Optimizing P300 Speller Performance through Uncertainty Quantification

Bernard Renardi Ivo Pascal de Jong Andreea Ioana Sburlea

university of groningen

Department of Artificial Intelligence, Bernoulli Institute – University of Groningen

Why is P300 speller so promising yet challenging?

P300? When we are asked to detect a target, our brain produces signals around **300 milliseconds** after seeing the target.

Why promising? Now let's imagine the target is a letter. We can type letters by detecting when the P300 appears. **P300 speller** is known for this application and enables users to type using brain signals for impaired motor movements.

What's the catch? It is a slow paradigm and the data distribution is imbalanced. It takes **4.8 minutes** to type "HELP"!

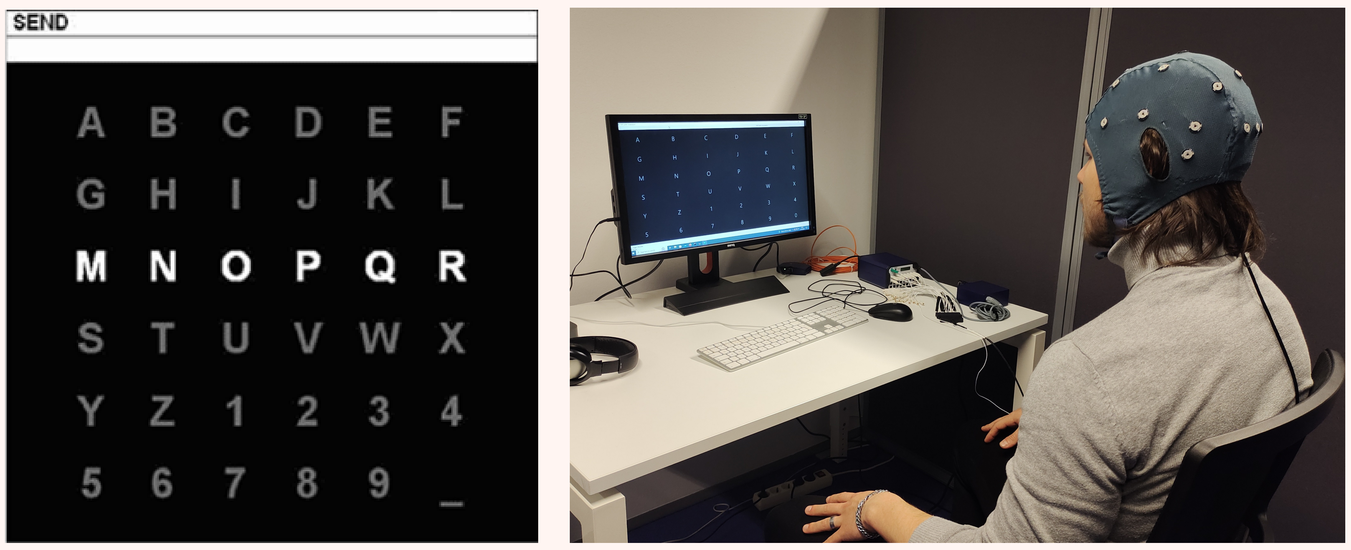

Figure 1. Left to right: A 6×6 matrix used in the P300 speller paradigm, followed by a demonstration of an experiment setup.

How does uncertainty quantification (UQ) help?

In BCI, UQ [3] can be used for: early stopping, which stops the process once the classifier is confident to produce an output, and trial rejection, which avoids making a decision when the system is uncertain.

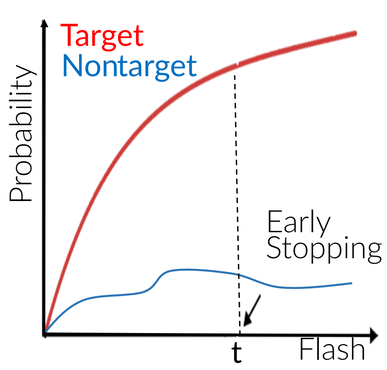

Figure 2. Continuous UQ based on updated posterior probabilities enables early stopping before all flashes are shown.

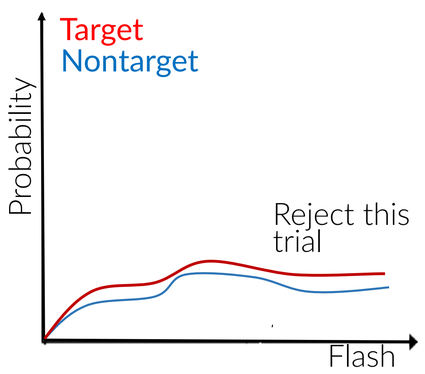

Figure 3. If no class shows sufficient confidence at the end of a trial, the system rejects the trial and withholds output.

Our methods

- **Dataset:** We used BNCI 2014-008 [2], which consists of data from 8 patients with Amyotrophic Lateral Sclerosis (ALS).
- **Classifier:** The Bayesian extension of the Riemannian MDM [1] was chosen for its invariance to linear transformations, non-parametric nature, and speed.

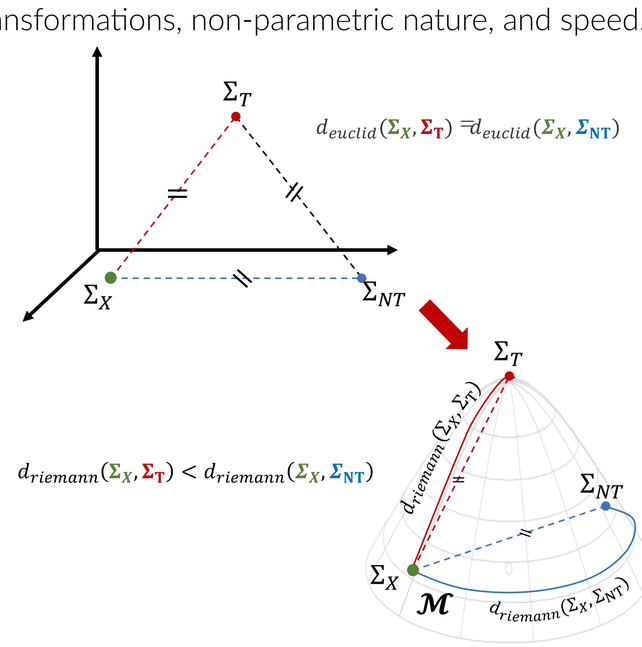

Figure 4. Suppose we have target (Σ_T) and non-target (Σ_{NT}) centroids, with (Σ_X) as the test sample. A Riemannian distance-based classifier has better sensitivity than a Euclidean-based classifier by measuring distances over a manifold $\mathcal{M}$. The test centroid (Σ_X) is classified as a target by Riemannian MDM, while the Euclidean classifier cannot, as the distances are the same.

Early Stopping

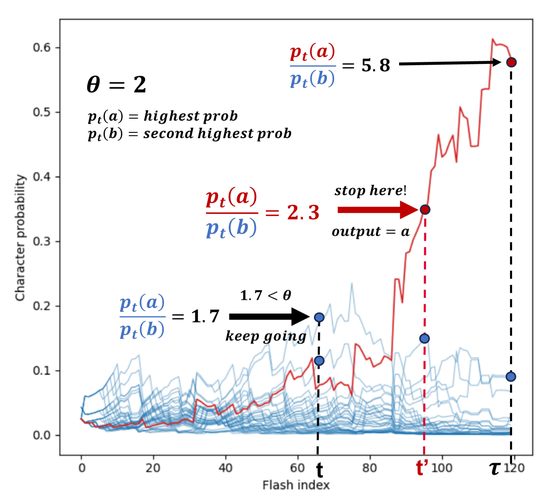

Figure 5. Bayesian accumulation from a single trial of a subject. With UQ, we are confident to stop the trial at an earlier flash t' rather than waiting until the end of the trial τ. θ is the ratio of the first and the second highest probability class.

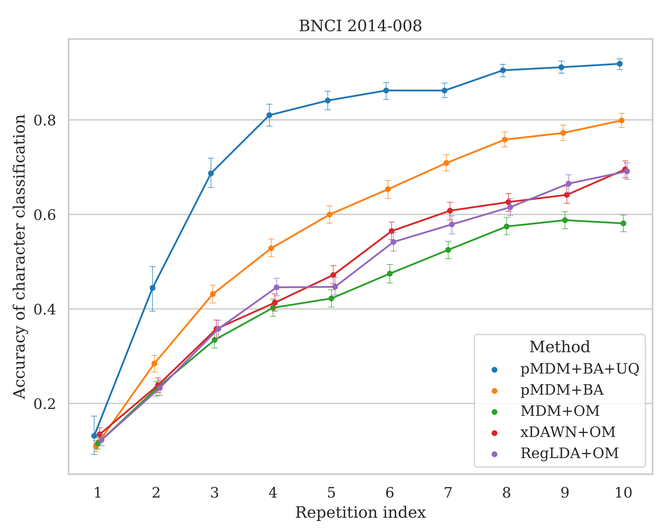

Figure 6. Average accuracy versus repetition index. It achieves the highest overall accuracy, saving 60% time to achieve the same accuracy as benchmark models.

Trial Rejection

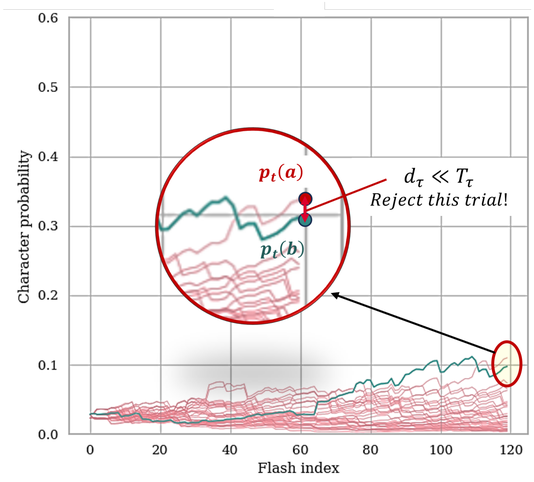

Figure 7. Bayesian accumulation from a single trial of another subject. With UQ, if uncertainty remains high by the end of the trial τ, the system rejects the trial instead of making a prediction.

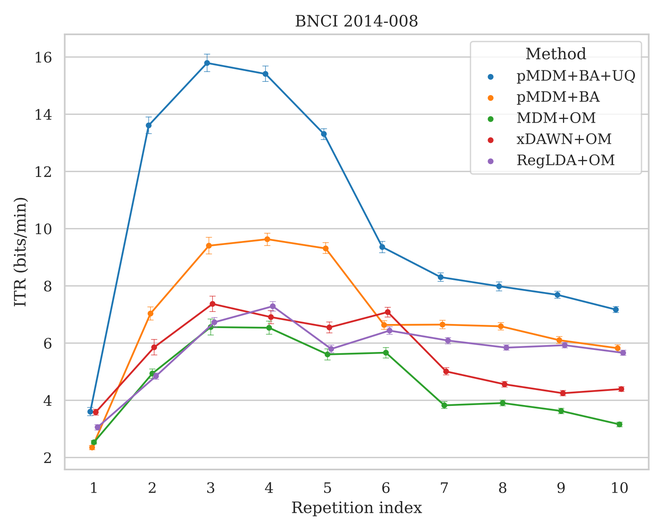

Figure 8. Average information transfer rate (ITR) comparison between pMDM+BA+UQ and other methods from all subjects in the dataset.

Table 1. Performance comparison of each subject within the dataset. Accuracy and F1-score are obtained from Bayesian Riemannian MDM model, and "+UQ" columns show results from Baseline + UQ. This highlights how UQ helps tackling imbalanced data distribution.

Subject	Centroids Dist	Coverage	Acc	Acc+UQ	F1-score	F1-score+UQ
1	0.542	0.64	0.82	0.89	0.20	0.31
2	0.545	0.75	0.71	0.86	0.12	0.25
3	0.610	0.89	0.86	0.84	0.25	0.23
4	0.373	0.11	0.39	1.00	0.03	1.00
5	0.667	0.79	0.79	0.95	0.17	0.54
6	0.664	0.82	0.86	1.00	0.25	1.00
7	0.669	0.96	0.96	0.96	0.60	0.59
8	0.695	1.00	1.00	1.00	1.00	1.00
Average	0.596 ± 0.21	0.75 ± 0.56	0.80 ± 0.38	0.94 ± 0.13	0.33 ± 0.64	0.62 ± 0.69

Conclusions & Future Works

- These results demonstrate the potential of UQ to enhance both the speed and accuracy of the P300 speller, offering a more efficient and robust solution.
- Our method could be further improved by combining early stopping and trial rejection techniques, as both use a similar approach based on the highest probability outputs.
- Adopting a data-driven approach to optimize these parameters could minimize manual intervention.

References

[1] Quentin Barthélemy, Sylvain Chevallier, Raphaëlle Bertrand-Lalo, and Pierre Clisson. End-to-end p300 bci using bayesian accumulation of riemannian probabilities. *Brain-Computer Interfaces*, 10(1):50–61, November 2022.

[2] Clemens Brunner et al. Bnci horizon 2020: towards a roadmap for the bci community. *Brain-Computer Interfaces*, 2(1):1–10, 2015.

[3] Ivo Pascal de Jong, Andreea Ioana Sburlea, and Matias Valdenegro-Toro. Uncertainty quantification in machine learning for biosignal applications – a review, 2023.

bernard.renardi111@gmail.com, ivo.de.jong@rug.nl, a.i.sburlea@rug.nl