**A multivariate approach to analyze connectivity matrices with individual-specific parcellation**

Ju-Chi Yu1, Micaela Chan1,2, Liang Han1,2, Phillip Agres1,2 & Hervé Abdi1

1. School of Behavior and Brain Sciences, The University of Texas at Dallas

2. Center for Vital Longevity, The University of Texas at Dallas

[Ju-Chi.Yu@utdallas.edu](mailto:Ju-Chi.Yu@utdallas.edu), [mchan@utdallas.edu](mailto:mchan@utdallas.edu), [Liang.Han@utdallas.edu](mailto:Liang.Han@utdallas.edu), [Phillip.Agres@utdallas.edu](mailto:Phillip.Agres@utdallas.edu), [herve@utdallas.edu](mailto:herve@utdallas.edu)

**Abstract**

DiSTATIS (Abdi, et al., 2012)—a three-way multidimensional scaling (MDS)—is used to analyze resting-state functional magnetic resonance imaging (rsfMRI) data and to visualize the pattern of dissimilarity between networks. These dissimilarities are obtained by analyzing multiple distance matrices that have matching rows and columns (i.e., distinct functional regions in rsfMRI analysis). These functional regions, referred to as parcels or networks, are usually derived from a shared template onto which all data are mapped. However, using a shared template could bias results against participants that vary greater from this template, and this issue is particularly problematic amongst participants with diverse brain structures and functions (e.g., elderly, lesion patients, children). For such participants, the optimal parcellation should maximize the homogeneity of the signal within each parcel region and of each participant. Therefore, recent work in rsfMRI analysis has developed techniques to derive individual-specific parcellations and sub-networks; the goal of this project is to propose a multivariate approach that accommodates such parcellation which allows differing parcel numbers and organization across participants.

The proposed multivariate approach first extracts the upper-triangle of the connectivity matrix of each participant session (Fig. 1A) and vectorizes it to form the rows of a data table. Sessions from the same participant are stacked to form a block of columns. Each participant's block is placed adjacent to one another to form the final data table, where the rows are sessions and the columns are network edges (Fig. 1B). Next, hierarchical multiple factor analysis (Fig. 1C) (Abdi, et al., 2013; Le Dien & Pagès, 2003) is used to preprocess the data table by normalizing each participant’s column-block by its first singular value, then dividing the column-block of each edge-type (i.e., between- or within-network) within each participant’s column block by its first singular value. Finally, a PCA is performed to visualize the edges, participants, and sessions.

To test this new technique, we obtained the Midnight Scan Club (MSC) dataset (Gordon, et al., 2017) from the OpenfMRI database (ds000224) of which the individual-specific functional parcellation for each participant was available. The functional network of each participant’s session is represented by a symmetric, region-by-region connectivity matrix, where connectivity (i.e., edge) is given by the Fisher’s *Z*-transformed correlation of the BOLD signals between two regions. From the original MSC data, four sessions from each participant were selected for the analysis (sessions 1 – 4). For Sessions 2 and 4, the connectivity of three types of edges was manipulated to simulate commonly observed changes in functional brain networks: (1) decreases within the default mode network (DMN) and increases (2) between DMN and the frontoparietal network (FPN) and (3) between the DMN and the dorsal attention network (DAN). Simulations of changes were restricted to specific networks to test whether our technique could detect where between-session changes are located.

This new multivariate approach to analyzing the MSC data table showed the session effect (simulated vs. non-simulated sessions) on the first component (Fig. 2A). This separation is driven by the within DMN connectivity and several other between-network edges, including edges between DMN and FPN (Fig. 2B). A comparison analysis was conducted using DiSTATIS, where data were mapped to a shared template (Gordon, et al., 2016). Although DiSTATIS also revealed the session effect within DMN (Fig. 2C), between-network effects could not be detected.

In conclusion, this technique provides a multivariate approach to analyze functional connectivity with individual-specific parcellation and is particularly useful for participant group of diverse brain sizes, functional parcellation or organization schemes.

**References**

Abdi, H., Williams, L.J. and Valentin, D. (2013) 'Multiple factor analysis: principal component analysis for multitable and multiblock data sets.' *Wiley Interdisciplinary reviews: computational statistics*, **vol. 5**, no. 2, pp. 149-179.

Abdi, H., Williams, L.J., Valentin, D. and Bennani-Dosse, M. (2012) 'STATIS and DISTATIS: optimum multitable principal component analysis and three way metric multidimensional scaling.' *Wiley Interdisciplinary Reviews: Computational Statistics*, **vol. 4**, no. 2, pp. 124-167.

Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M. and Petersen, S.E. (2016) 'Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations.' *Cerebral Cortex*, **vol. 26**, no. 1, pp. 288-303.

Gordon, E.M., Laumann, T.O., Gilmore, A.W., Newbold, D.J., Greene, D.J., Berg, J.J., Ortega, M., Hoyt-Drazen, C., Gratton, C., Sun, H., Hampton, J.M., Coalson, R.S., Nguyen, A.L., McDermott, K.B., Shimony, J.S., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., Nelson, S.M. and Dosenbach, N.U.F. (2017) 'Precision Functional Mapping of Individual Human Brains.' *Neuron*, **vol. 95**, no. 4, pp. 791-807.e7.

Le Dien, S. and Pagès, J. (2003) 'Hierarchical Multiple Factor Analysis: application to the comparison of sensory profiles.' *Food Quality and Preference*, **vol. 14**, no. 5, pp. 397-403.

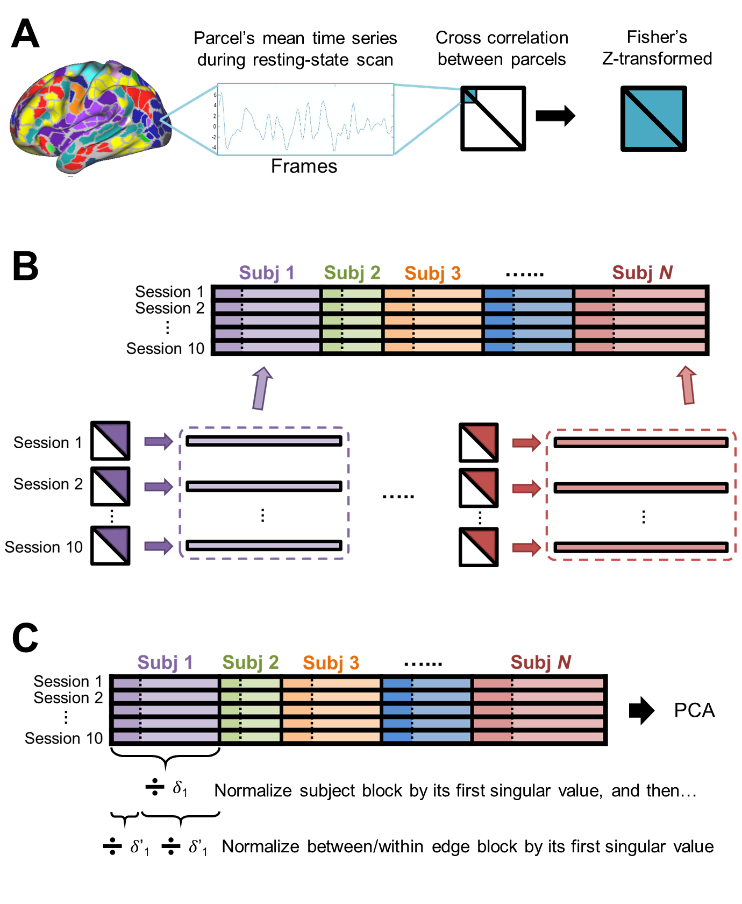


Fig. 1

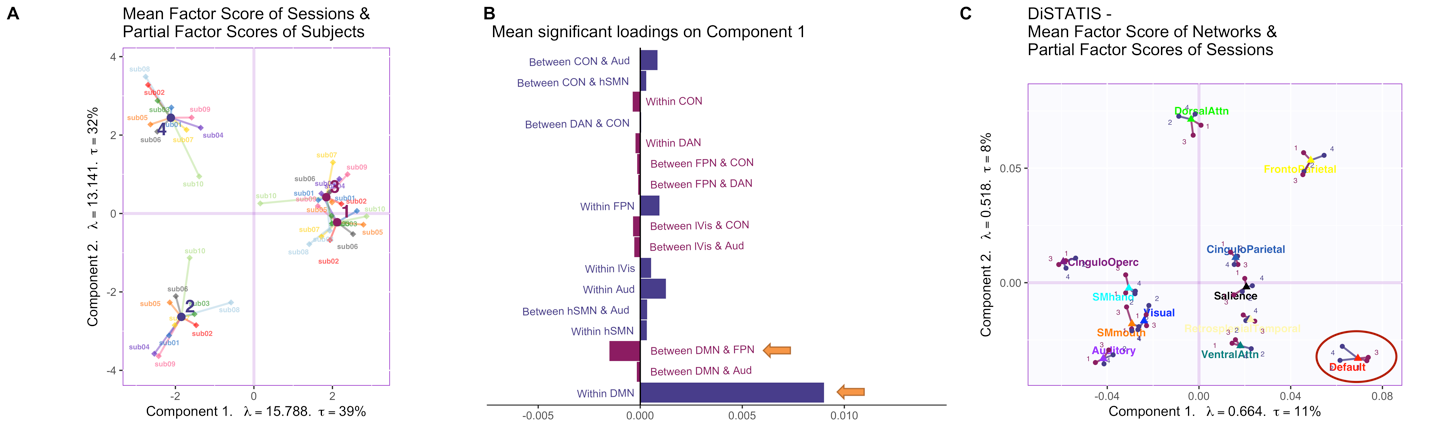


Fig. 2

1. The row factor scores from the hierarchical multiple factor analysis (HMFA) show that the first component distinguishes the simulated (blue dots; Sessions 2 and 4) and non-simulated (purple dots; Sessions 1 and 3) sessions. The small round dots that extend from each factor score are the partial factor scores, which represent how each session is viewed from the perspectives of different tables (i.e., different participants). (B) The mean loadings of all edges that significantly contribute to the first component and drive the session effect. The orange arrows pointing towards the simulated edges (i.e., the functional connectivity within the DMN as well as the functional connectivity between the default mode network (DMN) and the fronto-parietal network (FPN)). (C) The mean factor scores of all networks from DiSTATIS with parcellation based on a shared template. The four round dots that extend from each network factor score are the partial factor scores, which represent how each network is viewed from the perspectives of different sessions. Only the partial factor scores of the DMN distinguish the simulated (blue dots) and the non-simulated (purple dots) sessions.