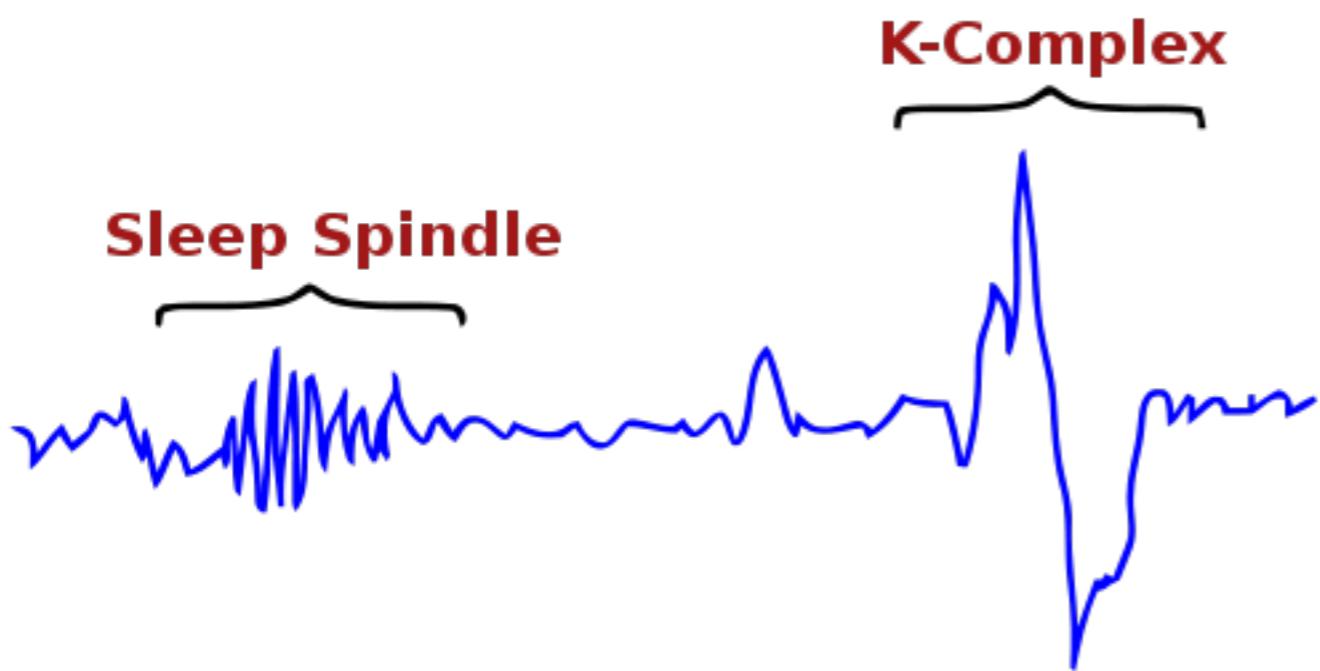


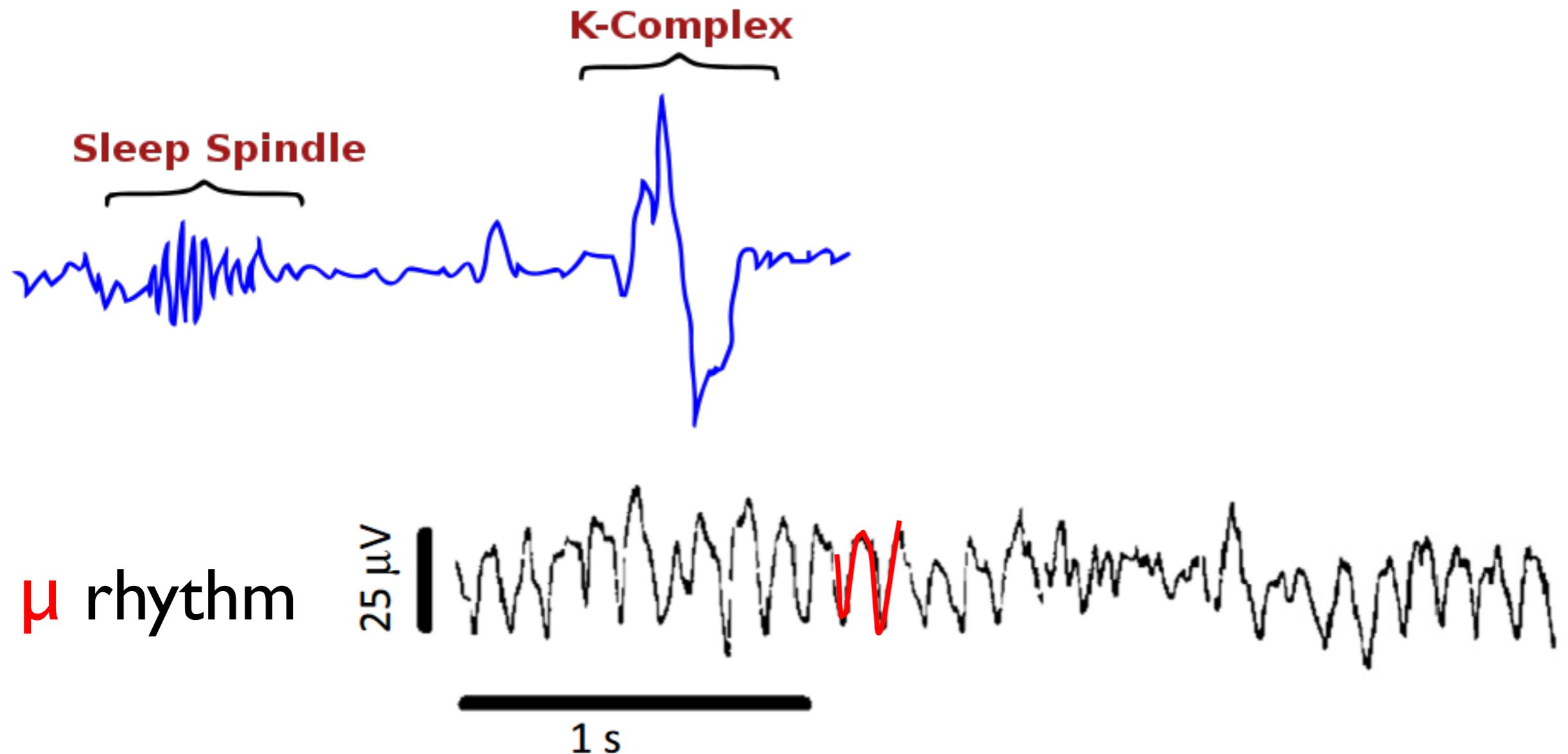
Machine Learning on EEG: From sleep to brain age

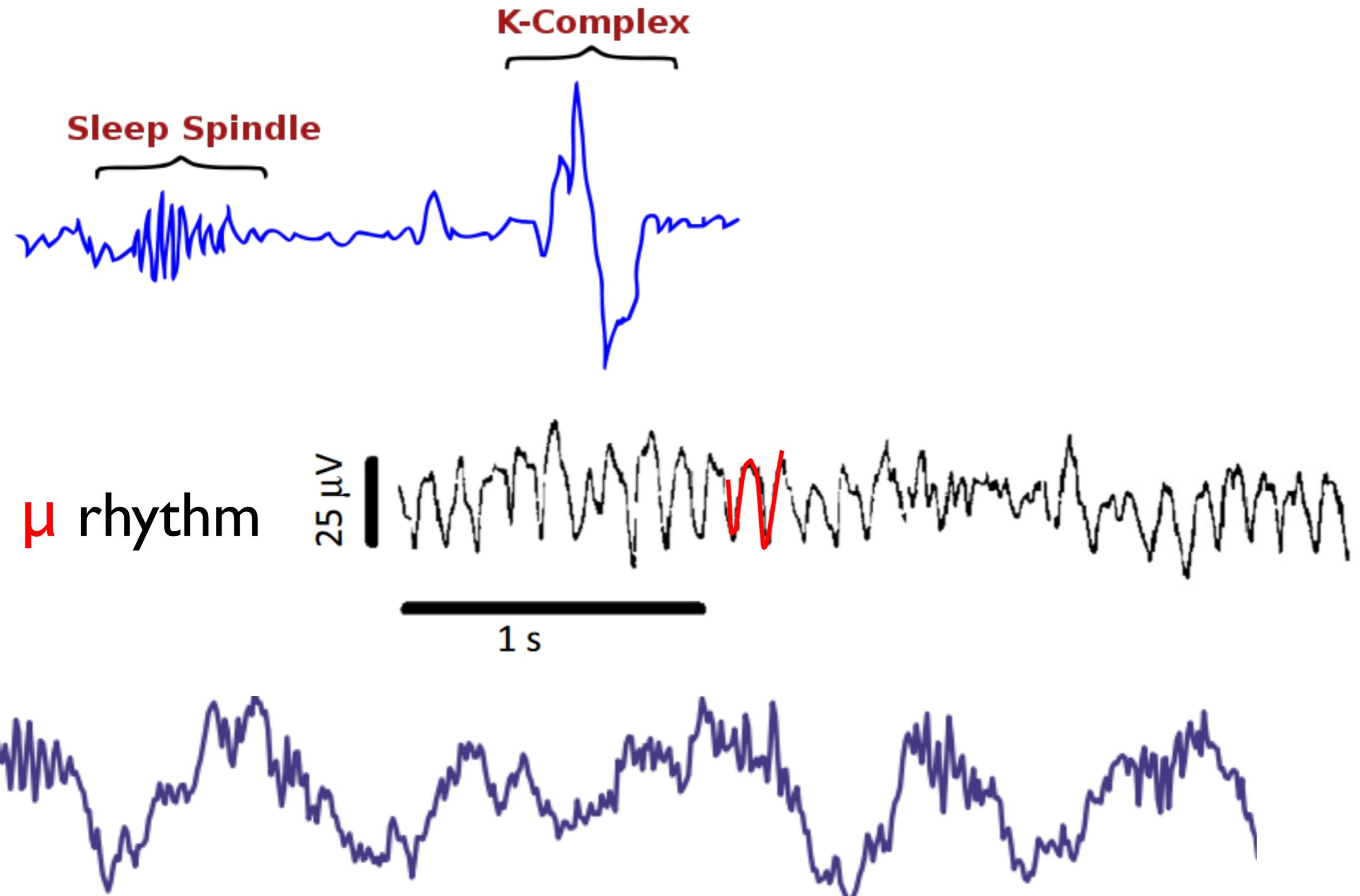
Alexandre Gramfort
<http://alexandre.gramfort.net>

Inria, Parietal Team

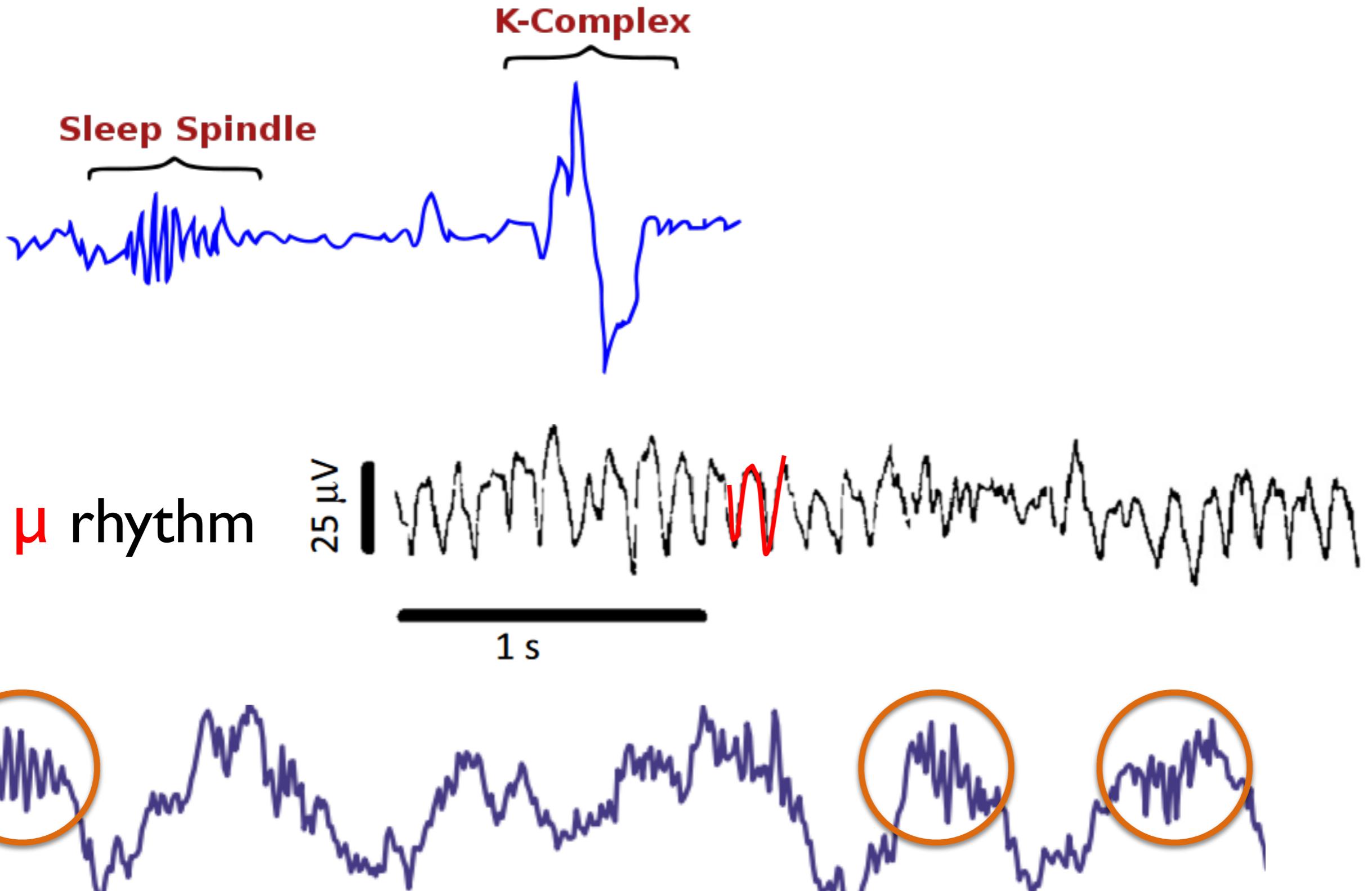




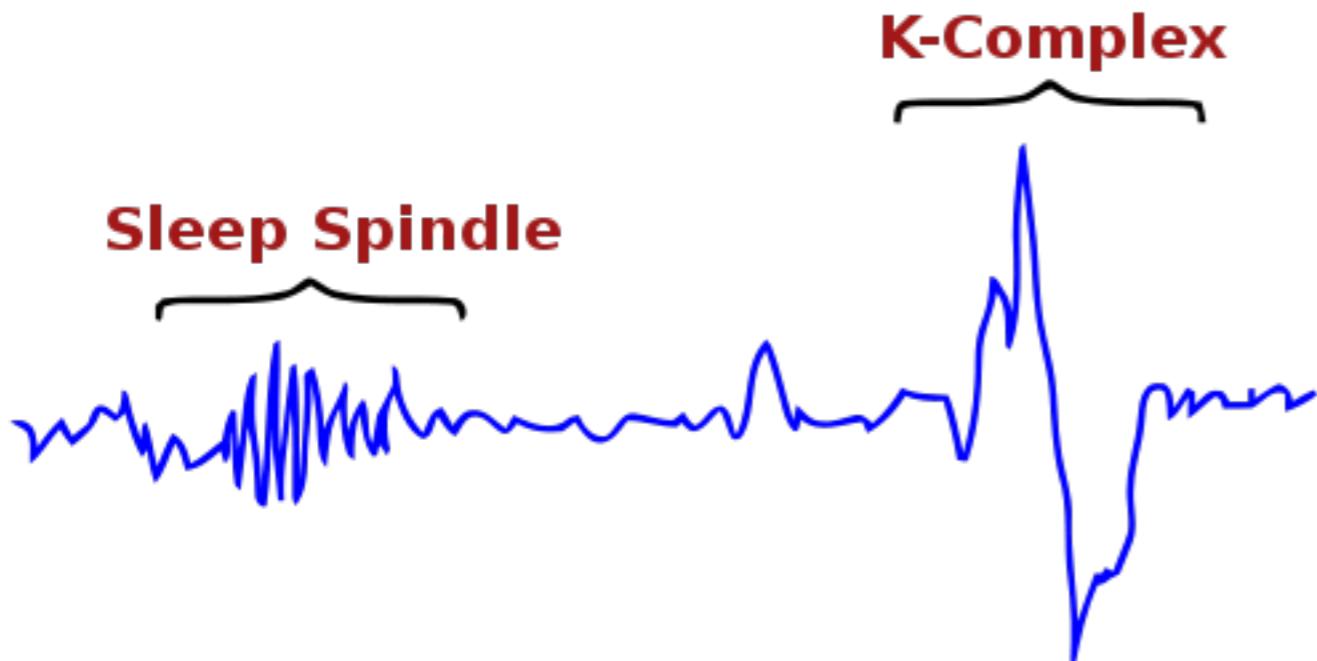




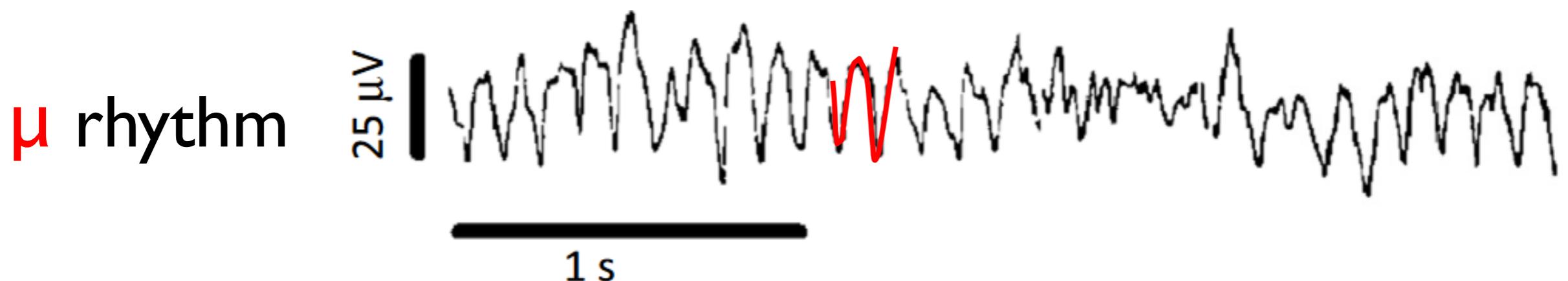
[T. Dupré la Tour, L. Tallot, L. Grabot, V. Doyère, V. van Wassenhove, Y. Grenier, A. Gramfort,
(2017) PLOS Computational biology]



CFC: High frequency bursts coupled with slow waves

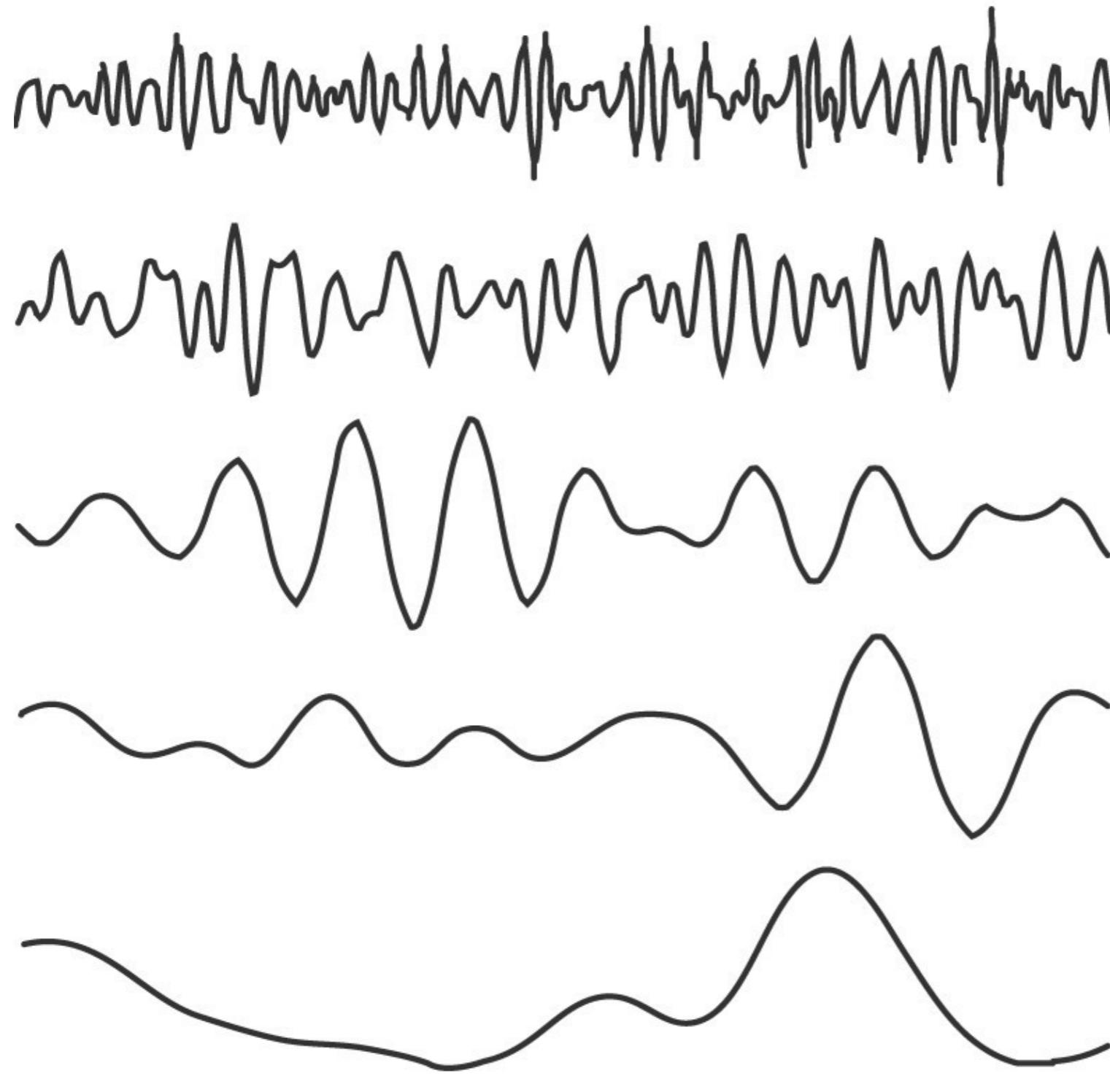


Neural signals exhibit diverse and complex morphologies



CFC: High frequency bursts coupled with slow waves

“Bandology”



Gamma
(> 25 Hz)

Beta
(12-25 Hz)

Alpha
(8-12 Hz)

Theta
(4-8 Hz)

Delta
(1-4 Hz)

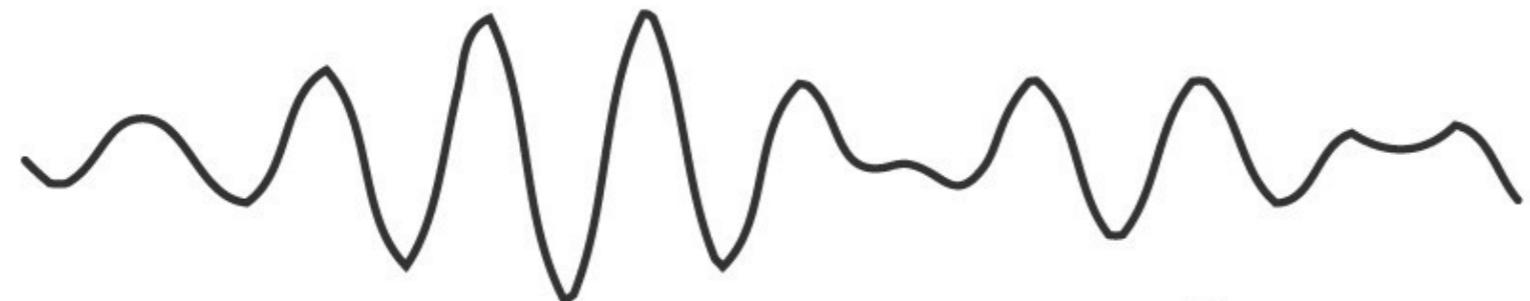
“Bandology”



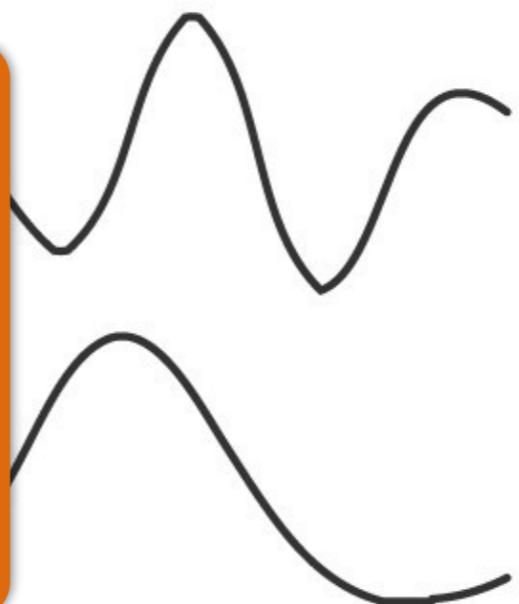
Gamma
(> 25 Hz)



Beta
(12-25 Hz)



Alpha
(8-12 Hz)



Theta
(4-8 Hz)



Delta
(1-4 Hz)

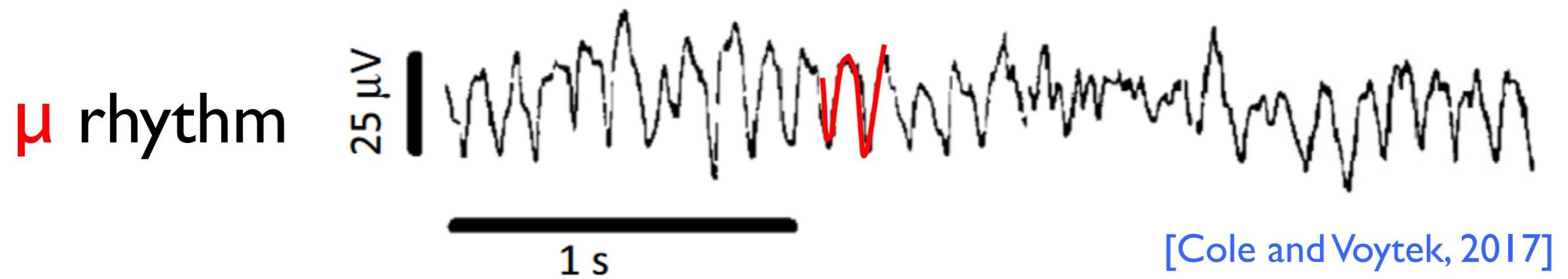
There is a lot more in
EEG/MEG signals
than we believe

Convolutional Sparse Coding (CSC) for learning the morphology of neural signals

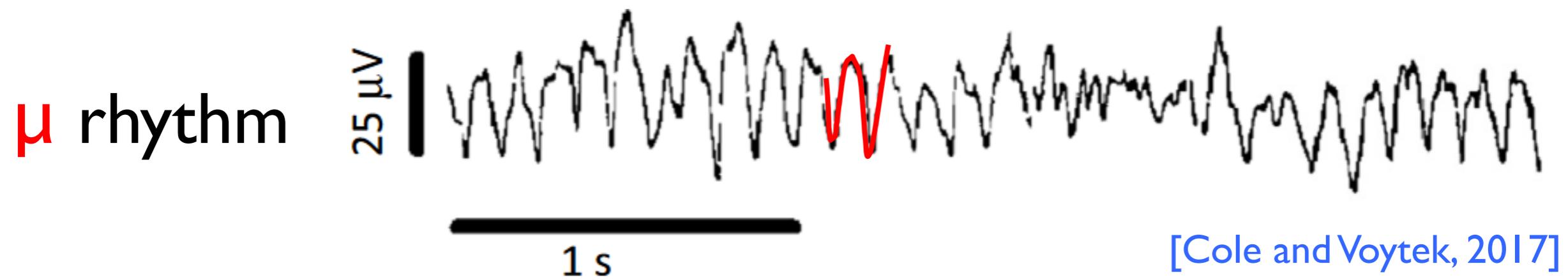


*Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals, (2018),
T. Dupré la Tour, T. Moreau, M. Jas, A. Gramfort, Proc. NIPS Conf.
Learning the Morphology of Brain Signals Using Alpha-Stable Convolutional Sparse Coding,
(2017), M. Jas, T. Dupré la Tour, U. Simsekli, A. Gramfort, Proc. NIPS Conf.*

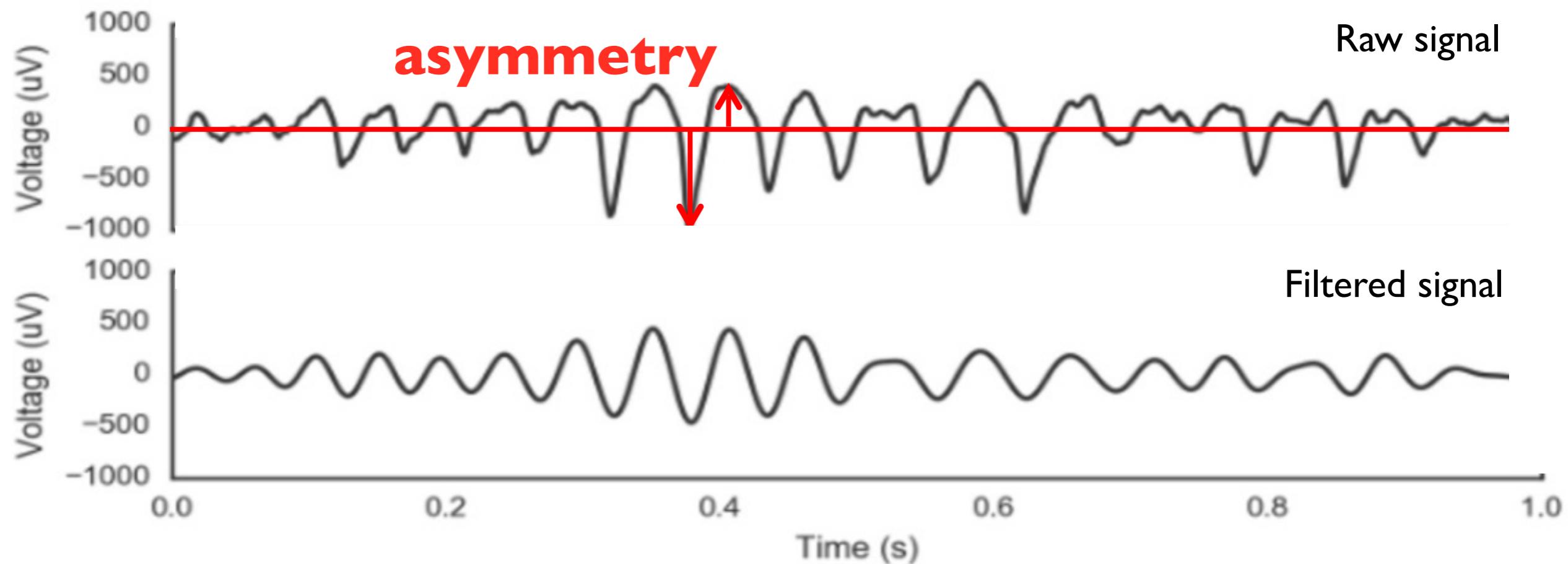
Shape of brain rhythms matter



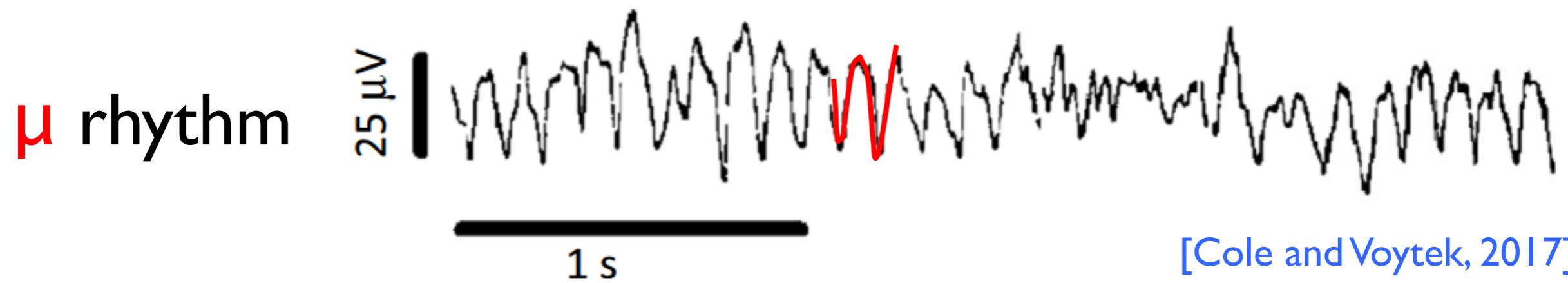
Shape of brain rhythms matter



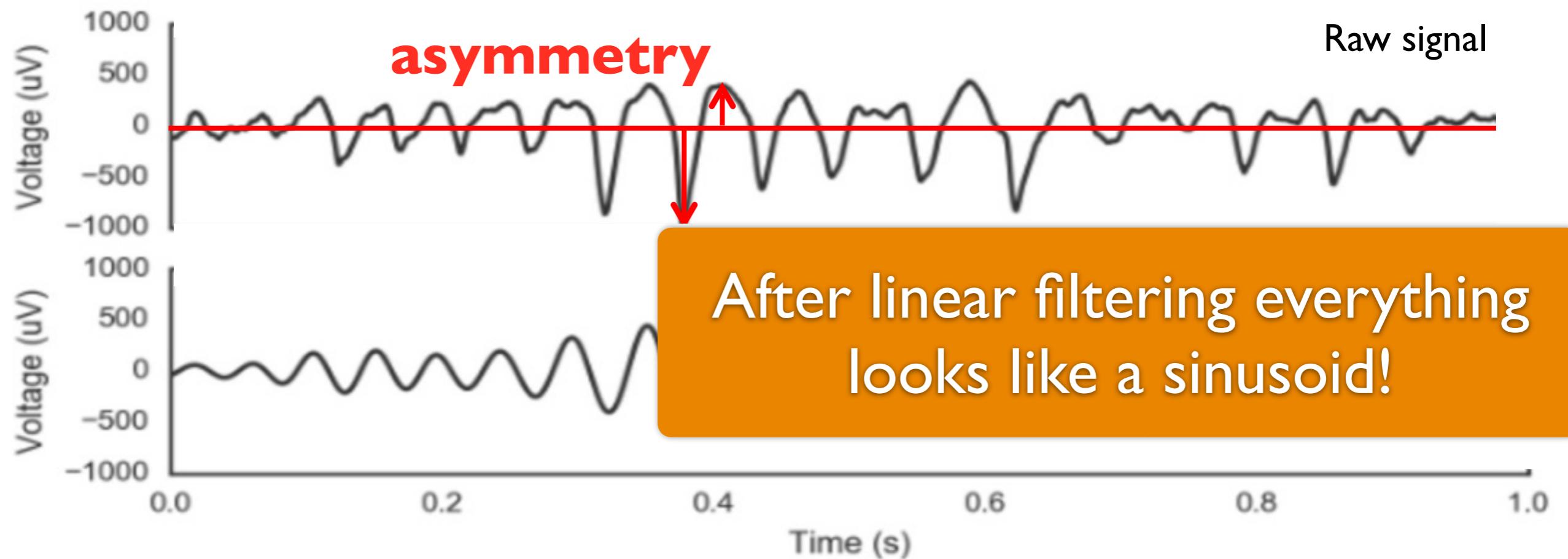
Problem of linear filtering:



Shape of brain rhythms matter



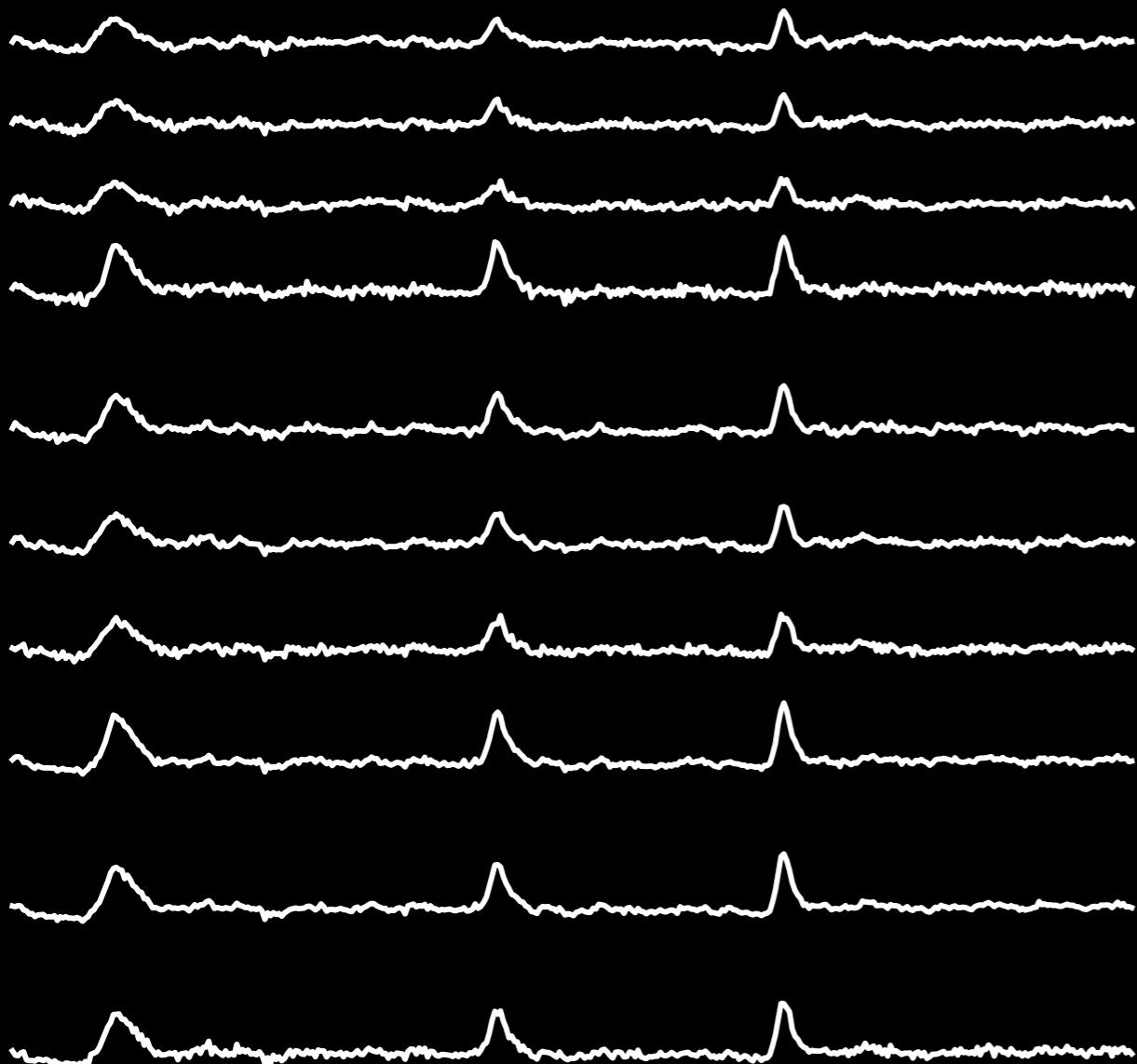
Problem of linear filtering:



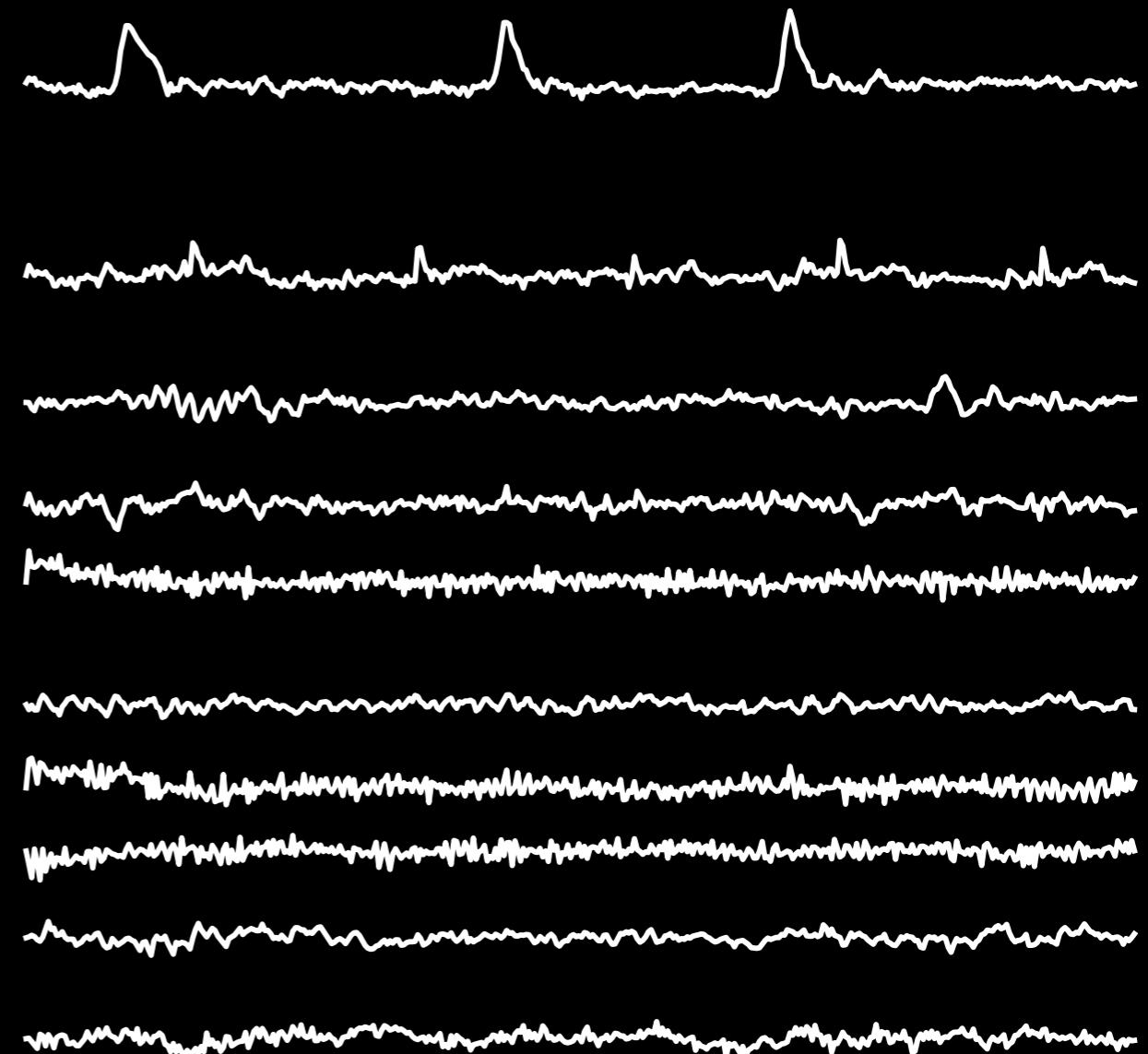
From ICA to CSC

Independent Component Analysis (ICA)

Observations (raw EEG)



ICA recovered sources

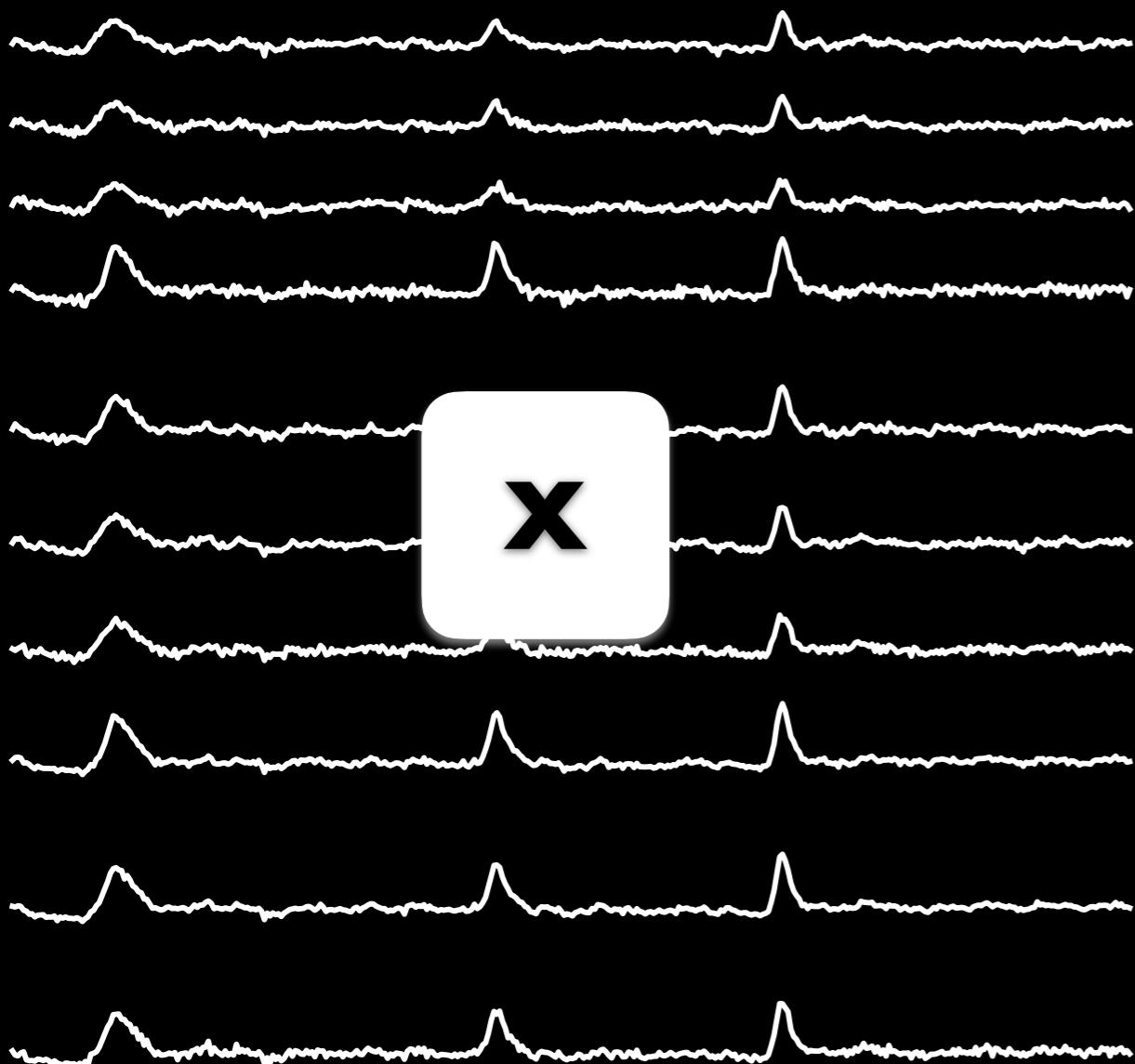


<https://pypi.python.org/pypi/python-picard/0.1>

From ICA to CSC

Independent Component Analysis (ICA)

Observations (raw EEG)



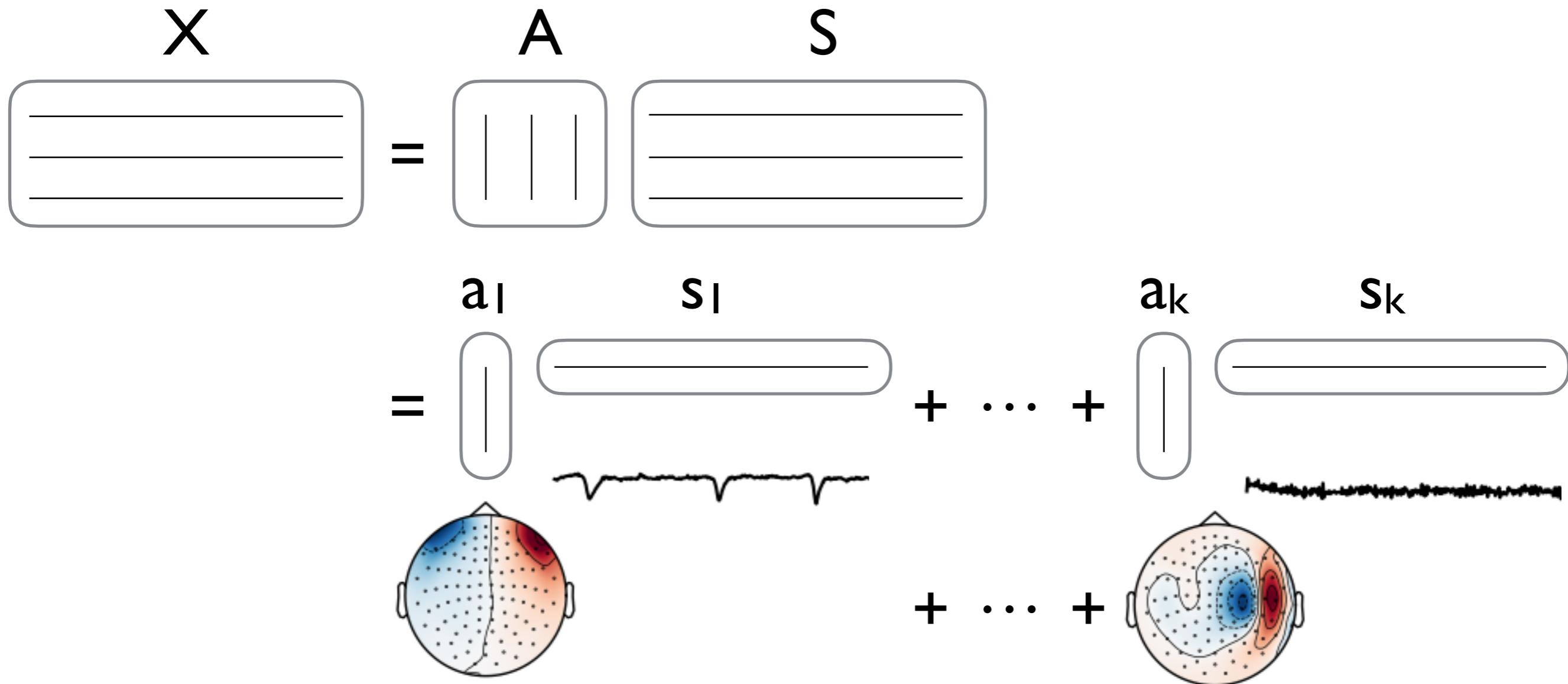
X

ICA recovered sources



<https://pypi.python.org/pypi/python-picard/0.1>

From ICA...



https://mne.tools/stable/auto_tutorials/preprocessing/plot_40_artifact_correction_ica.html

https://pierreablin.github.io/picard/auto_examples/plot_ica_eeg.html

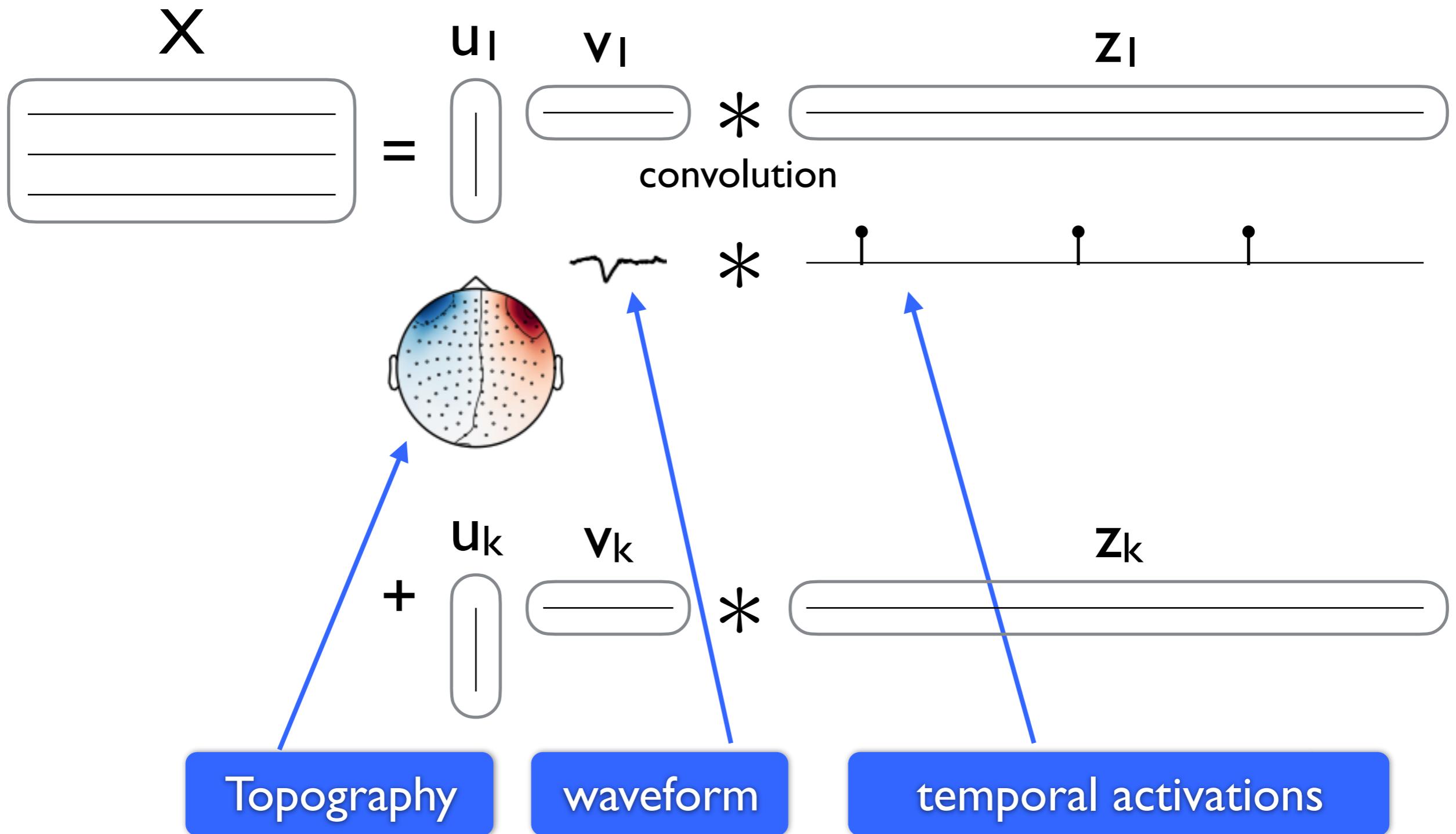
... to CSC

$$X = u_1 v_1 * z_1$$

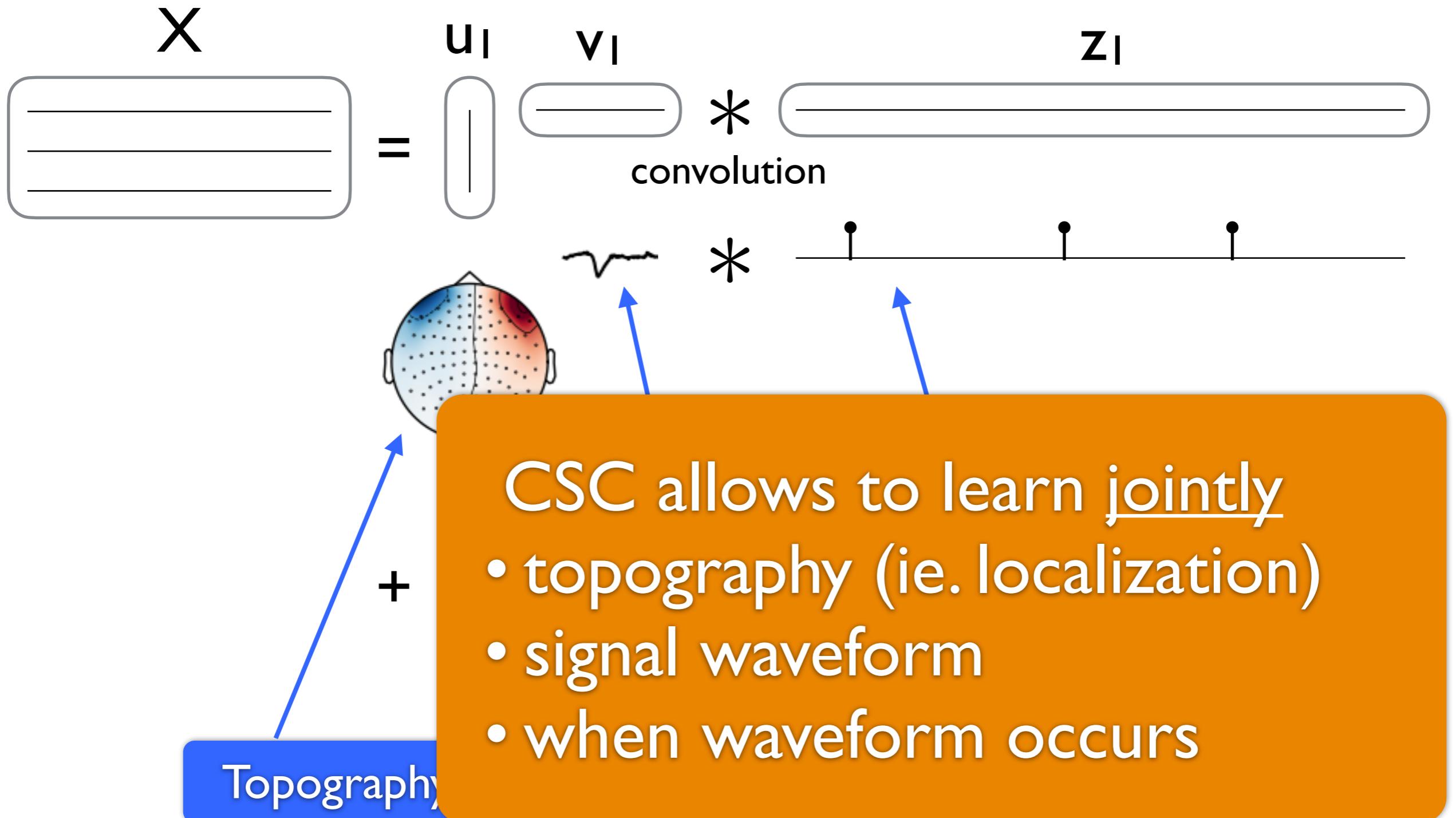
+ \dots

$$+ u_k v_k * z_k$$

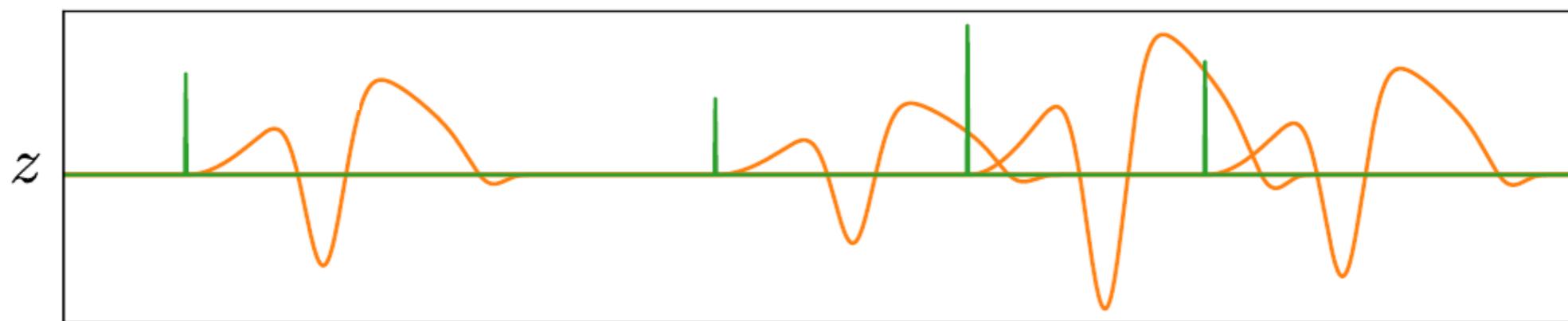
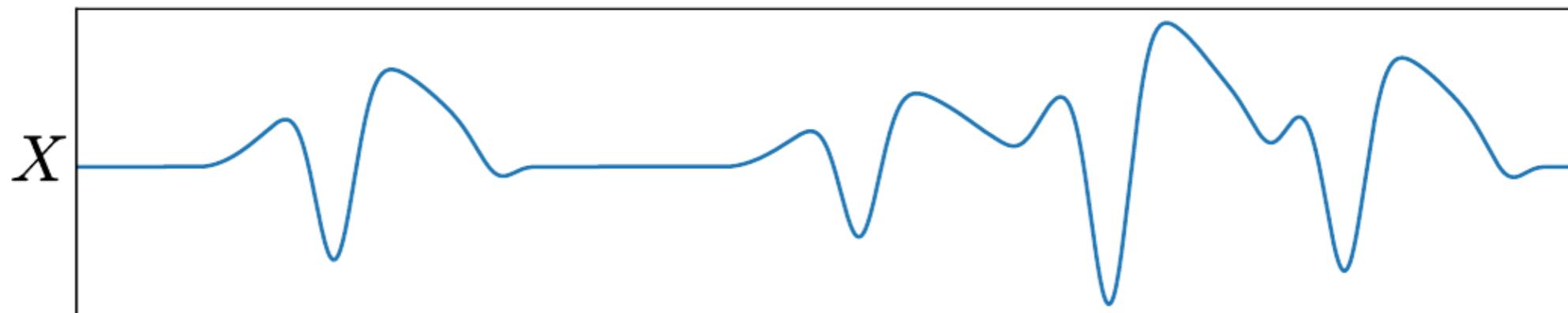
... to CSC



... to CSC



Multivariate CSC

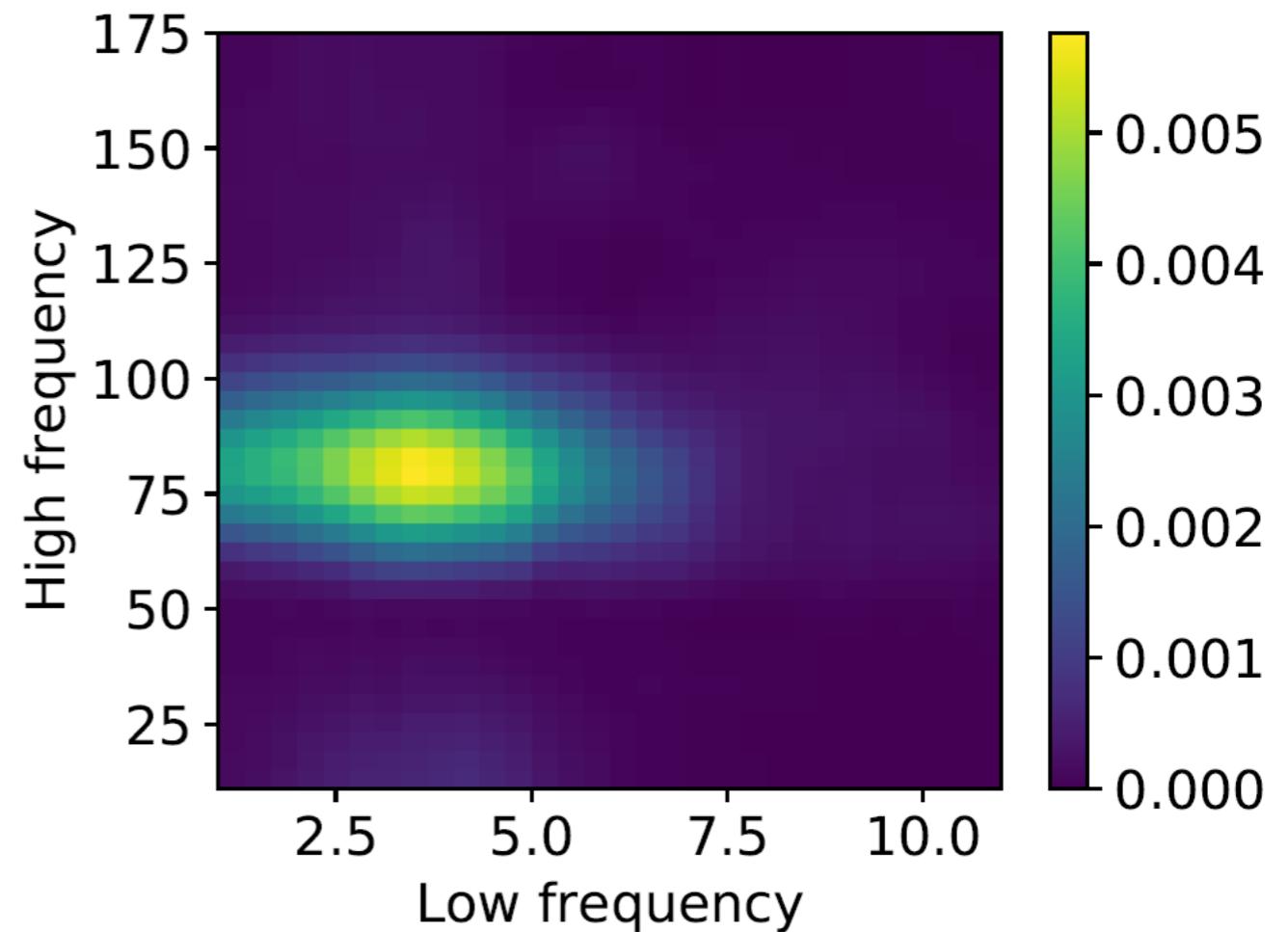
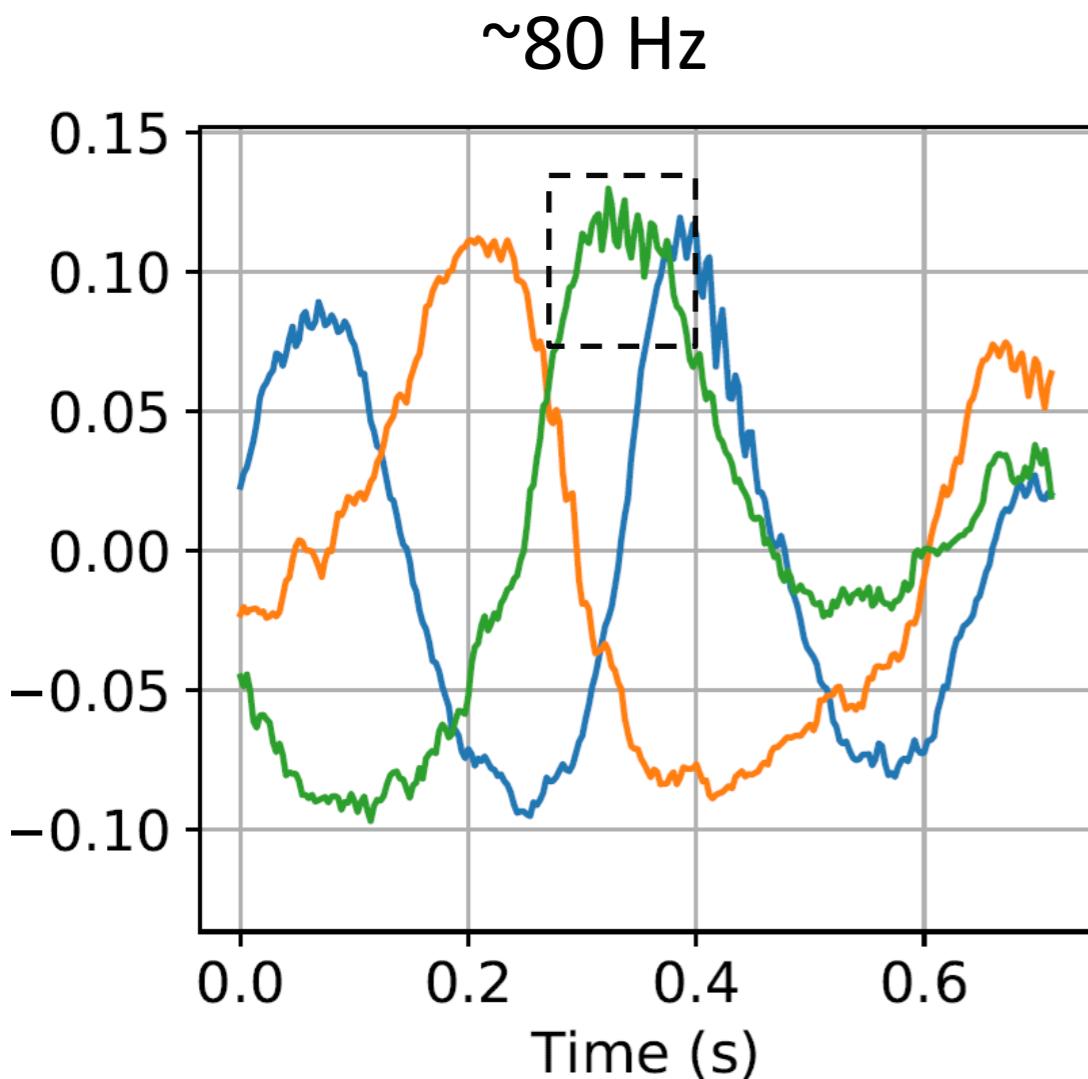


$$\min_{u,v,z} \sum_{n=1}^N \frac{1}{2} \left\| \boxed{X^n} - \sum_{k=1}^K \boxed{z_k^n} * \boxed{(u_k v_k^\top)} \right\|_2^2 + \lambda \sum_{k=1}^K \|z_k^n\|_1,$$

s.t. $\|u_k\|_2^2 \leq 1$, $\|v_k\|_2^2 \leq 1$ and $z_k^n \geq 0$.

[*Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals, (2018), T. Dupré la Tour, T. Moreau, M. Jas, A. Gramfort, Proc. NeurIPS Conf.*]

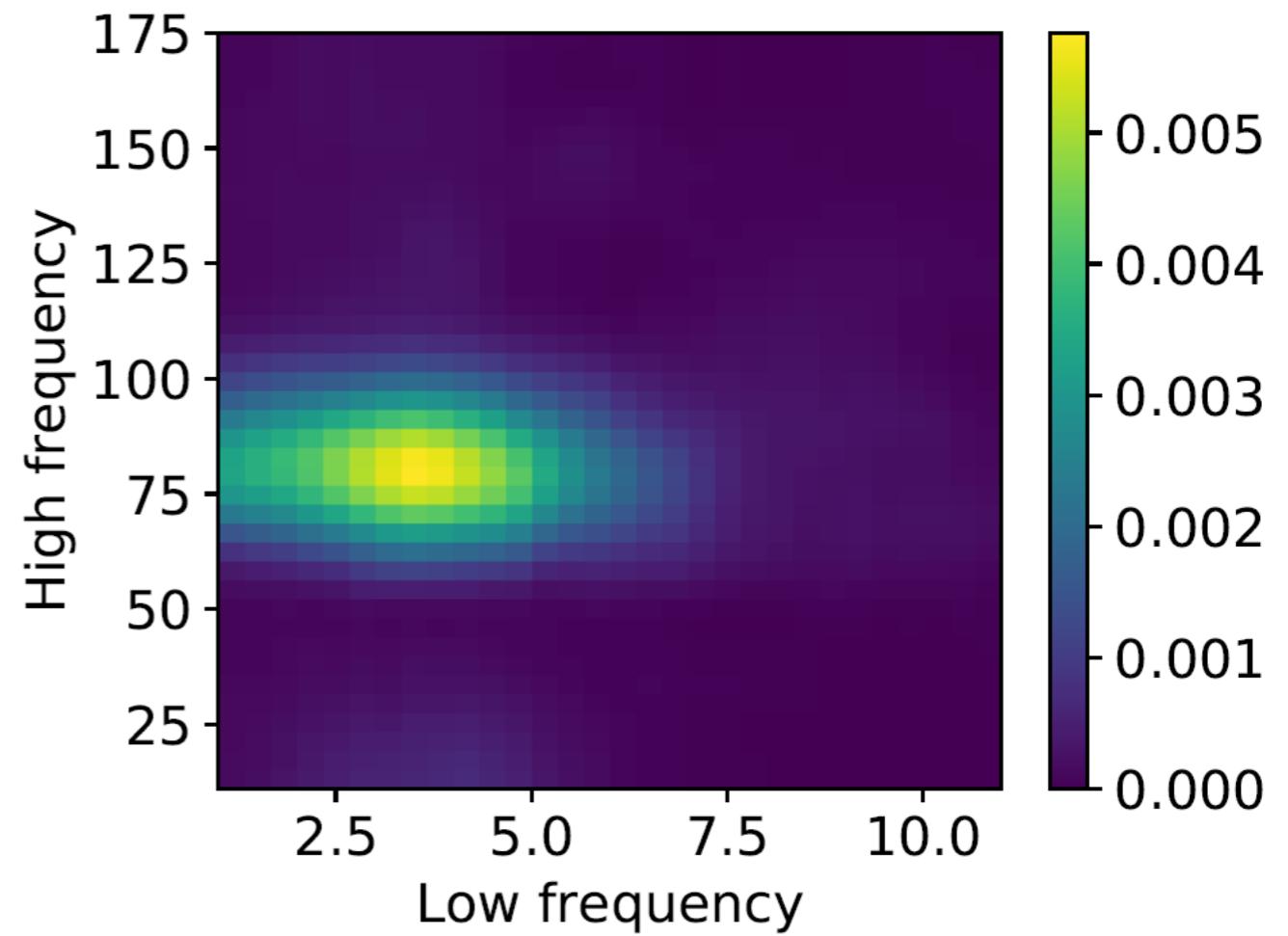
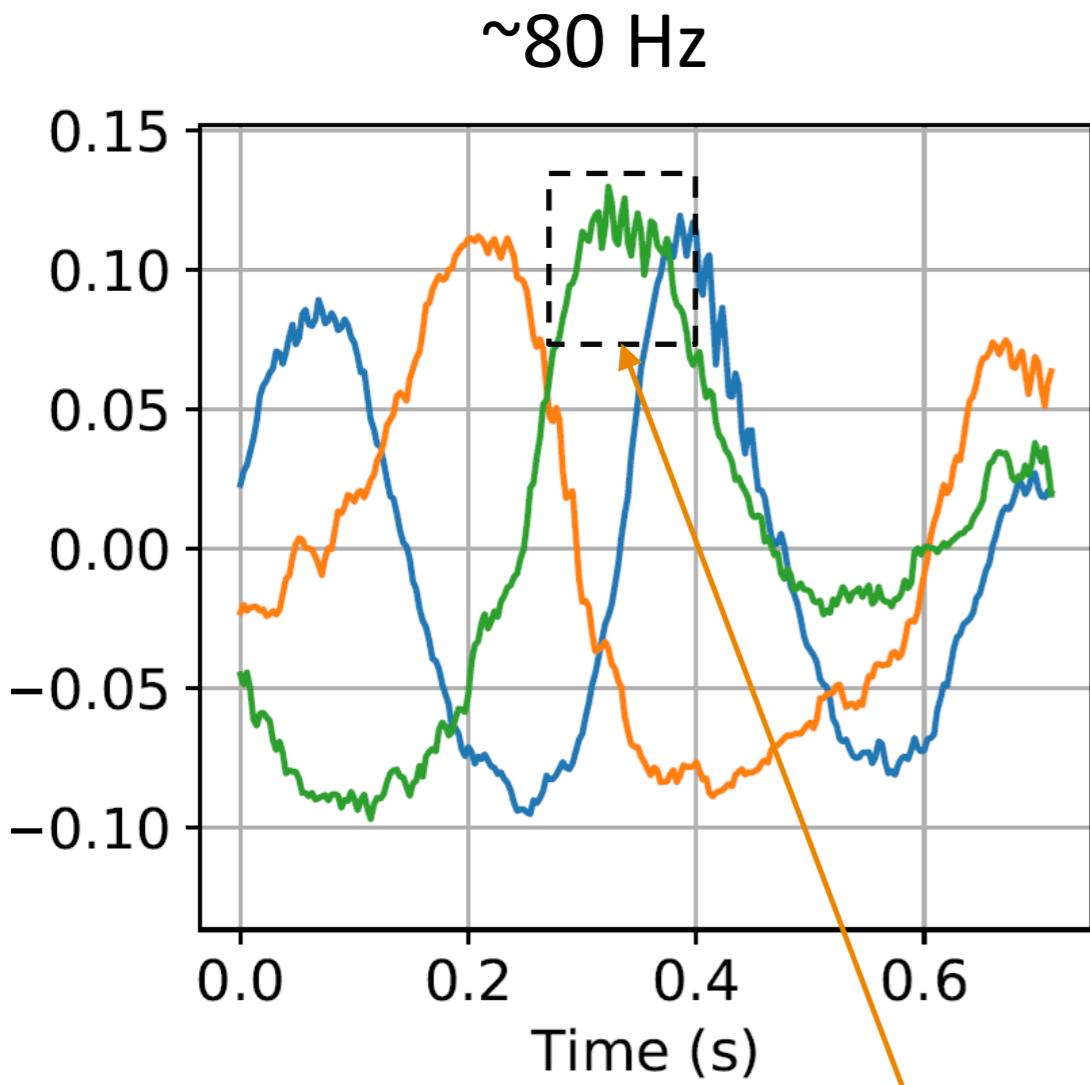
CSC on LFP



Signal from the striatum of a rodent

[*Learning the Morphology of Brain Signals Using Alpha-Stable Convolutional Sparse Coding*,
(2017), M. Jas, T. Dupré la Tour, U. Simsekli, A. Gramfort, Proc. NIPS Conf.]

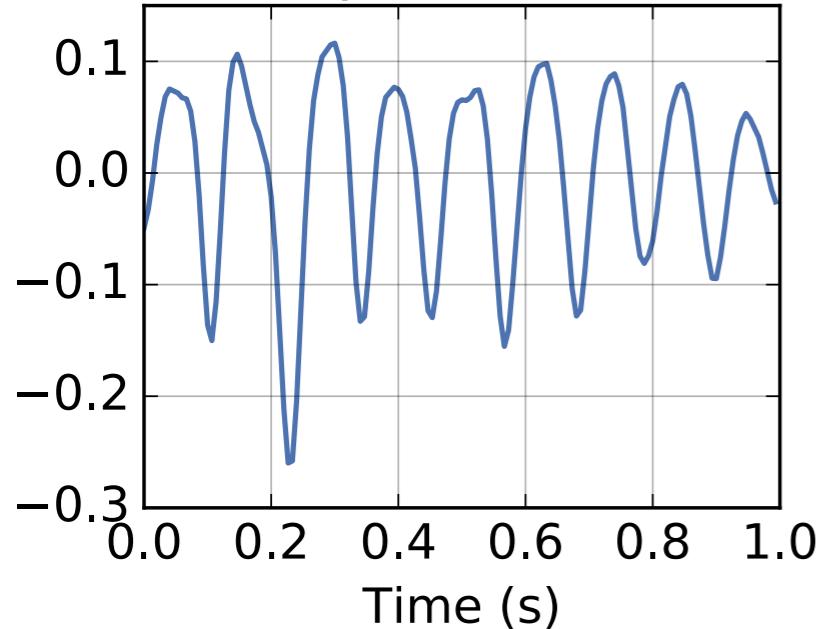
CSC on LFP



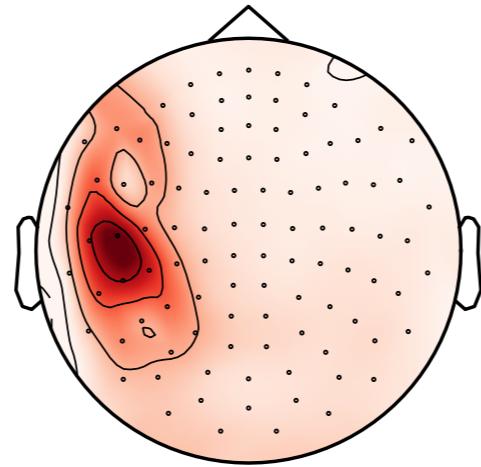
[Learning the Morphology of Brain Signals Using Alpha-Stable Convolutional Sparse Coding, (2017), M. Jas, T. Dupré la Tour, U. Simsekli, A. Gramfort, Proc. NIPS Conf.]

CSC on MEG

A. Temporal waveform

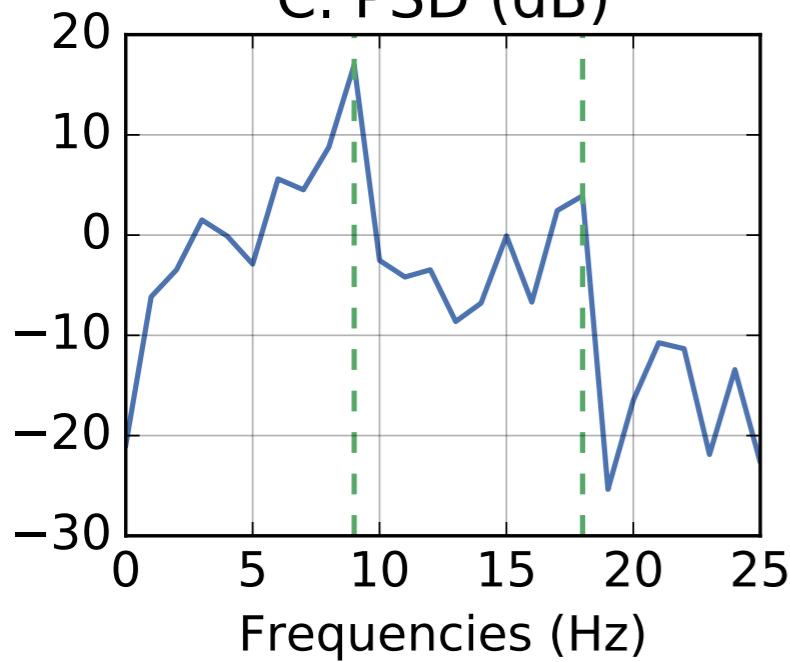


B. Spatial pattern

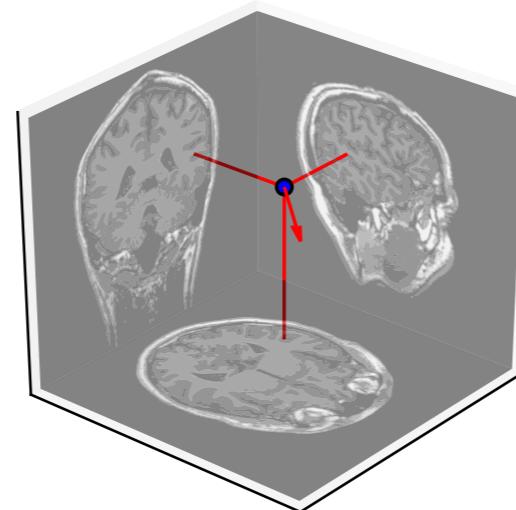


- MEG vectorview
- Median nerve stim.

C. PSD (dB)



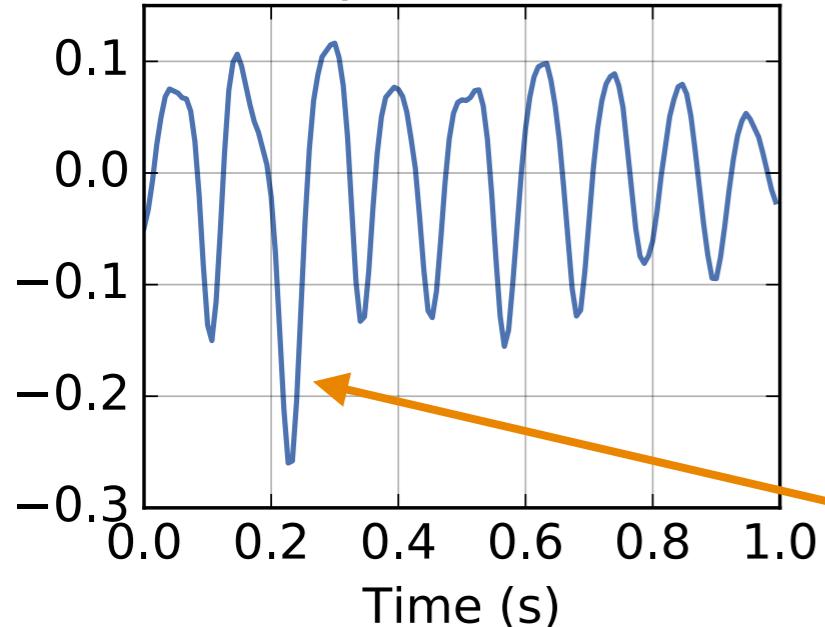
D. Dipole fit



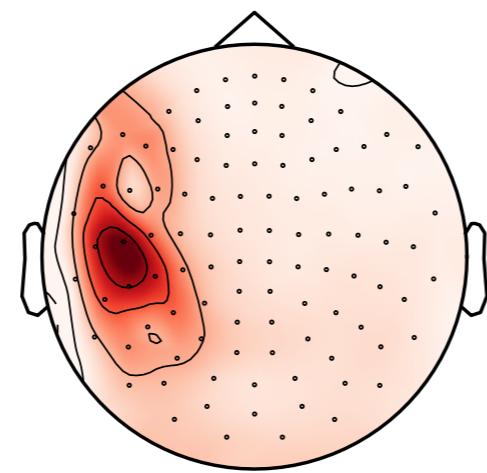
[*Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals, (2018),
T. Dupré la Tour, T. Moreau, M. Jas, A. Gramfort, Proc. NIPS Conf.*]

CSC on MEG

A. Temporal waveform

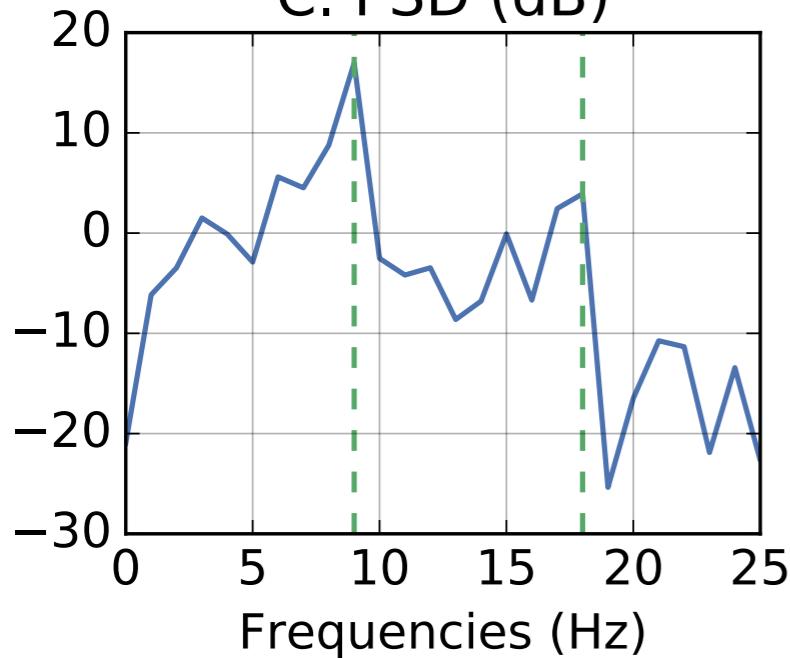


B. Spatial pattern

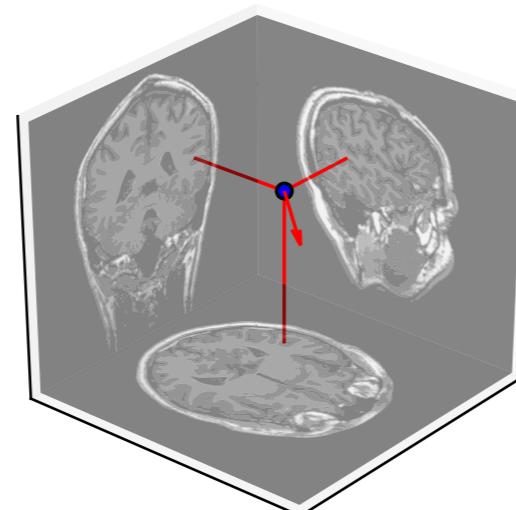


- MEG vectorview
- Median nerve stim.

C. PSD (dB)



D. Dipole fit

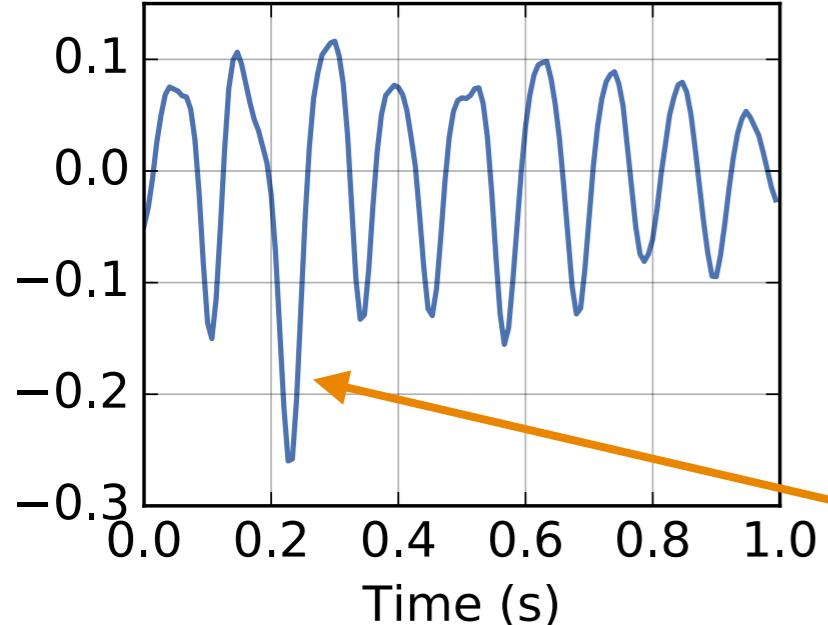


CSC reveals mu-shaped waveforms

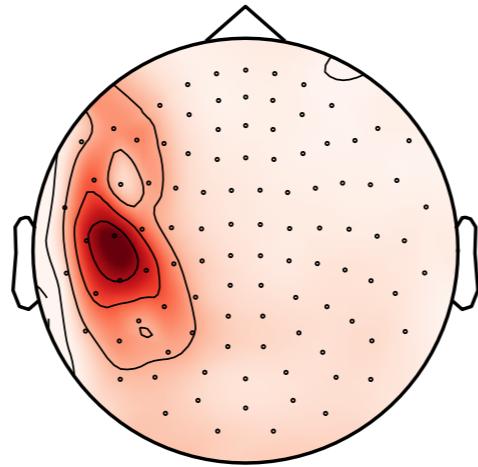
[*Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals, (2018), T. Dupré la Tour, T. Moreau, M. Jas, A. Gramfort, Proc. NIPS Conf.*]

CSC on MEG

A. Temporal waveform

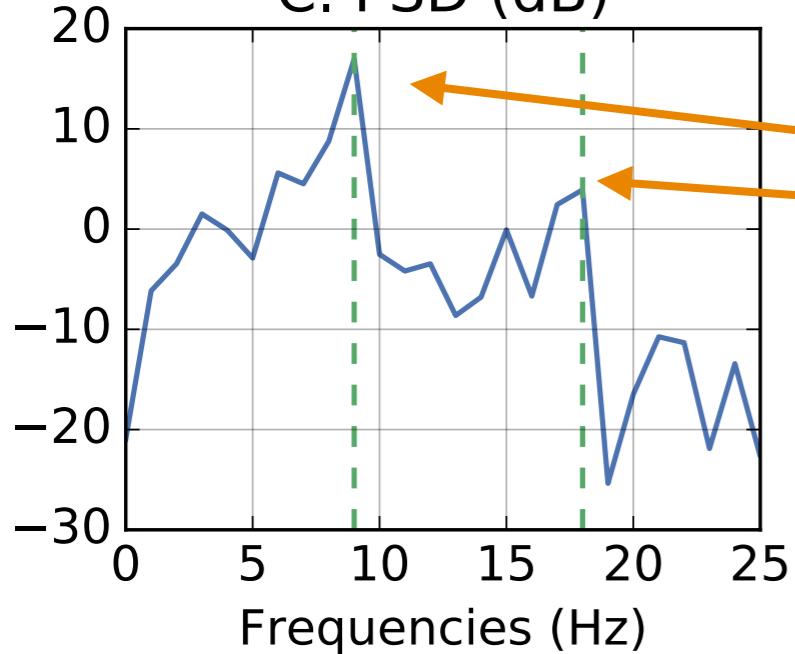


B. Spatial pattern

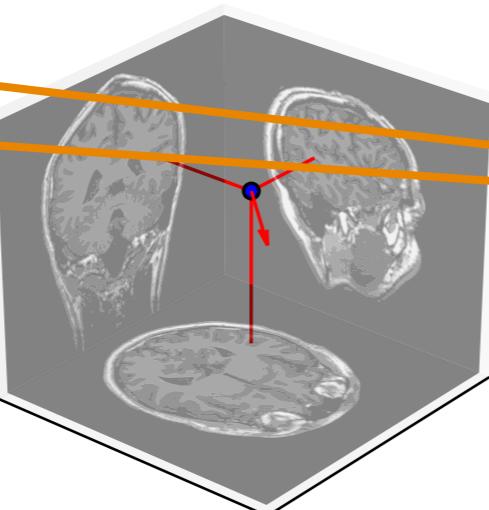


- MEG vectorview
- Median nerve stim.

C. PSD (dB)



D. Dipole fit



CSC reveals mu-shaped waveforms

See the frequency harmonics

[*Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals, (2018), T. Dupré la Tour, T. Moreau, M. Jas, A. Gramfort, Proc. NIPS Conf.*]

alphaCSC: Convolution sparse coding for time-series

build passing codecov 81%

This is a library to perform shift-invariant [sparse dictionary learning](#), also known as convolutional sparse coding (CSC), on time-series data. It includes a number of different models:

1. univariate CSC
2. multivariate CSC
3. multivariate CSC with a rank-1 constraint [\[1\]](#)
4. univariate CSC with an alpha-stable distribution [\[2\]](#)

A mathematical descriptions of these models is available [in the documentation](#).

Installation

To install this package, the easiest way is using [pip](#). It will install this package and its dependencies. The [setup.py](#) depends on [numpy](#) and [cython](#) for the installation so it is advised to install them beforehand. To install this package, please run

```
pip install numpy cython
pip install git+https://github.com/alphacsc/alphacsc.git#egg=alphacsc
```

If you do not have admin privileges on the computer, use the [--user](#) flag with [pip](#). To upgrade, use the [--upgrade](#) flag provided by [pip](#).

To check if everything worked fine, you can run:

```
python -c 'import alphacsc'
```

and it should not give any error messages.

Quickstart

Here is an example to present briefly the API:

<https://alphacsc.github.io>

```
import numpy as np
```

Extracting artifact and evoked response atoms from the sample dataset

This example illustrates how to learn rank1 atoms on the sample dataset from `mne`. We display a selection of atoms, featuring heartbeat and eyeblink artifacts, three atoms of evoked responses, and a non-sinusoidal oscillation.

```
# Authors: Thomas Moreau <thomas.moreau@inria.fr>
#          Mainak Jas <mainak.jas@telecom-paristech.fr>
#          Tom Dupre La Tour <tom.duprelatour@telecom-paristech.fr>
#          Alexandre Gramfort <alexandre.gramfort@telecom-paristech.fr>
#
# License: BSD (3-clause)
```

Let us first define the parameters of our model.

```
# sample frequency
sfreq = 150.

# Define the shape of the dictionary
n_atoms = 40
n_times_atom = int(round(sfreq * 1.0)) # 1000. ms

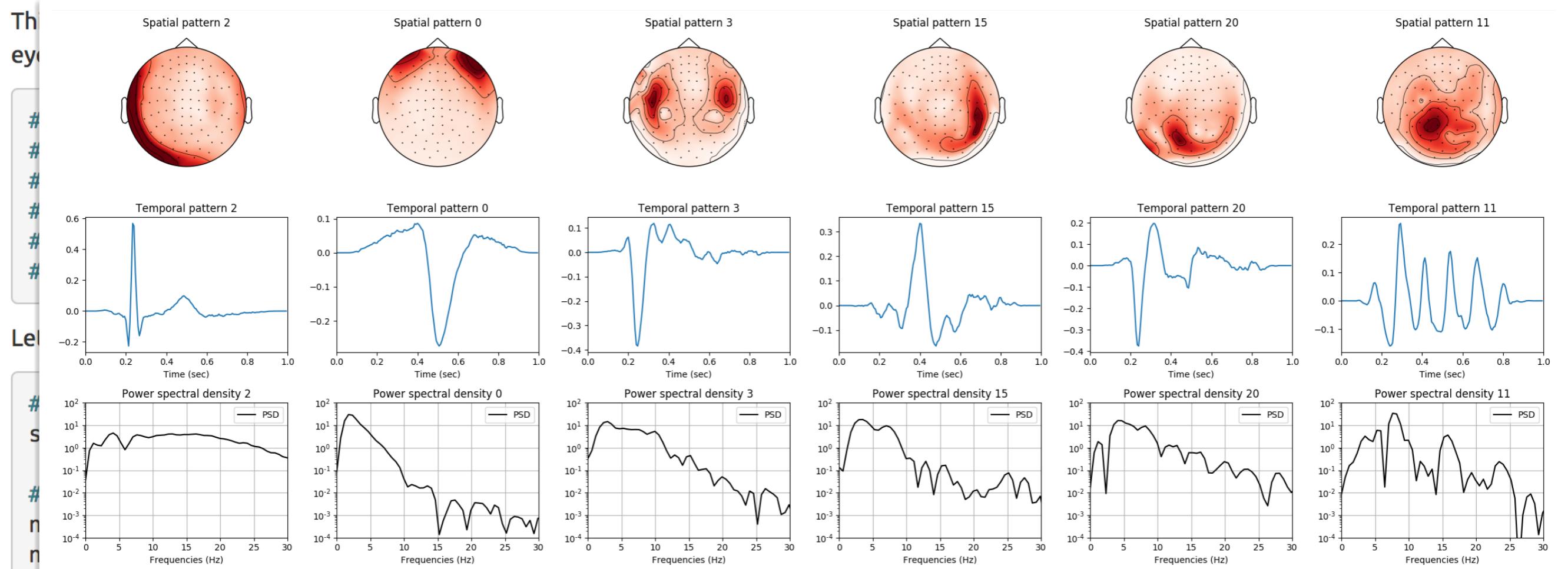
# Regularization parameter which control sparsity
reg = 0.1

# number of processors for parallel computing
n_jobs = 5
```

Next, we define the parameters for multivariate CSC

https://alphacsc.github.io/auto_examples/multicsc/plot_sample_evoked_response.html

Extracting artifact and evoked response atoms from the sample dataset



```
# Regularization parameter which control sparsity
```

```
reg = 0.1
```

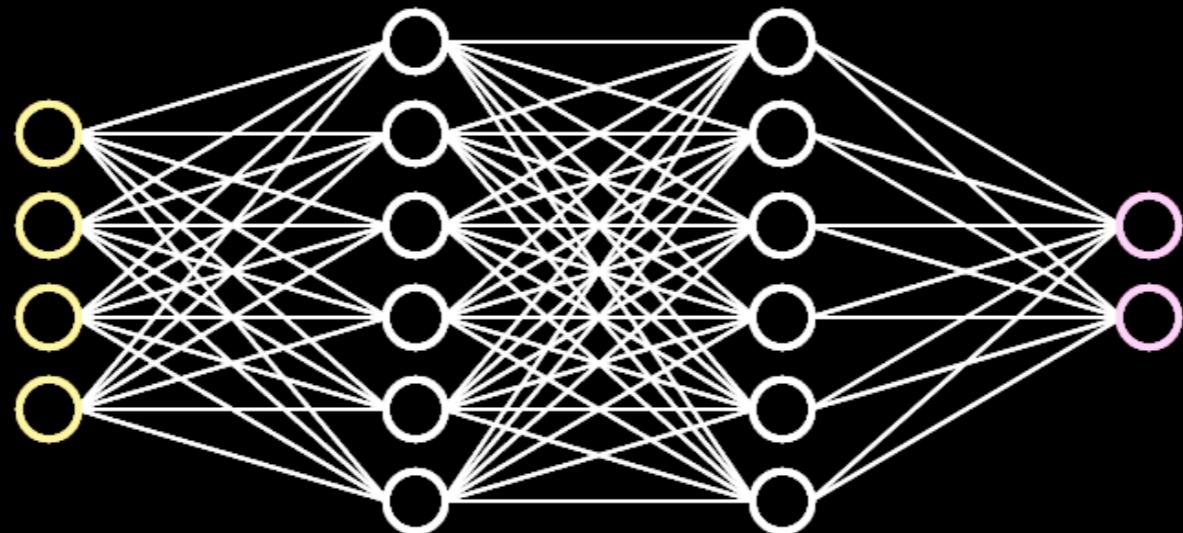
```
# number of processors for parallel computing
```

```
n_jobs = 5
```

Next, we define the parameters for multivariate CSC

https://alphacsc.github.io/auto_examples/multicsc/plot_sample_evoked_response.html

Self-supervised learning on EEG

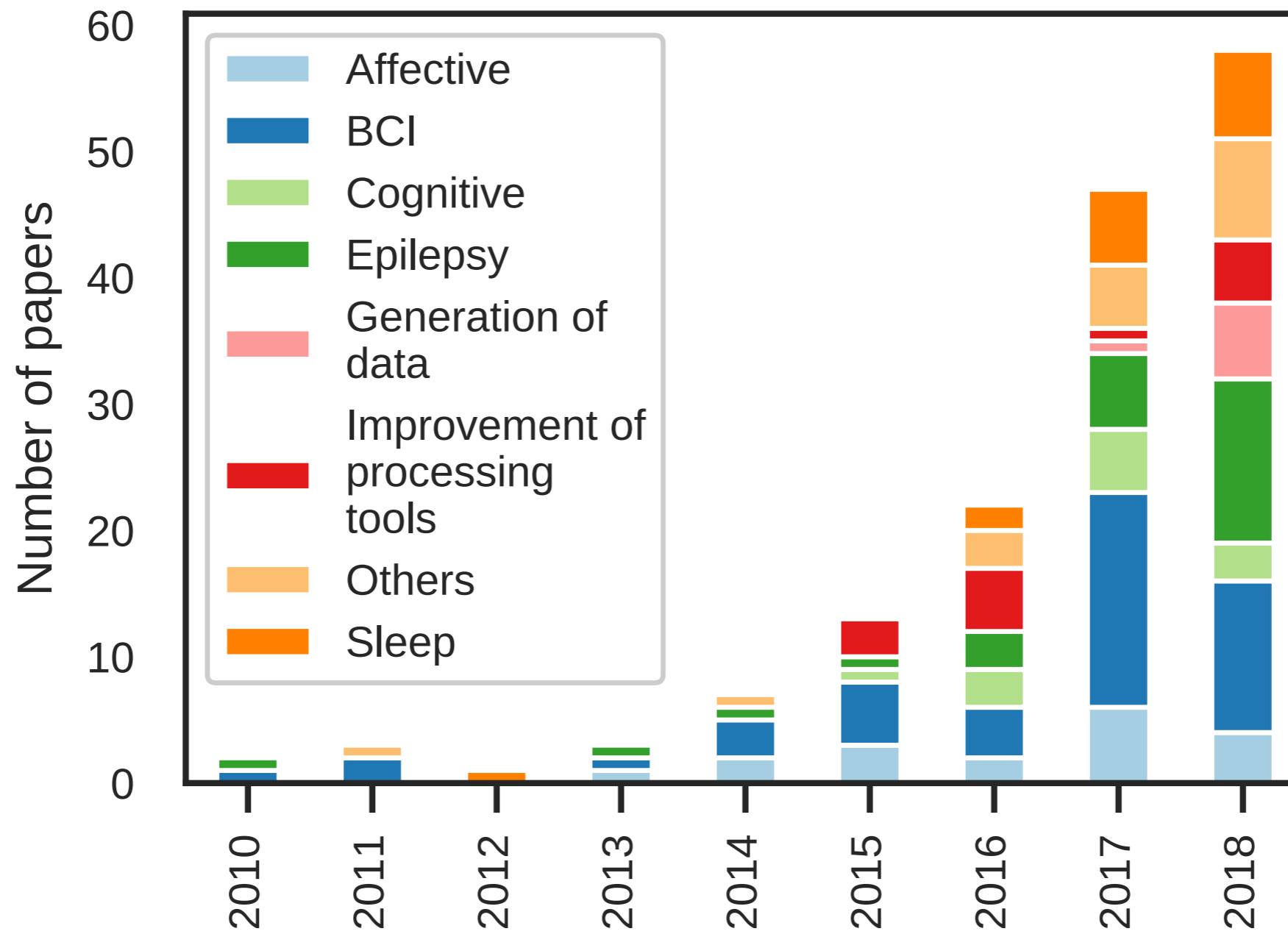


Uncovering the structure of clinical EEG signals with self-supervised learning
Banville, H., Chehab, O., Hyvärinen, A., Engemann, D. and Gramfort, A. (2020)

ArXiv abs/2007.16104

Self-supervised representation learning from electroencephalography signals
Banville, H., Albuquerque, I., Moffat, G., Engemann, D. and Gramfort, A. (2019)
Proc. Machine Learning for Signal Processing (MLSP) .

Deep Learning papers on EEG



[Deep learning-based electroencephalography analysis: a systematic review
Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. and Faubert, J. (2019)
Journal of Neural Engineering 16: (051001).]

Challenges of ML on EEG

- Low signal-to-noise ratio (SNR)
- Strong variability (intra/inter-subjects)
- Expensive/Time-consuming to collect & annotate

Challenges of ML on EEG

- Low signal-to-noise ratio (SNR)
- Strong variability (intra/inter-subjects)
- Expensive/Time-consuming to collect & annotate

Example: Sleep recording (polysomnography)

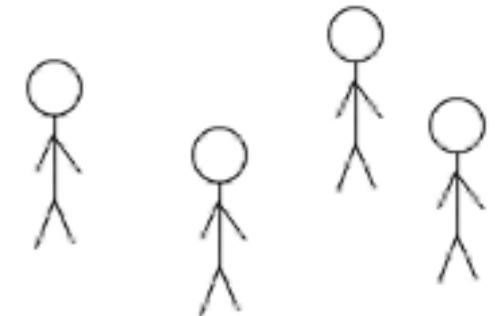
Requires controlled environment (e.g., clinic)



Need experts (e.g., sleep technologists) to annotate the data manually



Variability between subjects and for the same subject from day-to-day



Challenges of ML on EEG

- Low signal-to-noise ratio (SNR)
- Strong variability (intra/inter-subjects)
- Expensive/Time-consuming to collect & annotate

Example: Sleep recording (polysomnography)

Requires controlled environment (e.g., clinic)



Need experts (e.g., sleep technologists) to annotate the data manually



Question:
Can we
reduce our
dependence
on labels?

Self-supervision



[Noroozi & Favaro 2016] use a deep neural network to solve the Jigsaw puzzle

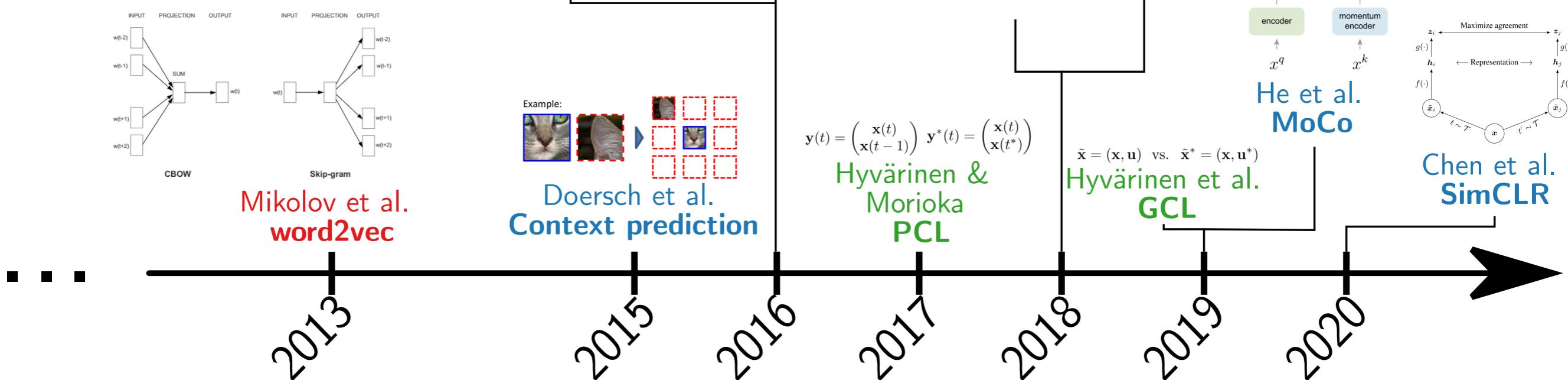


Use the **structure** of the data to pretrain a **feature extractor** with a supervised *pretext task* – then use the features on a *downstream task*.

(Partial) History of SSL beyond EEG

Application domains

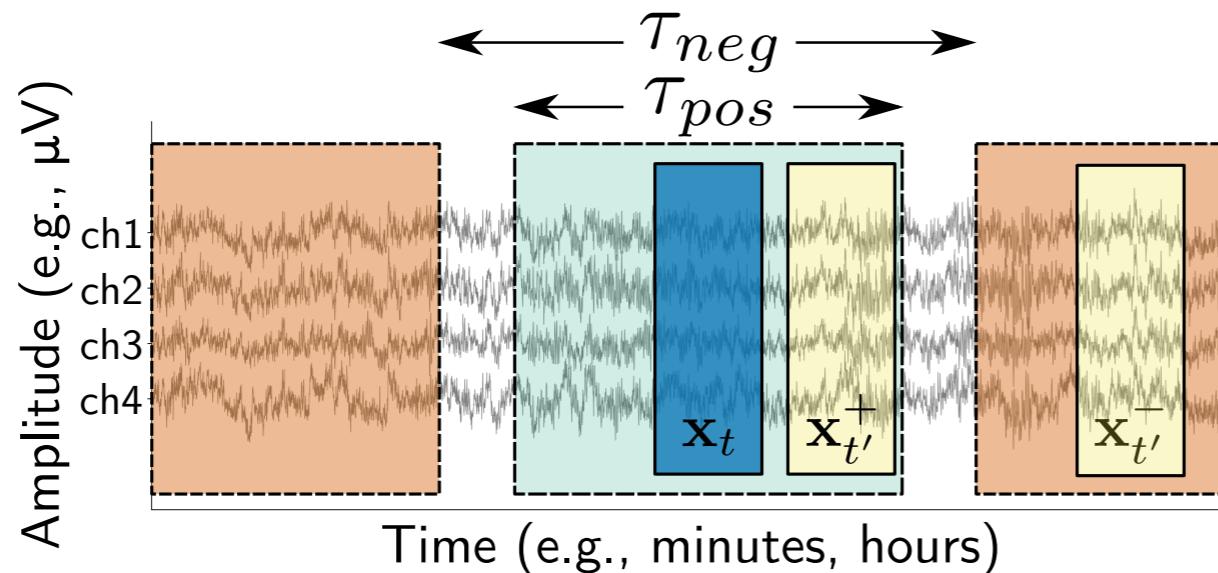
- Computer vision
- NLP
- Nonlinear ICA



Pretext Task

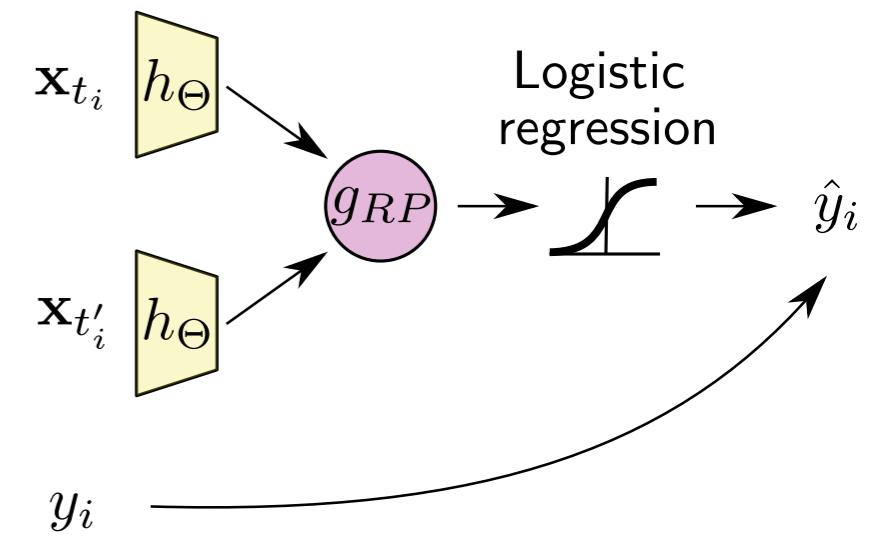
Relative positioning (RP)

1 Sampling



$$y_i = \begin{cases} 1, & \text{if } |t_i - t'_i| \leq \tau_{pos} \\ -1, & \text{if } |t_i - t'_i| > \tau_{neg} \end{cases}$$

2 Training



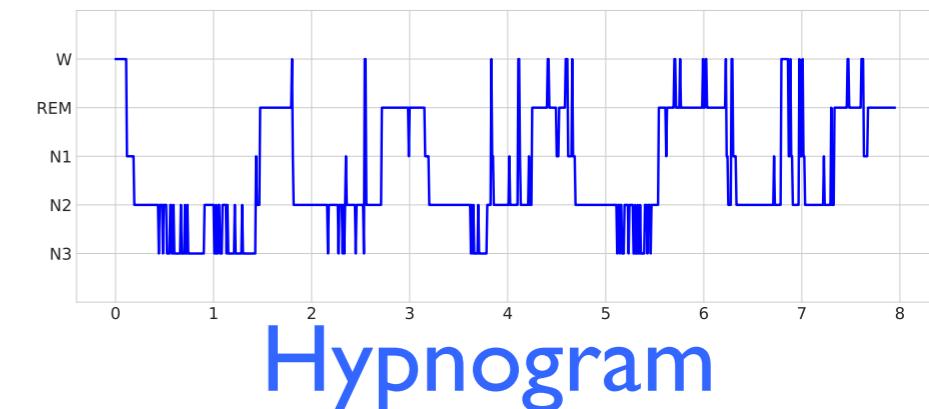
Predict if 2 windows of data are close in time

Other approaches: CPC [Oord et al. 2018],
PCL [Hyvärinen et al. 2017] etc.

Downstream tasks on clinical EEG

Sleep staging:

Predict sleep stage from EEG
(5-class: W, N1, N2, N3, R)



Pathology detection:

Is someone's EEG pathological?
(2-class: normal, abnormal)



Datasets

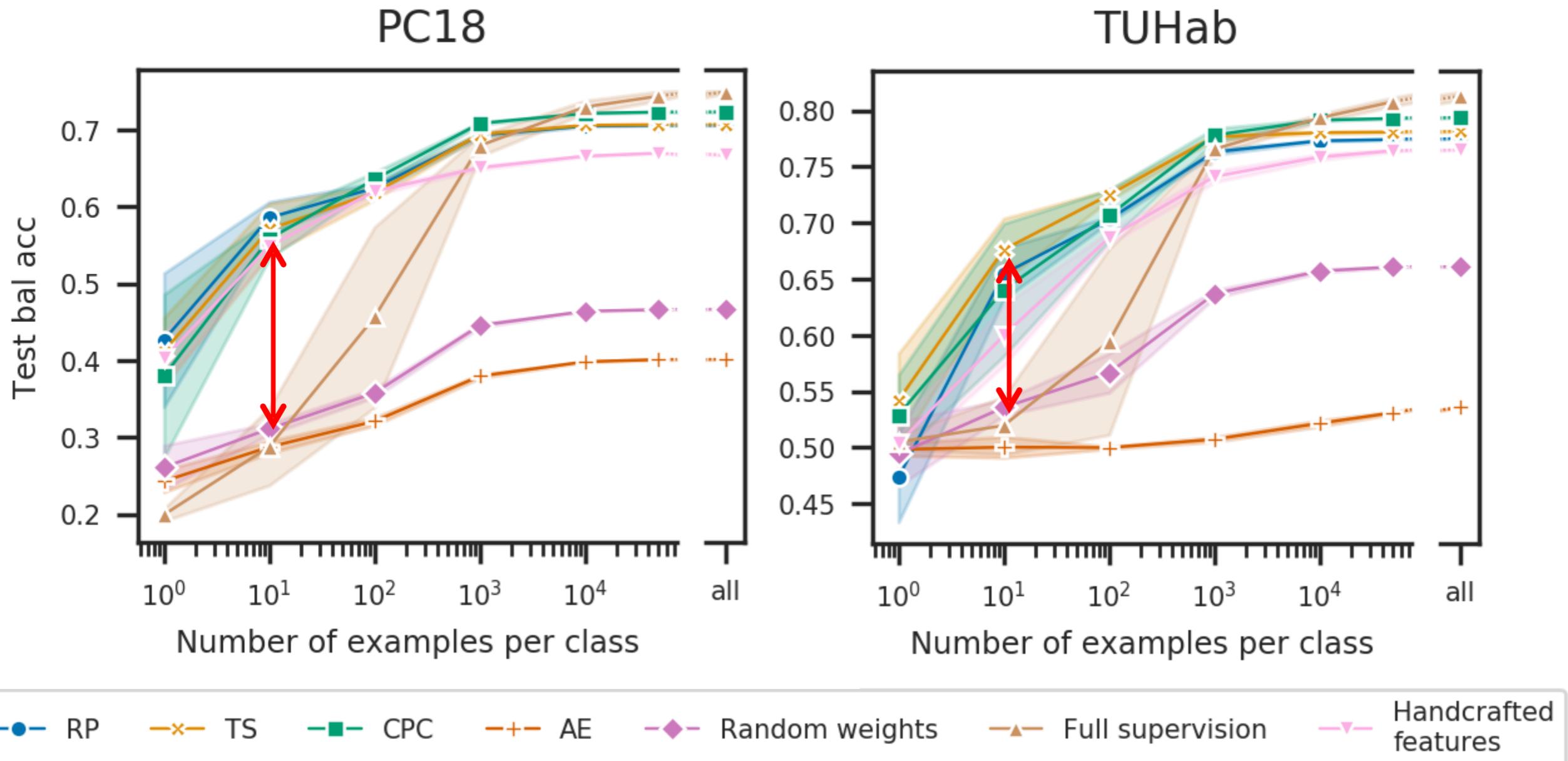
Two public datasets with hundreds of recordings:

- **Sleep staging:** Physionet Challenge 2018 (PC18) [Ghassemi et al. 2018]
- **Pathology detection:** TUH Abnormal EEG (TUHab) [López 2017]

PC18 (train)	
	# windows
W	158,020
N1	136,858
N2	377,426
N3	102,492
R	116,872
Total	891,668
# unique subjects	994
# recordings	994
Sampling frequency	200 Hz
# EEG channels	6
Reference	M1 or M2

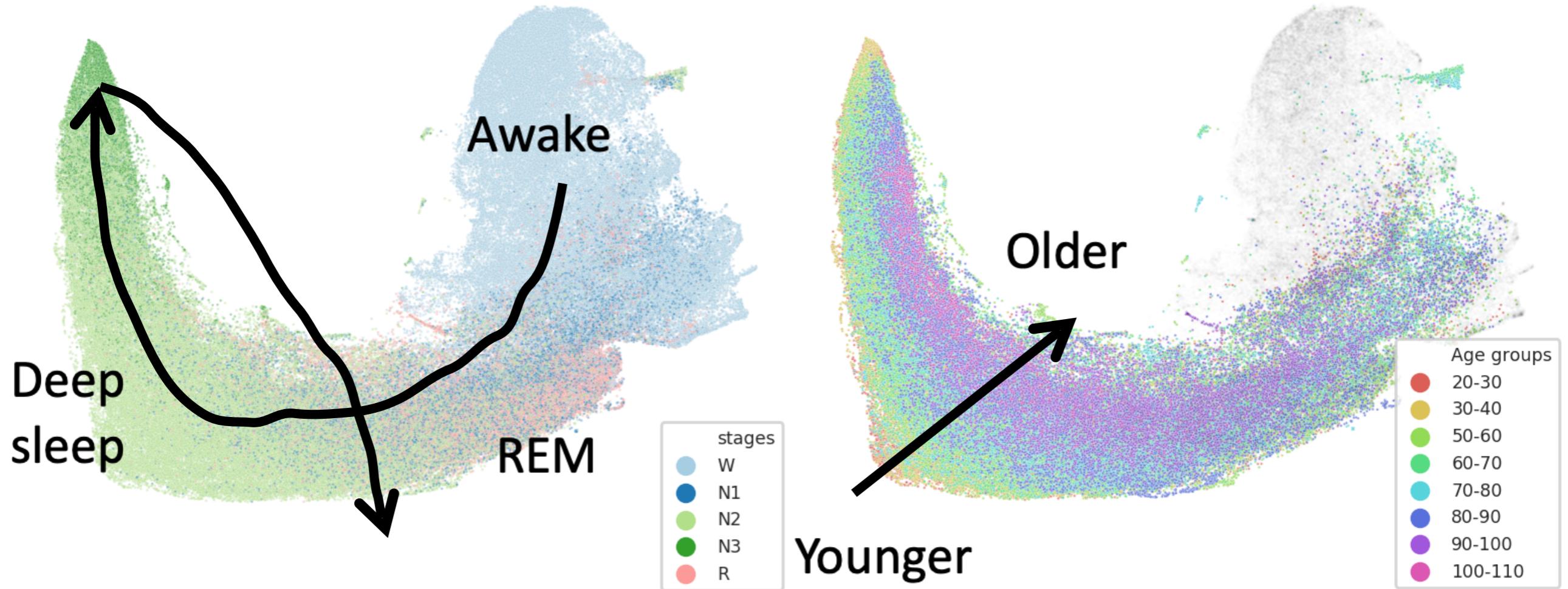
	TUHab (train)	TUHab (eval)
# recordings	1371	150
Normal	1346	126
Abnormal	2717	276
Total		
# unique subjects	2329	
# recordings	2993	
Sampling frequency	250, 256, 512 Hz	
# EEG channels	27 to 36	
Reference	Common average	

Results: Prediction accuracy



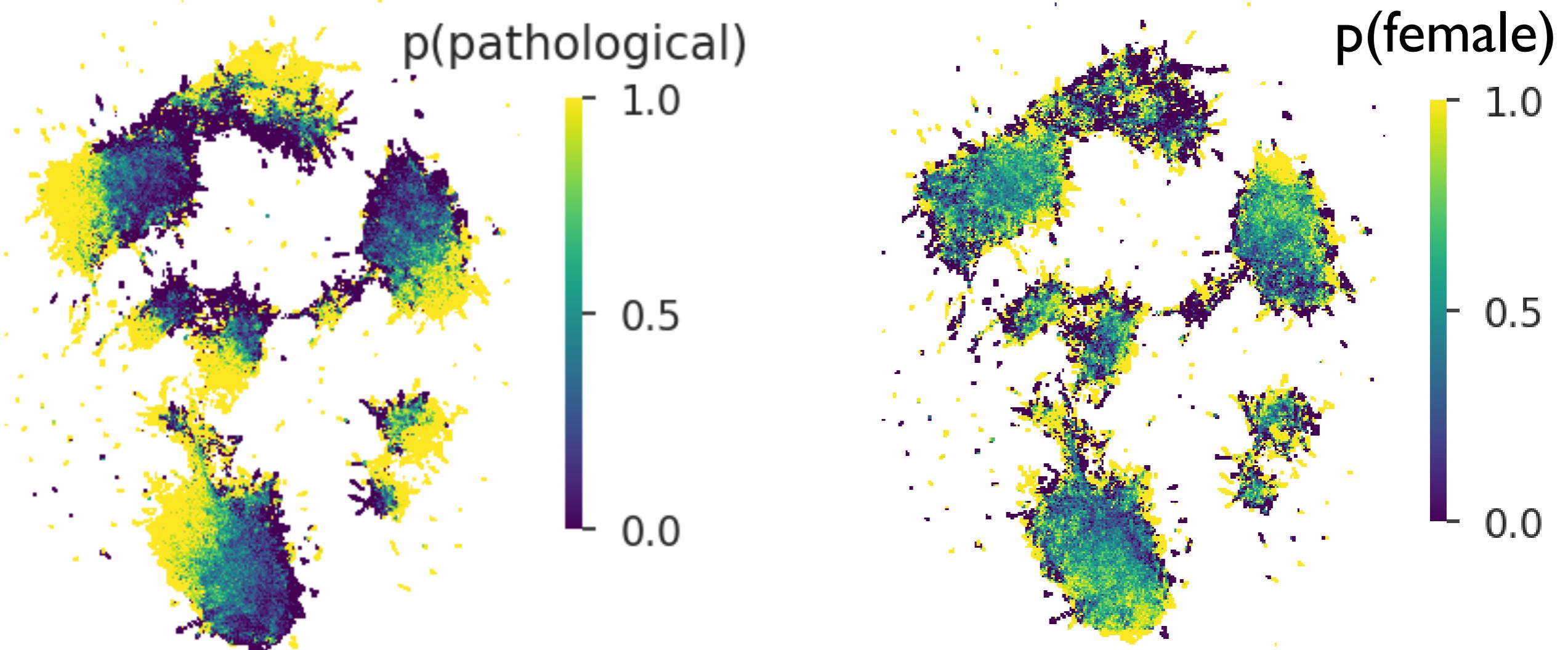
SSL is **better than full supervision** when limited data is available, and **competitive** when all data is available.

Results on sleep EEG



SSL can **uncover structure** without human supervision

Results on TUH data



SSL can **uncover clinically-relevant structure**
without human supervision

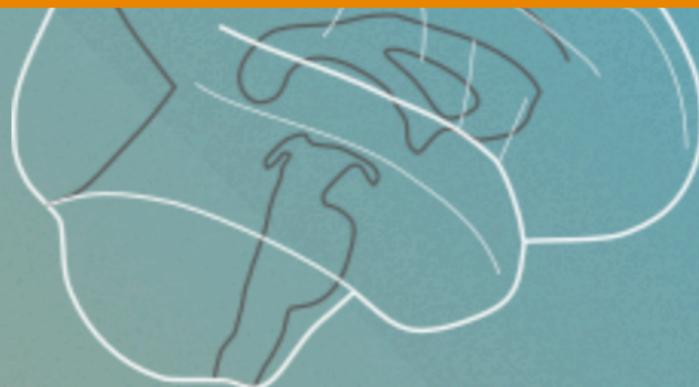
MEG and EEG for precision medicine and biomarkers

Manifold-regression to predict from MEG/EEG brain signals without source modeling
Sabbagh, D., Ablin, P., Varoquaux, G., Gramfort, A. and Engemann, D. (2019)
Advances in Neural Information Processing Systems (NeurIPS)

Predictive regression modeling with MEG/EEG: from source power to signals and cognitive states, Sabbagh, D., Ablin, P., Varoquaux, G., Gramfort, A. and Engemann, D. (2020), *NeuroImage*

Combining magnetoencephalography with magnetic resonance imaging enhances learning of surrogate-biomarkers, Engemann, D., Kozynets, O., Sabbagh, D., Lemaître, G., Varoquaux, G., Liem, F. and Gramfort, A. (2020), *eLife*

Precision medicine / Biomarkers



IMPAC

IMaging-PsychiAtry Challenge: predicting autism
A data challenge on Autism Spectrum Disorder detection

Deadline: July 1, 2018 - 8 pm (UTC)

https://paris-saclay-cds.github.io/autism_challenge/

Days

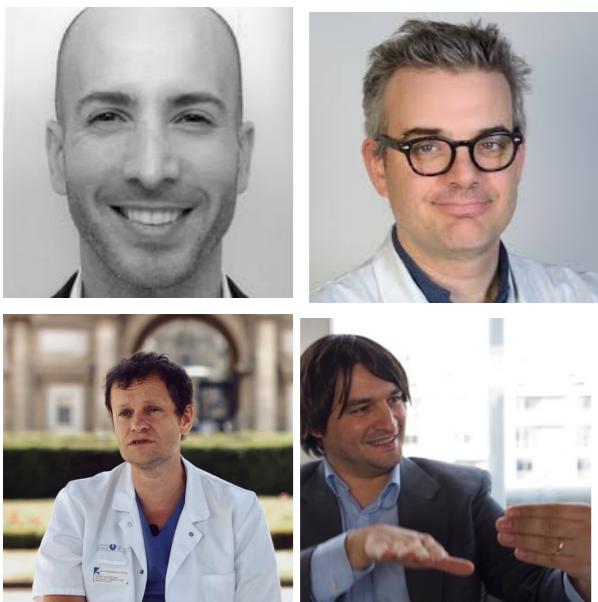
Hours

Minutes

Seconds

Predict of brain “fragility” for optimal drug dosage across age

Joint work with:



Institut national
de la santé et de la recherche médicale

Hôpitaux Universitaires

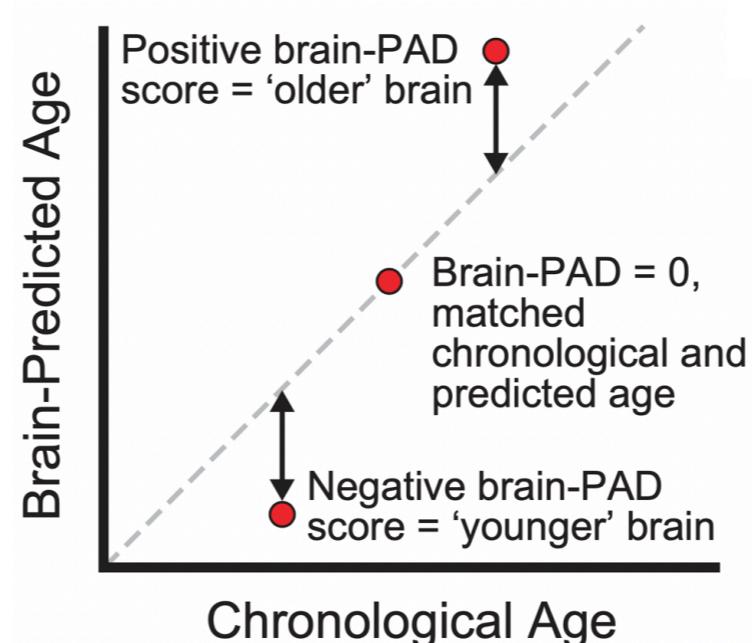


LARIBOISIÈRE



Solution: Surrogate Biomarker e.g. Brain Age

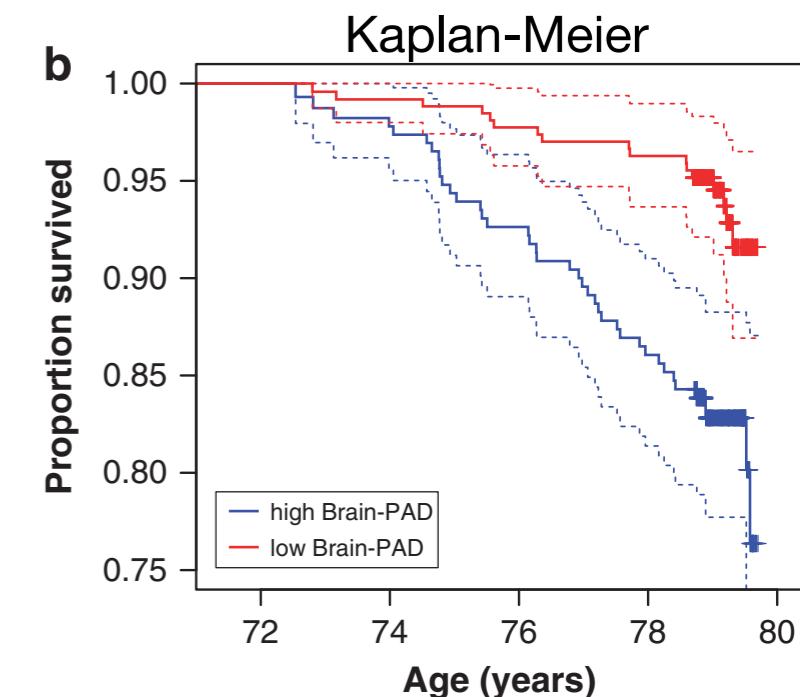
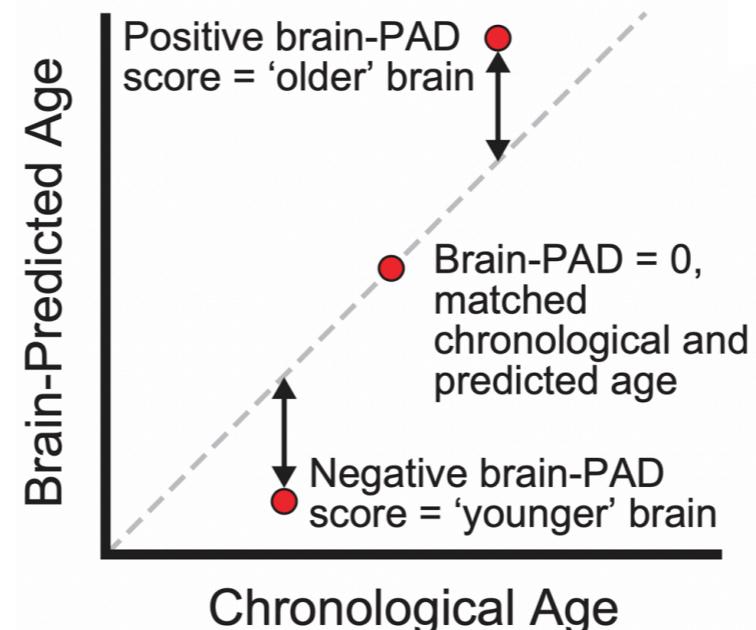
- **Problem:** few / expensive data on outcome e.g. cognitive decline
- **Idea:** Predict widely available outcome; exploit correlation with the outcome of interest, e.g. age
- Brain Age Delta = predicted age (PAD) - passport age



Cole et al. Mol. Psych. 2018

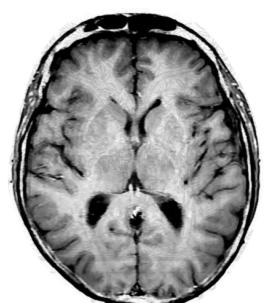
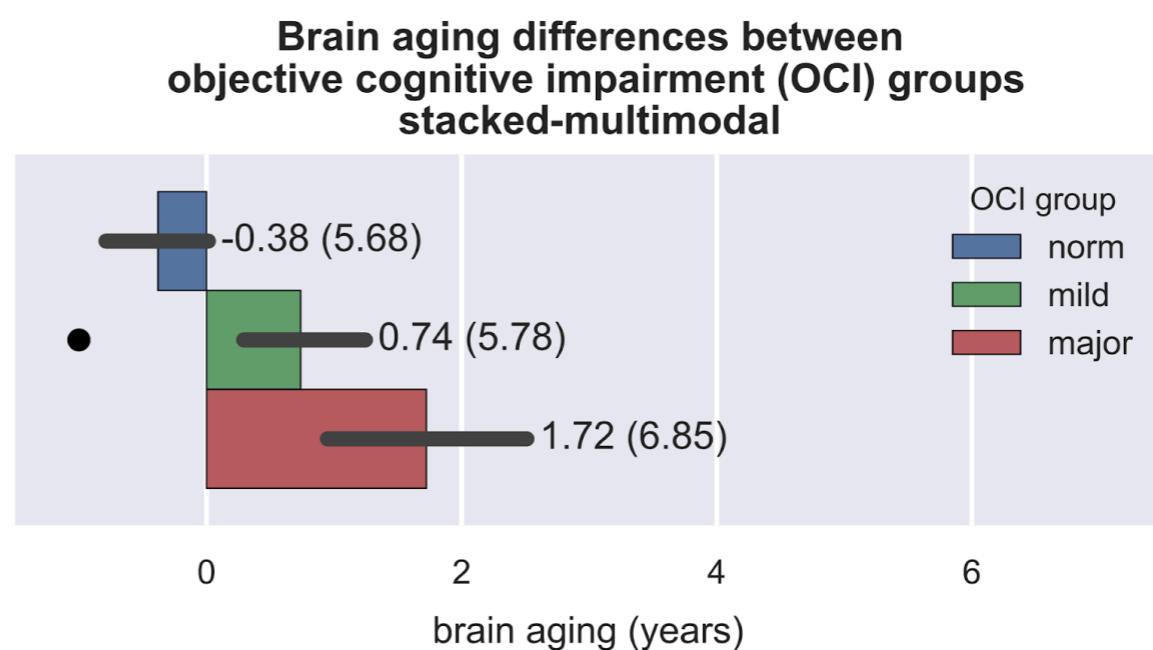
Solution: Surrogate Biomarker e.g. Brain Age

- **Problem:** few / expensive data on outcome e.g. cognitive decline
- **Idea:** Predict widely available outcome; exploit correlation with the outcome of interest, e.g. age



Cole et al. Mol. Psych. 2018

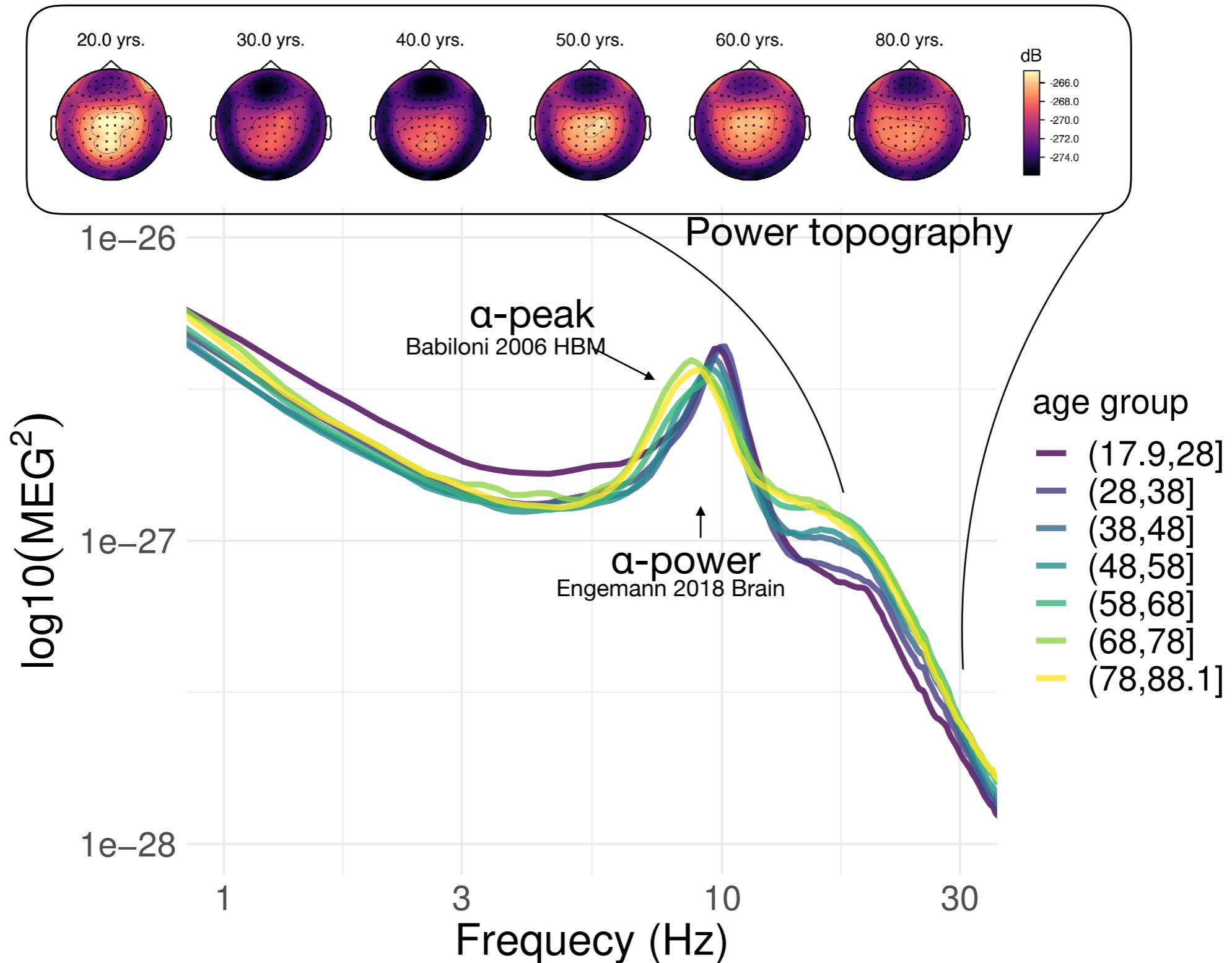
- Brain Age Delta = predicted age (PAD) - passport age
- High PAD is not good ...
- **Typically estimated with MRI!**



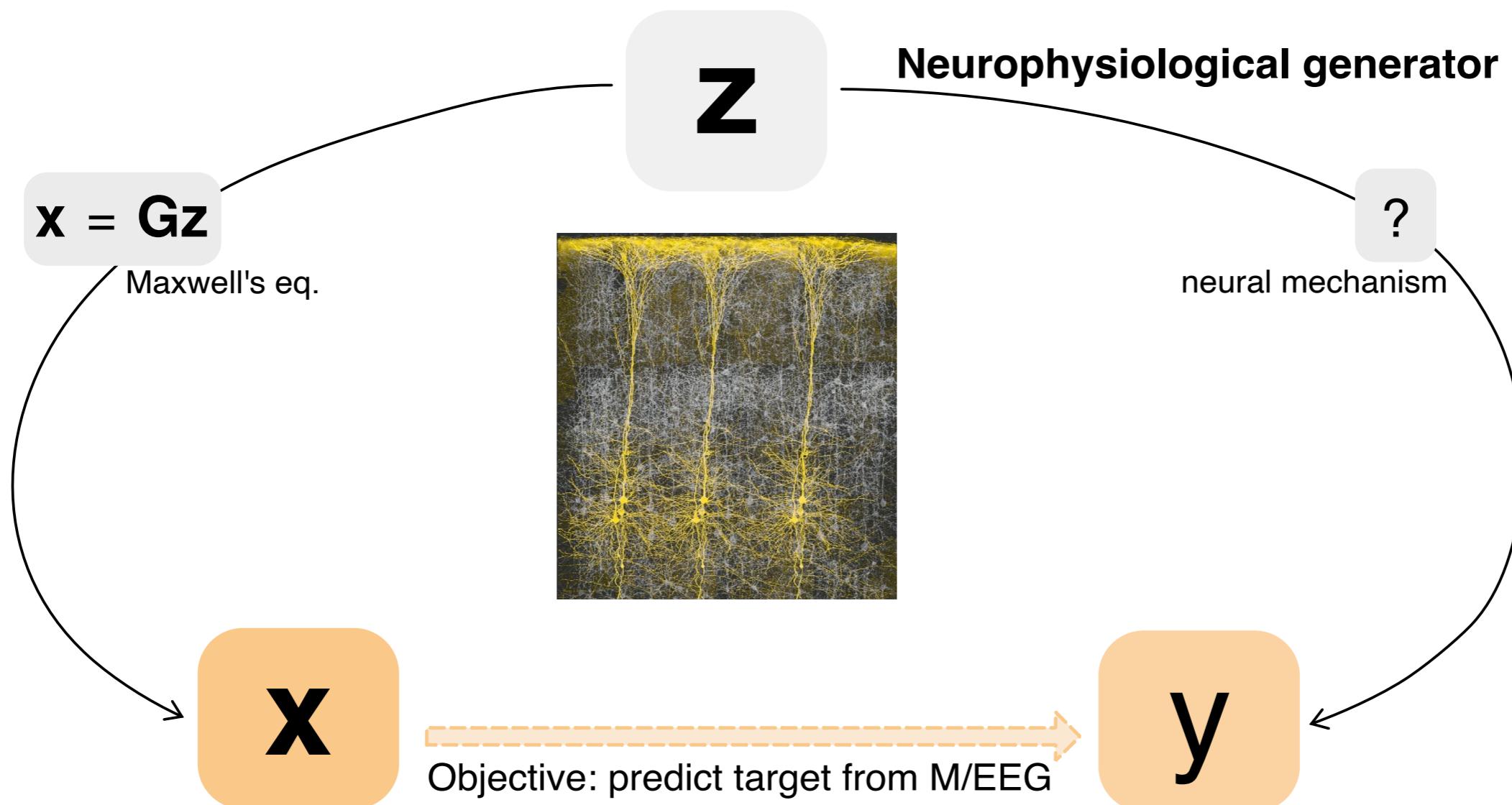
Using
MRI

Liem et al 2017 NIMG

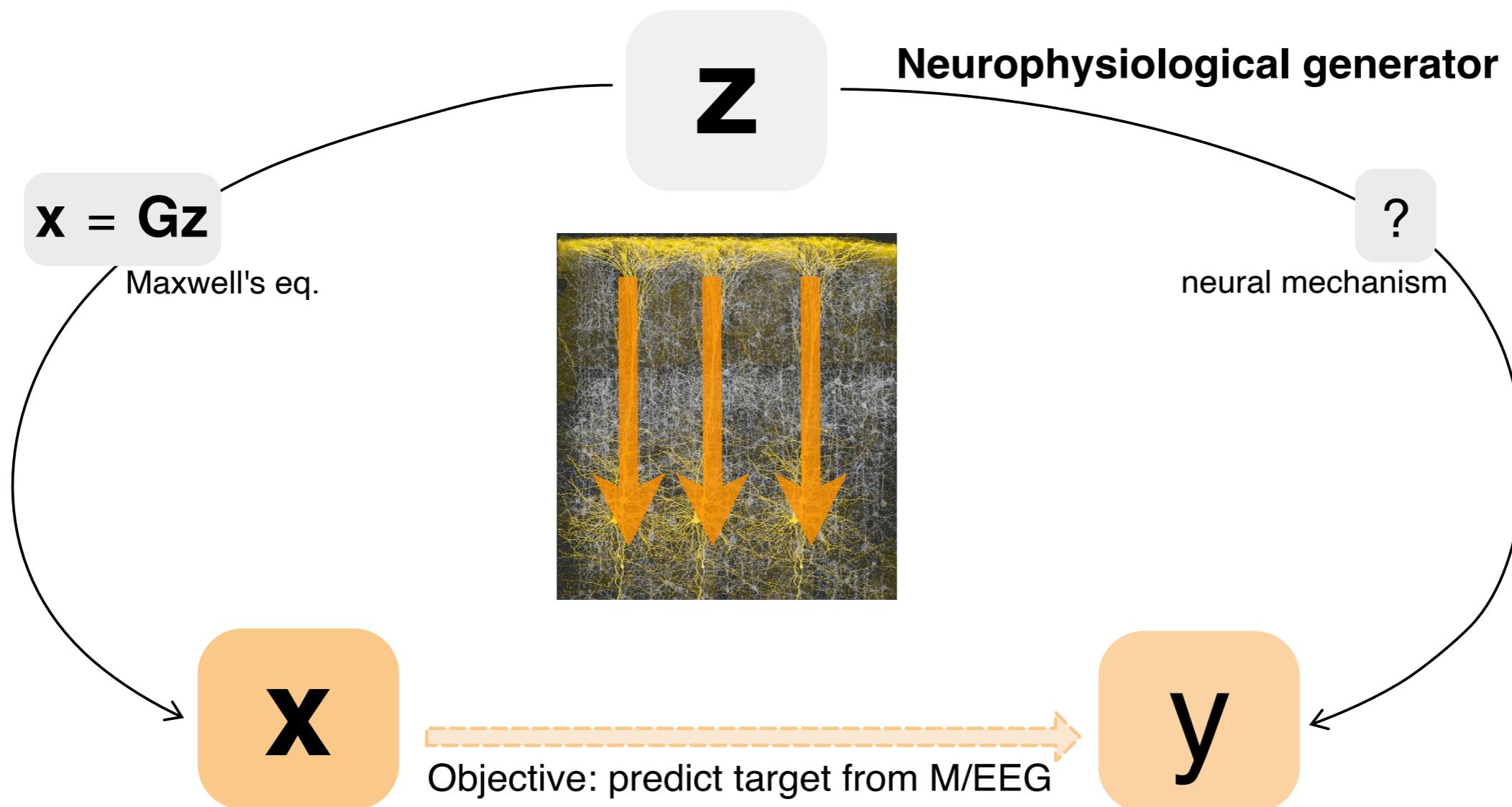
Age prediction: Which M/EEG features?



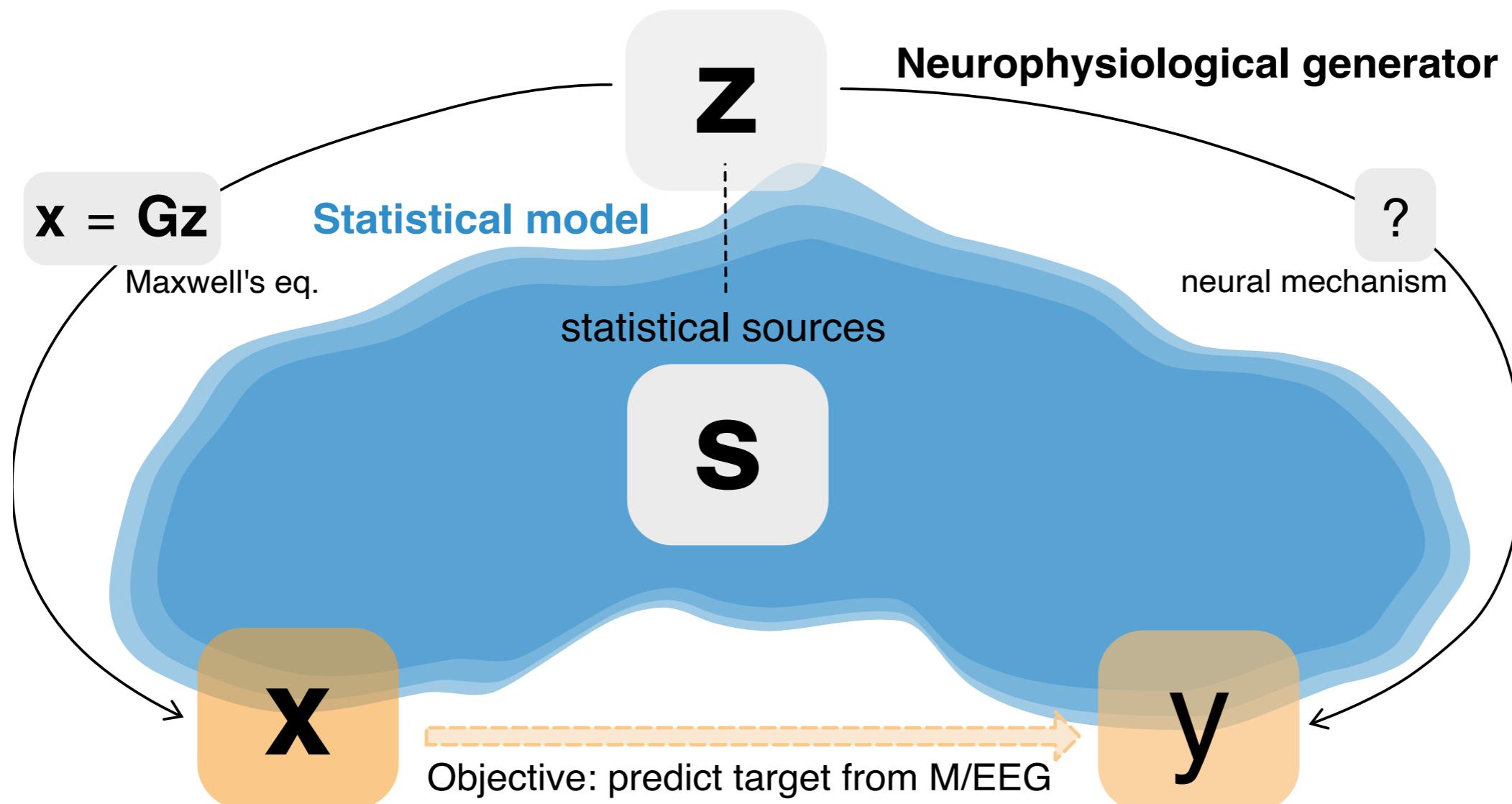
How to build MEG-based regression models?



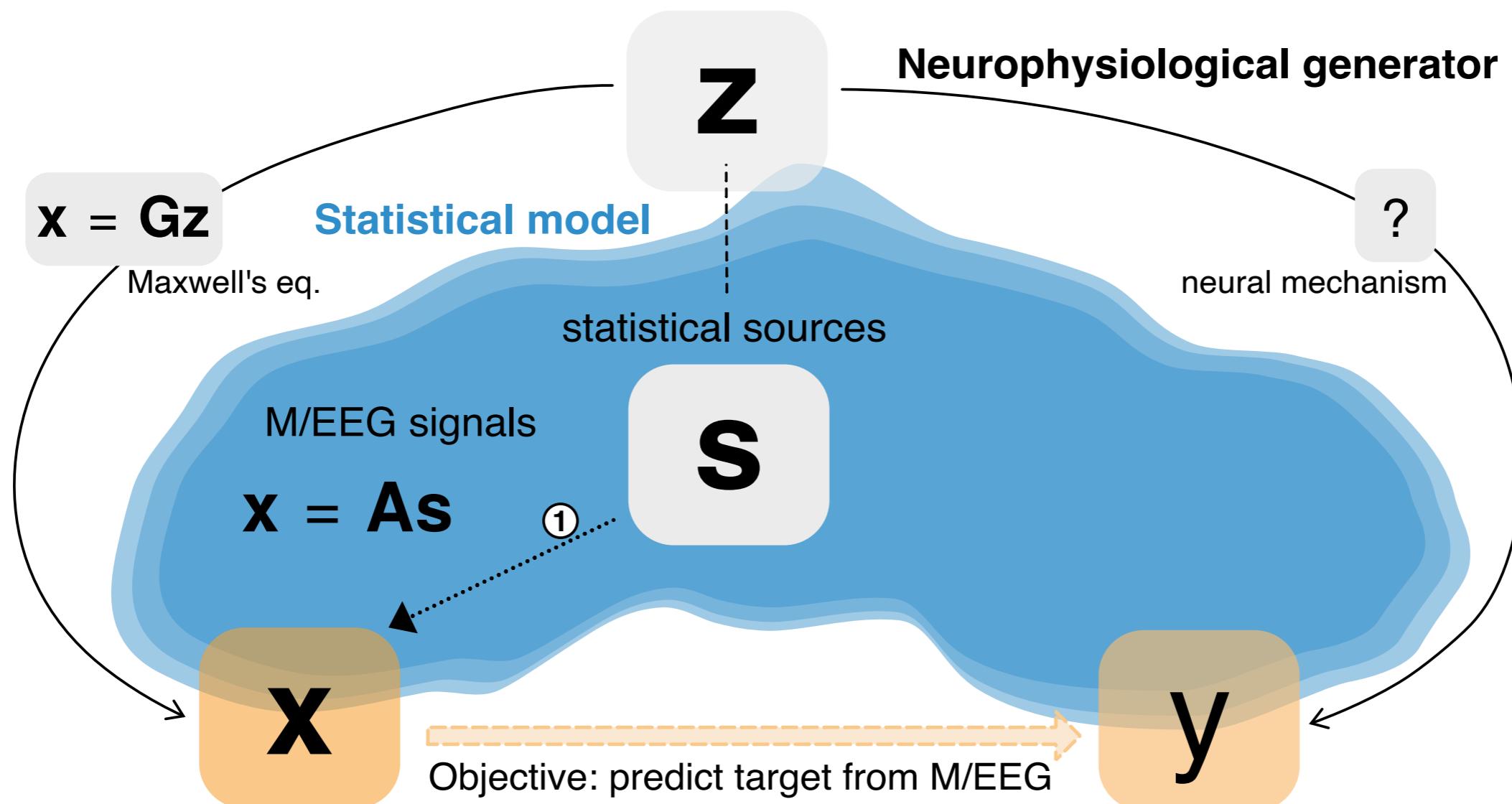
How to build MEG-based regression models?



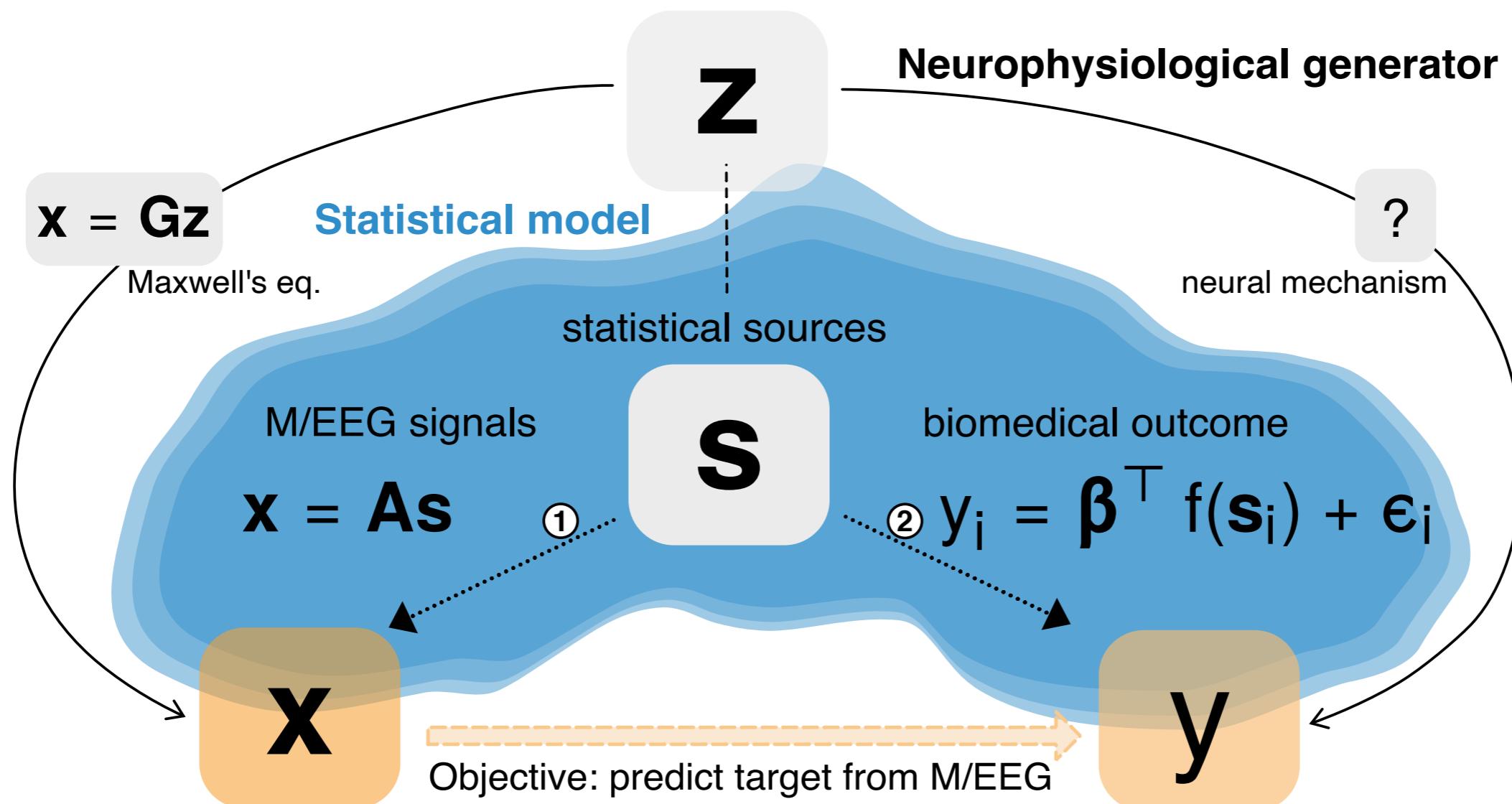
How to build MEG-based regression models?



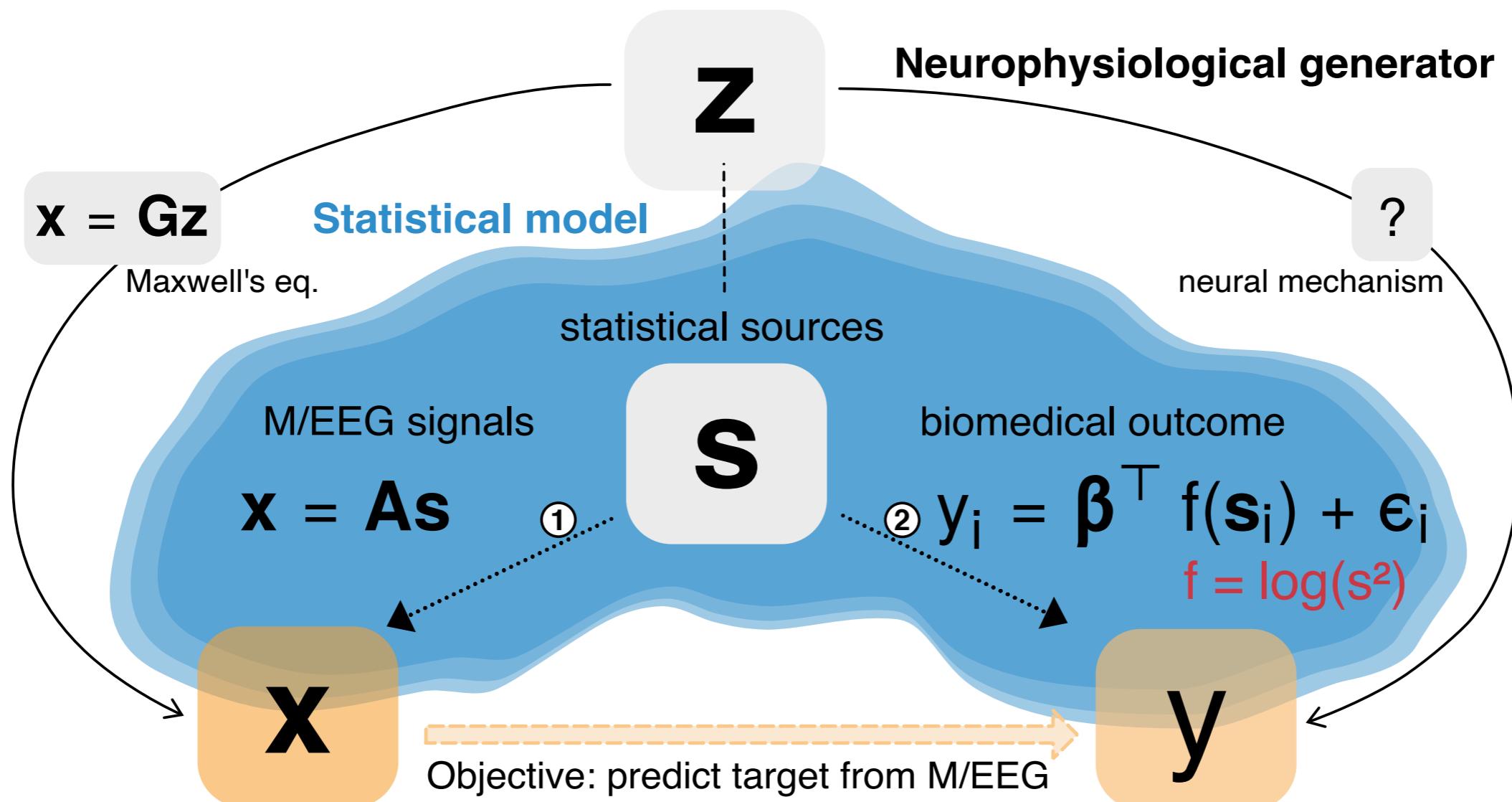
How to build MEG-based regression models?



How to build MEG-based regression models?



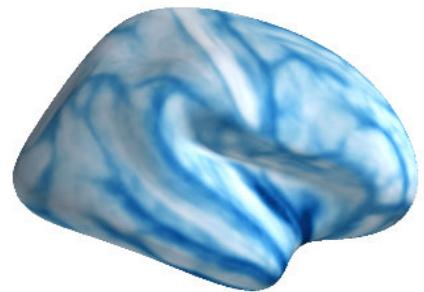
How to build MEG-based regression models?



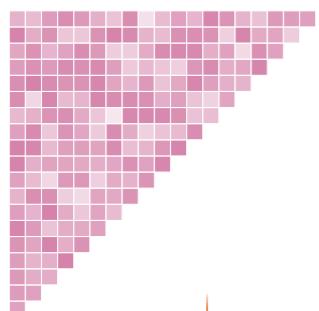
How shall we combine MEG with MRI? And is it worth the effort?

Multimodal input data

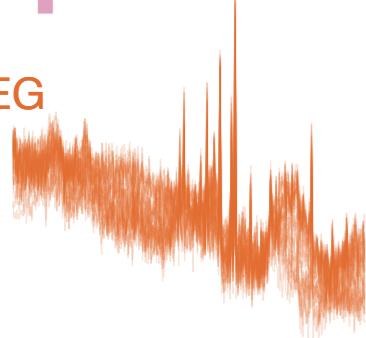
anatomical MRI



functional MRI



MEG

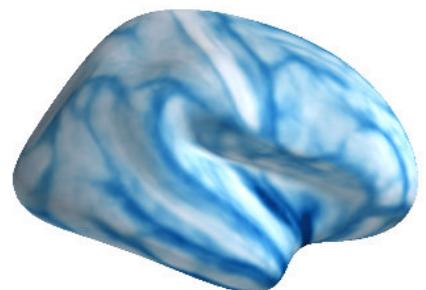


Stacking method: Wolpert 1992, Liem et al. 2017, NIMG; Karrer et al. 2019, HBM, ...

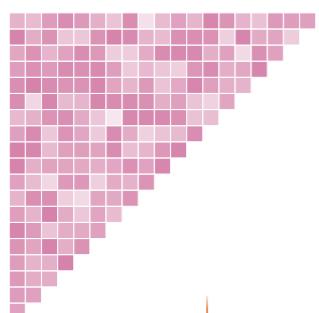
How shall we combine MEG with MRI? And is it worth the effort?

Multimodal input data Layer I: Ridge Regression

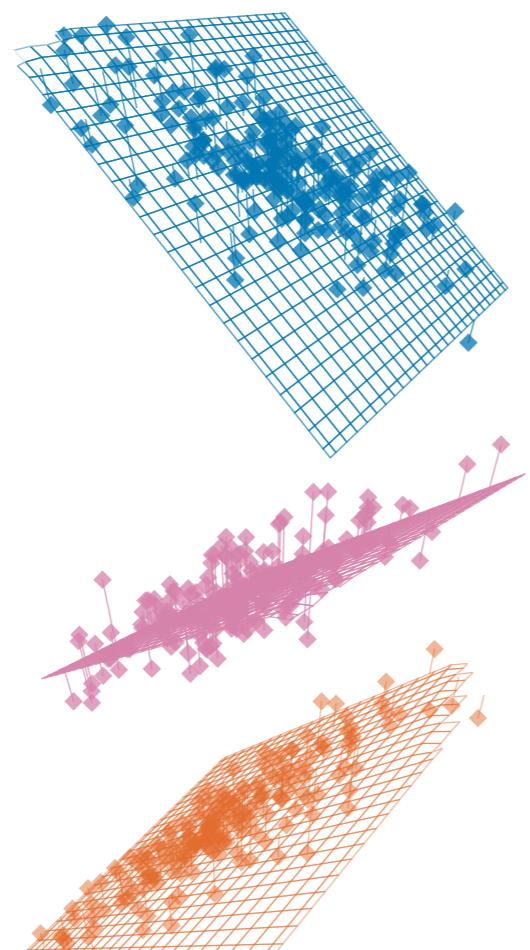
anatomical MRI



functional MRI



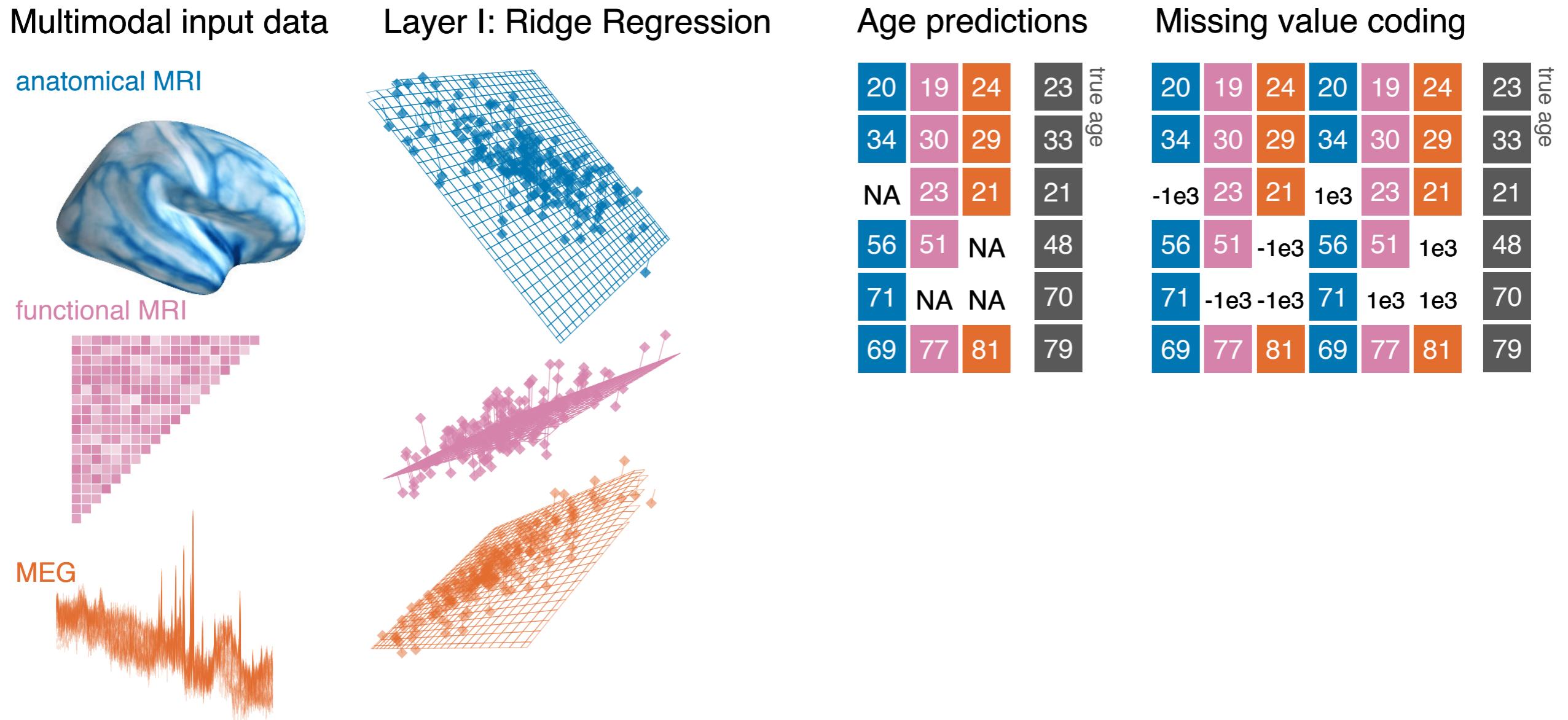
MEG



Stacking method: Wolpert 1992, Liem et al. 2017, NIMG; Karrer et al. 2019, HBM, ...

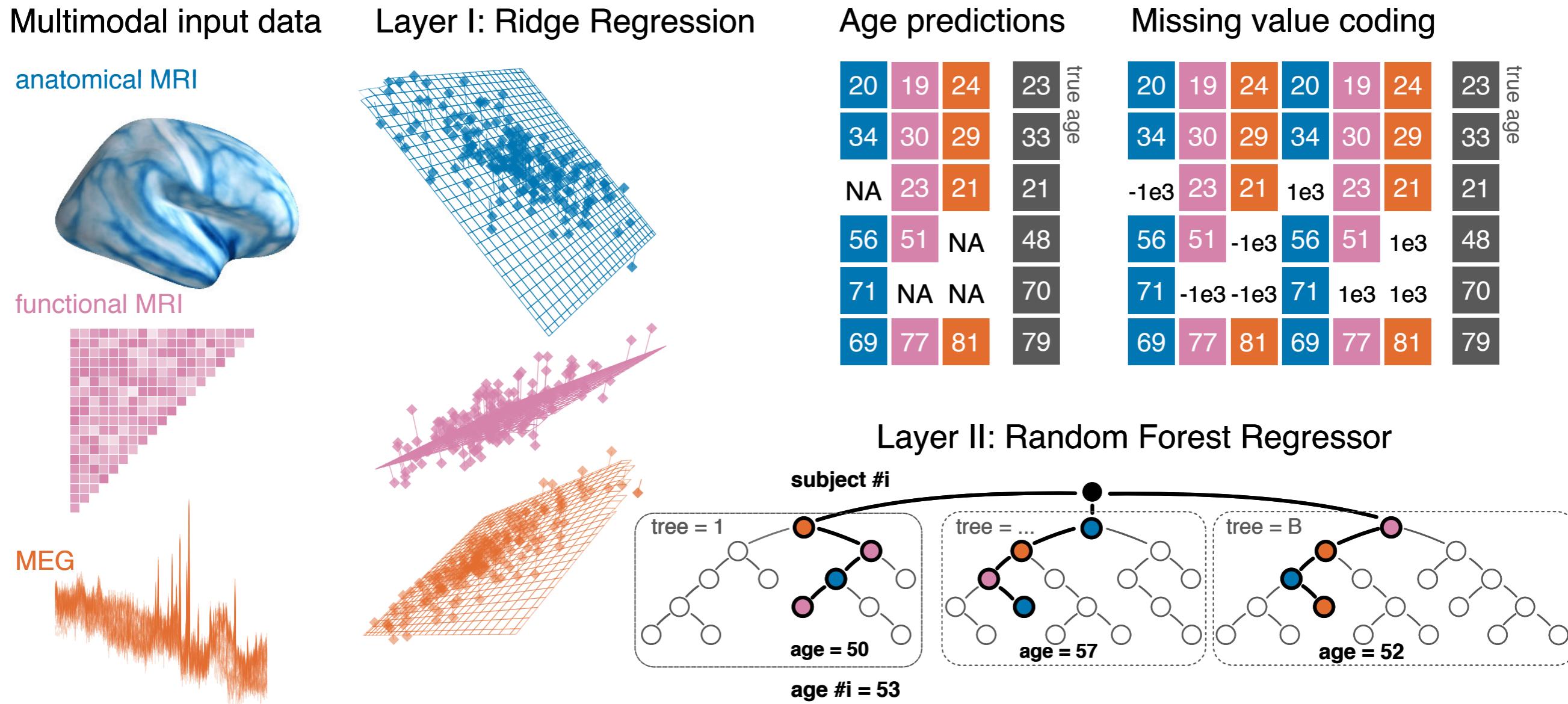
Engemann et al. (2020) eLife

How shall we combine MEG with MRI? And is it worth the effort?



Stacking method: Wolpert 1992, Liem et al. 2017, NIMG; Karrer et al. 2019, HBM, ...

How shall we combine MEG with MRI? And is it worth the effort?



Stacking method: Wolpert 1992, Liem et al. 2017, NIMG; Karrer et al. 2019, HBM, ...

Overview on all features

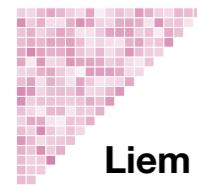
Table 1. Frequency band definitions.

Name	Low	δ	θ	α	β_1	β_2	γ_1	γ_2	γ_3
range (Hz)	0.1 - 1.5	1.5 - 4	4 - 8	8 - 15	15 - 26	26 - 35	35 - 50	50 - 74	76 - 100

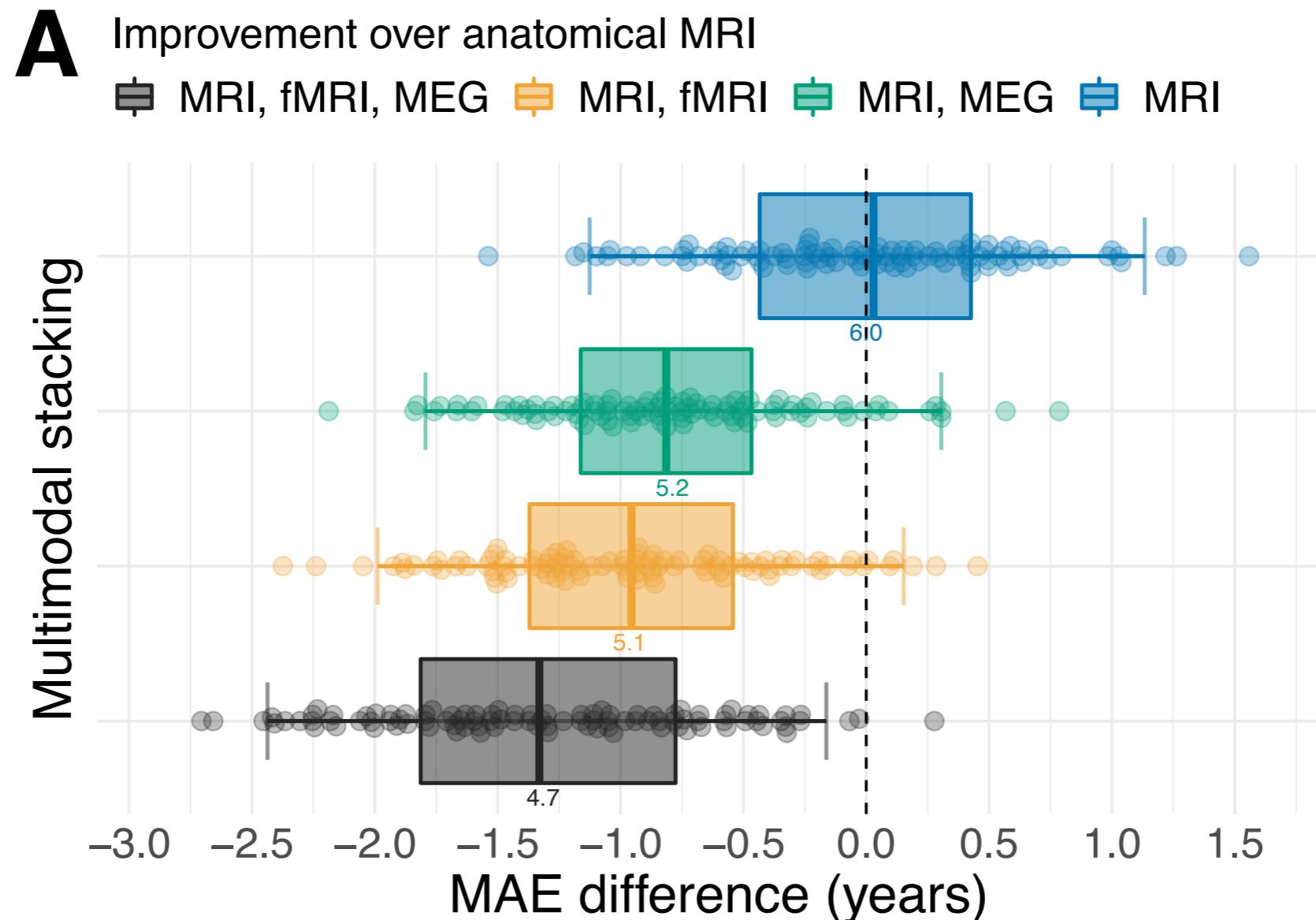
Table 2. Summary of extracted features.

#	Modality	Family	Input	Feature	Variants	Spatial selection
1	MEG	sensor mixed	ERF	latency	aud, vis, audvis	max channel
2	PSD α	peak		max channel
3	PSD	1/f slope	low, γ	max channel in ROI
4	...	source activity	signal	power	low, $\delta, \theta, \alpha, \beta_{1,2}, \gamma_{1,2,3}$	MNE, 448 ROIs
5	envelope
6	...	source connectivity	signal	covariance
7	envelope
8	env.	corr.
9	env.	corr. ortho.
10	fMRI	connectivity	time-series	correlation	...	256 ROIs
11	MRI	anatomy	volume	cortical thickness		5124 vertices
12	surface	cortical surface area		5124 vertices
13	volume	subcortical volumes		66 ROIs

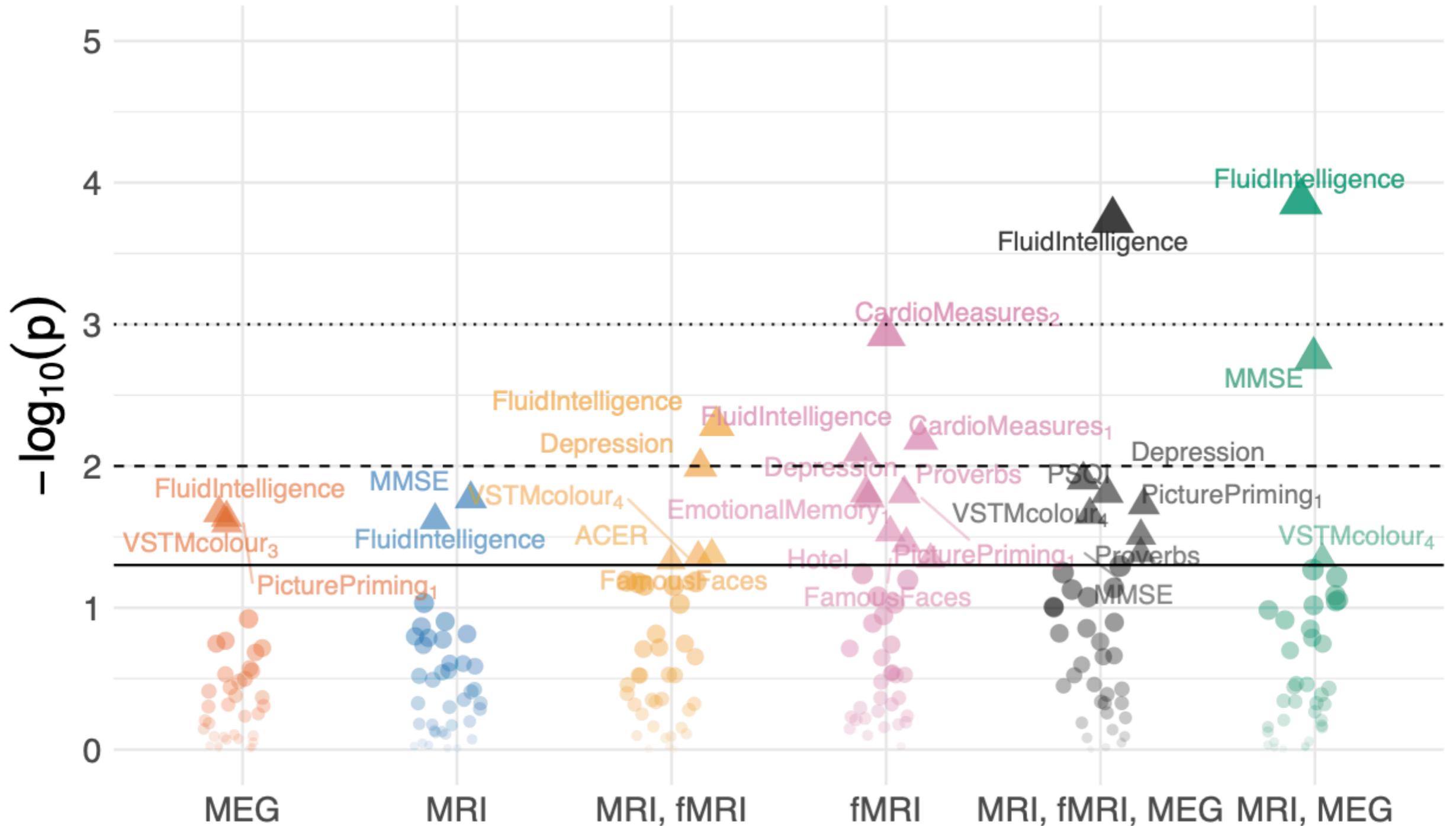
Note. ERF = event related field, PSD = power spectral density, MNE = Minimum Norm-Estimates, ROI = region of interest, corr. = correlation, ortho. = orthogonalized.



Predicting brain age from MRI & MEG enhances predictive performance



Predicting brain age from MRI & MEG enhances predictive performance & cognitive phenotyping





What if I don't have
MRI and only
electrophysiology?

Covariance

$X_i \in \mathbb{R}^{C \times T}$: MEG/EEG signals

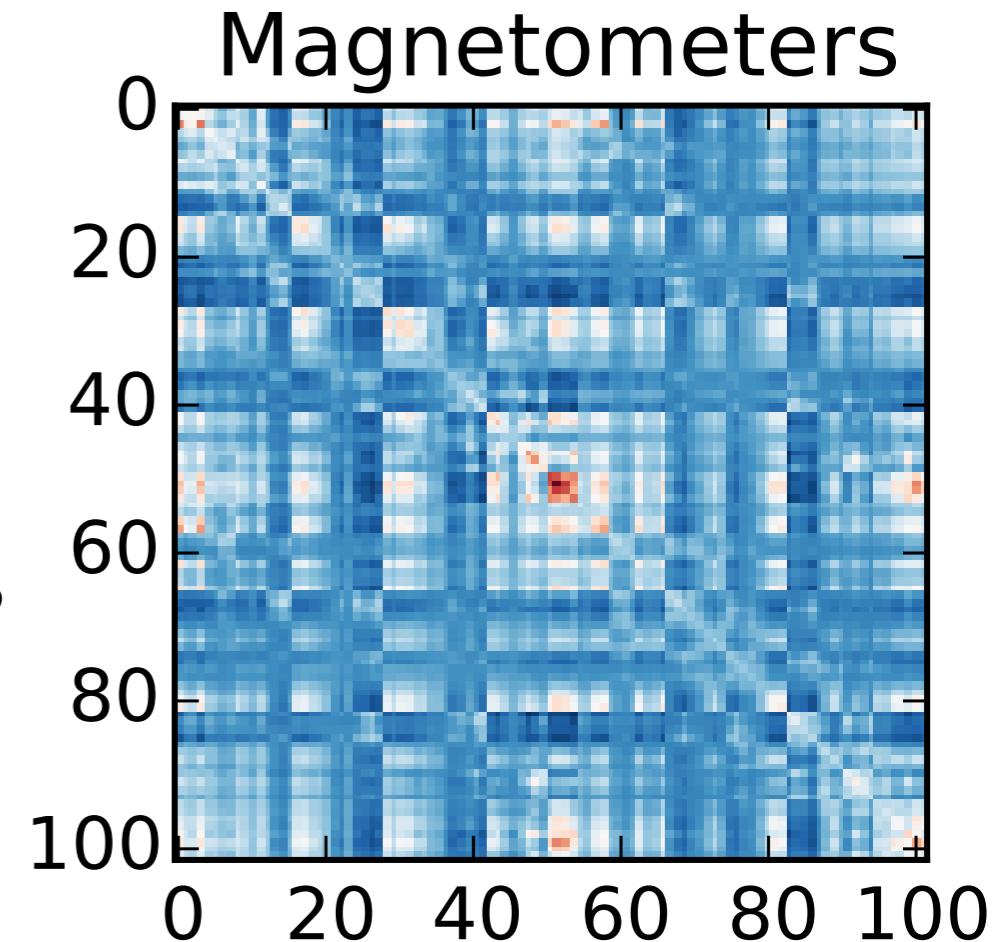
C : number of channels

T : number of time samples

Definition [Covariance]:

$$\Sigma_i = \frac{X_i X_i^\top}{T} \in \mathbb{S}_+^C \subset \mathbb{R}^{C \times C}$$

Set of non-negative
symmetric matrices



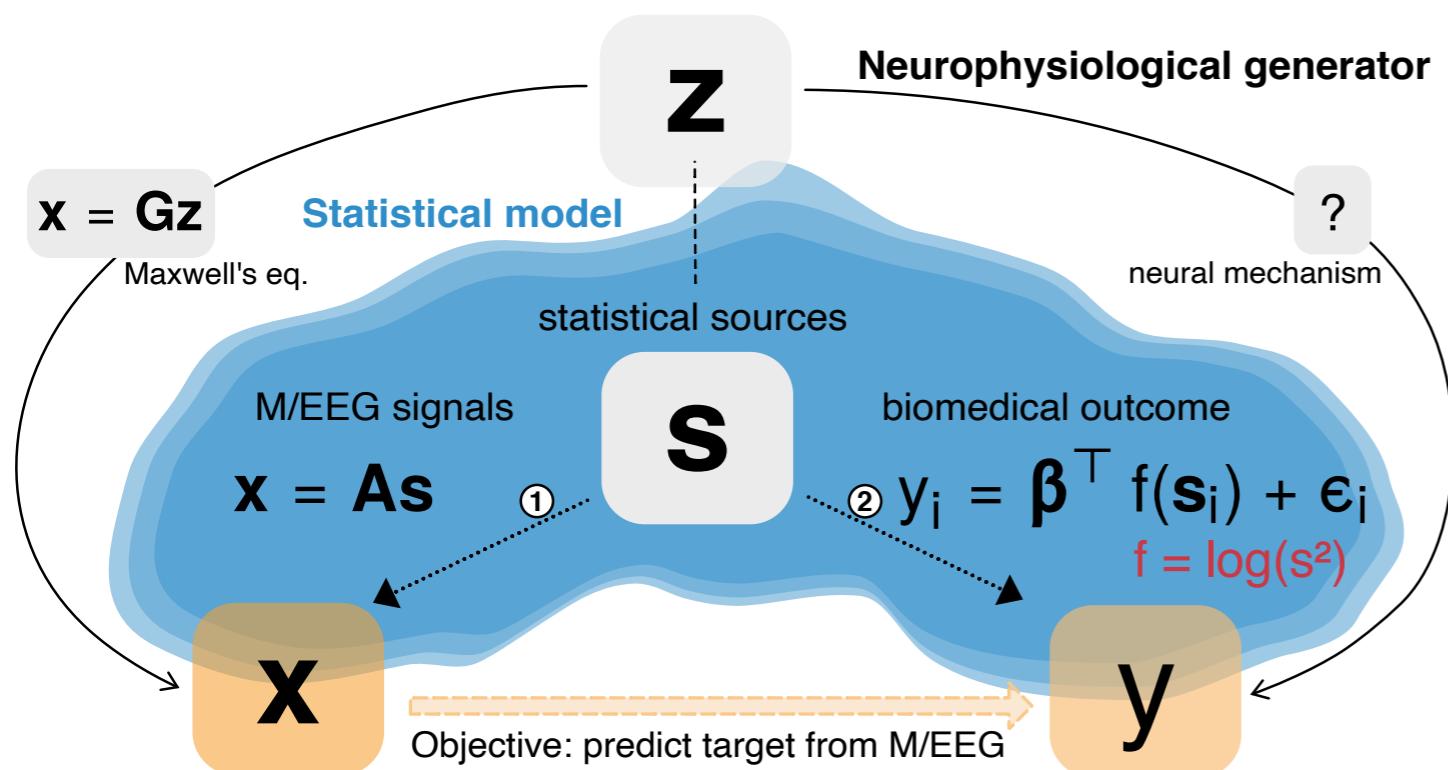
Remark: Estimation
of cov. requires care

Predicting from M/EEG source power

Without biophysical source localization (estimation of z)

Problem: A is unknown

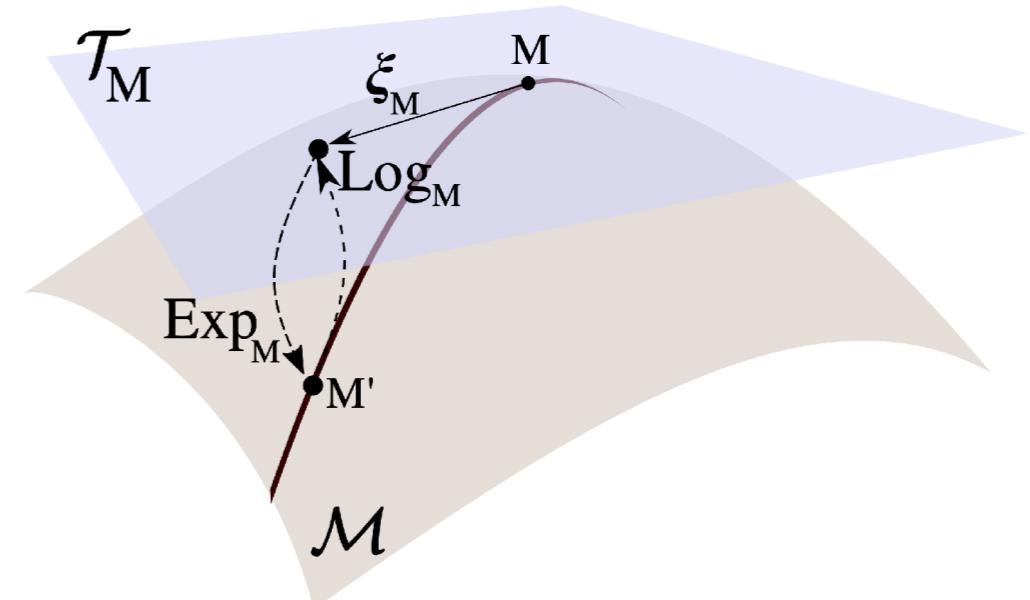
Idea: use affine invariant distances between covariances



Use cov. as representation

$$\Sigma_i = \frac{S_i S_i^\top}{T} \in \mathbb{S}_+^C \subset \mathbb{R}^{C \times C}$$

$$d_G(\Sigma_1, \Sigma_2) = \|\log(\Sigma_1^{-\frac{1}{2}} \Sigma_2 \Sigma_1^{-\frac{1}{2}})\|_F$$

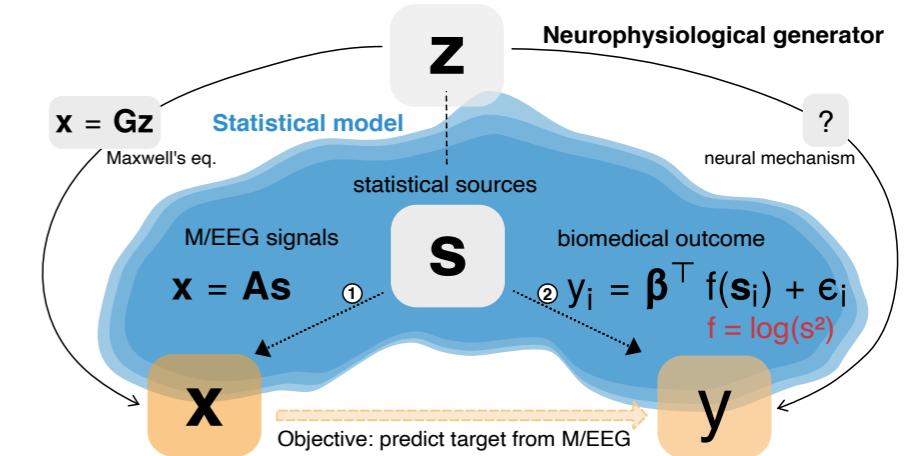
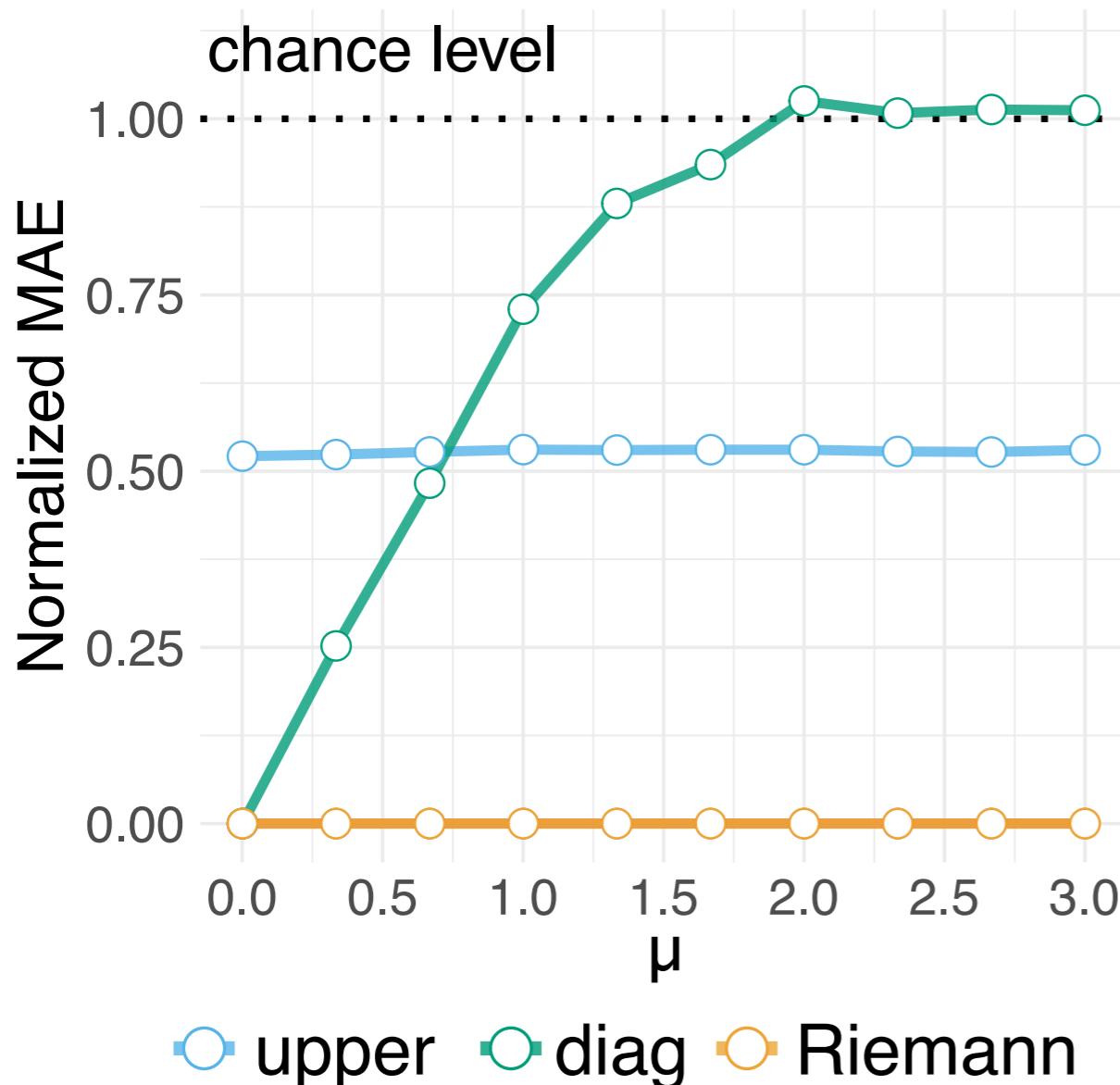


Take out volume conduction
using Riemannian distance

Simulations

Are shortcuts – in principle – possible?

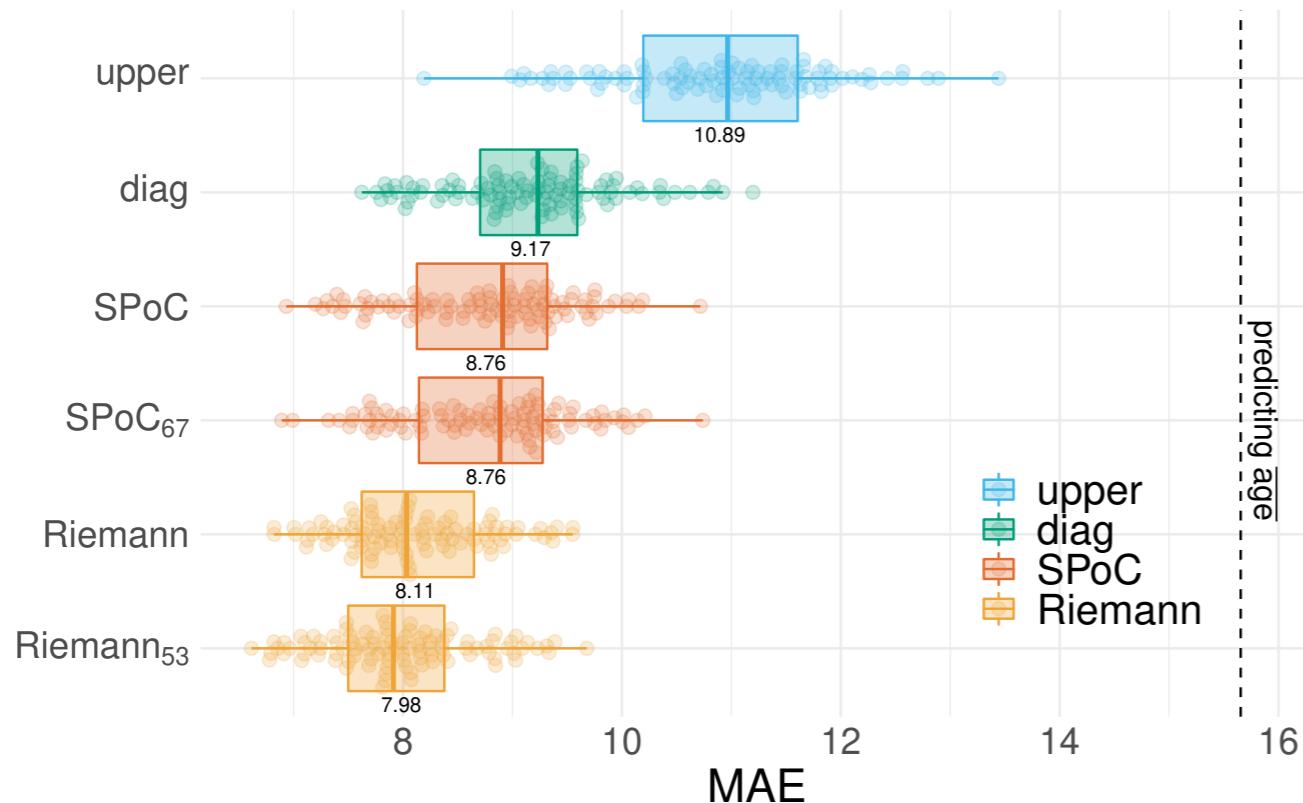
$$A = \exp(\mu B) \text{ with } B \text{ random.}$$



We have empirical and mathematical proven statistical consistency

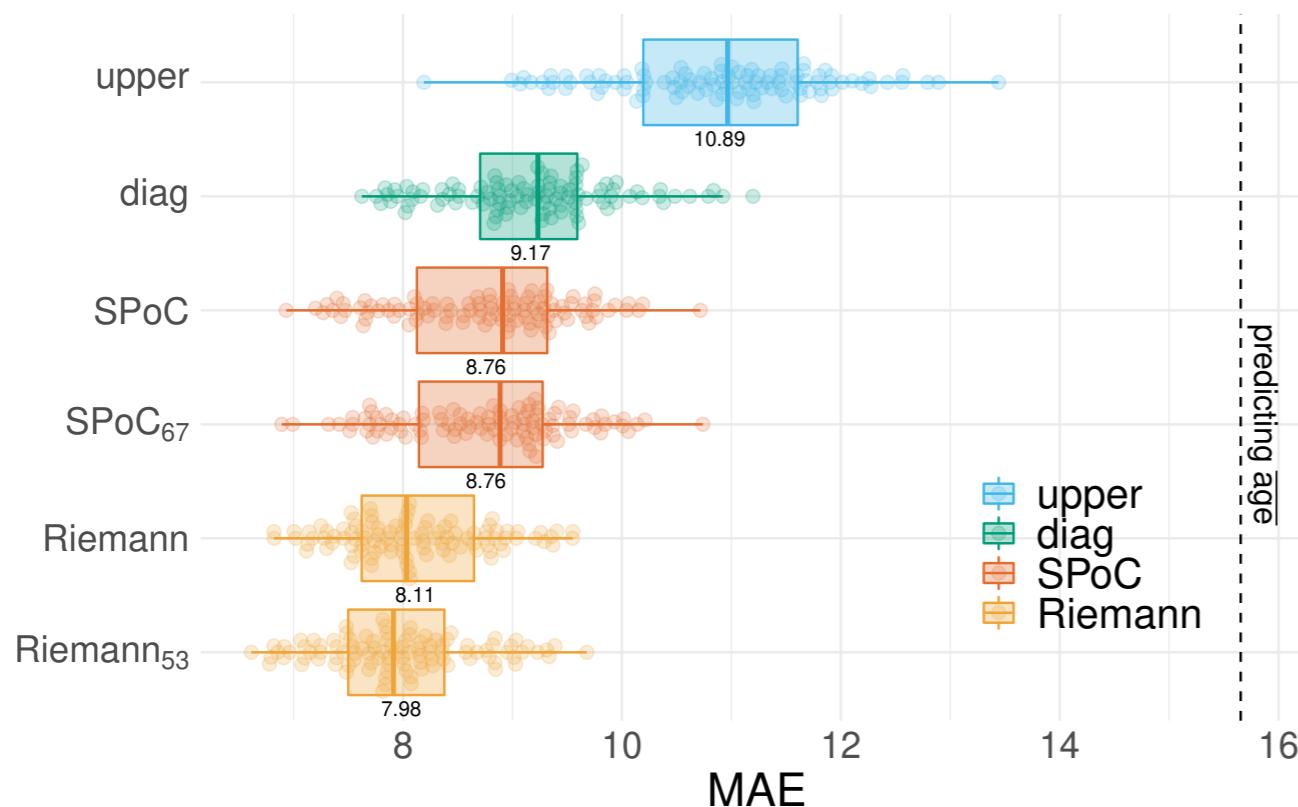
Empirical benchmarks

MEG (Cam-CAN, n=650)

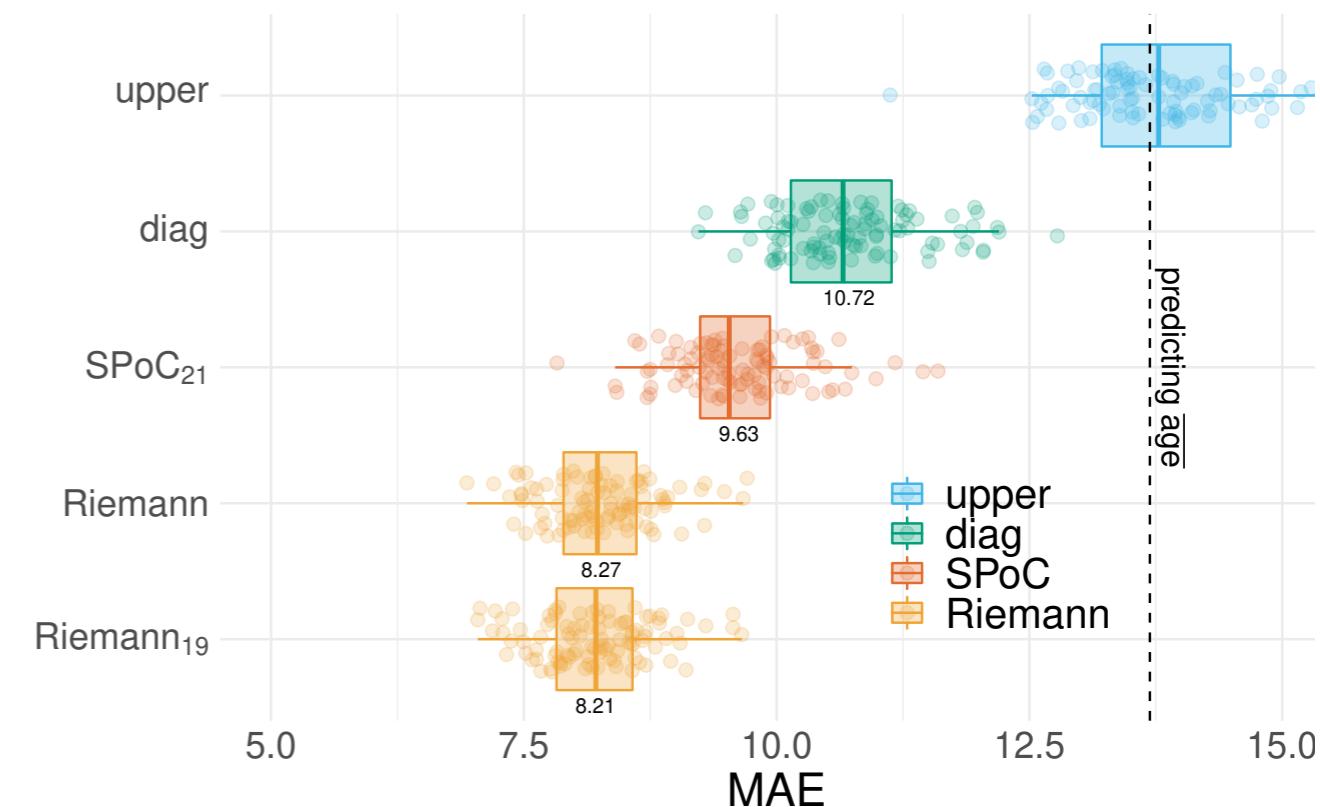


Empirical benchmarks

MEG (Cam-CAN, n=650)

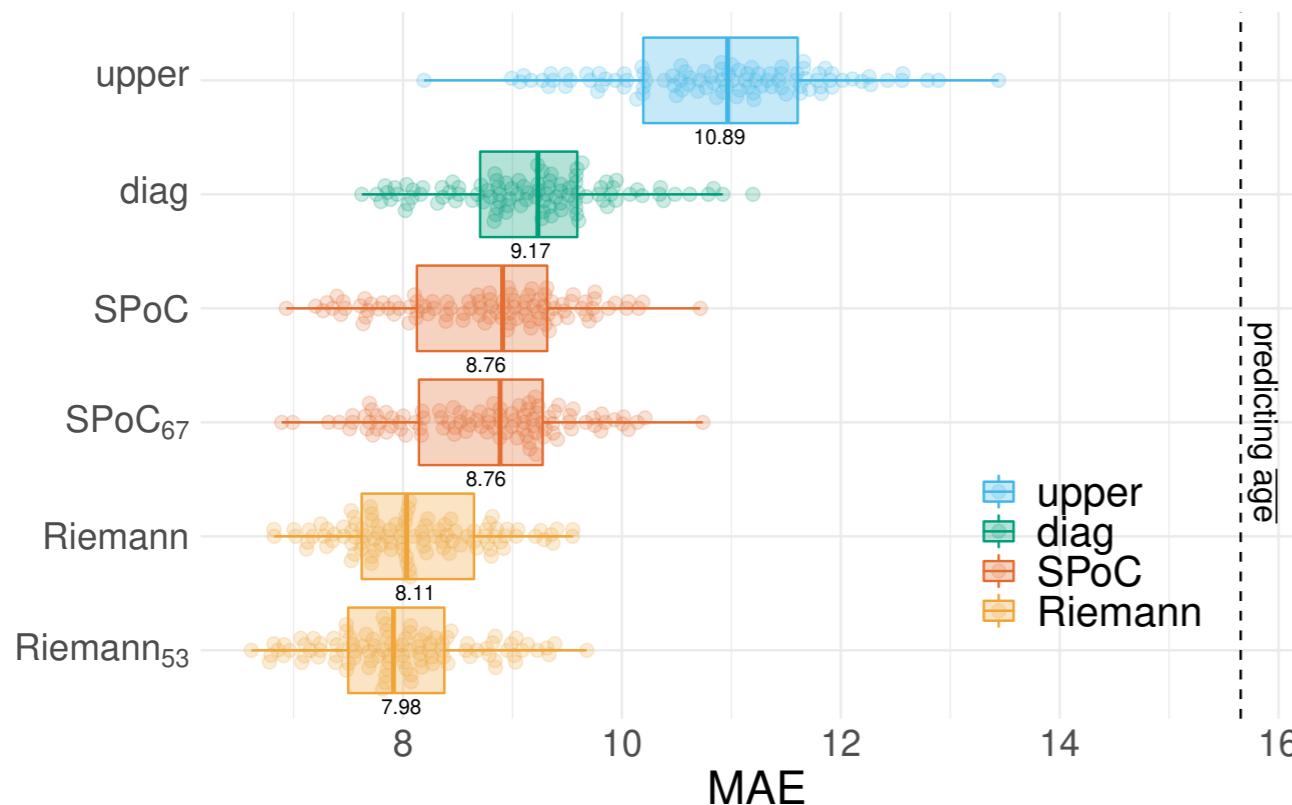


EEG (TUH data, n=1385)

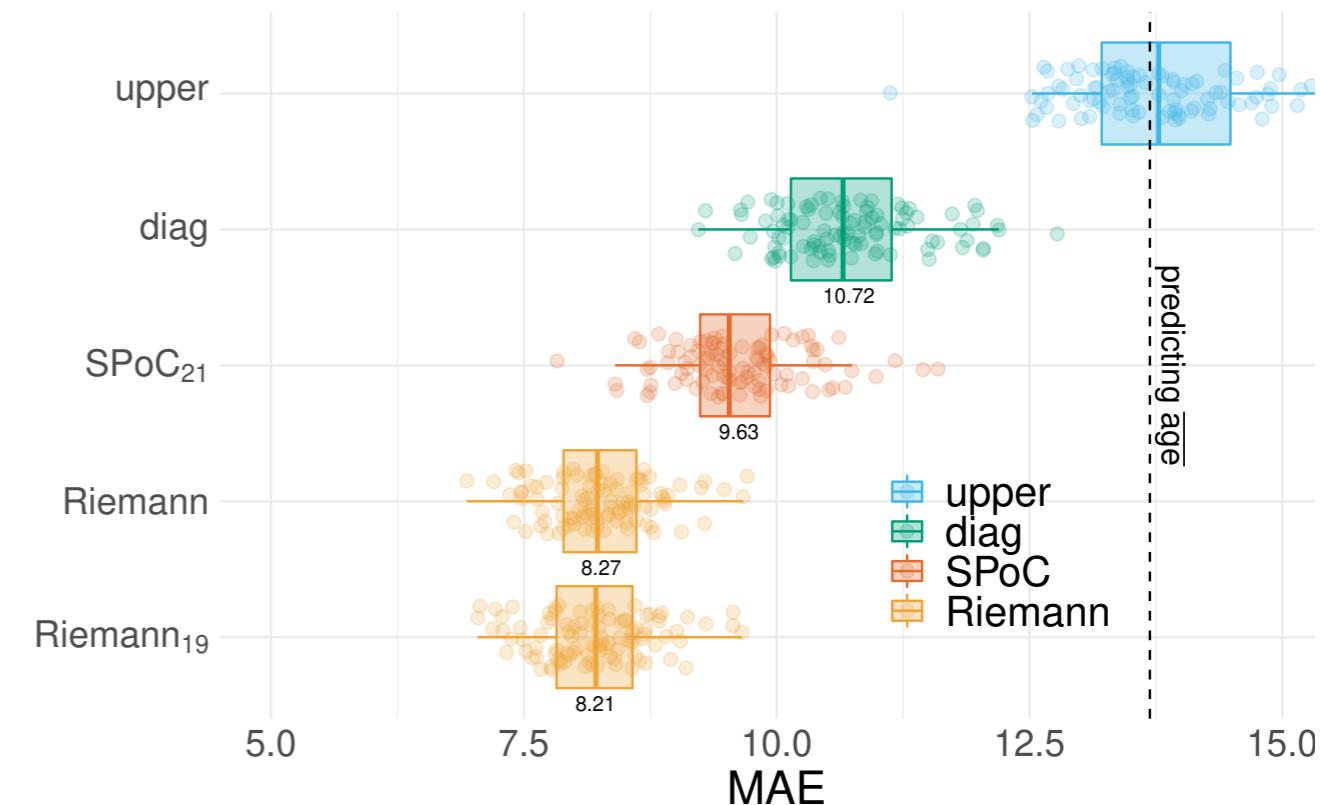


Empirical benchmarks

MEG (Cam-CAN, n=650)



EEG (TUH data, n=1385)



EEG can in principle be substituted for MEG

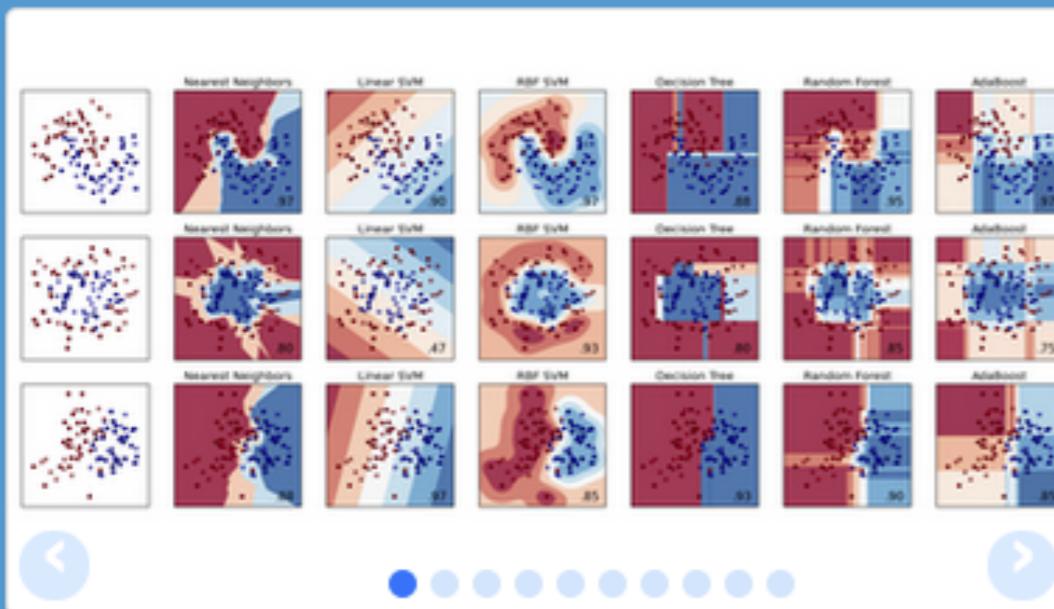
Conclusion

- Neuroscience signals are under exploited
- Need for better models and tools
- Need more interdisciplinary work (CS, ML, stats, neuro, physics...)
- If you want the maths look at papers...



“All models are wrong but some come with good open source implementation and good documentation so use those.”





scikit-learn

Machine Learning in Python

- Simple and efficient tools for data mining and data analysis
- Accessible to everybody, and reusable in various contexts
- Built on NumPy, SciPy, and matplotlib
- Open source, commercially usable - BSD license

Classification

Identifying to which category an object belongs to.

Applications: Spam detection, Image recognition.

Algorithms: *SVM, nearest neighbors, random forest, ...*

[— Examples](#)

Regression

Predicting a continuous-valued attribute associated with an object.

Applications: Drug response, Stock prices.

Algorithms: *SVR, ridge regression, Lasso,*

...

[— Examples](#)

Clustering

Automatic grouping of similar objects into sets.

Applications: Customer segmentation, Grouping experiment outcomes

Algorithms: *k-Means, spectral clustering, mean-shift, ...*

[— Examples](#)

Dimensionality reduction

Reducing the number of random variables to consider.

Applications: Visualization, Increased efficiency

Algorithms: *PCA, feature selection, non-*

Model selection

Comparing, validating and choosing parameters and models.

Goal: Improved accuracy via parameter tuning

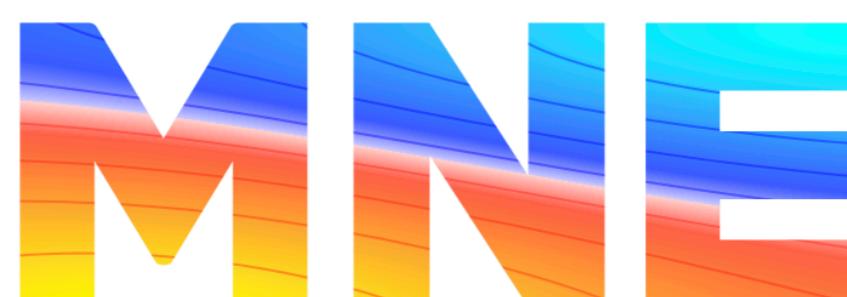
Modules: *grid search, cross validation*

Preprocessing

Feature extraction and normalization.

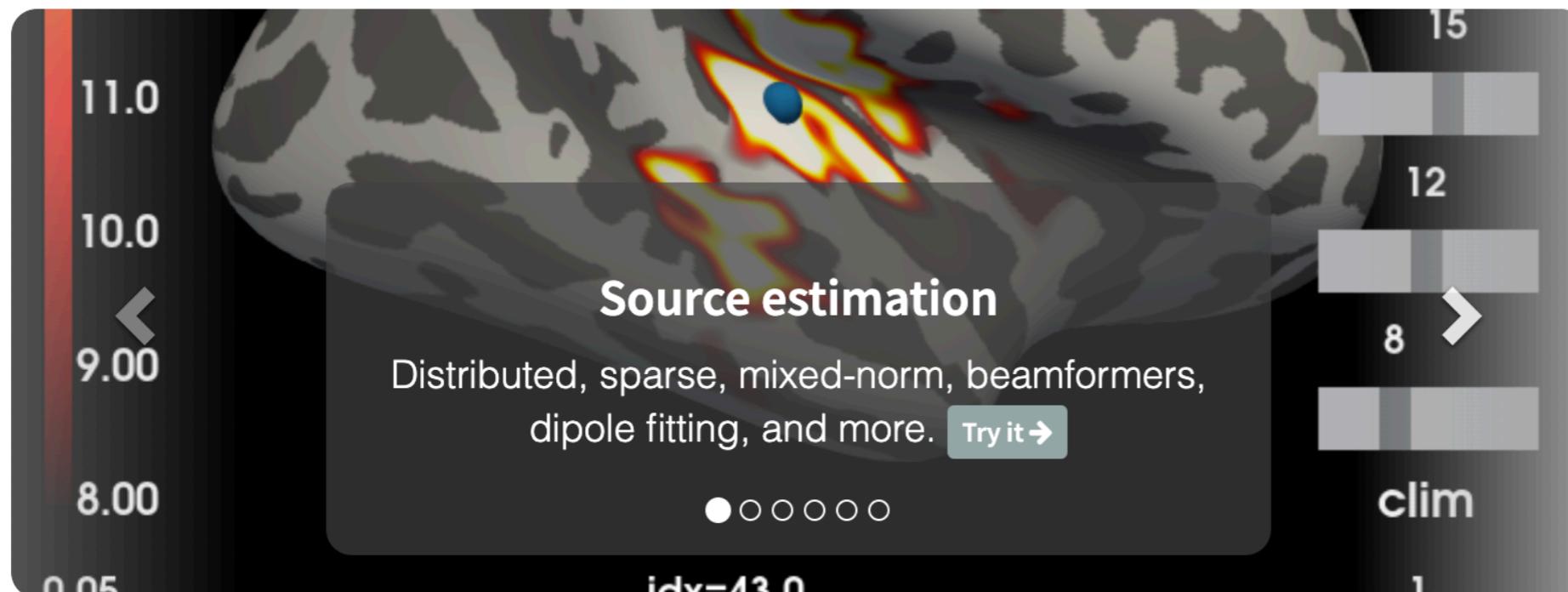
Application: Transforming input data such as text for use with machine learning algorithms.

Modules: *preprocessing, feature extraction.*



MEG + EEG ANALYSIS & VISUALIZATION

Open
human
<https://mne.tools/>



Version 0.21.dev0

- [What's new](#)
- [Installation](#)
- [Documentation](#)
- [Cite](#)

Direct financial support

- NIH** National Institutes of Health: R01-EB009048, EB009048, EB006385, HD40712, NS44319, NS37462, NS104585, P41-EB015896, RR14075-06
- US National Science Foundation:** 0958669, 1042134
- European Research Council:** YStG-263584, 676943
- US Department of Energy:** DE-FG02-99ER62764 (MIND)
- Agence Nationale de la Recherche:** 14-NEUC-0002-01
IDEX Paris-Saclay, 11-IDEX-0003-02
- Paris-Saclay Center for Data Science:** PARIS-SACLAY
- Google:** Summer of code (x6)
- Amazon:** AWS Research Grants
- Chan Zuckerberg Initiative:** Essential Open Source Software for Science



Thanks !



D. Sabbagh



D. Engemann



G. Varoquaux



B. Thirion



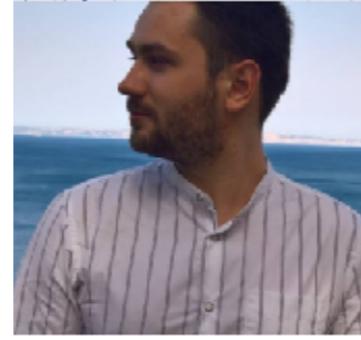
T. Dupré La Tour



M. Jas



F. Liem



O. Kozynets



G. Lemaitre



P. Ablin



T. Moreau



H. Banville

as well as A. Hyvärinen, U. Simsekli, O. Chehab

Contact

<http://alexandre.gramfort.net>

GitHub : @agramfort



Twitter : @agramfort



Support

ANR-NSF Grant Thalameeg
European Research Council (ERC SLAB-YStG-676943)



Main references

Learning the Morphology of Brain Signals Using Alpha-Stable Convolutional Sparse Coding, M. Jas, T. Dupré la Tour, U. Simsekli, A. Gramfort, (2017) Advances in Neural Information Processing Systems (NeurIPS)

Uncovering the structure of clinical EEG signals with self-supervised learning
Banville, H., Chehab, O., Hyvärinen, A., Engemann, D. and Gramfort, A. (2020)
Journal of Neural Engineering – ArXiv abs/2007.16104

Manifold-regression to predict from MEG/EEG brain signals without source modeling
Sabbagh, D., Ablin, P., Varoquaux, G., Gramfort, A. and Engemann, D. (2019)
Advances in Neural Information Processing Systems (NeurIPS)

Predictive regression modeling with MEG/EEG: from source power to signals and cognitive states, Sabbagh, D., Ablin, P., Varoquaux, G., Gramfort, A. and Engemann, D. (2020), *NeuroImage*

Combining magnetoencephalography with magnetic resonance imaging enhances learning of surrogate-biomarkers, Engemann, D., Kozynets, O., Sabbagh, D., Lemaître, G., Varoquaux, G., Liem, F. and Gramfort, A. (2020), *eLife*

MNE software for processing MEG and EEG data, A. Gramfort, M. Luessi, E. Larson, D. Engemann, D. Strohmeier, C. Brodbeck, L. Parkkonen, M. Härmäläinen, (2013) *Neuroimage*