Novel SPD matrix representations considering cross-frequency coupling for EEG classification using Riemannian geometry

Maria Sayu Yamamoto, Apolline Mellot, Sylvain Chevallier, Fabien Lotte

Novel SPD matrix representations considering cross-frequency coupling for EEG classification using Riemannian geometry

Maria Sayu Yamamoto[†], Apolline Mellot[‡], Sylvain Chevallier[†], and Fabien Lotte[§]
[†]LISN, University Paris-Saclay, Gif-sur-Yvette, France. [‡]Inria, CEA, University Paris-Saclay, Palaiseau, France.
[§] Inria center at the University of Bordeaux, Talence, France. Email: maria-sayu.yamamoto@universite-paris-saclay.fr

***Abstract*—Accurate classification of cognitive states from Electroencephalographic (EEG) signals is crucial in neuroscience applications such as Brain-Computer Interfaces (BCIs). Classification pipelines based on Riemannian geometry are often state-of-the-art in the BCI field. In this type of BCI, covariance matrices based on EEG signals of independent frequency bands are used as classification features. However, there is significant neuroscience evidence of neural interactions across frequency bands, such as cross-frequency coupling (CFC). Therefore, in this paper, we propose novel symmetric positive definite (SPD) matrix representations considering CFC for Riemannian geometry-based EEG classification. The spatial interactions of phase and amplitude within and between frequency bands are described in three different CFC SPD matrices. This allows us to include additional discriminative neurophysiological features that are not available in the conventional Riemannian EEG features. Our method was evaluated using a mental workload classification task from a public passive BCI dataset. Our fused model of the three CFC covariance matrices showed statistically significant improvements in average classification accuracies from the conventional covariance matrix in the theta and alpha bands by 18.32% and in the beta and gamma bands by 4.34% with smaller standard deviations. This result confirmed the effectiveness of considering more diverse neurophysiological interactions within and between frequency bands for Riemannian EEG classification.**

***Index Terms*—Electroencephalography (EEG), Riemannian geometry, Brain-computer interfaces (BCI), cross-frequency coupling (CFC)**

I. Introduction

Electroencephalographic (EEG) signals analysis is at the heart of numerous research disciplines. Thanks to its high temporal resolution, non-invasiveness, and relatively low cost, EEG is extensively used for applications monitoring brain activity, such as medical diagnosis, neuroscience research or Brain-Computer Interfaces (BCIs) [1]. BCIs are technology that can translate human brain activity into commands to control external devices without any physical movement [2]. EEG-based BCIs are notably promising as assistive devices for motor-impaired users [3], detector of intraoperative awareness during general anesthesia [4], real-time monitoring of pilots' mental states [5] and many more. The state-of-the-art of classification pipeline for EEG-based BCIs is the Riemannian BCIs [6]. In this pipeline, features are extracted by cross-channel covariance matrices (Cov) considering the variance and covariance of EEG signals across channels. Those matrices are known to be Symmetric Positive-Definite (SPD). Then, classification is conducted on the Riemannian manifold of SPD matrices, by defining a manifold with a metric considering the space curvature. Riemannian BCIs have shown significant improvements from conventional BCI system relying on spatial filters and won many EEG classification competitions [6], [7].

To further improve BCIs, various Riemannnian specific approaches have been developed including EEG channel selection [8], dimensionality reduction [9], artifact/outlier detection [10] or frequency band (FB) selection [11], among other. All these methods have in common that they operate after the covariance matrix estimations, and all apply on conventional cross-channel Cov only. Here, we take a different angle and aim to improve Riemannian BCIs by proposing new forms of SPD matrix representing novel neurophysiological features that have not been explored in conventional Riemannian BCIs.

More specifically, in this paper, we propose three different SPD matrices considering phase and amplitude within and between FBs. In fact, an oscillatory signal, such as EEG, consists of amplitudes and phases. Amplitudes and phases between different FBs are known to play a crucial role in organizing large-scale networks and integration across functional brain regions [12]. However, conventional Riemannian BCIs do not consider phases nor interactions between different FBs. To the best of our knowledge, this paper is the first work considering CFC, particularly phase-amplitude coupling (PAC), for EEG classification in Riemannian BCIs. Our CFC Cov were evaluated individually or in fusion and their effectiveness were confirmed on a public mental workload dataset.

This paper is organized as follows: Section II explains the principles of Riemannian geometry while Section III introduces related work. Section IV describes our new approach, and Section V the experimental evaluations. Then, Sections VI and VII describe the results and discussion, respectively, while Section VIII concludes the paper.

II. Riemannian geometry for EEG classification

In this section, we briefly introduce Riemannian geometry from the practical point of view of EEG signal classification.

MS.Y was supported by the French National Research Agency (ANR) with program UDOPIA (grant ANR-20-THIA-0013-01).

A. EEG covariance matrix (EEG Cov)

Let $X \in \mathbb{R}^{M \times T}$ be a filtered EEG trial, with M channels and T time samples. Its Cov C_X is defined as:

$$C_X = \frac{1}{T-1} X X^\top \in \mathbb{R}^{M \times M}. \tag{1}$$

The diagonal entries of this matrix represent the variance of each channel signal, *i.e.*, the band power of each channel, while the off-diagonal terms contain the covariance of each pair of channels. When the number of time samples is large enough with respect to the number of EEG channels, the estimated Cov is empirically full rank, and thus SPD. The set of SPD matrices forms a differentiable manifold called the Riemannian manifold. On such curved spaces, distances between matrices are computed along geodesics. The affine-invariant Riemannian metric (AIRM) δ_r is a geodesic distance that respects the original curvature of the manifold and is given as follows:

$$\delta_r(C_1, C_2) = \|\log({C_1}^{-\frac{1}{2}} C_2 {C_1}^{-\frac{1}{2}})\|_F = \left(\sum_{i=1}^{m} \log^2 \lambda_i \right)^{1/2} \tag{2}$$

where λ_i are the positive eigenvalues of ${C_1}^{-\frac{1}{2}} C_2 {C_1}^{-\frac{1}{2}}$ and $\|.\|_F$ is the Frobenius norm of the matrix.

B. Tangent space mapping (TSM) based classification

On a differentiable manifold $\mathcal{M}$, tangent vectors can be defined at $p \in \mathcal{M}$, and the set of tangent vectors spans a Euclidean space called the tangent space $T_p\mathcal{M}$. When the set of SPD matrices is concentrated locally, the tangent space is the local approximation of the space of the set of SPD matrices. Let $\{C_i\}$ $(i = 1, 2, \cdots, N)$ is a set of SPD matrices, the projection of C_i on the tangent space at $\overline{C}$ is defined as:

$$\boldsymbol{v}_i = \mathrm{Upper}\left(\log\left(\overline{C}^{-1/2} C_i \overline{C}^{-1/2} \right) \right) \in \mathbb{R}^{M(M+1)/2} \tag{3}$$

where $\mathrm{Upper}(\cdot)$ is a half-vectorization operator applying unit weights to diagonal entries and $\sqrt{2}$ to off-diagonal entries of a given SPD matrix and $\log(\cdot)$ is the principal matrix logarithm [13]. In this work, $\overline{C}$ is estimated by the Riemannian mean of the whole set of SPD matrices. As a tangent space is a Euclidean space, TSM allows us to classify vectorized SPD matrices by applying any vector-based classifier, such as Linear Discriminant Analysis (LDA). TSM-based classification showed the best performance in Riemannian BCIs [6].

III. Related work for proposed SPD matrices

This section briefly reviews the existing work that inspired our proposal for new spatial CFC covariance matrices.

A. Phase-amplitude coupling (PAC)

PAC is a type of CFC, which is a coupling phenomenon between bursts of high-frequency oscillations and the phase of lower-frequency rhythms. In other words, the fast oscillation is modulated by the cycle of the slower oscillation [14]. PAC is speculated to have high relevance for cognitive processing, such as memory processes and decisions making [14], [15]. Various indicators have been developed to quantify PAC, but no gold-standard exists yet [15]. Each metric has different pros and cons, thus, should be chosen according to the study goal. As our aim was developing a new SPD representation based on PAC, we focused on the direct PAC (dPAC), formulated as:

$$\mathrm{dPAC} := \frac{1}{\sqrt{T}} \frac{\left| \sum_{t=1}^{T} a_{f_\mathrm{H}} e^{i\phi_{f_\mathrm{L}}(t)} \right|}{\sqrt{\sum_{t=1}^{T} a_{f_\mathrm{H}}(t)^2}} \tag{4}$$

where $t = 1, 2, \cdots, T$ is time samples, ϕ_{f_L} is the phase of lower FB f_L and a_{f_H} is the amplitude of higher FB f_H. The dPAC is reported as a more accurate estimator than others [16] and can estimate PAC strength for concise data segments [17].

B. Cross-frequency covariance matrix (crossFB Cov)

CrossFB Cov, C_{crossFB}, was proposed for brain age prediction from EEG in [18]. To describe interactions of five FBs $f_1, f_2, \cdots, f_5$ in a Cov, filtered EEG are first concatenated as:

$$X_{\text{crossFB}} = \begin{pmatrix} X_{f_1}(t) \\ X_{f_2}(t) \\ \vdots \\ X_{f_5}(t) \end{pmatrix} \in \mathbb{R}^{5M \times T} \tag{5}$$

Then, $C_{\text{crossFB}} \in 5M \times 5M$ was estimated using Eq.1. In this Cov, the block matrices within the FBs are aligned diagonally (*i.e.*, conventional Riemannian Cov), and the cross-FB block matrices between band pairs are off-diagonal blocks. C_{crossFB} showed better regression performance than a model estimating covariance matrices for each FBs separately. However, C_{crossFB} has not been explored for BCI classification, thus, we include it as one of the CFC Cov in our study with two FBs, *i.e.* $C_{\text{crossFB}} \in 2M \times 2M$.

IV. Proposed methods

Our proposed CFC Cov capture PAC in two different ways, described below.

A. Phase-amplitude covariance matrix (PA Cov)

PA Cov, C_{PA}, aims to quantify the relationship between the phase of the low-FB and the amplitude of the high-FB. Once these two components extracted using Hilbert transforms, a data matrix is defined as:

$$X_{\text{PA}} = \begin{pmatrix} \cos(\phi_{f_\mathrm{L}(t)}) \\ \sin(\phi_{f_\mathrm{L}(t)}) \\ a_{f_\mathrm{H}(t)} \end{pmatrix} \in \mathbb{R}^{3M \times T}. \tag{6}$$

As phase is a circular variable, the phase value is divided into cosine and sine. Then the covariance matrix of X_{PA}, $C_{\text{PA}} \in \mathbb{R}^{3M \times 3M}$, is estimated with Eq. 1.

B. Direct PAC covariance matrix (dPAC Cov)

dPAC Cov, C_{dPAC}, aims to quantify the channel-to-channel interaction of dPAC values measured based on phase and

amplitude within a channel. To this end, we modified dPAC given by Eq. 4 to be a time-resoluted index as follows:

$$X_{\text{dPAC}} = \begin{pmatrix} \frac{\left|a_{f_{\text{H}}} e^{i\phi_{f_{\text{L}}}(t)}\right|_{\text{ch}1}}{\sqrt{\sum_{t=1}^{T} a_{f_{\text{H}}}(t)^2}_{\text{ch}1}} \\ \frac{\left|a_{f_{\text{H}}} e^{i\phi_{f_{\text{L}}}(t)}\right|_{\text{ch}2}}{\sqrt{\sum_{t=1}^{T} a_{f_{\text{H}}}(t)^2}_{\text{ch}2}} \\ \vdots \\ \frac{\left|a_{f_{\text{H}}} e^{i\phi_{f_{\text{L}}}(t)}\right|_{\text{ch}M}}{\sqrt{\sum_{t=1}^{T} a_{f_{\text{H}}}(t)^2}_{\text{ch}M}} \end{pmatrix} \in \mathbb{R}^{M \times T} \tag{7}$$

where $e^{i\phi_{f_{\text{L}}}(t)}$ is euler formula of phase $\phi_{f_{\text{L}}}$. The numerator is the absolute value of the complex composite signal of each channel and the denominator is a normalizing factor based on the amplitude of each channel. C_{dPAC} is estimated with Eq. 1 and it quantifies the variance and covariance of dPAC values.

V. Experimental evaluation

A. Data description

We evaluated our method on the COG-BCI dataset [19]. This dataset comprises EEG data from 29 subjects who performed N-back tasks, a widely used mental workload induction protocol. Participants were instructed to memorize the order in which the numbers appeared and to respond by pressing the space key on their keyboard if the number presented was the same as the N^{th} previously presented number. In a trial, the number was first displayed for 500 ms, then a blank screen for 1500 ms. Therefore, the each trial duration was 2 seconds. Participants performed totally three different N-back conditions ($N = 0, 1, 2$). and for three sessions (*i.e.*, three days). One session was divided into three recording blocks and each block took about two minutes consisting of 48 trials for each N-back condition. In this study, we used 0 and 2-back conditions EEG data from the first session. The first two trials were removed from each block of each task because the cognitive state in those trials in the 2-back condition was expected to be different from other trials. Thus 46 trials per block were used for each task in our analysis. We used the first two blocks as the training set and the third block as the testing set without any shuffle. This guarantees the cognitive states independence between the training and testing sets.

EEG were recorded using 62 channels. Among them, we used 16 channels [F3, Fz, F4, FC5, FC1, FC2, FC6, C3, C4, CP5, CP1, CP2, CP6, P3, Pz, P4] for the first nine subjects and 17 channels (the same 16 plus Cz) for the remaining subjects. Cz was missing for the first nine subjects in the dataset.

B. Pre-processing

The preprocessing was performed on training and test data separately. EEG data was first referenced using right mastoid electrode. Secondly, a high-pass filter was applied at 1 Hz using a FIR filter using $pop\_eegfiltnew.m$ from EEGLAB (v2022.1) [20] with default parameters. Next, electrodes detected as outlier based on their standard deviation (with a mean + two standard deviation threshold on the training set) were rejected and then interpolated using spherical interpolation using $pop\_interp.m$ from EEGLAB. Then, independent component analysis (ICA) was applied using second-order blind identification (SOBI) [21], and components with a likelihood greater than 95% to be muscular, ocular or cardiac artifacts, according to EEGLAB plugin *ICLabel* [22], were rejected. After that, data was low-pass filtered at 90 Hz and re-referenced using common average reference.

Then, bandpass filtering was applied in five FBs, $\delta : 1-4$ Hz, $\theta : 4-8$ Hz, $\alpha : 8-13$ Hz, $\beta : 13-30$, $\gamma : 30-90$ Hz, using non-causal forward-backward FIR filter from MNE-python (v1.1.0) [23] with the default filter order. Then, the data was epoched into each trial, *i.e.*, two second-long epochs without any overlap. All Cov in this study were estimated with the oracle approximating shrinkage [24].

C. Frequency band pair selection

For proper PAC quantification, two main constraints should be satisfied. First, the FB for phase should always be lower than the one for amplitude because PAC is a phenomena for which the phase of a low-frequency oscillation modulates the power of a high-frequency oscillation. Secondly, the bandwidth for the high-FB should be at least twice as large as the bandwidth of the low-frequency component, in order to capture the amplitude modulation effect [25]. Thus, to respect those two conditions, for C_{PA} and C_{dPAC}, the best FB pair was selected among $[f_{\text{L}}, f_{\text{H}}] = \{[\delta, \beta], [\theta, \beta], [\alpha, \beta], [\delta, \gamma], [\theta, \gamma], [\alpha, \gamma], [\beta, \gamma]\}$. For C_{crossFB}, the best FB pair (*i.e.*, two bands) was selected from all possible pairs from δ to γ. The best pair was selected for each CFC Cov type, for each subject. To this end, we selected the FB pair whose CFC Cov became the most discriminative on the manifold, using the *ClassDis* metric [26] on the training set.

D. Compared models

We compared the following three types of pipelines with totally six different classification models:

a) Baselines: We used the conventional Riemannian Cov given by Eq. 1 with two different FB separately. The Cov of each FB was estimated individually, then after TSM, the two vectorized Cov were concatenated into a single vector. Thus, the feature vector size is $M(M+1)$. It does not contain any cross-frequency component. The FB pairs used were:

- $C_{[\theta,\alpha]}$: $[\theta, \alpha]$ bands pair was chosen as they are known to be collaboratively associated with mental workload changes [5]. We set this model as the most basic baseline.
- $C_{[\beta,\gamma]}$: In the FB pair selection for C_{crossFB} (Section V-C), $[\beta, \gamma]$ bands pair was always selected as the best FB pair for all subjects. Thus, to investigate how much benefit the inclusion of cross-frequency components would bring, we also included this model as baseline.

b) Individual CFC pipelines: In this pipeline, we evaluated three different CFC Cov independently: C_{PA}, C_{dPAC} (both proposed in this paper) and C_{crossFB}. The FB pair used in each model was selected by the selection method described in Section V-C. For C_{crossFB}, we concatenated the filtered EEG

data as Eq. 5 using two different FBs f_1 and f_2. Therefore, $C_{\text{crossFB}} \in \mathbb{R}^{2M\times 2M}$ contains conventional Cov as diagonal block matrices, and in addition to this, the Cov between f_1 and f_2 are included as off-diagonal block matrices.

c) Fusion CFC pipeline: To combine the benefits of all CFC Cov, the vectorized features of each Cov (*i.e.*, $C_{\text{PA}} + C_{\text{dPAC}} + C_{\text{crossFB}}$) were fused, as described in Fig. 1. First, all CFC Cov were vectorized after TSM with Eq. 3. Then, those vectorized features were concatenated to form one large vector, as done in [18], and its size became $(7M+3)M$. This concatenated vector contains different feature scales, including multiple FBs of phase, amplitude and PAC values. Thus, we re-scaled the vector by subtracting its mean and scaling it to unit variance. This re-scaling was applied for all feature vectors of all methods, to make a fair comparison.

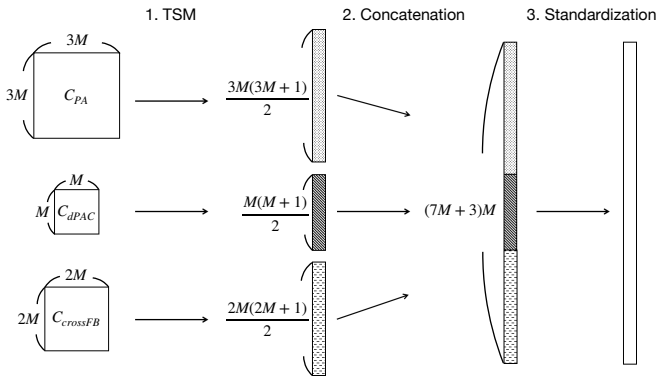

Fig. 1. Illustration of fusion pipeline of three different SPD representations

For all pipelines, a Ridge Regression was applied as a classifier using *RidgeClassifier* from Scikit-learn (v1.2.1) [27] with default parameters. To investigate statistical differences, we performed a Friedman test for Model ($C_{[\theta,\alpha]}$, $C_{[\beta,\gamma]}$, C_{PA}, C_{dPAC}, C_{crossFB}, and $C_{\text{PA}} + C_{\text{dPAC}} + C_{\text{crossFB}}$) after checking the normality assumption. In case where a statistical significance was observed, a Post-hoc analysis was performed using Durbin-Conover pairwise tests.

VI. Results

Average classification accuracies for each model are summarized in Table I, while Fig. 2 shows the distribution of classification accuracy of each model. Our fusion model showed the highest average performance and smallest standard deviation among all models. Friedman test revealed a statistical significant difference in Model [$X^2(5) = 75.4; p < 0.001$]. Post-hoc analysis with Durbin-Conover pairwise tests showed that all proposed models except C_{dPAC} were significantly better than $C_{[\theta,\alpha]}$ ($p < 0.001$). Compared to $C_{[\beta,\gamma]}$, fusion model showed significant improvement ($p = 0.034$), while C_{dPAC} was significantly worse ($p < 0.001$). Among our proposed methods, C_{PA} was significantly better than C_{dPAC} ($p < 0.001$). Also, the fusion model was significantly better than C_{PA}, C_{dPAC} and C_{crossFB} ($p < 0.001$).

The selected best FB pair for C_{PA} and C_{crossFB} was $[\beta, \gamma]$ for all subjects. On the other hand, C_{dPAC} had more variety in the selection result. Among the 29 subjects, $[\delta, \gamma]$ was selected for 10 subjects, $[\theta, \gamma]$ for 10, $[\alpha, \gamma]$ for five and $[\beta, \gamma]$ for four.

VII. Discussion

The aim of this experiment was to investigate the validity of our proposed CFC Cov, either individually or in fusion compared to conventional Cov. We set two different baseline models. Among them, $C_{[\beta,\gamma]}$ showed significant improvement from $C_{[\theta,\alpha]}$ simply by changing the FBs to use. Those high-FBs, β or γ are typically filtered out in most BCI designs. For instance, in the recent passive BCI classification competition, the dataset was initially low-pass filtered at 40 Hz [7]. Our results suggest that such high-FBs may be worth considering if artifacts are removed carefully, as we did here with ICA.

The fusion model showed the best average accuracy among all models. This suggests that combining various CFC covariance matrix features results in enhanced class discriminability. However, the resulting accuracy was not statistically significantly better than that of the best individual models. This may be due to the overly large dimension of the concatenated feature vector. To overcome this, future work should explore sparse classification of the concatenated features.

Regarding the individual pipelines, our proposed methods showed globally higher average accuracy than two baseline methods, except C_{dPAC}. The main difference between C_{dPAC} and the other CFC Cov was that C_{PA} and C_{crossFB} included amplitude block matrices independently in one covariance matrix, while C_{dPAC} did not. This result may imply that the independent inclusion of amplitude features is essential to maximize classification accuracy. However, we stress that it is premature to conclude that including PAC is ineffective. The current C_{dPAC} only considers PAC within channels, *i.e.*, the amplitude of the high-frequency component and the phase of the low-frequency component are calculated from the same channel, even though PAC plays a crucial role in both local and long-distance neuronal communication across brain regions. For instance, in [28], a correlation was observed between the phase from the visual cortex Electrocorticography (ECoG) electrodes and the amplitude in the medial frontal lobes during visuomotor tasks. As such, it is worth exploring ways to embed cross-channel PAC in SPD matrices in the future.

The only difference between C_{crossFB} and $C_{[\beta,\gamma]}$ was the inclusion of cross-frequency terms since the $[\beta, \gamma]$ pair was selected for all subjects in C_{crossFB}. The average accuracy of C_{crossFB} was slightly higher than $C_{[\beta,\gamma]}$ but not significantly. It may be worthwhile to estimate C_{crossFB} with more than two bands, as done in [18].

VIII. Conclusion

In this paper, we proposed novel SPD matrix representations considering CFC for EEG classification using Riemannian geometry. In most EEG classification approaches, including Riemannian BCIs, frequency band powers have been used as the major features. Here, we incorporated more diverse

TABLE I
AVERAGE CLASSIFICATION ACCURACY ACROSS THE 29 SUBJECTS, FOR EACH MODEL

	baseline		individual			fusion
	$C_{[\theta,\alpha]}$	$C_{[\beta,\gamma]}$	PA	dPAC	crossFB	PA+dPAC+crossFB
Ave. std [%]	73.58 ± 15.5	87.56 ± 18.2	88.27 ± 11.0	74.06 ± 16.1	88.53 ± 10.9	$\mathbf{91.90 \pm 9.75}$

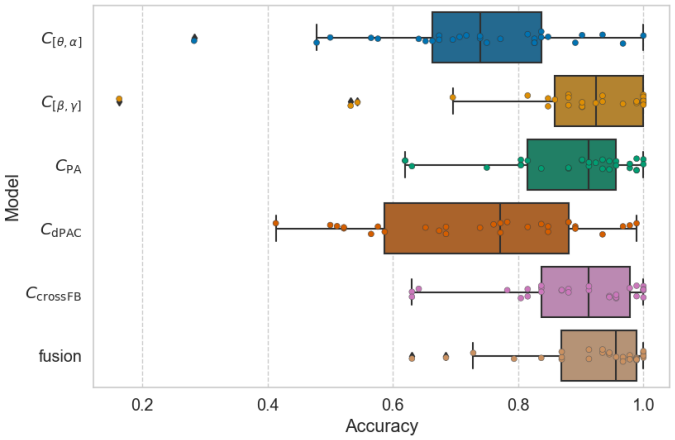

Fig. 2. Distribution of classification accuracies for each model.

neural interactions, such as inter- and intra-frequency band power and phase relationships, into SPD matrices. Through comparative experiments, their effectiveness was demonstrated and especially the fusion model of our proposed CFC matrices showed the highest average accuracy. Additionally, CFC matrices including block matrices of amplitude (*i.e.*, C_{PA} and C_{crossFB}) showed globally higher average accuracy than baseline. Those results thus suggested that to improve Riemannian BCIs, we need to combine amplitude block matrices with other neurophysiological features within the SPD matrix.

Future works will deepen neurophysiological interpretation of the empirical effectiveness of our new SPD representations. Furthermore, we will consider how to integrate CFC Cov features with conventional features more effectively, notably by using sparse modeling. We will also explore ways to quantify cross-channel PAC into SPD matrices efficiently. Finally, we will investigate our approach with non-linear classifiers or in more complex BCI setups, such as cross-session or multi-class.

REFERENCES

[1] E. Niedermeyer and F. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.

[2] M. Clerc, L. Bougrain, and F. Lotte, *Brain-Computer Interfaces 1*. 2016.

[3] G. Pfurtscheller, G. R. Müller-Putz, R. Scherer, and C. Neuper, "Rehabilitation with Brain-Computer Interface systems," *Computer*, vol. 41, no. 10, pp. 58–65, 2008.

[4] S. Rimbert, P. Riff, N. Gayraud, D. Schmartz, and L. Bougrain, "Median nerve stimulation based BCI: a new approach to detect intraoperative awareness during general anesthesia," *Front. Neurosci.*, vol. 13, 2019.

[5] F. Dehais, A. Duprès, S. Blum, N. Drougard, S. Scannella, R. N. Roy, and F. Lotte, "Monitoring pilot's mental workload using ERPs and spectral power with a six-dry-electrode EEG system in real flight conditions," *Sensors*, vol. 19, no. 6, p. 1324, 2019.

[6] F. Yger, M. Berar, and F. Lotte, "Riemannian approaches in Brain-Computer Interfaces: a review," *IEEE Trans. Neural Sys. Rehab.*, vol. 25-10, pp. 1753–1762, 2016.

[7] R. N. Roy, M. F. Hinss, L. Darmet, S. Ladouce, E. S. Jahanpour, B. Somon, X. Xu, N. Drougard, F. Dehais, and F. Lotte, "Retrospective on the first passive brain-computer interface competition on cross-session workload estimation," *Front. Neuroerg.*, 2022.

[8] K. Sadatnejad and F. Lotte, "Riemannian channel selection for BCI with between-session non-stationarity reduction capabilities," *IEEE Trans. Neural Syst. Rehab.*, vol. 30, pp. 1158–1171, 2022.

[9] M. S. Yamamoto, F. Yger, and S. Chevallier, "Subspace oddity-optimization on product of stiefel manifolds for EEG data," in *Proc. ICASSP*, pp. 1080–1084, 2021.

[10] M. S. Yamamoto, K. Sadatnejad, T. Tanaka, R. Islam, Y. Tanaka, and F. Lotte, "Detecting EEG outliers for BCI on the Riemannian manifold using spectral clustering," in *EMBC2020*, pp. 438–441, IEEE, 2020.

[11] M. S. Yamamoto, F. Lotte, F. Yger, and S. Chevallier, "Class-distinctiveness-based frequency band selection on the Riemannian manifold for oscillatory activity-based BCIs: preliminary results," in *Proc. IEEE/EMBS EMBC*, pp. 3690–3693, 2022.

[12] V. Jirsa and V. Müller, "Cross-frequency coupling in real and virtual brain networks," *Front. Comput. Neurosci.*, vol. 7, p. 78, 2013.

[13] A. Barachant, A. Andreev, and M. Congedo, "The Riemannian Potato: an automatic and adaptive artifact detection method for online experiments using Riemannian geometry," in *TOBI Workshop lV*, pp. 19–20, 2013.

[14] M. X. Cohen, C. E. Elger, and J. Fell, "Oscillatory activity and phase–amplitude coupling in the human medial frontal cortex during decision making," *J. Cog. Neur.*, vol. 21, no. 2, pp. 390–402, 2008.

[15] R. T. Canolty and R. T. Knight, "The functional role of cross-frequency coupling," *Trends Cog. Sci.*, vol. 14, no. 11, pp. 506–515, 2010.

[16] T. E. Özkurt and A. Schnitzler, "A critical note on the definition of phase–amplitude cross-frequency coupling," *Journal of Neuroscience methods*, vol. 201, no. 2, pp. 438–443, 2011.

[17] S. Samiee and S. Baillet, "Time-resolved phase-amplitude coupling in neural oscillations," *NeuroImage*, vol. 159, pp. 270–279, 2017.

[18] D. Sabbagh, "Building clinical biomarkers from cerebral electrophysiology brain age as a measure of neurocognitive disorders.," 2021.

[19] M. F. Hinss, E. S. Jahanpour, B. Somon, L. Pluchon, F. Dehais, and R. N. Roy, "Open multi-session and multi-task EEG cognitive dataset for passive brain-computer interface applications," *Scientific Data*.

[20] A. Delorme and S. Makeig, "Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis," *Journal of neuroscience methods*, vol. 134, no. 1, pp. 9–21, 2004.

[21] A. Belouchrani, K. Abed-Meraim, J. Cardoso, and E. Moulines, "Second-order blind separation of temporally correlated sources," in *Proc. Int. Conf. Digital Signal Processing*, pp. 346–351, Citeseer, 1993.

[22] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, "Iclabel: An automated electroencephalographic independent component classifier, dataset, and website," *NeuroImage*, vol. 198, pp. 181–197, 2019.

[23] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, *et al.*, "MEG and EEG data analysis with MNE-Python," *Front. Neurosci.*

[24] Y. Chen, A. Wiesel, Y. C. Eldar, and A. O. Hero, "Shrinkage algorithms for MMSE covariance estimation," *IEEE Trans. Sig. Proces.*

[25] D. Dvorak and A. A. Fenton, "Toward a proper estimation of phase–amplitude coupling in neural oscillations," *J. Neurosci. Methods*, vol. 225, pp. 42–56, 2014.

[26] F. Lotte and C. Jeunet, "Defining and quantifying users' mental imagery-based BCI skills: a first step," *J. Neural Eng.*, 2018.

[27] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, and et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[28] B. Voytek, M. D'Esposito, N. Crone, and R. T. Knight, "A method for event-related phase/amplitude coupling," *Neuroimage*, vol. 64.