

Dynamic Topological Subtype Analysis of Healthy Brain Activity

Jisong Sun^{1,a,*}

¹Chongqing University, Shapingba, Chongqing, 400044, China

a. jisongsun@outlook.com

*corresponding author

Abstract: The functional organization of the human brain exhibits certain heterogeneity; however, the topological subtype characteristics of brain networks in healthy populations have not been fully explored. This study, based on resting-state functional magnetic resonance imaging data from 48 healthy adults in the UK Biobank dataset, employs persistent homology analysis to identify and characterize different subtypes of brain network topology. By constructing individual functional connectivity matrices and extracting their zero-dimensional (connected components) and one-dimensional (circular structures) topological features, unsupervised clustering analysis revealed three brain network subtypes with distinct functional organizational characteristics: the mainstream subtype (52.1%) exhibited moderate network connectivity and moderate modularity; the high modularity subtype (27.1%) displayed the most persistent circular structures and the highest network segregation; the integrative subtype (20.8%) featured extensive functional connectivity but a simple topological structure. Further statistical analysis confirmed significant differences between these subtypes across all topological feature dimensions ($p < 0.001$). These findings, for the first time, reveal the heterogeneity of brain functional organization in healthy populations from a topological perspective, providing important evidence for constructing more accurate brain network models and developing individualized clinical applications.

Keywords: Resting-state functional magnetic resonance imaging, Persistent homology analysis, Topological data analysis, Healthy brain topological subtypes

1. Introduction

The human brain is a highly organized complex system whose function depends on the synchronization of neuronal activity and information exchange between different brain regions. Through resting-state functional magnetic resonance imaging (fMRI), researchers have discovered that even at rest, the brain exhibits stable functional connectivity patterns, which reflect the temporal correlation of spontaneous neural activity between brain regions [1]. These functional connections constitute the brain's basic operational modes, supporting human cognitive functions and behavioral performance [2]. Current research indicates that several neurological disorders are associated with specific changes in functional connectivity: the severity of motor symptoms in Parkinson's disease patients is significantly related to the functional connectivity strength of certain brain regions [3]; social cognitive impairments in autism spectrum disorder can be traced to abnormalities in specific functional networks [4]. However, these studies primarily rely on group comparisons between patient

populations and healthy controls, a method that struggles to capture finer patterns of functional connectivity changes. Some studies suggest that the brain's functional organization in healthy individuals may exhibit multiple stable patterns [5], and treating the healthy population as a single standard for comparison may obscure these important biological features. A deeper understanding of functional connectivity subtypes in healthy individuals not only helps reveal the basic principles of brain functional organization but also provides a critical foundation for developing more precise diagnostic and treatment strategies for diseases [6].

To systematically investigate the functional organization of the brain, researchers introduced the concept of brain networks, viewing the brain as a complex system composed of multiple functional nodes connected through relationships. Traditional graph theory methods represent brain functional connectivity as a network model of nodes and edges, used to analyze the functional interactions between brain regions [7]. This approach has revealed the modular structure of brain networks, small-world properties, and their association with cognitive functions, laying an important foundation for brain network research [8]. Topological Data Analysis (TDA) is an emerging mathematical tool that reveals the structural information of complex systems by studying the geometric and topological features of data. Unlike graph theory, which focuses only on bilateral relationships between nodes, TDA can capture higher-order interaction features between multiple brain regions [9]. These features not only reflect the efficiency of information flow but also embody the dynamic integration and segregation characteristics of brain networks [10]. Persistent Homology analysis, which tracks the “birth” and “death” of topological features, quantifies the stability of these higher-order features and provides a powerful tool for understanding the organizational principles of brain networks.

Therefore, based on Topological Data Analysis, this study designs a systematic analytical approach to explore the brain network organizational characteristics of healthy individuals. First, persistent homology analysis is used to extract the topological features of each subject's brain functional network, including the temporal dynamics of connected components and circular structures. Second, these topological features are used for unsupervised clustering to identify potential brain network subtypes. Third, differences in topological features between subtypes are analyzed to validate the rationality of subtype classification. Finally, typical samples are selected for in-depth analysis of their functional connectivity patterns and topological features, revealing the neurophysiological significance of different subtypes.

2. Methods

2.1. Dataset and Construction of Functional Connectivity Matrices

This study utilized resting-state fMRI data from 48 healthy adult volunteers in the UK Biobank public dataset. The data was based on the brain network parcellation proposed by Yeo et al., using the 7-networks, 400-regions parcellation [11]. The fMRI data preprocessing was carried out using the DPABI toolkit [12]. The specific process included the following steps: First, the first 10 time points of the functional images were discarded to eliminate data instability. Then, slice timing correction and head motion correction were applied. Functional images were standardized to MNI space using the Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) method. Next, spatial smoothing was performed using a Gaussian kernel (full-width at half-maximum of $6 \times 6 \times 6$ mm³). To reduce noise, signals from white matter, cerebrospinal fluid, and Frison 24 head motion parameters were regressed out. Additionally, a band-pass filter was applied to extract signals within the range of 0.009 to 0.08 Hz. Following preprocessing, the Pearson correlation coefficient of the time series between the 400 regions of interest (ROIs) was computed for each subject to construct the functional connectivity matrix. The positive part of the matrix, A^+ , was then extracted as the adjacency matrix. Based on the theory of brain network optimization for information transmission

efficiency and connection cost by Bullmore and Sporns [13], the positive functional connectivity matrix A^+ was converted into a distance matrix $D=1-A^+$.

2.2. Topological Data Analysis - Persistent Homology Computation

We applied the theory of persistent homology in computational topology [14] to perform topological data analysis on the distance-related matrix. First, the GUDHI library [15] was used to construct a Vietoris-Rips complex with a maximum dimension of 2 [14], and the 0-dimensional (connected components) and 1-dimensional (circular structures) persistent homology were computed. The results of persistent homology are represented by “birth-death” pairs, where “birth” indicates the threshold at which the topological feature begins to exist, and “death” indicates the threshold at which the feature disappears. The difference between these values reflects the lifetime of the topological feature. Topological structures with shorter lifetime are generally considered noise [16], so a minimum lifetime threshold of 0.025 was set for filtering purposes.

The results of persistent homology computation were visualized in two ways: barcode plots and persistence diagrams [14]. The barcode plot uses the length of horizontal bars to represent the birth and death of topological features (as shown in Figure 1b), while the persistence diagram (as shown in Figure 1c) maps the intervals from the barcode onto points in a plane, with the x-axis representing the birth time and the y-axis representing the death time. The further a point is from the diagonal line, the more persistent the topological feature is.

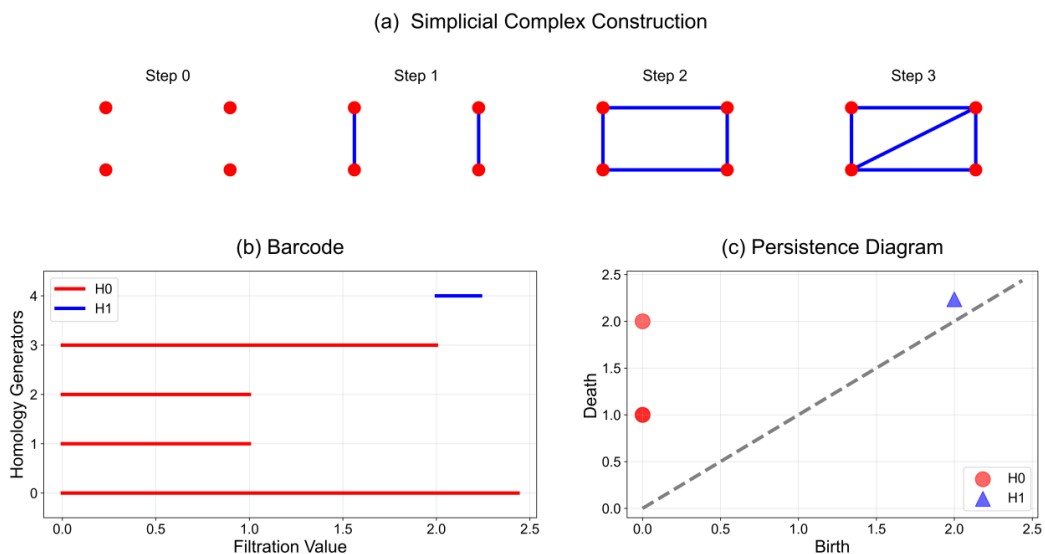


Figure 1: Example of persistent homology computation for a rectangular complex. (a) Schematic of the complex construction process: starting from 4 discrete vertices, short edges, long edges, and finally the diagonal are added. (b) Barcode plot, with red indicating connected components (H0) and blue indicating circular structures (H1). (c) Persistence diagram, with the x-axis representing birth time and the y-axis representing death time.

In the persistent homology computation, we focus on the evolutionary process of the 0-dimensional homology group (connected components) and the 1-dimensional homology group (circular structures). For the 0-dimensional homology features, we concentrate on the statistical properties of the middle 50% of connected components, as these features best represent the primary connectivity patterns of the network. Specifically, a larger mean lifetime (M_0) indicates greater stability in the network’s basic connectivity structure; a smaller standard deviation of lifetime (S_0) suggests more stable fluctuations in network connectivity. For the 1-dimensional homology features,

we construct a hierarchical feature system. First, we extract the average birth time (B1) and average death time (D1) of the circular structures in the time dimension. Later birth times indicate that circular structures form under higher connection strengths, while later death times suggest that these structures can be maintained under stronger connection conditions. Next, we calculate the statistical features of the circular structures: the average lifetime (M1), which reflects the overall persistence of the circular structures, and the standard deviation of the lifetime (S1), which quantifies the variability of the stability of these structures. These two indicators jointly describe the dynamic features of the network's circular structures. Finally, we introduce two distributional features: the Shannon entropy of the lifetime (E1) and the average lifetime of the top 25% most persistent circular structures (T1). A higher Shannon entropy of the lifetime indicates that the network has richer multiscale topological features, while the average lifetime of the top 25% most persistent circular structures specifically highlights the most prominent circular structures, providing a quantitative measure of the network's key topological features.

Based on the extracted topological features, we first perform MinMax normalization on all the features, then apply the K-means algorithm [17] to cluster the samples of the 48 participants. The optimal number of clusters is determined using the silhouette coefficient [18], thereby identifying the topological subtypes in the healthy population.

2.3. Statistical Analysis

For the identified topological subtypes, we calculated the sample distribution of each subtype and the descriptive statistics (mean and standard deviation) of their topological features. We then quantitatively evaluated the contribution of each topological feature to subtype differentiation using a random forest classifier (with 100 trees) [19]. In the statistical analysis, given the relatively small sample size, we performed pairwise comparisons of the subtypes obtained from clustering using the non-parametric Mann-Whitney U test [20] to assess the significance of differences in topological features between subtypes. To control for the false positive rate due to multiple comparisons, all comparisons were corrected using the Bonferroni method [21], with the correction factor based on the number of comparisons. Statistical analysis was implemented in Python, with the significance level set at 0.05. Finally, typical samples of each category were identified by calculating the Euclidean distance between the sample and the cluster center, providing a basis for the subsequent analysis of functional connectivity patterns and topological features.

3. Results

3.1. Identification of Brain Network Subtypes

K-means clustering analysis identified three distinct brain network subtypes, with a silhouette coefficient of 0.5621, indicating moderate to good separation of the clusters. Among the 48 participants, subtype 1 included 25 subjects (52.1%), subtype 2 included 13 subjects (27.1%), and subtype 3 included 10 subjects (20.8%). To further understand the characteristics of these subtypes, we conducted analyses from multiple dimensions (Figure 2).

Principal component analysis showed that the first two principal components explained 92.2% and 3.7% of the total variance, respectively (cumulative 95.9%). The three subtypes exhibited clear separation in the reduced-dimensional space (Figure 2A). The feature radar chart further revealed the differences in topological features across the three subtypes: subtype 2 exhibited the highest feature values, subtype 1 showed a balanced distribution of intermediate values, while subtype 3 displayed lower feature values across most dimensions (Figure 2B). Random forest classification analysis indicated that the most important features for distinguishing between subtypes were the average birth

time (importance 0.229), average death time (importance 0.163), and life cycle standard deviation (importance 0.156) from the one-dimensional ring topological features (Figure 2C).

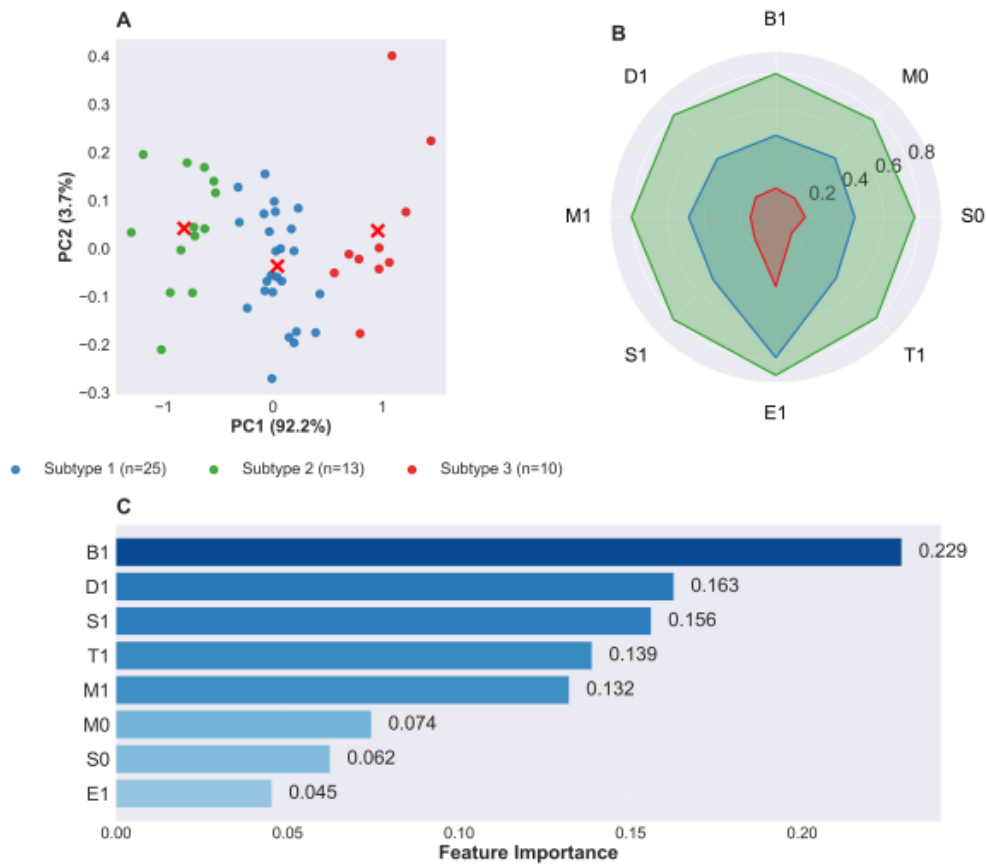


Figure 2: Feature analysis of brain network subtypes. (A) Principal component analysis, showing the distribution of the three subtypes in the 2D space (blue: subtype 1, green: subtype 2, red: subtype 3), with red cross marks indicating cluster centers; (B) Feature radar chart, displaying the distribution of each subtype in the topological feature dimensions; (C) Feature importance ranking derived from random forest classification.

3.2. Topological Subtype Feature Analysis

As shown in Figure 3, the comparison results of the differences in topological features between subtypes indicate that all subtypes exhibit significant differences across all topological feature dimensions ($p < 0.001$, Bonferroni correction). In terms of zero-dimensional topological features, the statistical characteristics of the middle 50% of connected components show a clear hierarchical pattern: subtype 2 exhibits the highest average life cycle and standard deviation, followed by subtype 1, and subtype 3 shows the lowest values. In one-dimensional topological features, temporal feature analysis revealed significant organizational differences between subtypes: subtype 2 has the latest birth and death times, forming stable and persistent ring structures; subtype 1 maintains an intermediate level; while subtype 3 shows the earliest birth and death times, indicating that its network topology is relatively unstable. This hierarchical pattern is further corroborated by life cycle-related indicators. Additionally, analysis of the standard deviation and entropy of the life cycle reveals that subtype 2 not only has the longest life cycle but also exhibits the most stable and complex ring structure features, while subtype 3 shows the weakest performance in these aspects.

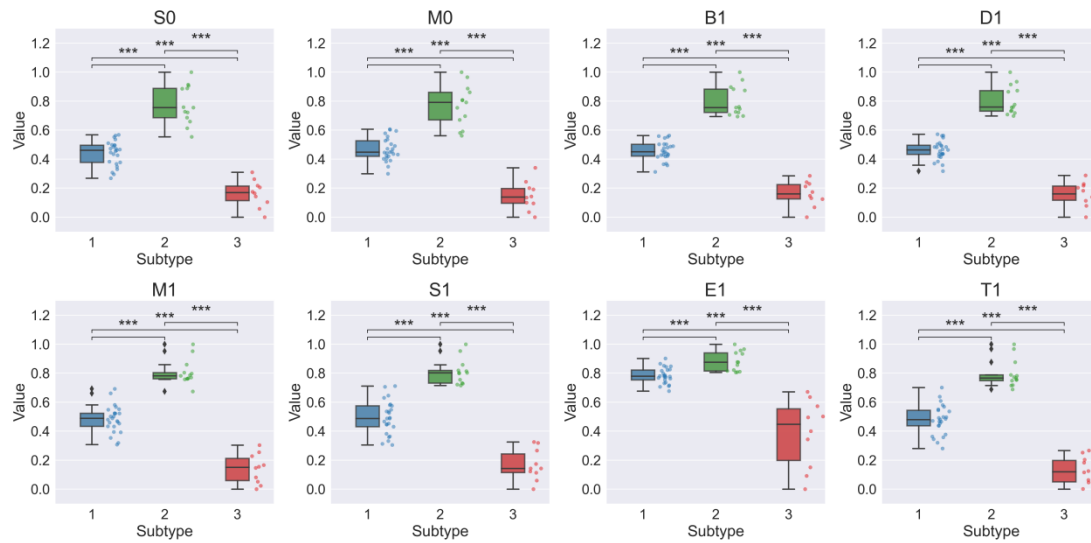


Figure 3: Comparison of topological features of brain network subtypes. Boxplots display the distribution of eight topological features across the three subtypes (blue: subtype 1, green: subtype 2, red: subtype 3), including connected component features (S0, M0) and ring structure features (B1, D1, M1, S1, E1, T1). Asterisks indicate statistical significance of group differences (***) $p < 0.001$. The scatter plots on the right side of the boxplots show the distribution of individual samples within each group.

3.3. Functional Connectivity and Topological Feature Analysis of Typical Samples

The typical samples of the three subtypes show significant differences in functional connectivity patterns and topological features (Figure 4). The representative sample of subtype 1 (subject 40) demonstrates moderate network organizational characteristics. The functional connectivity matrix displays a symmetric block pattern between the two hemispheres, with moderate connectivity ($FC \approx 0.6$) observed within the visual (VIS) and somatomotor (SMT) networks. This characteristic is reflected in the topological analysis: the barcode plot shows a moderate number of ring structures (blue bars), concentrated in the filter value range of 0.2–0.4; the feature points in the persistence diagram are located at a moderate distance from the diagonal. The typical sample of subtype 2 (subject 32) exhibits clear modular characteristics: the functional connectivity matrix shows a distinct block structure, with moderate connectivity ($FC > 0.4$) within the visual (VIS), somatomotor (SMT), and default mode (DMN) networks, while the connections between networks are generally weaker ($FC < 0.2$), forming a typical sparse connectivity pattern. This highly selective network organizational feature is reflected in the topological analysis: the barcode plot shows a large number of persistent ring structures, distributed over a wider filter value range (0.2–0.6); the feature points in the persistence diagram are distant from the diagonal and densely distributed. The typical sample of subtype 3 (subject 26) exhibits a unique network organizational pattern. Its functional connectivity matrix shows widespread high-intensity connections within networks, particularly between the visual (VIS) and somatomotor (SMT) networks on both sides, with FC values generally exceeding 0.8. Additionally, the strength of inter-network connectivity is higher than in the other two subtypes. However, its topological analysis reflects a relatively simple structure: the barcode plot shows the fewest ring structures, mainly concentrated in the low filter value range (0.1–0.2); the feature points in the persistence diagram are tightly clustered near the diagonal and sparsely distributed.

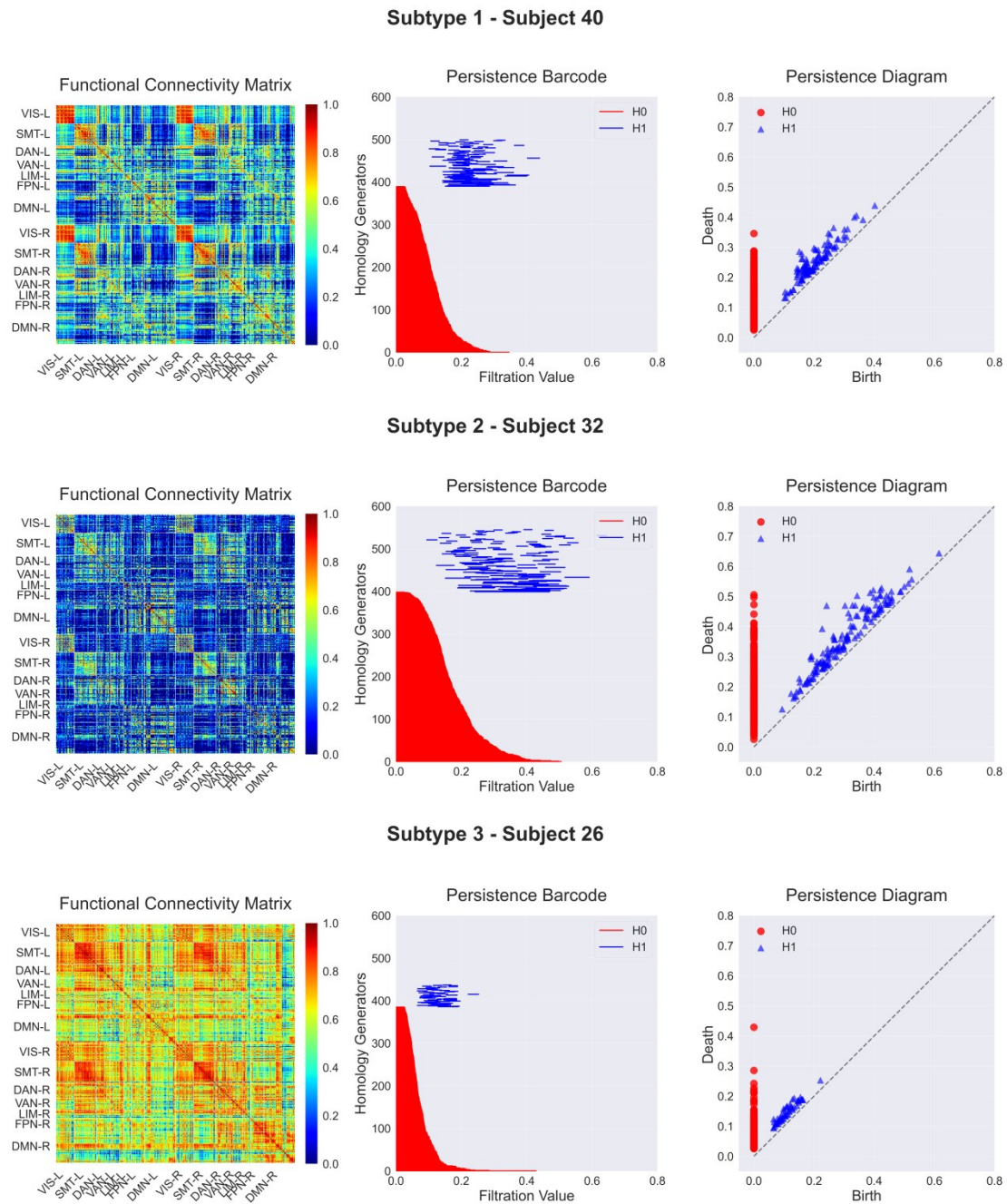


Figure 4: Comparison of typical samples across the three subtypes. Each row displays the representative sample of a subtype: the functional connectivity matrix (left) shows the correlation strength (0–1) between 400 brain regions, with seven functional networks (VIS, SMT, DAN, VAN, LIM, FPN, DMN) in each hemisphere; the persistence barcode (middle) shows red for connected components (H0) and blue for ring structures (H1); the persistence diagram (right) shows the distribution of life cycles of topological features, where the horizontal and vertical axes represent the birth and death times of the features.

4. Conclusion

This study, based on persistent homology analysis, identified three significantly different brain network subtypes in a healthy population and further analyzed the differences in topological features

and functional connectivity among these subtypes. The results indicate that these three subtypes exhibit significant differences in topological features, with subtype 1 having the highest proportion of subjects, and its topological feature statistics concentrated in the moderate range. In contrast, subtypes 2 and 3 exhibit distinct distributions of topological features, representing more complex higher-order interaction patterns. Previous studies, such as those by Gratton et al. [22], found that brain network structures in healthy individuals exhibit both group stability and significant individual differences, suggesting that brain functional networks may possess more complex higher-order interactive properties. This perspective aligns with the multi-scale complexity revealed in this study through one-dimensional ring topological features. Additionally, Martínez-Riaño et al. [23] used resting-state fMRI data and showed that the H1 feature lifetime of most individuals is concentrated within the range of 0.05 to 0.24, which is consistent with the findings for subtype 1 in this study. They also observed that some individuals had significantly longer H1 feature lifetime (>0.5), which corresponds with the features of subtype 2 in this study. These results further support the potential of persistent homology analysis in revealing brain network complexity and individual differences.

Despite the use of persistent homology-based topological analysis to reduce the limitations of threshold selection in traditional graph theory methods, there are still several limitations. First, the small sample size may restrict the generalizability of the results, so future studies with larger sample sizes are needed to validate the universality of these subtypes. Second, the choice of distance metric for calculating topological properties may influence the analysis results [24]. Different distance metrics could lead to bias in the representation of topological features, potentially affecting the precise characterization of brain network structures. Furthermore, the reliance of persistent homology analysis on data preprocessing could introduce the influence of non-neuronal noise [25]. However, the use of strict signal preprocessing procedures and multiple comparison correction strategies has mitigated these issues to some extent.

Future research could combine cognitive and behavioral indicators to explore the functional significance of the different subtypes in more detail, further validating the potential associations between subtypes and individual characteristics. This could help identify abnormal brain network patterns and provide quantitative bases for clinical diagnosis and intervention. Additionally, integrating other multimodal data (such as cognitive assessments and genetic information) with persistent homology analysis could further reveal the complexity of brain function and its individual specificity.

References

- [1] Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med*, 34(4), 537-541. doi:10.1002/mrm.1910340409
- [2] Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*, 8(9), 700-711. doi:10.1038/nrn2201
- [3] Ma, Q., Huang, B., Wang, J., Seger, C., Yang, W., Li, C., . . . Huang, R. (2017). Altered modular organization of intrinsic brain functional networks in patients with Parkinson's disease. *Brain Imaging Behav*, 11(2), 430-443. doi:10.1007/s11682-016-9524-7
- [4] Hong, S. J., Vogelstein, J. T., Gozzi, A., Bernhardt, B. C., Yeo, B. T. T., Milham, M. P., & Di Martino, A. (2020). Toward Neurosubtypes in Autism. *Biol Psychiatry*, 88(1), 111-128. doi:10.1016/j.biopsych.2020.03.022
- [5] Seitzman, B. A., Gratton, C., Laumann, T. O., Gordon, E. M., Adeyemo, B., Dworetzky, A., . . . Petersen, S. E. (2019). Trait-like variants in human functional brain networks. *Proc Natl Acad Sci U S A*, 116(45), 22851-22861. doi:10.1073/pnas.1902932116
- [6] Vriend, C., Wagenmakers, M. J., van den Heuvel, O. A., & van der Werf, Y. D. (2020). Resting-state network topology and planning ability in healthy adults. *Brain Struct Funct*, 225(1), 365-374. doi:10.1007/s00429-019-02004-6
- [7] Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, 52(3), 1059-1069. doi:10.1016/j.neuroimage.2009.10.003

- [8] Bassett, D. S., & Bullmore, E. (2006). Small-world brain networks. *Neuroscientist*, 12(6), 512-523. doi:10.1177/1073858406293182
- [9] Petri, G., Expert, P., Turkheimer, F., Carhart-Harris, R., Nutt, D., Hellyer, P. J., & Vaccarino, F. (2014). Homological scaffolds of brain functional networks. *J R Soc Interface*, 11(101), 20140873. doi:10.1098/rsif.2014.0873
- [10] Giusti, C., Ghrist, R., & Bassett, D. S. (2016). Two's company, three (or more) is a simplex : Algebraic-topological tools for understanding higher-order structure in neural data. *J Comput Neurosci*, 41(1), 1-14. doi:10.1007/s10827-016-0608-6
- [11] Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X. N., Holmes, A. J., . . . Yeo, B. T. T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex*, 28(9), 3095-3114. doi:10.1093/cercor/bhx179
- [12] Yan, C. G., Wang, X. D., Zuo, X. N., & Zang, Y. F. (2016). DPABI: Data Processing & Analysis for (Resting-State) Brain Imaging. *Neuroinformatics*, 14(3), 339-351. doi:10.1007/s12021-016-9299-4
- [13] Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*, 10(3), 186-198. doi:10.1038/nrn2575
- [14] Otter, N., Porter, M. A., Tillmann, U., Grindrod, P., & Harrington, H. A. (2017). A roadmap for the computation of persistent homology. *EPJ Data Sci*, 6(1), 17. doi:10.1140/epjds/s13688-017-0109-5
- [15] Maria, C., Boissonnat, J.-D., Glisse, M., & Yvinec, M. (2014, 2014//). *The Gudhi Library: Simplicial Complexes and Persistent Homology*. Paper presented at the Mathematical Software – ICMS 2014, Berlin, Heidelberg.
- [16] Kumar, S., Shovon, A. R., & Deshpande, G. (2023). The robustness of persistent homology of brain networks to data acquisition-related non-neural variability in resting state fMRI. *Hum Brain Mapp*, 44(13), 4637-4651. doi:10.1002/hbm.26403
- [17] Liu, H., Chen, J., Dy, J., & Fu, Y. (2023). Transforming Complex Problems Into K-Means Solutions. *IEEE Trans Pattern Anal Mach Intell*, 45(7), 9149-9168. doi:10.1109/tpami.2023.3237667
- [18] Lovmar, L., Ahlford, A., Jonsson, M., & Syvänen, A. C. (2005). Silhouette scores for assessment of SNP genotype clusters. *BMC Genomics*, 6, 35. doi:10.1186/1471-2164-6-35
- [19] Degenhardt, F., Seifert, S., & Szymczak, S. (2019). Evaluation of variable selection methods for random forests and omics data sets. *Brief Bioinform*, 20(2), 492-503. doi:10.1093/bib/bbx124
- [20] Tai, K. Y., Dhaliwal, J., & Balasubramaniam, V. (2022). Leveraging Mann-Whitney U test on large-scale genetic variation data for analysing malaria genetic markers. *Malar J*, 21(1), 79. doi:10.1186/s12936-022-04104-x
- [21] Curtin, F., & Schulz, P. (1998). Multiple correlations and Bonferroni's correction. *Biol Psychiatry*, 44(8), 775-777. doi:10.1016/s0006-3223(98)00043-2
- [22] Gratton, C., Laumann, T. O., Gordon, E. M., Adeyemo, B., & Petersen, S. E. (2016). Evidence for Two Independent Factors that Modify Brain Networks to Meet Task Goals. *Cell Rep*, 17(5), 1276-1288. doi:10.1016/j.celrep.2016.10.002
- [23] Martínez-Riaño, D. E., González, F., & Gómez, F. (2023). $\beta(1)$ persistent features of the resting-state connectome in healthy subjects. *Netw Neurosci*, 7(1), 234-253. doi:10.1162/netn_a_00280
- [24] Skaf, Y., & Laubenbacher, R. (2022). Topological data analysis in biomedicine: A review. *J Biomed Inform*, 130, 104082. doi:10.1016/j.jbi.2022.104082
- [25] Liu, T. T. (2016). Noise contributions to the fMRI signal: An overview. *Neuroimage*, 143, 141-151. doi:10.1016/j.neuroimage.2016.09.008