Through the looking-glass: how analytical flexibility shapes subject space in resting state fMRI
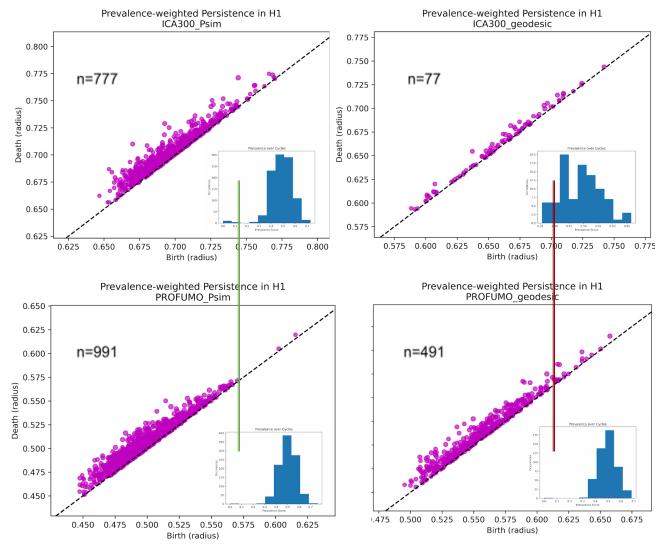
Ty Easley[1], Elizabeth Munch[2,3], Kevin Freese[4†], Janine Bijsterbosch[1]

1. Mallinckrodt Institute of Radiology, Washington University in St. Louis
2. Department of Computational Mathematics, Science, and Engineering, Michigan State University; 3. Department of Mathematics, Michigan State University
4. Cloud Object Storage, IBM† († The views expressed herein are those of the author and do not represent the views of IBM.)

Reliable characterizations of between-subject differences are crucial to neuroimaging studies, especially those aiming to relate neural organization and mental health. However, varying analytical practices across researchers create silos in the field, producing different understandings of between-subject variability that can translate into disparate findings (Botvinik-Nezer et al. 2020; Bijsterbosch et al. 2020). We use topological data analysis to articulate differences in between-subject covariance structure induced by different analytical approaches.

Minimally-processed grayordinate data from N=1003 Human Connectome Project (HCP-YA) young adult subjects were analyzed through six different dimensionality reduction and feature selection pipelines ("brain representations"). From each brain representation, subject-pairwise dissimilarity matrices were computed for geodesic (Pennec, Fillard, and Ayache 2006), inner product-, and Pearson similarity-derived distance (pre)metrics. We then computed the persistent homology (Dey and Wang 2022) of each subject-pairwise dissimilarity matrix and measured its stability over 1000 cycle-registered bootstraps (Reani and Bobrowski 2022).

We find significant ($p < 0.05$) differences for different representations (Figure; green line denotes significant score distribution difference, n is the number of matched cycles) in at least either the number of prevalent homology generators (matched cycles) and prevalence score distributions (Reani and Bobrowski 2022), where significance is determined by pairwise Kolmogorov-Smirnov (KS) tests between per-method prevalence score distributions. It follows that different modes of analysis produce characterizations of subject variability that do not admit lossless translations between all pairs. Our results demonstrate that brain representation is a central choice in the rfMRI analysis pipeline: the choice of brain representation can engender irreconcilable interpretive lenses on rfMRI data. Further study is required to characterize shared and divergent topological structure between representations of subject space and the ensuing consequences for translation between representations.

Bijsterbosch, Janine, Samuel J. Harrison, Saad Jbabdi, Mark Woolrich, Christian Beckmann, Stephen Smith, and Eugene P. Duff. 2020. "Challenges and Future Directions for Representations of Functional Brain Organization." *Nature Neuroscience*, October, 1–12.
Botvinik-Nezer, Rotem, Felix Holzmeister, Colin F. Camerer, Anna Dreber, Juergen Huber, Magnus Johannesson, Michael Kirchler, et al. 2020. "Variability in the Analysis of a Single Neuroimaging Dataset by Many Teams." *Nature* 582 (7810): 84–88.
Dey, Tamal Krishna, and Yusu Wang. 2022. *Computational Topology for Data Analysis*. Cambridge University Press.
Pennec, Xavier, Pierre Fillard, and Nicholas Ayache. 2006. "A Riemannian Framework for Tensor Computing." *International Journal of Computer Vision* 66 (1): 41–66.
Reani, Yohai, and Omer Bobrowski. 2022. "Cycle Registration in Persistent Homology with Applications in Topological Bootstrap." *IEEE Transactions on Pattern Analysis and Machine Intelligence* PP (October). https://doi.org/10.1109/TPAMI.2022.3217443.