Exploring new territory: Calibration-free decoding for c-VEP BCI

Jordy Thielen[1], Jan Sosulski[2] & Michael Tangermann[1]

[1] Radboud University, Donders Institute for Brain, Cognition and Behaviour, Machine Learning and Neural Computing, Data-Driven Neurotechnology Lab, Nijmegen, the Netherlands

[2] University of Freiburg, Department of Computer Science, Brain State Decoding Lab, Freiburg, Germany

DONDERS INSTITUTE

INTRODUCTION

Background

A code-modulated visual evoked potential (c-VEP) is brain activity recorded by EEG, triggered by rapid visual stimulation using a pseudo-random sequence of flashes[1]. Thanks to the rapid and optimized nature of its stimulus protocol, c-VEP-based BCIs are currently among **the fastest non-invasive BCIs** for communication and control[1].

Challenge

Before using a BCI, a machine learning model must typically be calibrated on labeled EEG data from the user, as each individual exhibits unique brain activity patterns. While this trained model is essential for accurate classification, its **calibration delays deployment**. Overall, the need for calibration can hinder the acceptance and widespread adoption of BCIs.

Approach

We evaluated two decoding approaches (CCA and UMM) that enable **instantaneous classification of unseen data without the need for a calibration session**. Instantaneous methods do not learn from past data, but instead only consider information from the current trial. While learning from multiple trials can improve a method's performance under stationary conditions, instantaneous approaches have the advantage that they naturally adapt to changes over time.

METHODS

CCA

Instantaneous classification: treat each trial separately

Reconvolution canonical correlation analysis (CCA)[2–3] fits a decoding-encoding model for each of the $N = 20$ candidate stimulus sequences $i \in \{1,\dots,N\}$. Stimuli are described by a structure matrix $\mathbf{M} \in \mathbb{R}^{M\times T}$ denoting the onset, duration M and overlap of events across time T. For the current single-trial $\mathbf{X} \in \mathbb{R}^{C\times T}$ of C channels, CCA learns sequence-specific spatial filters $\mathbf{w}_i \in \mathbb{R}^C$ and temporal responses $\mathbf{r}_i \in \mathbb{R}^M$ by optimizing the correlation ρ:

$$\arg\max_{\mathbf{w}_i,\mathbf{r}_i} \rho(\mathbf{w}_i^\top \mathbf{X}, \mathbf{r}_i^\top \mathbf{M}_i) \quad (1)$$

Determining the attended target symbol $\hat{y}$ is performed by maximizing the correlation ρ:

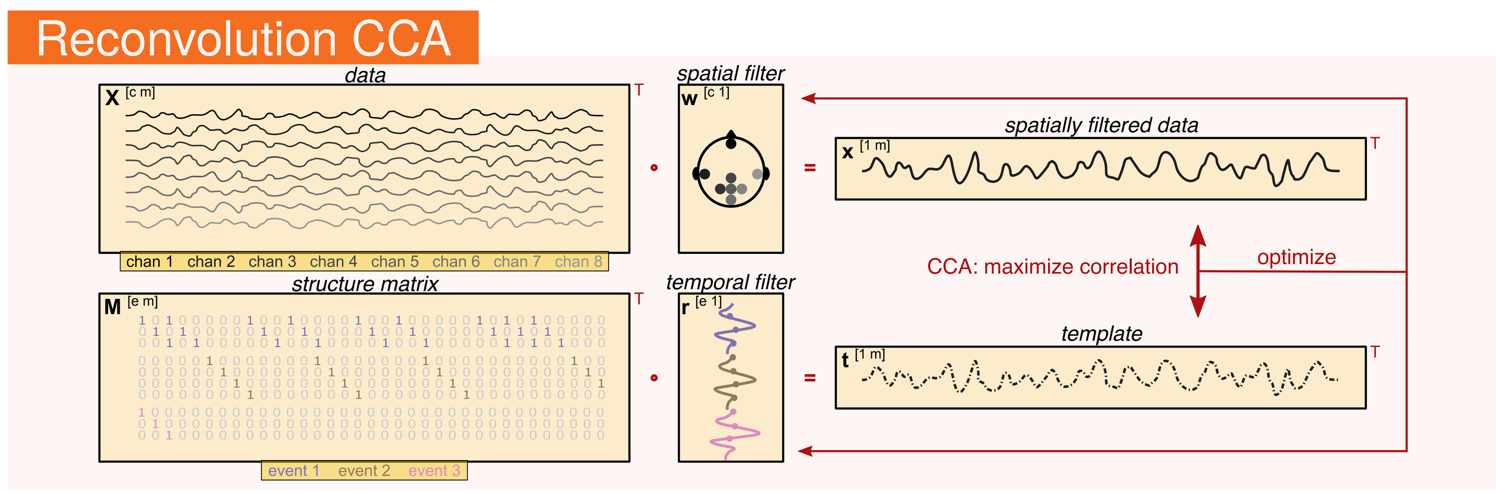
$$\hat{y} = \arg\max_i \rho(\mathbf{w}_i^\top \mathbf{X}, \mathbf{r}_i^\top \mathbf{M}_i) \quad (2)$$

The instantaneous CCA (CCA_e1) uses an empirical covariance matrix (denoted by postfix *e*) and the information provided by only the current trial (postfix *1*).

Cumulative classification: learn from previous trials

Previously classified trials can be included to improve the estimates for the current trial. Specifically, for the kth trial, $\mathbf{X} = [\mathbf{X}_0, \dots, \mathbf{X}_k]$ and $\mathbf{M} = [\mathbf{M}_{\hat{y}_0}, \dots, \mathbf{M}_{\hat{y}_{k-1}}, \mathbf{M}_i]$. The cumulative CCA (CCA_ec) used an empirical covariance matrix (postfix *e*) and cumulative (postfix *c*) information collected from previously classified trials.

Reconvolution CCA



The code for this CCA is available at: `https://github.com/thijor/pyntbci`.

UMM

Instantaneous classification: treat each trial separately

Unsupervised mean-difference maximization (UMM) identifies an attended target symbol by comparing distances between target and non-target ERPs[4–5] for all possible target outcomes. It slices the current trial into K-many epochs $\mathbf{x} \in \mathbb{R}^D$, synchronized to each bit in the stimulus sequences. Assuming symbol i to be target defines a set of target (A_i^+) and non-target epochs (A_i^-). UMM estimates the mean-difference vector $\Delta\boldsymbol{\mu}_i \in \mathbb{R}^D$ for each target hypothesis i:

$$\Delta\boldsymbol{\mu}_i = \frac{1}{|A_i^+|}\sum_{j\in A_i^+} \mathbf{x}_j - \frac{1}{|A_i^-|}\sum_{j\in A_i^-} \mathbf{x}_j \quad (3)$$

Obtaining the attended target symbol $\hat{y}$ is performed by maximizing the Mahalanobis distance:

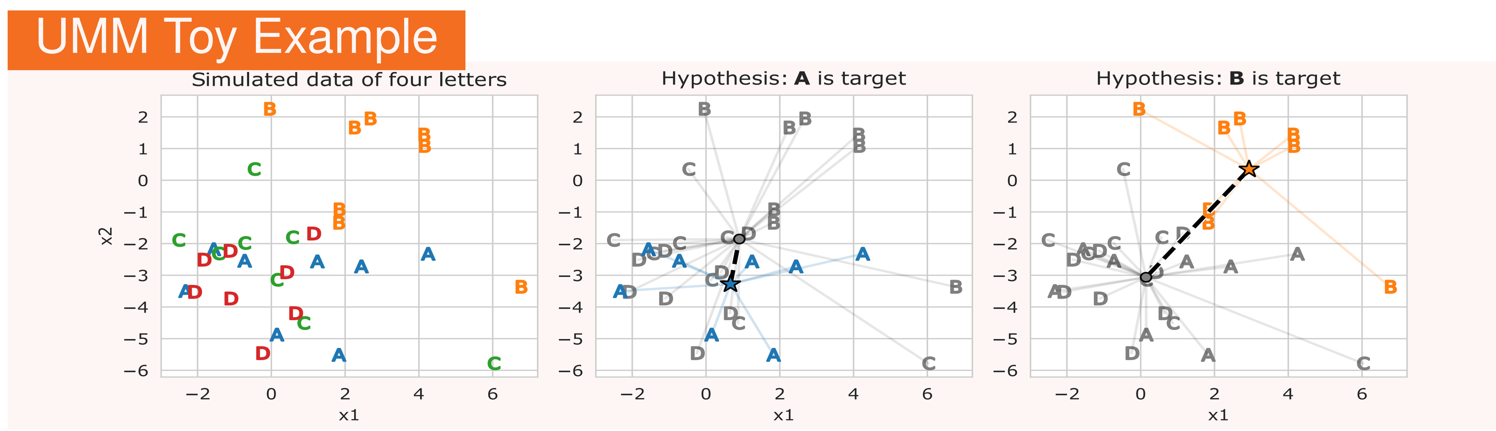
$$\hat{y} = \arg\max_i (\Delta\boldsymbol{\mu}_i)\boldsymbol{\Sigma}^{-1}(\Delta\boldsymbol{\mu}_i) \quad (4)$$

Instantaneous UMM (UMM_e11) uses an empirical covariance matrix (postfix *e*) and the current trial only to estimate the covariance matrix (postfix *1*) and class means (postfix *1*).

Cumulative classification: learn from previous trials

Information from previously classified trials can improve the estimate for the current trial. The covariance matrix can be updated without label information. The target and non-target means are updated using the predicted labels, weighted by their confidences[5]. This cumulative UMM (UMM_tcw) uses a block-Toeplitz regularized[4] covariance matrix (postfix *t*), previous trials for the covariance matrix (postfix *c*), and weighted ERP means (postfix *w*).

UMM Toy Example



Code for this UMM is available at: `https://github.com/jsosulski/umm_demo`.

RESULTS

c-VEP dataset

The open-access c-VEP dataset[6] of a previously published study[3] contains EEG data of 30 participants, acquired with a Biosemi Active2 amplifier at 512 Hz. Participants used a 4 × 5 matrix speller. Its $N = 20$ cells/symbols alternated between black and white at 60 Hz following 126-bit modulated Gold codes[2]. A total of 100 (5 per class) 31.5-second trials were collected.

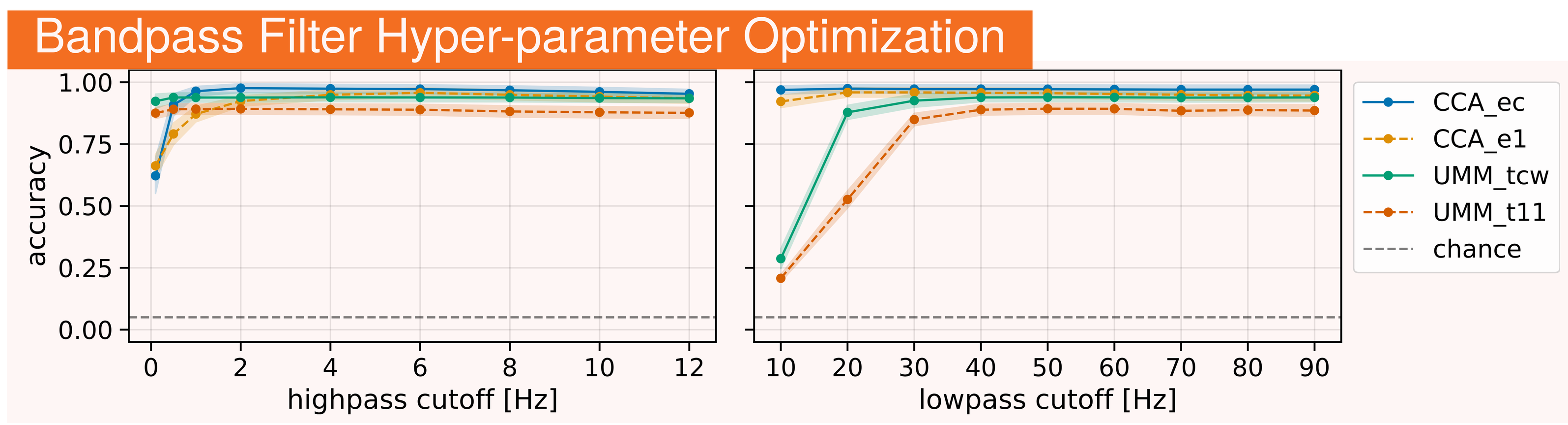
Optimization of bandpass hyper-parameters

Acknowledging the potential of each decoding method to reveal distinct responses for varying bandpass filters, we determined the optimal filter hyper-parameters for each method. The following observations were made:

- CCA_e1 and CCA_ec were more sensitive to a low highpass value than UMM_e11 and UMM_tcw.
- UMM_e11 and UMM_tcw were more sensitive to the lowpass value than CCA_e1 and CCA_ec.

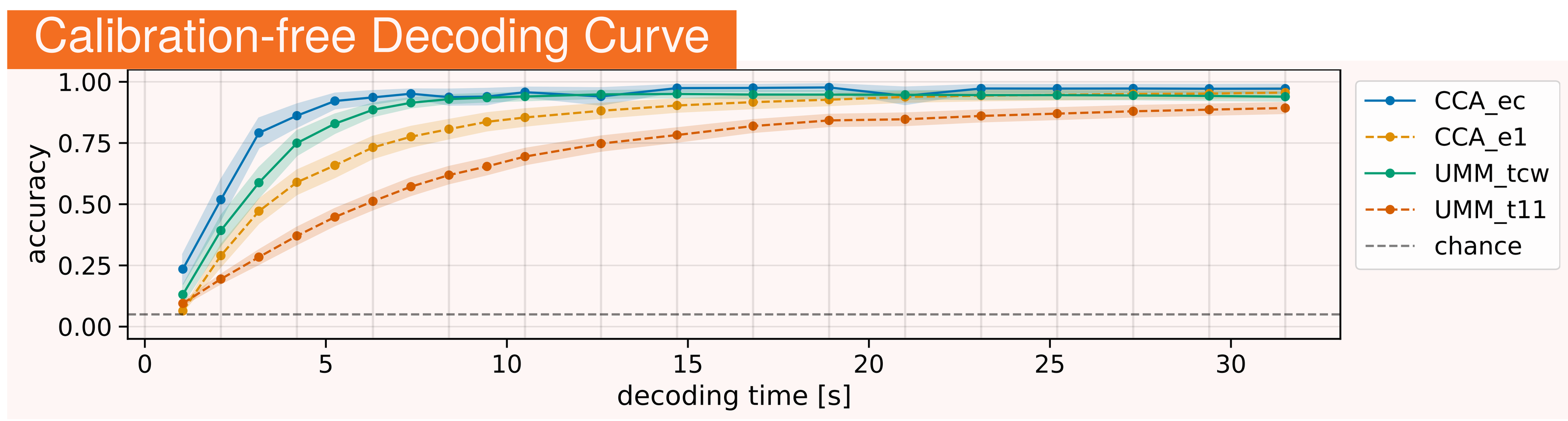
Based on these results, we used a generally optimal passband of 6 to 50 Hz.

Bandpass Filter Hyper-parameter Optimization



Calibration-free classification accuracy

Calibration-free Decoding Curve



We analyzed the performance of the methods by estimating decoding curves, retaining the chronological order of the 100 trials in the dataset for all 30 participants. We observed:

- The cumulative versions always outperformed the instantaneous version for both, CCA and UMM
- CCA_ec significantly outperformed UMM_tcw at all time-points except at 1.05 and 2.1 s.
- CCA_e1 significantly surpassed the accuracy of UMM_t11 at all timepoints except at 1.05 s.

Calibration-free Decoding Accuracy

	1.05 s	2.1 s	4.2 s	10.5 s	31.50 s
CCA_ec	**0.24**	**0.52**	**0.86**	**0.96**	**0.97**
CCA_e1	0.06	0.29	0.59	0.85	0.96
UMM_tcw	0.13	0.39	0.75	0.94	0.94
UMM_t11	0.09	0.19	0.37	0.69	0.89

DISCUSSION

Realizing instantaneous and cumulative zero-training

We observed 90 % at 14.70 s with CCA and 89 % at 29.40 s with UMM, so **instantaneous classification, using no other information than the current trial, was realized**. With cumulative classification, using information from previously classified trials, this was furhter improved to reaching 90 % at 5.25 s for CCA and at 7.35 s for UMM. This relatively high performance is surprising, given that UMM has not been designed with c-VEP data in mind.

Restraining the search space using the application domain

Both CCA and UMM exploit that **selecting a single symbol from a set of candidates is considerably simpler than reconstructing the stimulus sequence**. Here, with $N = 20$ symbols, one has to evaluate 20 candidate stimulus sequences only, while the exhaustive reconstruction would involve exponential growth, reaching 2^{60} candidates for a 1-second stimulus at 60 Hz.

Leveraging prior knowledge about the data domain

Estimating the empirical covariance matrix can be challenging with limited data. **UMM improved the covariance estimates by using domain-specific regularization** like shrinkage and block-Toeplitz regularization. These techniques repeatedly show superior results compared to vanilla counterparts.

OPEN QUESTIONS

Can CCA and UMM be unified?

CCA and UMM make different assumptions motivated by the c-VEP and P300/ERP protocols respectively. These assumptions may be met by different datasets to various degrees. Understanding the strengths and limitations of both methods bears the potential to develop refined versions tailored to specific characteristics of novel datasets.

Generalization to other BCI paradigms

CCA and UMM necessitate knowledge about the timing and sequence of stimuli within a trial. While such information is typically available in BCI protocols using evoked responses, it may not seamlessly extend to other protocols like those based on sensorimotor rhythms.

REFERENCES

[1] Martinez-Cagigal et al. (2021) *J Neural Eng* doi:10.1088/1741-2552/ac38cf

[2] Thielen et al. (2015) *PLOS ONE* doi:10.1371/journal.pone.0133797

[3] Thielen et al. (2021) *J Neural Eng* doi:10.1088/1741-2552/abecef

[4] Sosulski & Tangermann (2022) *J Neural Eng* doi:10.1088/1741-2552/ac9c98

[5] Sosulski & Tangermann (2023) *arXiv* doi:10.48550/arXiv.2306.11830

[6] Thielen et al. (2023) *Radboud Data Repository* doi:10.34973/9txv-z787

E-mail: `jordy.thielen@donders.ru.nl`
Twitter: `https://twitter.com/ThielenJordy`
Lab: `https://neurotechlab.socsci.ru.nl/`

Poster PDF:

Radboud University Radboudumc

Presented at the 9th Graz Brain-Computer Interface Conference 2024 in Graz, Austria on Thursday 12th September 2024