Using convex feature selection to improve offline movement decoding

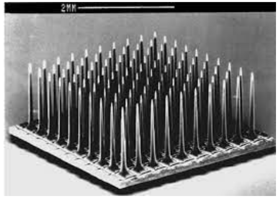
Lauren Peterson, Si Jia Lee, Leo Scholl, Amy Orsborn

Introduction

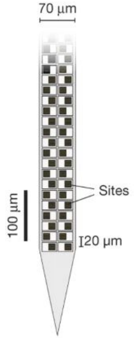
- Graduated in 2022 from the University of Washington in Electrical Engineering
- Part of the A3D3 institute as a postbac researcher
- Currently work with Dr. Amy Orsborn to improve real-time neuroscience computations
- Will start my PhD in Electrical & Computer Engineering at Rice in August

We can record a lot of neural signals

Utah Array (1997)
96 electrodes

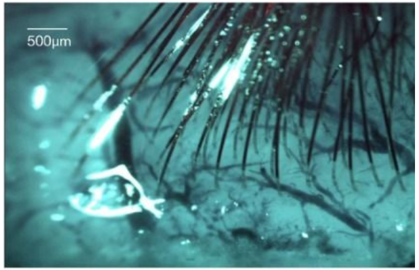


Neuropixels
960 sites (2017)

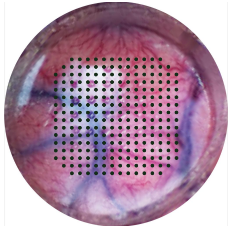


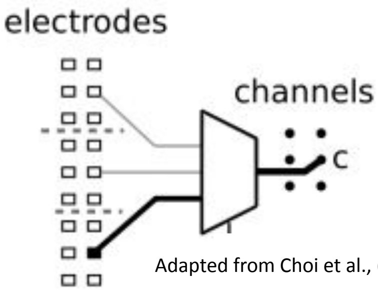
Surface
Electromyography
32 to 256
electrodes

Neuralink (2019)
3074 sites



Electrocorticography
244 Electrodes

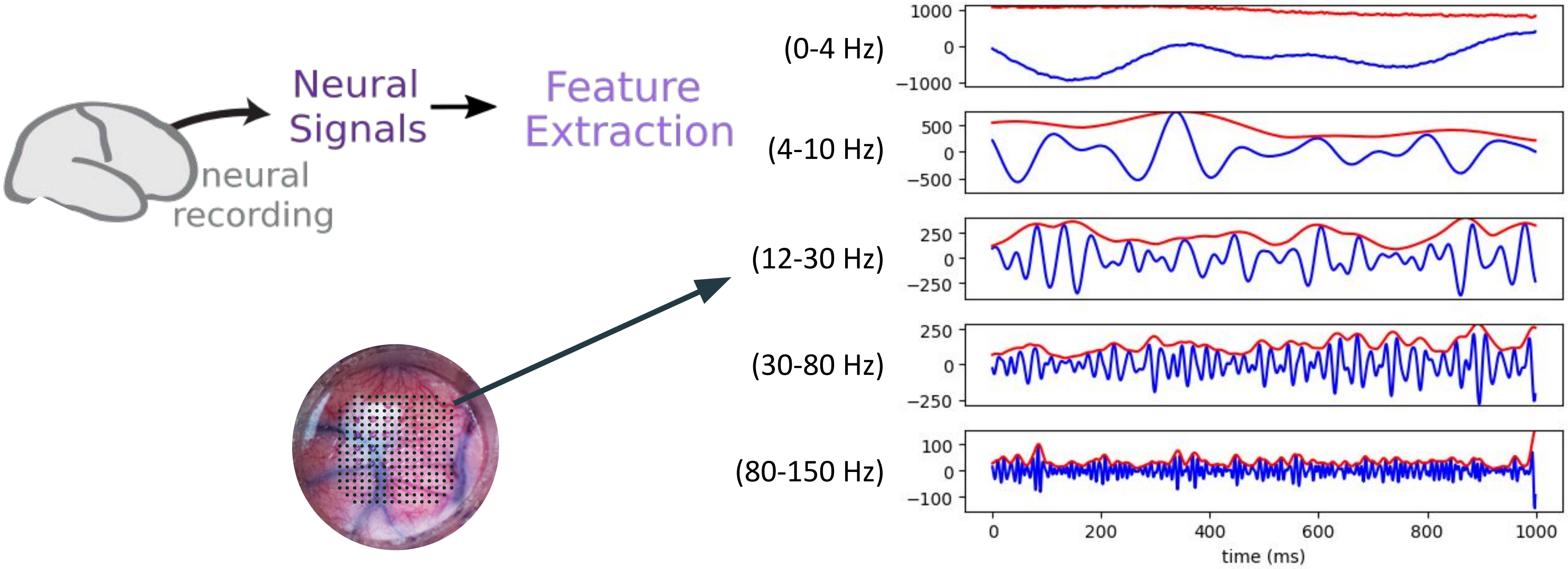




Adapted from Choi et al., (2020)

Need to select a small set of electrodes for streaming

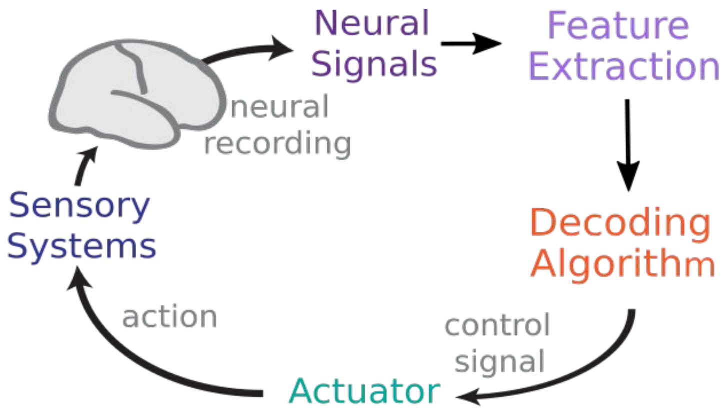
The number of neural features multiplies



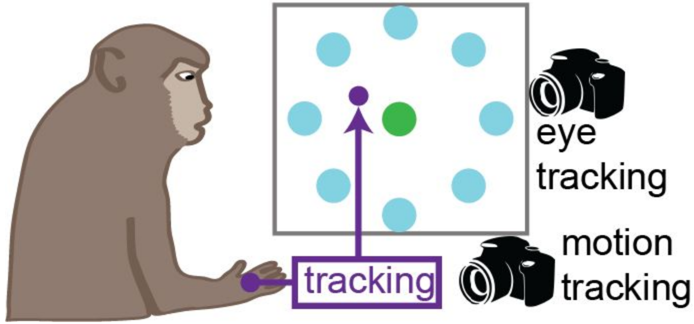
e.g. 236 electrodes x 5 frequency bands = 1180 features

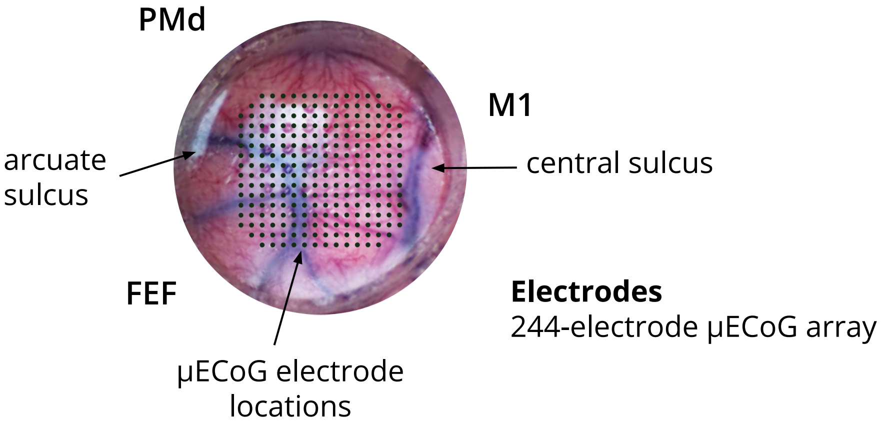
Adaptive Feature Selection

- Can we adapt to a BCI’s user learning strategy in trying out different features?
  - Brain engages in **sparsification** of learning
  - Could we free up features for other tasks?
- Goal: develop an online adaptive feature selection scheme for neural interface learning
  - Relevance
  - Sparsity
  - Smoothness



Movement Task





Subject: 1 NHP rhesus macaque

Data: continuous time, broadband, sampled at 25kHz

Kalman Filter

State-transition:

$$x_t = \boxed{A}x_{t-1} + \boxed{w_t}$$

State transformation

State covariance

$$w_t \sim N(0, \boxed{W})$$

Observation-model:

$$y_t = \boxed{C}x_{t-1} + \boxed{q_t}$$

Observation transformation

Observation covariance

$$q_t \sim N(0, \boxed{Q})$$

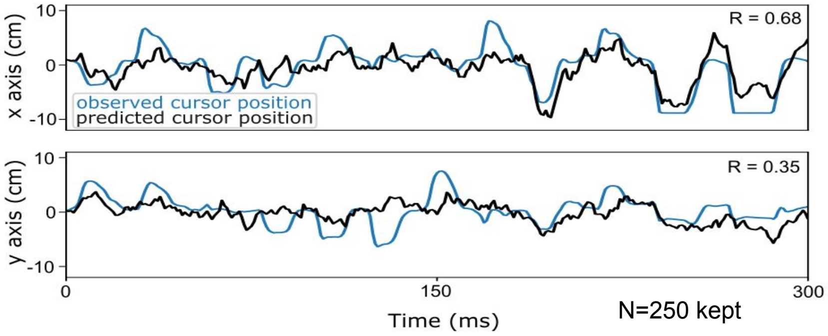
Convex Feature Objective

- We introduce a smoothness component to the objective
- Existing convex objective algorithms also can’t work in real time (LASSO)

$$\min_{\theta_i} \; \underbrace{-\mathbf{logdet}(C_i^{T} Q^{-1} \Theta_i C_i)}_{\text{relevance}} + \underbrace{\lambda \theta_i^{T} \mathbf{1}}_{\text{sparsity}} - \underbrace{\mu \theta_i^{T} [\, \theta_{i-1} \dots \theta_{i-k} \,][\, p \dots p^{k} \,]^{T}}_{\text{smoothness}}$$

$$\text{s.t. } 0 \leq \theta_i \leq 1$$

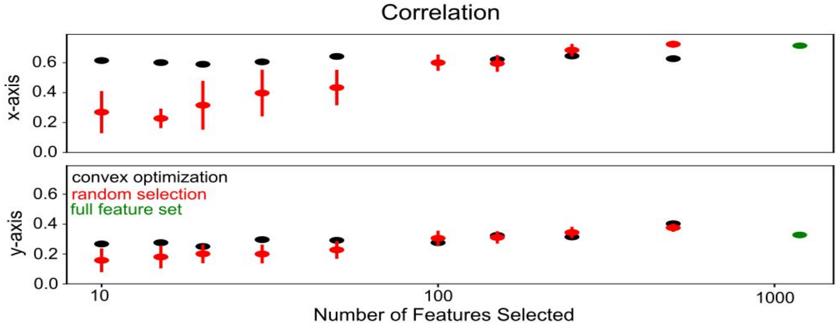
Correlation Analysis



N=250 kept

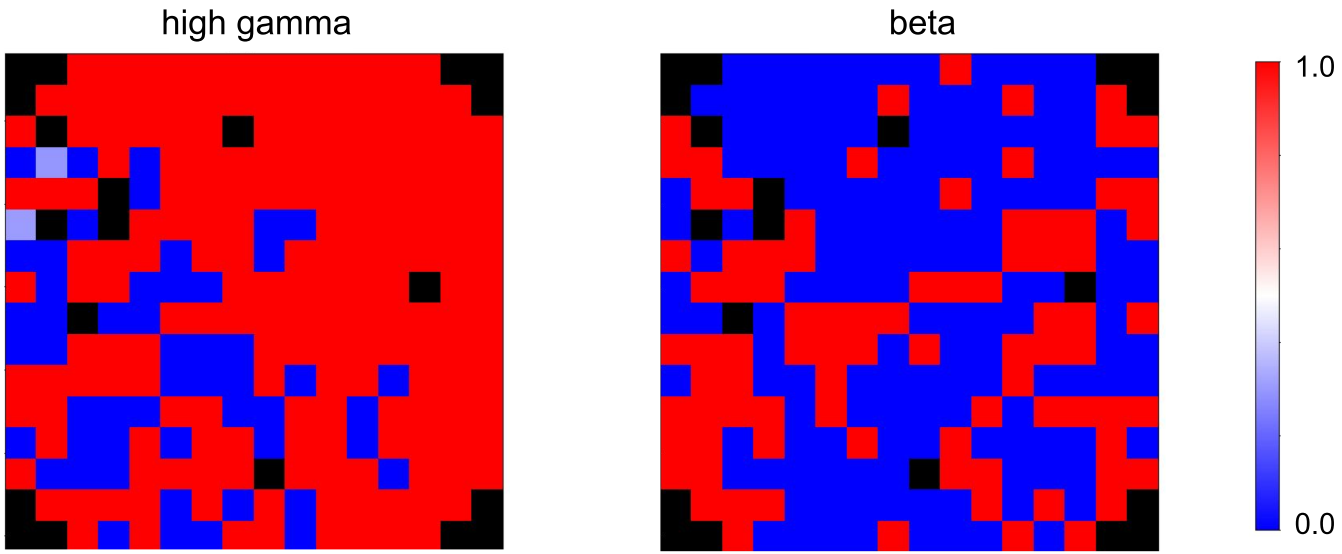
Removing a subset of the initial feature set does not decrease decoder performance

The algorithm removes redundant features and out-performs a random selection at smaller feature subsets



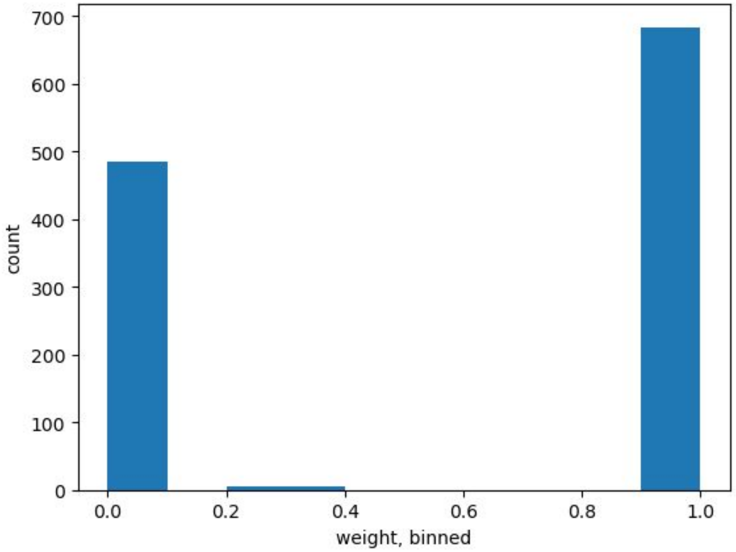
Which features were selected?

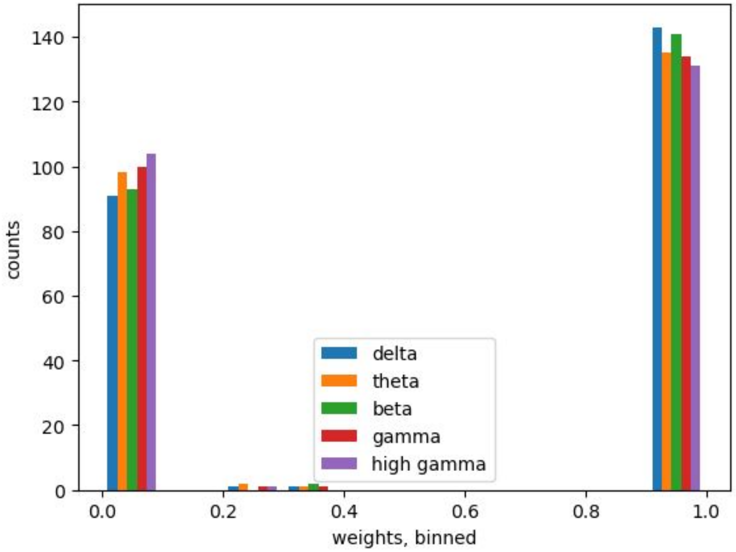
- Anecdotally, choosing many features from high gamma (80-150 Hz) and fewer features from beta (12-30 Hz)



n features = 1180, n kept = 590

Features by frequency bands



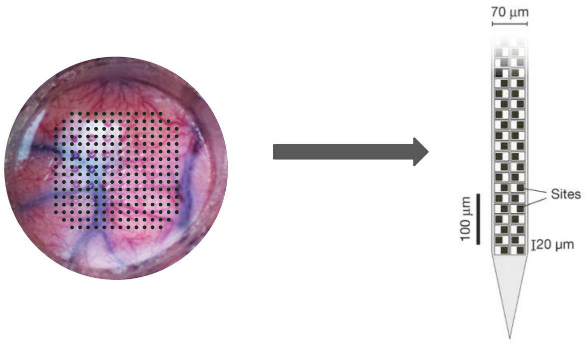


However, on a closer look, the number of features selected appears to be consistent across the frequency bands.

n features = 1180, n kept = 590

Next steps…

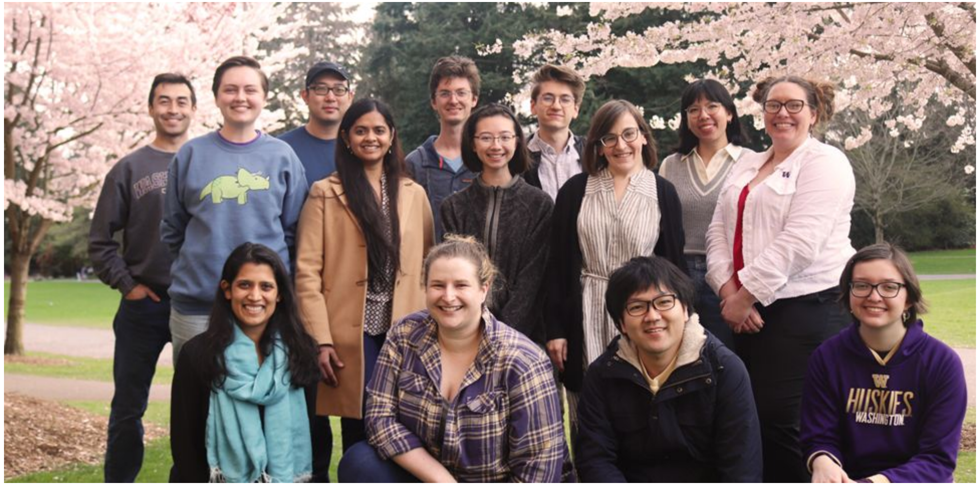
- This was an offline verification of the algorithm, with one dataset from one subject
- We will expand this analysis to another dataset (neuropixels) and another subject
- Currently working to verify results for paper submission



A3D3 Feedback

- An amazing opportunity, and the potential for collaboration is high
- Sometimes I felt out of my depth because of the emphasis on physics, but everything is so new that no group was too insulated either
- I was able to do research full-time, and determine if I liked it enough to continue
- Some constructive feedback: I would appreciate knowing about program events, and the responsibilities associated with them, further in advance (everything is so new that very little is previously established)

Acknowledgements



Orsborn Lab
Dr. Amy Orsborn
Dr. Lydia Smith, Dr. Leo Scholl
Si Jia Li, Pavithra Rajeswaran,
Maneeshika Madduri,
Augusto Millevolte, Tomohiro Ouchi,
Ryan Canfield, Katherine Perks,
Sasha Burchardt

Thanks to A3D3 for funding my work, and Pavi, Leo, and Lydia for collecting the data.