



# A multi-class EEG-based BCI classification using multivariate empirical mode decomposition based filtering and Riemannian geometry



Pramod Gaur<sup>a,\*</sup>, Ram Bilas Pachori<sup>b</sup>, Hui Wang<sup>c</sup>, Girijesh Prasad<sup>a</sup>

<sup>a</sup> Intelligent Systems Research Centre, Ulster University, L'Derry, UK

<sup>b</sup> Discipline of Electrical Engineering, Indian Institute of Technology Indore, Indore, India

<sup>c</sup> School of Computing and Mathematics, Ulster University, Jordanstown, UK

## ARTICLE INFO

### Article history:

Received 20 June 2017

Revised 2 November 2017

Accepted 2 November 2017

Available online 7 November 2017

## ABSTRACT

A brain-computer interface (BCI) facilitates a medium to translate the human motion intentions using electrical brain activity signals such as electroencephalogram (EEG) into control signals. EEG signals are non-stationary and subject specific. A major challenge in BCI research is to classify human motion intentions from non-stationary EEG signals. We propose a novel subject specific multivariate empirical mode decomposition (MEMD) based filtering method, namely, SS-MEMDBF to classify the motor imagery (MI) based EEG signals into multiple classes. The MEMD method simultaneously decomposes the multichannel EEG signals into a group of multivariate intrinsic mode functions (MIMFs). This decomposition enables us to extract the cross-channel information and also localize the specific frequency information. The MIMFs are considered as narrow-band, amplitude and frequency modulated (AFM) signals. The statistical measure, mean frequency has been used to automatically filter the MIMFs to obtain enhanced EEG signals which better represent motor imagery related brainwave modulations over  $\mu$  and  $\beta$  rhythms. The sample covariance matrix has been computed and used as a feature set. The feature set has been classified into multiple MI tasks using Riemannian geometry. The proposed method has helped achieve mean Kappa value of 0.60 across nine subjects of the BCI competition IV dataset 2A which is superior to all the reported methods.

© 2017 Elsevier Ltd. All rights reserved.

## 1. Introduction

Brain computer interface (BCI) allows a medium for healthy or disabled people to communicate with external environment using various neuroimaging technology such as electroencephalogram (EEG)/magnetoencephalography (MEG) (Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002). The BCI aims to translate the intent of a subject into control commands for a neuroprosthetics application. A popular paradigm for BCI communication is motor imagery (MI) (Pfurtscheller, Neuper, Flotzinger, & Pregenzer, 1997; Pfurtscheller, Brunner, Schlögl, & Da Silva, 2006). A subject is required to perform an imagination of a movement with a particular limb. Moreover, EEG signals may be recorded during multiple MI tasks (e.g. right hand, left hand, foot and tongue).

Several attempts have been made to understand the dynamics of EEG signals by exploring different frequency bands, such as

$\mu$  (8–13 Hz) and  $\beta$  (13–25 Hz) rhythms. It is a fact that the responses and topographies obtained from beta ( $\beta$ ) rhythm is distinct as compared to the mu ( $\mu$ ) rhythm corresponding to limb movements. It has been shown during limb movements, there is an increase in the oscillatory power of the beta rhythm observed in ipsilateral sensorimotor cortex and simultaneously there is a decrease in oscillatory power of mu rhythm observed in the contralateral sensorimotor cortex (Herman, Prasad, McGinnity, & Coyle, 2008; Gandhi, Prasad, Coyle, Behera, & McGinnity, 2014; Gaur, Pachori, Wang, & Prasad, 2015; 2016a). The BCI systems identify these changes to provide some meaningful command. One of the major issues in BCI systems is intrinsic non-stationarity present in the EEG signals which happens when these signals originate from different sources. In addition, these signals have a very low signal-to-noise ratio (SNR) (Nicolas-Alonso & Gomez-Gil, 2012). Several factors like artifacts coming from electrooculogram (EOG), electromyogram (EMG) interference and electrical power line may contribute to low SNR. In order to increase SNR, the most useful step would be to enhance the EEG signals by eliminating these distortions or artifacts from the raw EEG signals in the preprocessing stage. This step helps to obtain better separability in the feature

\* Corresponding author.

E-mail addresses: [gaur-p@email.ulster.ac.uk](mailto:gaur-p@email.ulster.ac.uk) (P. Gaur), [pachori@iiti.ac.in](mailto:pachori@iiti.ac.in) (R.B. Pachori), [h.wang@ulster.ac.uk](mailto:h.wang@ulster.ac.uk) (H. Wang), [g.prasad@ulster.ac.uk](mailto:g.prasad@ulster.ac.uk) (G. Prasad).

set corresponding to multiple MI tasks (Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002).

Recently, single-channel empirical mode decomposition (EMD) based filtering has been studied for two-class classification problems (Gaur, Pachori, Wang, & Prasad, 2015). In addition, EMD has been also explored for the classification of epileptic seizures by Sharma and Pachori (2015). Gaur et al. (2015) use mean frequency of intrinsic mode functions (IMFs) to achieve enhanced EEG signals for BCI. Also, the multichannel variation of EMD (MEMD) has been explored (Park, Looney, Ahrabian, & Mandic, 2013) to handle the adverse effect of intrinsic non-stationarity present in EEG signals. This method also utilizes cross-channel information available in multichannel EEG signals. A multi-variate EMD based filtering has been studied with common spatial pattern (CSP) features to classify two-class classification problem for subject independent BCI (Gaur, Pachori, Wang, & Prasad, 2016b). Motivated by these decomposition techniques, here we extend our work in (Gaur, Pachori, Wang, & Prasad, 2016a; 2016b) and propose a novel multichannel subject specific MEMD based filtering called SS-MEMDBF. The SS-MEMDBF filtering range has been identified for every subject based on the mu and beta band activities observed in MI related brainwaves. A subject specific training method has been developed to resolve the multi-class classification problem, namely, right hand, left hand, foot and tongue MI tasks using the BCI competition IV dataset 2A (Brunner, Leeb, Müller-Putz, Schlögl, & Pfurtscheller, 2008). The MEMD method decomposes the multivariate time-series into a group of multichannel IMFs (MIMFs) such that the IMFs across the multichannel give the exact same number of oscillations. These IMFs are considered as narrow-band amplitude and frequency modulated (AFM) signals.

In BCI research community, there are different variants of CSP algorithm studied and used by several groups (Ang, Chin, Wang, Guan, & Zhang, 2012; Zhang, Yang, & Guan, 2013) to extract more separable spatial patterns as features. In this work, we exploit the structure of sample covariance matrix as feature set, as the sample covariance matrix contains the spatial information present in EEG signal. The main objective is to devise a unique step by combining the spatial filtering and the classification. A similar concept has been explored in (Farquhar, 2009). However, sample covariance matrices structure needs to be handled carefully in Riemannian manifold. In this respect, a rich framework is facilitated by Riemannian geometry (Barachant, Bonnet, Congedo, & Jutten, 2012) to handle these matrices.

Here, the process is two-fold, first we perform SS-MEMDBF filtering to filter the single trials of each subject. This step enables us to use the cross-channel information and localize the frequency component across the channels. The second step involves manipulating EEG spatial covariance matrices in their native space and utilize the strength of Riemannian distance to identify different MI tasks. We have not applied any kind of spatial filtering because spatial covariance matrices already contain spatial information. To send covariance matrices into Euclidean space, we have also used tangent space mapping of these matrices, where they may be considered as vectors belonging to a manifold (Barachant, Bonnet, Congedo, & Jutten, 2012).

In this paper, we thus propose a new SS-MEMDBF for classification of multi-class MI tasks by extending our previously proposed filtering technique (Gaur et al., 2015) based on EMD which is restricted to two-class MI tasks in BCI. The enhanced EEG signals have been obtained corresponding to  $\mu$  and  $\beta$  rhythms before extracting features in order to classify right hand, left hand, foot, and tongue MI tasks. Particularly here, we utilize the capability of existing MEMD method (Davies & James, 2013; 2014; Park, Looney, Ahrabian, & Mandic, 2013; Park et al., 2014) to decompose the signal into a set of narrow-band MIMFs and perform automatic selection of MIMFs based on a statistical measure namely,

mean frequency, corresponding to  $\mu$  and  $\beta$  rhythms of each subject. Thereafter, the summation of selected MIMFs is performed to attain enhanced EEG signals corresponding to a particular subject. This automatic selection method will enhance EEG signals for all MI tasks. A block diagram of the proposed SS-MEMDBF methodology is shown in Fig. 1.

The aims of the work reported in the paper are thus as follows:

1. To study the intra- and inter-subject non-stationarity present in the EEG signals;
2. To identify whether the same frequency or different frequency components are involved in MI activity when EEG signals are recorded from the same cortical areas across the subjects;
3. To automatically identify the subject specific MIMFs based on mean frequency measure;
4. To present the results in terms of classification accuracy and Kappa value when single trials are classified.

This paper is organized as follows: Section 2 discusses the BCI competition IV dataset 2A details. In Section 3, we review the MEMD decomposition technique. The feature extraction method is discussed in Section 4. A brief introduction of Riemannian geometry used has been provided in Section 5. The Section 5 also presents the details about the Riemannian framework for the MI based EEG signals. The results obtained from the proposed method along with the discussion are described in Section 6 and comparison with the methodologies developed by other research groups are discussed in Section 7 and finally, the paper concludes in Section 8.

## 2. BCI competition IV dataset 2A description

We have made use of publically available BCI competition IV dataset 2A (Brunner, Leeb, Müller-Putz, Schlögl, & Pfurtscheller, 2008). It contains EEG signals from 22 EEG channels and 3 EOG channels with left mastoid as reference, while performing multiple MI tasks: right hand, left hand, foot, and tongue movements. The dataset contains nine healthy subjects and each subject has two sessions, one training session and one test session. Each session has 288 trials of MI data with 72 trials for each MI task. The EEG signals were sampled at the sampling rate of 250 Hz and band-pass filtered between 0.5 Hz and 100 Hz. An additional 50 Hz notch filter has been applied to quench line noise. The selection of a specific time-interval within a trial period plays a crucial role in MI classification. We have taken 2 s data between 0.5 s and 2.5 s after the onset of the visual cue in the training stage as studied by the competition winner (Ang, Chin, Wang, Guan, & Zhang, 2012). For more details about BCI competition IV dataset 2A refer to (Brunner et al., 2008). We have considered twenty-two channels in our study as shown in Fig. 2. The enhanced EEG signals from all channels have been used to extract the sample covariance matrix as a feature set.

## 3. Multivariate empirical mode decomposition: a review

As discussed earlier, the raw EEG signals have low SNR as they may be affected with interference from electrosurgical units (ESUs), EOG, and EMG (Pfurtscheller, Neuper, Flotzinger, & Pregeizer, 1997). The frequency bands of interest (e.g. mu and beta rhythms) may contain a lot of contaminating noise in the raw EEG signal which may produce erroneous results. Therefore, a time-frequency decomposition technique is required, which can remove noise and artifacts but does not affect the original EEG signal. Huang et al. (1998) proposed EMD, a single-channel decomposition technique which decomposes the original signal into a group

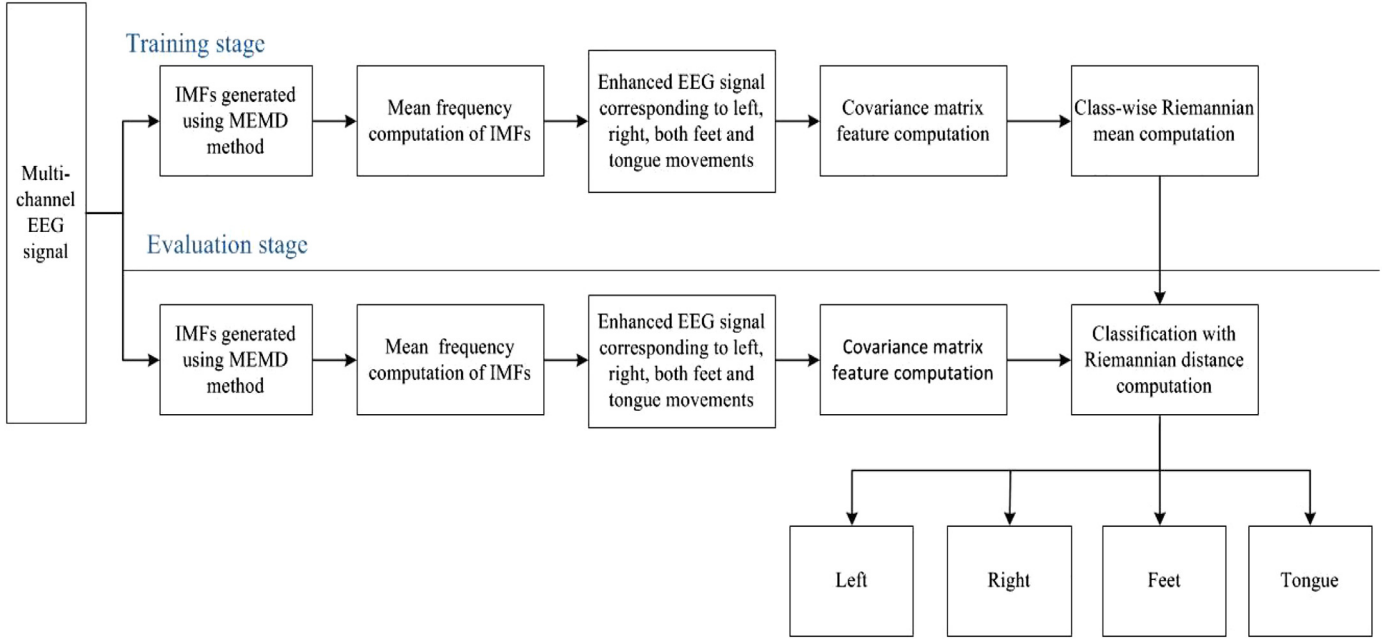


Fig. 1. Block diagram for the proposed method.

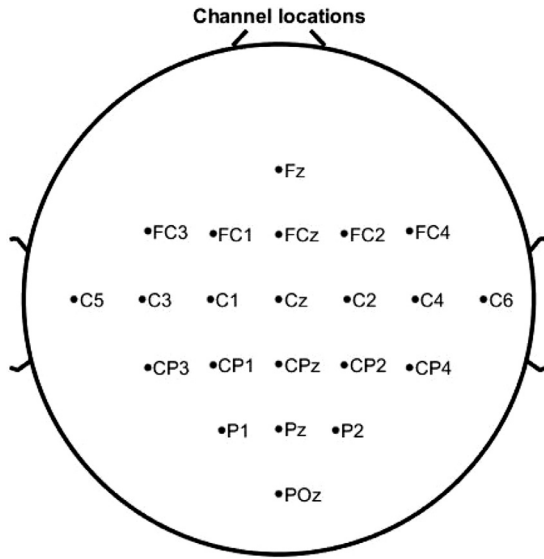


Fig. 2. Headplot showing all the channel locations.

of multivariate IMFs (MIMFs), described as follows:

$$Y(t) = \sum_{a=1}^n Q_a(t) + r_{s_a}(t) \quad (1)$$

where  $Y(t)$  denotes the original EEG signal in time-domain,  $Q_a(t)$  is  $a$ th IMF, and  $r_{s_a}(t)$  denotes the residual. Henceforth, a combination of IMFs can be selected to reconstruct the enhanced EEG signal of our interest and rest of the IMFs contributing to artifacts and noise can be discarded. However, single-channel EMD decomposition technique loses the cross-channel information and suffers from the mode mixing problem wherein specific frequency information is not localized. To handle this issue an ensemble empirical mode decomposition (EEMD) method has been studied in (Wu & Huang, 2009). It is a fact that this decomposition technique is computationally expensive and time-consuming. Furthermore, a multivariate version of updated EMD has been pro-

posed by Park, Looney, Ahrabian, and Mandic (2013), known as MEMD. This decomposition technique utilizes the cross-channel information and provides high localization of the specific frequency component. They proposed a noise-assisted MEMD (N-A MEMD) method (ur Rehman, Park, Huang, & Mandic, 2013) and addressed the mode mixing issue by adding white Gaussian noise to multiple channels. Here, we have used the MEMD method for the decomposition of EEG signals (Rehman & Mandic, 2010; Rilling, Flandrin, & Gonçalves, 2003; Huang et al., 2003). In calculation of MEMD, the mean  $m(t)$  is computed by means of the multivariate (multichannel) envelope curves, given as follows (Rehman & Mandic, 2010):

$$m(t) = \frac{1}{n} \sum_{b=1}^n e^{\theta_b}(t) \quad (2)$$

where  $b$  denotes the length of vectors and  $e^{\theta_b}(t)$  represents the multichannel envelope curves in  $n$ -dimensions of direction vectors. Then, the candidate IMF  $c(t)$  by  $c(t) = Y(t) - m(t)$  is calculated. If the stoppage criterion is satisfied then a candidate IMF becomes the multichannel IMF. Otherwise, the input  $Y(t)$  will be assigned to the remainder  $c(t)$  and the process of determining IMF for the remainder again is performed. The entire process is repeated until all multichannel IMFs are extracted (Park, Looney, Ahrabian, & Mandic, 2013; Rehman & Mandic, 2010). For more details refer to (Rehman & Mandic, 2010).

The main contribution of this work is to automatically identify the subject specific MIMFs and subject specific filtering range which contribute to  $\mu$  and  $\beta$  rhythms for a particular subject and discard the remaining of the MIMFs which are not identified. A mean frequency measure has been computed for all the MIMFs corresponding to the right hand, left hand, both feet and tongue MI tasks. The Mean frequency of each IMF is computed as the sum of a product of IMF power spectrum and the frequency divided by the total sum of IMF power spectrum in the frequency domain (Phinyomark, Phukpattaranont, & Limsakul, 2012; Pachori, 2008). The mathematical expression of mean frequency is denoted as

$$MF_{IMF} = \frac{\sum_{b=1}^n P_b f_b}{\sum_{b=1}^n P_b} \quad (3)$$

where  $n$  denotes the length of frequency bin, and  $P_b$  gives the power spectrum at the frequency bin  $b$ .  $f_b$  represents the frequency value at the frequency bin  $b$ . These computed mean frequencies represent centroids of the IMF in the spectrum in frequency domain. This mean frequency has been used to first automatically identify the subject specific MIMFs which provide a major contribution to  $\mu$  and  $\beta$  rhythms and then the sum of these identified MIMFs provides enhanced EEG signals corresponding to the left hand, right hand, foot and tongue MI tasks.

#### 4. Feature extraction

In BCI literature, the most studied EEG features are Hjorth parameters, band power, power spectral density (PSD) and bispectrum (BSP) features (Brunner et al., 2008; Davies & James, 2013; 2014; Leeb, Brunner, Müller-Putz, Schlögl, & Pfurtscheller, 2008; Lotte, Congedo, Lécuyer, & Lamarche, 2007; Shahid & Prasad, 2011). The most commonly used features in MI based BCI applications for classification of multiple MI tasks are Hjorth features and band power (Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002; Park, Looney, Kidmose, Ungstrup, & Mandic, 2011; Gaur, Pachori, Wang, & Prasad, 2015). Also, some of the research groups investigated spatial filters technique namely, CSP and its extension for multi-class MI tasks classification problems. This method enhances class separability by estimating the eigenvectors. Thereafter, the spatial patterns are derived from these eigenvectors (Ang et al., 2008).

Let  $x_j \in \mathbb{R}^r$  represent the enhanced EEG signal vector at a particular time instant  $j$  where  $r$  represents the number of electrodes. The formal definition of spatial covariance matrix is given by  $\text{Cov} = E\{(x_j - E\{x_j\})(x_j - E\{x_j\})^T\}$ , wherein  $E\{\cdot\}$  depicts the expected value and superscript  $T$  represents matrix transposition. For designing a BCI, a short-time segments are extracted from continuous EEG signals of a trial. They may be denoted in the form of a matrix  $X_i = [x_{T_i} \cdots x_{T_i+T_n-1}] \in \mathbb{R}^{r \times T_n}$  corresponding to  $i$ -th trial of an MI task wherein the segment began at time  $T_i$ .

Furthermore, the spatial covariance matrix (SCM) corresponding to the  $i$ -th trial is computed using the unbiased SCM  $Q_i \in \mathbb{R}^{r \times r}$  such as,

$$Q_i = \frac{1}{T_n - 1} X_i X_i^T \quad (4)$$

where  $T_n$  represents the number of time instants taken from each trial.

#### 5. Riemannian geometry framework

As shown in Fig. 3, the space of symmetric and positive definite (SPD) matrices  $Q(r)$  is denoted by a differentiable Riemannian manifold  $Z$  (Moakher, 2005). The tangent space for the derivatives of a matrix  $Q$  lies in a vector space  $T_Q$  over the manifold. The dimensions of the tangent space and the manifold are  $d_{ts} = r(r+1)/2$ .

The geodesic distance on the manifold is defined as the minimum length curve joining these two points and is denoted by,

$$\delta_R(Q_1, Q_2) = \|\log(Q_1^{-1}Q_2)\|_F = \left[ \sum_{i=1}^r (\log^2 \chi_i) \right]^{1/2} \quad (5)$$

where  $\chi_i$ ,  $i = 1, \dots, r$  denotes the real eigenvalues of  $Q_1^{-1}Q_2$  and  $r$  represents the number of channels. Each tangent vector  $P_i$  is seen as the derivative at  $t = 0$  of the geodesic  $D_i(t)$  between exponential mapping  $Q_i = \text{Exp}_Q(P_i)$  and  $Q$ , defined as,

$$\text{Exp}_Q(P_i) = Q_i = Q^{1/2} \exp(Q^{-1/2} P_i Q^{-1/2}) Q^{1/2} \quad (6)$$

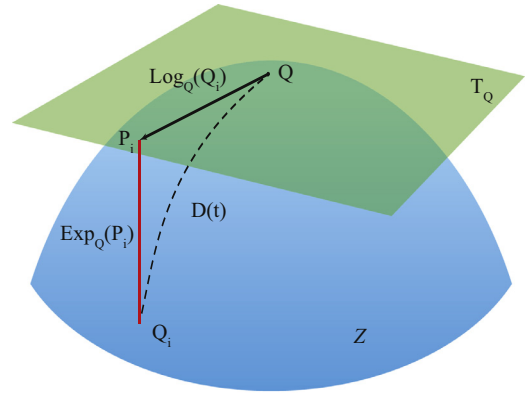


Fig. 3. The tangent space at point  $Q$ ,  $P_i$  is a tangent vector at  $Q$ ,  $D(t)$  is the geodesic between  $Q$  and  $Q_i$ .

The logarithmic mapping gives the inverse mapping which is defined as,

$$\text{Log}_Q(Q_i) = P_i = Q^{1/2} \log(Q^{-1/2} Q_i Q^{-1/2}) Q^{1/2} \quad (7)$$

The geometrical procedure has been presented in Fig. 3. Refer to (Barachant et al., 2012) for more details.

The Riemannian mean (geometric mean) of  $J \times 1$  SPD matrices using Riemannian geodesic distance (5) is denoted by,

$$G(Q_1, \dots, Q_J) = \underset{Q \in Q(r)}{\text{argmin}} \sum_{i=1}^J \delta_R^2(Q, Q_i) \quad (8)$$

Considering  $1 \times 1$  SPD matrices  $y_i > 0_{1 \times J}$ , this definition gives  $G(y_1, \dots, y_J) = \sqrt{y_1, \dots, y_J}$ . For each of the MI tasks, the class-wise covariance matrices  $Q_G^{(c)}$  have been calculated with the equations discussed above, where  $c = [1 : K]$  depicts the class indices. A Riemannian distance algorithm has been used to compute the distance between class-wise covariance matrix  $Q_G^{(c)}$  and an unknown test trial. The class giving the minimum distance is assigned to the unknown test trial  $X^{ts}$ . This procedure has been applied to assign the class to each of the new test trials. More details can be obtained from Barachant et al. (2012).

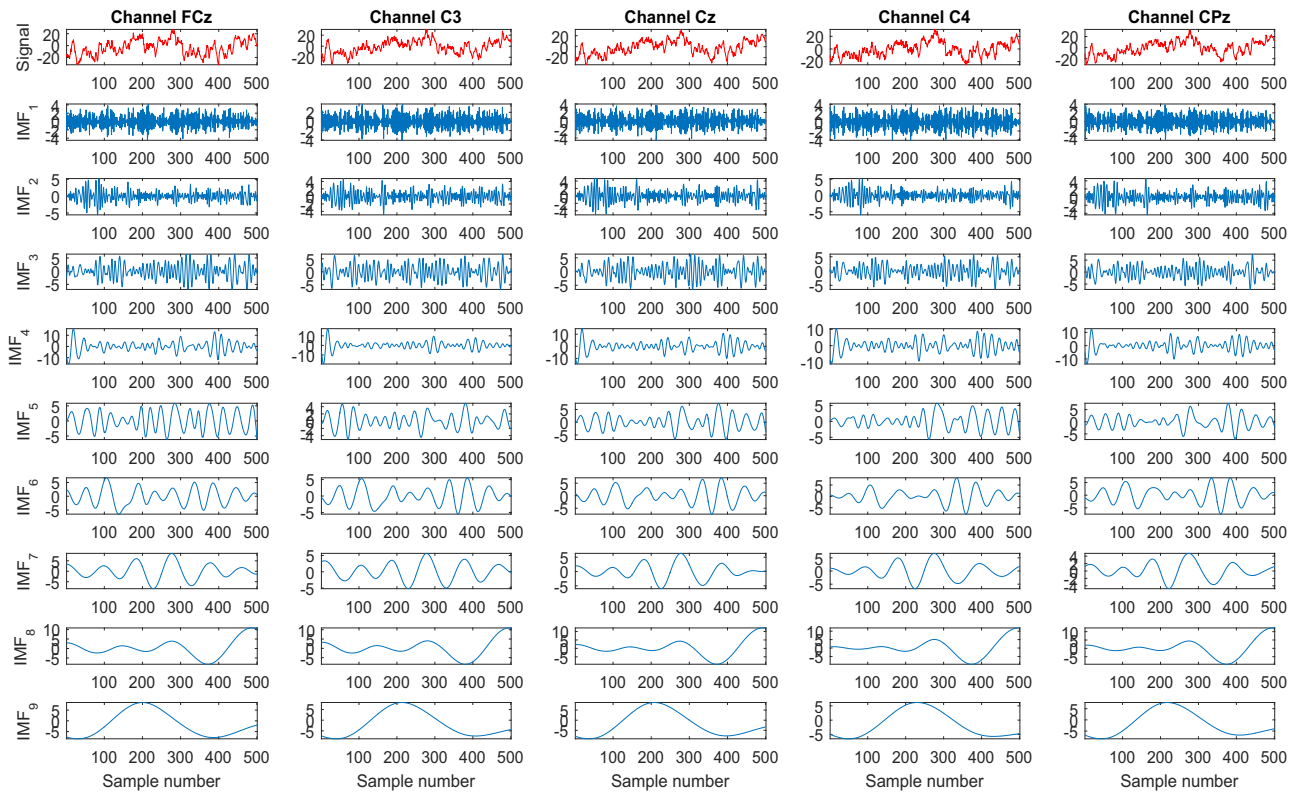
#### 6. Results and discussion

The decomposition dynamics of the MEMD technique may be explained using single trials of left hand and right hand MI EEG signals. Fig. 4(a) and (b) display the obtained MIMFs from the subject A01T. Similarly, the foot and tongue MI EEG signals and their obtained MIMFs are shown in Fig. 5(a) and (b) respectively. In training session, a five-fold cross-validation has been done to avoid the overfitting issue during the classification of multi-class MI EEG signals.

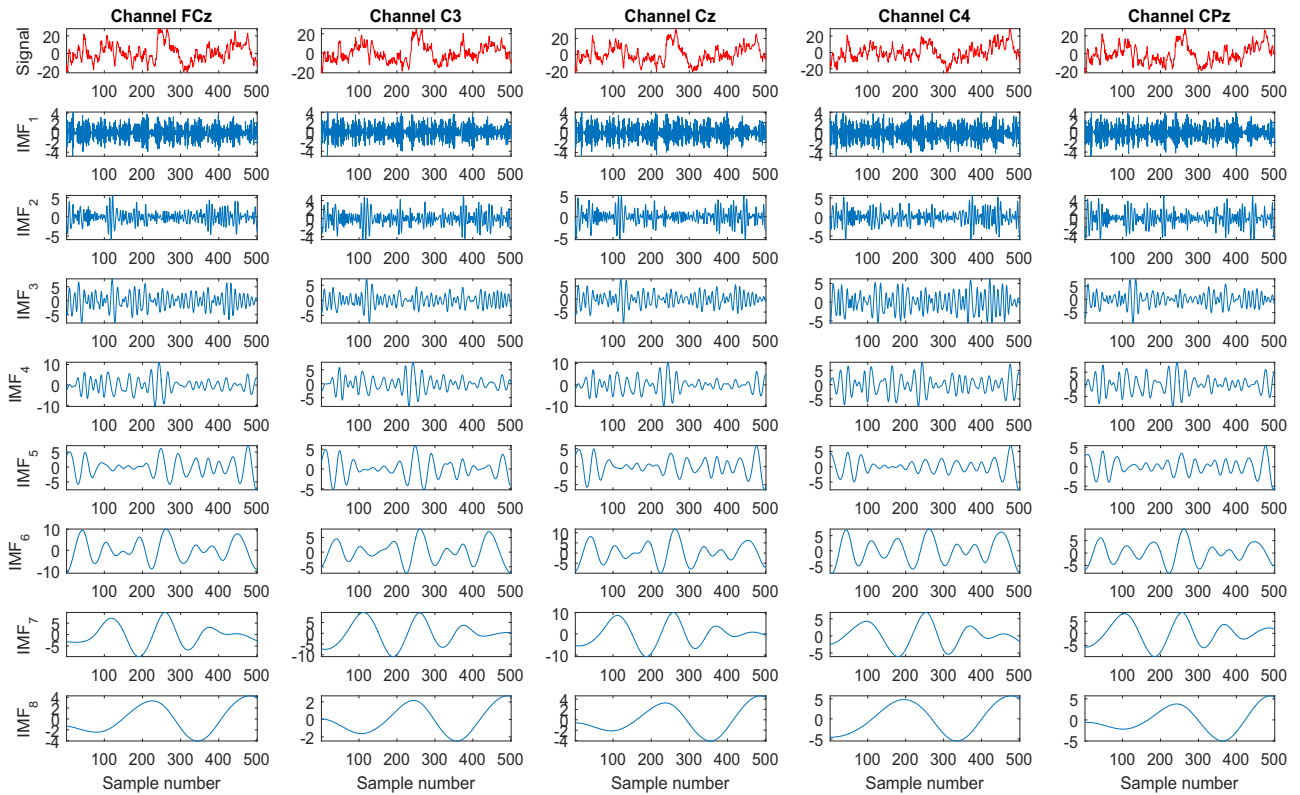
For the evaluation session, the Riemannian mean covariance matrix for each of the classes has been obtained with all training trials from the A0ST and evaluated on the all unknown test trials of the corresponding evaluation session A0SE on the trial-by-trial basis, where  $S$  represents the subject number. As the MI task began at 2 s, the covariance matrix feature is computed for the EEG signals from 2.5 s to 4.5 s time-interval of the MI paradigm. The Kappa value (between 0 and 1) and classification accuracy for all the nine subjects have been computed.

The results in Fig. 6(a) and (b) show the feature distribution of first five best features. Fig. 6(a) depicts the boxplot obtained using Kruskal-Wallis test (McKnight & Najab, 2010) for the first five best features with the SS-MEMDBF. After applying the SS-MEMDBF, the first five best features extracted have shown statistically significant



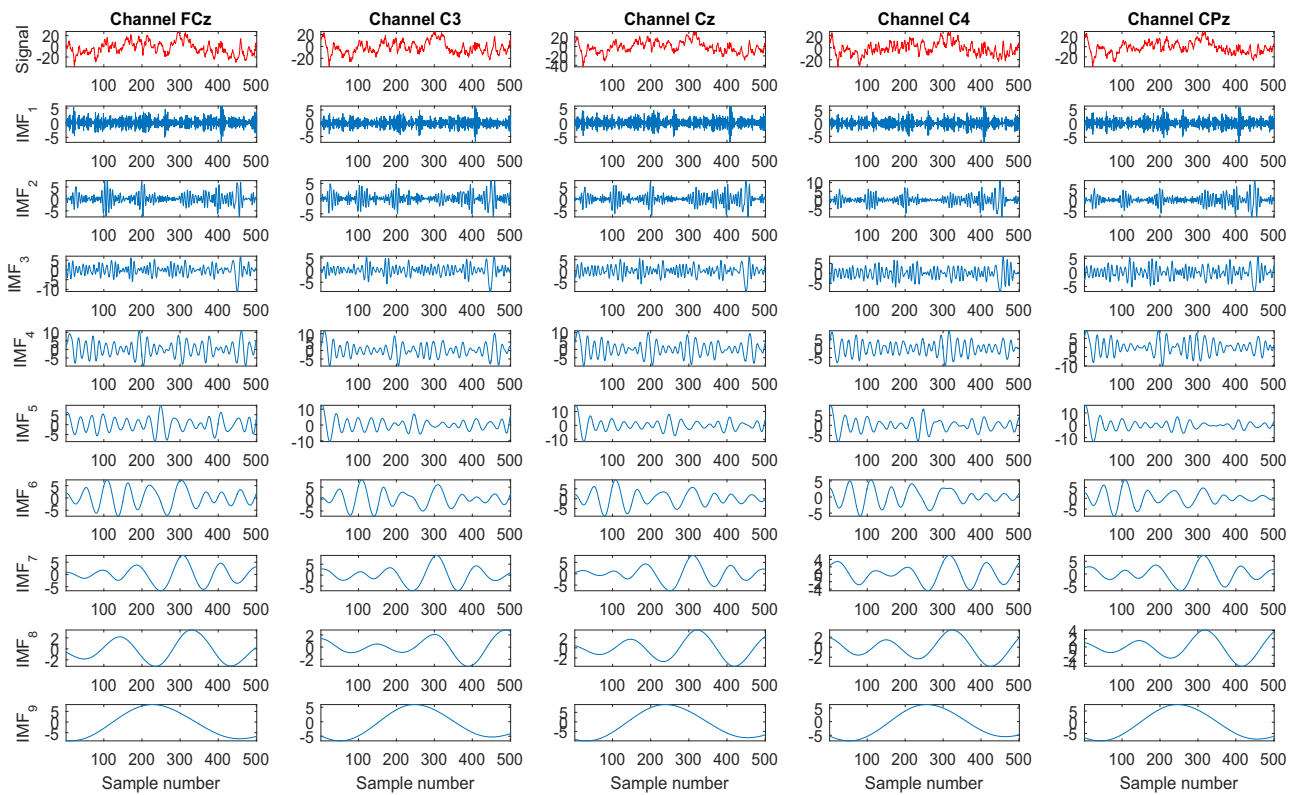


(a)

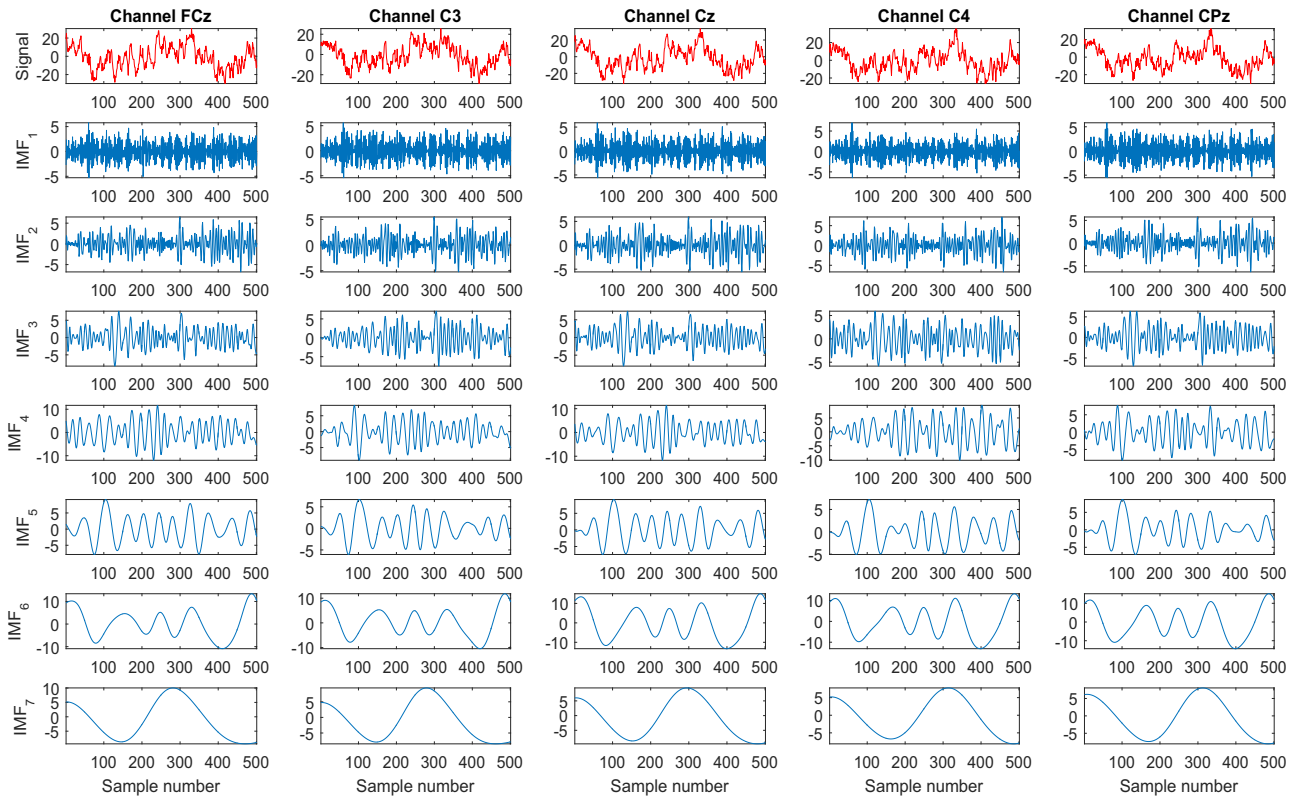


(b)

**Fig. 4.** The EEG signals and their MIMFs corresponding to channels FCz, C3, Cz, C4, and CPz of A01T for (a) Left hand MI task, (b) Right hand MI task.

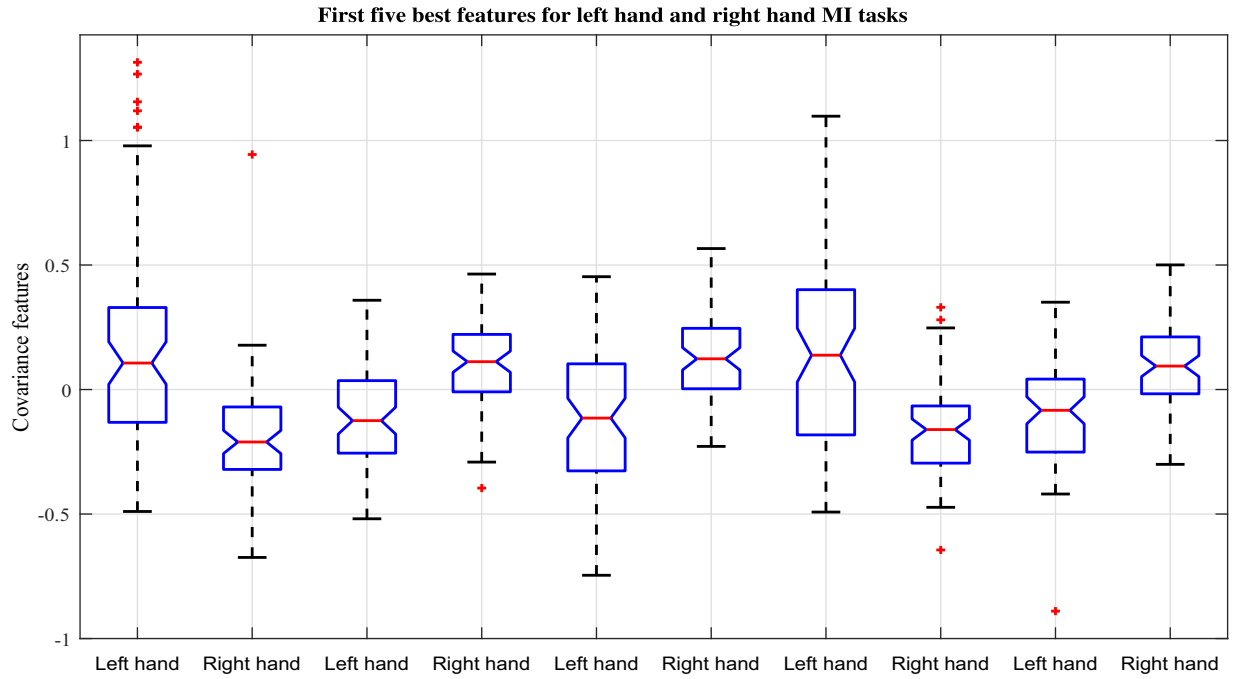


(a)

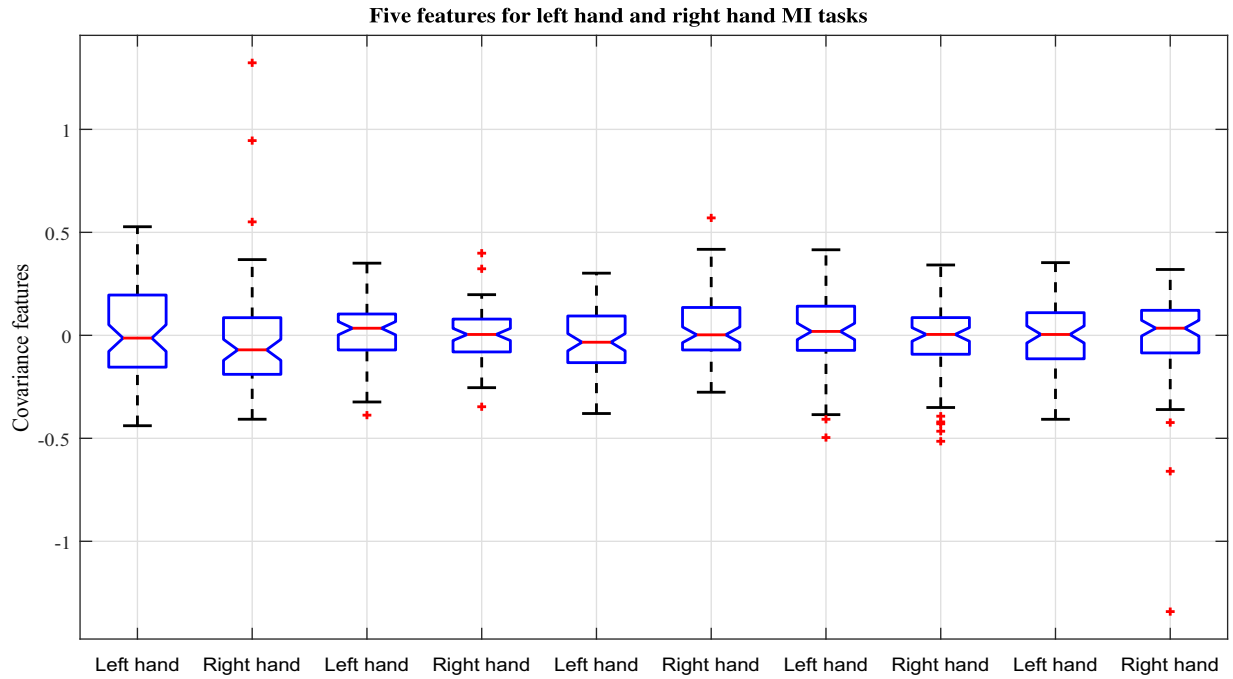


(b)

**Fig. 5.** The EEG signals and their MIMFs corresponding to channels FCz, C3, Cz, C4, and CPz of A01T for (a) Foot MI task (b) Tongue MI task.



(a) Kruskal-Wallis Test on enhanced EEG signals with SS-MEMDBF for first five best features.



(b) Kruskal-Wallis Test on raw EEG signals for same five features.

**Fig. 6.** (a) The box plot depicts the first five best features obtained in the training session for left hand and right hand MI tasks using SS-MEMDBF. With the proposed SS-MEMDBF method, the features obtained are statistically significant with  $p$ -values  $< 0.05$  in terms of separability. (b) The same five features computed from the raw EEG signals are not statistically significant with  $p$ -values (0.1065, 0.2861, 0.0643, 0.2111 and 0.6983) in terms of separability. All the results shown are for subject A01. The five best features have been obtained using non-parametric Wilcoxon test. The  $p$ -value is computed using the Kruskal–Wallis test.

improvement in the separability with  $p$ -value  $< 0.05$  in the training session for the left hand and right hand MI tasks. The feature selection has been done by ranking the available features with the Wilcoxon test method for left hand and right hand MI tasks. The feature set has been arranged based on the rank in the decreasing order of class separability. Then, we have selected the best five features giving the highest class discrimination for left hand and right

hand MI task classification as shown in Fig. 6(a) with the proposed method. For illustration, the first five features are shown although there are more features in the training session which are statistically significant in terms of features separability. Fig. 6(b) shows the same five features computed from the raw EEG signals giving  $p$ -values of 0.1065, 0.2861, 0.0643, 0.2111, and 0.6983. The  $p$ -values indicate that the two classes have no significant difference in their

**Table 1**  
Trials rejected from all subjects.

Subject	Total trials	Correct trials	Rejected trials	Left hand	Right hand	Foot	Tongue
A01	288	281	7	1	2	3	1
A02	288	283	5	1	1	3	0
A03	288	273	15	5	2	4	4
A04	288	228	60	13	15	13	19
A05	288	276	12	2	7	0	3
A06	288	215	73	19	17	18	19
A07	288	277	11	1	3	1	6
A08	288	271	17	6	4	3	4
A09	288	264	24	7	7	3	7

**Table 2**  
Classification accuracy (in %) for the proposed method with and without SS-MEMDBF with one vs one scheme applied on BCI competition IV dataset 2A.

Subject	Accuracy with SS-MEMDBF (in %)					
	LvR	LvF	LvT	RvF	RvT	FvT
A01	91.49	97.14	98.59	98.56	100	75
A02	60.56	78.57	67.13	80.71	68.53	74.47
A03	94.16	91.11	88.89	97.1	95.65	69.85
A04	76.72	91.53	87.5	92.24	86.36	66.07
A05	58.52	68.31	71.94	67.88	70.9	70.92
A06	68.52	69.16	73.58	66.97	75	67.29
A07	78.57	92.25	89.05	93.57	90.37	78.83
A08	97.01	88.89	83.58	91.97	83.09	81.02
A09	93.85	92.54	98.46	79.85	83.85	89.55
Average	79.93	85.5	84.3	85.43	83.75	74.78

feature distribution for left hand and right hand MI tasks. However, with the proposed method, the  $p$ -values show statistically significant difference in feature distribution.

**Table 1** gives the details of the trials rejected from each subject in the evaluation session as indicated with the event 1023 (Brunner et al., 2008). Subject A06 has maximum number of the trials rejected. The rejected trials corresponding to each MI tasks are as follow: left hand 19 trials, right hand 17 trials, foot 18 trials and tongue 19 trials respectively. This gives the total of 73 trials rejected for subject A06 in the evaluation session.

**Table 2** presents the results obtained using the pairwise binary classification for multiple MI tasks. Since there are four classes, we can have six possible pairs of imagery tasks: left vs right (LvR), left vs foot (LvF), left vs tongue (LvT), right vs foot (RvF), right vs tongue (RvT), foot vs tongue (FvT). The features obtained with the SS-MEMDBF have helped to achieve superior classification accuracy.

**Table 3** displays the comparison of classification accuracy using SS-MEMDBF and bandpass filtering (8–30 Hz) along with Riemannian geometry framework. It shows the comparison of SS-MEMDBF with CSP features as well as other state-of-the-art methods. Method-1 shows the results obtained by performing the bandpass filtering in the range of 8–30 Hz with the same time window from which features are extracted. These features have been classified using Riemannian geometry framework. With the proposed method, the group mean accuracy has improved by  $> 1.3\%$  ( $p = 0.7422$ ) and four of the nine subjects have shown improvement in the accuracy, while the subject A01 has shown no change. Method-2 displays the classification accuracy in the evaluation session using the proposed SS-MEMDBF with CSP features. In this case, we have obtained an improvement of  $> 1.6\%$  ( $p = 0.7422$ ) across the nine subjects. Method-3, method-4 and method-5 display the comparison of classification accuracy obtained in the evaluation session with other published results. We have achieved superior average classification accuracy across nine subjects in comparison to those obtained by three most recent advanced techniques, namely

the method-3 ( $p = 0.20$ ) (Lotte & Guan, 2011), the method-4 ( $p = 0.027$ ) (Raza, Cecotti, Li, & Prasad, 2015), and the method-5 ( $p = 0.0039$ ) (Raza, Cecotti, Li, & Prasad, 2015). These methods studied the binary classification to classify the left hand and right hand MI tasks. The method-3 implemented the CSP on bandpass-filtered EEG between 8 and 30 Hz. Then it computed the log variance from three pairs of filters in the feature extraction, further classified by linear discriminant analysis (LDA) (Lotte & Guan, 2011). Method-4 and method-5 studied the CSP feature and identified the covariate shift followed by adaptive learning and transductive learning respectively (Raza, Cecotti, Li, & Prasad, 2015). There is a substantial improvement in four of the nine subjects in terms of classification accuracy as reported in **Table 3** marked in boldface. The  $p$ -value has been computed using the Wilcoxon signed rank test using signrank command in Matlab. We have identified subject specific filtering range for all nine subjects for each pair of MI tasks presented in **Table 4**.

We have also computed the Kappa value using Riemannian geometry with one vs rest (OVR) scheme for the multi-class classification of MI tasks on the same dataset. The results obtained using the proposed methodology are presented in **Table 5**. The Kappa value measures the agreement across the two outcomes (Schlogl, Kronegg, Huggins, & Mason, 2007; Bakeman & Gottman, 1997). We have obtained subject specific filtering range providing best Kappa value shown in **Table 5** for the multi-class problem. As seen in the **Table 6** with subject specific MEMD based filtering, the average Kappa value across all the nine subjects has improved by 0.08 ( $p = 0.0508$ ) in the evaluation sessions as compared to 0.52 reported using the Riemannian geometry in (Barachant et al., 2012). The average Kappa value of 0.54 has been obtained with 5-fold cross-validation in training stage. In addition, there was the maximum improvement of 0.17 for the subject A08 when compared to that reported using Riemannian geometry in (Barachant et al., 2012). Subjects A04 and A09 have improved by 0.15 and 0.14 in Kappa value. There is an improvement in Kappa value with the SS-MEMDBF filtering in eight of the nine subjects considering only the evaluation session compared to that reported using Riemannian geometry in (Barachant et al., 2012).

## 7. Comparison with other published results

**Table 6** displays the Kappa values computed with the proposed methodology and comparison with other similar works based on  $p$ -value computation. The proposed SS-MEMDBF has shown substantial performance improvement in multi-class MI based BCI classification when compared to that reported using Riemannian geometry in (Barachant et al., 2012) and also BCI competition IV dataset 2A winners (Ang et al., 2008). The SS-MEMDBF done at preprocessing stage has thus helped to achieve a mean Kappa of 0.60 which is superior to all the results reported so far. The results obtained have been computed using a one versus rest mechanism to classify the MI EEG signals into multiple classes. Moreover, we



**Table 3**

Comparison of classification accuracy (in %) for LvR task of the proposed method with other published results using one vs one scheme applied on BCI competition IV dataset 2A.

Subject	Evaluation session comparison with other groups (in %)					
	Proposed method	Method-1	Method-2	Method-3 (Lotte & Guan, 2011)	Method-4 (Raza et al., 2015)	Method-5 (Raza et al., 2015)
A01	91.49	91.49	<b>92.91</b>	88.89	90.28	90.28
A02	60.56	<b>61.27</b>	59.86	51.39	54.17	57.64
A03	94.16	94.89	95.62	<b>96.53</b>	93.75	95.14
A04	<b>76.72</b>	75.86	66.38	70.14	64.58	65.97
A05	58.52	60	<b>62.96</b>	54.86	57.64	61.11
A06	68.52	70.37	68.52	<b>71.53</b>	65.28	65.28
A07	78.57	65	68.57	<b>81.25</b>	62.5	61.11
A08	<b>97.01</b>	96.27	96.27	93.75	90.97	91.67
A09	<b>93.85</b>	92.31	93.08	93.75	85.42	86.11
Average	79.93	78.61	78.24	78.01	73.84	74.92
p-value		0.7422	0.7422	0.2031	0.0039	0.0273

**Table 4**

Subject specific filtering range for all six possible MI tasks with the proposed method applied on BCI competition IV dataset 2A.

Subject	LvR		LvF		LvT		RvF		RvT		FvT	
	Low	high	Low	high	Low	high	Low	high	Low	high	Low	high
A01	9	30	9	25	9	25	4	26	8	25	9	26
A02	6	25	9	29	9	29	7	26	8	22	9	30
A03	7	27	7	23	4	25	8	24	5	26	10	26
A04	7	30	5	25	10	26	9	27	9	25	8	22
A05	10	26	10	30	4	30	9	28	4	26	4	25
A06	4	25	6	28	5	22	7	28	8	27	5	22
A07	10	27	10	29	6	29	10	30	6	27	7	27
A08	6	28	5	23	8	22	8	24	10	28	10	26
A09	5	24	6	22	5	23	6	25	5	25	4	26

**Table 5**

Kappa values of the proposed method with and without SS-MEMDBF applied on BCI competition IV dataset 2A.

Subject	Subject specific filtering range (%)		Kappa value with SS-MEMDBF		Kappa value with raw EEG	
	Lower	higher	Training	Evaluation	Training	Evaluation
A01	8	26	0.78	0.86	0.58	0.53
A02	7	26	0.16	0.24	0.25	0.21
A03	7	26	0.7	0.7	0.58	0.54
A04	6	28	0.4	0.68	0.22	0.38
A05	4	29	0.24	0.36	0.18	0.15
A06	6	28	0.22	0.34	0.33	0.28
A07	6	29	0.79	0.66	0.49	0.4
A08	8	22	0.77	0.75	0.77	0.61
A09	5	25	0.81	0.82	0.67	0.63
Average			0.54	0.60	0.45	0.41
p-value					0.0781	0.0039

**Table 6**

Kappa value comparison with other published results. The best result for each subject is displayed in bold characters.

Subjects	SS-MEMDBF	MDRM (Barachant et al., 2012)	Winner 1 (Ang et al., 2008)	Winner 2 (Guangquan et al., 2008)	Winner 3 (Song, 2008)	Winner 4 (Coyle, 2008)	Winner 5 (Jin, 2008)
A01	<b>0.86</b>	0.75	0.68	0.69	0.38	0.46	0.41
A02	0.24	0.37	<b>0.42</b>	0.34	0.18	0.25	0.17
A03	0.70	0.66	<b>0.75</b>	0.71	0.48	0.65	0.39
A04	<b>0.68</b>	0.53	0.48	0.44	0.33	0.31	0.25
A05	0.36	0.29	<b>0.40</b>	0.16	0.07	0.12	0.06
A06	<b>0.34</b>	0.27	0.27	0.21	0.14	0.07	0.16
A07	0.66	0.56	<b>0.77</b>	0.66	0.29	0.00	0.34
A08	0.75	0.58	<b>0.75</b>	0.73	0.49	0.46	0.45
A09	<b>0.82</b>	0.68	0.61	0.69	0.44	0.42	0.37
Average	<b>0.60</b>	0.52	0.57	0.52	0.31	0.30	0.29
p-value		0.0508	0.4844	0.0469	0.0039	0.0078	0.0039

have exploited the frequency domain information as done by the competition winner.

The top competition winner implemented filter bank CSP (FBCSP) and multiple one-against-the rest classifiers (Ang et al., 2008) and reported average Kappa value 0.57 across nine subjects; while the runner up implemented CSP on bandpass filtered data (between 8 and 30 Hz) and computed log variance of best eight components as the features and then classified by LDA and Bayesian classifiers and computed average Kappa 0.52 (Guangquan, Gan, & Xiangyang, 2008). The third group (Song, 2008) have applied bandpass filtering between 8 and 25 Hz and used a recursive channel elimination for the channels selection. They, thereafter extracted the CSP feature and used ensemble multi-class classifier using three SVM classifiers and computed mean Kappa value 0.31. The fourth group (Coyle, 2008) applied CSP on spectrally filtered neural time-series prediction pre-processing (NTSPP) signals at pre-processing stage and used the log variance of each filtered channel with a one second sliding window as features and then the best classifier among two variants of support vector machine (SVM) and three variants of LDA was chosen for each subject individually for classification purposes and calculated mean Kappa value 0.30.

It should be noted that there is a significant difference in the methodology followed between competitors and us: we have obtained the enhanced EEG signal using SS-MEMDBF and computed feature set as SCM for each trial across all the nine subjects, then did the classification using Riemannian mean distance. A research group however did study the usage of the Riemannian geometry but their pre-processing approach was very different and has reported mean kappa value of 0.52 (Barachant et al., 2012). Comparing these results, we have achieved highest Kappa value in four subjects and same Kappa value in one subject of the provided nine subjects in the evaluation session. These results clearly show that without the prior knowledge about the non-stationary characteristics of the EEG signals, the proposed SS-MEMDBF method has shown promising performance and thus, has potential to enhance the feature separability when incorporated as a pre-processing method and significantly enhance BCI applications.

## 8. Conclusion

The SS-MEMDBF method has been explored with sample covariance as a feature set to enhance the performance of multiple class MI based BCI. The main idea in the proposed method is to provide subject specific MEMD based filtering range in the preprocessing stage reducing the effect of the intra- and inter-subject non-stationarity present in the EEG signals. This preprocessing step provides the enhanced EEG signals from which the extracted feature's distributions have statistically significant differences. Also, mean frequency ranges have been identified when EEG signals are recorded from the same cortical areas across the subjects. We have obtained the results in terms of Kappa value and classification accuracy when single trials are classified. The filtering method is demonstrated to reduce the effect of intrinsic non-stationarity in the EEG signals to some extent. Also, it reduces the noise, and artifacts. Moreover, we have obtained highly significant performance improvement in binary class and multi-class classification problems of MI based BCI by enhancing the EEG signals at the pre-processing stage. In the SS-MEMDBF, a selection of one or multiple MIMFs is done when the MIMFs have mean frequencies falling in the frequency range of  $\mu$  and  $\beta$  rhythms specific to a subject. Further, the enhanced EEG signal obtained has helped to achieve improvement in the Kappa value while classifying EEG signals into left hand, right hand, foot, and tongue MI tasks when compared to raw EEG signals. The Kappa value obtained with the SS-MEMDBF has shown significant improvement

in both the training session and the evaluation session across the multiple subjects. Overall the results compare favourably with a research group that used Riemannian geometry and the BCI competition IV dataset 2A winners for the multi-class and binary class classification problems. Although SS-MEMDBF has helped to obtain enhanced feature separability and reduce the error rates due to intrinsic non-stationarities present in EEG signals to a large extent, adaptive classification methods may also be explored to handle the non-stationarities more efficiently.

The proposed method in this paper is studied on publicly available BCI competition IV dataset 2A for discrimination of two-class and four-class MI based EEG signals in offline mode. In future, it may be interesting to evaluate the proposed method in online classification problem in real time BCI using EEG or MEG recording techniques and also, to further extend the proposed method to classify the EEG or MEG into more than four classes. In future, a genetic algorithm may be applied to reduce the number of channels without compromising the classification accuracy.

## References

- Ang, K. K., Chin, Z. Y., Wang, C., Guan, C., & Zhang, H. (2012). Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. *Frontiers in Neuroscience*, 6.
- Ang, K. K., Chin, Z. Y., Wang, C., Guan, C., Zhang, H., Phua, K. S., Hamadicharef, B., & Tee, K. P. (2008). BCI competition IV results. [http://bbci.de/competition/iv/results/ds2a/KaiKengAng\\_desc.pdf](http://bbci.de/competition/iv/results/ds2a/KaiKengAng_desc.pdf).
- Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis*. Cambridge University Press.
- Barachant, A., Bonnet, S., Congedo, M., & Jutten, C. (2012). Multiclass brain-computer interface classification by Riemannian geometry. *IEEE Transactions on Biomedical Engineering*, 59, 920–928.
- Brunner, C., Leeb, R., Müller-Putz, G., Schlögl, A., & Pfurtscheller, G. (2008). BCI competition 2008–graz data set a. *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces)*, Graz University of Technology, 136–142.
- Coyle, D. (2008). BCI competition IV results. [http://bbci.de/competition/iv/results/ds2a/DamienCoyle\\_desc.pdf](http://bbci.de/competition/iv/results/ds2a/DamienCoyle_desc.pdf).
- Davies, S. R., & James, C. J. (2013). Novel use of empirical mode decomposition in single-trial classification of motor imagery for use in brain-computer interfaces. In *35th annual international conference of the IEEE engineering in medicine and biology society* (pp. 5610–5613).
- Davies, S. R., & James, C. J. (2014). Using empirical mode decomposition with spatio-temporal dynamics to classify single-trial motor imagery in BCI. In *36th annual international conference of the IEEE engineering in medicine and biology society* (pp. 4631–4634).
- Farquhar, J. (2009). A linear feature space for simultaneous learning of spatio-spectral filters in BCI. *Neural Networks*, 22, 1278–1285.
- Gandhi, V., Prasad, G., Coyle, D., Behera, L., & McGinnity, T. M. (2014). Quantum neural network-based EEG filtering for a brain-computer interface. *IEEE Transactions on Neural Networks and Learning Systems*, 25, 278–288.
- Gaur, P., Pachori, R. B., Wang, H., & Prasad, G. (2015). An empirical mode decomposition based filtering method for classification of motor-imagery EEG signals for enhancing brain-computer interface. In *Neural networks (IJCNN), 2015 International Joint Conference on* (pp. 1–7). IEEE.
- Gaur, P., Pachori, R. B., Wang, H., & Prasad, G. (2016a). Enhanced motor imagery classification in EEG-BCI using multivariate EMD based filtering and CSP features. In *Proceedings of the sixth international brain-computer interface meeting: BCI past, present, and future*.
- Gaur, P., Pachori, R. B., Wang, H., & Prasad, G. (2016b). A multivariate empirical mode decomposition based filtering for subject independent BCI. In *27th Irish signals and systems conference (ISSC) IEEE* (pp. 1–7).
- Guangquan, L., Gan, H., & Xiangyang, Z. (2008). BCI competition IV results. [http://bbci.de/competition/iv/results/ds2a/LiuGuangquan\\_desc.txt](http://bbci.de/competition/iv/results/ds2a/LiuGuangquan_desc.txt).
- Herman, P., Prasad, G., McGinnity, T. M., & Coyle, D. (2008). Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16, 317–326.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., et al. (1998). The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the royal society of london a: Mathematical, physical and engineering sciences*: 454 (pp. 903–995).
- Huang, N. E., Wu, M.-L. C., Long, S. R., Shen, S. S., Qu, W., Gloersen, P., & Fan, K. L. (2003). A confidence limit for the empirical mode decomposition and hilbert spectral analysis. *The Royal Society. Proceedings of the royal society of London A: Mathematical, physical and engineering sciences*, 2317–2345, 459.
- Jin, W. (2008). BCI competition IV results. [http://bbci.de/competition/iv/results/ds2a/JinWu\\_desc.txt](http://bbci.de/competition/iv/results/ds2a/JinWu_desc.txt).
- Leeb, R., Brunner, C., Müller-Putz, G., Schlögl, A., & Pfurtscheller, G. (2008). BCI competition 2008–graz data set b. *Graz University of Technology, Austria*.

- Lotte, F., Congedo, M., Lécuyer, A., & Lamarche, F. (2007). A review of classification algorithms for EEG-based brain–computer interfaces. *Journal of Neural Engineering*, 4.
- Lotte, F., & Guan, C. (2011). Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms. *IEEE Transactions on Biomedical Engineering*, 58, 355–362.
- McKnight, P. E., & Najab, J. (2010). Kruskal–Wallis test. *Corsini Encyclopedia of Psychology*.
- Moakher, M. (2005). A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 26, 735–747.
- Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors*, 12, 1211–1279.
- Pachori, R. B. (2008). Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition. *Research Letters in Signal Processing*, 2008, 14.
- Park, C., Looney, D., Ahrabian, A., Mandic, D. P., et al. (2013). Classification of motor imagery BCI using multivariate empirical mode decomposition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 21, 10–22.
- Park, C., Looney, D., Kidmose, P., Ungstrup, M., & Mandic, D. P. (2011). Time-frequency analysis of EEG asymmetry using bivariate empirical mode decomposition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 19, 366–373.
- Park, C., Plank, M., Snider, J., Kim, S., Huang, H. C., Gepshtein, S., et al. (2014). EEG gamma band oscillations differentiate the planning of spatially directed movements of the arm versus eye: Multivariate empirical mode decomposition analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22, 1083–1096.
- Pfurtscheller, G., Brunner, C., Schlögl, A., & Da Silva, F. L. (2006). Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks. *Neuroimage*, 31, 153–159.
- Pfurtscheller, G., Neuper, C., Flotzinger, D., & Pregenzer, M. (1997). EEG-based discrimination between imagination of right and left hand movement. *Electroencephalography and Clinical Neurophysiology*, 103, 642–651.
- Phinyomark, A., Phukpattaranont, P., & Limsakul, C. (2012). Feature reduction and selection for EMG signal classification. *Expert Systems with Applications*, 39, 7420–7431.
- Raza, H., Cecotti, H., Li, Y., & Prasad, G. (2015). Adaptive learning with covariate shift-detection for motor imagery-based brain–computer interface. *Soft Computing*, 1–12.
- Rehman, N., & Mandic, D. P. (2010). Multivariate empirical mode decomposition. The Royal Society. Proceedings of the royal society of London A: Mathematical, physical and engineering sciences, 1291–1302, 466.
- ur Rehman, N., Park, C., Huang, N. E., & Mandic, D. P. (2013). EMD via MEMD: Multivariate noise-aided computation of standard EMD. *Advances in Adaptive Data Analysis*, 5, 1350007.
- Rilling, G., Flandrin, P., & Gonçalves, P., et al. (2003). On empirical mode decomposition and its algorithms. IEEE. IEEE-EURASIP workshop on nonlinear signal and image processing, 8–11, 3.
- Schlogl, A., Kronegg, J., Huggins, J. E., & Mason, S. G. (2007). Evaluation criteria for BCI research. *Toward Brain-Computer Interfacing*.
- Shahid, S., & Prasad, G. (2011). Bispectrum-based feature extraction technique for devising a practical brain–computer interface. *Journal of Neural Engineering*, 8, 025014.
- Sharma, R., & Pachori, R. B. (2015). Classification of epileptic seizures in EEG signals based on phase space representation of intrinsic mode functions. *Expert Systems with Applications*, 42, 1106–1117.
- Song, W. (2008). BCI competition IV results. [http://bbci.de/competition/iv/results/ds2a/WeiSong\\_desc.txt](http://bbci.de/competition/iv/results/ds2a/WeiSong_desc.txt).
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, 113, 767–791.
- Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1, 1–41.
- Zhang, H., Yang, H., & Guan, C. (2013). Bayesian learning for spatial filtering in an EEG-based brain–computer interface. *IEEE Transactions on Neural Networks and Learning Systems*, 24, 1049–1060.