Toward everyday BCI:

Augmented Covariance Method in a reduced dataset setting

Igor CARRARA[1,2], Théodore PAPADOPOULO[1,2]

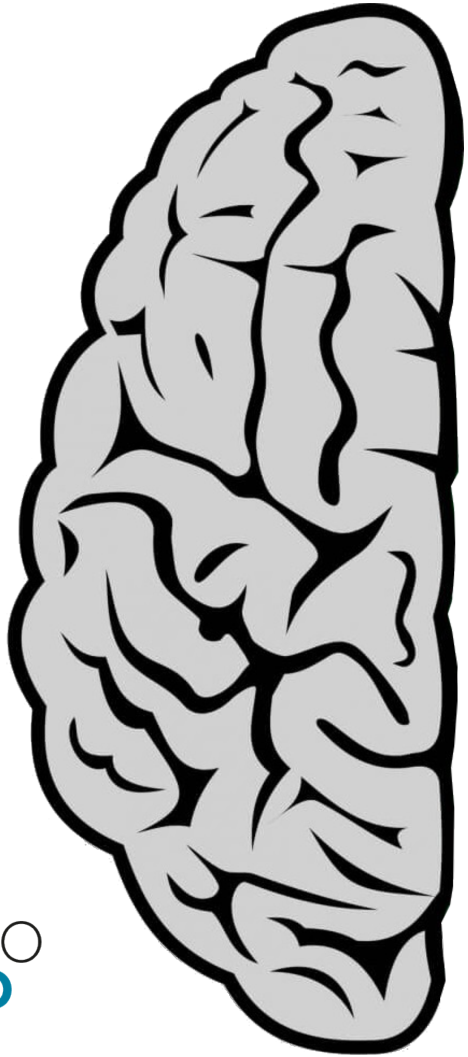
[1] Université Côte d'Azur (UCA), Nice, France,
[2] Inria centre at Université Côte d'Azur, Sophia Antipolis, France

Inria Cronos

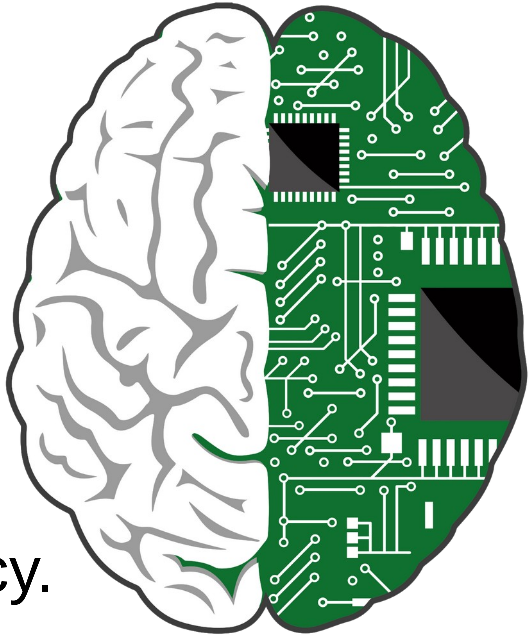
DIGITAL SYSTEMS FOR HUMANS
GRADUATE SCHOOL AND RESEARCH

UNIVERSITÉ CÔTE D'AZUR

NEURO MOD



Goals



Improve BCI effectiveness:

- Cost of equipment, setup or comfort
  → Use less electrodes.
- Efficiency / Usability
  → Reduce the training time…
  → … while keeping a reasonable accuracy.

Study how the augmented covariance approach can help.

- Within-session.
- Cross-session.

Tests done with MOABB on motor imagery benchmarks.

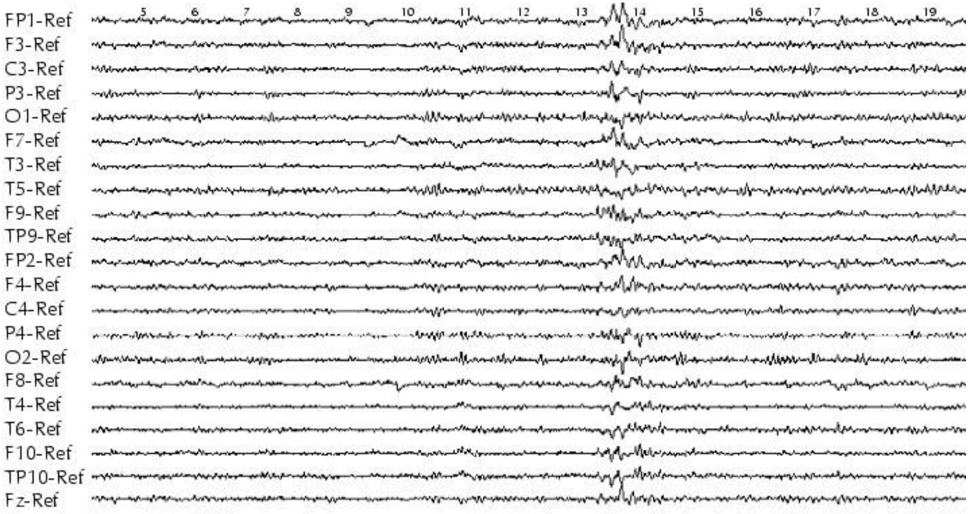
Augmented Covariance (ACM)

Autoregressive model formulation

$$\mathbf{X}_t = \sum_{i=1}^{p} \mathbf{A}_i \mathbf{X}_{t-i\tau} + \varepsilon_{\mathbf{t}}$$

Compute the Yule-Walker equations

$$\begin{cases} \mathbf{\Gamma}(\mathbf{0}) = \sum_{k=1}^{p} \mathbf{A}_k \mathbf{\Gamma}(-k) + \mathbf{U} \quad for \quad i = 0 \\ \boxed{\mathbf{\Gamma}(\mathbf{i}) = \sum_{k=1}^{p} \mathbf{A}_k \mathbf{\Gamma}(i-k) \quad for \quad i \neq 0} \end{cases}$$



With:

- $\mathbf{\Gamma}(\mathrm{i}) = \mathrm{E}(\mathrm{X}_t, \mathbf{X}_{t-i\tau}^T)$ is autocovariance matrix of lag i.
- $\boldsymbol{U}$ is the covariance matrix of the noise.
- 2 new parameters p and τ.

Autoregressive model formulation

$$\mathbf{X}_t = \sum_{i=1}^{p} \mathbf{A}_i \mathbf{X}_{t-i\tau} + \varepsilon_{\mathbf{t}}$$

Compute the Yule-Walker equation

$$\begin{cases} \mathbf{\Gamma(0)} = \sum_{k=1}^{p} \mathbf{A}_k \mathbf{\Gamma}(-k) + \mathbf{U} \quad for \quad i = 0 \\ \boxed{\mathbf{\Gamma(i)} = \sum_{k=1}^{p} \mathbf{A}_k \mathbf{\Gamma}(i-k) \quad for \quad i \neq 0} \end{cases}$$

SPD matrix

$$\boxed{\begin{bmatrix} \Gamma_0 & \Gamma_{-1} & \Gamma_{-2} & \cdots \\ \Gamma_1 & \Gamma_0 & \Gamma_{-1} & \cdots \\ \Gamma_2 & \Gamma_1 & \Gamma_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \Gamma_{p-1} & \Gamma_{p-2} & \Gamma_{p-3} & \cdots \end{bmatrix}} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_p \end{bmatrix} = \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \\ \Gamma_3 \\ \vdots \\ \Gamma_p \end{bmatrix}$$

Augmented Covariance Matrix (ACM)

Two strategies:

- Compute $(\mathbf{A}_i)_{i=1\ldots p}$ and classify.
- Classify the ACM directly.

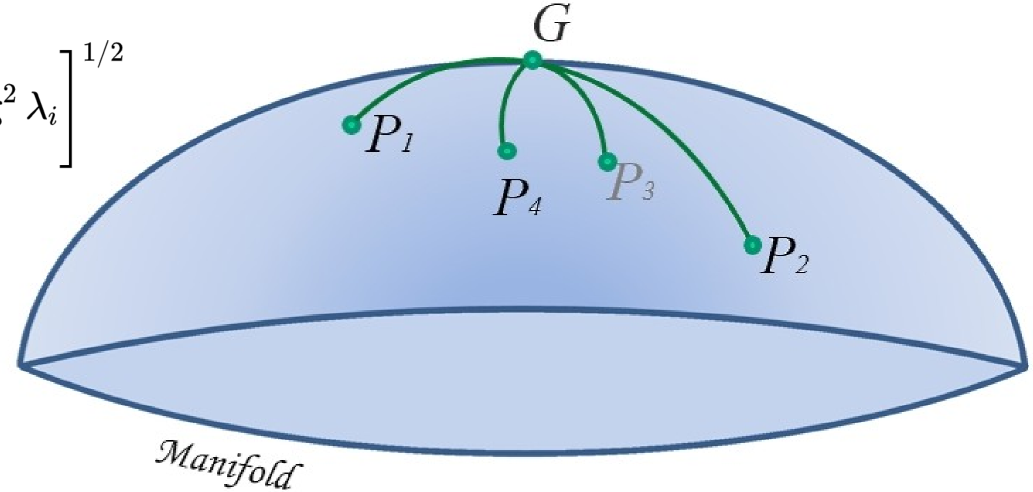
Riemann distance based classification

Affine Invariant metrics

$$\delta_R(\mathbf{P}_1, \mathbf{P}_2) = \left\| \log\left(\mathbf{P}_1^{-\frac{1}{2}} \mathbf{P}_2 \mathbf{P}_1^{-\frac{1}{2}}\right) \right\|_F = \left[\sum_{i=1}^{n} \log^2 \lambda_i\right]^{1/2}$$

Fréchet mean

$$\mathfrak{G}(\mathbf{P}_1, \dots, \mathbf{P}_m) = \underset{G \in P(n)}{\operatorname{argmin}} \sum_{i=1}^{m} \delta_R^2(\mathbf{G}, \mathbf{P}_i)$$



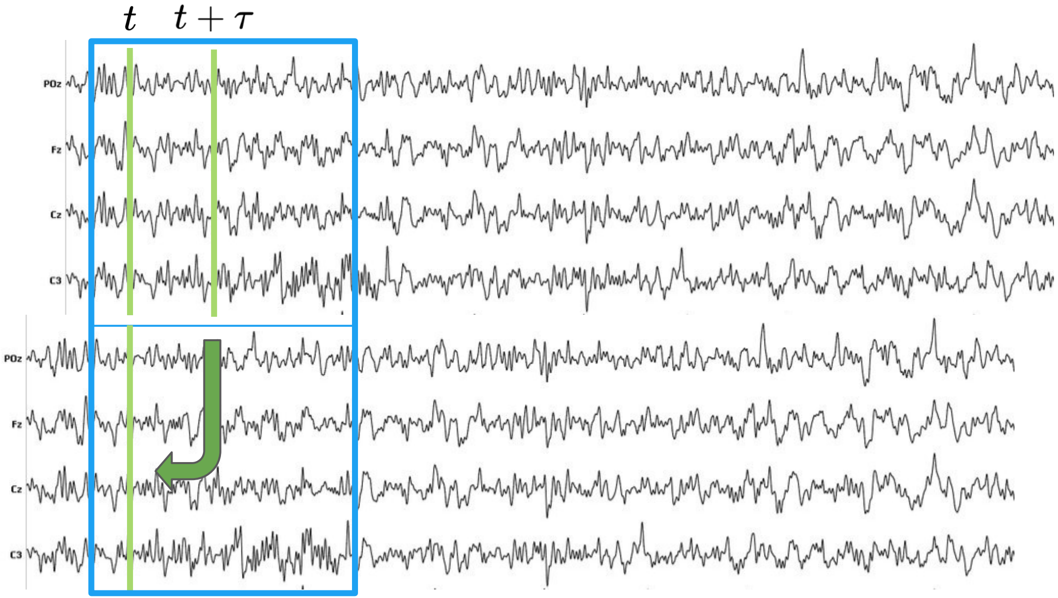
SPD matrix considered

$$Cov(X) = \frac{1}{Time - 1} \sum_{i=1}^{Time} X_i X_i^T$$

Alexandre Barachant et al. “Riemannian geometry applied to BCI classification”. In: *International conference on latent variable analysis and signal separation*. Springer. 2010, pp. 629–636.

Maher Moakher. “A differential geometric approach to the geometric mean of symmetric positive-definite matrices”. In: *SIAM Journal on Matrix Analysis and Applications* 26.3 (2005), pp. 735–747.

Alternative formulation



Taken’s Theorem

$$\Gamma_{aug} = \begin{bmatrix} \Gamma_0 & \Gamma_{-1} & \Gamma_{-2} & \cdots \\ \Gamma_1 & \Gamma_0 & \Gamma_{-1} & \cdots \\ \Gamma_2 & \Gamma_1 & \Gamma_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \Gamma_{p-1} & \Gamma_{p-2} & \Gamma_{p-3} & \cdots \end{bmatrix}$$

Sample Covariance

$$Cov(X) = \frac{1}{Time - 1} \sum_{i=1}^{Time} X_i X_i^T$$

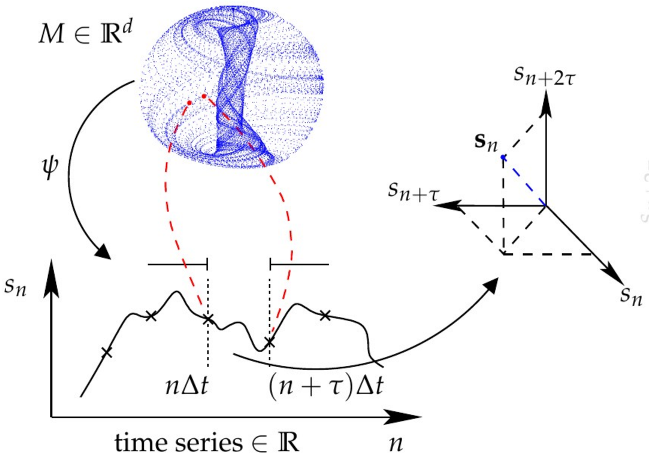
Introducing block-Toeplitz covariance matrices to remaster linear discriminant analysis for event-related potential brain–computer interfaces.
Jan Sosulski and Michael Tangermann. JNE, 2022.

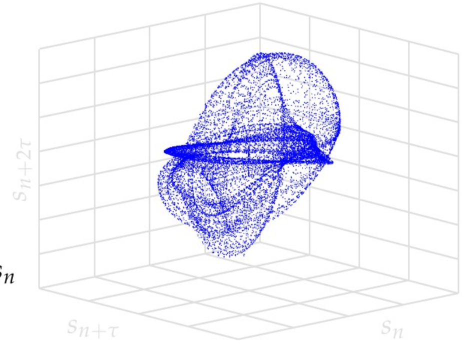
Takens Theorem

Let M to be a compact manifold of dimension m. For pairs (Φ, Y), $\Phi : M \to M$ a smooth diffeomorphism and $Y : M \to \mathbb{R}$ a smooth function, it is a generic property that the map $\varphi_{(\Phi,Y)} : M \to \mathbb{R}^{2m+1}$, defined by

$$\varphi_{(\Phi,Y)}(x) = (Y(x), Y(\Phi(x)), \dots, Y(\Phi^{2m}(x)))$$

is an embedding; by "smooth" we mean at least $\mathbb{C}^2$. We refer to the function Y as measurement function. By embedding we mean that the reconstructed system and the original system have the same dynamic invariants.





reconstructed space $\in \mathbb{R}^m$

$s(n)$

$$s_{R(n)} = [s(n),\, s(n-\tau),\, \dots,\, s(n-(D-1)\tau)]^T$$

Javier de Pedro-Carracedo et al. "Phase space reconstruction from a biological time series: A photoplethysmographic signal case study". In: *Applied Sciences* 10.4 (2020), p. 1430.

Floris Takens. "Detecting strange attractors in turbulence". In: *Dynamical systems and turbulence, Warwick 1980*. Springer, 1981, pp. 366–381.

Experiments

ID	subjects	channels	sampling rate	sessions	tasks	trials/class	Epoch (s)
Zhou2016	3	14	250 Hz	3	3	160	[0, 5]
BNCI2014001	9	22	250 Hz	2	4	144	[2, 6]
BNCI2015001	12	13	512 Hz	1	2	200	[0, 5]
BNCI2014002	14	15	512 Hz	5	2	80	[3, 8]
BNCI2014004	9	3	250 Hz	1	2	360	[3, 7.5]

Bandpass [8, 35]Hz

Classification of different task

- Right hand vs Left hand.
- Right hand vs Feet.
- Right Hand vs Left Hand vs Feet.

Different evaluation

- Within-session (5 folds Cross Validation).
- Cross-session (LOOCV),

Classification Algorithm

- On Riemann Surface (MDM).
- On Tangent Space (using SVM),

Hyper-parameters

- Grid search.

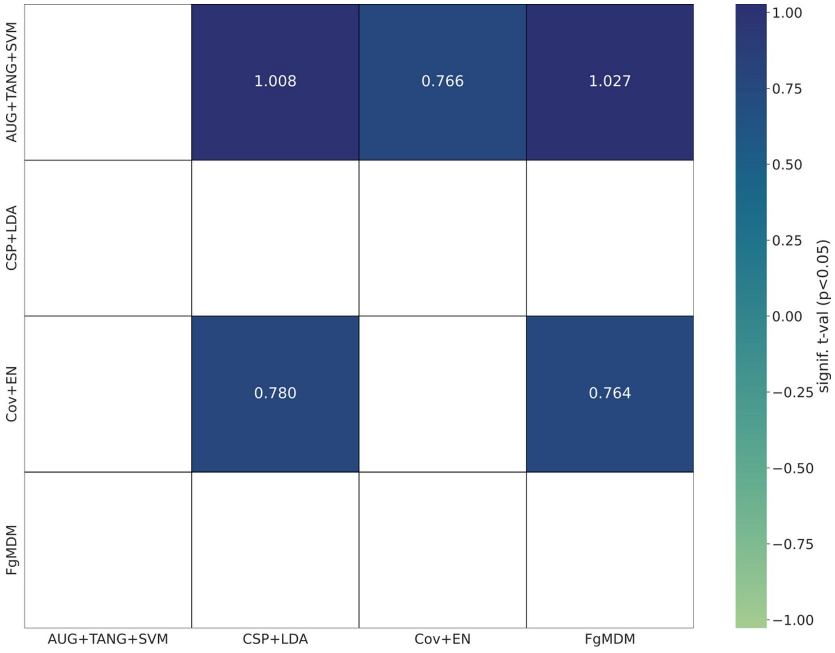
Carrara, I., & Papadopoulo, T. (2023). Classification of BCI-EEG based on augmented covariance matrix. *arXiv preprint arXiv:2302.04508*.

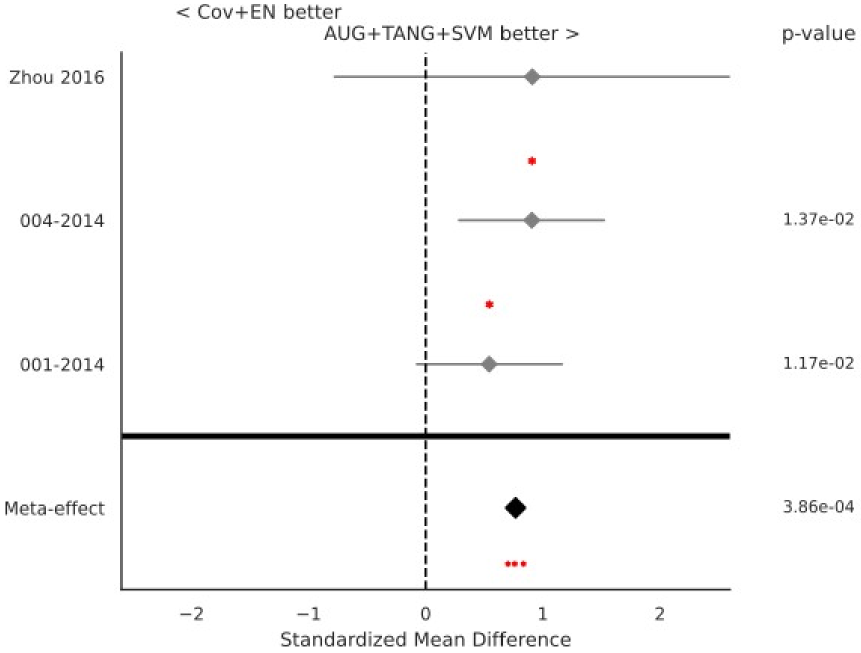
MOABB

Vinay Jayaram and Alexandre Barachant. "MOABB: trustworthy algorithm benchmarking for BCIs". In: *Journal of neural engineering* 15.6 (2018), p. 066011.

Classification results on **Right Hand vs Left Hand** task w.r.t the **state of the art** using a **within-session** evaluation procedure.

Dataset	CSP+LDA	AUG+TANG+SVM (Grid Search)	COV+EN	FgMDM
BNCI2014001	0.85 ± 0.14	$\mathbf{0.93 \pm 0.09}$	0.88 ± 0.11	0.88 ± 0.11
BNCI2014004	0.80 ± 0.15	$\mathbf{0.83 \pm 0.15}$	0.80 ± 0.15	0.79 ± 0.15
Zhou2016	0.91 ± 0.09	$\mathbf{0.95 \pm 0.06}$	0.93 ± 0.07	0.91 ± 0.08





Classification results on **Right Hand vs Left Hand** task w.r.t the **state of the art** using a **within-session** evaluation procedure.

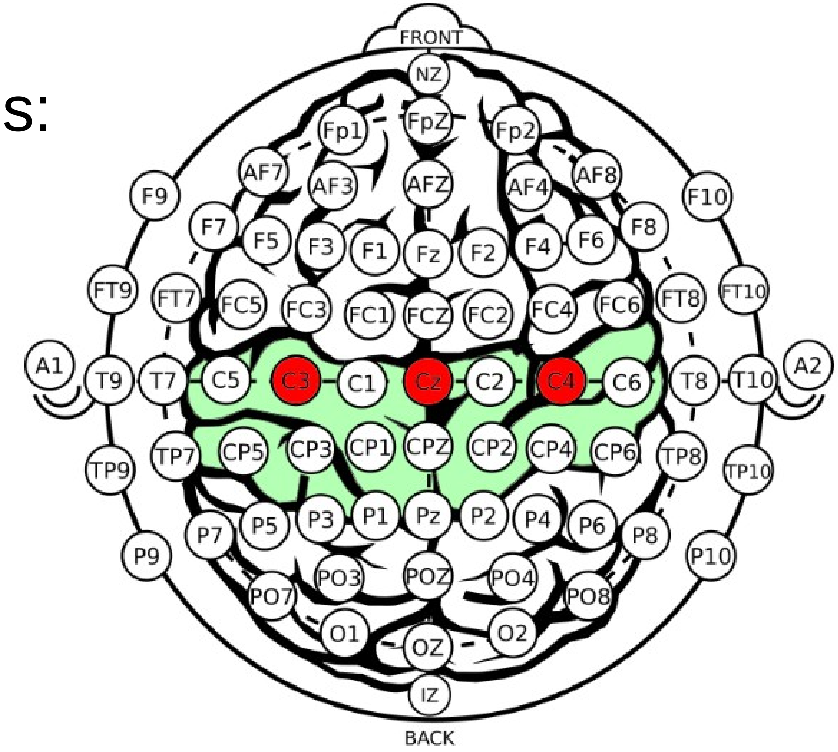
Dataset	CSP+LDA	AUG+TANG+SVM (Grid Search)	COV+EN	FgMDM
BNCI2014001	0.85 ± 0.14	$\mathbf{0.93 \pm 0.09}$	0.88 ± 0.11	0.88 ± 0.11
BNCI2014004	0.80 ± 0.15	$\mathbf{0.83 \pm 0.15}$	0.80 ± 0.15	0.79 ± 0.15
Zhou2016	0.91 ± 0.09	$\mathbf{0.95 \pm 0.06}$	0.93 ± 0.07	0.91 ± 0.08

- 2% to 5% over the second best method !!!
- Easily obtained.

Reduced Dataset setting

1. Reduced Number of Electrodes

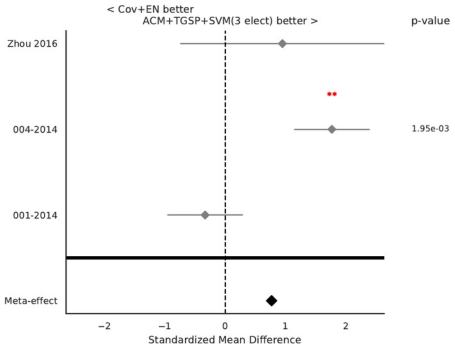
Only use the “motor” electrodes:
C3, Cz, C4

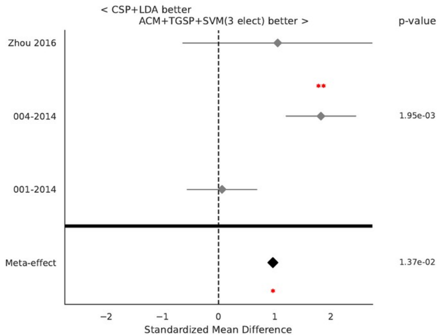


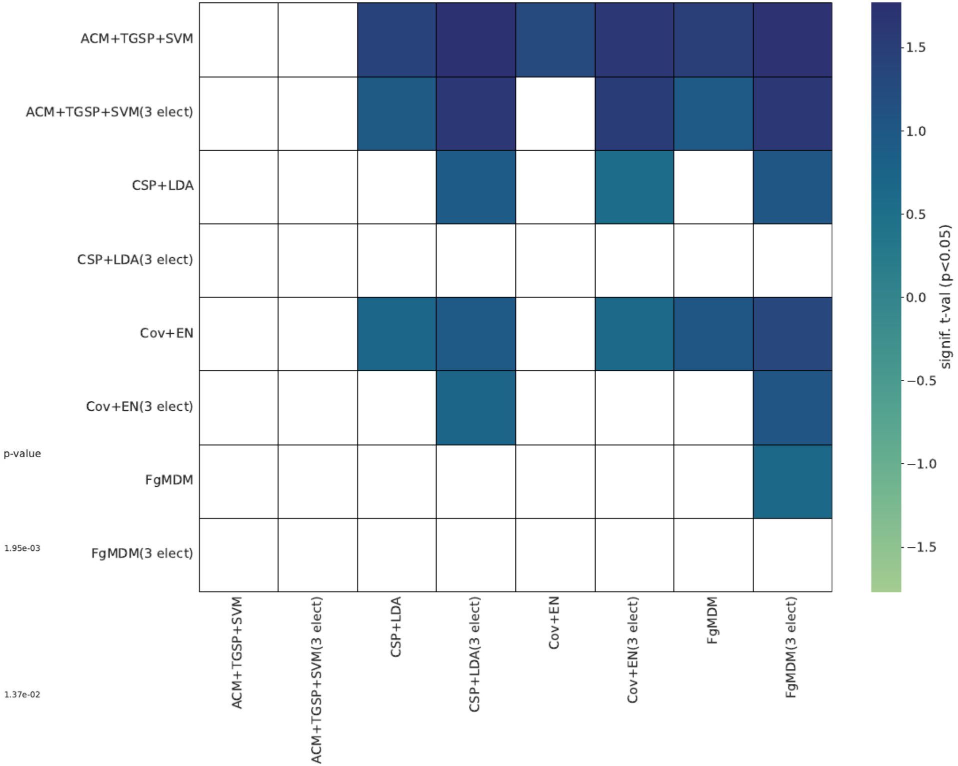
1. Reduced Number of Electrodes (C3, C4, Cz)

Different datasets, different tasks and evaluation procedures.

Pipeline	Eval	BNCI2014001	BNCI2014004	Zhou2016
MDM	WS	0.76 ± 0.15	0.78 ± 0.16	0.90 ± 0.05
COV+EN	WS	0.78 ± 0.15	0.80 ± 0.15	0.93 ± 0.05
FgMDM	WS	0.77 ± 0.15	0.79 ± 0.15	0.91 ± 0.05
ACM + TANG + SVM	WS	$\mathbf{0.86 \pm 0.10}$	$\mathbf{0.89 \pm 0.11}$	$\mathbf{0.98 \pm 0.03}$
CSP + LDA	WS	0.77 ± 0.16	0.80 ± 0.15	0.91 ± 0.06
ACM + CSP + LDA	WS	0.83 ± 0.12	0.85 ± 0.13	0.97 ± 0.02
VECTORIZE + LDA	WS	0.73 ± 0.15	0.80 ± 0.16	0.86 ± 0.08
ACM + VECTORIZE + LDA	WS	0.79 ± 0.13	0.83 ± 0.15	0.92 ± 0.04
MDM	CS	0.75 ± 0.16	0.79 ± 0.15	0.90 ± 0.04
COV+EN	CS	0.78 ± 0.14	0.81 ± 0.14	0.92 ± 0.06
FgMDM	CS	0.78 ± 0.13	0.80 ± 0.14	0.91 ± 0.06
ACM + TANG + SVM	CS	$\mathbf{0.84 \pm 0.11}$	$\mathbf{0.86 \pm 0.14}$	$\mathbf{0.97 \pm 0.02}$
CSP + LDA	CS	0.77 ± 0.15	0.81 ± 0.14	0.92 ± 0.05
ACM + CSP + LDA	CS	0.81 ± 0.13	0.84 ± 0.14	$\mathbf{0.97 \pm 0.02}$
VECTORIZE + LDA	CS	0.74 ± 0.14	0.79 ± 0.15	0.85 ± 0.11
ACM + VECTORIZE + LDA	CS	0.80 ± 0.12	0.84 ± 0.14	0.93 ± 0.03

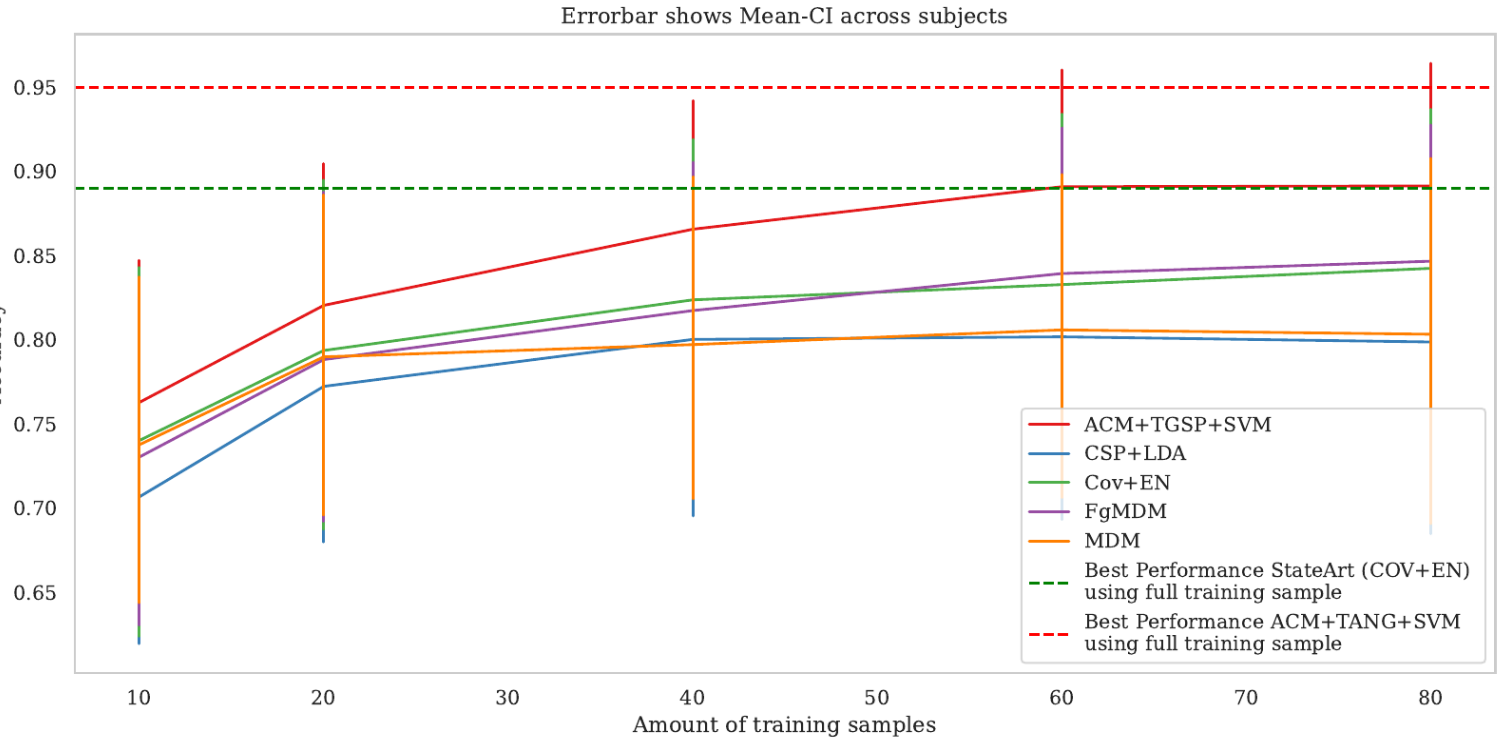






2. Reduced Training Set

Different datasets, tasks in the Within-Session evaluation framework.



Conclusion

- Augmented Covariance Method
  is an “easy” way to improve BCI MI classification.
- Introduces only 2 hyper-parameters.
- Classification improvements can be used to reduce the number of channels or the number of training sample, while maintaining an accuracy comparable to the non-ACM state of the art.
  → More classification improvements.

Future Directions

- Explore different distances.
- Investigate what makes the Riemannian distance so effective.
- Find better ways than Grid Search to estimate hyper-parameters.
- Use the tools of AR models or dynamical system theory.

Thanks for the Attention

To know more, see Igor demo at the end of the session…
And also poster 1-F-51

Cronos

We are looking for 1 (2) postdocs and 1 engineer → Contact me for details.