Center of Functionally Integrative Neuroscience
Aarhus University / Aarhus University Hospital - DENMARK

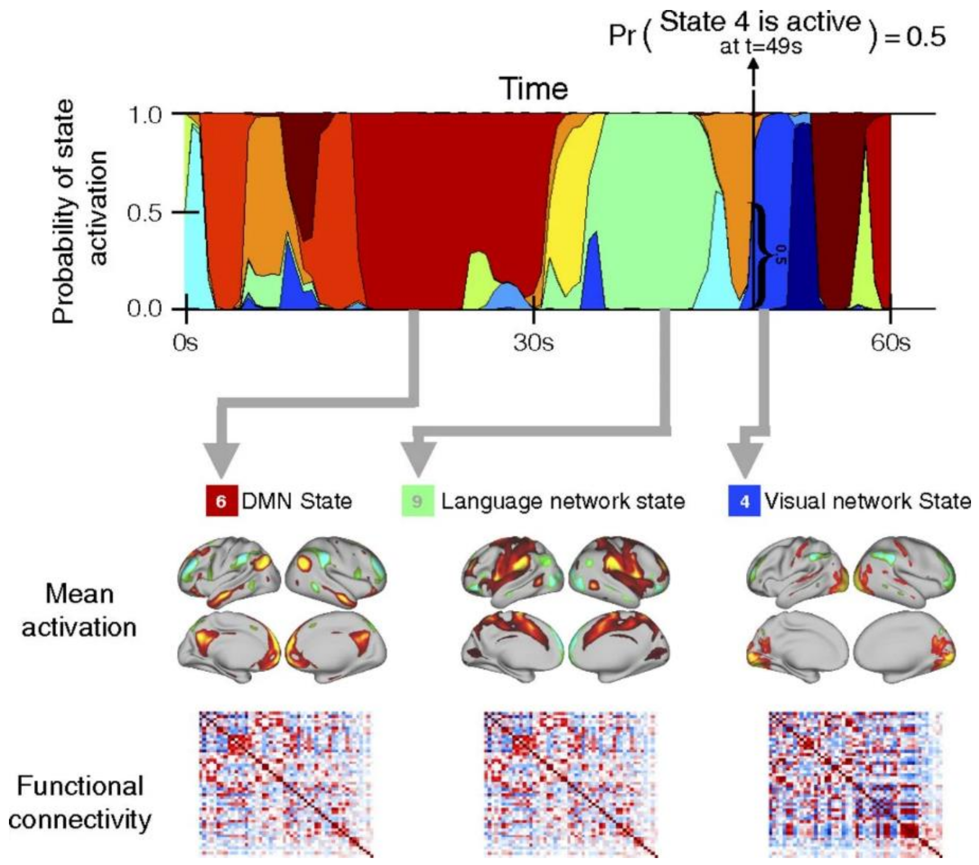
GLHMM toolbox: Out-of-sample prediction

Christine Ahrends, PhD

Analysis workshop CFIN – 26.05.2025

How can we predict an individual's traits or behaviour from the patterns in which their brain function changes over time?

Challenge: HMMs for phenotype prediction



Model of brain dynamics (Hidden Markov Model)

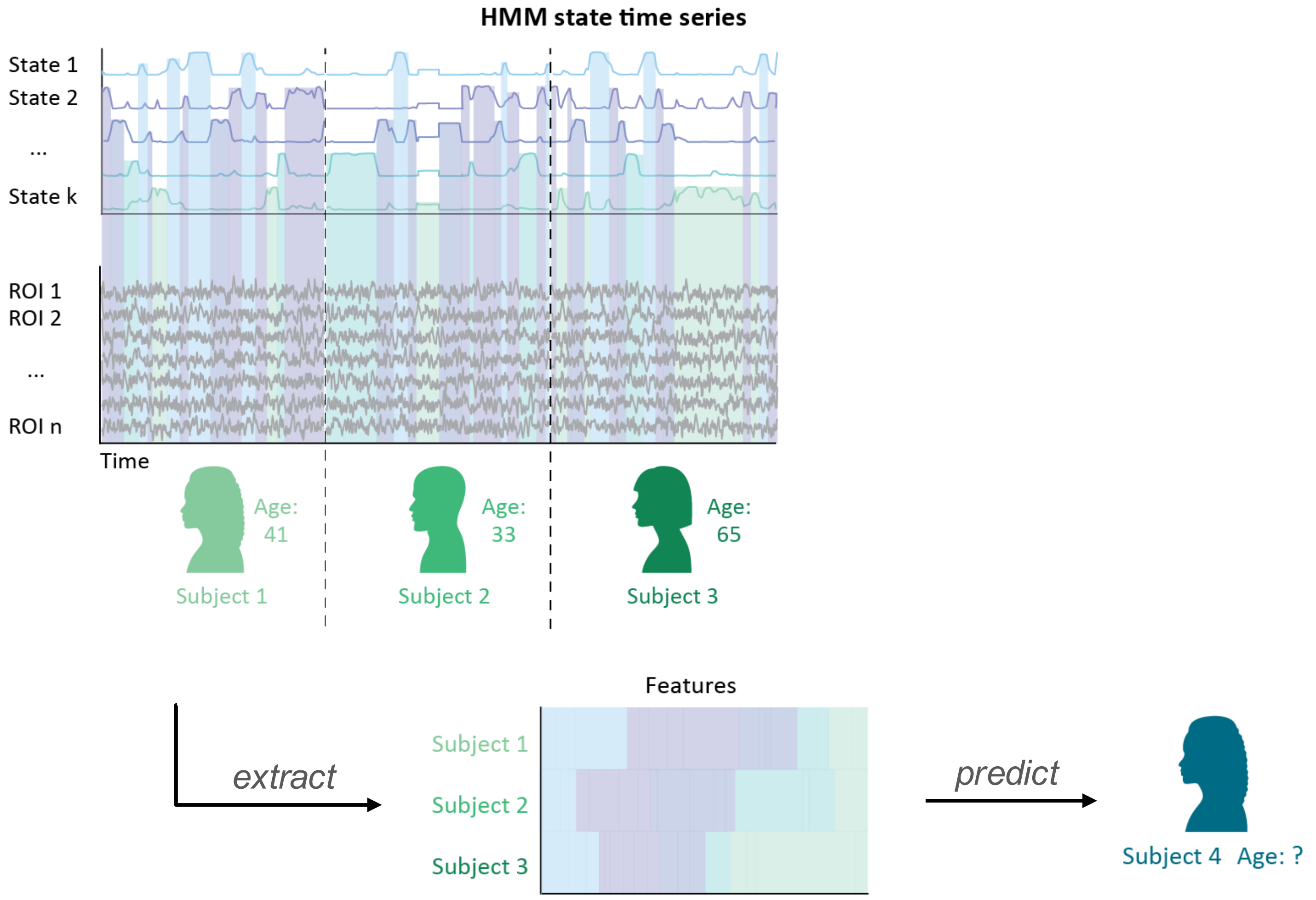
Age?

Summary features

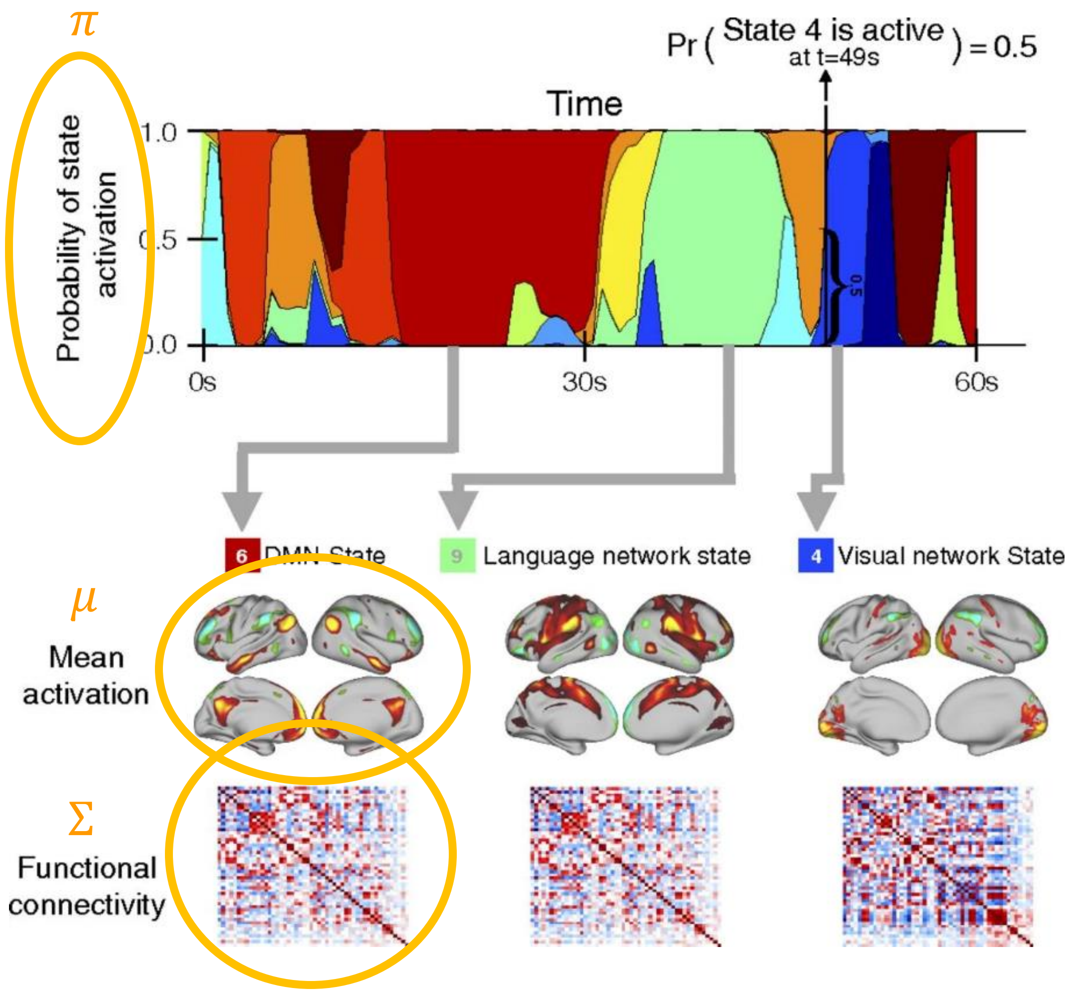
- **Feature engineering approach:**
- Extract features of interest from HMM
    - Fractional occupancies
    - Switching rates
    - State lifetimes
    - …
- Use features to predict individual traits/performance/...

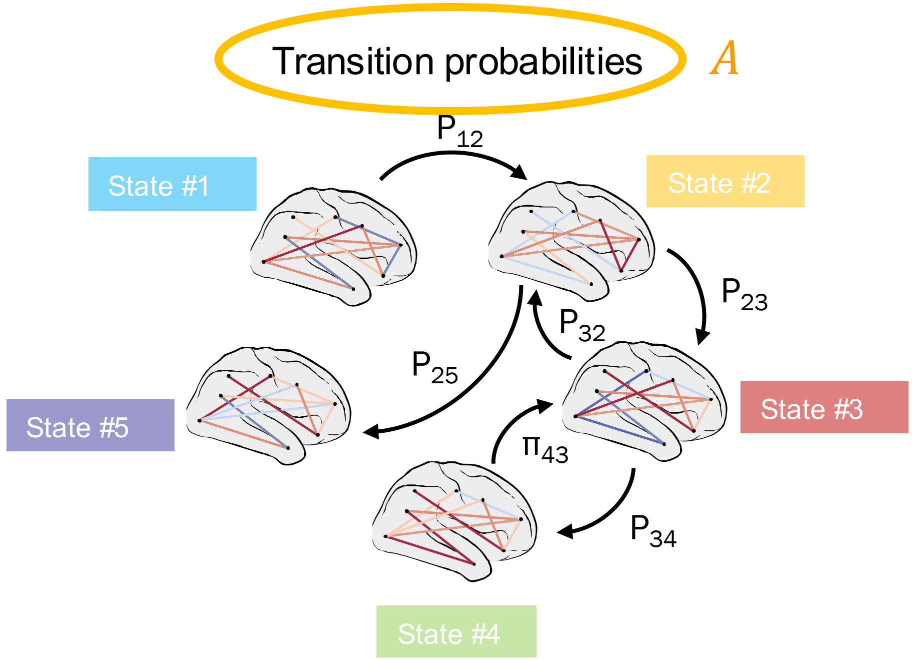
\+ Hypothesis-driven

\- Losing information



Challenge: HMMs for phenotype prediction



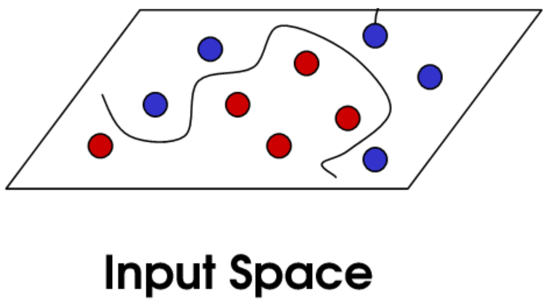


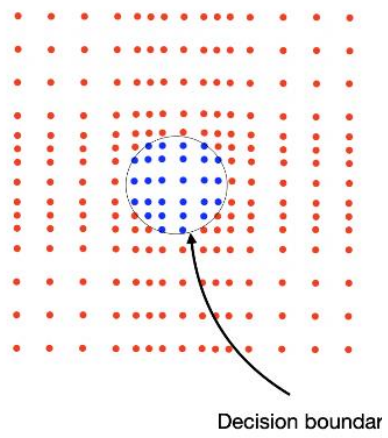
HMM parameters:

$$\theta = [\pi, A, \mu, \Sigma]$$

→ **construct kernel**

Kernel methods

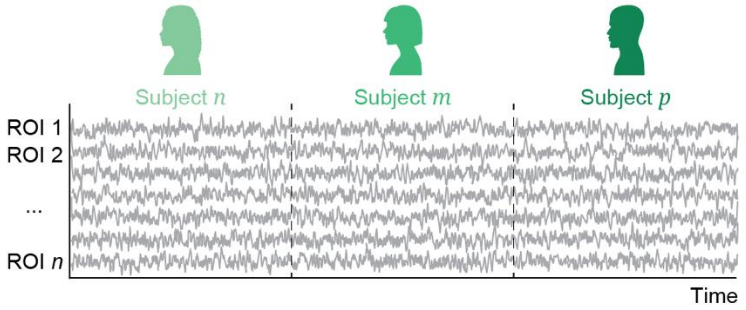




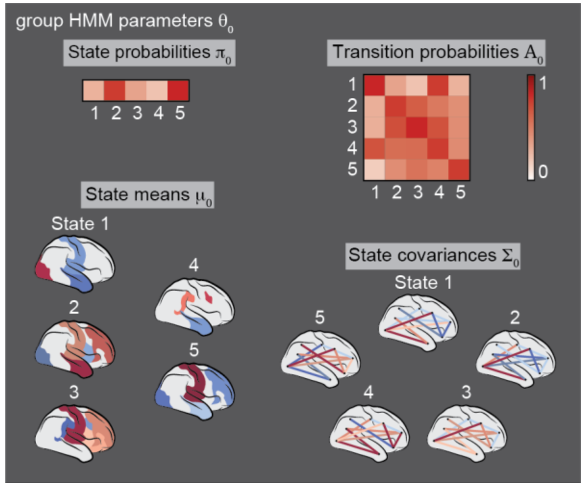
- Find nonlinear decision boundaries using linear model
    - Transform input space into another embedding space
    - Find linearly separating hyperplane in embedding space
- Kernel trick: use kernel function instead of each data point in embedding space
- Similarity functions (e.g. similarity between pair of subjects)
- Computationally efficient, no feature selection

Kernels from HMM: General approach

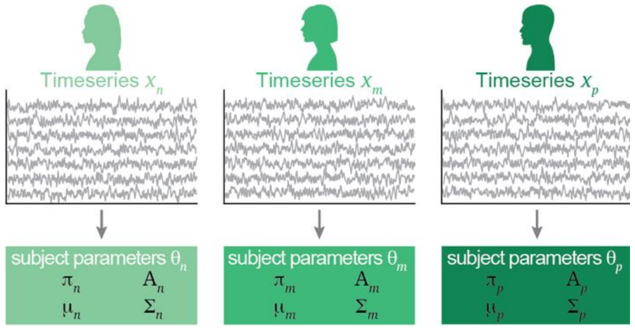
1. Group neuroimaging timeseries (e.g. fMRI, MEG, EEG)



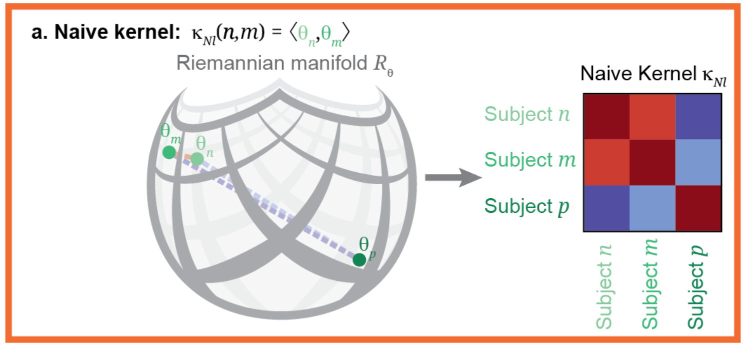
2. Estimate Hidden Markov model (group-level HMM)

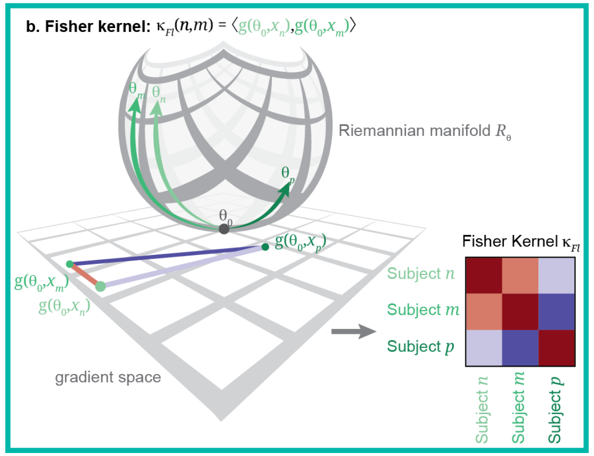


3. Dual estimation (subject-level HMM)



4. Map examples into feature space and construct kernel



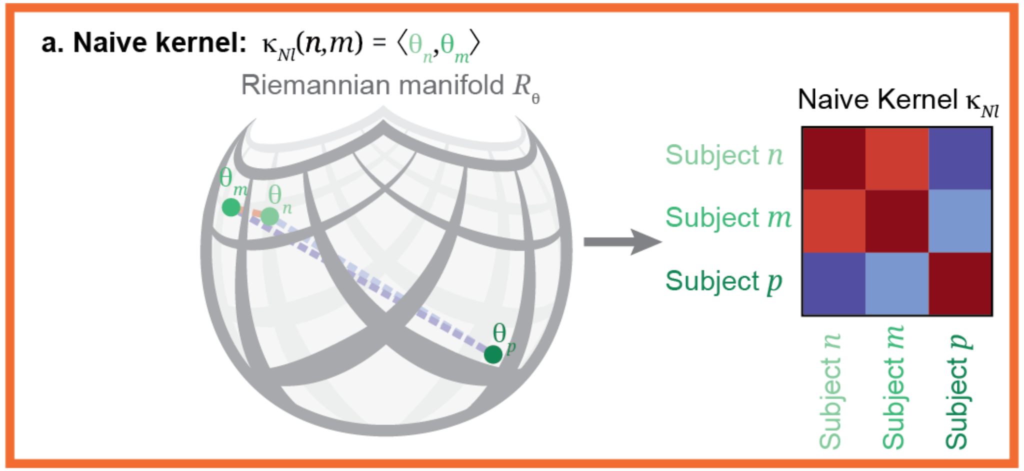


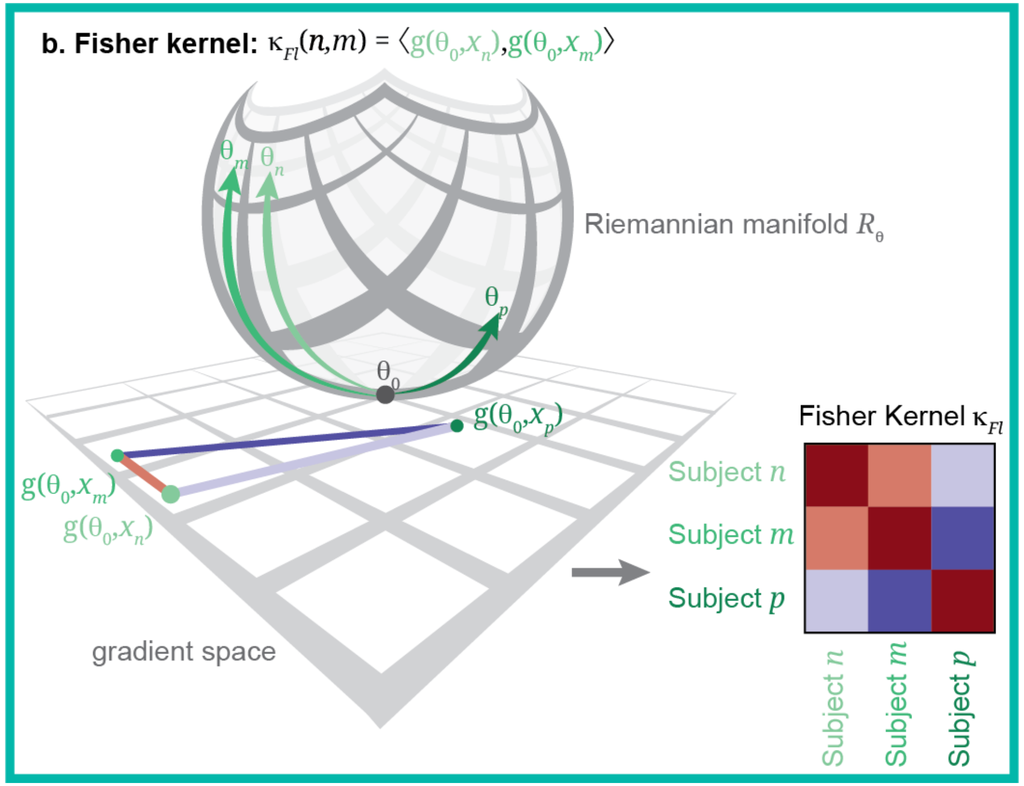
5. Kernel-based prediction

	Age	Item A	Item B	Item C	Item D
Subject n	28	102	115	5	87
Subject m	32	110	113	4	33
Subject p	54	99	125	6	18

Ahrends, Woolrich & Vidaurre (elife, 2025)

Kernels from HMM: General approach





Ahrends, Woolrich & Vidaurre (elife, 2025)

The Fisher Kernel

- Fisher score: gradient of the log-likelihood w.r.t. each model parameter:

$$g(\theta^0, x^n) = \left(\frac{\partial \log \mathcal{L}_{\theta^0}(x^n)}{\partial \theta^0}\right) \qquad g \in \mathbb{R}^{1x(K+K*K+K*M+K*M*M)}$$

- how we need to change the group-level model to better explain an individual's timeseries

- (Practical) linear Fisher kernel:

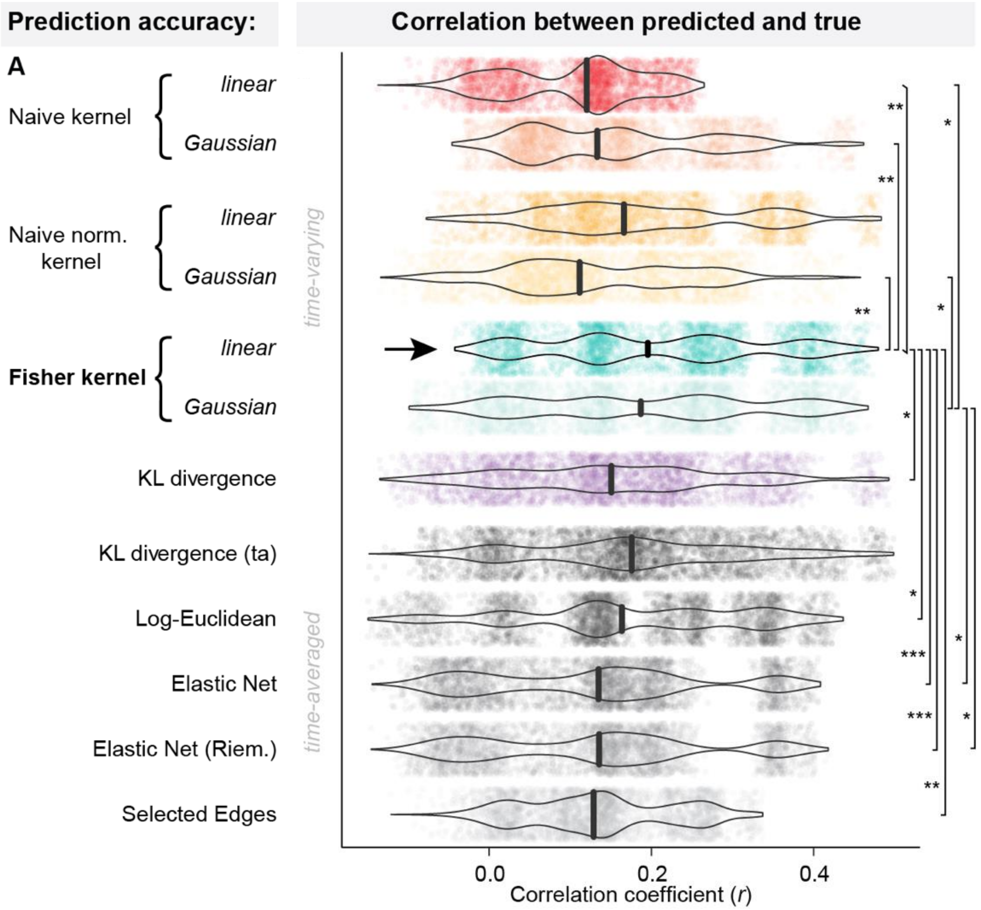
$$\kappa_F(n,m) = g(\theta^0, x^n)^{\mathrm{T}} g(\theta^0, x^m) \qquad \kappa_F \in \mathbb{R}^{NxN}$$

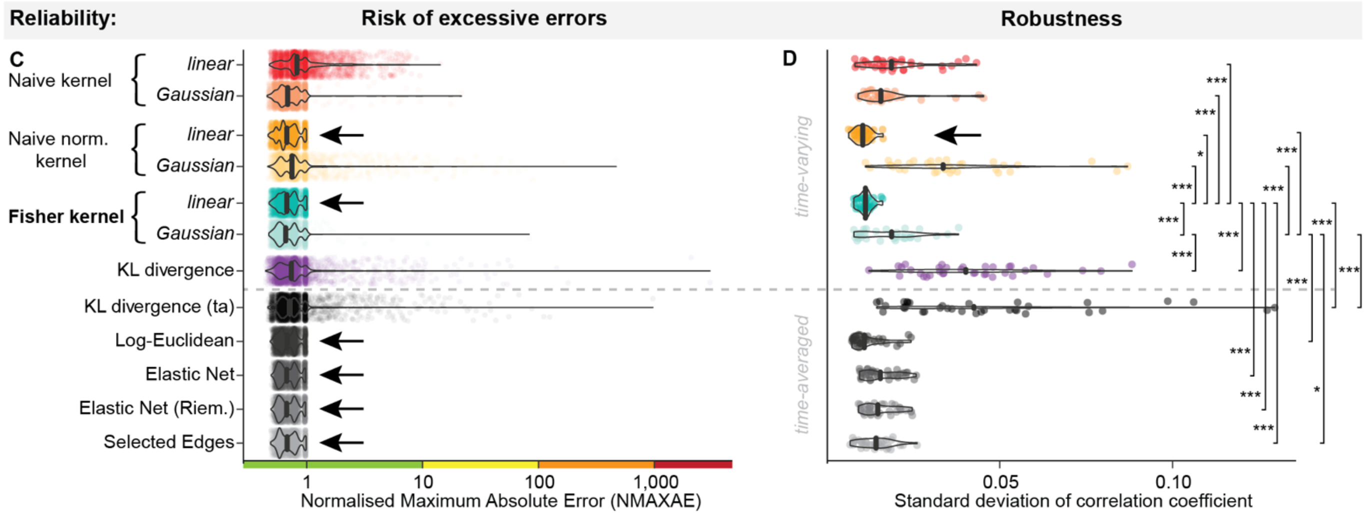
- Gaussian Fisher kernel:

$$\kappa_{Fg}(n,m) = \exp(-\frac{\left\|g(\theta^0,x^n)-g(\theta^0,x^m)\right\|^2}{2\tau^2}) \qquad \kappa_{Fg} \in \mathbb{R}^{NxN}$$

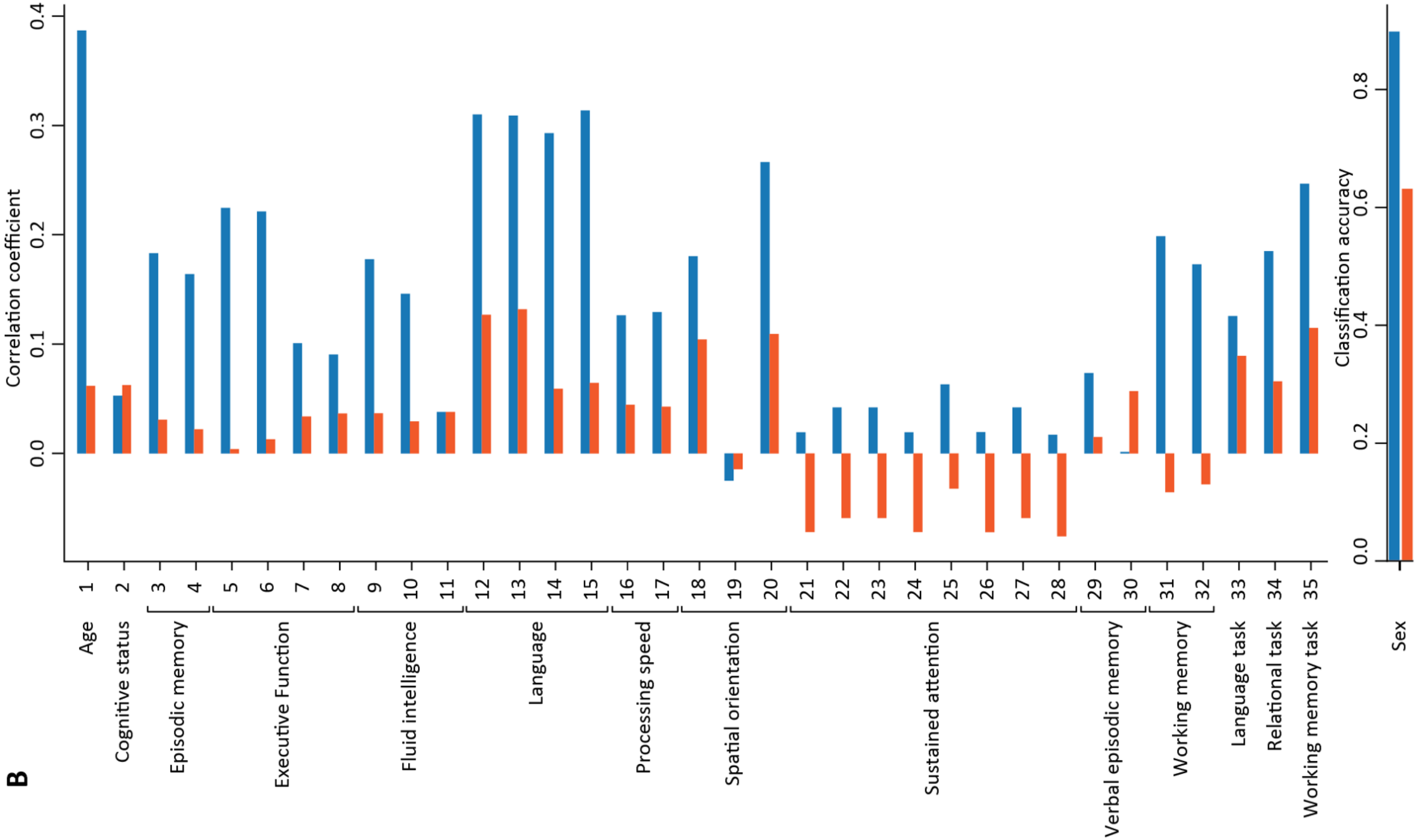
Jaakkola & Haussler (1998) *NIPS*
Jaakkola et al. (1999) *PICISMB*

Fisher kernel: Accuracy and reliability





Kernels vs. feature engineering



Fisher kernel

Summary features

Try it yourself:

- Notebooks:
  - Gaussian HMM `2_1_GaussianHMM.ipynb`
  - Out-of-sample prediction `2_2_Predction.ipynb`

Additional reading:
Ahrends, Woolrich & Vidaurre (*elife*, 2025) Predicting individual traits from models of brain dynamics accurately and reliably using the Fisher kernel.

Code:
https://github.com/ahrends/FisherKernel