Leveraging robust statistics for EEG classification

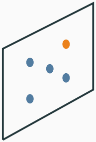
Imen Ayadi, **Florent Bouchard**, Frédéric Pascal

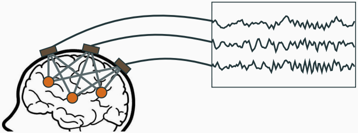
Université Paris Saclay, CNRS, CentraleSupélec, laboratoire des signaux et systèmes

Workshop ILLS - DATAIA – May, 24th 2023

Context – EEG classification

- Electrical brain activity recorded with electrodes placed on the scalp [Berger, 1929]
- BCI: interact with computer through brain signals [Lotte et al., 2018]

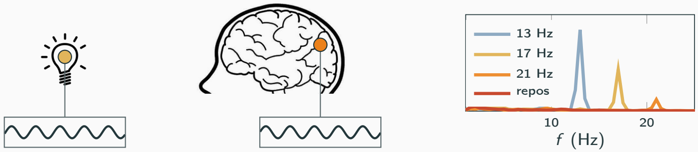




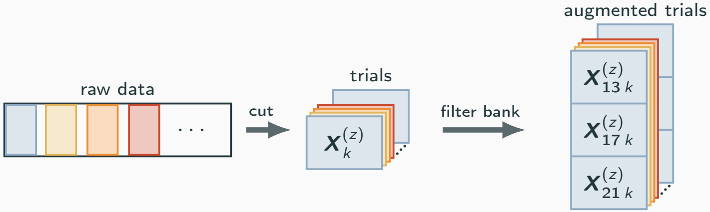
- Pipeline:

EEG recording → Pre-processing → Feature extraction → Machine learning

- Principle:

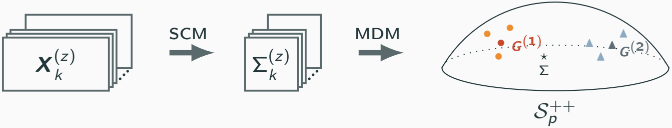


- Feature extraction:



Context – State-of-the-art MDM classifier

[Barachant et al., 2011]



- SCM: $\Sigma_k^{(z)} = \frac{1}{n} \boldsymbol{X}_k^{(z)} \boldsymbol{X}_k^{(z)\,T}$
- Fréchet mean: $\boldsymbol{G}^{(z)} = \underset{\boldsymbol{G} \in \mathcal{S}_p^{++}}{\text{argmin}} \sum_k \delta^2(\boldsymbol{G}, \Sigma_k^{(z)})$
- MDM decision rule for unknown $\boldsymbol{X}$:

$$z^* = \underset{z \in \{1, \ldots, Z\}}{\text{argmin}} \{\delta^2(\boldsymbol{G}^{(z)}, \Sigma)\}$$

Context – EEG Classification

Why BCI not yet out of the lab ?

- EEG data noisy, contain outliers
- Recorded in controlled environment to limit noise
- Preprocessing: data manually curated

Our take on these issues: leverage robust statistics theory

[Ollila et al., 2012, Maronna et al., 2019]

First try – Centered multivariate t-distribution

First Try – Centered multivariate t-distribution [Ollila et al., 2012]

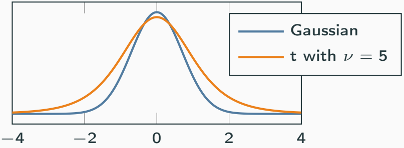
- Probability density function:

$$f(\{\mathbf{x}(i)\}_{i=1}^{n}|\Sigma) \propto \det(\Sigma)^{-n/2} \prod_{i=1}^{n} \left(\nu + \mathbf{x}(i)^T \Sigma^{-1} \mathbf{x}(i)\right)^{-(\nu+p)/2}$$

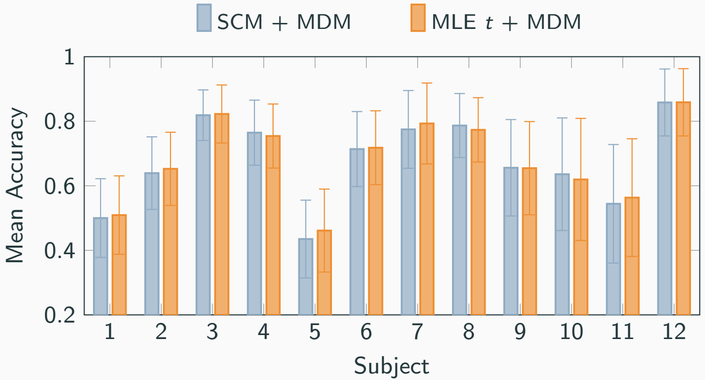
  - ν: degree of freedom – controls heaviness of tail

- MLE solution to:

$$\hat{\Sigma} = \frac{\nu + p}{n} \sum_{i=1}^{n} \frac{\mathbf{x}(i)\mathbf{x}(i)^T}{\nu + \mathbf{x}(i)^T \hat{\Sigma}^{-1} \mathbf{x}(i)}$$



First try – Results on SSVEP data



First try – Why no real improvement ?

- **Robustness not included at the right place**

Empirical cumulative density function of $\{\text{trace}(\Sigma_k^{(z)})\}_k$ for SSVEP data:

Second try – t-Wishart distribution

Second try – t-Wishart distribution

[Maronna et al., 2019]

- **Robustness: on distribution of covariance matrices**

- **Wishart distribution** $\mathcal{W}(n, \boldsymbol{G})$:
  - distribution of $\Sigma_{\boldsymbol{k}} = \boldsymbol{X}_k \boldsymbol{X}_k^T, \quad \boldsymbol{X}_k = [\boldsymbol{x}_k(i)] \in \mathbb{R}^{p \times n}, \quad \boldsymbol{x}_k(i) \sim \mathcal{N}(0, \boldsymbol{G})$
  - pdf:

$$f(\{\Sigma_k\}_{k=1}^K | \boldsymbol{G}) \propto \det(\boldsymbol{G})^{-nK/2} \prod_{k=1}^K \det(\Sigma_{\boldsymbol{k}})^{n-p-1/2} \exp(-\frac{1}{2} \operatorname{trace}(\boldsymbol{G}^{-1}\Sigma_{\boldsymbol{k}}))$$

  - MLE:

$$\hat{\boldsymbol{G}} = \frac{1}{nK} \sum_{k=1}^K \boldsymbol{S}_k$$

Second try – t-Wishart distribution

[Maronna et al., 2019]

t-Wishart distribution $t\text{-}\mathcal{W}(n, \boldsymbol{G}, \nu)$

$\Sigma_{\boldsymbol{k}} = \boldsymbol{X}_k \boldsymbol{X}_k^T \in \mathcal{S}_p^{++}, \quad \boldsymbol{X}_k \in \mathbb{R}^{p \times n},$ with pdf:

$$f(\{\Sigma_k\}_{k=1}^K | \boldsymbol{G}) \propto \det(\boldsymbol{G})^{-nK/2} \prod_{k=1}^{K} \det(\Sigma_{\boldsymbol{k}})^{n-p-1/2} \left(\nu + \operatorname{trace}(\boldsymbol{G}^{-1}\Sigma_{\boldsymbol{k}})\right)^{-(\nu+np)/2}$$

- ν: degree of freedom of t density generator, controls tail

▶ **Keep SCM, change their distribution**

Second try – t-Wishart distribution

MLE of t-$\mathcal{W}(n, \boldsymbol{G}, \nu)$ [Ayadi et al., 2023a]

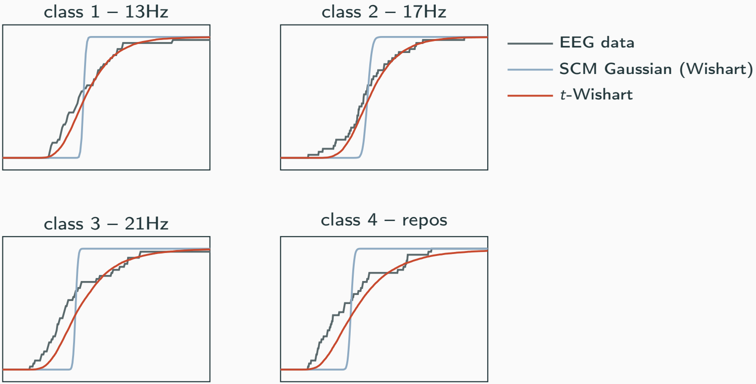
Samples $\{\Sigma_k\}_{k=1}^K$, MLE solution to:

$$\hat{\boldsymbol{G}} = \frac{\nu + np}{nK} \sum_{k=1}^{K} \frac{\Sigma_K}{\nu + \mathrm{trace}(\hat{\boldsymbol{G}}^{-1}\Sigma_k)}$$

- No analytical solution – iterative algorithm
- Developed estimation method based on Riemannian optimization

Second try – SSVEP data fitting

Empirical cumulative density function of $\{\mathrm{trace}(\Sigma_k^{(z)})\}_k$ for SSVEP data:



- From training data $\{\Sigma_k^{(z)}\}$, learn centers $\{\hat{\boldsymbol{G}}^{(z)}\}$

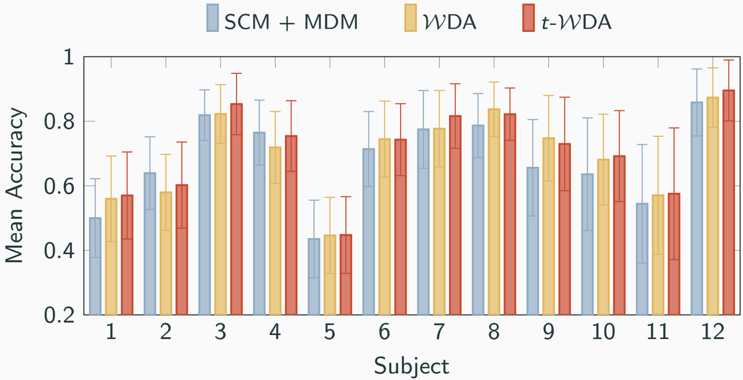
Decision rule of t-$\mathcal{W}$DA

For $\Sigma = \boldsymbol{X}\boldsymbol{X}^T$:

$$\hat{z} = \underset{z \in \{1,\dots,Z\}}{\operatorname{argmax}} \left\{ \log \hat{\pi}^{(z)} - \frac{n}{2} \log \det(\hat{\boldsymbol{G}}^{(z)}) - \frac{\nu + np}{2} \log(\nu + \operatorname{trace}(\hat{\boldsymbol{G}}^{(z)\,-1} \Sigma)) \right\}$$

- $\hat{\pi}^{(z)}$: proportion of class z in training set

Second try – Results on SSVEP data



Conclusions and perspectives

Conclusions and perspectives

- New estimator for t-Wishart distribution, new discriminant analysis
- Leveraging robustness for EEG classification can improve performance
- The way to include robustness matters a lot
- More validation still needed: different datasets, different paradigms
- Extension to P300 paradigm required: mean of $\boldsymbol{X}_k$ differs from zero

Ayadi, I., Bouchard, F., and Pascal, F. (2023a).
Elliptical Wishart distribution: maximum likelihood estimator from information geometry.
In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE.

Ayadi, I., Bouchard, F., and Pascal, F. (2023b).
t-$\mathcal{W}$da: A novel discriminant analysis applied to eeg classification.
In *submitted to European Conference on Signal Processing (EUSIPCO)*.

Barachant, A., Bonnet, S., Congedo, M., and Jutten, C. (2011).
Multiclass brain–computer interface classification by Riemannian geometry.
IEEE Transactions on Biomedical Engineering, 59(4):920–928.

Berger, H. (1929).
Über das elektrenkephalogramm des menschen.
Archiv für psychiatrie und nervenkrankheiten, 87(1):527–570.

Chevallier, S., Kalunga, E., Barthélemy, Q., and Yger, F. (2018).
Riemannian classification for SSVEP based BCI: offline versus online implementations.
In *BCI Handbook: Technological and Theoretical Advances*. CRC Press.

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., and Yger, F. (2018).
A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update.
Journal of neural engineering, 15(3).

Maronna, R. A., Martin, R. D., Yohai, V. J., and Salibián-Barrera, M. (2019).
Robust statistics: theory and methods (with R).
John Wiley & Sons.

Ollila, E., Tyler, D. E., Koivunen, V., and Poor, H. V. (2012).
Complex elliptically symmetric distributions: Survey, new results and applications.
IEEE Transactions on Signal Processing, 60(11):5597–5625.