Accepted Manuscript



(a) CSP



(b) PGC

Figure 3. Mean AUC measures (across subjects) estimated by 10-fold cross validation classification with session-1, session-2, and the combined data involving 16 pairwise comparisons for: **(a)** CSP method, and **(b)** PGC method. The first six comparisons from left side belong to dataset-1 while the rest belong to dataset-2. The error bars represent the standard error across subjects.

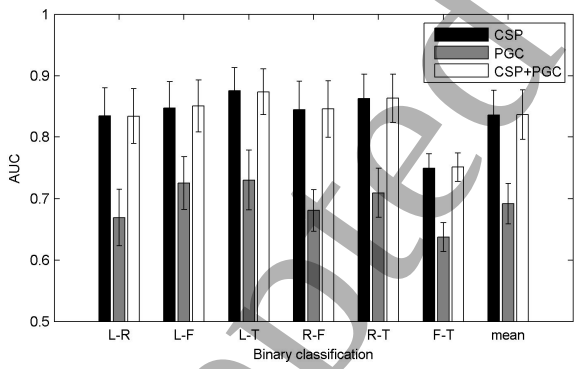


Figure 4. Mean AUC measures (across subjects) estimated for dataset-1 involving six pairwise binary comparisons with only CSP features, PGC features, and with combination of CSP and PGC features. The error bars represent the standard error across subjects.

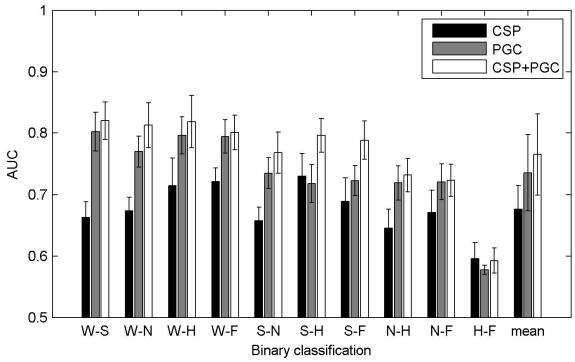


Figure 5. Mean AUC measures (across subjects) estimated for dataset-2 involving ten pairwise binary comparisons with only CSP features, PGC features, and with combination of CSP and PGC features. The error bars represent the standard error across subjects.

subtraction, and spatial navigation tasks instead of MI tasks (i.e. Left hand, right hand, feet, and tongue). For these imagery tasks, PGC significantly outperformed a state-of-the-art method. However, connectivity features related to MI tasks failed to provide high separability between different MI tasks. We also found higher AUCs for mixed imagery tasks (i.e. adding CI and MI) as compared to single imagery type tasks (i.e. either MI or CI only) with both CSP and PGC features (see Table 2) which is consistent

Accepted Manuscript

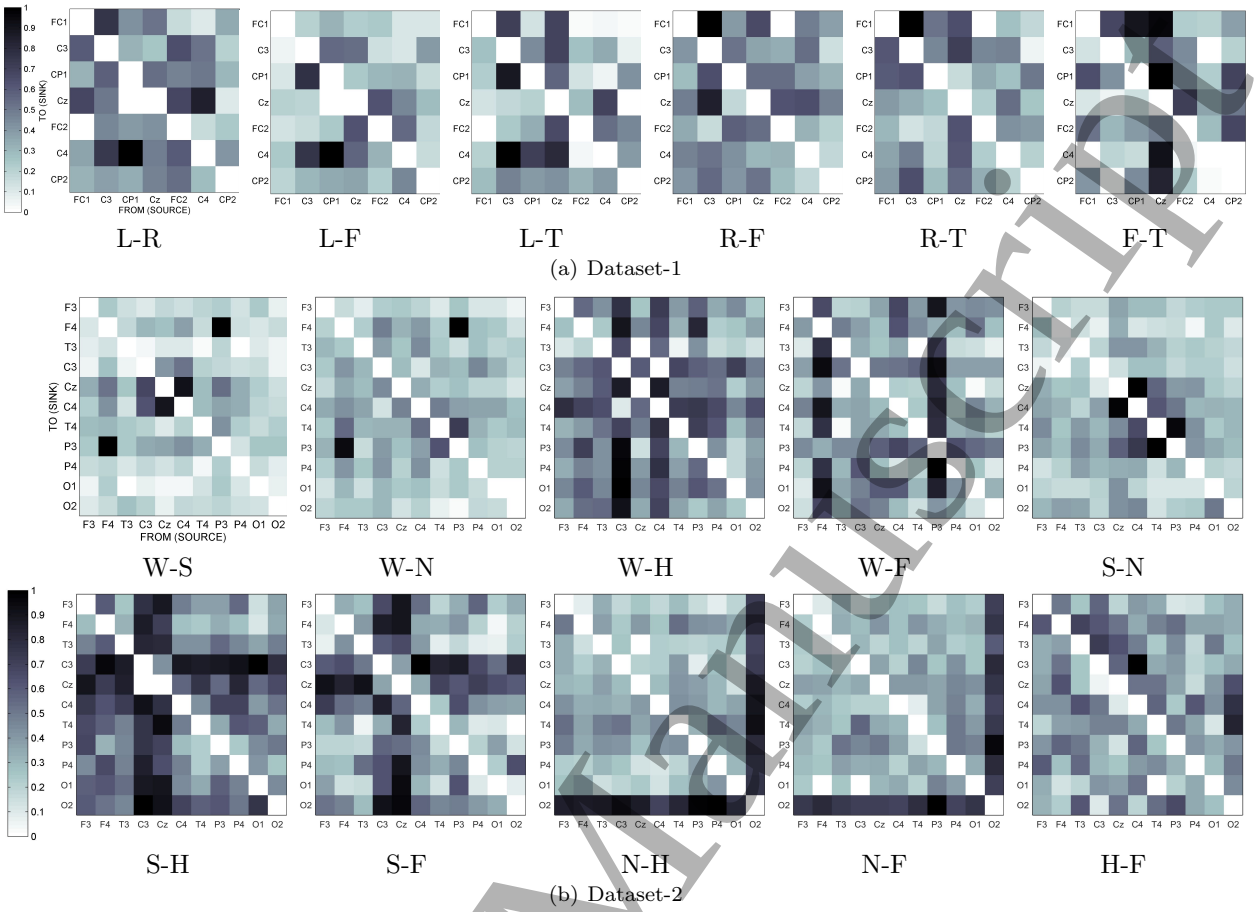


Figure 6. Class-correlation matrices for each pairwise comparison for both datasets. The absolute values of the class correlation coefficients for all subjects and sessions were averaged and normalised between 0 and 1 for each binary classification task pair separately. Each matrix element represents directional connectivity feature from the ith column (source) to the jth row (sink).

with findings of several previous studies [29, 48, 49]. However, as per our knowledge, this is the first study to compare the CI, MI, and the combined tasks pairs for connectivity features. There are several limitations to this study. Firstly, dataset-2 consists of small number of trials, and hence future work may involve validation of the proposed hypothesis with a larger number of trials and also with other non-invasive imaging modalities, such as magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI). Secondly, the linearly mixed noise due to volume conduction is the most challenging problem in scalp-based connectivity analysis. Although we computed CSDs from the raw EEG signal to reduce the amount of spurious interactions before estimating the connectivity matrices, several other methods, including ICA decomposition [50] and source localization [51] may provide more robust estimations in the face of higher computational complexity. Thirdly, although we estimated the class-correlation coefficients for all binary classification tasks to find significant connectivity features, the interpretation of neurophysiological mechanisms based on these findings is not absolute and requires further analysis in the source space. The EEG data acquired from the sensors can not be directly related to the underlying neuronal sources as the brain signal undergoes several spatio-temporal transformation before reaching the scalp surface [52, 53]. Thus, for studying actual physiological mechanisms the sensor space data must be source localized before estimation of connectivity features, however, it requires to solve the ill-posed EEG inverse problem. Nevertheless, future work may involve in-depth study of the mixed imagery related physiological interactions at source level.

A major challenge in BCI is to increase the information transfer rate (ITR) as the brain computer interaction competes with other modalities (e.g. eyetracking) particularly for people who are not completely locked-in. The increase of the number of commands, i.e. the increase of the imagery tasks, that can be accurately selected at any moment can improve the ITR. This study has shown that while the performance of features based on effective connectivity

Accepted Manuscript

remains below the existing methods for MI tasks, this type of features provides a robust performance for other types of imagery tasks such as CI tasks. Finally, because the choice of the easiest imagery tasks may be subject dependent, the results show that the choice of the feature extraction method must be based on the selected imagery tasks.

5. Conclusion

Discrimination of brain evoked responses corresponding to imagery tasks is a challenge for both the person who has to imagine a specific task, and the signal processing methods that have to extract robust discriminant features. Most of the previous studies have focused on motor-imagery tasks (e.g., left vs. right) by using features based on the bandpower or log-variance of bandpassed signals after spatial filtering. While this approach stays successful for the classification of motor-imagery, we have shown that the use of features based on effective connectivity using Partial Granger Causality in the gamma band can lead to more reliable performance in certain imagery tasks: word generation, numerical subtraction, and spatial navigation tasks. The results showed that the performance of the proposed system is superior to what can be achieved with a state-of-the-art method for motor imagery. Finally, the results also suggest that the combination of different methods opens the scope for improved brain-computer interface based on the detection of imagery tasks by increasing the number of possible classes. Further work will deal with the combination of methods to increase the number of commands in BCI.

Acknowledgment

D.R. was supported by Ulster University Vice Chancellor's Research scholarship (VCRS). G.P. and H.C. were supported by the Northern Ireland Functional Brain Mapping Facility project (1303/101154803), funded by InvestNI and the University of Ulster. G.P. was also supported by the UKIERI DST Thematic Partnership project "A BCI operated hand exoskeleton based neuro-rehabilitation system" (UKIERI-DST-2013-14/126).

Appendix A.

The demographic details of the participants for dataset-2 are summarized in Table A1. For further exploration, the readers are referred to the original study [29].

Table A1. Participant details: The ID, age in years, months since occurrence and the type of event are shown for each individual of dataset-2 [29].

ID	Age	Month	Event
P1	42	6	Locked-in syndrome due to brainstem stroke
P2	31	5	Locked-in syndrome due to brainstem stroke
P3	33	2	Spinal cord injury C5, ASIA C
P4	40	255	Spinal cord injury C5, ASIA A
P5	57	5	Hemorrhagic stroke (HS) in left hemisphere
P6	43	27	Spinal cord injury C5, ASIA C
P7	20	6	HS parieto-temporal, right central no cranium
P8	36	53	Spinal cord injury C5, ASIA A
P9	38	15	Spinal cord injury C4, ASIA A

References

[1] D. Valbuena, M. Cyriacks, O. Friman, I. Volosyak, and A. Graser. Brain-computer interface for high-level control of rehabilitation robotic systems. In *Proc. of the IEEE 10th Int. Conf. on Rehab. Robotics*, pages 619–625, 2007.

[2] H. Cecotti. Spelling with non-invasive brain-computer interfaces - current and future trends. *J. of Physiology-Paris*, 105(1-3):106–114, 2011.

[3] F. Velasco-Alvarez, R. Ron-Angevin, and M. A. Lopez-Gordo. BCI-based navigation in virtual and real environments. In *12th Int. Work-Conf. on Artificial Neural Net., LNCS*, volume 7903, pages 404–412, 2013.

[4] M. Rodríguez, R. Giménez, P. Diez, E. Avila, E. Laciar, L. Orosco, and A. G. Correa. Playing with your mind. *J. of Physics: Conference Series*, 477(1):012038, 2013.

[5] T. O. Zander and C. Kothe. Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *J. of Neural Engineering*, 8(2):025005, 2011.

[6] Y. K. Meena, H. Cecotti, K. Wong-Lin, and G. Prasad. Towards increasing the number of commands in a hybrid brain-computer interface with combination of gaze and motor imagery. In *Proc. of 37th Annual Int. Conf. of the IEEE Eng. in Medicine and Biology Society (EMBC)*, pages 506–509. IEEE, 2015.

[7] J. Wu, E. B. Quinlan, L. Dodakian, A. McKenzie, N. Kathuria, R. J. Zhou, R. Augsburger, J. See, V. H. Le, R. Srinivasan, and S. C. Cramer. Connectivity measures are robust biomarkers of cortical function and plasticity after stroke. *Brain*, 138:2359–2369, 2015.

[8] G. Prasad, P. Herman, D. Coyle, S. McDonough, and J. Crosbie. Applying a brain-computer interface to support motor imagery practice in people with stroke for upper limb recovery: a feasibility study. *J. of neuroengineering and rehabilitation*, 7(1):1, 2010.

[9] Jun Lu, Dennis J McFarland, and Jonathan R Wolpaw. Adaptive laplacian filtering for sensorimotor rhythm-based brain–computer interfaces. *Journal of neural engineering*, 10(1):016002, 2012.

[10] Ou Bai, Peter Lin, Sherry Vorbach, Mary Kay Floeter, Noriaki Hattori, and Mark Hallett. A high performance sensorimotor beta rhythm-based brain–computer interface associated with human natural motor behavior. *Journal of neural engineering*, 5(1):24, 2007.

[11] Jonathan R Wolpaw and Dennis J McFarland. Control

Accepted Manuscript

of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 101(51):17849–17854, 2004.

[12] Nazlar Ghassemzadeh and Siamak Haghipour. A review on EEG based brain computer interface systems feature extraction methods. *International Journal of Advanced Biological and Biomedical Research*, 4(2):117–123, 2016.

[13] M. Hamedi, S. Salleh, and A. M. Noor. Electroencephalographic motor imagery brain connectivity analysis for BCI: A review. *Neural computation*, 2016.

[14] Hikaru Takeuchi, Yasuyuki Taki, Yuko Sassa, Hiroshi Hashizume, Atsushi Sekiguchi, Ai Fukushima, and Ryuta Kawashima. Brain structures associated with executive functions during everyday events in a non-clinical sample. *Brain Structure and Function*, 218(4):1017–1032, 2013.

[15] D. J. Krusienski, D. J. McFarland, and J. R. Wolpaw. Value of amplitude, phase, and coherence features for a sensorimotor rhythm-based brain–computer interface. *Brain research bulletin*, 87(1):130–134, 2012.

[16] M. Grosse-Wentrup. Understanding brain connectivity patterns during motor imagery for brain-computer interfacing. In *Advances in neural information processing systems*, pages 561–568, 2009.

[17] M. Billinger, C. Brunner, and G. R. Müller-Putz. Single-trial connectivity estimation for classification of motor imagery data. *Journal of neural engineering*, 10(4):046006, 2013.

[18] H. Zhang, R. Chavarriaga, and J. del R Millán. Discriminant brain connectivity patterns of performance monitoring at average and single-trial levels. *NeuroImage*, 120:64–74, 2015.

[19] T. Schreiber. Measuring information transfer. *Physical review letters*, 85(2):461, 2000.

[20] M. J. Kaminski and K. J. Blinowska. A new method of the description of the information flow in the brain structures. *Biological cybernetics*, 65(3):203–210, 1991.

[21] L. A. Baccalá and K. Sameshima. Partial directed coherence: a new concept in neural structure determination. *Biological cybernetics*, 84(6):463–474, 2001.

[22] L. Xu, H. Zhang, M. Hui, Z. Long, Z. Jin, Y. Liu, and L. Yao. Motor execution and motor imagery: a comparison of functional connectivity patterns based on graph theory. *Neuroscience*, 261:184–194, 2014.

[23] D. Rathee, H. Cecotti, and G. Prasad. Estimation of effective fronto-parietal connectivity during motor imagery using partial granger causality analysis. In *Proc. of Int. Joint Conf. on Neural Networks*, pages 1–7, 2016.

[24] S. Guo, A. K. Seth, K. M. Kendrick, C. Zhou, and J. Feng. Partial granger causality-eliminating exogenous inputs and latent variables. *J. of Neuroscience Methods*, 172(1):79–93, 2008.

[25] Angeliki Papana, Catherine Kyrtsou, Dimitris Kugiumtzis, and Cees Diks. Simulation study of direct causality measures in multivariate time series. *Entropy*, 15(7):2635–2661, 2013.

[26] Elzbieta Olejarczyk, Laura Marzetti, Vittorio Pizzella, and Filippo Zappasodi. Comparison of connectivity analyses for resting state eeg data. *Journal of Neural Engineering*, 14(3):036017, 2017.

[27] H. Hytti, R. Takalo, and H. Ihalainen. Tutorial on multivariate autoregressive modelling. *Journal of clinical monitoring and computing*, 20(2):101–108, 2006.

[28] M. Tangermann, K.-R. Müller, A. Aertsen, N. Birbaumer, C. Braun, C. Brunner, R. Leeb, C. Mehring, K. J. Miller, G. Mueller-Putz, and G. Nolte. Review of the BCI competition IV. *Frontiers in Neuroscience*, 6:55, 2012.

[29] R. Scherer, J. Faller, E. VC Friedrich, E. Opisso, U. Costa, A. Kübler, and G. R. Müller-Putz. Individually adapted imagery improves brain-computer interface performance in end-users with disability. *PloS one*, 10(5):e0123727, 2015.

[30] F. Perrin, O. Bertrand, and J. Pernier. Scalp current density mapping: value and estimation from potential data. *IEEE Trans. on Biomedical Engineering*, 4:283–288, 1987.

[31] M. X. Cohen. Comparison of different spatial transformations applied to EEG data: A case study of error processing. *International Journal of Psychophysiology*, 97(3):245–257, 2015.

[32] M. Ding, S. L. Bressler, W. Yang, and H. Liang. Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. *Biological cybernetics*, 83(1):35–45, 2000.

[33] M. Morf, A. Vieira, D. Lee, and T. Kailath. Recursive multichannel maximum entropy spectral estimation. *IEEE Trans. Geosci. Electron*, 16(2):85–94, 1978.

[34] Camillo Porcaro, Gianluca Coppola, Francesco Pierelli, Stefano Seri, Giorgio Di Lorenzo, Leo Tomasevic, Carlo Salustri, and Franca Tecchio. Multiple frequency functional connectivity in the hand somatosensory network: an eeg study. *Clinical Neurophysiology*, 124(6):1216–1224, 2013.

[35] Camillo Porcaro, Filippo Zappasodi, Paolo Maria Rossini, and Franca Tecchio. Choice of multivariate autoregressive model order affecting real network functional connectivity estimate. *Clinical Neurophysiology*, 120(2):436–448, 2009.

[36] G. Schwarz. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.

[37] H. Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974.

[38] J. Durbin and G. S. Watson. Testing for serial correlation in least squares regression. II. *Biometrika*, 38(1/2):159–177, 1951.

[39] A. K. Seth. A MATLAB toolbox for granger causal connectivity analysis. *Journal of neuroscience methods*, 186(2):262–273, 2010.

[40] G. Niso, R. Bruña, E. Pereda, R. Gutiérrez, R. Bajo, F. Maestú, and F. del Pozo. HERMES: towards an integrated toolbox to characterize functional and effective brain connectivity. *Neuroinformatics*, 11(4):405–434, 2013.

[41] Tom Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27:861–874, 2006.

[42] S. Hetu, M. Gregoire, A. Saimpont, M. P. Coll, F. Eugene, P. E. Michon, and P. L. Jackson. The neural network of motor imagery:an ALE meta-analysis. *Neurosci. Biobehav. Rev.*, 37, 2013.

[43] K. J. Miller, G. Schalk, E. Fetz, M. den Nijs, J. G. Ojemann, and R. P. N. Rao. Cortical activity during motor execution, motor imagery, and imagery-based online feedback. *Proc. of the National Academy of Sciences*, 107(9):4430–4435, 2010.

[44] M. M. Smith, K. E. Weaver, T. J. Grabowski, R. P. N. Rao, and F. Darvas. Non-invasive detection of high gamma band activity during motor imagery. *Frontiers in human neuroscience*, 8, 2014.

[45] D. Hermes, K. J. Miller, M. J. Vansteensel, E. J. Aarnoutse, F. S. S. Leijten, and N. F. Ramsey. Neurophysiologic correlates of fMRI in human motor cortex. *Human brain mapping*, 33(7):1689–1699, 2012.

[46] C. R. Conner, T. M. Ellmore, T. A. Pieters, M. A. DiSano, and N. Tandon. Variability of the relationship between electrophysiology and BOLD-fMRI across cortical regions in humans. *The Journal of Neuroscience*, 31(36):12855–12865, 2011.

Accepted Manuscript

[47] V. Youssofzadeh, G. Prasad, M. Naeem, and K. Wong-Lin. Temporal information of directed causal connectivity in multi-trial ERP data using partial granger causality. *Neuroinformatics*, 14(1):99–120, 2016.

[48] Elisabeth VC Friedrich, Reinhold Scherer, and Christa Neuper. The effect of distinct mental strategies on classification performance for brain–computer interfaces. *International Journal of Psychophysiology*, 84(1):86–94, 2012.

[49] Elisabeth VC Friedrich, Reinhold Scherer, and Christa Neuper. Long-term evaluation of a 4-class imagery-based brain–computer interface. *Clinical Neurophysiology*, 124(5):916–927, 2013.

[50] S. Haufe, V. V. Nikulin, K.-R. Müller, and G. Nolte. A critical assessment of connectivity measures for EEG data: a simulation study. *Neuroimage*, 64:120–133, 2013.

[51] M. A. Jatoi, N. Kamel, A. S. Malik, I. Faye, and T. Begum. A survey of methods used for source localization using EEG signals. *Biomedical Signal Processing and Control*, 11:42–52, 2014.

[52] Sebastianus Petrus van den Broek, F Reinders, M Donderwinkel, and MJ Peters. Volume conduction effects in EEG and MEG. *Electroencephalogr. Clin. Neurophysiol.*, 106(6):522–534, 1998.

[53] Sylvain Baillet, John C Mosher, and Richard M Leahy. Electromagnetic brain mapping. *IEEE Signal processing magazine*, 18(6):14–30, 2001.

Accepted Manuscript