Anatomically-informed spatial noise models improve inference for multi-voxel pattern analysis

Mahdiyar Shahbazi*, Lingling Lin*, Jingyu Cui and Jörn Diedrichsen

Western Institute for Neuroscience, Western University, London, Canada

Western University · Canada

Poster # 804

Introduction

What factors determine the voxel-by-voxel noise covariance in fMRI?

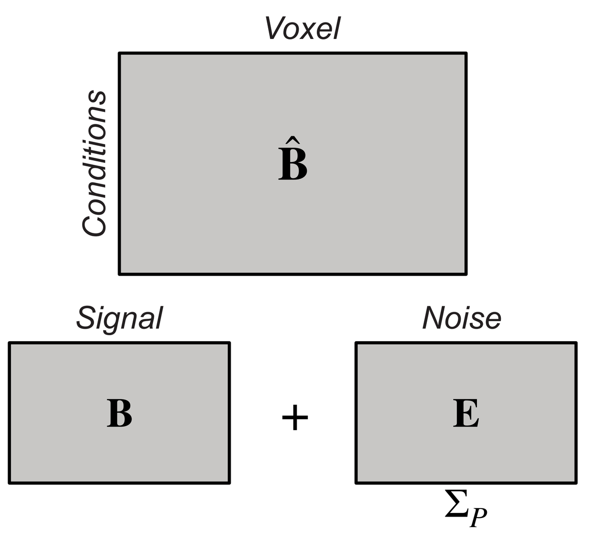
- The spatial covariance of fMRI measurement noise has a strong and reliable structure
- Assuming that voxels are independent leads to sub-optimal multivariate inference (Walther et al., 2016, Diedrichsen et al., 2021)
- Estimating the spatial noise covariance is hard: Often the #observations is smaller (or close to) the #voxels -> Empirical estimate has *high variance*.
- Solution: Shrink the estimate towards the diagonal (Ledoit and Wolf, 2003). This biases the estimate towards the incorrect assumption that voxels are independent.

Goals:

- *Build a model that predicts noise correlations based on anatomical information*
- *Use the model prediction to integrate with empirical estimate*
- *Improve inference for multi-voxel pattern analysis*

Methods: empirical noise correlation

Estimating the noise correlation from the residual of the first-level GLM



First-level linear model: $\mathbf{Y} = \mathbf{XB} + \mathbf{R}_{T\times P}$

*Voxels have **different noise-levels*** $\hat{\sigma}_p^2 = \frac{1}{T-Q}\sum_{t=1}^{T} r_{t,p}^2$

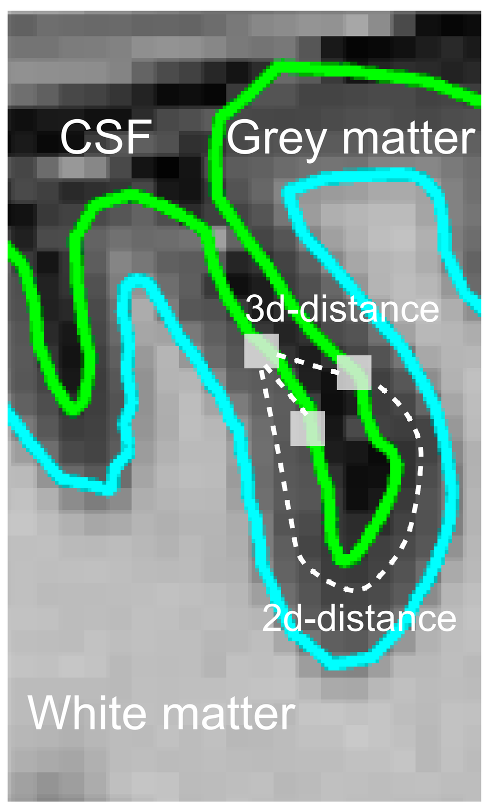
*Voxels are **correlated:** estimate the correlation from standardized residuals* $\hat{\Sigma}_P = \mathbf{\grave{R}\grave{R}}^T$

We used **6** fMRI datasets with voxel resolution ranging from 1.4mm-3mm

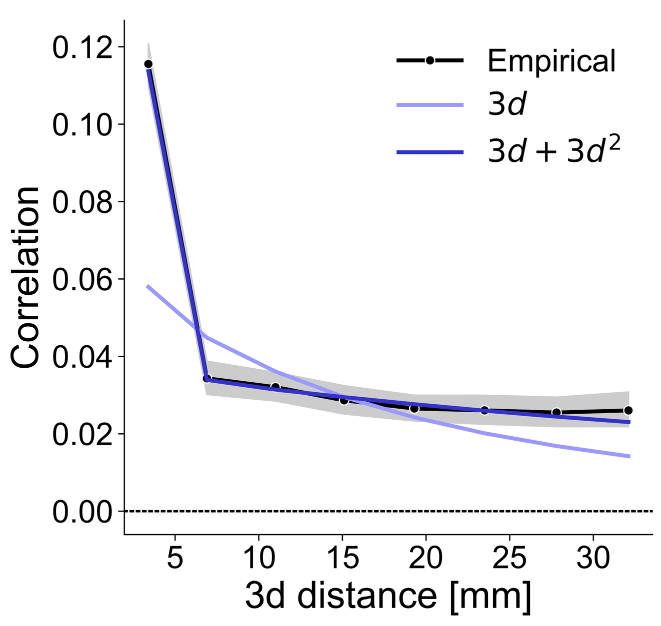
Voxel correlation depends on their spatial distance

We expect noise correlation to decay as a function of spatial distance.

- Measurement noise is expected to fall off with distance in the volume (3d)
- Neural variability is expected to fall off with the distance on the cortical surface (2d)



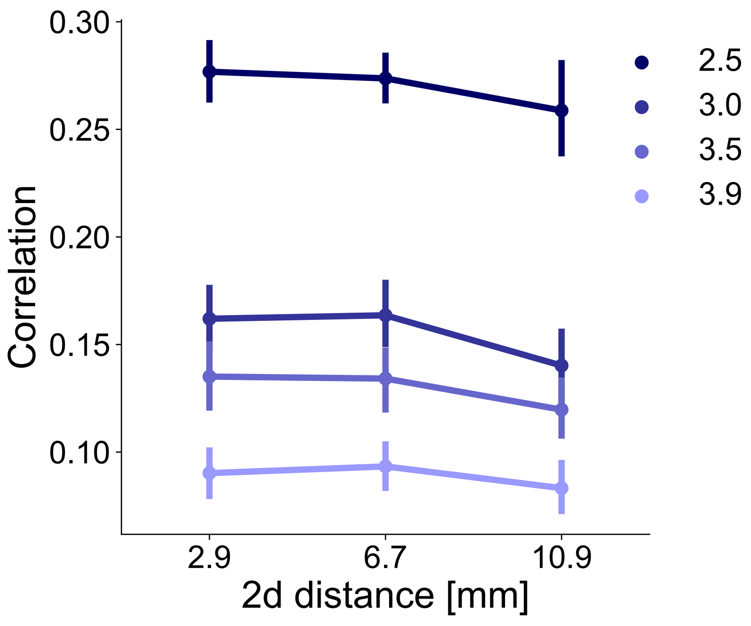
Correlation vs. 3d



Correlation falls off as a double exponential. Large correlations between neighbouring voxels - persistent correlations over long distances.

- In most brain regions, 3d and 2d are correlated
- Accounting for 3d, does the correlation decay with 2d distance?

Correlation vs. 2d



Within each bin, we calculated the slope and average the slopes across bins

2/6 datasets showed significant decreasing pattern

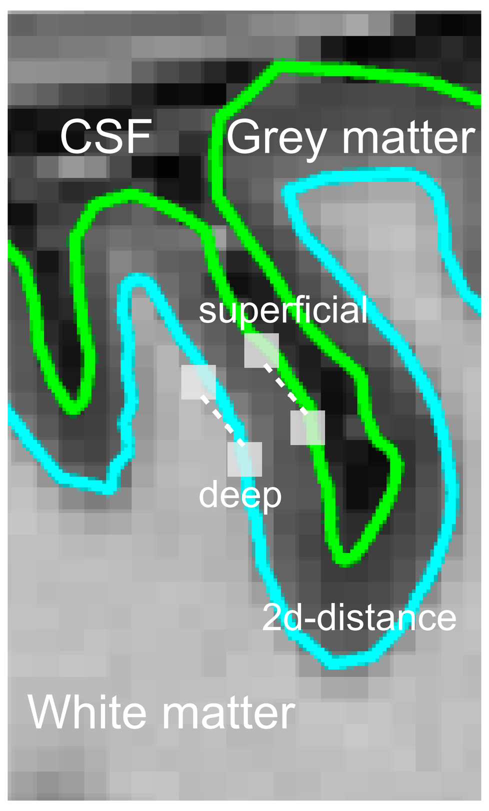
4/6 datasets showed numerically negative slope

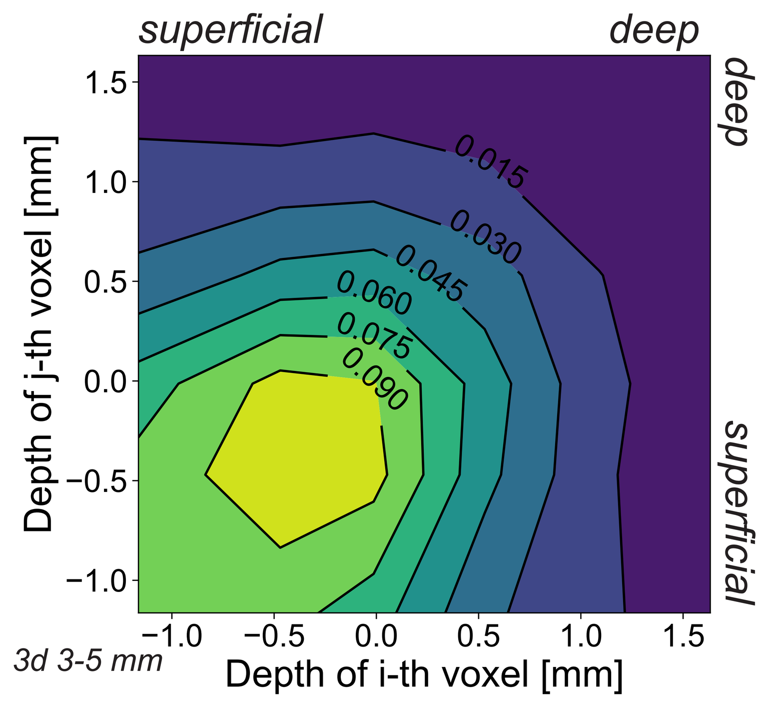
Different distance models:

3d	$\gamma_1 \exp(-\alpha d_{ij})$
3d+3d²	$\gamma_1 \exp(-\alpha d_{ij}) + \gamma_2 \exp(-\beta d_{ij}^2)$
2d+2d²	$\gamma_1 \exp(-\alpha g_{ij}) + \gamma_2 \exp(-\beta g_{ij}^2)$
3d+3d²+2d+2d²	$\gamma_1 \exp(-\alpha_1 d_{ij}^2) + \gamma_2 \exp(-\alpha_2 g_{ij}^2) + \gamma_3 \exp(-\beta_1 d_{ij}) + \gamma_4 \exp(-\beta_2 g_{ij})$

Voxel correlation also depends on cortical depth

Voxels lie deeper in the grey matter tend to have smaller correlation





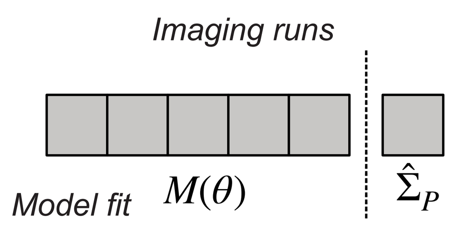
2 voxels both located in deep, have small correlation

2 voxels both located near CSF, have larger correlation

3d+3d²+depth $\gamma_1 \exp\{-\alpha d_{ij} - \theta((h_i-\mu)^2 + (h_j-\mu)^2)\} + \gamma_2 \exp(-\beta d_{ij}^2)$

Evaluating model predictions

How well does each model predict the noise-correlation in a left-out run?



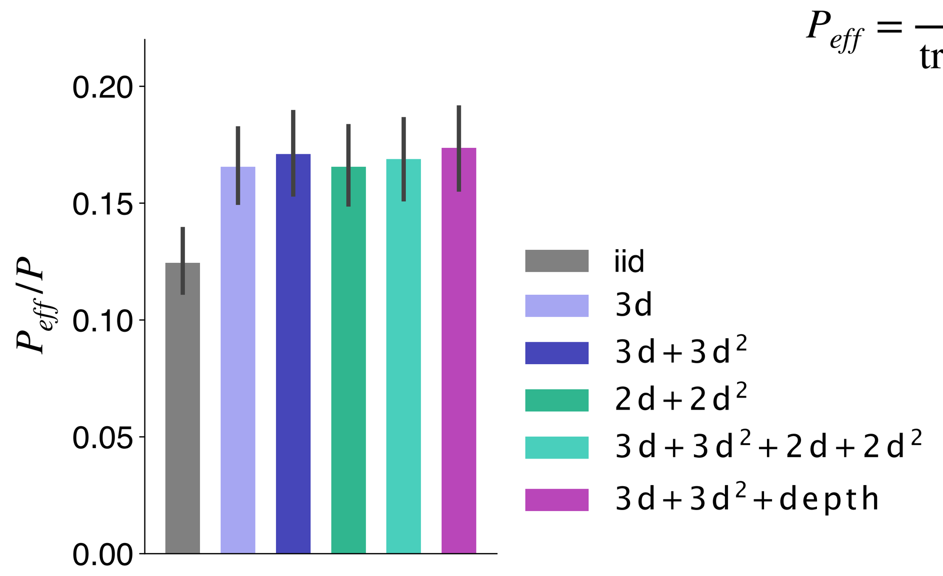
Evaluation: Predicted variance of a multivariate measure (distance, correlation, second moment) depends on the effective number of voxels after pre-whitening:

$$\Sigma_R = M^{-\frac{1}{2}}\hat{\Sigma}_P M^{-\frac{1}{2}}$$

Voxel correlation after pre-whitening

$$P_{eff} = \frac{P^2}{\mathrm{trace}(\Sigma_R \Sigma_R)}$$

Effective number of voxels. Perfect prediction $P_{eff} = P$. Large deviation lead to $P_{eff} = 1$.

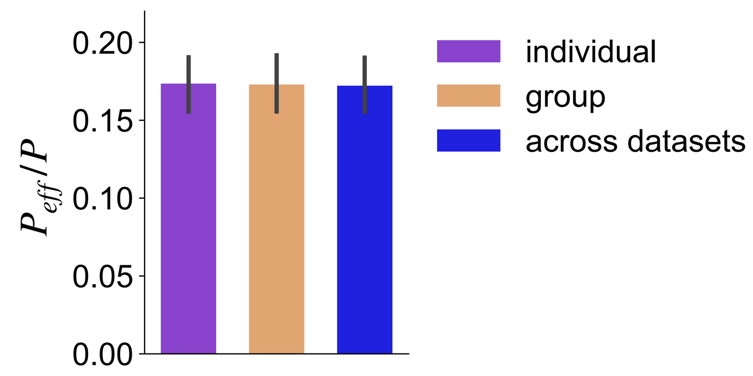


2d-distance does not add much

Model with depth and 3d-distance is the best

Model parameters are stable across subjects and datasets

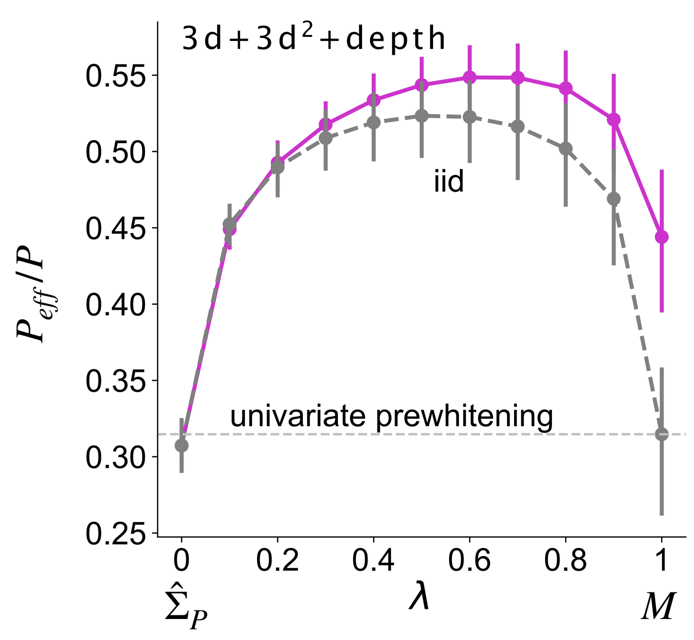
Practically, we can use the common parameters estimated in one study to apply to a completely new dataset.



Regularized estimate

Using model prediction as a shrinkage target in regularization:

$$\tilde{\Sigma}_P = (1-\lambda)\hat{\Sigma}_P + \lambda M$$



Regularization improves estimate generally

Optimal λ depends on dataset and #Time points / #voxels

Model-based shrinkage target is always better than identity

Conclusion and open questions

- Models with only 3d distance predict noise correlation better than models with 2d distance. Is 2d distance a better predictor for signal variabilities?
- Better model of noise correlation leads to smaller variance multivariate measures
- Pre-whitening emphasizes high spatial frequencies. What if signal variabilities are in low spatial frequencies?

Ledoit, O., & Wolf, M. (2003). Honey, I shrunk the sample covariance matrix. UPF economics and business working paper

Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., & Diedrichsen, J. (2016). Reliability of dissimilarity measures for multi-voxel pattern analysis. Neuroimage

Diedrichsen, J., Berlot, E., Mur, M., Sch, H., Shahbazi, M., Kriegeskorte, N. (2021). Comparing representational geometries using whitened unbiased-distance-matrix similarity. Neurons, Behavior, Data and Theory

@_mshahbazi