Towards Adaptive Classification using Riemannian Geometry approaches in Brain-Computer Interfaces

Satyam Kumar, Florian Yger, Fabien Lotte

▶ **To cite this version:**

HAL Id: hal-01924646

https://hal.inria.fr/hal-01924646v2

Submitted on 21 Apr 2020

Towards Adaptive Classification using Riemannian Geometry approaches in Brain-Computer Interfaces

Satyam Kumar
Indian Institute of technology, Kanpur
Inria
Kanpur, India / Bordeaux, France
satyamjuve@gmail.com

Florian Yger
LAMSADE, CNRS
Université Paris-Dauphine, Université PSL
Paris, France
florian.yger@dauphine.fr

Fabien Lotte
Inria / LaBRI
Bordeaux, France
fabien.lotte@inria.fr

Abstract—The omnipresence of non-stationarity and noise in Electroencephalogram signals restricts the ubiquitous use of Brain-Computer interface. One of the possible ways to tackle this problem is to adapt the computational model used to detect and classify different mental states. Adapting the model will possibly help us to track the changes and thus reducing the effect of non-stationarities. In this paper, we present different adaptation strategies for state of the art Riemannian geometry based classifiers. The offline evaluation of our proposed methods on two different datasets showed a statistically significant improvement over baseline non-adaptive classifiers. Moreover, we also demonstrate that combining different (hybrid) adaptation strategies generally increased the performance over individual adaptation schemes. Also, the improvement in average classification accuracy for a 3-class mental imagery BCI with hybrid adaption is as high as around 17% above the baseline non-adaptive classifier.

Index Terms—Remannian Geometry, BCI, Adaptive classifier

I. Introduction

ElectroEncephaloGraphy (EEG)-based Brain-Computer Interfaces (BCIs) have proven promising for many applications, ranging from communication and control for severely motor-impaired users, entertainment, mental state monitoring to stroke rehabilitation [1]. Despite this promising potential, BCIs are still scarcely used outside laboratories, arguably due to their poor reliability. Indeed, the mental commands from the users are often incorrectly recognized by the BCI, due to the low signal-to-noise ratio of EEG signals, to their non-stationarity and to the limited amount of calibration data available, among other [2]. Therefore, there is a pressing need for new approaches to deal with these limitations.

To do so, a variety of machine learning methods have been proposed, among which the most efficient ones include Riemannian Geometry-based Classifiers (RGC) and adaptive classifiers [2]. RGC represent EEG signals as covariance matrices, and can classify such matrices based on dedicated distance measures between them, known as Riemannian distances, see [3], [4] for reviews. RGC have been shown to be very effective, due to their affine invariance properties, their formulation removing the need for a separate spatial filter optimization, and their ability to be calibrated with little data [4]. RGC were actually used to win several international brain signal classification competitions [2], [3].

Another approach that proved effective as well to improve BCI performance is adaptive classifiers. Such classifiers are updating their parameters incrementally, according to incoming EEG data during BCI use, see [2], [5] for reviews. By doing so, such classifiers can adapt to EEG non-stationarities that can degrade performances when the calibration and testing data come from different distributions, as often with BCI. Adaptive classifiers have proven to be superior to non-adaptive ones (with fixed parameters) in BCI, both in offline and online study, and with both supervised (i.e., with the knowledge of the incoming EEG data label) and unsupervised adaptation [2].

Overall, both RGC and adaptive classifiers proved useful to improve BCI performances. A promising direction to improve BCIs further would thus be to explore adaptive RGC, combining the benefits of both approaches. This is the objective of this paper. It should be mentioned that a couple of adaptation strategies have been already explored for RGC. Notably, [6] explored an unsupervised adaptive re-referencing of covariance matrices, that was shown to improve performances as compared to non-adaptive RGC. Similar re-referencing methods have been also studied in [7], although not for RGC, as well as in [8] and [9] for RGC, although they were not compared to standard non-adaptive methods. Supervised adaptation was also explored for RGC in [10] for P300-BCI. In this paper we aim to go further. Indeed, the results we present in this paper suggest that we can improve RGC in Mental Imagery-based BCIs by making them adaptive, and even more so by combining adaptation strategies.

This paper is organized as follows: Section II presents in more details Riemannian geometry principles, while their use to design RGC in BCI is presented in Section III. Then Section IV describes the adaptation strategies for RGC that we explore, as well as the data sets on which we assess and compare them. Then, Section V presents the results, which are discussed in Section VI. Finally, Section VII concludes this paper.

II. Riemannian Geometry in brief

In this section, we discuss basic tools of Riemannian geometry for manipulating symmetric positive definite (SPD) matrices. Indeed, with RGC, EEG signals are represented as covariance matrices, which are SPD matrices. Furthermore,

we elaborate on Riemannian geometry based approaches for classification of EEG signals.

Let $X \in \mathbb{R}^{N_c \times N_s}$ denotes a Mental imagery trial of EEG signals, where N_c is the number of channels and N_s the number of temporal samples. Moreover, its normalized sample covariance matrix (SCM) is denoted by $C \in \mathbb{R}^{N_c \times N_c}$ and can be estimated as follows, as in [11]:

$$C = \frac{XX^T}{trace(XX^T)} \qquad (1)$$

The covariance matrices are symmetric positive definite, i.e., they have strictly positive eigenvalues. We denote the set of $n \times n$ symmetric matrices by $\mathcal{S}_n$ and $n \times n$ symmetric positive definite (SPD) matrices by $\mathcal{P}_n$.

Riemmanian Manifold: A Riemannian manifold is defined as a smooth manifold equipped with finite dimensional Euclidean tangent space (homogeneous to $\mathcal{S}_n$) at each point. Due to the constraint in positive definiteness, the SPD matrices $\mathcal{P}_n$ are restricted inside a cone of dimension $n(n+1)/2$. The shortest path (called *geodesic* for curved spaces) between two matrices P_1 and $P_2 \in \mathcal{P}_n$ can be written as [12]:

$$\gamma(P_1, P_2, t) = P_1^{\frac{1}{2}} \left(P_1^{-\frac{1}{2}} P_2 P_1^{-\frac{1}{2}} \right)^t P_1^{\frac{1}{2}} \quad t \in [0, 1] \qquad (2)$$

using Eq. 2 the distance between two matrices P_1 and P_2 on the Riemmanian manifold can be defined as [13]

$$\delta_r(P_1, P_2) = \int_0^1 \gamma(P_1, P_2, t)dt = \left\| \log(P_1^{-\frac{1}{2}} P_2 P_1^{-\frac{1}{2}}) \right\|_{\mathcal{F}} \qquad (3)$$

where the $\log(.)$ corresponds to matrix logarithm and $\|\cdot\|_{\mathcal{F}}$ is the frobenius norm of the matrix. The Riemmanian distance $\delta_r(.,.)$ in Eq. 3 possesses Affine invariance property [6] and is commonly known as Affine Invariant Riemannian Metric (AIRM) distance.

Tangent Space: The tangent space of $\mathcal{P}_n$ at P can be understood as a linearization of the manifold (see [3] for more details). For $\mathcal{P}_n$, the tangent space at any point P is homogeneous to $\mathcal{S}_n$. As for any riemannian manifold, there exists a mapping from manifold to tangent space at any given point and vice versa. Moreover, in the tangent space, we are allowed to use classical techniques for estimating means and other Euclidean tools [12].

Riemannian Mean: Similar to Euclidean mean, Karcher/Fréchet means extends the notion of mean/center of mass to $\mathcal{P}_n$ by estimating the SPD matrix which minimizes the sum of squared AIRM distances to all the SPD matrices in the set. Mathematically it is written as:

$$\bar{P} = \underset{P \in \mathcal{P}_n}{\arg\min} \sum_{i=1}^{N} \delta_r^2(P_i, P) \qquad (4)$$

Although the optimization problem in Eq. 4 has a unique minimum, it does not have any closed-form solution for $N > 2$ and it is estimated using several different optimization methods [12], [14]. In our study we use a gradient descent based method proposed by Barachant et al. [12].

III. RIEMANNIAN GEOMETRY-BASED CLASSIFIERS IN BCI

Using the mathematical tools presented above to manipulate SPD matrices, different classifiers can be constructed to discriminate SPD matrices. In particular, for BCI, EEG signals can be represented as covariance matrices (using Eq. 1 for Mental Imagery BCIs), and these matrices can further be used for classification. In this section we describe two different riemannian geometry based classifiers: Minimum distance to the Mean (MDM) and Fisher geodesic MDM (FgMDM).

A. MDM Classifier

Barachant et al. [12] proposed the Minimum Distance to the Mean (MDM) approach for classification of EEG trials in the Riemannian framework. Precisely it can be characterized with the following two steps:

1) **Training**: A Class prototype $\bar{C}_k$, $k \in K$ is computed for each class k using the karcher mean of the labelled trials collected from class k in the training session.

2) **Prediction**: For an incoming EEG trial i, its normalized spatial covariance matrix C_i is computed (Eq. 1). The MDM classifier assigns a label to the covariance matrix C_i, corresponding to the closest class prototype in $\bar{C}_1, \bar{C}_2, ..., \bar{C}_K$ for a K class classification problem, according to the AIRM distance.

B. FgMDM classifier

One of the main drawbacks of the MDM classifier is that it does not take into account the inter-class distribution. Barachant et al., [12] proposed Fisher Geodesic Discriminant Analysis for performing Geodesic filtering to make the classes more separable along the geodesics. Precisely FgMDM can be characterized as follows:

1) **Training**: Generally, training of the classifier is performed in the following three steps:

I Estimating the reference covariance matrix $\bar{C}_{train}$ (Karcher mean of all the training data from all classes) for projecting the covariance matrices to tangent space at $\bar{C}_{train}$.

II Estimating the discriminant (Euclidean) Fisher filters (W) in tangent space followed by filtering the tangent space features using the estimated filters. See [12] for details.

III Projecting back the filtered tangent space features to the Riemannian manifold and applying MDM training to estimate class prototypes.

2) **Prediction**: First an incoming trial covariance matrix is projected onto the tangent space of the reference covariance matrix $\bar{C}_{train}$ (estimated during the training). Next, the tangent space representation of this trial is filtered using the filter (W). In the last step, the filtered feature vector is projected back onto the manifold and classified according to the MDM prediction rule presented in the previous section.

IV. METHODS

A. Adaptive Riemannian classifiers

Shenoy et al. [15] introduced the notion of RETRAIN and REBIAS to study the adaptation of classifiers in BCIs.

In RETRAIN adaptation, the classifier is retrained (updated) using the data from the calibration session together with the labeled data acquired during the feedback stage. In REBIAS adaptation, the classifier trained on the calibration session data is used. However, the output of the classifier is then shifted in a way to adapt the shifts occurring in data distribution due to between-session changes. In this section, we present various methods to incorporate REBIAS and RETRAIN in the Riemannian geometry framework. Specifically we propose different adaptation strategies to adapt the MDM and FgMDM classifier in an (simulated) online BCI setting.

1) *Unsupervised Adaptive MDM*: This method is motivated by the RETRAIN approach. Initially, an MDM classifier is trained on the training/calibration session data. Then, during the testing session, the classifier is retrained after each prediction. More precisely, the class covariance prototype $\bar{C}^k$ corresponding to the predicted label k of the incoming trial is updated using geodesic interpolation (as introduced in Eq. 2) and the update rule is as follows:

$$\bar{C}_i^k = \begin{cases} \gamma\left(\bar{C}_{i-1}^k, C^k, \frac{1}{N_{\bar{C}_{i-1}^k}+1}\right) & i > 0 \\ \bar{C}_{train}^k & i = 0 \end{cases} \quad (5)$$

where $\bar{C}_i^k$ is the class prototype of the kth class after a total number of i trials has been used for estimating the class prototype corresponding to class k. C^k is the spatial covariance matrix of the incoming trial with predicted label k and $N_{\bar{C}_{i-1}^k}$ is number of trials used to estimate $\bar{C}_{i-1}^k$.

2) *Supervised Adaptive MDM*: This method is similar to that of the unsupervised framework of retraining the classifier. However, in supervised MDM approach, the classifier is updated according to the ground truth of the incoming trial. The update rule for class prototypes remains the same as that of Eq. 5.

3) *Rebias MDM*: Zanini et al. [6] proposed a transfer learning based framework to align the covariance matrices, which generally are shifted on the Riemannian manifold due to inter-session and inter-subject variability. Put simply, reference covariance matrices R (here the Karcher means of covariance matrices of all mental imagery trials, as suggested by [16]) from both the training and testing sessions, are moved to a common reference point (Identity matrix) using the following equation:

$$C_i^{Rebiased} = R^{-\frac{1}{2}} C_i R^{-\frac{1}{2}} \quad R \in \{R^{test}, R^{train}\} \quad (6)$$

Where R is a reference matrix used for shifting the covariance matrices and C_i corresponds to the covariance matrix of the ith trial. This thus reduces changes between training and testing covariance matrices distribution. To maintain the causality and mimic the scenario of an online BCI we propose an update rule for the online estimation of reference matrix as follows:

$$R_i^{test} = \begin{cases} R^{train} & i = 1 \\ C_{i-1} & i = 2 \\ \gamma\left(R_{i-1}^{test}, C_i, \frac{1}{i-1}\right) & i \geq 3 \end{cases} \quad (7)$$

After shifting the incoming trial of the testing set using Eq. 6 and 7, we use the classifier trained on affine transformed calibration data for prediction. In this adaptation scheme, we just shift the data without updating the class prototypes, thus we refer to it as REBIAS adaptation.

4) *Supervised Rebias MDM*: In supervised rebias MDM framework, first we train the MDM classifier on affine transformed calibration data. For prediction, we sequentially perform the REBIAS and RETRAIN in an online fashion for trials belonging to the testing set. For every new trial, rebiasing is done according to the update equation in Eq. 6 and 7 to align the incoming data with training data distribution. Then, the incoming trial is classified using MDM prediction framework. Finally, we update the class prototypes according to the ground truth of the incoming trial using Eq. 5.

5) *Unsupervised Adaptive FgMDM*: With Unsupervised Adaptive FgMDM, for an incoming trial (C) we first predict its label k using the FgMDM classifier. Next we update the reference covariance matrix ($\bar{C}_i$ at i^{th} incoming trial) used to estimate the tangent space elements using the proposed geodesic interpolation in Eq. 8

$$\bar{C}_i = \begin{cases} \gamma\left(\bar{C}_{i-1}, C, \frac{1}{N_{\bar{C}_{i-1}}+1}\right) & i \geq 1 \\ \bar{C}_{train} & i = 0 \end{cases} \quad (8)$$

Then we re-estimate the geodesic filters W^{new} by incorporating trial (C) and its predicted label k . Finally we update the class prototypes by geodesic filtering (*calibration and feedback data*) and recalculating the karcher mean of geodesically filtered trials corresponding to each class.

6) *Supervised Adaptive FgMDM*: This method is similar to unsupervised FgMDM adaptation. However, in supervised framework, we update the parameters (geodesic filters (W)) of the classifier using the ground truth label of the incoming trial. Moreover, we also use the ground truth label of the incoming trial for updating the geodesically filtered class prototypes.

7) *Rebias FgMDM*: This adaptation strategy is similar to that of Rebias MDM. First the calibration data is shifted using affine transform in Eq. 6. Next we calibrate an FgMDM on the shifted data. We use the geodesic adaptation scheme as proposed in Eq. 7 for estimating the reference covariance matrix for shifting the incoming trial. Finally, the FgMDM prediction framework is used to classify the shifted trial.

8) *Supervised Rebias FgMDM*: We do a supervised adaptation to update the parameters of FgMDM classifier trained on affine transformed calibration data. For an incoming trial, we first do an affine transformation using Eq. 6

and 7. The label of affine transformed incoming trial is predicted using FgMDM prediction framework. To update the parameters (W and class prototypes) of FgMDM classifier, the affine transformed trial is projected to tangent space of Identity matrix and discriminant filters (W) in tangent space are re-estimated (W^{new}) using the new trial and its true label. Finally the class prototypes are updated through filtering the calibration data using the new discriminant filters (W^{new}) followed by re-estimation of class prototypes (karcher mean of class specific filtered data).

B. Datasets

We evaluate our proposed framework on a public motor imagery dataset and an In-house mental imagery dataset.

1) BCI competition IV dataset IIa [17]: This dataset is composed of EEG signal recordings from 9 different subjects. EEG signals were recorded using 22 electrodes. In the experimental paradigm, subjects were asked to perform four different motor imagery tasks, i.e., left hand, right hand, foot, and tongue motor imagery. Training (session-1) and testing (session-2) sets were available for every subject. The same number of trials for all the MI tasks were provided for testing and training session. Each of the session had 72 trials for each of the four motor imagery classes. At the beginning of trial (t=0s) a fixation cross appeared on the screen, After two seconds (t=2s) a cue instructing motor imagery was presented. The subjects were asked to perform motor imagery until the fixation cross disappeared at t=6s.

2) In-House dataset: 18 BCI-naive subjects took part in this study, for 6 different sessions each (each on a different day). Subjects had to perform three different mental imagery (MI) tasks: 1) left-hand motor imagery, 2) mental rotation of a 3D geometric figure and 3) mental subtraction of a 2 digit number from a 3-digit number (both displayed on screen). EEG were recorded from 30 channels. Each session comprised 5 runs. During each run, subjects had to perform 45 trials (15 trials per task), each trial lasting 8s. At t=0s, a cross was displayed on screen. At t=2s, a "beep" announced the coming instruction and at t=3s, an arrow was displayed, the direction of which informed the subject which task to perform. Finally, at t=4.250s, for 4s, a visual feedback was provided in the shape of bar, whose length reflected the classifier output. More details about this dataset can be found in [18].

C. Preprocessing and Evaluation Strategy

In both the datasets, the EEG signals are band-pass filtered in 8-30Hz (using 5th order butterworth filter for In-house dataset and 50th order Finite Impulse Response (FIR) filter for BCI competition dataset) containing both the mu and beta rhythms, which are key for mental imagery classification. Furthermore, EEG signals are extracted from 0.5s to 3.5s after the instruction cue was presented. The spatial covariance matrix for motor imagery trials is then estimated using shrinkage based covariance estimator [19] to avoid any numerical problems and provide better estimates, and thus possibly better

BCI performances [20]. Moreover, the covariance matrix of tangent space features in FgMDM is also estimated using shrinkage based estimator [21]. All the implementations are performed in Matlab (R2018a) running on Intel i7-6500U CPU @ 2.50GHz processor and 16GB of RAM using covariance[1] and RCSP [2] toolboxes. We use session 2 as evaluation data and session 1 as calibration data on BCI competition data set. For the In-House data set, session 1 is used as calibration data and session 2,3,4,5,6 are independently used as evaluation data.

V. RESULTS

We present the classification performance of the different methods discussed in the paper on the two datasets. Table I and II present the mean and standard deviation of the classification accuracy (averaged across all the subjects) on a session to session transfer evaluation for MDM and FgMDM schemes.

BCI Competition dataset: A one factor *Method* repeated measure ANOVA (Methods: MDM {Baseline, Unsupervised, Supervised, Rebias, Supervised Rebias} Dependent variable: Accuracy) with Greenhouse-Geisser Correction determined that classification accuracies differed significantly across different methods (F(1.788,14.308)=18.564, $p < 0.05$). Post-hoc tests using the Bonferroni correction revealed that baseline performance (62.81% ± 17.71 %) is significantly better compared to unsupervised adaptation (53.62% ± 21.23%). Moreover, Supervised and Rebias MDM outperformed the baseline by 2% and 1.88% but do not show a statistically significant difference. Furthermore Supervised Rebias adaptation (66.82% ± 16.61%) outperformed the baseline MDM by 4.01% and showed a tendency towards statistically significant difference ($p = 0.052$).

A one factor *Method* repeated measure ANOVA (Methods: FgMDM, same dependent and independent variable as previous ANOVA) with Greenhouse-Geisser Correction determined that classification accuracies differed significantly across different adaptive methods for FgMDM (F(1.356,10.847)=6.105, $p < 0.05$). Post-hoc tests using the Bonferroni correction revealed that baseline FgMDM (68.94% ± 15.77%) outperformed the unsupervised adaptation of FgMDM (66.70% ± 22.69%) however not statistically significantly. Additionally, Supervised (74.30% ± 15.91%) and Rebias adaptation (71.10% ± 14.55%) outperformed the baseline by (5.36% $p = 0.008, 2.160\%$ $p = 0.267$). Moreover Supervised Rebias adaptation (73.418% ± 14.89%) of FgMDM outperformed the baseline by 4.4715% and showed a statistically significant difference ($p = 0.012$).

In-House dataset: Tables I and II present the average classification accuracy across subjects and testing sessions for different adaptive strategies of FgMDM and MDM classifier. A two way repeated ANOVA method*session (Methods: MDM {Baseline, Unsupervised, Supervised, Rebias, Supervised Rebias} Session: 2 to 6, Dependent variable: Accuracy) showed no significant interaction. However, main effect *Method*, using
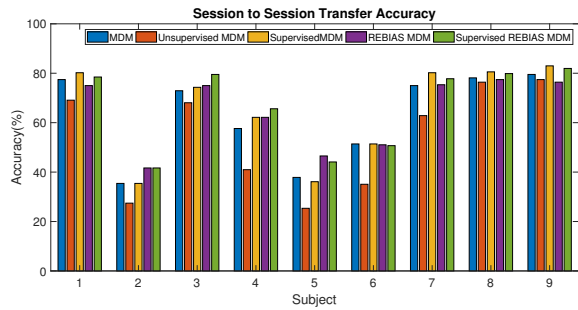
---

[1]https://github.com/alexandrebarachant/covariancetoolbox
[2]https://sites.google.com/site/fabienlotte/research/code-and-softwares

Fig. 1: Subject wise testing set classification accuracy on BCI competition dataset for different adaptation strategies of MDM classifier
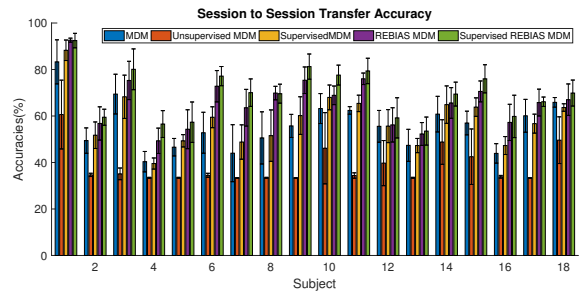

Fig. 2: Experimental results of Average classification accuracy across session 2-6 for different adaptation methods derived from MDM classifier on In-House dataset.

TABLE I: Mean Classification accuracies and Standard Deviation (std) in % obtained across subjects for MDM and different frameworks proposed in the paper

	BCI competition Dataset IIa		In-House Data	
	Mean	Std	Mean	Std
MDM	62.81	17.71	56.02	10.67
UnSupervised MDM	53.62	21.23	38.53	7.89
Supervised MDM	64.81	19.42	58.33	11.14
Rebias MDM	64.69	14.63	66.08	10.68
Supervised Rebias MDM	66.82	16.61	69.71	10.63

TABLE II: Mean Classification accuracies and Standard Deviation(std) in % obtained across subjects for MDM and different frameworks proposed in the paper

	BCI competition Dataset IIa		In-House Data	
	Mean	Std	Mean	Std
FgMDM	68.94	15.77	58.50	9.95
UnSupervised FgMDM	66.70	22.69	57.63	12.47
Supervised FgMDM	74.30	15.91	72.67	8.94
Rebias FgMDM	71.10	14.55	70.35	9.46
Supervised Rebias FgMDM	73.41	14.89	76.19	9.13

Greenhouse-Geisser correction showed a significant difference on classification accuracy between different methods $(F_{(1.90,32.30)}=118.89$, $p<0.05)$. Post-hoc tests using Bonferroni correction revealed that baseline $(56.02\% \pm 10.67\%)$ outperformed the Unsupervised adaptation $(38.53\% \pm 7.89\%)$ and the improvement is statistically significant $(p < 0.005)$. Supervised and Rebias adaptation outperformed the baseline by $(2.31\% \quad p = 0.049)$ and $(10.06\% \quad p < 0.005)$ respectively and the improvements are statistically significant at 5% level. Finally the Supervised rebias adaptation outperformed the baseline by 13.69% and shows a statistically significant improvement $p < 0.005$.

Concerning the FgMDM accuracies, a two way repeated ANOVA method*session (Methods: FgMDM {Baseline, Unsupervised, Supervised, Rebias, Supervised Rebias} Session: 2 to 6, Dependent variable: Accuracy) showed no significant interaction. However, the main effect *Method*, using Greenhouse-Geisser correction showed a significant difference on classification accuracy between different methods $(F_{(1.567,26.643)}=49.81$, $p<0.005)$. Post-hoc tests using Bonferroni correction confirmed that baseline $(58.50\% \pm 9.95\%)$ outperformed the unsupervised adaptation $(57.63\% \pm 12.47\%)$ slightly by 0.87% and the improvement is statistically insignificant. Supervised adaptation and Rebias adaptation outperformed the baseline by 14.17% and 12.55% and the improvement over baseline in both of them is statistically significant $(p < 0.05)$. Finally, Supervised rebias adaptation $(76.19\% \pm 9.13\%)$ showed a statistically significant improvement of $(17.69\% \quad p < 0.05)$ in classification performance over baseline.

VI. DISCUSSION

The results obtained using different methods gave several insights concerning the adaptation of classification algorithms during BCI experiments of mental imagery. The unsupervised adaptation of both FgMDM and MDM RGCs results in a decrease in classification performance compared to baseline on both the datasets. As we continuously update the classifier using the label from the predicted trial, a wrong prediction exacerbates the drop in classification accuracy and propagates the error to the upcoming classification models. Moreover, the unsupervised adaptive FgMDM suffers from a lesser decrease in contrast to MDM. This is probably due to improved classification performance of baseline FgMDM over MDM.

The increase in performance for the Supervised MDM classifier compared to baseline MDM on both data sets is similar, and around 2%. The improvement in classification accuracy is probably mainly due to an increase in the number of samples used for calibration and hence to a more precise estimate of class prototypes for classification. Moreover, the supervised adaptation might also helped in adapting the classifier to capture the class-specific change in EEG signals, due to fatigue and/or user training. The performances of supervised and unsupervised frameworks validate our proposal of the online update technique Eq. 5 of class prototypes. Also, the increase in performance of supervised FgMDM adaptation is much higher in comparison to that of supervised MDM. This observation can be possibly attributed to supervised adaptation in geodesically filtered space and thus allowing to have more precise class prototypes as well as better separability after discriminant filtering.

Supervised MDM outperformed Rebias MDM on BCI competition data maybe because subjects were not naive BCI users and hence produced more stable EEG patterns. Thus increasing the training data size helped to get more precise estimates of class prototypes as compared to shifting the incoming trials. However, on the In-house data set as the subjects were naive, they may have produced more shifts in MI trials (due to learning and/or trying out various strategies), which may explain why Rebias MDM outperformed the supervised adaptation. Interestingly, Supervised FgMDM outperformed Rebias FgMDM on both datasets. One possible reason is that complete retraining in supervised FgMDM would have lead to a better separation of geodesically filtered covariance matrices and the shifts occurring due to intersession changes are not much elicited in comparison to supervised adaptation.

Supervised Rebias MDM approach outperformed all the other algorithms on both the BCI competition and In-House datasets. This adaptation framework is superior to others because we are doing the rebiasing of incoming data thus making the distribution of incoming trials similar to calibration data and then a supervised adaptation which increases the number of trials for RETRAIN. Supervised Rebias FgMDM adaptation outperformed all the algorithms on the In-House data. However, it is outperformed by supervised FgMDM adaptation on BCI competition data. A possible reason is that when doing a supervised adaptation, we update the reference point to estimate the tangent space features of all the trials continuously and hence a precise estimate of geodesic filters is obtained. However, in supervised Rebias adaption of the FgMDM framework, we always estimate the tangent space features at Identity and thus with time we deteriorate the estimation of tangent space leading to worst performances. However, the supervised Rebias adaptation has much lower computational time in comparison to Supervised FgMDM adaptation as we just need to estimate the tangent space feature of the incoming trial at Identity compared to the recalculation of tangent space features for every trial (calibration+feedback).

From results in section V we demonstrate that adaptation schemes are efficient and results in improvement over non-adaptive MDM and FgMDM RGC. Our results are also in line with literature depicting the superior performance of Fg-MDM compared to MDM. Furthemore, for all the adaptation strategies, adaptation of FgMDM classifier outperformed the similar adaptation on MDM classifier. This further reinforces the main idea behind the formulation of FgMDM i.e. using the between class information for classification.

VII. CONCLUSION

In this paper, we proposed several schemes for adaptation of Riemannian geometry based classifiers. We validated the effectiveness of our proposal of adaptation schemes on two different types of Mental imagery datasets. Precisely we demonstrated that the different adaptation strategies (except the unsupervised one) outperformed their corresponding baseline RGCs. Moreover, we also demonstrated the effectiveness of hybrid adaptation schemes. Our proposed approaches for adaptation

can also be directly used in an online setting for classification of mental imagery data. One possible direction of future work would be to use this adaptive classifier for feedback training for control of real-time, online mental imagery-based BCIs. Another direction could be to optimize the forgetting factor and geodesic adaptation parameter t in Eq. 2 to perform an optimal rebias, with a speed adapted to each user and context.

REFERENCES

[1] A. Nijholt C. S. Nam and F. Lotte. *Brain–Computer Interfaces Handbook: Technological and Theoretical Advances*. CRC Press, 2018.

[2] F. Lotte, L. Bougrain, A. Cichock, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger. A review of classification algorithms for EEG-based brain-computer interfaces: A 10-year update. *J Neur Eng*, 2018.

[3] F. Yger, M. Berar, and F. Lotte. Riemannian approaches in brain-computer interfaces: a review. *IEEE Trans Neur Syst Rehab*, 2017.

[4] M. Congedo, A. Barachant, and R. Bhatia. Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review. *Brain-Computer Interfaces*, pages 1–20, 2017.

[5] J. Mladenovic, J. Mattout, and F. Lotte. A generic framework for adaptive EEG-based BCI training and operation. In CS. Nam, A. Nijholt, and F. Lotte, editors, *Brain–Computer Interfaces Handbook: Technological and Theoretical Advances*. Taylor & Francis, 2017, in press.

[6] P. Zanini, M. Congedo, C. Jutten, S. Said, and Y. Berthoumieu. Transfer learning: a riemannian geometry framework with applications to brain–computer interfaces. *IEEE Trans Biomed Eng*, 2018.

[7] B. Reuderink, J. Farquhar, M. Poel, and A. Nijholt. A subject-independent brain-computer interface based on smoothed, second-order baselining. In *Proc. IEEE EMBC*. IEEE, 2011.

[8] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten. Classification of covariance matrices using a riemannian-based kernel for bci applications. *Neurocomputing*, 2013.

[9] F. Yger and M. Sugiyama. Supervised logeuclidean metric learning for symmetric positive definite matrices. *arXiv:1502.03505*, 2015.

[10] A. Barachant and M. Congedo. A plug&play P300 BCI using information geometry. *arXiv preprint arXiv:1409.0107*, 2014.

[11] H. Ramoser, J. Muller-Gerking, , and G. Pfurtscheller. Optimal spatial filtering of single trial eeg during imagined hand movement. *IEEE Trans Rehabil Eng*, 2000.

[12] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten. Riemannian geometry applied to bci classification. In *Proc LVA-ICA*, 2010.

[13] M. Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM J Matrix Anal Appl*, 2005.

[14] B. Jeuris, R. Vandebril, and B. Vandereycken. A survey and comparison of contemporary algorithms for computing the matrix geometric mean. *Electronic Transactions on Numerical Analysis*, 2012.

[15] P. Shenoy, M. Krauledat, B. Blankertz, R. PN Rao, and K. R. Müller. Towards adaptive classification for bci. *J Neur Eng*, 2006.

[16] H. He and D. Wu. Transfer learning for brain-computer interfaces: An euclidean space data alignment approach. *arXiv:1808.05464*, 2018.

[17] M. Naeem, C. Brunner, R. Leeb, B. Graimann, and G. Pfurtscheller. Seperability of four-class motor imagery data using independent components analysis. *J Neur Eng*, 2006.

[18] C. Jeunet, B. NKaoua, S. Subramanian, M. Hachet, and F. Lotte. Predicting mental imagery-based bci performance from personality, cognitive profile and neurophysiological patterns. *PloS one*, 2015.

[19] O. Ledoit and M. Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *J Multivar Anal*, 2004.

[20] F. Lotte. Signal processing approaches to minimize or suppress calibration time in oscillatory activity-based brain–computer interfaces. *Proc IEEE*, 2015.

[21] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller. Single-trial analysis and classification of erp componentsa tutorial. *NeuroImage*, 2011.