ENHANCING COMPUTATIONAL EFFICIENCY OF MOTOR IMAGERY BCI CLASSIFICATION WITH BLOCK-TOEPLITZ AUGMENTED COVARIANCE MATRICES AND SIEGEL METRIC

Igor Carrara[1], Theodore Papadopoulo[1]

[1]Université Côte d'Azur, INRIA, Cronos Team, France

E-mail: igor.carrara@inria.fr, theodore.papadopoulo@inria.fr

ABSTRACT: Electroencephalographic signals are represented as multidimensional datasets. We introduce an enhancement to the augmented covariance method (ACM), exploiting more thoroughly its mathematical properties, in order to improve motor imagery classification. Standard ACM emerges as a combination of phase space reconstruction of dynamical systems and of Riemannian geometry. Indeed, it is based on the construction of a Symmetric Positive Definite matrix to improve classification. But this matrix also has a Block-Toeplitz structure that was previously ignored. This work treats such matrices in the real manifold to which they belong: the set of Block-Toeplitz SPD matrices. After some manipulation, this set is can be seen as the product of an SPD manifold and a Siegel Disk Space. The proposed methodology was tested using the MOABB framework with a within-session evaluation procedure. It achieves a similar classification performance to ACM, which is typically better than – or at worse comparable to – state-of-the-art methods. But, it also improves consequently the computational efficiency over ACM, making it even more suitable for real time experiments.

INTRODUCTION

In electroencephalography (EEG) based Brain Computer Interfaces (BCI), state-of-the-art algorithms are often built on Riemannian distance based classification algorithms [1]. The basic idea underlying these methods is to treat the spatial covariance matrix (SCM), extracted from the EEG signal, as an element of the Riemannian manifold of Symmetric Positive Definite (SPD) matrices [2].

A recent extension of this work was obtained by using the Augmented Covariance Method (ACM) [3]. ACM relies on the concept of phase space reconstruction of dynamical systems to create an "ACM matrix" (also called ACM) that contains not only an average spatial representation of the signal but also a representation of its evolution in time. Consequently, the amount of information contained in this ACM matrix is increased w.r.t. the standard spatial covariance. As the ACM matrix also turns out to be an SPD matrix, it can be classified using the same Riemannian framework that was so successful for SCMs. However, it also possesses a structural property

of being Block-Toeplitz, that is, a block matrix with constant blocks over all diagonals. Recently, an approach has been proposed to better deal with such Block-Toeplitz SPD matrices [4], with applications in diverse fields such as audio processing or radar signal analysis [5].

The idea of this research is thus to endow the smooth manifold of Block-Toeplitz SPD matrices with a Riemannian metric, thus allowing the ACM matrix to be treated within its true manifold membership. It is actually possible to treat the Block-Toeplitz SPD matrix manifold as the product of an SPD manifold and a Siegel Disk Space, after applying an appropriate conversion of the blocks of the ACM matrix into the Verblusky coefficients [6].

This approach provides a new – more specific – metric to use for BCI classification algorithms. The strength of the approach lies in its ability to deconstruct the manifold into its constituent elements: the Symmetric Positive Definite (SPD) manifold and the Siegel Disk Space. By discerningly analyzing each component within its respective geometrical domain, this method significantly alleviates the computational demand traditionally associated with the ACM methodology. The resulting algorithm achieves performance that is, at worst, comparable with state-of-the-art BCI algorithms, and often provides quite better results (on par with those of ACM). However, it distinguishes itself by achieving this at substantial reduction in computational costs and carbon footprint compared with standard ACM.

The new – Siegel metric based – pipeline was tested and validated against several state-of-the-art algorithms (Machine Learning (ML) and Deep Learning (DL)) on several datasets for motor imagery (MI) classification using several subjects and on a right versus left hand task with the MOABB framework [7], and a within-session evaluation procedure.

MATERIALS AND METHODS

The EEG signal is represented as a multivariate time series $\mathbf{X} \in \mathbb{R}^{d \times T}$, where T represents the total number of sampled data points, and d indicates the number of electrodes used in the EEG recording. Since this paper focuses on MI task, we split the EEG signals into smaller sections known as epochs, each representing a snapshot

of brain activity during various tasks or mental states. The core aim of our research is to develop a method that can accurately identify the specific task or mental state associated with these EEG epochs.

The space of SPD matrices is composed by square real symmetric matrices that are positive definite, and this space form a smooth manifold that can be equipped with a Riemannian metric [1]. This space is defined as

$$\mathbb{SPD}_d = \{\mathbf{M} \in \mathbb{R}^{d \times d} \mid x^T \mathbf{M} x > 0 \ \forall x \in \mathbb{R}^d \backslash \{0\}\} \quad (1)$$

ACM [3] extends this methodology by combining it with the phase space reconstruction (PSR) approach that is grounded in the Takens theorem [8]. The ACM matrix thus obtained contains spatial and temporal information of the signal and remains an SPD matrix that can be classified using the same Riemannian metric that was so successful for SCMs. This enrichment with temporal features of the information extracted from the signal allows for an improvement of classification performance.

The idea of using Takens theorem is based on the idea that time series obtained from experimental observations that capture only a fraction of the complex dynamics of the underlying system, can nonetheless be utilized to reconstruct the system's full dynamical behavior. This is achieved using a uniform embedding procedure: consider a time series $s(n)$ created thought a measurement process, the PSR technique generates a point $\mathbf{s}_E(n)$ of a D-dimensional space constructed as

$$\mathbf{s}_E(n) = [s(n), s(n-\tau), ..., s(n-(D-1)\tau)]^T \quad (2)$$

where τ is a positive integer called the embedding delay and D is the embedding dimension. $\mathbf{s}_E(n) \in \mathbb{R}^D$ is an uniform embedding of the original phase space.

The ACM matrix (see Fig. 1) is obtained by expanding the original EEG signal using the PSR approach with an embedding dimension p to get a new $dp \times T$ time series, parameterized by the fixed delay τ. The Augmented Covariance Matrix Γ_{aug} is defined as the autocovariance matrix of this new time series:

$$\Gamma_{Aug} = \begin{bmatrix} \Gamma_0 & \Gamma_{-1} & \Gamma_{-2} & \cdots \\ \Gamma_1 & \Gamma_0 & \Gamma_{-1} & \cdots \\ \Gamma_2 & \Gamma_1 & \Gamma_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \Gamma_{p-1} & \Gamma_{p-2} & \Gamma_{p-3} & \cdots \end{bmatrix}, \quad (3)$$

where Γ_0 is the standard spatial covariance matrix and Γ_i is the lagged covariance matrix of the original signal with a delay of $i\tau$ and $\Gamma_{-i} = \Gamma_i^T$. As an autocovariance matrix, Γ_{aug} is symmetric and positive by construction. If not definite, it can be regularized [9], so that we consider it as SPD in the remainder of this article. But, the ACM matrix also has a specific Block-Toeplitz structure, with blocks of dimension $d \times d$ [5]. More formally, Γ_{Aug} belongs to the space $\mathbb{B}_{d \times p}$ of Block-Toeplitz and SPD matrices i.e., SPD matrices of size $dp \times dp$ with constant blocks of size $d \times d$ along all diagonals. This opens up new possibilities for enhancing the ACM formulation by

mapping the ACM matrix to the most suitable geometric space that fully captures both its Block-Toeplitz and SPD natures.

The blocks of the matrix Γ_{Aug}, have been demonstrated to belong to a specific mathematical space [4]

$$\Gamma_i \in \mathbb{D}_d \quad \mathbb{D}_d = \{\mathbf{M} \in \mathbb{C}^{d \times d} \mid \mathbf{I} - \mathbf{M}\bar{\mathbf{M}} > 0\} \quad (4)$$

with $\bar{\mathbf{M}} = \mathbf{J}\mathbf{M}^H\mathbf{J}$ where $\mathbf{J}$ denotes the anti-diagonal matrix and ^{H} is the conjugate transpose operator[1]. This space has no known Riemannian structure but, by implementing a minor adjustment to the coefficients, it is possible to ensure their belonging within the domain of the Siegel disk [6, 10], defined as

$$\Omega_i \in \mathbb{SD}_d \quad \mathbb{SD}_d = \{\mathbf{M} \in \mathbb{C}^{d \times d} \mid \mathbf{I} - \mathbf{M}\mathbf{M}^H > 0\}. \quad (5)$$

The coefficients that have undergone such modification are also known as Verblunsky coefficients [6].

The transformation allow the following conversion,

$$\begin{aligned} \mathbb{B}_{d \times p} &\to \mathbb{SPD}_d \times \mathbb{SD}_d^{p-1} \\ \Gamma_{Aug} &\mapsto (\Gamma_0, \Omega_1, ..., \Omega_{p-1}). \end{aligned} \quad (6)$$

Consider the initial matrix Γ_{Aug} decomposed in its constituent blocks $(\Gamma_0, ..., \Gamma_{p-1})$. The initialization of the recursive transformation is set to $\mathbf{P}_0 = \Gamma_0$. The subsequent coefficients are computed with

$$\Omega_{l+1} = \mathbf{L}_l^{-1/2}(\mathbf{R}_{l+1} - \mathbf{M}_l)\mathbf{K}_l^{-1/2}, \quad (7)$$

with $l = 0, ..., p-1$ and

$$\begin{aligned} \mathbf{L}_l &= \mathbf{P}_0 - (\Gamma_1, ..., \Gamma_l)\tilde{\Gamma}_{l-1}^{-1}(\Gamma_1, ..., \Gamma_l)^H \\ \mathbf{K}_l &= \mathbf{P}_0 - (\Gamma_1^H, ..., \Gamma_l^H)\tilde{\Gamma}_{l-1}^{-1}(\Gamma_1^H, ..., \Gamma_l^H)^H \\ \mathbf{M}_l &= (\Gamma_1, ..., \Gamma_l)\tilde{\Gamma}_{l-1}^{-1}(\Gamma_1^H, ..., \Gamma_l^H)^H \end{aligned}$$

where $\tilde{\Gamma}_{l-1}$ denotes the sub-matrix of Γ_{Aug} obtained by keeping only its first $l-1$ rows and columns. This transformation operates recursively, enabling the foundational blocks of the Γ_{Aug} matrix to be transformed into square matrices that are positioned within the domain of the Siegel Disk.

The smooth manifold of $\mathbb{B}_{d \times p}$ is thus identified as a Kähler manifold [4], on which is possible to define a Kähler potential Φ [11, 12], computed as:

$$\Phi(\Gamma_{Aug}) = -log(det(\Gamma_{Aug})) - log(\pi e) \quad (8)$$

After applying some decomposition properties of the determinant of Γ_{Aug}, it is possible to compute the metric of the manifold simply as the Hessian matrix of the Kähler potential

$$\begin{aligned} ds^2 = & p \operatorname{trace}\left(\mathbf{P}_0^{-1} d\mathbf{P}_0 \mathbf{P}_0^{-1} d\mathbf{P}_0\right) \\ & + \sum_{l=1}^{p-1} (p-l) \operatorname{trace}\left(\left(\mathbf{I} - \Omega_l\Omega_l^H\right)^{-1} d\Omega_l \left(\mathbf{I} - \Omega_l^H\Omega_l\right)^{-1} d\Omega_l^H\right) \end{aligned}$$

$$(9)$$

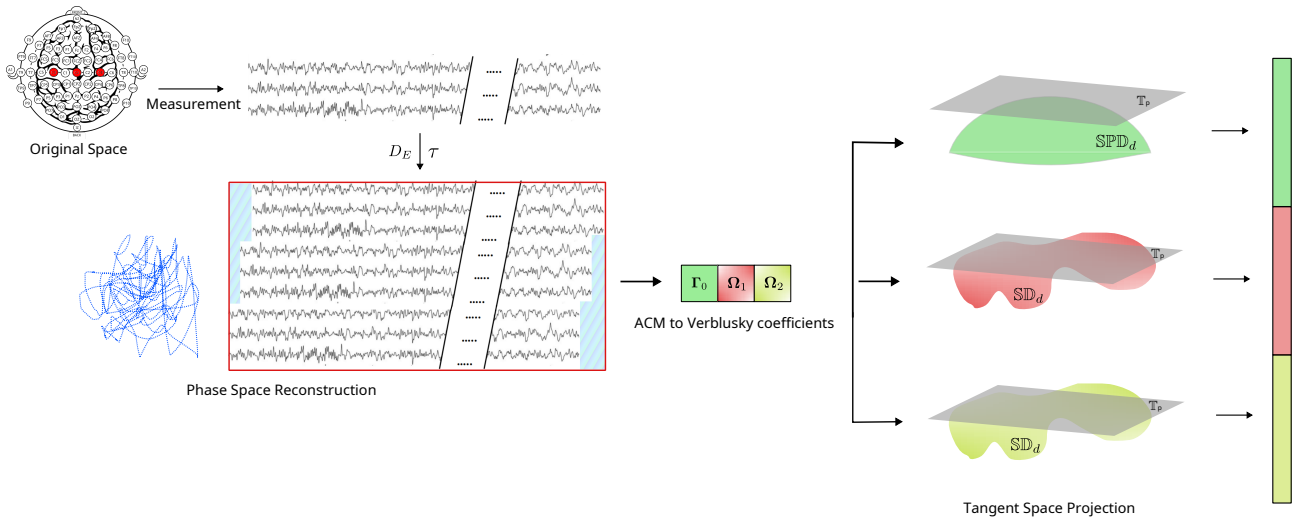[1]$A > B$ when $A - B$ is a positive definite matrix.

Figure 1: Schematic illustration of the BT-ACM + TG + SVM methodology. The presented example uses only 3 electrodes (in red on the top left plot). The measurement process of the original dynamic system is thus a 3-dimensional time series. The process then begins with the extraction of epoch signal representing left and right hand tasks. We then use the phase space reconstruction process to obtain a dynamic system equivalent to the original one (selection of hyper-parameters are made via grid search using the nested approach). In this figure, we see an embedding corresponding to $p = 3$ and $\tau = 10$). The BT-ACM matrix is computed as the autocovariance of this high-dimensional time series. Subsequently the main blocks are converted in Verblunsky coefficients. Then, each component is mapped to the tangent space using the appropriate Riemannian manifold computations and vectorized. The final step is the application of an SVM-based classification algorithm.

The first term of Equation (9) is identifiable as the metric for the $\mathbb{SPD}_d$ space. The other term represents the metric for the Siegel disk space $\mathbb{SD}_d$ repeated $p - 1$ times, i.e. the space of Block-Toeplitz SPD matrices is equipped with a product Riemannian metric over $\mathbb{SPD}_d \times \mathbb{SD}_d^{p-1}$.

The resulting algorithm (depicted in Fig. 1) using this new metric using SVM on the tangent space is called BT-ACM+TS+SVM: Block-Toeplitz Augmented Covariance Matrix (BT-ACM) with Tangent Space projection (TS) and SVM classifier.

Using fixed hyper-parameters p and τ, the PSR approach expands the signal (p and τ will subsequently be carefully chosen using a grid-search procedure). The spatial autocovariance matrix of the expanded signal is obtained using regularization through the Oracle Approximating Shrinkage Estimator (OAS) [9]. After the Verblunsky transformation, each component is mapped to the Tangent space using the Logarithmic map of each specific manifold. The final classification step is obtained with a Support Vector Machine (SVM) algorithm.

Dataset and Evaluation procedure: To validate the proposed methodology, we use open accessible datasets available from MOABB [7]. We selected 3 datasets with a total of 70 subjects. All the information regarding the considered datasets are presented in Tab. 1.

Table 1: Dataset considered during this study.

Dataset	subjects	channels	sampling rate	trials/class	Epoch (s)
BNCI2014001 [13]	9	22	250 Hz	144	[2, 6]
BNCI2014004 [14]	9	3	250 Hz	360	[3, 7.5]
Cho2017 [15]	52	64	512 Hz	100	[0, 3]

The duration of each epoch within our study is intentionally aligned with the task conditions' length, which is subject to variation across the datasets employed. On each dataset, we applied a standard band pass filter procedure for the Motor Imagery task, in the frequency range of 8 to 32 Hz.

We use a Within-Session (WS) evaluation procedure as provided in MOABB. This means that our analysis works on each session separately. The implementation is based on a Nested Cross-Validation methodology [16], structured with an outer loop of 5-Fold Cross validation and a inner one composed by a 3-fold Cross Validation. We use statistical tests provided by MOABB to confront the different pipelines, i.e., based on a t-test [17] for datasets with less than 20 subjects, or a Wilcoxon non-parametric signed-rank test [18] otherwise.

The state-of-the-art pipelines used in this research contain both Machine Learning (ML) and Deep Learning (DL) methods. Detail of the pipelines are listed in Tab. 2.

For DL pipelines, we used a standardization step that normalizes every channel to have a zero mean and unit standard deviation. Additionally, we employed a resampling procedure to ensure that each architecture integrates a temporal filter aligned with the state-of-the-art techniques' implementations. This procedure was added in order align to the state-of-the-art implementation and avoid the need of redoing hyper-parameter tuning. The DL pipelines are using a Sparse Categorical Cross-Entropy loss function and a standard Adam optimizer using 300 epochs and a batch size of 64. To avoid overfitting, we used an early stopping procedure with a patience parameter of 75.

RESULTS

Table 2: Pipelines considered in this study are organized into two distinct sections within the table: the first part is dedicated to the traditional classical ML pipelines, while the second part focuses on DL pipelines for MI.

Pipeline	Feature Extraction	Classifier
CSP + LDA [19]	Common Spatial Patterns (CSP) with OAS covariance estimator	Optimized Shrinkage LDA
MDM [1]	Spatial Covariance using OAS	Mean Distance to Mean (MDM)
FgMDM [1]	Spatial Covariance using OAS	Minimum Distance to Mean with geodesic filtering (FgMDM)
TS + EL [20]	Spatial Covariance using OAS mapped to TS	Optimized Elastic Network (EL)
TS + SVM [1]	Spatial Covariance using OAS mapped to TS	Optimized SVM
ACM + TS + SVM [3]	ACM with Sample Covariance Estimator mapped to TS	Optimized SVM
BT-ACM + TS + SVM (Proposed)	BT-ACM with Sample Covariance using OAS mapped to each respectively TS	SVM
ShallowConvNet [21]	Standardized and resample EEG signal at 250Hz	Convolutional Neural Network (CNN)
DeepConvNet [21]	Standardized and resample EEG signal at 250Hz	CNN
EEGNet 8 2 [22]	Standardized and resample EEG signal at 128Hz	CNN with architecture EEGNet

In this section, we describe the results obtained for the Right vs Left hand classification task.

Table 3: Performance (AUC) of Right hand vs Left hand classification. The table contains the results over all subjects (average plus or minus standard deviation).

Pipeline	BNCI2014001	BNCI2014004	Cho2017
CSP+LDA	0.82 ± 0.17	0.80 ± 0.15	0.71 ± 0.15
MDM	0.82 ± 0.15	0.78 ± 0.16	0.63 ± 0.14
FgMDM	0.87 ± 0.12	0.79 ± 0.15	0.73 ± 0.13
TS+EL	0.86 ± 0.13	0.80 ± 0.15	$\mathbf{0.76 \pm 0.14}$
TS+SVM	0.87 ± 0.14	0.79 ± 0.15	0.75 ± 0.14
ACM+TS+SVM	$\mathbf{0.92 \pm 0.10}$	$\mathbf{0.83 \pm 0.15}$	0.74 ± 0.15
BT-ACM+TS+SVM	0.89 ± 0.11	$\mathbf{0.83 \pm 0.14}$	$\mathbf{0.76 \pm 0.14}$
ShallowConvNet	0.86 ± 0.14	0.72 ± 0.18	0.74 ± 0.15
DeepConvNet	0.82 ± 0.16	0.72 ± 0.19	0.72 ± 0.13
EEGNet	0.77 ± 0.19	0.70 ± 0.20	0.67 ± 0.16

The various approaches of this study are compared in Tab. 3. A detailed picture of the results and the statistical significance of these results is provided in Fig. 2.

Overall, our method scores the best in 2 datasets - BNCI2014001 and Cho2017 - and obtains similar performance with respect to ACM+TS+SVM and TS+EL respectively for the third dataset. Across other datasets evaluated, our approach consistently delivers results that closely rival those of other leading algorithms, with a marginal performance deviation of no more than 1% in the AUC score with the only exception of BNCI2014001 dataset (where only the more costly ACM+TS+SVM does better).

Moreover, a comprehensive analysis across multiple datasets underscores a statistically significant performance enhancement achieved by BT-ACM+TS+SVM compared to all considered algorithms, with the sole exception of TS+EL, where the outcomes are remarkably similar (Fig. 2 (a) and (d)). This evidence collectively affirms the superiority of our method, not only in achieving high-performance benchmarks, but also in maintaining competitive results across different datasets.

We further explored the carbon emission and the computational time. To perform this analysis, we have run all the algorithms on the same hardware, a Dell C6420 dual-Xeon Cascade Lake SP Gold 6240 @ 2.60GHz. Furthermore, in order to conduct a fair comparison, we considered ACM+TS+SVM and BT-ACM+TS+SVM with the same number of parameter to optimize, i.e. we optimize the order and the lag of the augmentation procedure in the range $[1 - 10]$ without any optimization of the SVM parameter. Fig. 2 (b) shows the timings and

estimated carbon footprint for both the ACM+TS+SVM and the BT-ACM+TZ+SVM pipelines over the dataset BNCI2014001 composed by 9 subject over 2 sessions. The carbon footprint was estimated using Code Carbon [23] and expressed as gCO2 equivalent emission.

DISCUSSION

Our analysis across multiple datasets demonstrates (see Fig. 2 (a)) that the BT-ACM+TS+SVM algorithm not only competes with but frequently surpasses the performance of current state-of-the-art methodologies. The only exception to this trend is when compared to the TS+EL algorithm, where our results are statistically indistinguishable (Fig. 2 (d)). Despite their close relationship, the BT-ACM+TS+SVM algorithm shows a significant (even if small) superior performance compared to the ACM+TS+SVM algorithm (Fig. 2 (c)).

We also noticed an improved stability over changes of the SVM parameter, which thus does not need to be optimized as we did in the TS+SVM and ACM+TS+SVM cases (see Table 2).

In addition to its improved classification performance, the BT-ACM+TS+SVM algorithm also exhibits significant advancements in computational efficiency. The difference in computational time is statistically significant as we pass from a mean time of $(94.58 \pm 2.62)s$ for the ACM methodology to $(79.71 \pm 0.80)s$ for the BT-ACM+TS+SVM approach. The results have the same level of statistical significance for the carbon emission. Note that the comparison is done using the same number of parameters. The fact that we did not optimize the SVM regularization parameter is not the reason for this improvement.

It is also noteworthy that the ACM methodology exhibits a significant variability in computational times across different sessions and subjects, indicating a fluctuation in performance consistency. In contrast, the BT-ACM approach demonstrates enhanced stability, showcasing a more uniform and predictable computational time.

CONCLUSION

Throughout this research, we focused on the uses of the Block-Toeplitz Augmented Covariance matrix (BT-ACM) for Motor Imagery classification. This methodology extends the current ACM by comprehensively utilizing the Block-Toeplitz properties of the BT-ACM matrix.
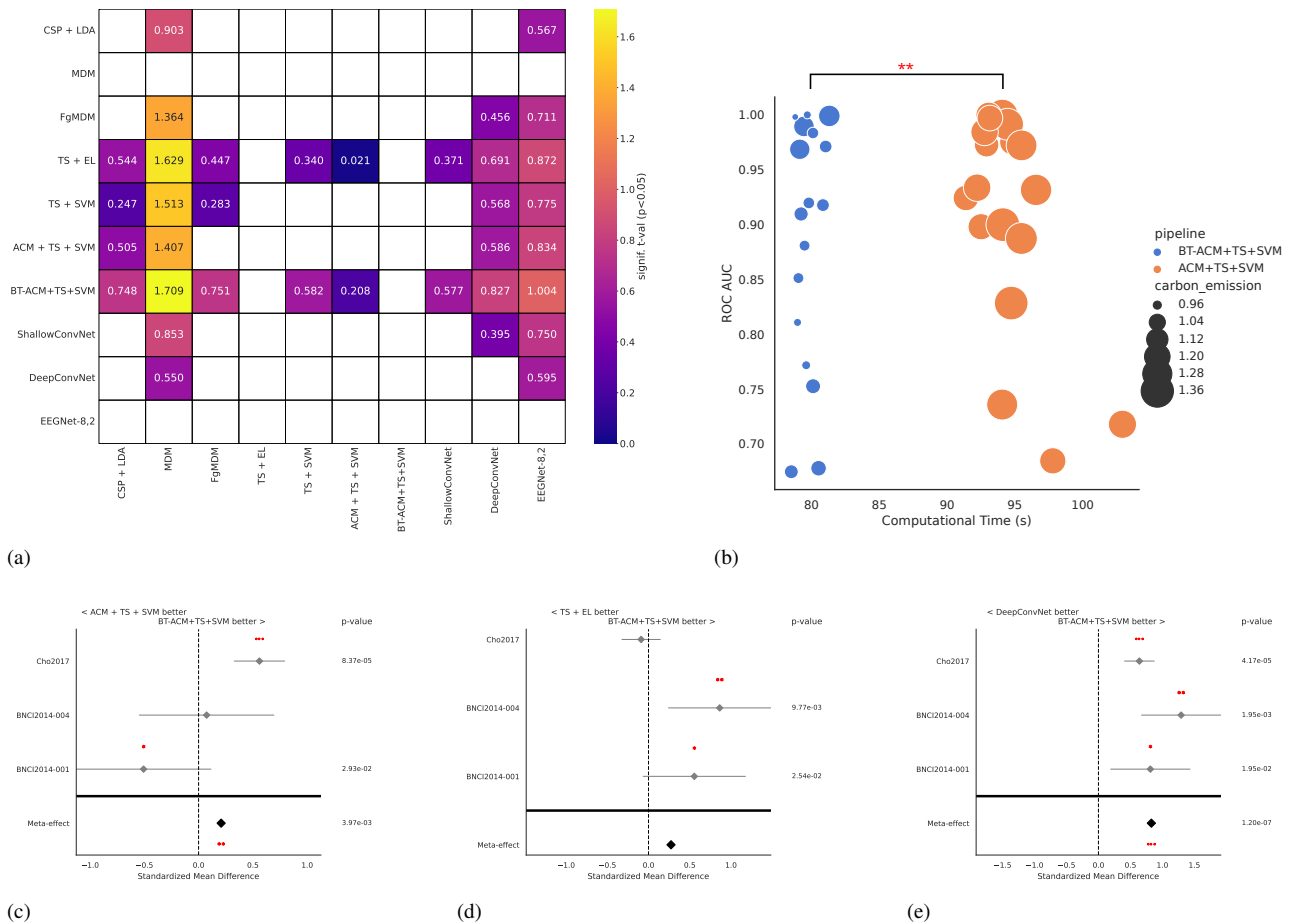
(a)

(b)

(c)

(d)

(e)

Figure 2: Results for Right vs Left hand classification, using WS evaluation. Plot (a) provides a combined meta analysis (over all datasets) of the different pipelines. It shows the significance that the algorithm on the y-axis is better than the one on the x-axis. The color represents the significance level of the difference of accuracy, in terms of t-values. We only show significant interactions ($p < 0.05$). Plots (b) summarizes the computational time and carbon footprint of ACM+TS+SVM vs BT-ACM+TS+SBM. Plots (c), (d) and (e) show the meta analysis of BT-ACM+TS+SVM against respectively ACM+TS+SVM (Grid), TS+EN, DeepConvNet. We show the standardized mean differences of p-values computed as one-tailed Wilcoxon signed-rank test for the hypothesis given as title of the plot. The gray bar denotes the 95% interval. * stands for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$.

This approach transforms the classification challenge, enabling a distinct analysis within the separate domain of SPD and Siegel Disk matrices.

This procedure achieves performance generally superior to the state-of-the-art, or at worst comparable. But, the improvement over the standard ACM is not only in terms of ROC-AUC but also in terms of significant reductions in computational costs and carbon emissions.

An interesting future exploration emerges from the fact that we are not using directly the property of the blocks contained in the BT-ACM matrix since the metric does not belong to any known manifold. In order to treat the problem, we were forced to use the Verblusky coefficient transform that introduces possible errors and complications. This means that it might be interesting to develop the mathematical framework for the direct treatment of such BT matrices.

ACKNOWLEDGMENT

* Data and Code Availability

The codes used to produce the results of this study are publicly available in this Github repository: `https://github.com/carraraig/BTACM_BCI`.

References

[1] Barachant A, Bonnet S, Congedo M, Jutten C. Riemannian geometry applied to bci classification. In: International conference on latent variable analysis and signal separation. 2010, 629–636.

[2] Förstner W, Moonen B. A metric for covariance matrices. In: Geodesy-the Challenge of the 3rd Millennium. Springer, 2003, 299–309.

Published by
Verlag der Technischen Universität Graz

[3] Carrara I, Papadopoulo T. Classification of bci-eeg based on augmented covariance matrix. arXiv preprint arXiv:2302.04508. 2023.

[4] Jeuris B, Vandebril R. The kahler mean of block-toeplitz matrices with toeplitz structured blocks. SIAM Journal on matrix analysis and applications. 2016;37(3):1151–1175.

[5] Cabanes Y. Multidimensional complex stationary centered gaussian autoregressive time series machine learning in poincaré and siegel disks: Application for audio and radar clutter classification. Ph.D. dissertation. Université de Bordeaux. 2022.

[6] Dette H, Wagener J. Matrix measures on the unit circle, moment spaces, orthogonal polynomials and the geronimus relations. Linear algebra and its applications. 2010;432(7):1609–1626.

[7] Aristimunha B *et al.* Mother of all bci benchmarks. Zenodo. 2023.

[8] Takens F. Detecting strange attractors in turbulence. In: Dynamical systems and turbulence, Warwick 1980. Springer, 1981, 366–381.

[9] Chen Y, Wiesel A, Eldar YC, Hero AO. Shrinkage algorithms for mmse covariance estimation. IEEE transactions on signal processing. 2010;58(10):5016–5029.

[10] Fritzsche B, Kirstein B. An extension problem for non-negative hermitian block toeplitz matrices. ii. Mathematische Nachrichten. 1987;131(1):287–297.

[11] Ballmann W. Lectures on kähler manifolds. European mathematical society (2006).

[12] Yang L. Medians of probability measures in riemannian manifolds and applications to radar target detection. Ph.D. dissertation. Université de Poitiers. 2011.

[13] Tangermann M *et al.* Review of the bci competition iv. Frontiers in neuroscience. 2012:55.

[14] Leeb R, Lee F, Keinrath C, Scherer R, Bischof H, Pfurtscheller G. Brain–computer communication: Motivation, aim, and impact of exploring a virtual apartment. Neural Systems and Rehabilitation Engineering, IEEE Transactions on. 2007;15(4):473–482.

[15] Cho H, Ahn M, Ahn S, Kwon M, Jun SC. Eeg datasets for motor imagery brain–computer interface. GigaScience. 2017;6(7):gix034.

[16] Cawley GC, Talbot NL. On over-fitting in model selection and subsequent selection bias in performance evaluation. The Journal of Machine Learning Research. 2010;11:2079–2107.

[17] Student. The probable error of a mean. Biometrika. 1908:1–25.

[18] Wilcoxon F. Individual comparisons by ranking methods. In: Breakthroughs in statistics. Springer, 1992, 196–202.

[19] Lotte F, Guan C. Learning from other Subjects Helps Reducing Brain-Computer Interface Calibration Time. In: International Conference on Audio Speech and Signal Processing (ICASSP). Dallas, United States, Mar. 2010, 614–617.

[20] Corsi MC, Chevallier S, Fallani FDV, Yger F. Functional connectivity ensemble method to enhance bci performance (fucone). IEEE Transactions on Biomedical Engineering. 2022.

[21] Schirrmeister RT *et al.* Deep learning with convolutional neural networks for eeg decoding and visualization. Human brain mapping. 2017;38(11):5391–5420.

[22] Lawhern VJ, Solon AJ, Waytowich NR, Gordon SM, Hung CP, Lance BJ. Eegnet: A compact convolutional neural network for eeg-based brain–computer interfaces. Journal of neural engineering. 2018;15(5):056013.

[23] Benoit C *et al.* Codecarbon. Zenodo. 2024.