

Supplementary Information for State-transition dynamics of resting-state functional magnetic resonance imaging data: Model comparison and test-to-retest analysis

Saiful Islam¹, Pitambar Khanra², Johan Nakuci⁴, Sarah F. Muldoon^{1,2,3}, Takamitsu Watanabe⁵, Naoki Masuda^{1,2*}

¹*Institute for Artificial Intelligence and Data Science, University at Buffalo, State University of New York at Buffalo, USA*

²*Department of Mathematics, University at Buffalo, State University of New York, USA*

³*Neuroscience Program, University at Buffalo, State University of New York at Buffalo, USA*

⁴*School of Psychology, Georgia Institute of Technology, USA*

⁵*International Research Centre for Neurointelligence, The University of Tokyo Institutes for Advanced Study, Tokyo, Japan*

*Corresponding author: naokimas@buffalo.edu

S1. Within-participant and between-participant reproducibility of the state-transition dynamics with $K = 7$ and $K = 10$

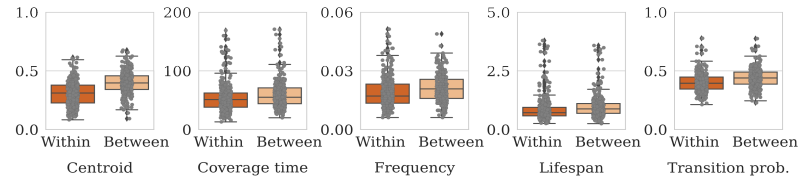
We show the distributions of the discrepancy in terms of the five observables for the within-participant and between-participant comparisons in Figs. S1 and S2 for the number of clusters $K = 7$ and $K = 10$, respectively.

S2. Robustness test with the Euclidean distance used as dissimilarity between sessions

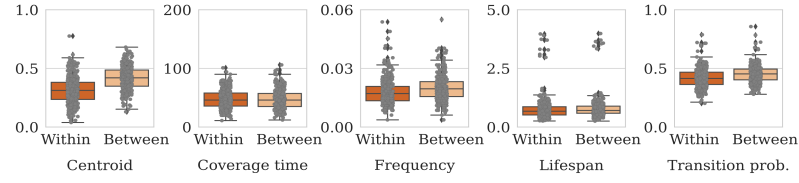
We mainly used the dissimilarity based on the $\text{corr}(\mathbf{c}_\ell, \mathbf{c}_{\ell'})$ defined by Eq. (7) in the main text to match and compare the set of the centroid positions, $\{\mathbf{c}_1, \dots, \mathbf{c}_L\}$, between pairs of sessions. In this section, we investigate the robustness of the results of the test-retest reproducibility analysis by running the same analysis with the Euclidean distance between the centroid positions, defined by Eq. (14) in the main text, to match and compare the set of centroid positions between pairs of sessions. We also use the Euclidean distance to measure the discrepancy in terms of the centroid position between the two optimally matched sessions.

We show the p values for the permutation test for the three clustering methods, five discrepancy measures for the state-transition dynamics, and $K \in \{2, \dots, 10\}$ in Tables S1, S2, and Table S3 for the DMN extracted from the MSC data, the whole-brain network extracted from the MSC data, and the whole-brain network extracted from the HCP data, respectively. In the tables, $p = 0$ indicates that the deviation of the ND value for the original session-to-session comparisons from 1 is larger than the deviation

(a) K-means



(b) TAAHC



(c) Bisecting K-means

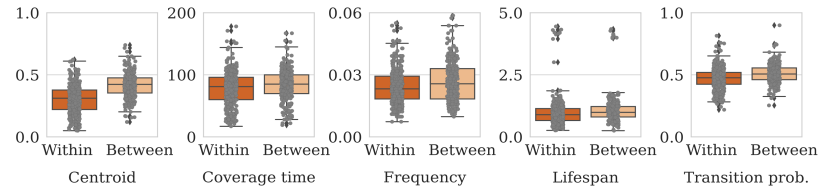


Figure S1: Within-participant and between-participant reproducibility of the state-transition dynamics with $K = 7$ states. (a) K-means. (b) TAAHC. (c) Bisecting K-means. “Within” and “Between” indicate the within-participant and between-participant comparisons, respectively. Each box plot shows the minimum, maximum, median, first quartile, and third quartile of the measurements. Each dot represents a session. “Centroid” abbreviates the centroid position, and “Transition prob.” abbreviates the transition probability matrix.

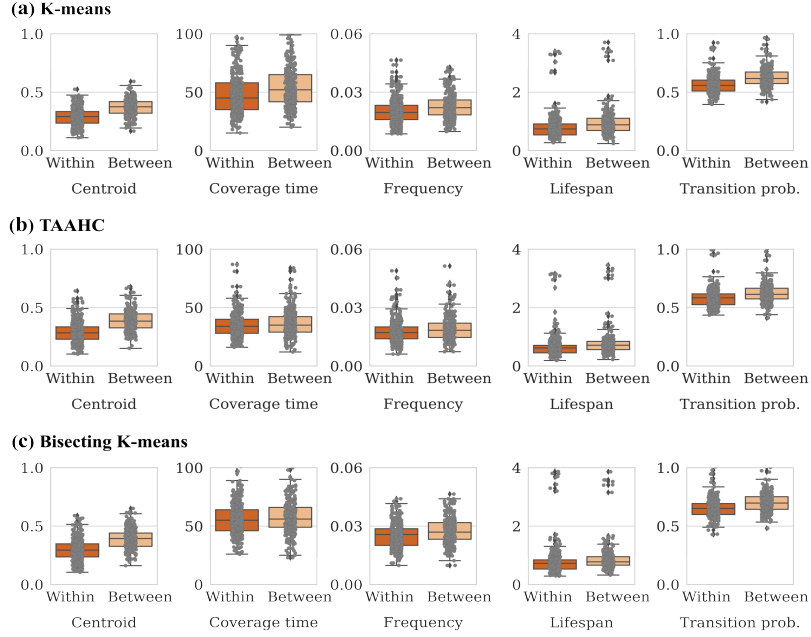


Figure S2: Within-participant and between-participant reproducibility of the state-transition dynamics with $K = 10$ states. (a) K-means. (b) TAAHC. (c) Bisecting K-means.

of the ND value for all the 10^4 randomized comparisons from 1, corresponding to $p < 0.5 \times 10^{-4}$. We find that the permutation test results obtained with the Euclidean distance are similar to those obtained with the cosine distance shown in the main text.

Quantitatively, for the DMN extracted from the MSC data, most p values (i.e., 128 out of the 135 comparisons; 94.81%) were smaller than 0.05 (indicated by *), and 111 of them (i.e., 83.07% of the 135 comparisons) remained significant after the Bonferroni correction (indicated by ***; equivalent to $p < 0.00037$, uncorrected). Moreover, we obtained $p < 0.001$, indicated with **, in 113 out of the 135 comparisons (i.e., 83.15%). We also found that the number of significant p values (with $p < 0.05$) for the K-means (45 out of 45) and bisecting K-means (44 out of 45), which differ only by 1 from each other, was somewhat larger than that for the TAAHC (42 out of 45). After the Bonferroni correction, the number of the significant p values remained larger for the K-means (43 out of 45) than for bisecting K-means (36 out of 45), which is larger than for the TAAHC (32 out of 45).

Table S1: Results of the permutation test for the DMN extracted from the MSC data when we used the Euclidean distance to calculate the distance between the centroids of clusters. The entries with 0 mean that $p < 0.5 \cdot 10^{-4}$, uncorrected. *: $p < 0.05$, uncorrected; **: $p < 0.001$, uncorrected; ***: $p < 0.05$, Bonferroni corrected (which is equivalent to $p < 0.00037$, uncorrected). We remark that “Centroid” and “Transition prob.” abbreviate the centroid position and transition probability matrix, respectively.

	K	Centroid	Coverage	Frequency	Lifespan	P
K-means	2	0***	0.0186*	0***	0***	0***
	3	0***	0***	0***	0***	0***
	4	0***	0.0014*	0***	0***	0***
	5	0***	0***	0***	0***	0***
	6	0***	0***	0***	0***	0***
	7	0***	0***	0***	0***	0***
	8	0***	0***	0***	0***	0***
	9	0***	0***	0***	0***	0***
	10	0***	0***	0***	0***	0***
TAAHC	2	0***	0.5528	0.0035*	0.0036*	0.002*
	3	0***	0.1027	0***	0.0026*	0.0001***
	4	0***	0.0435	0***	0***	0***
	5	0***	0.0842	0.0001***	0***	0***
	6	0***	0.0076*	0***	0***	0***
	7	0***	0.5548	0***	0***	0***
	8	0***	0.0718*	0***	0***	0***
	9	0***	0.0434*	0***	0***	0***
	10	0***	0.2352	0***	0***	0***
Bisecting K-means	2	0***	0.0287*	0***	0***	0***
	3	0***	0.0001***	0***	0***	0***
	4	0***	0.0584	0***	0***	0***
	5	0***	0.0419*	0.0003***	0***	0***
	6	0***	0.018*	0.0001***	0.0004**	0***
	7	0***	0.0005**	0***	0***	0***
	8	0***	0.0037*	0.0002***	0.0001***	0***
	9	0***	0.0072*	0.0001***	0***	0***
	10	0***	0.0067*	0***	0***	0***

Table S2: Results of the permutation test for the whole-brain network extracted from the MSC data when we used the Euclidean distance to calculate the distance between the centroids of clusters. The entries with 0 mean that $p < 0.5 \cdot 10^{-4}$, uncorrected. *: $p < 0.05$, uncorrected; **: $p < 0.001$, uncorrected; ***: $p < 0.05$, Bonferroni corrected.

	K	Centroid	Coverage	Frequency	Lifespan	P
K-means	2	0***	0.0368*	0***	0***	0***
	3	0***	0.0402*	0***	0***	0***
	4	0***	0.0353*	0***	0***	0***
	5	0***	0***	0***	0***	0***
	6	0***	0.0007**	0***	0***	0***
	7	0***	0.0016*	0***	0***	0***
	8	0***	0.0004**	0***	0***	0***
	9	0***	0***	0***	0***	0***
	10	0***	0.0003***	0***	0***	0***
TAAHC	2	0***	0.0776	0.0001***	0***	0.0001***
	3	0***	0.0463*	0***	0.0003***	0***
	4	0***	0.0940	0***	0.0011*	0.0004**
	5	0***	0.0813*	0***	0***	0***
	6	0***	0.7868	0***	0.0011*	0***
	7	0***	0.8686	0***	0***	0***
	8	0***	0.0875	0***	0***	0***
	9	0***	0.0403*	0***	0***	0***
	10	0***	0.036*	0***	0***	0***
Bisecting K-means	2	0***	0.0303*	0***	0***	0***
	3	0***	0.2482	0***	0.0041*	0.0174*
	4	0***	0.0576	0***	0***	0***
	5	0***	0.2599	0.0002***	0.0003***	0.0483*
	6	0***	0.0658	0***	0***	0***
	7	0***	0.6427	0.0002***	0***	0.0001***
	8	0***	0***	0***	0***	0***
	9	0***	0.0905	0.0018*	0***	0***
	10	0***	0.0031*	0***	0***	0***

Table S3: Results of the permutation test for the whole-brain network extracted from the HCP data when we used the Euclidean distance to calculate the distance between the centroids of clusters. The entries with 0 mean that $p < 0.5 \cdot 10^{-4}$, uncorrected. *: $p < 0.05$, uncorrected; **: $p < 0.001$, uncorrected; ***: $p < 0.05$, Bonferroni corrected.

	K	Centroid	Coverage	Frequency	Lifespan	P
K-means	2	0***	0.0023*	0***	0***	0***
	3	0***	0***	0***	0***	0***
	4	0***	0***	0***	0***	0***
	5	0***	0***	0***	0***	0***
	6	0***	0***	0***	0***	0***
	7	0***	0***	0***	0***	0***
	8	0***	0***	0***	0***	0***
	9	0***	0***	0***	0***	0***
	10	0***	0***	0***	0***	0***
TAAHC	2	0***	0***	0***	0***	0***
	3	0***	0.0952	0***	0***	0***
	4	0***	0.0159*	0***	0***	0***
	5	0***	0.1318	0***	0***	0***
	6	0***	0.0191*	0***	0***	0***
	7	0***	0.0096*	0***	0***	0***
	8	0***	0.0007**	0***	0***	0***
	9	0***	0.0003***	0***	0***	0***
	10	0***	0.0008**	0***	0***	0***
Bisecting K-means	2	0***	0.0016*	0***	0***	0***
	3	0***	0.013*	0***	0***	0***
	4	0***	0***	0***	0***	0***
	5	0***	0***	0***	0***	0***
	6	0***	0.0001***	0***	0***	0***
	7	0***	0.0118*	0***	0***	0***
	8	0***	0***	0***	0***	0***
	9	0***	0***	0***	0***	0***
	10	0***	0***	0***	0***	0***