Current source density estimates improve the discriminability of scalp-level brain connectivity features related to motor-imagery tasks

Dheeraj Rathee*, Hubert Cecotti *Senior Member, IEEE*, Girijesh Prasad *Senior Member, IEEE*

Abstract— Recent progress in the number of studies involving brain connectivity analysis of motor imagery (MI) tasks for brain-computer interface (BCI) systems has warranted the need for pre-processing methods. The objective of this study is to evaluate the impact of current source density (CSD) estimation from raw electroencephalogram (EEG) signals on the classification performance of scalp level brain connectivity feature based MI-BCI. In particular, time-domain partial Granger causality (PGC) method was implemented on the raw EEG signals and CSD signals of a publicly available dataset for the estimation of brain connectivity features. Moreover, pairwise binary classifications of four different MI tasks were performed in inter-session and intra-session conditions using a support vector machine classifier. The results showed that CSD provided a statistically significant increase of the AUC: 20.28% in the inter-session condition; 12.54% and 13.92% with session 01 and session 02, respectively, in the intra-session condition. These results show that pre-processing of EEG signals is crucial for single-trial connectivity features based MI-BCI systems and CSD can enhance their overall performance.

I. INTRODUCTION

Alternative methods for communication and/or control applications can be developed using brain-computer interfaces (BCIs) which involve the non-invasive acquisition of electroencephalogram (EEG) signals with scalp electrodes [1], [2], [3]. Motor imagery (MI) related BCI systems record EEG signals during kinesthetic imagination of one or more motor tasks, extract task-related features of brain activity and may utilize them to generate control inputs for the efferent devices. Most of these BCI systems focused on the sensorimotor rhythms that are specific to the MI tasks (i.e., event-related desynchronization/synchronization (ERD/ERS)) [4], [5]. However, recent studies have provided sufficient empirical evidence that brain connectivity features can reliably discriminate the MI tasks [6], [7], [8], [9]. Current BCIs related to MI detection are primarily based on features obtained from individual EEG channels i.e. power measure of band-passed data and common spatial pattern based features. Yet, there exists still a high number of people who are potential BCI users and are not able to use current BCIs. Brain connectivity measurements can, therefore, provide key insights of brain responses associated to the MI tasks that can aid to the improvement of these BCIs.

Several factors drive the overall performance of a typical brain connectivity based MI-BCI. These include the pre-processing algorithms implemented for cleaning the noisy data, the connectivity measures for estimating the interactions between EEG channels and finally, the classifiers for discriminating the connectivity features. Yet, the majority of the previous studies focused on different connectivity estimation methods and their performances with different MI tasks, however, EEG signal pre-processing have been largely overlooked [10], [7], [11]. A typical pre-processing step involves the reduction of unwanted components from the raw EEG data. These artifacts can occur due to head/electrode movement, eye blinks, and volume conduction (VC). This step leads to enhanced signal-to-noise ratio (SNR) of the data and thus, reduce the spurious connections in the estimated brain networks. Imagination of different motor tasks induce significant changes in the physiological interactions between various cortical regions. However, these brain regions are situated in the close neighborhood and thus, making scalp-level MI-related causal interactions a significantly challenging task. Moreover, the high sensitivity of connectivity features to the presence of VC makes the task more difficult [12]. Therefore, to improve the performance of connectivity-based MI-BCI systems, a reliable pre-processing of the EEG data must be performed.

In the present work, in line with our previous study [6], we estimate the effect of a spatial pre-processing method (i.e. current source density (CSD)) on the discriminability of single-trial directed functional connectivity features related to the EEG signals during four MI tasks. In particular, we compare the classification performance of the time-domain Partial Granger Causality (PGC) measures with and without CSD pre-processing. The remainder of this paper is organized as follows: Section II provides a detailed description of the MI related EEG dataset, estimation of CSD and time-domain PGC, and the applied methodology. Section III illustrates the comparison of performances of connectivity based MI-BCI for inter-session and intra-session conditions with both raw data and CSD pre-processed data. Finally, In section IV the significance and limitations of the study are discussed of along with possible future work.

II. MATERIAL AND METHODS

A. EEG Dataset

This study utilized an open-access MI related EEG dataset: Dataset-2a from the BCI competition-IV [13]. EEG signals were recorded from nine participants, namely [A01-A09] with monopolar electrodes placed at 22 positions over the

scalp involving four different MI classes, i.e., left hand, right hand, both feet, and tongue. The dataset consists of two sessions (i.e., S01 and S02) acquired on two different days. Each session consists of 288 trials (72 for each class). For further analysis, we have selected seven channels (i.e., FC1, FC2, C3, CZ, C4, CP1, and CP2) over the motor cortex region.

B. Estimation of Current Source Density (CSD)

Spherical spline interpolation method, presented by Perrin et al. [14], [15], was implemented for the estimation of scalp-level CSD values from the raw EEG data. The estimation of CSD is formally defined by the following expressions:

Let's assume an EEG data with N number of electrodes with v_i be the amplitude estimated for i^{th} electrode. Now, the first step is to calculate the function $g(\cos(L_i, L_j))$ using the following mathematical expressions:

$$g(x) = \frac{1}{4\pi} \sum_{p=1}^{\infty} \left(\frac{2p+1}{p^q(p+1)^q} \right) \cdot E_p(x) \quad (1)$$

$$\cos(L_i, L_j) =$$

$$1 - \frac{(X_{L_i} - X_{L_j})^2 + (Y_{L_i} - Y_{L_j})^2 + (Z_{L_i} - Z_{L_j})^2}{2} \quad (2)$$

where, (X,Y,Z) are cartesian coordinates; L_i and L_j are spherical projections of ith and j^{th} electrodes; q is the spline flexibility index and $E_p(x)$ is the Legendre polynomial of degree p. The second step is to estimate the transformation constants $k_i s$ using the solutions of the following equations:

$$(G + \lambda)K + Dk_0 = V \quad (3)$$
$$D'K = 0$$

where $D' = [1, 1, \ldots, 1]$, $K' = [k_1, k_2, ..., k_N]$, $V' = [v_1, v_2, ..., v_N]$, $G = g(\cos(L_i, L_j))$, and the value of λ (smoothing constant) is equal to 1.0e - 5. Next, similar to Equation 1, we can get function $h(\cos(L_i, L_j))$ using:

$$h(x) = -\frac{1}{4\pi} \sum_{p=1}^{\infty} \left(\frac{(2p+1)^2}{p^q(p+1)^q} \right) \cdot E_p(x) \quad (4)$$

$$(5)$$

where x is $\cos(L_i, L_j)$. Finally, the CSD estimates can be generated using:

$$U(L) = \sum_{i=1}^{N} k_i h(\cos(L, L_i)) \quad (6)$$

C. Time-domain Partial Granger Causality Analysis

The concept of PGC, provided by Guo et al. [16], is based on eliminating the effects of exogenous (environmental) input and latent variables during the estimation of conditional GC. Initially, the representations of exogenous (environmental) input and latent variables are considered during the multivariate autoregressive (MVAR) modelling of the data. Later, their effect is mitigated by partitioning

the noise covariance matrices generated from the model. For generating mathematical expression for PGC, let's consider time series data from two electrodes as $P(t)$ and $Q(t)$ for which causality analysis have to be performed. Also, consider another time series representing the data from third electrode (i.e., a common input) as $R(t)$. Further, two separate MVAR models are considered for predicting the values of $P(t)$ from its previous values and the common input (i.e., reduced model) and for predicting the values of $P(t)$ from the previous values of itself, $Q(t)$ and the common input (i.e., full model). The reduced model can be presented as:

$$P(t) = \sum_{o=1}^{k} (a_{(1,o)} P(t-o)) + \sum_{o=1}^{k} (c_{(1,o)} R(t-o)) + \quad (7)$$
$$\epsilon_1(t) + \epsilon_1^E(t) + \beta_1(L)\epsilon_1^L(t)$$

$$R(t) = \sum_{o=1}^{k} (b_{(1,o)} R(t-o)) + \sum_{o=1}^{k} (d_{(1,o)} P(t-o)) + \quad (8)$$
$$\epsilon_2(t) + \epsilon_2^E(t) + \beta_2(L)\epsilon_2^L(t)$$

where o is the model order, $\epsilon_i(t)$, $\epsilon_i^E(t)$, and $\beta(L)\epsilon_i^L(t)$ are the prediction errors related to model, exogenous inputs (E) and latent variables (L), respectively. Likewise, the full model can be represented as:

$$P(t) = \sum_{o=1}^{k} (a_{(2,o)} P(t-o)) + \sum_{o=1}^{k} (b_{(2,o)} Q(t-o)) + \quad (9)$$
$$\sum_{o=1}^{k} (c_{(2,o)} R(t-o)) + \epsilon_3(t) + \epsilon_3^E(t) + \beta_3(L)\epsilon_3^L(t)$$

$$Q(t) = \sum_{o=1}^{k} (d_{(2,o)} P(t-o)) + \sum_{o=1}^{k} (e_{(2,o)} Q(t-o)) + \quad (10)$$
$$\sum_{o=1}^{k} (f_{(2,o)} R(t-o)) + \epsilon_4(t) + \epsilon_4^E(t) + \beta_4(L)\epsilon_4^L(t)$$

$$R(t) = \sum_{o=1}^{k} (g_{(2,o)} P(t-o)) + \sum_{o=1}^{k} (h_{(2,o)} Q(t-o)) + \quad (11)$$
$$\sum_{o=1}^{k} (i_{(2,o)} R(t-o)) + \epsilon_4(t) + \epsilon_4^E(t) + \beta_4(L)\epsilon_4^L(t)$$

For ease of notation, let's define the collective prediction error as:

$$\mu_i(t) = \epsilon_j(t) + \epsilon_j^E(t) + \beta_j(L)\epsilon_j^L(t) \quad (12)$$

with $1 \leq j \leq 5$.

Now, the noise covariance matrices for the reduced and full model can be obtained as:

$$X = \begin{bmatrix} \text{var}(\mu_1(t)) & \text{cov}(\mu_1(t), \mu_2(t)) \\ \text{cov}(\mu_2(t), \mu_1(t)) & \text{var}(\mu_2(t)) \end{bmatrix} \quad (13)$$

$$L =$$

$$\begin{bmatrix} \text{var}(\mu_3(t)) & \text{cov}(\mu_3(t), \mu_4(t)) & \text{cov}(\mu_3(t), \mu_5(t)) \\ \text{cov}(\mu_4(t), \mu_3(t)) & \text{var}(\mu_4(t)) & \text{cov}(\mu_4(t), \mu_5(t)) \\ \text{cov}(\mu_5(t), \mu_3(t)) & \text{cov}(\mu_5(t), \mu_4(t)) & \text{var}(\mu_5(t)) \end{bmatrix} \quad (14)$$

Finally, the directional connectivity measure from $Q(t)$ to $P(t)$ conditioned over $R(t)$ can be estimated as :

$$\mathcal{C}_{Q \to P|R} = ln(\frac{X_{1,1} - X_{1,2}X_{2,2}^{-1}X_{2,1}}{Y_{1,1} - Y_{1,3}Y_{3,3}^{-1}Y_{3,1}}) \qquad (15)$$

D. Data Analysis

The initial step was the estimation of scalp CSD values from the raw EEG data. The remaining processing steps were performed separately for the raw EEG data and CSD processed data. Next, the data were bandpassed in the lower γ frequency band (25-40 Hz) using a 4th order, zero-phase forward and backward bandpass Butterworth filter as our previous study showed better separability of MI related connectivity values for this frequency band [6]. Furthermore, the data related to the imagery tasks (i.e. 3000 ms to 6000 ms) were extracted for each trial. and the epoched data were segmented using a sliding window (segment length 1000 ms and 500 ms overlap) approach. Thus, we obtained 5 segments from each trial. The windowing strategy can effectively reduce the probability of spurious causal effects due to non-stationarity in EEG signal. Multi-trial time-domain PGC method was implemented to estimate the causal interactions between the scalp channel data. The coefficients of the MVAR model were estimated using the LWR algorithm [17]. Schwarz Bayesian Information criterion (SBIC) [18] was used for the estimation of the optimal value of the model order o. The analysis provides 42 ($N_s * N_s - N_s$, with $N_s = 7$ being the number of sensors) non-zero directed connectivity measures for each trial. Furthermore, SVM classifier was incorporated to calculate the classification performances (i.e., area under the ROC curve (AUC)) of the six binary classification tasks. The complete analysis was performed for two different conditions: Firstly, an inter-session condition where the classifier was trained and evaluated on feature-sets of S01 and S02, respectively; Second, an intra-session condition where 10 times 10-fold cross-validation (CV) was performed for S01 and S02, separately. The models were implemented using custom MATLAB (V8.6) scripts using functions from GCCA (V2.9) toolbox [19] on an Intel Core i7-4790 processor with 16 GB of memory.

III. RESULTS

The data processing and analysis yielded to 6 binary classifications tasks i.e., left vs right (L↔R), left vs feet (L↔F), left vs tongue (L↔T), right vs feet (R↔F), right vs tongue (R↔T), and feet vs tongue (F↔T) for both inter-session and intra-session conditions. Fig. 1 and 2 provide the mean AUC values obtained with six classification tasks and their grand mean values for session S01 and S02 (intra-session condition), respectively. For this analysis, CSD provided statistically significant improvements of 12.54% ($p = 0.00014$) and 13.92% ($p = 0.00017$) in overall performances (i.e., grand mean AUCs) of session S01 and S02, respectively. The average ($\pm$SD) AUC values obtained across nine subjects for session S01 are 0.65 ± 0.11, 00.68 ± 0.09, 0.66 ± 0.15, 0.64 ± 0.09, 00.63 ± 0.15, and 0.61 ± 0.10 and session S02

are 0.67 ± 0.14, 00.70 ± 0.14, 0.66 ± 0.14, 0.64 ± 0.10, 00.65 ± 0.13, and 0.61 ± 0.08 for L↔R, L↔F, L↔T, R↔F, R↔T, and F↔T.

For inter-session analysis, the mean AUC values obtained with six classification tasks and their grand mean values are presented in Fig. 3. For this condition, CSD provided statistically significant improvements of 20.28% ($p = 0.00007$) in overall performance (i.e., grand mean AUCs). The average ($\pm$SD) AUC values obtained across nine subjects are 0.69 ± 0.14, 0.73 ± 0.13, 0.73 ± 0.14, 0.68 ± 0.10, 00.71 ± 0.12, and 0.64 ± 0.07 for L↔R, L↔F, L↔T, R↔F, R↔T, and F↔T.
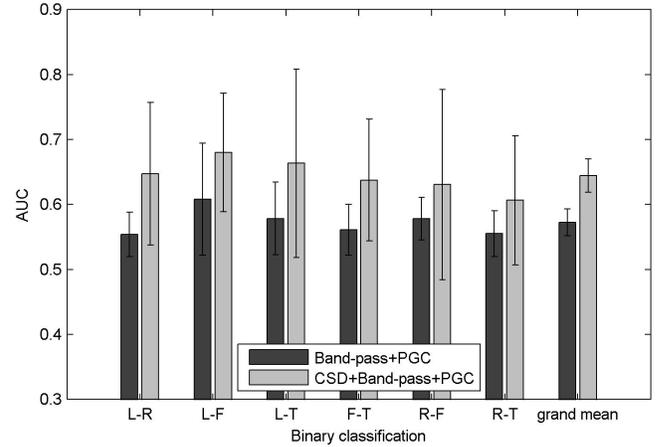

Fig. 1.   Mean AUC measures obtained from 10 times 10-fold CV with Session (S01) data for six pairwise comparisons. Each tick on x-axis provides mean AUCs for the band-passed data and CSD processed band-passed data. The error bar for each of the six pairwise comparison represents the standard deviation (SD) across 9 subjects whereas the error bar for grand mean is the SD across 6 comparisons.
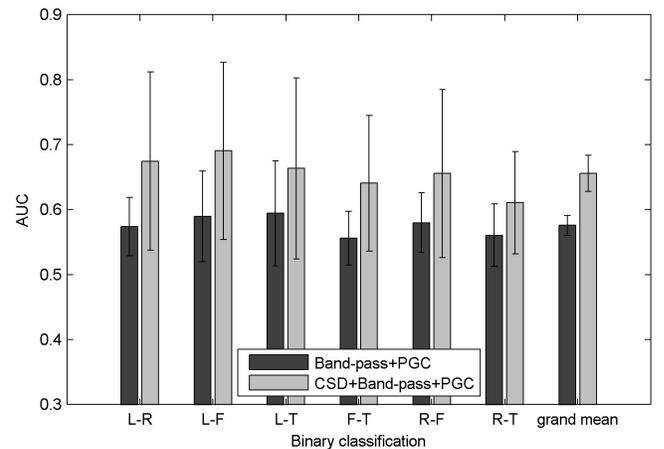

Fig. 2.   Mean AUC measures obtained from 10 times 10-fold CV with Session (S02) data for six pairwise comparisons. Each tick on x-axis provides mean AUCs for the band-passed data and CSD processed band-passed data. The error bar for each of the six pairwise comparison represents the SD across 9 subjects whereas the error bar for grand mean is the SD across 6 comparisons.

IV. DISCUSSION AND CONCLUSION

The complete signal processing pipeline of a pattern recognition system includes multiple stages such as the feature
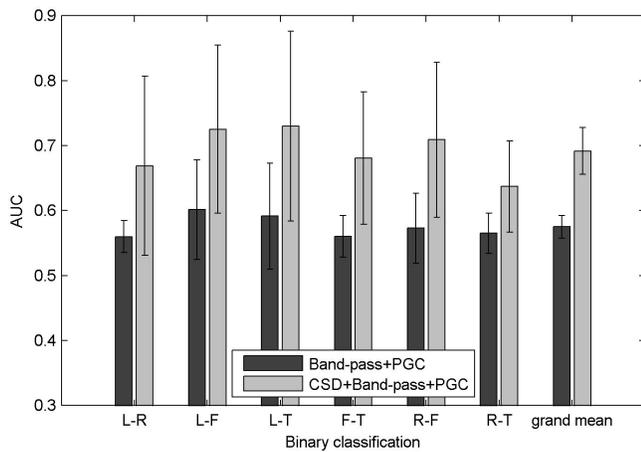
Fig. 3. Mean AUC measures obtained from training with data of session S01 and evaluation with session S02 for six pairwise comparisons. Each tick on x-axis provides mean AUCs for the band-passed data and CSD processed band-passed data. The error bar for each of the six pairwise comparison represents the SD across 9 subjects whereas the error bar for grand mean is the SD across 6 comparisons.

extraction step and the classification stage. For features such as connectivity measurements, pre-processing methods are often overlooked and the focus for improving brain connectivity based MI-BCI has been mainly on feature extraction (i.e., connectivity estimation) techniques that maximize the difference between given classes. Although we acknowledge the importance of using reliable connectivity measures, we found that the implementation of efficient pre-processing technique can further enhance the performance of such BCI systems, as shown in the present study. In particular, brain connectivity measures were estimated by implementing time-domain PGC method on the raw EEG signals and CSD processed EEG signals of a publicly available dataset. Furthermore, the effect of CSD processing was evaluated by comparing the grand mean AUC values obtained for both intra-session and inter-session analyses. The results confirmed significant improvements in the dicriminability of the connectivity features for both analyses using CSD as a pre-processing method. Our findings complement a recent study which showed that better interpretation of brain networks can be obtained through CSD processed EEG data. The analysis found a significant reduction in the number of spurious connections with CSD [20].

Despite the interest of the approach for improving the discriminability of brain connectivity based MI-BCI, several limitations can be found. First, the effect of CSD has been evaluated only with one connectivity measure (i.e., PGC) while further analysis must involve several other connectivity measures such as Transfer Entropy and coherence-based methods. Second, several studies showed a reduction of low spatial frequencies with surface Laplacian implementation, which may result in reduced sensitivity of the EEG signals to the deep neural source generators [21]. MI related activations are however majorly associated with the cortical regions and the effect of high-spatial filtering due to CSD may not affect

their SNR. Further work will deal with the online analysis of brain connectivity during real-time BCIs.

REFERENCES

[1] N. Birbaumer et al., "A spelling device for the paralysed," Nature, vol. 398, no. 6725, pp. 297–298, 1999.
[2] G. Pfurtscheller et al., "Brain-computer interface – a new communication device for handicapped persons," J. Microcomput. Appl, vol. 16, no. 3, pp. 293–299, 1993.
[3] Y. Meena et al., "Simultaneous gaze and motor imagery hybrid bci increases single-trial detection performance: a compatible incompatible study." IEEE EMBS, 2015.
[4] H. Raza et al., "Adaptive learning with covariate shift-detection for motor imagery-based brain–computer interface," Soft Computing, vol. 20, no. 8, pp. 3085–3096, 2016.
[5] E. Thomas et al., "Investigating brief motor imagery for an erd/ers based bci," in Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE. IEEE, 2012, pp. 2929–2932.
[6] D. Rathee et al., "Single-trial effective brain connectivity patterns enhance discriminability of mental imagery tasks," Journal of neural engineering, vol. 14, no. 5, p. 056005, 2017.
[7] M. Hamedi et al., "Electroencephalographic motor imagery brain connectivity analysis for BCI: A review," Neural computation, 2016.
[8] D. Rathee et al., "Current source density estimation enhances the performance of motor-imagery-related brain–computer interface," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 25, no. 12, pp. 2461–2471, 2017.
[9] A. Chowdhury et al., "Online covariate shift detection based adaptive brain-computer interface to trigger hand exoskeleton feedback for neuro-rehabilitation," IEEE Transactions on Cognitive and Developmental Systems, 2017.
[10] M. Billinger et al., "Single-trial connectivity estimation for classification of motor imagery data," Journal of neural engineering, vol. 10, no. 4, p. 046006, 2013.
[11] M. Grosse-Wentrup, "Understanding brain connectivity patterns during motor imagery for brain-computer interfacing," in Advances in neural information processing systems, 2009, pp. 561–568.
[12] S. P. van den Broek et al., "Volume conduction effects in EEG and MEG," Electroencephalogr. Clin. Neurophysiol., vol. 106, no. 6, pp. 522–534, 1998.
[13] M. Tangermann et al., "Review of the BCI competition IV," Front. Neurosci., vol. 6, p. 55, 2012.
[14] F. Perrin et al., "Spherical splines for scalp potential and current density mapping," Electroencephalogr. Clin. Neurophysiol., vol. 72, no. 2, pp. 184–187, 1989.
[15] J. Kayser, "Current source density (CSD) interpolation using spherical splines–CSD toolbox (version 1.1)," New York State Psychiatric Institute: Division of Cognitive Neuroscience, 2009.
[16] S. Guo et al., "Partial granger causality-eliminating exogenous inputs and latent variables," J. of Neuroscience Methods, vol. 172, no. 1, pp. 79–93, 2008.
[17] M. Morf et al., "Recursive multichannel maximum entropy spectral estimation," IEEE Trans. Geosci. Electron, vol. 16, no. 2, pp. 85–94, 1978.
[18] G. Schwarz, "Estimating the dimension of a model," The annals of statistics, vol. 6, no. 2, pp. 461–464, 1978.
[19] A. K. Seth, "A MATLAB toolbox for granger causal connectivity analysis," Journal of neuroscience methods, vol. 186, no. 2, pp. 262–273, 2010.
[20] M. X. Cohen, "Comparison of different spatial transformations applied to eeg data: a case study of error processing," International Journal of Psychophysiology, vol. 97, no. 3, pp. 245–257, 2015.
[21] P. L. Nunez et al., "EEG coherency-I: statistics, reference electrode, volume conduction, laplacians, cortical imaging, and interpretation at multiple scales," Electroencephalogr. Clin. Neurophysiol., vol. 103, no. 5, pp. 499–515, 1997.