RIEMANNIAN GEOMETRY ON CONNECTIVITY FOR CLINICAL BCI

Marie-Constance Corsi[‡] *Florian Yger*[§*] *Sylvain Chevallier*[¶] *Camille Noûs*[†$]

[‡] Inria Paris, Aramis project-team, Paris Brain Institute, Paris, France
[§] LAMSADE, Univ. Paris-Dauphine, Paris, France
[¶] LISV, Univ. Paris-Saclay, Versailles, France
[$] Cogitamus, CNRS, Paris, France

ABSTRACT

Riemannian BCI based on EEG covariance have won many data competitions and achieved very high classification results on BCI datasets. To increase the accuracy of BCI systems, we propose an approach grounded on Riemannian geometry that extends this framework to functional connectivity measures. This paper describes the approach submitted to the Clinical BCI Challenge-WCCI2020 and that ranked 1st on the task 1 of the competition.

***Index Terms*—** Riemannian geometry, functional connectivity, ensemble learning, BCI

1. INTRODUCTION

Using a brain-computer interfaces (BCIs) is a learned skill that requires time to reach high performance [1]. Despite its clinical applications [2, 3], one of the main drawbacks is the high inter-subject variability that could be noticed for performance. This is sometimes referred in the literature as the "BCI inefficiency" phenomenon [4] and affects its usability. Among the approaches adopted to tackle these issues are the search for neuromarkers, that potentially capture better the neurophysiological mechanisms underlying the BCI performance [5, 6], and the optimization of classification pipelines [7], that could be robust enough to be applied to any subject.

In this work, we proposed an original approach that combines functional connectivity estimators, Riemannian geometry and ensemble learning to ensure a robust classification. This article not only describes the proposed approach but also presents the methodology and the results that were conducted for our submission.

2. RIEMANNIAN GEOMETRY

As pointed out in [7], the use of Riemannian geometry for Motor Imagery BCI is one of the breakthroughs of the last 10 years of research in BCI and is now the golden standard.

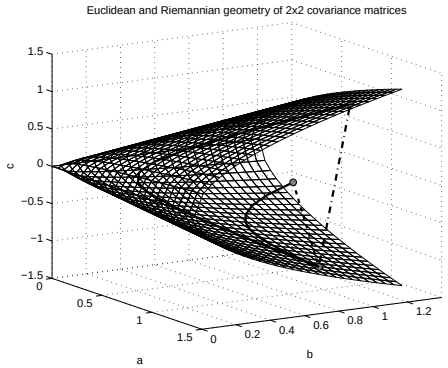

Fig. 1. Comparison of Euclidean and Riemannian geometries for 2×2 SPD matrices.

The approach consists in extracting Symmetric Positive Definite (SPD) matrices that are symmetric matrices with strictly positive eigenvalues, usually the covariance matrices among sensors, for each epoch and then in considering this space as a curved (i.e. Riemannian) space. As illustrated in Fig. 1 for 2×2 matrices, the space of SPD matrices could be considered as a Euclidean space (as a subspace of the Symmetric matrices) but several drawbacks occur (e.g. *swelling effect* - see [8]). Those drawbacks are leveraged when the Riemannian geometry is used and the distance between two SPD matrices A and B is expressed as :

$$\delta_R(A, B) = || \log \left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}} \right) ||_{\mathcal{F}} \quad (1)$$

with $\log(\cdot)$ the matrix logarithm and $|| \cdot ||_{\mathcal{F}}$ the Frobenius norm.

However, in practice, we will favor another Riemannian geometry with similar properties but being faster to compute[1], the LogEuclidean distance :

$$\delta_{LE}(A, B) = || \log(A) - \log(B) ||_{\mathcal{F}} \quad (2)$$

In both geometries, the Kärcher average of a set of matri-

*FY acknowledges the support of the ANR as part of the "Investissements d'avenir" program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute).

†For more information,see https://www.cogitamus.fr/camilleen.html

[1]The relationship between those geometries is developed in [9].

ces $\{X_1, \cdots X_n\}$ is defined as :

$$\min_X \sum_{i=1}^{n} \delta^2(X_i, X) \tag{3}$$

A closed-form solution exists for δ_{LE} (but not for δ_R)

$$\bar{X}_{LE} = \exp\left(\frac{1}{n}\sum_{i=1}^{n} \log(X_i)\right) \tag{4}$$

A simple, yet efficient classifier for SPD matrices consists in computing the Kärcher average of each class and then in predicting for a given test sample the class which average is the closest (using δ).

The interested reader can refer to [8, 10] for more details. Until now, the Riemannian geometry was applied on SPD matrices extracted from covariances among sensors but other characteristics extracted from the EEG signal could produce SPD matrices. The next section will describe an alternative way to obtain SPD matrices based on functional connectivity.

3. FUNCTIONAL CONNECTIVITY

Functional connectivity (FC), which consists of assessing the interaction between different brain areas [11], can be a valuable tool to provide alternative features to discriminate subjects' mental states [12] and to study neural mechanisms underlying BCI learning [13]. Here, as an exploratory study, we considered complementary undirected FC estimators to assess which of them, associated to Riemannian geometry, could best classify the data. For a given FC estimator, we took into account a time window of [3, 7.5 s] and we averaged the FC values within the alpha-beta band [8, 30 Hz]. Computations were made using the Brainstorm toolbox [14]. In the following subsections, we defined the metrics computed between two given signals referred as $s_1(t)$ and $s_2(t)$ between two EEG sensors.

3.1. Spectral estimation

We computed two spectral estimators: the coherence (Coh) and the Imaginary coherence (ICoh). Coh and ICoh are both computed from the coherency, defined as the normalized cross-spectral density obtained from two given signals. More specifically, they are obtained as follows:

$$Coh_{12}(f) = \frac{|S_{12}(f)|^2}{S_{11}(f).S_{22}(f)} \tag{5}$$

$$ICoh_{12}(f) = \frac{\Im S_{12}(f)}{\sqrt{S_{11}(f).S_{22}(f)}} \tag{6}$$

with $S_{12}(f)$ the cross-spectral density and $S_{11}(f)$ the auto-spectral density.

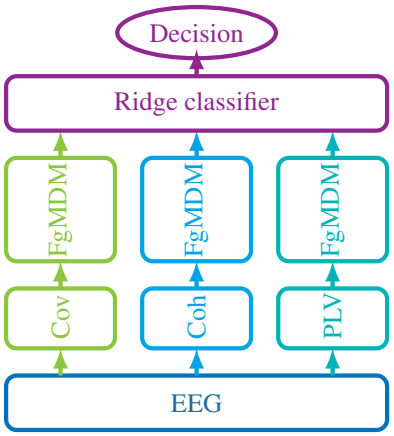

Fig. 2. Classification pipeline: coherence, phase locking value and spatial covariances are estimated from the EEG signal. A first level of classification was performed by FgMDM classifiers, that yielded output decision probabilities to train a second level classifier, a ridge regression classifier, that provided the final decision.

3.2. Phase estimation

As a phase estimator method, we worked with the Phase Locking Value (PLV), which assesses phase synchrony between two signals in a specific frequency band [15]. More specifically, it corresponds to the absolute value of the mean phase between s_1 and s_2, defined as follows:

$$PLV = |e^{i\Delta\phi(t)}| \tag{7}$$

where $\Delta\phi(t) = arg(\frac{z_1(t).z_2^*(t)}{|z_1(t)|.|z_2(t)|})$ $\Delta\phi(t)$ represents the associated relative phase computed between signals and $z(t) = s(t) + i.h(s(t))$ the analytic signal obtained by applying the Hilbert transform on the signal $s(t)$.

3.3. Amplitude coupling method

We computed the Amplitude Envelope Correlation (AEC) [16] which relies on the linear correlations of the envelopes of the band-pass filtered signals obtained from Hilbert transform. For the sake of completeness, we report the results of both AEC and ICoh, although those features were not used in the final submission. The generated matrices were not SPD and we had to pre-process them heavily in order to be able to apply the Riemmannian geometry. This may explain the poor results of those features in our setup.

4. PROPOSED APPROACH

The task discussed here consisted of classifying the EEG data from a 2-class paradigm (left versus right-hand motor

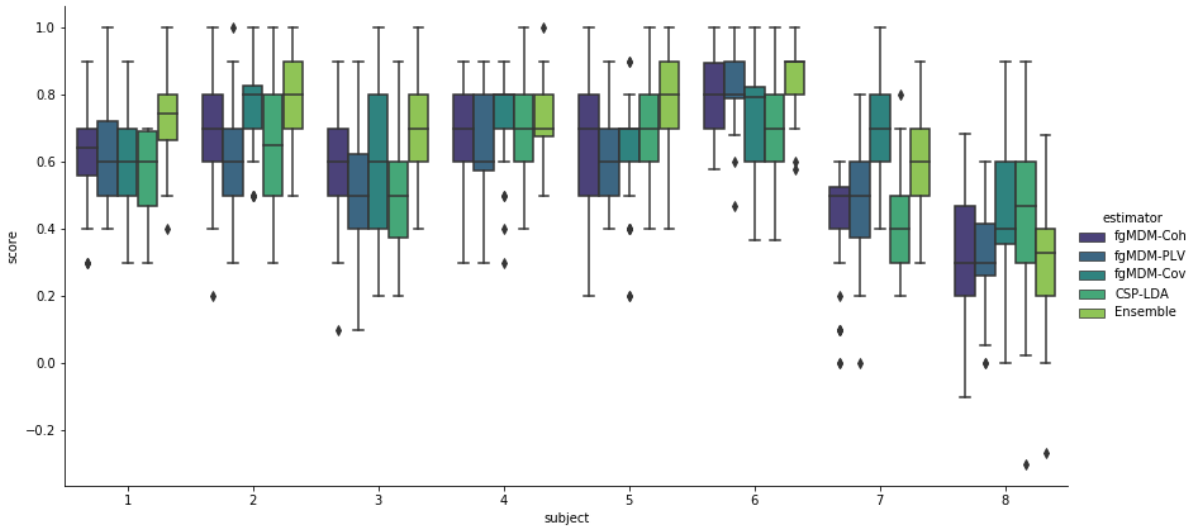

Fig. 3. Per subject Kappa score, comparing the separate pipelines, i.e. FgMDM estimated on covariance, spectral coherence and PLV, a CSP+LDA for the baseline and the submitted ensemble classifier.

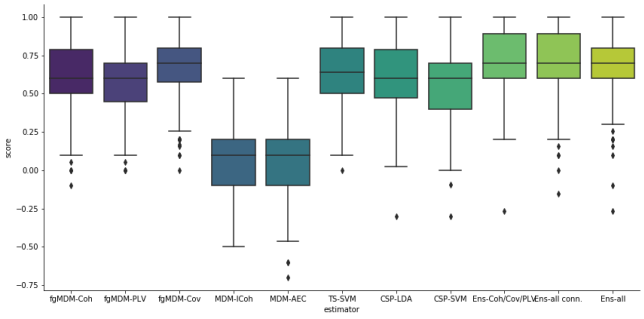

Fig. 4. Average kappa score for all subjects, for the different tested estimators.

imagery) recorded from 8 stroke patients[2]. The novelty of our approach consists of combining Riemannian classifiers trained on SPD matrices coming from both measures of FC and covariance estimation.

To estimate FC features, we used the computation detailed in the previous section implemented in Brainstorm software [14], and the sample covariance estimator of Matlab. The covariance estimators often include regularization using shrinkage approach to avoid ill-conditioned matrices. For FC, no shrinkage estimators had been defined yet. Thus, we used a simple algorithm to project FC matrices on the manifold of PSD matrices [17]. We applied the FgMDM algorithm [18], that computes filters from a Fisher Geodesic Discriminant Analysis before using a Minimum Distance to Mean (MDM) classifier. We used the LogEuclidean distance and its associated mean in the MDM for its robustness and its efficiency. To take into account the shift between training and test set, the test data were transported to the mean of training set[3], as described in [9, 19]. Each FgMDM classifier predicted a probability for the output classes using the softmax function on distance to nearest mean. These probabilities were used to train a stacked classifier [20]. We tried several classifier and we chose the ridge classifier for its robustness. This classifier made the final decision for the prediction, as shown in Fig. 2.

The performance of our submission was estimated on training data and compared to a baseline, that was the Linear Discriminant Analysis LDA) with CSP spatial filters. The results are shown in Fig. 3, indicating the Kappa score estimated with repeated 5-fold cross-validation for each subject. The performance of each level 1 classifiers – the FgMDM-Coh, FgMDM-PLV, FgMDM-Cov – are provided, along with the ensemble classifier. Indeed, we tested several classifiers and a combination for the stacked classifier. The obtained results are summarized in Fig. 4. The FgMDM trained on ICoh and AEC features presented very low kappa score. We also used a popular RG classifier, a SVM trained on the tangent space (TS-SVM). We also tested the CSP-SVM. In both cases, the SVM was parametrized through a grid search on the parameter space. This figure displays the score of the chosen system and stacked classifiers trained on different features: one ensemble classifier trained on all FC features (Coh, ICoh, PLV, AEC, Cov) and one trained on all possible level one classifiers.

[2]For a more detailed presentation of the protocol, the reader can refer to `https://sites.google.com/view/bci-comp-wcci/`

[3]This transductive setup was allowed by the rules of the competition but in a real-life scenario, the mean of the test data could be estimated with unlabelled data during the calibration.

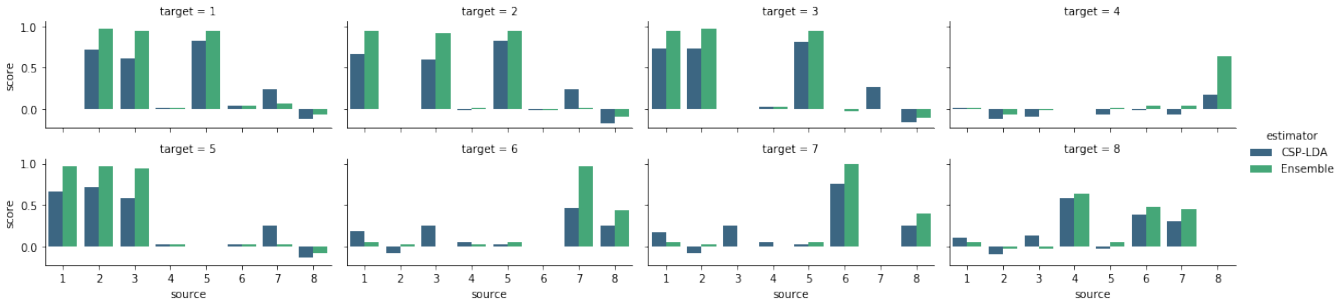

Fig. 5. Kappa score for each target subjects, comparing our system and CSP-LDA trained on each source subject.

5. AFTERMATH

There were 14 submissions to the competition from 12 different institutions around the world across 9 different countries spread across 3 continents. At the end of the competition[4].Our approach got the first position on this task with a substantial margin, the following teams having respectively kappa scores of 0.49 and 0.47 and accuracies of 74.69% and 73.75%. The kappa score obtained on validation is close to 0.68 and is close to the value obtained on training data with a 5-fold cross-validation.

For a full description of our pipeline and the requirements to use it, the reader can refer to the RIGOLETTO (RIemannian GeOmetry LEarning : applicaTion To cOnnectivity) GitHub repository[5]. Using FC estimators associated with an ensemble classifier gives the possibility to take into account the users' specificity. After participating to the competition, we elicited different approaches to improve our method depending on the adopted perspective: theory behind RG, features extraction and transfer learning. In the first case, further investigation should be done regarding the follow-up to the MDM [21] and the dimensionality reduction [22] (for other higher dimensionality datasets). Regarding the feature extraction, we plan to improve in particular the selection of the frequency band of interest [23]. Another promising lead would be to extract for each epoch several PSD matrices, each on a different frequency band, and to consider this set as a trajectory on the manifold, in the spirit of [24]. Other items, such as the agreement and variability among covariance and connectivity and the non-stationarity of connectivity features [25] will be considered.

Participating to the WCCI-Clinical BCI Competition has been the occasion to propose a novel approach and to start bridging the gap between Riemannian geometry and connectivity features. Nevertheless, it motivates the need to study more in depth the connectivity features under the lens of the Riemannian geometry. For instance, our preliminary results showed that covariance and connectivity features seems to produce similar average patterns but this raises as well the question of their individual variability. We plan to study those questions in a long version of this draft.

[4]For more details, the reader can access to the website of the competition : `https://sites.google.com/view/bci-comp-wcci/`

[5]available in https://github.com/sylvchev/wcci-rgcon

6. ENVIRONMENTAL IMPACT

The approach taken in this submission does not require lengthy computation on GPU clusters or HPC, in order to reduce its environmental impact. The team members relied mainly on Slack, git and overleaf to communicate. As there is no direct estimation of the footprint of these services, we use the email scenario of The Shift Project report [26] as a surrogate. We estimate that this submission generated the equivalent of 62 gCO_2. The Shift Project made a contested estimation for the environmental impact of watching a video in HD on a streaming service [27]. The impact of our submission lies thus between streaming the theater-released version and the extended version of the *"Lord of the Ring"* trilogy.

7. ACKNOWLEDGEMENTS

The authors would like to thank the organizers of the WCCI-Clinical BCI competition for giving them the opportunity to kickstart this promising collaboration. FY thanks F. Lotte for suggesting several years ago the use of connectivity features in lieu of covariances in Riemannian classifiers.

8. REFERENCES

[1] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain–computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, 2002.

[2] F. Pichiorri, G. Morone, M. Petti, J. Toppi, I. Pisotta, M. Molinari, S. Paolucci, M. Inghilleri, L. Astolfi, F. Cincotti, and D. Mattia, "Brain-computer interface boosts motor imagery practice during stroke recovery," *Annals of Neurology*, vol. 77, no. 5, 2015.

[3] C. King, P. Wang, L. Chui, A. Do, and Z. Nenadic, "Operation of a brain-computer interface walking simulator for individuals with spinal cord injury," *Journal of NeuroEngineering and Rehabilitation*, vol. 10, no. 1, 2013.

[4] M. Thompson, "Critiquing the Concept of BCI Illiteracy," *Science and Engineering Ethic*, 2018.

[5] B. Blankertz, C. Sannelli, S. Halder, E. Hammer, A. Kübler, K. Müller, G. Curio, and T. Dickhaus, "Neurophysiological predictor of SMR-based BCI performance," *NeuroImage*, vol. 51, no. 4, 2010.

[6] M. Ahn, H. Cho, S. Ahn, and S. Jun, "High theta and low alpha powers may be indicative of BCI-illiteracy in motor imagery," *PLoS ONE*, vol. 8, no. 11, 2013.

[7] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A Review of Classification Algorithms for EEG-based Brain-Computer Interfaces: A 10-year Update," *Journal of Neural Engineering*, 2018.

[8] F. Yger, M. Berar, and F. Lotte, "Riemannian approaches in brain-computer interfaces: a review," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 10, 2016.

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

[10] M. Congedo, A. Barachant, and R. Bhatia, "Riemannian geometry for eeg-based brain-computer interfaces; a primer and a review," *Brain-Computer Interfaces*, vol. 4, no. 3, 2017.

[11] F. de Vico Fallani, J. Richiardi, M. Chavez, and S. Achard, "Graph analysis of functional brain networks: practical issues in translational neuroscience," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 369, no. 1653, 2014.

[12] T. Cattai, S. Colonnese, M. Corsi, D. Bassett, G. Scarano, and F. De Vico Fallani, "Phase/amplitude synchronization of brain signals during motor imagery BCI tasks," *arXiv:1912.02745*, 2019.

[13] M. Corsi, M. Chavez, D. Schwartz, N. George, L. Hugueville, A. Kahn, S. Dupont, D. Bassett, and F. De Vico Fallani, "Functional disconnection of associative cortical areas predicts performance during BCI training," *NeuroImage*, vol. 209, 2020.

[14] F. Tadel, S. Baillet, J. Mosher, D. Pantazis, and R. Leahy, "Brainstorm: A User-Firendly Application for MEG/EEG Analysis," *Computational Intelligence and Neuroscience*, 2011.

[15] J. Lachaux, E. Rodriguez, J. Martinerie, and F. Varela, "Measuring phase synchrony in brain signals," *Human Brain Mapping*, vol. 8, no. 4, 1999.

[16] J. Hipp, D. Hawellek, M. Corbetta, M. Siegel, and A. Engel, "Large-scale cortical correlation structure of spontaneous oscillatory activity," *Nature Neuroscience*, vol. 15, no. 6, 2012.

[17] N. Higham, "Computing a nearest symmetric positive semidefinite matrix," *Linear algebra and its applications*, vol. 103, 1988.

[18] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Riemannian geometry applied to bci classification," in *International Conference on Latent Variable Analysis and Signal Separation*, 2010.

[19] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Multiclass brain–computer interface classification by riemannian geometry," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, 2011.

[20] D. Wolpert, "Stacked generalization," *Neural networks*, vol. 5, no. 2, 1992.

[21] M. Congedo, P. Rodrigues, and C. Jutten, "The riemannian minimum distance to means field classifier," in *8th Graz BCI Conference*, 2019.

[22] I. Horev, F. Yger, and M. Sugiyama, "Geometry-aware principal component analysis for symmetric positive definite matrices," in *ACML*, 2016.

[23] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain Research Reviews*, vol. 29, no. 2, 1999.

[24] Y. Li, K. Wong, and H. de Bruin, "Electroencephalogram signals classification for sleep-state decision–a riemannian geometry approach," *IET signal processing*, vol. 6, no. 4, 2012.

[25] A. Balzi, F. Yger, and M. Sugiyama, "Importance-weighted covariance estimation for robust common spatial pattern," *Pattern Recognition Letters*, pp. 139–145, 2015.

[26] H. Ferreboeuf, F. Berthoud, P. Bihouix, P. Fabre, D. Kaplan, L. Lefèvre, A. Monnin, O. Ridoux, S. Vaija, M. Vautier, X. Verne, A. Ducass, M. Efoui-Hess, and Z. Kahraman, "Lean ICT: Towards digital sobriety," Tech. Rep., The Shift Project, 2019.

[27] J Geist, "Did The Shift Project really overestimate the carbon footprint of online video? Our analysis of the IEA and CarbonBrief articles," 2020.