

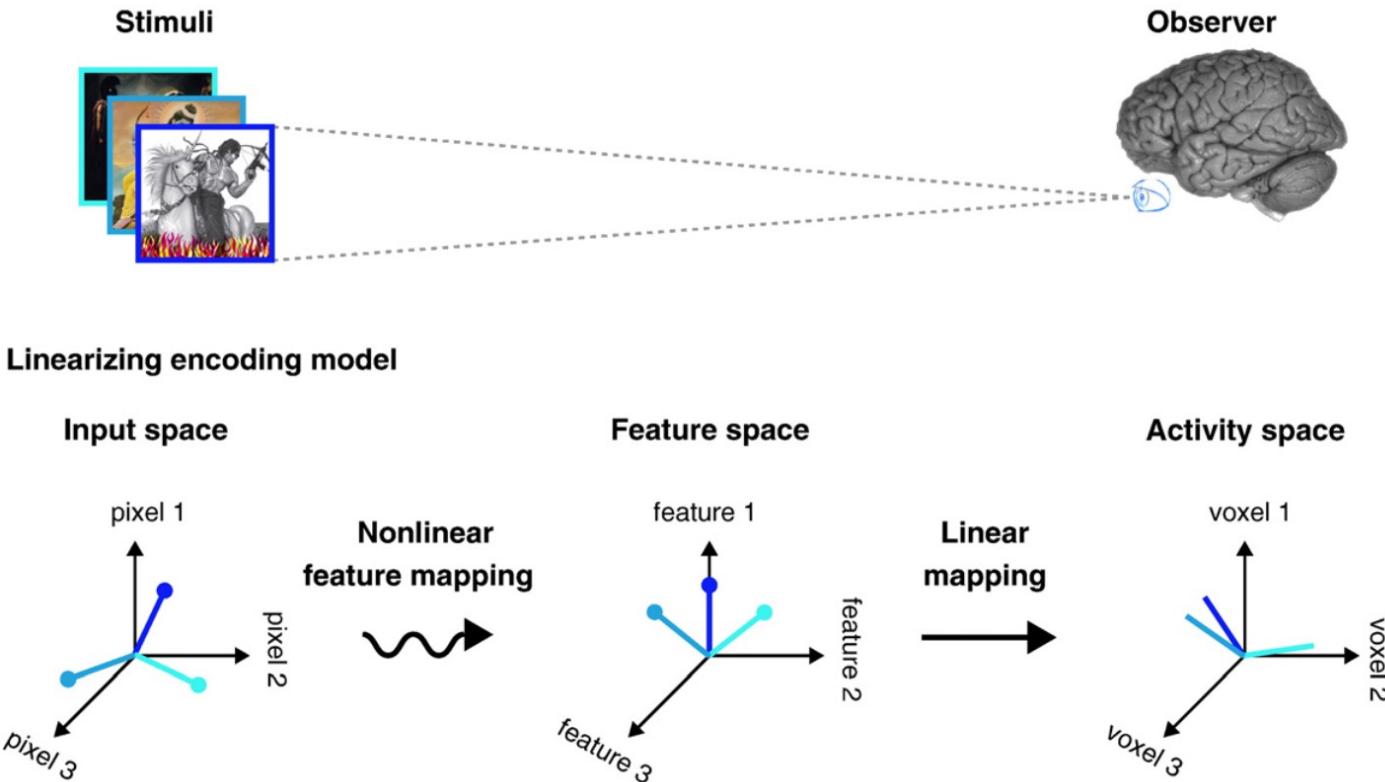
Voxelwise encoding modeling tutorial

Dasom Kwon
2024. 07. 19.

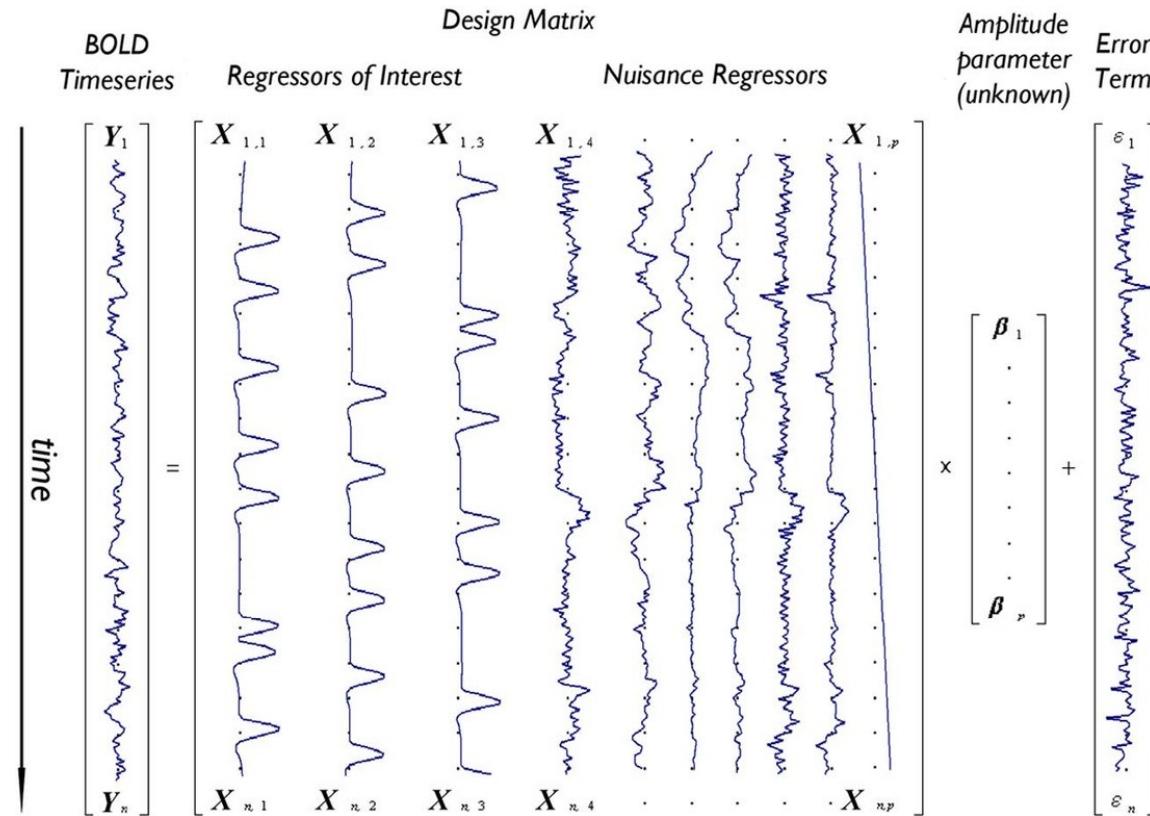
References

1. Dupré la Tour, T., Eickenberg, M., Nunez-Elizalde, A. O., & Gallant, J. L. (2022). Feature-space selection with banded ridge regression. *NeuroImage*, 264, 119728. <https://gallantlab.org/himalaya/>
2. Gao, J. S., Huth, A. G., Lescroart, M. D., & Gallant, J. L. (2015). Pycortex: an interactive surface visualizer for fMRI. *Frontiers in neuroinformatics*, 9, 23. <https://gallantlab.org/pycortex/>
3. Dupré la Tour, T., Visconti di Oleggio Castello, M., & Gallant, J. L. (2024). The Voxelwise Modeling framework: a tutorial introduction to fitting encoding models to fMRI data. https://gallantlab.org/voxelwise_tutorials/
4. Huth, A. G., De Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600), 453-458.
5. Kumar, S., Sumers, T. R., Yamakoshi, T., Goldstein, A., Hasson, U., Norman, K. A., ... & Nastase, S. A. (2024). Shared functional specialization in transformer-based language models and the human brain. *Nature Communications*, 15(1), 5523.

Encoding model?

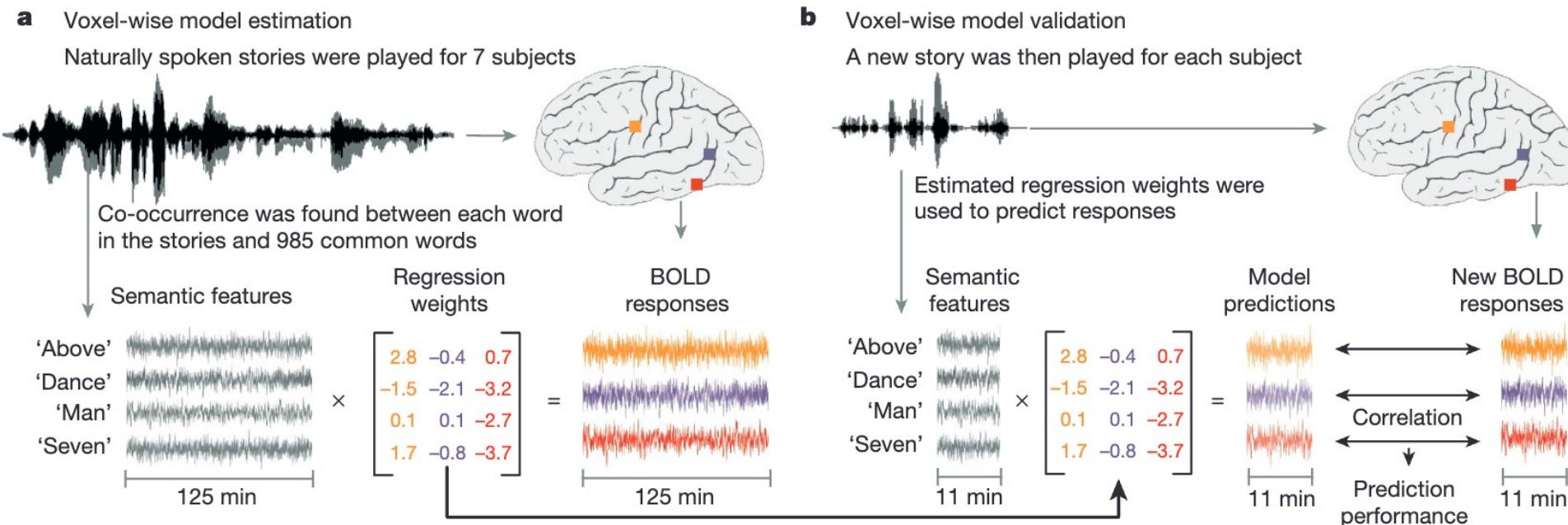


Encoding model? GLM?



Difference: Brain to features / Brain to conditions

Encoding modeling

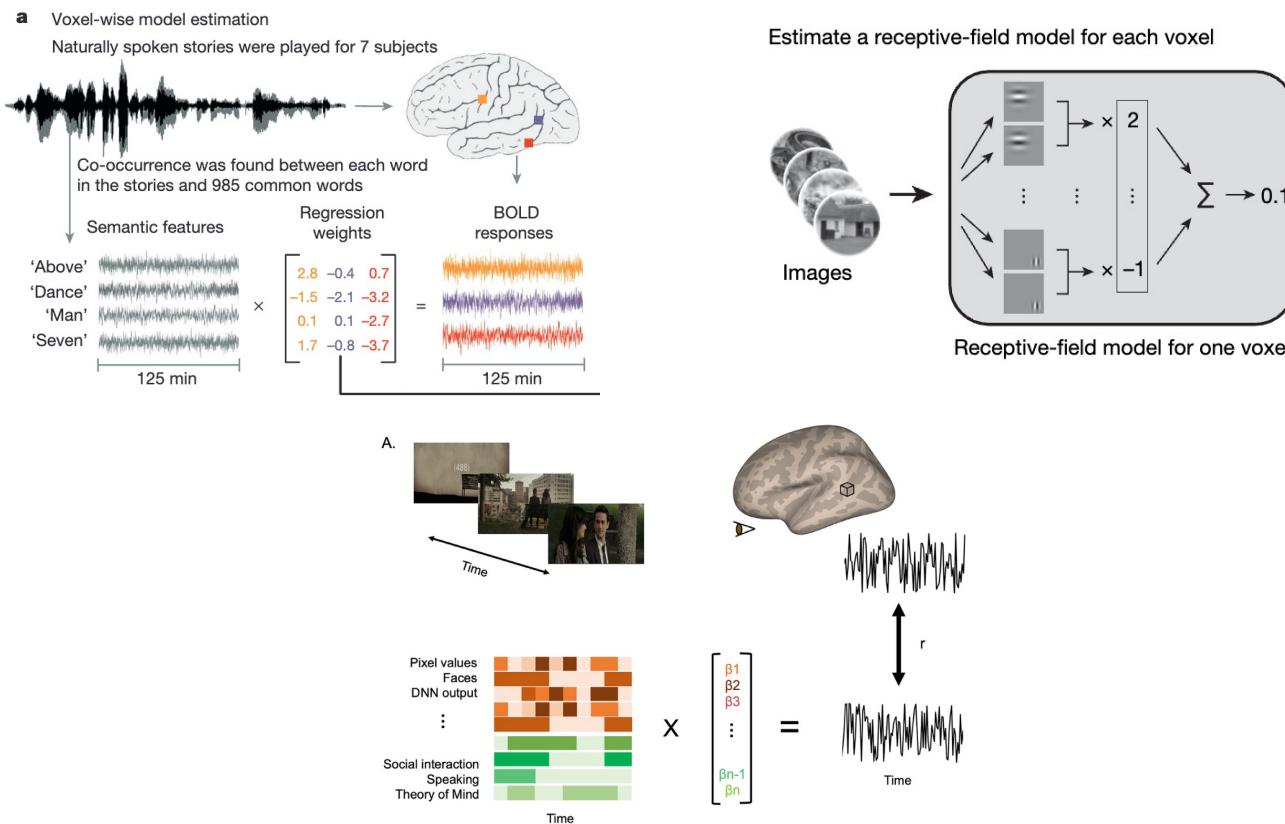


Encoding modeling

1. Extracting features
2. Training model (cross-validation)
3. Selecting an algorithm
4. Assessing model performance

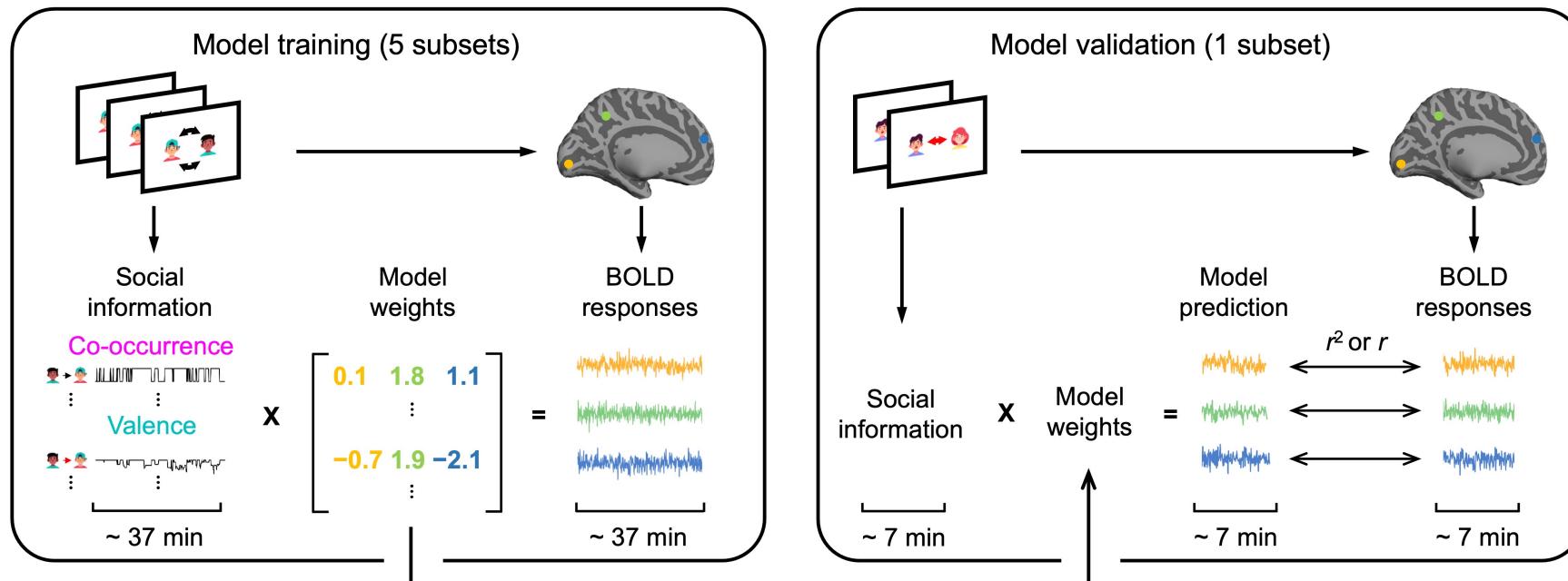
Encoding modeling

1. Extracting features



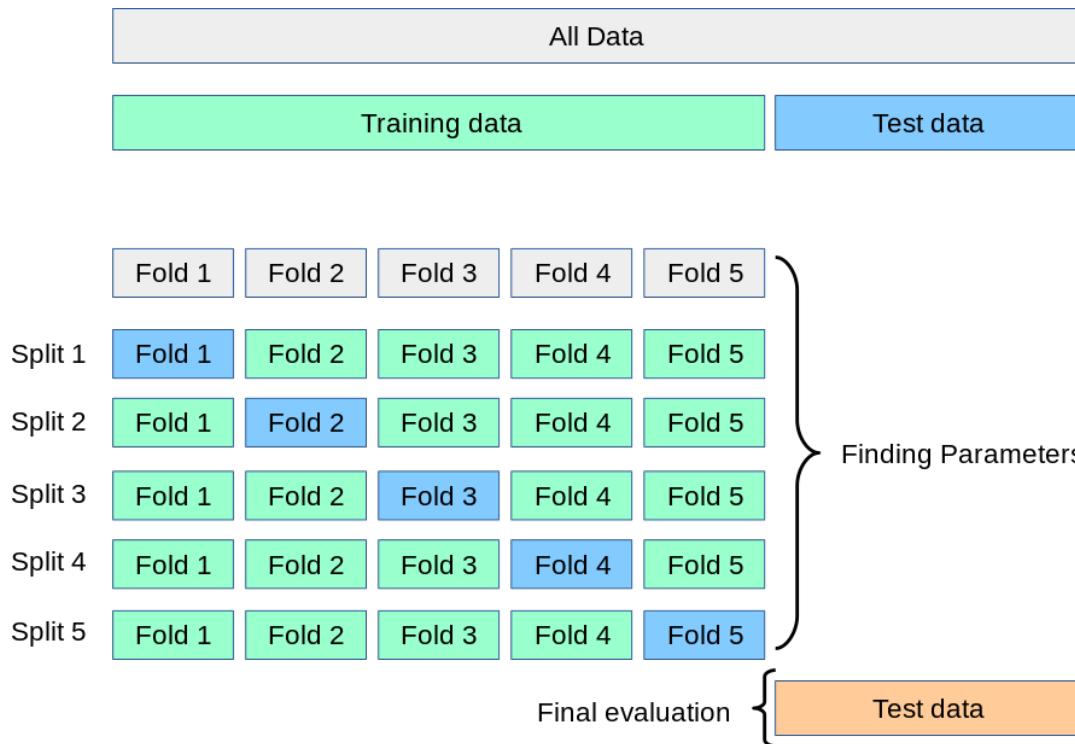
Encoding modeling

2. Training model (cross-validation)



Encoding modeling

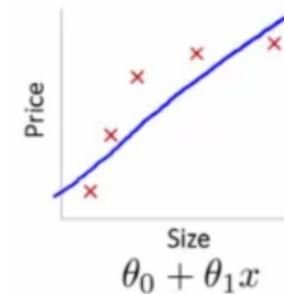
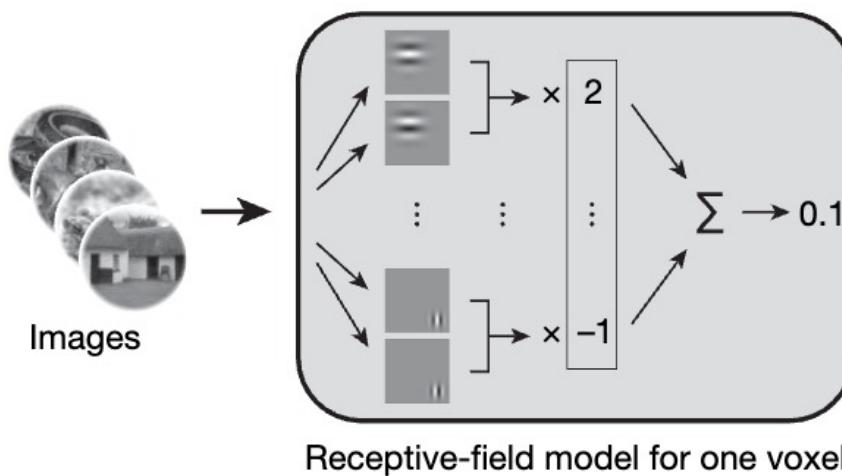
2. Training model (cross-validation)



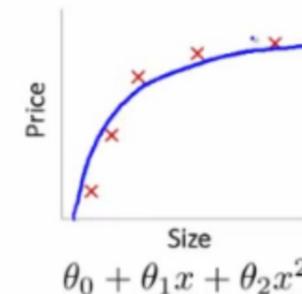
Encoding modeling

3. Selecting an algorithm

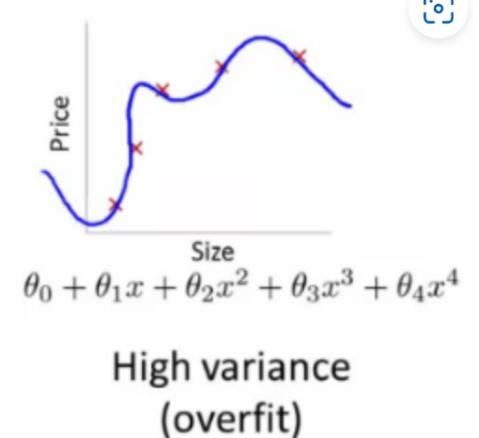
Estimate a receptive-field model for each voxel



High bias
(underfit)



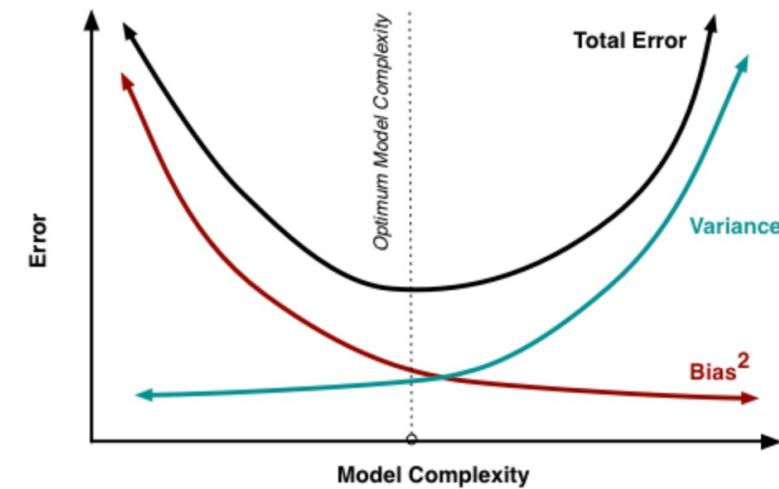
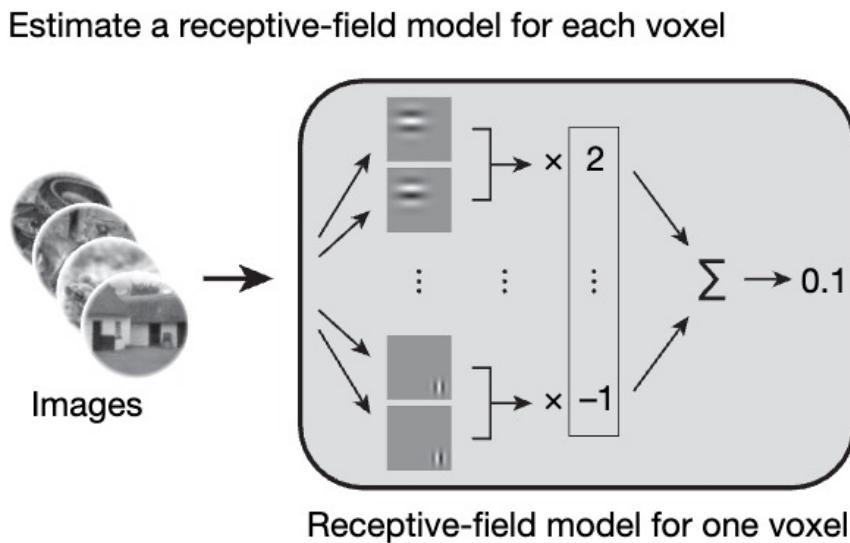
"Just right"



High variance
(overfit)

Encoding modeling

3. Selecting an algorithm



Encoding modeling

3. Selecting an algorithm

Ridge regression

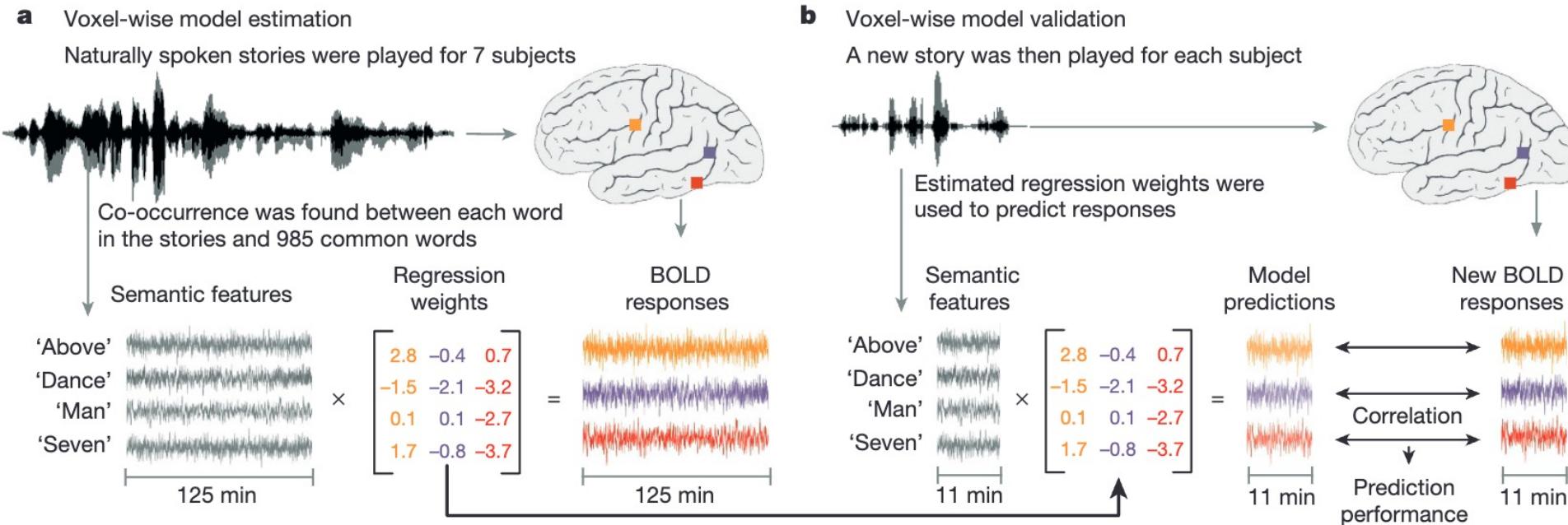
$$SSE_{L_2} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^P \beta_j^2$$

Hyperparameter

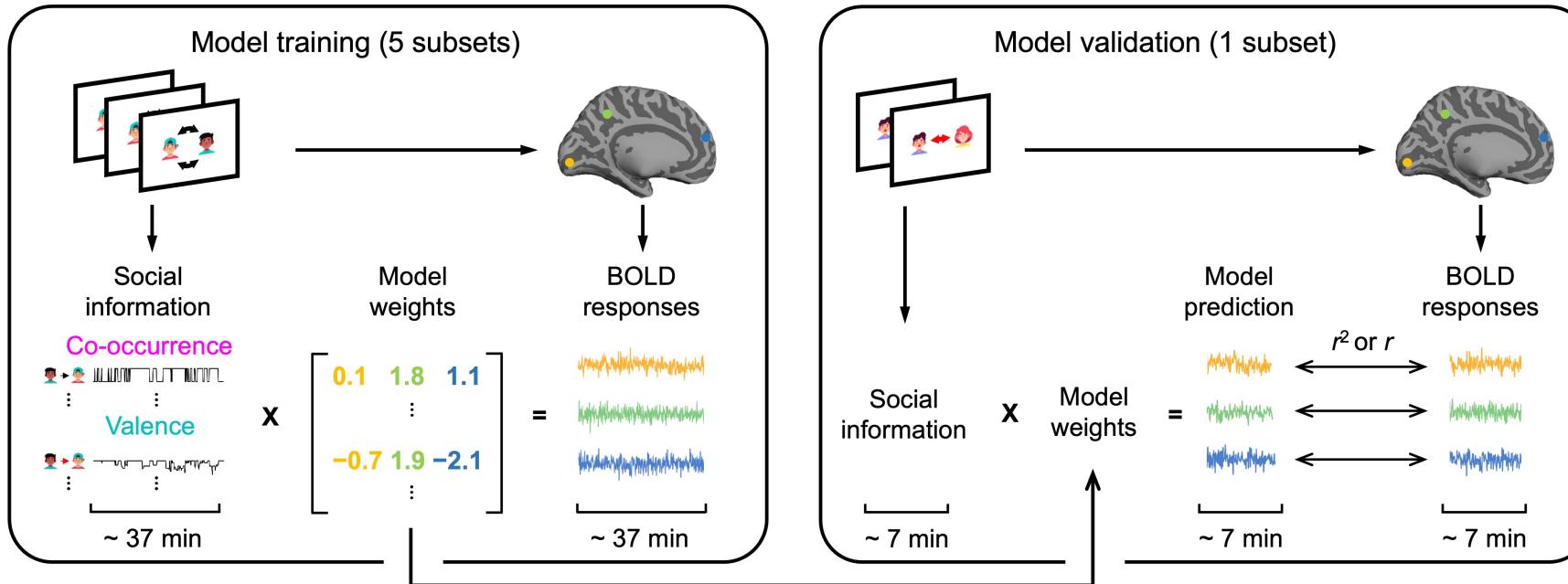
Why not use ridge regression in GLM?

Encoding modeling

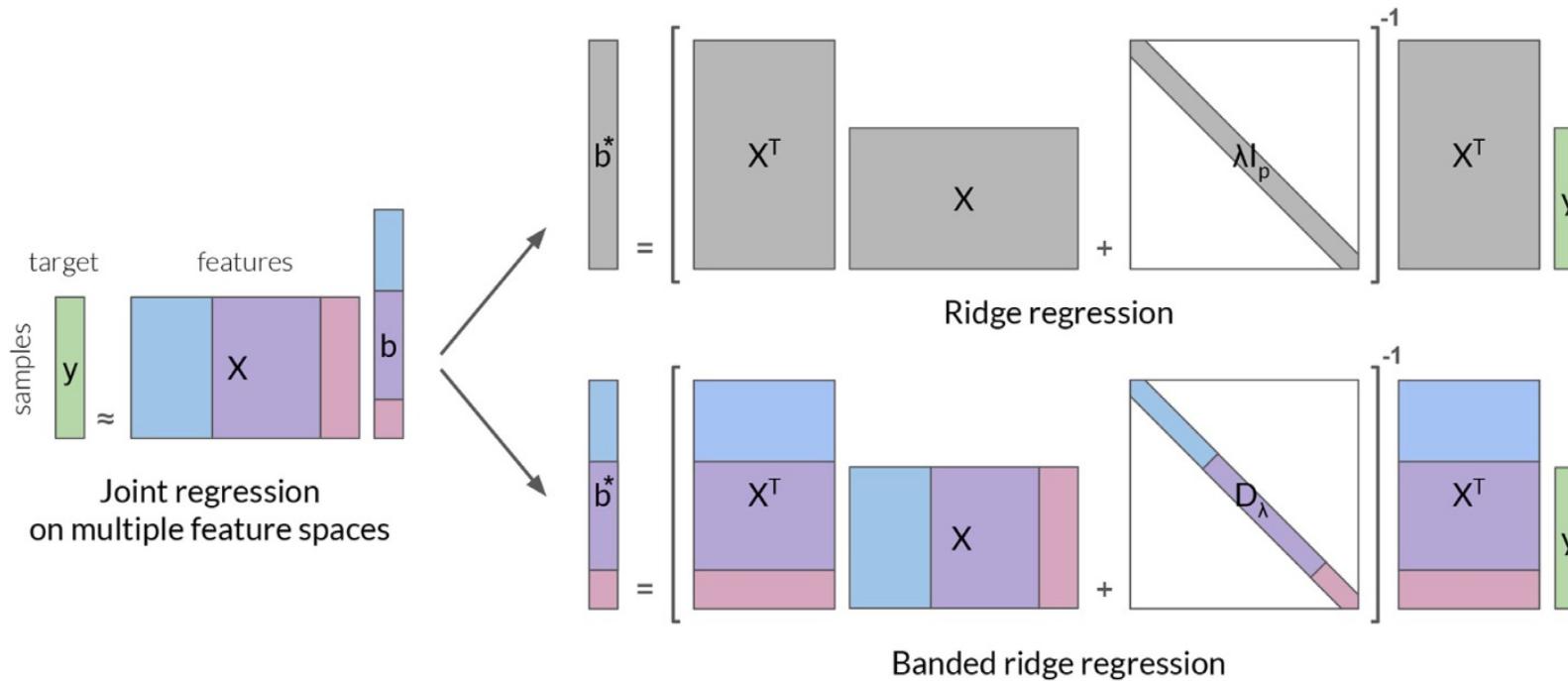
4. Assessing model performance



Banded ridge regression

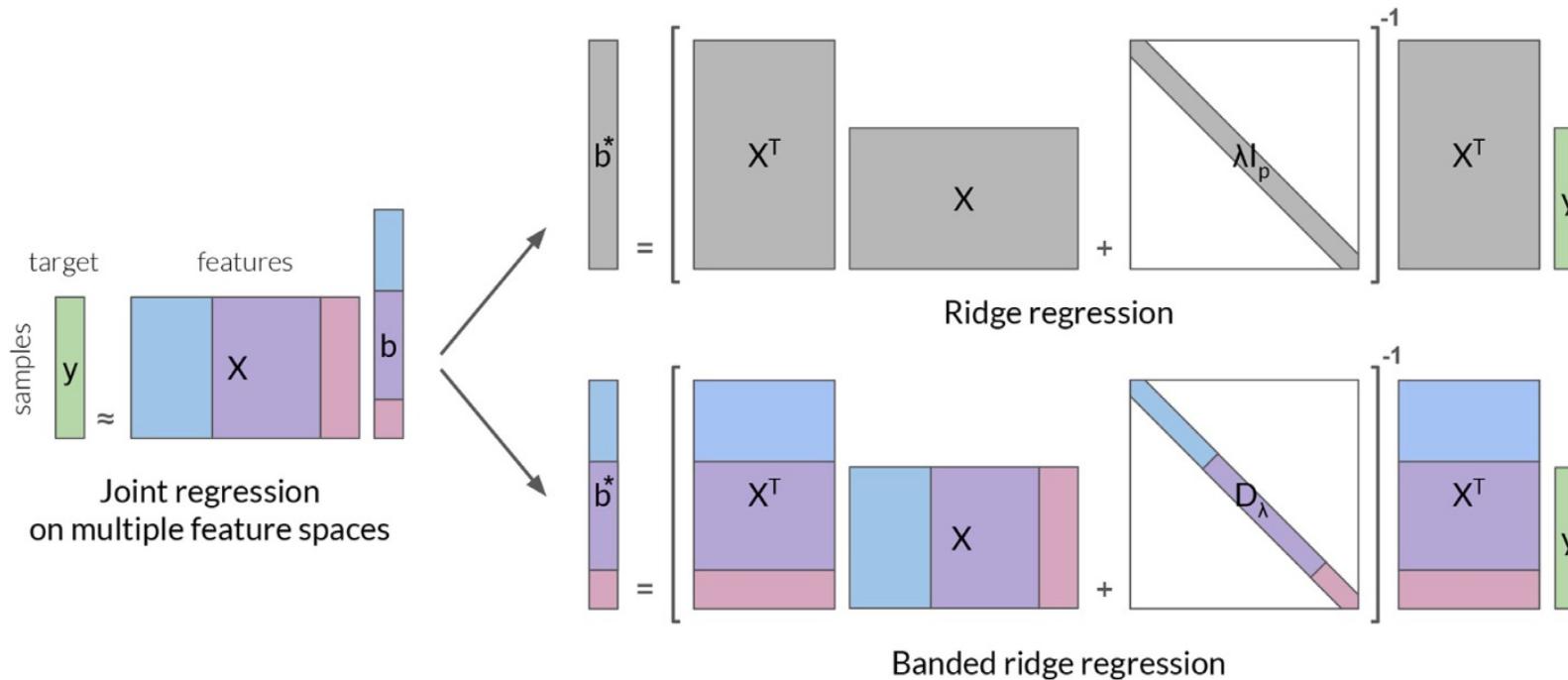


Banded ridge regression



Adapting the regularization hyperparameter for each feature space.

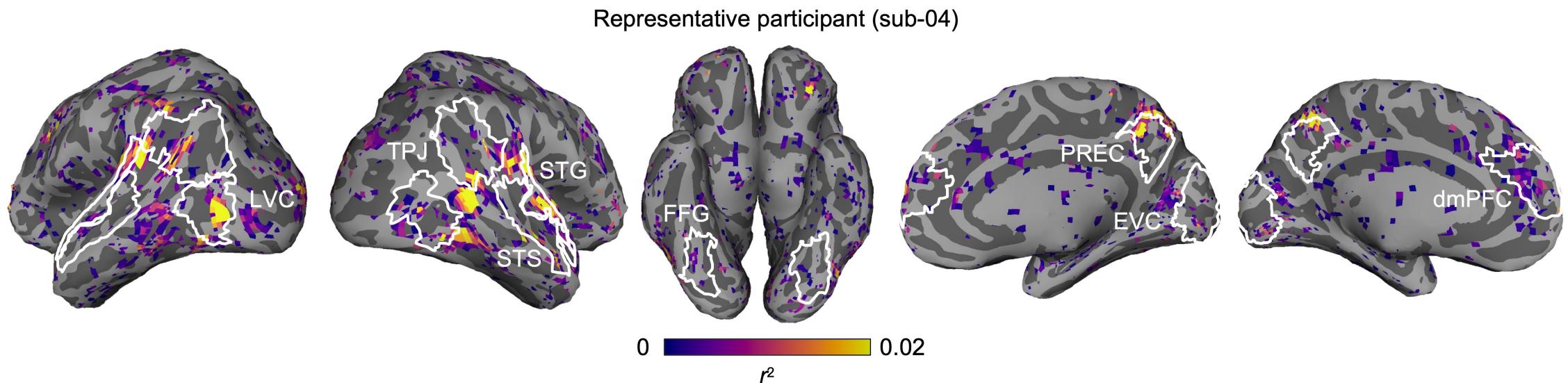
Banded ridge regression



A feature-space selection mechanism?

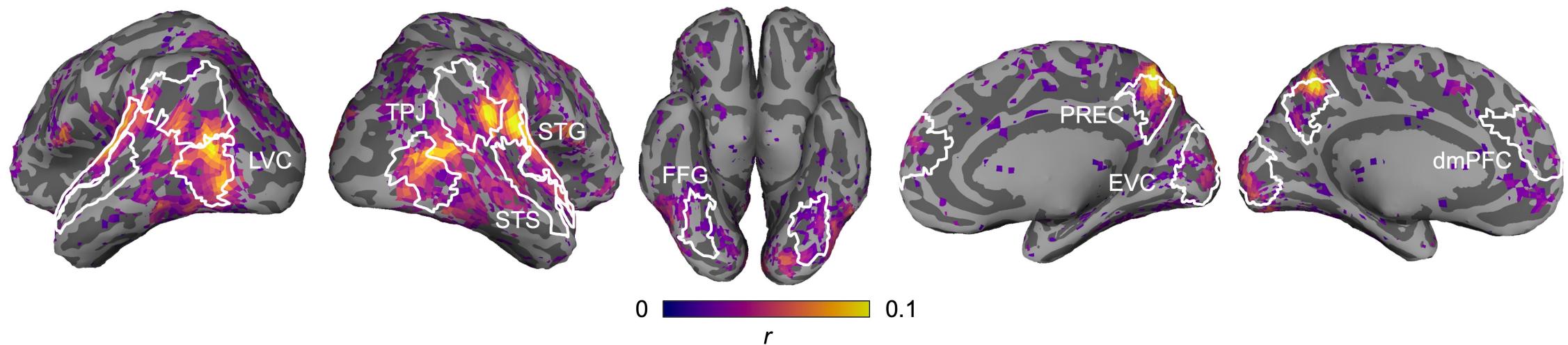
EM results

b Prediction score map of DASOM encoding model



EM results

C Group-level statistics

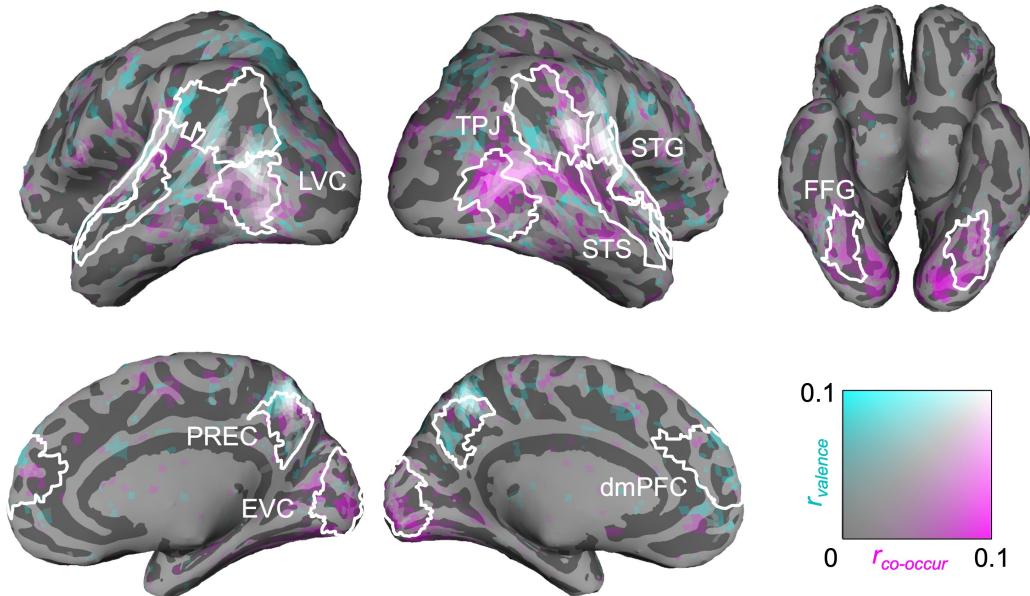


EM results

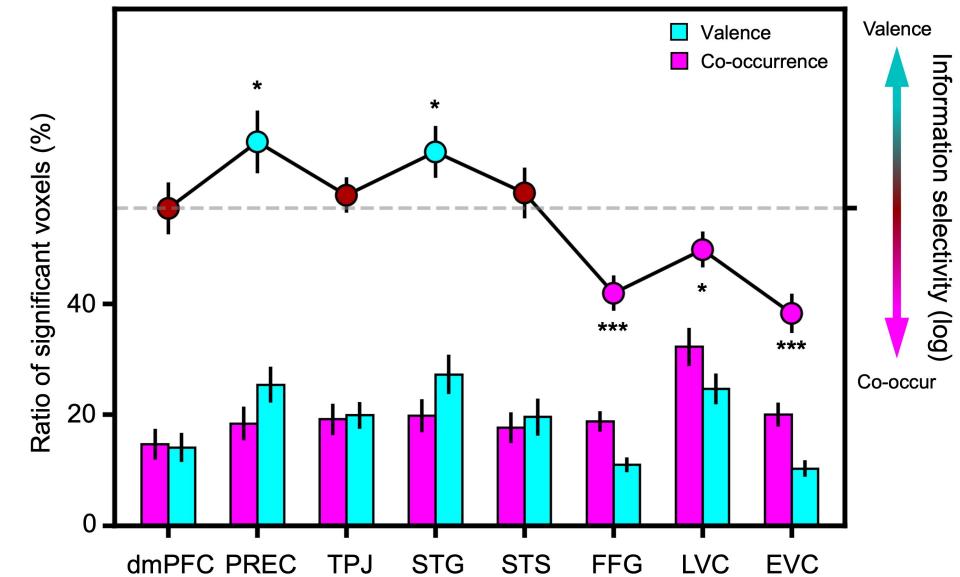
Banded ridge regression

a Co-occurrence vs. Valence information encoding

Group-level encoding performance



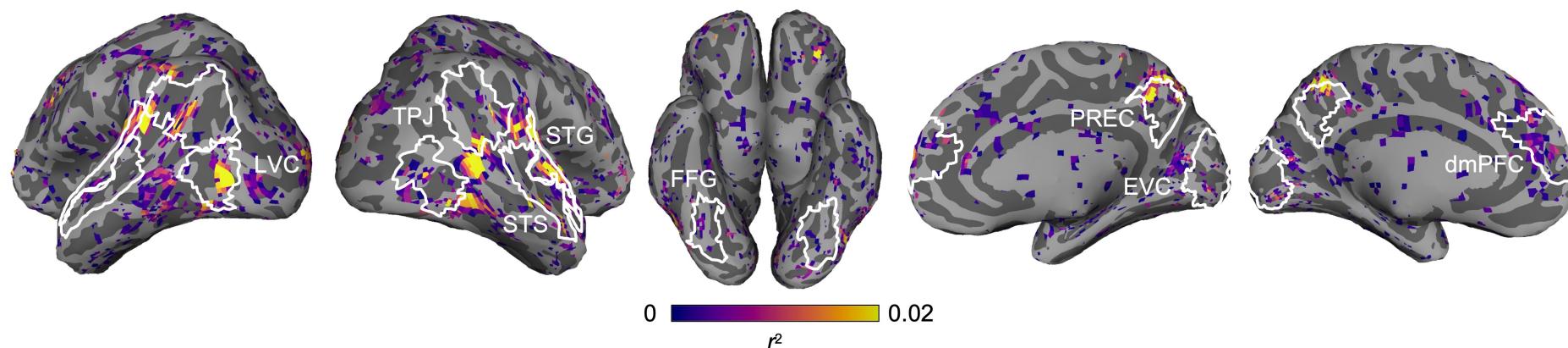
b Information selectivity



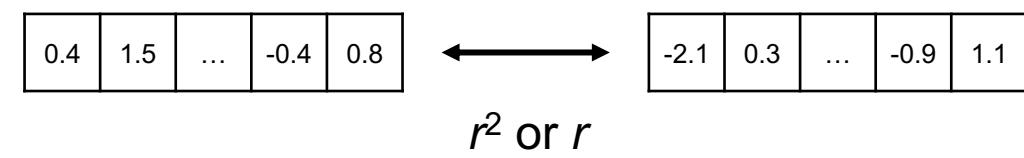
Statistical test

Individual participant

Representative participant (sub-04)



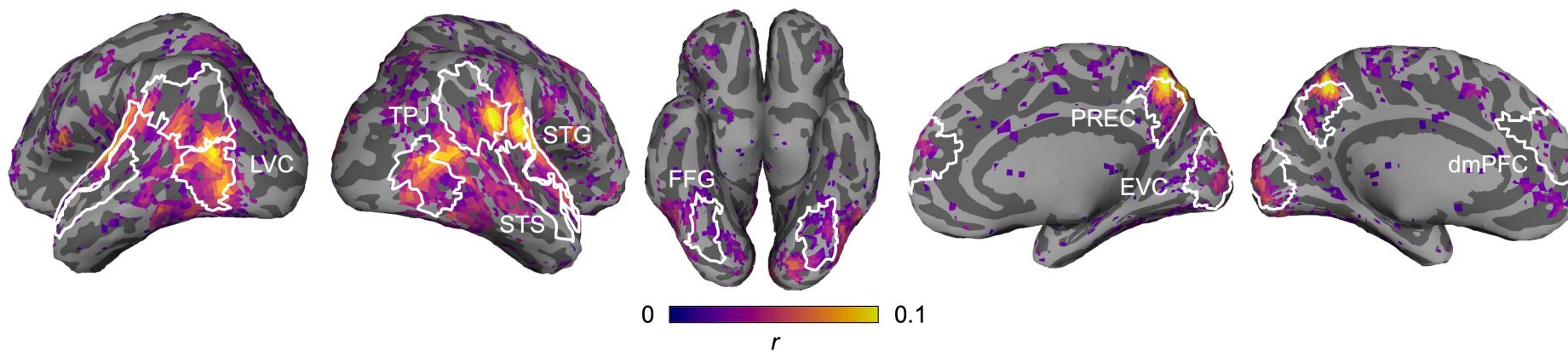
Random Gaussian vectors



$\times 1000$ iteration

Statistical test

Group-level statistics



One-sample bootstrap hypothesis test