Network decoding for multi-subject archetypal analysis

Alex Shepherd, Emily Stone, and Lucy L. W. Owen
University of Montana

Overview

- We hypothesize that distinct network patterns correspond to characteristic archetypal configurations of large-scale brain networks.
- We extended classical archetypal analysis (AA) to multi-subject fMRI data, optimizing archetypes jointly across subjects while preserving individual coefficients. This approach captures high-variance, condition-specific network configurations beyond mean-based methods like ICA or PCA.
- We applied our approach to an fMRI dataset collected as participants listened to a 10 minute story with different listening conditions as well as rest.

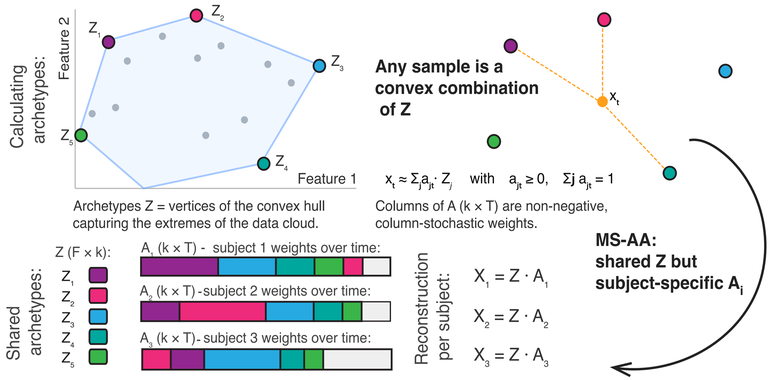


Fig. 1 Schematic of the AA model. Each subject's data matrix X_i (timepoints × nodes) is represented as a convex combination of k archetypes (Z), shared across participants. Subject-specific coefficient matrices A_i capture how each archetype fluctuates over time. Multi-subject AA (Hinrich et al., 2016) constrains archetypes to span a shared convex hull across subjects, yielding group-level components that capture population-level extremes in spatial and temporal patterns.

Methods

- 108 participants in an fMRI study each participated in 1 of 3 experimental conditions: intact, word-scrambled, or rest (Simony et al., 2016)
- We calculated MS-AA on both within and across conditions.
- We assessed our model using cross validated timepoint decoding using an increasing number of archetypes.

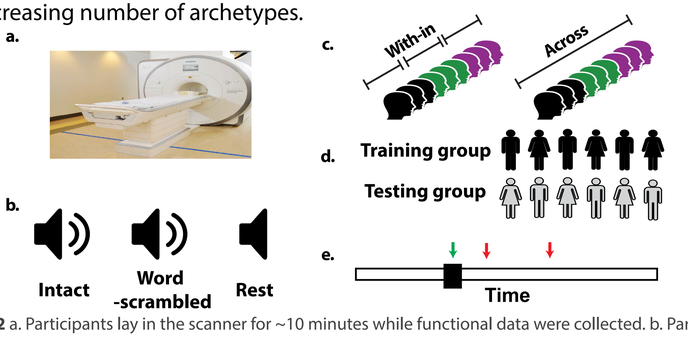


Fig. 2 a. Participants lay in the scanner for ~10 minutes while functional data were collected. b. Participants were randomly assigned to1 of 4 experimental conditions (paragraph-scrambled was excluded due to shorter scan time.) The experimental conditions varied systemically in cogntive "richness" with audio recording of the story Pie Man by Jim O'Grady. We applied HTFA (Manning et al., 2018) to obtain 700 node activities for every participant. c. We computed MS-AA for with-in (spatial and tempora AAI) and across (spatial AA) conditions, and then we randomly assigned participants in each condition to two groups. d. We compared the groups' activity patterns (using Pearson correlations) to estimate the story times for each corresponding pattern. Timepoint decoding was performed using correlation-based matching between in-fold and out-of-fold averages.

Approach

- Spatial and temporal archetypes represent complementary facets of network organization.

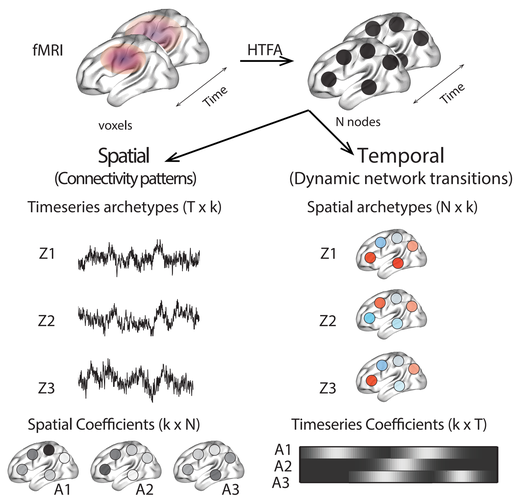


Fig. 3 (Left) Spatial archetypes capture prototypical temporal patterns. Each archetype corresponds to an extremal time series with associated brain-wide coefficients (gray-scale node maps) indicating how strongly each region expresses that temporal profile. (Right) Temporal archetypes describe spatially distinct patterns of network coactivation. Each archetype is defined by a characteristic spatial map (red–blue node weights) and a corresponding set of time-varying coefficients (gray coefficient matrix) showing how expression of each archetype fluctuates over time. Coefficient matrices sum to one.

Results

- We performed spatial and temporal MS-AA for k archetypes, with-in conditions.
- We used the resulting reconstructed data to assess decoding accuracy.

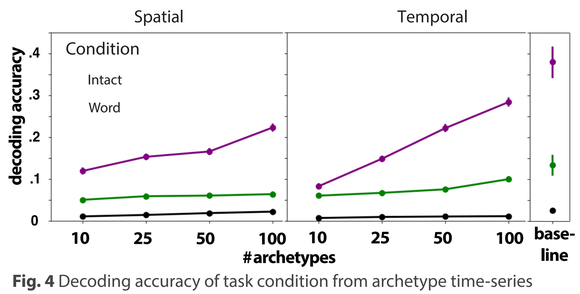


Fig. 4 Decoding accuracy of task condition from archetype time-series (**spatial; left**) and spatial reconstructions (**temporal; middle**) as a function of model order. Accuracy increased with k, with maximal decoding at k = 100 using temporal archetypes. Each line shows mean ± SE across repetitions (n = 50). Decoding was also performed on the z-scored data (**baseline, right**).

Please visit our website: the-brain-lab.github.io

Results

- We can also look across all subjects (and all conditions) to perform archetypes. This allows us to compare coefficients across conditions.

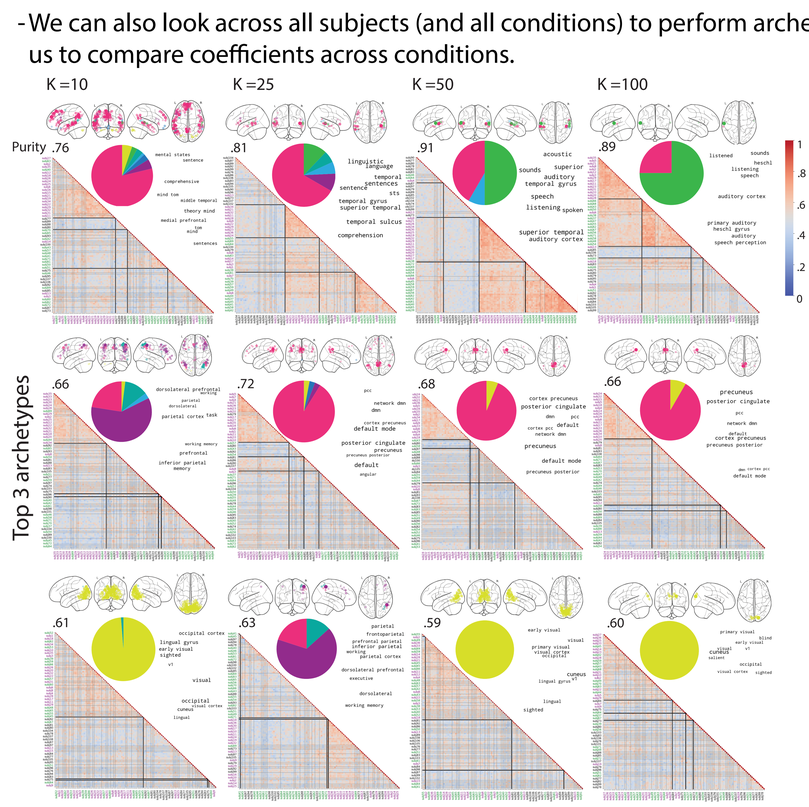


Fig. 5 Each heatmap shows subject-by-subject similarity of temporal coefficients for a given archetype. Similarity was computed as the Pearson correlation between subjects' time-varying archetype weights after within-subject z-scoring. The resulting similarity matrices were hierarchically clustered using average linkage on Euclidean distances between correlation rows. Cluster boundaries (black lines) mark subgroups of subjects with correlated temporal expression patterns. Purity was quantified as the proportion of subjects within each cluster sharing the same experimental condition; archetypes exceeding purity > 0.55 were considered condition-specific "winners." Associated brain maps depict spatial weights of these high-purity archetypes, colored by Yeo-7 network assignment and scaled by relative node contribution. Adjacent pie charts summarize each archetype's network composition. Reverse inference results from Neurosynth reported, and word size is scaled by term similarity.

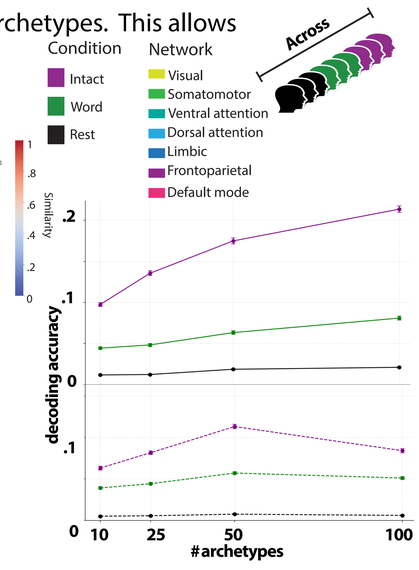


Fig. 6 (Top) Timepoint decoding was performed on across-condition MS-AA reconstructions using correlation-based template matching between in-fold and out-of-fold subject averages. Accuracy was computed as the proportion of correctly matched timepoints between folds. Decoding was repeated for each K and averaged across 50 repetitions. **(Bottom)** Combined "winning" archetypes were reconstructed by concatenating all high-purity (>.55) archetypes across conditions before decoding.

Summary of findings

- Multi-Subject Archetypal Analysis (MS-AA) captures population-level extremes in brain network configurations, decomposing fMRI data into shared spatial (network maps) and subject-specific temporal (time-varying expression patterns) archetypes.
- Condition-specific archetypes emerged for intact, word-scrambled, and rest states—identified through subject-by-subject similarity clustering and network purity analyses. High-purity "winning" archetypes revealed distinct large-scale network engagements validated via Neurosynth functional decoding.
- Temporal archetypes outperform spatial archetypes in decoding task within-in conditions for subjects and model orders, showing that dynamic network transitions—not static connectivity patterns—carry the most discriminative information about ongoing cognitive state.

References:
Hinrich, J. L., Hansen, L. K., & Mørup, M. (2016). Archetypal analysis for modeling multi-subject fMRI data. Neuroinformatics, 14(2), 163-176.
Manning, J. R., ,et al. (2018). A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. NeuroImage,180, 243–252.
Simony, E., Honey, C. J., Chen, J., & Hasson, U. (2016). Uncovering stimulus-locked network dynamics during narrative comprehension. Nature Communications, 7(12141), 1–13
Stone, E., & Cutler, A. (1996). Archetypal analysis of spatio-temporal dynamics. Physica D: Nonlinear Phenomena, 90(3), 209-224.