

# Functional brain networks reflect spatial and temporal autocorrelation

Received: 27 April 2022

Accepted: 14 March 2023

Published online: 24 April 2023

 Check for updates

Maxwell Shinn<sup>1,2</sup>, Amber Hu<sup>3</sup>, Laurel Turner<sup>3</sup>, Stephanie Noble<sup>1,4</sup>,  
Katrín H. Preller<sup>1,5</sup>, Jie Lisa Ji<sup>1,2</sup>, Flora Moujaes<sup>5</sup>, Sophie Achard<sup>1,6</sup>,  
Dustin Scheinost<sup>1,4</sup>, R. Todd Constable<sup>1,4</sup>, John H. Krystal<sup>1,2</sup>,  
Franz X. Vollenweider<sup>1,5</sup>, Daeyeol Lee<sup>7,8,9,10</sup>, Alan Anticevic<sup>1,2</sup>,  
Edward T. Bullmore<sup>11,13</sup> & John D. Murray<sup>1,2,12,13</sup> 

High-throughput experimental methods in neuroscience have led to an explosion of techniques for measuring complex interactions and multi-dimensional patterns. However, whether sophisticated measures of emergent phenomena can be traced back to simpler, low-dimensional statistics is largely unknown. To explore this question, we examined resting-state functional magnetic resonance imaging (rs-fMRI) data using complex topology measures from network neuroscience. Here we show that spatial and temporal autocorrelation are reliable statistics that explain numerous measures of network topology. Surrogate time series with subject-matched spatial and temporal autocorrelation capture nearly all reliable individual and regional variation in these topology measures. Network topology changes during aging are driven by spatial autocorrelation, and multiple serotonergic drugs causally induce the same topographic change in temporal autocorrelation. This reductionistic interpretation of widely used complexity measures may help link them to neurobiology.

As neuroscience data become more complex, so do the methods used to analyze them. Are sophisticated methods necessary for sophisticated data? We focus here on network analysis of functional connectivity (FC)<sup>1</sup>, the matrix of pairwise correlations across brain regions in resting-state functional magnetic resonance imaging (rs-fMRI). FC is an extremely successful tool for understanding brain function<sup>2</sup>. Non-linear methods from network neuroscience build upon the success of FC by unlocking properties observable only at higher levels of abstraction<sup>3</sup>. In these methods, FC is interpreted as a network using graph theory, where

nodes represent brain regions and edges represent the correlation in activity between those regions over time. Network analysis enables the study of brain topology at a higher level of abstraction, using statistics called ‘graph metrics’ to look at properties such as clustering, modular organization and regional influence of brain-wide activity<sup>4</sup>.

Network neuroscience has profoundly influenced how we understand the organization of rs-fMRI activity across a variety of different domains<sup>5</sup>. One such example is aging, a complex process that impacts many regions of the brain differently. A rich literature shows that

<sup>1</sup>Interdepartmental Neuroscience Program, Yale University, New Haven, CT, USA. <sup>2</sup>Department of Psychiatry, Yale University, New Haven, CT, USA.

<sup>3</sup>Yale College, Yale University, New Haven, CT, USA. <sup>4</sup>Department of Radiology and Biomedical Imaging, Yale University, New Haven, CT, USA.

<sup>5</sup>Department of Psychiatry, Psychotherapy and Psychosomatics, University Hospital for Psychiatry, Zurich, Switzerland. <sup>6</sup>University of Grenoble Alpes, CNRS, Inria, Grenoble INP, LJK, Grenoble, France. <sup>7</sup>Zanvyl Krieger Mind/Brain Institute, Johns Hopkins University, Baltimore, MD, USA. <sup>8</sup>Kavli Discovery Neuroscience Institute, Johns Hopkins University, Baltimore, MD, USA. <sup>9</sup>Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD, USA. <sup>10</sup>Department of Neuroscience, Johns Hopkins University, Baltimore, MD, USA. <sup>11</sup>Department of Psychiatry, University of Cambridge, Cambridge, UK. <sup>12</sup>Department of Physics, Yale University, New Haven, CT, USA. <sup>13</sup>These authors contributed equally: Edward T. Bullmore, John D. Murray.

 e-mail: [john.murray@yale.edu](mailto:john.murray@yale.edu)

network topology reflects numerous changes in functional organization as we age, including changes in local and global connectivity across the brain<sup>5,6</sup>. In addition to observational and cross-sectional variation in network topology, causal pharmacological manipulations have also been shown to impact within-subject topology, an essential ingredient for translational applications. For example, the psychedelic serotonergic receptor agonist lysergic acid diethylamide (LSD) causally impacts the modular organization and global integration of brain networks<sup>7,8</sup>. However, because graph theoretical analysis involves several levels of abstraction, it is unclear if the differences in topology within or across individuals are associated with differences in simpler time series properties. This can make it challenging to draw neurobiological conclusions about the brain using changes in MRI-derived graph metrics. Is it possible that lower-level factors, such as basic statistical properties of rs-fMRI time series, may mediate the changes in network topology?

The basic statistical properties of rs-fMRI time series vary within and across individuals<sup>9</sup>. Two of these properties, spatial autocorrelation (SA) and temporal autocorrelation (TA), are highly present in neuroimaging data and have striking consequences for the statistical analysis of data<sup>10,11</sup>. Qualitatively, SA captures the idea that two data points have similar values if they are nearby in space, whereas TA captures the idea that they have similar values if they are close in time. As a result, SA and TA represent statistical dependencies across variables, which often manifest in ways that are difficult to predict. Although many different methods exist to measure SA, we quantify it as the rate at which similarity decays across space, detailed below. Likewise, despite many competing ways to measure TA, we use the correlation coefficient of neighboring time points in a time series, also detailed below. SA and TA are attractive features because they reflect a myriad of physical and biological mechanisms, including molecular<sup>11,12</sup>, structural<sup>13,14</sup>, activity state<sup>15,16</sup> and organizational<sup>17,18</sup> properties of the brain. However, despite the importance of SA and TA on the mechanistic and organizational level, and the known consequences for statistical data analysis, it is largely unknown whether individual variation in SA and TA reliably relates to individual variation in more complex measures, such as those from network neuroscience.

Here we show that two simple time series statistics, SA and TA, explain a large fraction of individual variation in network topology. We demonstrate that variation in SA and TA is reliable across subjects and brain regions and that individual variation in SA and TA corresponds to individual variation in graph metrics.

To understand the association among SA, TA and graph metrics, we developed a spatiotemporal surrogate time series model that is parameterized only by SA and TA. After fitting the model to time series from individual subjects, we found a large similarity in graph topology between the subject-derived and model-derived networks, demonstrating the direct influence of SA and TA on graph theoretical analysis. To examine how this impacts data analysis in practice, we measure the changes in network topology with aging. Not only are SA and TA more sensitive to age-related changes than graph metrics, but our model can decompose the age-related changes in graph metrics into changes in SA and TA. They are also sensitive to age-related cognitive decline.

Finally, we show that TA captures within-subject variation and is causally linked to biological processes beyond a graph theoretical framework. Causal pharmacological manipulations with two different serotonergic receptor agonists show the same topographic patterns of decrease in TA, which graph metrics do not capture. Additionally, a serotonin receptor antagonist shows an increase in TA along the same topographic pattern, demonstrating a common but opposite effect from a common but opposite molecular mechanism of action. Overall, we show that SA and TA are important properties of rs-fMRI with close ties to functional networks, offering a reductionistic perspective on complex brain network topology.

## Results

### Low dimensionality of graph topology measures

We found that many measures of graph topology are correlated with each other across subjects in networks constructed from rs-fMRI. We analyzed multiple neuroimaging datasets using diverse methodologies implemented by different teams, focusing on the Human Connectome Project (HCP) dataset<sup>19</sup>, with validation in the Yale Test–Retest (Yale-TRT)<sup>20</sup> and the Cambridge Centre for Ageing and Neuroscience (Cam-CAN)<sup>21</sup> datasets. In line with previous work, networks were constructed by thresholding the FC matrix to maintain only the 10% strongest connections and ensuring that no nodes are disconnected from the rest of the network. We used several graph metrics to quantify network topology: assortativity, global efficiency, transitivity, modularity, mean clustering coefficient and mean local efficiency. We also considered two nodal graph metrics: degree and centrality (see Methods for details).

In all datasets, most graph metrics are highly correlated with most other graph metrics across subjects (Fig. 1b)<sup>22,23</sup>. However, unweighted graph metrics, derived from binarized networks, are also highly correlated with the mean (mean-FC), variance (var-FC) and kurtosis (kurt-FC) of FC (Fig. 1b). This is surprising because the FC binarization procedure destroys all explicit information about mean-FC, var-FC and kurt-FC (see Methods for proof). These observations were consistent across datasets (Extended Data Fig. 1). Thus, some unobserved underlying factors must influence both the statistical moments of FC and the topology of the unweighted graph. In what follows, we demonstrate that SA and TA are two such factors.

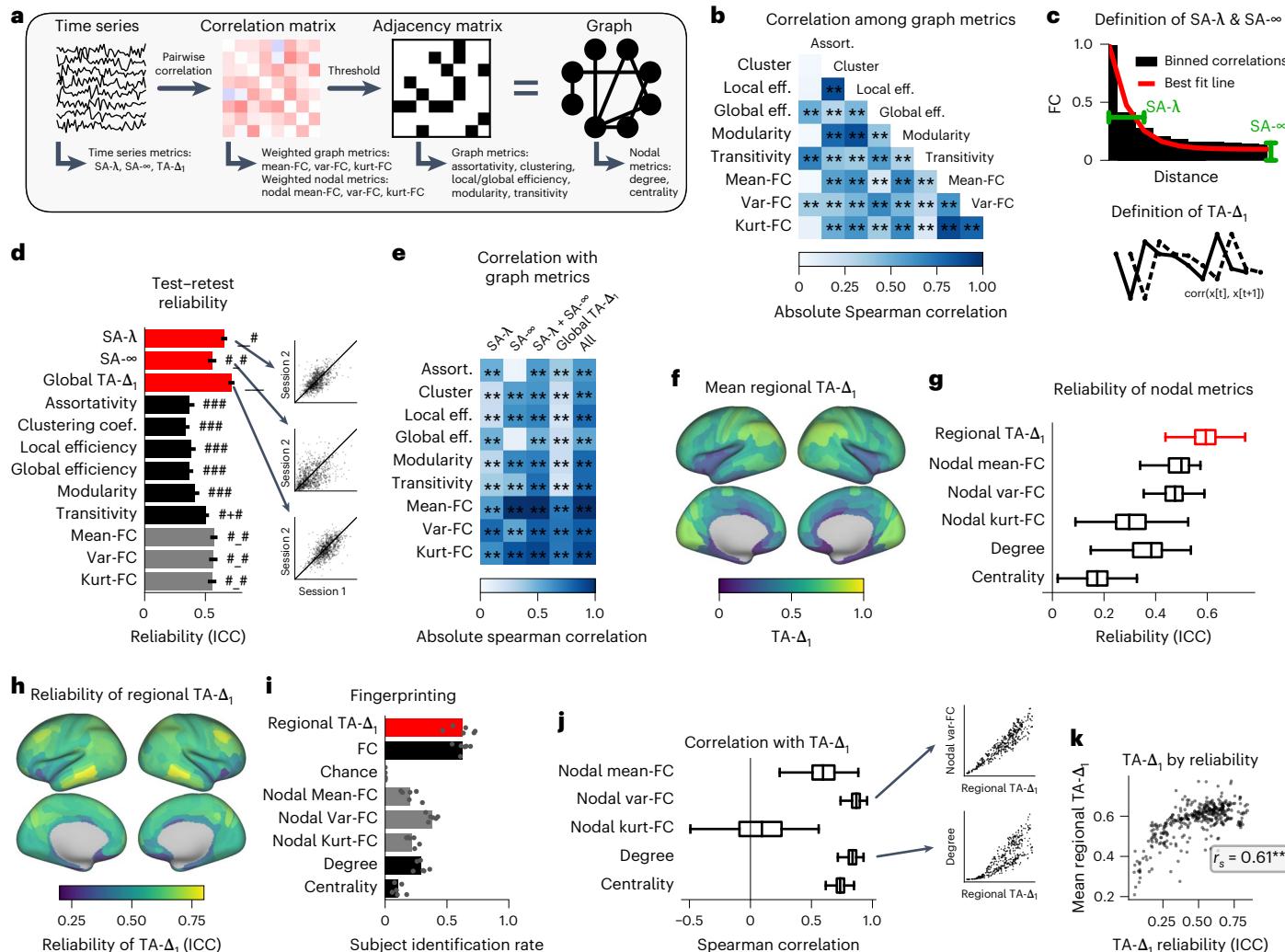
### Autocorrelation in space

SA is the ubiquitous but often ill-defined phenomenon in neuroscience that nearby regions are more similar than distant regions<sup>18,24</sup>. To examine SA's test–retest reliability—that is, how well SA is preserved across different rs-fMRI sessions from the same subject—we developed a method to quantify SA on a single-subject level by decomposing it into two components: the rate at which FC falls off with physical distance (SA-λ) and the average correlation between two distant brain regions (SA-∞). Our method bins FC by distance and finds the best fit SA-λ and SA-∞ for each subject's FC versus distance curve (Fig. 1c), using Euclidean distance to accommodate both hemispheres and subcortex (Equation 1) (Methods). Test–retest reliability is quantified with intraclass correlation coefficient (ICC), which compares the variability of multiple observations of the same subject to the variability across all subjects, where 1 is perfect reliability and 0 is chance. Historically, different qualitative assessments have been assigned to ICCs<sup>25,26</sup>, with values from 0.5 to 0.7 usually being called ‘moderate’ to ‘good’.

We found that SA is both reliable and correlated with graph metrics across subjects. Both SA-λ and SA-∞ have high test–retest reliability compared to graph metrics (Fig. 1d) (SA-λ greater than all metrics,  $P < 0.01$ , and SA-∞ greater than unweighted graph metrics,  $P < 0.05$ , bootstrap resampling). The effect cannot be explained by head motion within the scanner (accounting for motion, partial correlation  $> 0.65$ ,  $P < 10^{-10}$  for both SA-λ and SA-∞) or global signal regression (GSR) (Extended Data Fig. 1) and is strongest in the HCP dataset, which had the smallest motion and parcel size confounds (Extended Data Fig. 2). Both SA-λ and SA-∞ are highly correlated with weighted and unweighted graph metrics (Fig. 1e). To test whether the combined influence of SA-λ and SA-∞ is greater than either one alone, we constructed a linear model including SA-λ, SA-∞ and a constant term, training it on 50% of subjects selected randomly to fit each graph metric. This model significantly predicts graph metrics on the held-out data (Fig. 1e). These results hold in all datasets (Extended Data Fig. 1). Therefore, SA is reliable and topologically informative.

### Autocorrelation in time

TA describes the smoothness, or memory, of the rs-fMRI time series over time and is known to vary heterogeneously across brain regions<sup>27,28</sup> and



**Fig. 1 | SA and TA are important features of rs-fMRI time series.** **a**, Diagram describing the connectomic pipeline and graph metrics. **b**, Correlation across all subjects in the HCP dataset, with Bonferroni FWER-corrected two-sided *P* values. **c**, Schematic demonstrating the calculation of SA-λ and SA-∞ (top) and global TA-Δ<sub>1</sub> (bottom). To compute SA-λ and SA-∞, FC values are binned across distances *D*, and the curve  $SA-\infty + (1 - SA-\infty) \exp(-D/SA-\lambda)$  is fit to the binned data. To compute global TA-Δ<sub>1</sub>, the Pearson correlation is taken between the time series and the time series shifted by one time point. **d**, Test-retest reliability of graph metrics, quantified by ICC. Error bars indicate 95% CI. After each bar is a string of three characters indicating significance: the first indicates significantly less than SA-λ, the second SA-∞ and the third global TA-Δ<sub>1</sub>, where # indicates *P* < 0.01, + indicates *P* < 0.05 and – indicates *P* > 0.5, by a two-sided bootstrap resampling procedure. Inset scatter plots show correlation across subjects for two example sessions. **e**, Correlation across subjects between graph metrics and SA-λ, SA-∞ or

global TA-Δ<sub>1</sub>. ‘SA-λ + SA-∞’ and ‘all’ indicate a cross-validated linear model with two or three terms, respectively, using *n* = 334 unrelated subjects to avoid relatedness confounds and Bonferroni FWER-corrected two-sided *P* values. **f**, The brain map depicting regional TA-Δ<sub>1</sub>, averaged across all subjects. **g**, Distribution of reliability for each brain region. **h**, Reliability of regional TA-Δ<sub>1</sub> is plotted across brain regions. **i**, Mean fraction of subjects correctly identified by a fingerprinting analysis. Points indicate identification performance on each of six possible test-retest pairs from the four sessions. **j**, Correlation across regions of regional TA-Δ<sub>1</sub> with nodal graph metrics for each subject. **k**, The regional TA-Δ<sub>1</sub> for each region, from **f**, is plotted against its reliability, from **h**. For all subpanels, unless otherwise indicated, *n* = 883 subjects, \* indicates *P* < 0.05 and \*\* indicates *P* < 0.01, and box plots indicate the median, first and third quartiles and range, with outliers hidden for visualization. FWER, family-wise error rate.

influence both FC<sup>16,29</sup> and graph topology<sup>30,31</sup>. We quantify TA for each parcellated region in each subject as the Pearson correlation between adjacent time points of the region’s time series—that is, the lag-1 temporal autocorrelation (TA-Δ<sub>1</sub>) (Fig. 1c). TA-Δ<sub>1</sub> is a simple non-parametric measure of TA. Although TA-Δ<sub>1</sub>, in theory, measures only correlation across a single time point in the time series, in practice, it was effective in measuring correlations at much longer time scales, including parametric estimates of long-memory dynamics (Extended Data Fig. 3)<sup>32</sup>. Because pre-processing methodology and TR both influence TA-Δ<sub>1</sub>, comparisons between TA-Δ<sub>1</sub> must be made within a single dataset. TA-Δ<sub>1</sub> can be measured at either the level of individual regions or as an average of these at the whole-brain level, so we refer to these as ‘regional

TA-Δ<sub>1</sub>’ and ‘global TA-Δ<sub>1</sub>’, respectively. We observed highest regional TA-Δ<sub>1</sub> in occipital and parietal regions and lowest in limbic regions (Fig. 1f).

TA is reliable across subjects, both at the whole-brain level and the regional level. At the whole-brain level, we computed each subject’s global TA-Δ<sub>1</sub> by averaging regional TA-Δ<sub>1</sub> across all regions and found that global TA-Δ<sub>1</sub> is highly reliable compared to graph metrics (Fig. 1d). For each region, we also computed the reliability across subjects and found that median regional reliability was higher for regional TA-Δ<sub>1</sub> than for other nodal graph metrics (Fig. 1g). The reliability of regional TA-Δ<sub>1</sub> varies heterogeneously across the brain (Fig. 1h)<sup>33</sup>.

This reliability and heterogeneity suggests that regional TA-Δ<sub>1</sub> could be used to identify individual subjects across the population.

To identify subjects, we used a ‘fingerprinting’ analysis<sup>34</sup>. For a given measure, such as regional TA- $\Delta_1$ , we matched each session of each subject to the session with the highest Pearson correlation in that measure across regions, selecting among all sessions from all subjects. Then, we counted the number of pairs for which both sessions belonged to the same subject. We found that fingerprinting using regional TA- $\Delta_1$  identified the single matching session from the pool of 1,765 sessions in over 62.4% of subjects, compared to 62.6% with a more traditional fingerprinting analysis using FC (Fig. 1*i*). Although this is lower than reported in ref. 34, the number of possible matches was nearly an order of magnitude larger in our dataset and is consistent with previous findings<sup>35</sup>. We compared this to fingerprinting using several nodal graph metrics, which were unable to match the performance (Fig. 1*i*). Similar results were obtained in other datasets (Extended Data Fig. 1). We also tested whether this performance is due to differences in parcel boundaries or reliable inhomogeneity within the parcel by comparing regional TA- $\Delta_1$  to the parcel’s regional homogeneity. Regional homogeneity showed a moderate correlation with regional TA- $\Delta_1$ , and its ability to identify individual subjects through fingerprinting varied greatly between datasets but had no ability to identify subjects in the HCP dataset (Extended Data Fig. 4). This indicates that regional TA- $\Delta_1$  is reliable enough to identify an individual subject from a population.

TA- $\Delta_1$  is highly correlated with graph topology at the individual and regional levels. At the individual level, we found a strong correlation between global TA- $\Delta_1$  and graph metrics (Fig. 1*e*). To test its influence in conjunction with SA, we developed a linear model incorporating global TA- $\Delta_1$ , SA- $\lambda$ , SA- $\infty$  and a constant, training it on 50% of randomly selected subjects to predict each graph metric. This model significantly predicted all graph metrics (Fig. 1*e*) and almost all graph metrics in the other datasets (Extended Data Fig. 1). At the regional level, for each subject, we computed the Spearman correlation between regional TA- $\Delta_1$  and various graph metrics. We found that both weighted and unweighted nodal graph metrics were highly correlated with regional TA- $\Delta_1$  (Fig. 1*j*). Most notably, a node’s degree, or the total number of connections a node makes to other regions, was predicted by regional TA- $\Delta_1$  with a median correlation of 0.89 (Fig. 1*j*), and this relationship was not driven by parcel size (median partial correlation 0.83 between degree and regional TA- $\Delta_1$ , accounting for parcel size). Remarkably, this implies that network hubs can be discovered without examining the topology of network<sup>31,36</sup>.

### Multiple sources of TA

Heterogeneity in TA across brain regions can be shaped by multiple underlying factors. One such factor is the region’s intrinsic time scale, which measures the rate at which a signal in a region decays over time. Intrinsic time scales vary from region to region and have been previously described as a central factor of brain dynamics in both fMRI and electrophysiology<sup>27,28</sup>. If intrinsic time scale was the primary driver of TA heterogeneity, regions with longer time scales would lead to higher TA. However, brain maps of intrinsic time scale<sup>27,28</sup> do not correspond well to our brain map of regional TA- $\Delta_1$  (Fig. 1*f*). This suggests that other factors may be important in driving TA.

In addition to intrinsic time scale, the noise level of a brain region also influences its TA<sup>37</sup>. Decreasing the signal-to-noise ratio or increasing the noise level will decrease TA. More generally, adding white noise to any time series will cause its TA to approach 0 (see Supplementary Information for proof). Thus, because different regions of the brain are affected by different sources and quantities of noise, these differences may also influence TA. In the extreme case where noise is the only driving factor of TA, all regions would have an identical slow intrinsic time scale, and different amounts of noise in each region would determine the region’s TA. Consistent with the hypothesis that noise shapes regional TA, we found that regions with the lowest regional TA- $\Delta_1$  show the lowest reliability in regional TA- $\Delta_1$  (Fig. 1*k*). By this logic, TA may influence FC by altering the fraction of shared variance between pairs of regions<sup>38</sup>.

In what follows, we used this principle to build a spatiotemporal model for generating surrogate time series with regionally heterogeneous noise. We compare this with the ‘intrinsic time scale with SA’ model whereby intrinsic time scale is the primary driver of TA. The main practical difference between these models is that, in the spatiotemporal model, reductions in TA have a corresponding reduction in the local effect of SA (due to increased noise), whereas, in the ‘intrinsic time scale with SA’ model, SA is independent of TA by construction. We also compare the spatiotemporal model with models that use only SA or only TA, testing the importance of the joint action between SA and TA. Finally, we compare the spatiotemporal model to several popular null models from the graph theory literature.

### Spatiotemporal model for surrogate time series

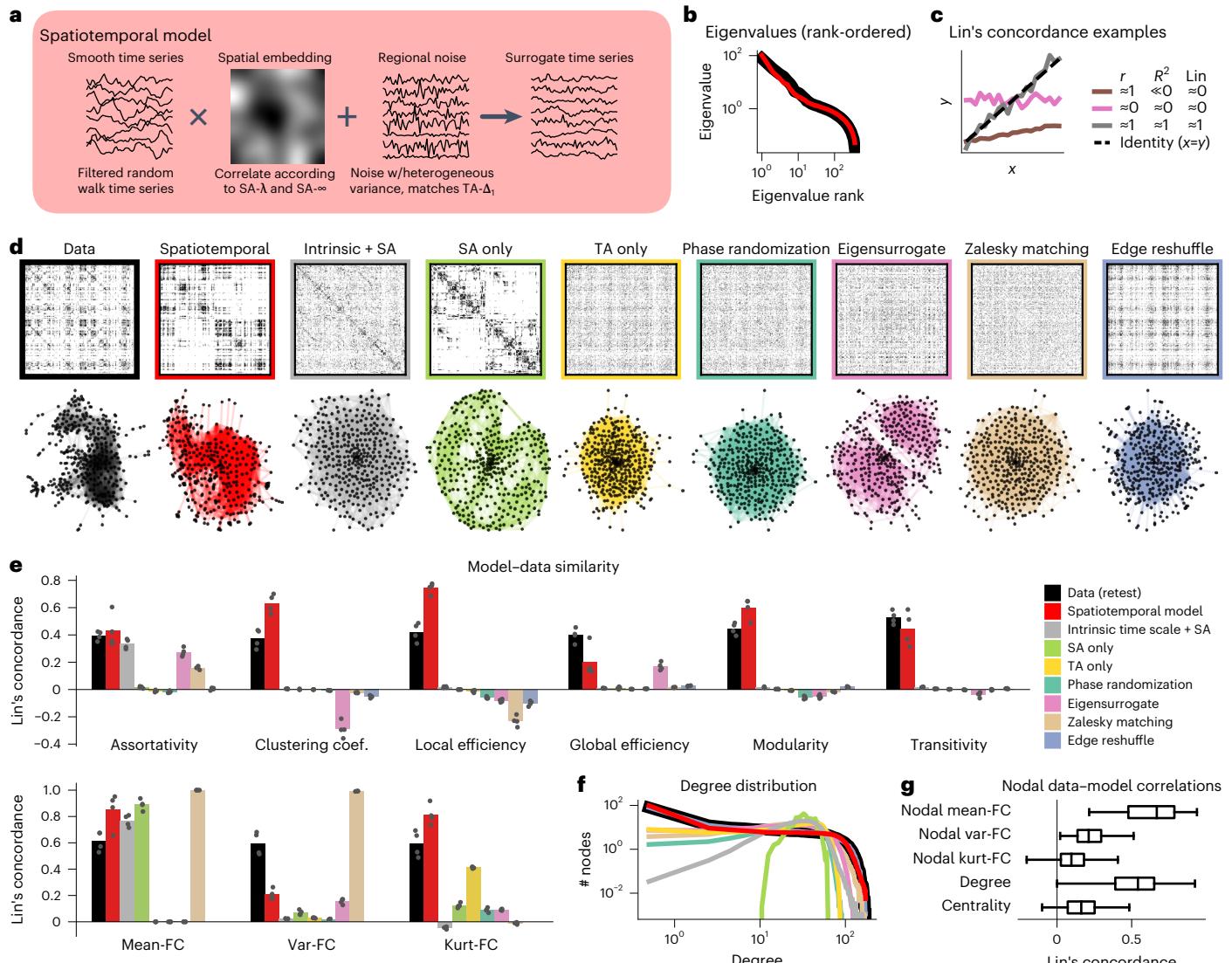
We designed a spatiotemporal surrogate time series model that, when fit to individual subjects, uses SA and TA to capture individual variation in network topology. The model operates at the level of the parcelated time series, meaning that comparisons may be drawn at multiple levels of the analysis pipeline (Fig. 1*a*). Our model can be summarized in a small number of steps (Fig. 2*a*). First, we generate long-memory (1/f<sup>a</sup> spectrum) time series that have uniformly high TA. Concurrently, we spatially embed these time series according to brain geometry, introducing SA by increasing the correlation of nearby regions, while preserving the frequency spectra. This is made possible with our correlated spectral sampling algorithm (Methods). Then, we add uncorrelated white noise with a region-specific variance, thereby lowering TA heterogeneously. The resulting time series can then be analyzed the same way as rs-fMRI time series. Thus, in our model, all variation in graph topology must be caused by variation in SA and/or TA.

Our model contains two SA parameters that are fit on the single-subject level, corresponding to noiseless SA- $\lambda$  (SA- $\lambda^{\text{gen}}$ ) and noiseless SA- $\infty$  (SA- $\infty^{\text{gen}}$ ). We also use the subject’s observed regional TA- $\Delta_1$  to determine the variance of noise to add. We fit SA- $\lambda^{\text{gen}}$  and SA- $\infty^{\text{gen}}$  through optimization to the distribution of eigenvalues of the FC matrix, a property that is not a graph metric yet is well captured by the model (Fig. 2*b*). This parameterization ensures that all variation in graph topology must be caused by variation in SA and/or TA.

### Model performance

We evaluated the ability of our model to capture weighted and unweighted graph metrics using surrogate time series. In principle, model fit could be assessed using several different criteria, such as the model’s ability to capture as much variance as possible (for example, using  $R^2$ ) or, alternatively, by its ability to capture individual variation (for example, using Pearson correlation) (Supplementary Table 1). Here, we assess model fit using Lin’s concordance, a stricter criterion that is 1 only when the model captures both variance and individual variation and 0 or negative when either assessment is poor. Fig. 2*c* demonstrates a schematic comparison among Lin’s concordance, correlation coefficient and  $R^2$  on artificial data. An example subject’s FC matrix and network diagram demonstrate that the spatiotemporal model is the only model we tested that produces the complex, asymmetric patterns observed in the data (Fig. 2*d*).

Our model captures important features of graph topology. The model exhibits a high Lin’s concordance to the data for both weighted and unweighted graph metrics, indicating that it matches both the individual variability as well as the precise values of the graph metrics (Fig. 2*e*). The model–data match is close to the total reliable variability, which we estimate using the Lin’s concordance between two independent sessions from the same subject (Fig. 2*e*). The high Lin’s concordance corresponds to a close fit to the diagonal of graph metrics plotted between the model and the data (Fig. 3). As expected, we also confirmed the model matches the SA and TA of the original subjects (Extended Data Fig. 5). Furthermore, it reproduces the degree distribution (Fig. 2*f*) and also correlates with several nodal graph metrics on the subject level (Fig. 2*g*).



**Fig. 2 | A spatiotemporal model captures connectome topology.** **a**, Schematic of our spatiotemporal modeling framework for surrogate time series with SA and regionally heterogeneous TA. **b**, log–log distribution of the eigenvalues of the FC matrix for an example subject (black) compared to the spatiotemporal model (red). **c**, Schematic demonstrating Lin’s concordance, our model fit statistic, on example data. Values close to 1 indicate both a correlation and a close match in value (gray). Values close to 0 indicate either no correlation (pink) or a large squared error (brown). **d**, Example FC matrices (top) and graphs (bottom) for the original data (black) and for each model (colors). Graphs are visualized using

a force-directed layout, which positions topographically neighboring nodes nearby. **e**, Lin’s concordance between model and data for each model. Bars represent the mean across the four scanning sessions, and points indicate the Lin’s concordance between the model and data for each session. For comparison, black indicates Lin’s concordance between separate sessions from the same subject, where dots indicate pairs of sessions. **f**, log–log degree distribution for each model compared to the data (black). **g**, Distribution of Lin’s concordance of nodal metrics between model and data for each region. Box plots show median, first and third quartiles and range, with outliers hidden for visualization,  $n = 883$ .

To test alternative ways that graph metrics might be related to time series properties, we fit several additional models, depicted graphically in Extended Data Fig. 6. First, we tested whether both SA and TA are necessary within our model. To do this, we fit the spatiotemporal model with TA set to 0 ('SA only' model) as well as the spatiotemporal model with an infinitely small SA- $\lambda$  and a SA- $\infty$  of 0 ('TA only' model) and compared their performance to the performance of our spatiotemporal model. Second, we designed a model to test the hypothesis that TA is determined by the region’s estimated intrinsic time scale in conjunction with SA ('Intrinsic time scale + SA'). This model estimates a pink noise  $1/f^\alpha$  exponent for each region, as well as SA- $\lambda$  and SA- $\infty$  for the full brain, and then generates surrogates using correlated spectral sampling (Methods). Third, we tested whether a full reconstruction of TA at all scales, via the power spectrum

amplitudes with randomized phases ('Phase randomization'), can match our model’s performance. Likewise, given our findings related to mean-FC and var-FC, we tested a model that matches these two statistics ('Zalesky matching'), previously shown to exhibit more brain-like network topology<sup>30</sup>. Additionally, we tested a model that matches the network’s degree distribution ('Edge reshuffle') due to its overwhelming popularity in the graph theory literature. Finally, to confirm whether our excellent fit was driven by our use of eigenvalues in the loss function, we also created an 'eigensurrogate' method for creating surrogate FC matrices. Eigensurrogates preserve the eigenvalue distribution of the FC matrix but randomize the eigenvectors, thereby perfectly duplicating the linear dimensionality of the FC matrix. This tests the null hypothesis that an effect is due to restrictions in dimensionality.



**Fig. 3 | Correlation of model and data graph metrics for all models.** For each model, each subject's empirical graph metrics are plotted against the graph metrics predicted by the spatiotemporal model. These scatter plots depict the

relationships summarized by Lin's concordance in Fig. 2e for the HCP dataset. Spearman correlation ( $r_s$ ) and Lin's concordance (Lin) are inset.

\* indicates Spearman correlation two-tailed  $P < 0.05$ , and \*\* indicates  $P < 0.01$ .

None of these alternative models captures network topology (Figs. 2e and 3 and Extended Data Fig. 5) or the degree distribution (Fig. 2f) as well as the spatiotemporal model. These results are not affected by the use of geodesic instead of Euclidean distance (Extended Data Fig. 7). Across all other datasets, only the eigensurrogate model

on the Yale-TRT dataset showed more explanatory power than the spatiotemporal model (Extended Data Fig. 8 and Supplementary Table 1). The ability of our spatiotemporal model to generally fit the Yale-TRT and Cam-CAN datasets better than alternative models underscores its generality.

## Linking autocorrelation to graph topology

Our results show that SA and TA predict graph topology, but establishing this relationship mathematically is difficult due to the non-linearity of constructing and analyzing a thresholded graph. Instead, we attempted to provide an intuition of why SA and TA may influence topology by considering their impact on individual edges. The impact of SA is relatively straightforward—SA increases the mean correlation between nearby regions and, thus, increases the probability of an edge. However, it is not immediately clear how TA might influence topology.

A statistical argument explains the high correlation between regional TA- $\Delta_1$  and degree (Fig. 1j)<sup>31,36</sup> and shows why strong TA creates hubs. Degree is determined by the number of correlations the node makes that exceed the binarization threshold (Fig. 1a). This means that, even if the expected value of the correlation (nodal mean-FC) is low, there may still be several correlations above the threshold if it has a high variance (nodal var-FC). In other words, a high variance may increase the probability of crossing the threshold more so than a high mean. Thus, any process that increases a node's variance should also increase the node's degree. Two temporally autocorrelated time series are expected to have a higher variance in their pairwise correlation<sup>31</sup> (Supplementary Information 1.5), and this relationship is reflected in our data (Fig. 1e,j). Thus, for individual nodes, one way that TA drives a node's degree is by increasing the var-FC. If region A is highly correlated with regions B and C, it is also more likely that B and C are correlated with each other<sup>30</sup>. This means that high-TA- $\Delta_1$  nodes in the graph are more likely to share an edge<sup>31,39</sup>, creating a clustered network topology.

We confirmed this reasoning using a two-parameter ‘economical clustering’ model<sup>23</sup>, which builds graphs directly by probabilistically connecting nodes based on their distance and clustering topology (Supplementary Information 2). This model is known to reproduce several topological features of brain networks<sup>23,40</sup> and is convenient because the procedure for constructing networks is distinct from the rs-fMRI pipeline. We found that changes in SA and TA correspond to an increased propensity for short and clustered edges, respectively, in the economical clustering model (Extended Data Fig. 9). In other words, the parameters of a graph-level generative process known to reproduce topological features of rs-fMRI networks are closely linked to the dimensions of network topology spanned by SA and TA. Thus, SA and TA can be interpreted directly in terms of graph topology.

## SA in healthy aging

We have shown that our surrogate model can trace differences in graph metrics among subjects down to differences in the SA and TA of the time series used to produce them. Therefore, we asked if the opposite is possible—if graph metrics change across a population, is SA or TA responsible for those changes in graph metrics? Due to the rich history of graph theoretical analysis of aging<sup>5,6</sup>, we looked for age-related changes in graph metrics, and we show that perturbations in the SA- $\infty$  parameter, but not the SA- $\lambda$  or global TA- $\Delta_1$  parameters, lead to a change in graph metrics that mirrors that of aging.

We analyzed the Cam-CAN dataset, containing cross-sectional rs-fMRI data from over 800 subjects ranging from age 18 years to 90 years. The relationships among SA, TA and graph metrics were unchanged despite the large variation in age (Extended Data Fig. 1), and our surrogate model remained effective in this dataset (Extended Data Fig. 8). Because motion was highly correlated with age in this dataset (Extended Data Fig. 2), we performed all analyses using partial correlation controlling for motion. Several weighted and unweighted graph metrics are correlated with age (Fig. 4), with global efficiency, var-FC and kurt-FC showing significant correlations. Age was positively correlated with SA- $\lambda$ , negatively correlated with SA- $\infty$  and uncorrelated with global TA- $\Delta_1$  (Fig. 4a).

Our spatiotemporal model can be used to determine the extent to which SA- $\lambda$  and SA- $\infty$  govern the change in network topology. Because both SA- $\lambda$  and SA- $\infty$  have significant partial correlations with

age, we ran the model ‘in reverse’ to understand which of these two parameters mediated the effects of graph metrics on aging. To do this, we perturbed each of these parameters in the direction predicted by aging. If graph metrics increase or decrease in the same direction as they do across age, this provides evidence that the perturbed parameter mediates the effects of aging. Because SA- $\infty$  decreases with age, we perturbed the model by decreasing SA- $\infty$ <sup>gen</sup> to obtain predictions for SA- $\infty$ -mediated aging. We found that all of the graph metrics that show significant correlation with age change in response to this perturbation, and the direction of this change matches the direction of the change over age (Fig. 4b). We also perturbed the model by increasing SA- $\lambda$ <sup>gen</sup> to obtain predictions for SA- $\lambda$ -mediated aging. However, graph metrics showed only a slight change in response to this perturbation, and the change was in the opposite direction as age (Fig. 4b). This suggests that the impact of aging on network structure is mediated by changes in SA- $\infty$  rather than SA- $\lambda$ . In other words, the effect of aging on graph metrics is driven by the baseline level of SA at long distances rather than the rate at which SA decays at short distances.

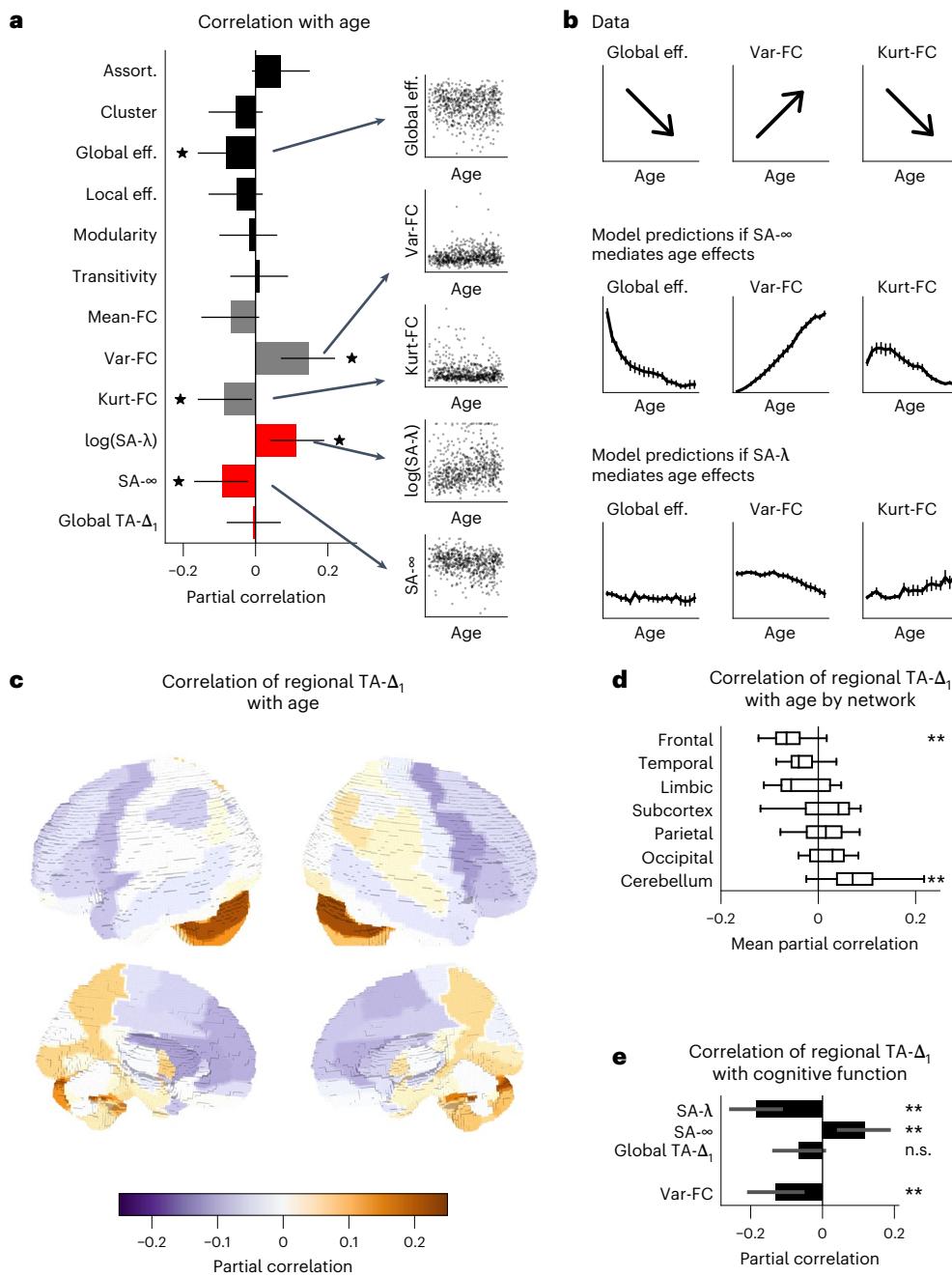
Because the global TA did not change with age, we asked whether this was also true on a regional level. We computed the partial correlation of age and regional TA- $\Delta_1$  across the brain (Fig. 4c). Regional TA- $\Delta_1$  decreased with age in the frontal subnetwork but increased with age in cerebellar regions (Fig. 4d), suggesting a difference in how these regions change over age. This difference may reflect age-related structural changes along an anterior-posterior gradient<sup>41</sup>. Despite small effect size, this demonstrates that age-related changes in SA and TA are consistent with age-related changes in graph metrics.

## Subclinical markers of dementia

Although SA and TA relate closely to network topology, are they useful in explaining clinical symptoms beyond network topology? We explored the relationship of SA and TA with early symptoms of dementia, asking whether SA or TA predict cognitive decline beyond the effect of healthy aging. To assess cognitive function, we used the Addenbrookes Cognitive Examination Revisited (ACE-R), a battery of cognitive tests for dementia screening in subclinical populations. Because dementia was an exclusion criterion for participation in the Cam-CAN study, we did not expect to find a relationship between dementia markers and SA or TA. We computed the partial correlation across subjects of ACE-R score with SA- $\lambda$ , SA- $\infty$  and global TA- $\Delta_1$ , partial on age and motion. Surprisingly, we found a significant negative partial correlation with SA- $\lambda$  ( $r = -0.184$ ,  $P < 10^{-5}$ ), associating wider SA with reduced cognitive function (Fig. 4e). There was also a weaker correlation with SA- $\infty$  ( $r = 0.118$ ,  $P = 0.002$ ) but no significant relationship with global TA- $\Delta_1$  ( $r = -0.067$ ,  $P = 0.09$ ) (Fig. 4e). Of all the graph metrics, only the weighted graph metric var-FC ( $r = -0.13$ ,  $P = 0.001$ ) showed a significant relationship (others,  $P > 0.1$ ). This highlights the sensitivity of SA- $\lambda$  and SA- $\infty$  compared to graph metrics. The SA- $\lambda$ -driven effect also contrasts with that of healthy aging, which was primarily driven by SA- $\infty$ . These findings hint at a relationship of clinical symptoms with SA and TA<sup>42</sup>.

## Pharmacological manipulation

Lastly, we tested whether SA and TA causally reflect neurobiological processes. In principle, it is possible that TA and SA are driven exclusively by noise, morphology or other reliable artifacts instead of differences in brain dynamics. A within-subject pharmacological study allows conclusions about acute changes with a causal mechanism. Therefore, in two human pharmacological fMRI experiments, we measured changes in SA and TA caused by manipulation of a neural circuit with the serotonin receptor agonists LSD<sup>7</sup> and psilocybin<sup>43</sup>. In both experiments, subjects were administered drug or placebo on separate days in a double-blind methodology, and rs-fMRI was performed at early and late time points after each administration. The same basic relationships among SA, TA and graph metrics were preserved under the drugs (Extended Data Fig. 10), despite a small sample size less than 3% that of HCP.



**Fig. 4 | SA links functional connectome topology to neurobiology during aging.** **a**, Partial correlation of graph metrics with age is shown, controlling for motion. Asterisks indicate two-tailed significance of partial correlation ( $P < 0.05$ ,  $n = 652$ ). Error bars show 95% CI. **b**, The spatiotemporal model can be perturbed to test whether each SA parameter mediates aging. Top: Subject data predict global efficiency and kurt-FC decrease with age, whereas var-FC increases with age. The SA-∞ and SA-λ parameters were separately perturbed in the spatiotemporal model according to how they change with age. The mean predicted change of each metric with age is shown if the age-related changes are due to SA-∞ alone (middle) or SA-λ alone (bottom). Error bars show standard error across ten runs of the model. The direction (positive or negative) of the change with age corresponds to the data for SA-∞. **c**, Partial

correlation of regional TA- $\Delta_1$  with age is shown for each region, partial on motion.

**d**, Mean partial correlation of regional TA- $\Delta_1$  with age across brain regions, partial on motion,  $n = 646$  subjects. Box plots indicate the median, first and third quartiles and outlier-excluded range of partial correlation distribution.

\*\* indicates  $P < 0.01$  significant partial correlation, after Bonferroni FWER correction. **e**, Partial correlation of measures with the ACE-R assessment, partial on motion and age. Top: partial correlation with SA and global TA. SA-λ and SA-∞ are significant after Bonferroni FWER correction, and global TA- $\Delta_1$  is not. Bottom: Var-FC was the only weighted or unweighted graph metric that was significant after Bonferroni FWER correction. Error bars show 95% CI. \*\* indicates  $P < 0.01$ , and \* indicates  $P < 0.05$ , two-sided Wilcoxon signed-rank test with Bonferroni FWER correction. FWER, family-wise error rate.

We found that both LSD and psilocybin caused robust overall reductions in TA across cortex. LSD and psilocybin reduced cortical TA at both early and late time points, in regional-averaged and subject-averaged TA (Wilcoxon rank-sum test  $P < 10^{-40}$  for all conditions) (Fig. 5a). Subject-level drug effects on global TA- $\Delta_1$  were not

due to differences in within-scanner motion, and no effect of drugs on SA and graph metrics was detectable for this sample size (Extended Data Fig. 10).

Psychedelic effects of LSD and psilocybin are predominantly attributed to agonism of serotonin receptors, in particular the

5-HT<sub>2A</sub> receptor<sup>7,43</sup>. A common serotonergic mechanism for LSD and psilocybin should produce similar cortex-wide topographies of the change in regional TA-Δ<sub>1</sub>. Indeed, we found significant positive correlation among the cortical topographies for both drugs and both time points (Fig. 5c).

Because both of these 5-HT<sub>2A</sub> agonists produced a specific topographic pattern of reductions in regional TA-Δ<sub>1</sub>, we predicted that a 5-HT<sub>2A</sub> antagonist would produce the same pattern but as an increase in regional TA-Δ<sub>1</sub> instead of a decrease. In the LSD study, on a third visit, subjects were pre-administered ketanserin, a selective 5-HT<sub>2A</sub> antagonist, before receiving LSD (LSD+Ket). We found that, at the late time point, ketanserin strongly attenuated the effect of LSD on regional TA-Δ<sub>1</sub> (Fig. 5a). At the early time point, there was a cortex-wide increase in regional TA-Δ<sub>1</sub> (Fig. 5a). Furthermore, regions with the strongest decrease under LSD showed the strongest increase with pre-administration of ketanserin (Fig. 5b). These time-dependent changes in TA are consistent with observed pharmacokinetics of ketanserin, which exhibits relatively fast decrease in plasma levels after initial administration<sup>44</sup>.

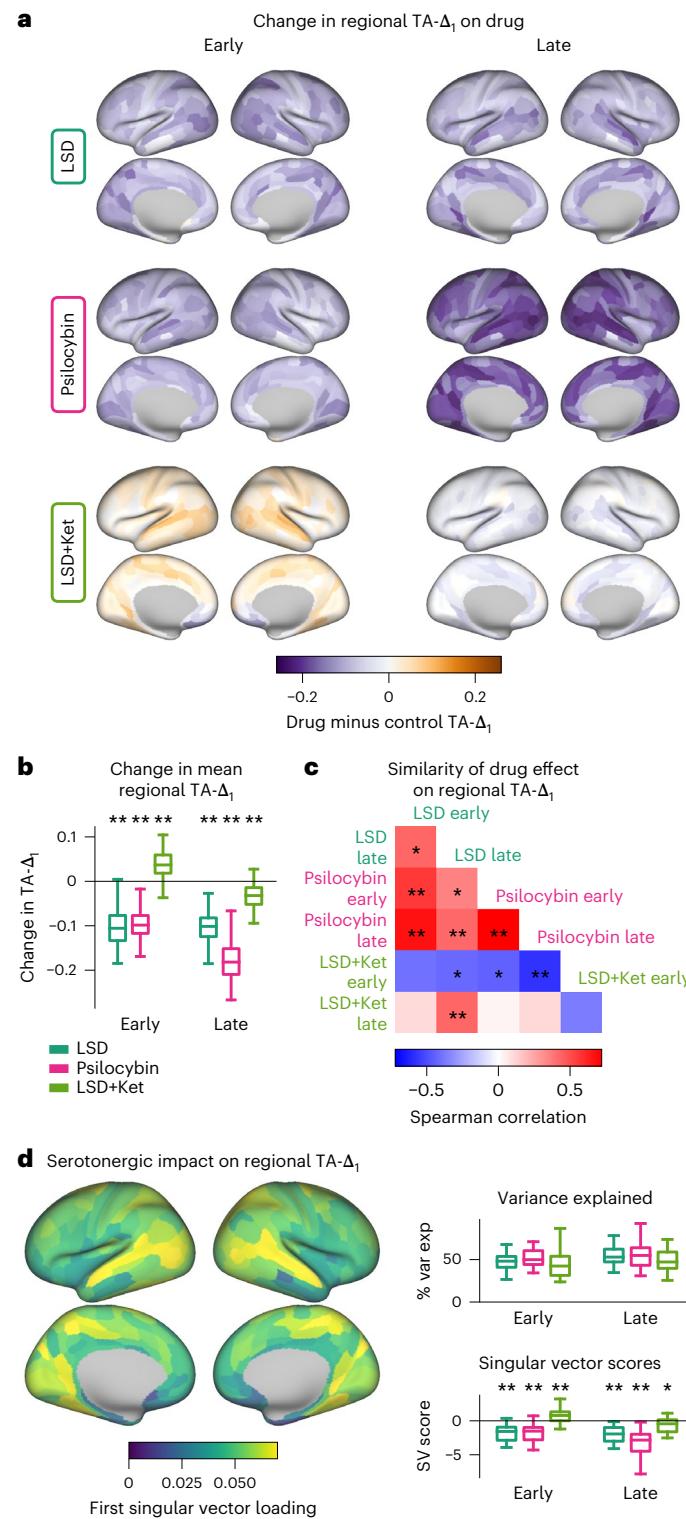
We used these experiments to construct a cortical map of overall regional TA-Δ<sub>1</sub> modulation by serotonergic drugs. To incorporate information from all experimental conditions, we used singular value decomposition (SVD) to compute the first singular vector across all drug versus control contrasts from all participants. We found that a single map (Fig. 5d, left) explained approximately 50% of the variance of individual subjects for each of the three experiments at both time points (Fig. 5d, top right). This map weighted negatively on the LSD and psilocybin conditions and on the late LSD+Ket condition but positively on the early LSD+Ket condition (Fig. 5d, bottom right), consistent with their correlational structure (Fig. 5c). These findings demonstrate that TA is sensitive to causal pharmacological perturbation by serotonergic drugs, showing that TA reflects meaningful differences in neurobiology.

## Discussion

Here we have shown that SA and TA—as parameterized by SA-λ, SA-∞ and TA-Δ<sub>1</sub>—are highly reliable properties of rs-fMRI time series that correlate with, and are predictive of, network topology. When we use a model to generate surrogate time series with subject-matched SA and TA, the time series produce networks that match the subject's graph metrics. The surrogate time series model was also used to track age-related changes in graph metrics, and SA and TA correlate with subclinical symptoms of dementia, even though graph metrics are not sensitive enough to detect this change. Furthermore, causal pharmaceutical manipulations with multiple serotonergic drugs modulate TA in a distinct and reliable pattern with a large within-subject effect size. We anticipate that comparable relationships with SA and TA will be present in other datasets stemming from a variety of condi-

tions, diseases and pharmacological states. The high reliability, interpretability, effect size, clinical relevance and sensitivity to neurobiology make SA and TA serious candidates for fMRI-based biomarkers<sup>33,42</sup>.

What factors drive SA and TA? We showed SA and TA can be influenced by biological factors, such as aging; pharmacological factors, such as serotonergic drugs; and methodological factors, such as motion, parcel size and noise. These findings are consistent with previous work in brain signal variability during aging<sup>45</sup> as well as in non-serotonin neuromodulators, such as norepinephrine<sup>12,46</sup>,



**Fig. 5 | Pharmacological administration of serotonergic drugs causes changes in TA.** **a**, Cortical maps showing the mean difference in regional TA-Δ<sub>1</sub> between drug and placebo at early (left) and late (right) time points. Purple indicates lower regional TA-Δ<sub>1</sub> in the drug than placebo conditions (LSD  $n = 24$ , psilocybin  $n = 23$ ). **b**, Change in the regional TA-Δ<sub>1</sub> mean across subjects. Box plots show the median, first and third quartiles and outlier-excluded range of the distribution of regional TA-Δ<sub>1</sub> mean across subjects for different regions. \* indicates  $P < 0.05$ , and \*\* indicates  $P < 0.01$ , two-sided Wilcoxon signed-rank test. **c**, Similarity, measured by Spearman correlation, between 'drug minus placebo' cortical maps for each drug and time point. \* indicates  $P < 0.05$ , and \*\* indicates  $P < 0.01$ , two-tailed permutation test correcting for SA<sup>11</sup> and Bonferroni–Holm correction for FWER. **d**, The first right singular vector (SV) across all drug minus control contrasts for all subjects. Right top: variance explained by the first SV averaged across each drug condition. Right bottom: the SV score averaged across each condition. Box plots show the median, first and third quartiles and outlier-excluded range of the distribution. \* indicates  $P < 0.05$ , and \*\* indicates  $P < 0.01$ , two-sided Wilcoxon signed-rank test,  $n = 646$  subjects. FWER, family-wise error rate.

dopamine<sup>47</sup> and acetylcholine<sup>48</sup>. They are also consistent with methodological studies, showing the statistical considerations of SA and TA<sup>10,49</sup> and the ability of null models to form complex networks<sup>30,50</sup>. SA is particularly important for the emerging literature on diverse cortical gradients<sup>47</sup>. Furthermore, SA and TA are influenced by local circuit connectivity and properties such as intrinsic time scale<sup>27,51</sup>. The influence of confounding physiological processes on SA and TA must be better understood<sup>52,53</sup>, including the link between TA and signal variability<sup>54</sup>, especially related to aging<sup>45</sup>. The hemodynamic response function imposes not only temporal but also spatial filtering to the BOLD signal<sup>55</sup>, which could contribute to changes in SA and TA due to age<sup>56</sup> or serotonergic drugs<sup>57</sup>. Because anatomical structure serves as a scaffold for SA and TA<sup>13,14</sup>, our test-retest reliability and fingerprinting performance could reflect individual differences in brain morphology<sup>58,59</sup>. Similarly, differences in functional network organization across participants could lead to reliable misalignment between parcel boundaries<sup>60,61</sup>, which would contribute to the reliability that we observed in regional TA- $\Delta_1$ . Our results showed that the regional homogeneity within a parcel, associated with many of these structural and functional differences, was positively correlated with regional TA- $\Delta_1$  but had varied ability to explain fingerprinting performance. In the HCP dataset, it could not explain fingerprinting performance at all, whereas, in the Yale-TRT dataset, it had better fingerprinting performance than regional TA- $\Delta_1$ . This suggests that the strongest underlying influences on TA- $\Delta_1$  vary across datasets, which may explain the difference in model performance between the HCP and Yale-TRT datasets. TA also has a complex relationship with methodological factors, such as repetition time (TR). On the one hand, a shorter TR causes samples to be more closely spaced in time, meaning that, for smooth 1/f-like power spectra, TA will increase. On the other hand, shorter TR may also result in a noisier signal, causing TA to decrease. These differences, as well as the influence of pre-processing strategy, must be studied systematically to compare TA across studies with different TRs.

In general, SA and TA constrain the dimensionality of the neural signal, and these dimensionality constraints may explain their link to network topology. Previous work has shown that lower-dimensional subspaces capture network topology<sup>23,40</sup>, and we showed that SA and TA align with these dimensions through the economic clustering model. These dimensions may map onto other aspects of network topology—for example, a high-degree backbone and a lattice-like background<sup>39</sup>. Graph metrics provide complex non-linear projections of these dimensions, making it possible to compare topological properties between networks. The eigensurrogate model tested the broader question of whether another FC matrix with equivalent linear dimensionality could reproduce graph metrics. Although it did not perform as well as our time-series-based model on most datasets, it performed surprisingly well and even outperformed our model on the Yale-TRT dataset, emphasizing the role of constrained dimensionality in shaping network topology. Variations on the eigensurrogate method, including constraints on the eigenvectors, could yield insights into individual variation in brain topology.

Our present study used a parcellated analysis, which limits the precision with which SA can be measured. Our methods for computing both SA and TA scale well for voxel-level or vertex-level analyses, but fitting the spatiotemporal model is computationally intractable for parcellations with many nodes. SA- $\infty$  is conceptually similar to mean-FC, which has known links to aging<sup>58,62</sup>. It is also similar to global signal, but it remains reliable even after GSR and, thus, may represent spatial inhomogeneities in the global signal. Nevertheless, even this parcellated analysis is sensitive enough to reveal features inaccessible to graph theoretical analysis. For example, in our pharmacological experiments, our sample size of fewer than 25 participants did not show significance in graph-theoretical measures, but changes in TA revealed highly significant differences in functional organization.

Our work highlights the need to always study complex properties, such as graph metrics, alongside simpler properties, such as SA and TA. In future studies using graph metrics, specific hypotheses must guide the careful interpretation of graph metrics in light of SA and TA. Our work also highlights that graph metrics derived from rs-fMRI networks cannot be directly interpreted as signatures of regional communication or information processing. Our results may extend more generally beyond graph theory to other analyses of FC, including both resting-state networks<sup>32,42</sup> and task-related changes<sup>15,16</sup>. The use of our spatiotemporal surrogate model—and of previously published spatially informed<sup>24,63</sup> and temporally informed<sup>30</sup> null models—could be applied more generally to high-dimensional statistics in neuroimaging. Our work also highlights the informativeness of SA and TA about biological processes such as aging and serotonergic tone. Historically, SA and TA have been considered confounds that need to be corrected for. Our results suggest that SA and TA should not be treated as confounds but, rather, as essential informative properties of the connectome.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41593-023-01299-3>.

## References

- Buckner, R. L., Krienen, F. M. & Yeo, B. T. T. Opportunities and limitations of intrinsic functional connectivity MRI. *Nat. Neurosci.* **16**, 832–837 (2013).
- Smith, S. M. et al. Functional connectomics from resting-state fMRI. *Trends Cogn. Sci.* **17**, 666–682 (2013).
- Fornito, A., Zalesky, A. & Bullmore, E. *Fundamentals of Brain Network Analysis* (Academic Press, 2016).
- Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* **52**, 1059–1069 (2010).
- Betzel, R. F. et al. Changes in structural and functional connectivity among resting-state networks across the human lifespan. *Neuroimage* **102**, 345–357 (2014).
- Damoiseaux, J. S. Effects of aging on functional and structural brain connectivity. *Neuroimage* **160**, 32–40 (2017).
- Preller, K. H. et al. Changes in global and thalamic brain connectivity in LSD-induced altered states of consciousness are attributable to the 5-HT2A receptor. *eLife* **7**, e35082 (2018).
- Luppi, A. I. et al. LSD alters dynamic integration and segregation in the human brain. *Neuroimage* **227**, 117653 (2021).
- Murphy, K., Birn, R. M. & Bandettini, P. A. Resting-state fMRI confounds and cleanup. *Neuroimage* **80**, 349–359 (2013).
- Friston, K. J. et al. Analysis of fMRI time-series revisited. *Neuroimage* **2**, 45–53 (1995).
- Burt, J. B., Helmer, M., Shinn, M., Anticevic, A. & Murray, J. D. Generative modeling of brain maps with spatial autocorrelation. *Neuroimage* **220**, 117038 (2020).
- Huang, Z., Liu, X., Mashour, G. A. & Hudetz, A. G. Timescales of intrinsic BOLD signal dynamics and functional connectivity in pharmacologic and neuropathologic states of unconsciousness. *J. Neurosci.* **38**, 2304–2317 (2018).
- Sethi, S. S., Zerbi, V., Wenderoth, N., Fornito, A. & Fulcher, B. D. Structural connectome topology relates to regional BOLD signal dynamics in the mouse brain. *Chaos* **27**, 047405 (2017).
- Fallon, J. et al. Timescales of spontaneous fMRI fluctuations relate to structural connectivity in the brain. *Netw. Neurosci.* **4**, 788–806 (2020).

15. Fascianelli, V., Tsujimoto, S., Marcos, E. & Genovesio, A. Autocorrelation structure in the macaque dorsolateral, but not orbital or polar, prefrontal cortex predicts response-coding strength in a visually cued strategy task. *Cereb. Cortex* **29**, 230–241 (2019).
16. Arbabshirani, M. R. et al. Autoconnectivity: A new perspective on human brain function. *J. Neurosci. Methods* **323**, 68–76 (2019).
17. Honey, C. J. et al. Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* **76**, 423–434 (2012).
18. Song, H. F., Kennedy, H. & Wang, X.-J. Spatial embedding of structural similarity in the cerebral cortex. *Proc. Natl Acad. Sci. USA* **111**, 16580–16585 (2014).
19. Glasser, M. F. et al. A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).
20. Noble, S. et al. Influences on the test-retest reliability of functional connectivity MRI and its relationship with behavioral utility. *Cereb. Cortex* **27**, 5415–5429 (2017).
21. Shafto, M. A. et al. The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC Neurol.* **14**, 204 (2014).
22. Li, C., Wang, H., de Haan, W., Stam, C. J. & Mieghem, P. V. The correlation of metrics in complex networks with applications in functional brain networks. *J. Stat. Mech.* **2011**, P11018 (2011).
23. Vértes, P. E. et al. Simple models of human brain functional networks. *Proc. Natl Acad. Sci. USA* **109**, 5868–5873 (2012).
24. Roberts, J. A. et al. The contribution of geometry to the human connectome. *Neuroimage* **124**, 379–393 (2016).
25. Cicchetti, D. V. & Sparrow, S. A. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *Am. J. Ment. Defic.* **86**, 127–137 (1981).
26. Koo, T. K. & Li, M. Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **15**, 155–163 (2016).
27. Murray, J. D. et al. A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663 (2014).
28. Raut, R. V., Snyder, A. Z. & Raichle, M. E. Hierarchical dynamics as a macroscopic organizing principle of the human brain. *Proc. Natl. Acad. Sci. USA* **117**, 20890–20897 (2020).
29. Arbabshirani, M. R. et al. Impact of autocorrelation on functional connectivity. *Neuroimage* **102**, 294–308 (2014).
30. Zalesky, A., Fornito, A. & Bullmore, E. On the use of correlation as a measure of network connectivity. *Neuroimage* **60**, 2096–2106 (2012).
31. Afyouni, S., Smith, S. M. & Nichols, T. E. Effective degrees of freedom of the Pearson's correlation coefficient under autocorrelation. *Neuroimage* **199**, 609–625 (2019).
32. Shafiei, G. et al. Topographic gradients of intrinsic dynamics across neocortex. *eLife* **9**, e62116 (2020).
33. Keitel, A. & Gross, J. Individual human brain areas can be identified from their characteristic spectral activation fingerprints. *PLoS Biol.* **14**, e1002498 (2016).
34. Finn, E. S. et al. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* **18**, 1664–1671 (2015).
35. Waller, L. et al. Evaluating the replicability, specificity, and generalizability of connectome fingerprints. *Neuroimage* **158**, 371–377 (2017).
36. Baria, A. et al. Linking human brain local activity fluctuations to structural and functional network architectures. *Neuroimage* **73**, 144–155 (2013).
37. Noble, S., Scheinost, D. & Constable, R. T. A decade of test-retest reliability of functional connectivity: a systematic review and meta-analysis. *Neuroimage* **203**, 116157 (2019).
38. Cole, M. W., Yang, G. J., Murray, J. D., Repovš, G. & Anticevic, A. Functional connectivity change as shared signal dynamics. *J. Neurosci. Methods* **259**, 22–39 (2016).
39. Achard, S. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J. Neurosci.* **26**, 63–72 (2006).
40. Betzel, R. F. et al. Generative models of the human connectome. *NeuroImage* **124**, 1054–1064 (2016).
41. Sullivan, E. V. & Pfefferbaum, A. Diffusion tensor imaging and aging. *Neurosci. Biobehav. Rev.* **30**, 749–61 (2006).
42. Watanabe, T., Rees, G. & Masuda, N. Atypical intrinsic neural timescale in autism. *eLife* **8**, e42256 (2019).
43. Preller, K. H. et al. Psilocybin induces time-dependent changes in global functional connectivity. *Biol. Psychiatry* **88**, 197–207 (2020).
44. Persson, B., Heykants, J. & Hedner, T. Clinical pharmacokinetics of ketanserin. *Clin. Pharmacokinet.* **20**, 263–279 (1991).
45. Grady, C. L. & Garrett, D. D. Understanding variability in the BOLD signal and why it matters for aging. *Brain Imaging Behav.* **8**, 274–283 (2013).
46. Zerbi, V. et al. Rapid reconfiguration of the functional connectome after chemogenetic locus coeruleus activation. *Neuron* **103**, 702–718 (2019).
47. Shafiei, G. et al. Dopamine signaling modulates the stability and integration of intrinsic brain networks. *Cereb. Cortex* **29**, 397–409 (2018).
48. Turchi, J. et al. The basal forebrain regulates global resting-state fMRI fluctuations. *Neuron* **97**, 940–952 (2018).
49. Davey, C. E., Grayden, D. B., Egan, G. F. & Johnston, L. A. Filtering induces correlation in fMRI resting state data. *Neuroimage* **64**, 728–740 (2013).
50. Cantwell, G. T. et al. Thresholding normally distributed data creates complex networks. *Phys. Rev. E* **101**, 062302 (2020).
51. Ito, T., Hearne, L. J. & Cole, M. W. A cortical hierarchy of localized and distributed processes revealed via dissociation of task activations, connectivity changes, and intrinsic timescales. *Neuroimage* **221**, 117141 (2020).
52. Chen, J. E. et al. Resting-state ‘physiological networks’. *Neuroimage* **213**, 116707 (2020).
53. Drew, P. J., Mateo, C., Turner, K. L., Yu, X. & Kleinfeld, D. Ultra-slow oscillations in fMRI and resting-state connectivity: neuronal and vascular contributions and technical confounds. *Neuron* **107**, 782–804 (2020).
54. Baracchini, G. et al. Inter-regional BOLD signal variability is an organizational feature of functional brain networks. *Neuroimage* **237**, 118149 (2021).
55. Kriegeskorte, N., Cusack, R. & Bandettini, P. How does an fMRI voxel sample the neuronal activity pattern: compact-kernel or complex spatiotemporal filter? *Neuroimage* **49**, 1965–1976 (2010).
56. West, K. L. et al. BOLD hemodynamic response function changes significantly with healthy aging. *Neuroimage* **188**, 198–207 (2019).
57. Cohen, Z., Bonvento, G., Lacombe, P. & Hamel, E. Serotonin in the regulation of brain microcirculation. *Prog. Neurobiol.* **50**, 335–362 (1996).
58. Geerligs, L., Tsvetanov, K. A. & Henson, R. N. Challenges in measuring individual differences in functional connectivity using fMRI: the case of healthy aging. *Hum. Brain Mapp.* **38**, 4125–4156 (2017).
59. Honey, C. J. et al. Predicting human resting-state functional connectivity from structural connectivity. *Proc. Natl Acad. Sci. USA* **106**, 2035–2040 (2009).
60. Braga, R. M. & Buckner, R. L. Parallel interdigitated distributed networks within the individual estimated by intrinsic functional connectivity. *Neuron* **95**, 457–471 (2017).

61. Kong, R. et al. Individual-specific areal-level parcellations improve functional connectivity prediction of behavior. *Cereb. Cortex* **31**, 4477–4500 (2021).
62. Balthazar, M. L. F., de Campos, B. M., Franco, A. R., Damasceno, B. P. & Cendes, F. Whole cortical and default mode network mean functional connectivity as potential biomarkers for mild Alzheimer's disease. *Psychiatry Res.* **221**, 37–42 (2014).
63. Wiedermann, M., Donges, J. F., Kurths, J. & Donner, R. V. Spatial network surrogates for disentangling complex system structure from spatial embedding of nodes. *Phys. Rev. E* **93**, 042308 (2016).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2023

## Methods

Our research complies with all relevant ethical regulations, as approved by the Washington University institutional review board, the Yale University institutional review board, the Cambridgeshire 2 Research Ethics Committee, the Cantonal Ethics Committee of Zurich and the Swiss Federal Office of Public Health.

## Datasets

We analyzed the following datasets, which comprise a diversity of pre-processing steps and experimental methodologies, including different parcellations, sampling rates, spatial and temporal smoothings, covariate regressions and noise removal strategies. Pre-processing methodology for each dataset was intended to be as faithful as possible to the ‘standard’ pipeline of the dataset, showing the generality of our conclusions across pre-processing strategies. Because data were not collected specifically for our study, we used the maximum possible sample size for available data. Our team has substantial experience with these datasets<sup>7,20,43,64,65</sup>.

**HCP.** A total of 883 subjects aged 22–37 years (459 females) from the HCP1,200-subject data release underwent four resting-state scanning sessions spread across 2 days. Subjects provided informed consent and were compensated for participation. Resting-state scans lasted for 14.4 minutes with a TR of 0.72 seconds (sampling rate of 1.39 Hz). Data were pre-processed with the HCP minimal pre-processing pipeline. This includes distortion correction using the field map, realignment for head motion and registration to T1 images. Data were further denoised using ICA-FIX, a technique based on independent component analysis to remove structured noise. A high-pass filter was applied at 0.01 Hz. Data were parcellated into 360 regions (180 per hemisphere)<sup>19</sup> with multimodal surface matching based on MSMAll (ref. 66) and with 2-mm full width at half maximum (FWHM) surface spatial smoothing constrained to the parcel. The first 100 time points were removed to ensure steady state, consistent with prior work<sup>64</sup>. Because many HCP subjects are genetically related, we used a subset of 334 unrelated subjects where indicated to avoid this potential confound. When using geodesic distance instead of Euclidean distance, we used only the right hemisphere (180 regions).

**HCP with GSR.** Subjects and scan parameters are identical to the HCP dataset, but GSR was added as a pre-processing step before parcellation. Additionally, we did not truncate the first 100 TRs, for a total of 1,200 time points. Thirty-three additional subjects were excluded due to non-convergence in the GSR pipeline, for a total of 850 subjects.

**Yale-TRT.** A total of 12 subjects (six females) aged 27–56 years were scanned on four different days, and six 6-minute sessions were performed each day. Subjects provided informed consent and were compensated for participation. Each subject was scanned on two scanners, 2 days on each scanner, on days spaced approximately 1 week apart. Thus, half of the scanning sessions for each subject were performed on a different scanner. Scanning sessions lasted 6 minutes, each with a TR of 1.0 (sampling rate of 1.0 Hz). Motion correction was applied, and images were spatially smoothed to achieve uniform spatial smoothness of a 2.5-mm Gaussian kernel<sup>67,68</sup>. Images were co-registered to a common subject-specific space across days, and subsequently into MNI space, and parcellated using the Shen parcellation<sup>69</sup>.

High-pass filtering was performed by regressing out linear, quadratic and cubic trends, and low-pass filtering was performed with a Gaussian kernel with cutoff frequency of 0.19 Hz. Mean white matter, mean cerebrospinal fluid, mean global signal and a 24-parameter motion model were also regressed out of the data. No subjects, sessions or regions were excluded.

**Cam-CAN.** A total of 652 subjects aged 18–88 years (326 females) were scanned using a TR of 1.97 seconds (sampling rate of 0.508 Hz). Subjects provided informed consent and were compensated for participation. We used the standard pre-processing pipeline and parcellation provided by the Cam-CAN project<sup>21,70</sup>. In summary, scans underwent motion correction and slice time correction before co-registration to T1 images and normalization to MNI space with DARTEL. We used the default AAL parcellation provided by the Cam-CAN project<sup>71</sup>. We also applied a second-order Butterworth low-pass filter at half the Nyquist frequency (0.127 Hz) to account for high-frequency motion artifacts. We excluded six subjects and one cerebellar region due to missing data, for a grand total of 646 subjects with 115 regions in the parcellation.

**LSD dataset.** The study used a double-blind randomized design<sup>7</sup>. Subjects provided informed consent and were compensated for participation. On each of the 3 days,  $n = 24$  participants aged 20–34 years (five females) received one of the following treatments: (1) placebo (179 mg of mannitol and 1 mg of aerosil, orally) pre-treatment followed by placebo (179 mg of mannitol and 1 mg of aerosil, orally) treatment; (2) placebo (179 mg of mannitol and 1 mg of aerosil, orally) pre-treatment followed by LSD (100 µg, orally); and (3) ketanserin (40 mg, orally) pre-treatment followed by LSD (100 µg, orally) treatment. Pre-treatment was given 60 minutes before the treatment. The ‘early’ resting-state scan occurred 75 minutes after treatment and the ‘late’ scan 300 minutes after treatment. Data were acquired with TR = 2,500 ms (sampling rate of 0.4 Hz). Twenty-five subjects were enrolled in the study. One subject was excluded for technical faults in the data, for a total of 24 subjects.

We used the same pre-processing pipeline described in ref. 7, which is summarized below. First, the data were subject to the HCP minimal pre-processing pipeline. This involved correction for field inhomogeneities, phase-encoding directions and magnetic susceptibility artifacts as well as motion correction and registration to structure images with non-brain tissue masking. Data were high-pass filtered ( $>0.008$  Hz). Several nuisance variables were regressed out: average signal in the ventricles, average signal in deep white matter, motion parameters and the mean time series across gray matter (that is, the global signal) as well as the first derivative of each of these. Data were motion scrubbed, identifying outlier frames as either frames with a summed framewise displacement over 0.5 mm or a root mean square (RMS) of differences in intensity of subsequent frames over 1.6 times the median. All outlier frames, the one frame preceding them and two frames following them, were excluded from analysis. Lastly, data for cortex were parcellated into 360 regions according to the Glasser parcellation<sup>19</sup>. To ensure that scrubbed frames did not impact TA- $\Delta_1$  calculations, TA- $\Delta_1$  was computed only on consecutive segments involving no dropped frames, such that TA- $\Delta_1$  was a weighted average of the TA- $\Delta_1$  of each consecutive segment, weighted by the degrees of freedom of that segment. The long-distance correlation parameter  $d$  was not computed on scrubbed data.

**Psilocybin.** The study used a double-blind randomized design<sup>43</sup>, where 23 participants aged 20–40 years (11 females) were scanned two separate days. Subjects provided informed consent and were compensated for participation. On each day, participants received either placebo (179 mg of mannitol and 1 mg of aerosil) or psilocybin (0.2 mg kg<sup>-1</sup>), administered orally. rs-fMRI was performed at three time points after psilocybin administration: ‘immediate’, performed 20 minutes after administration; ‘early’, performed 40 minutes after administration; and ‘late’, performed 70 minutes after administration. Data were acquired with TR = 2,430 ms. Twenty-four participants were enrolled in the study. One participant was excluded due to a missing scan, for a total of 23 participants. This sample size was determined based on the prior LSD study<sup>7</sup>.

Data pre-processing was identical to that of the LSD dataset described above.

### Estimating SA

We quantified SA by decomposing it into two components: the rate at which correlations decrease exponentially with distance ( $SA\cdot\lambda$ ) and the spatially invariant level of correlation to which it decays ( $SA\cdot\infty$ ). This can be described quantitatively as

$$\text{corr}(x_i, x_j) = SA\cdot\infty + (1 - SA\cdot\infty) e^{-D_{ij}/SA\cdot\lambda} \quad (1)$$

where  $x_i$  is the time series for region  $i$ ;  $D_{ij}$  is the Euclidean distance between regions  $i$  and  $j$ ; and  $\text{corr}$  is the Pearson correlation. The quantities  $SA\cdot\lambda$  and  $SA\cdot\infty$  are constants and do not depend on  $i$  or  $j$ .

We estimated the parameters  $SA\cdot\lambda$  and  $SA\cdot\infty$  for each subject as follows. For each pair of brain regions, we computed both their physical Euclidean distance (from the region's centroid) as well as their Pearson correlation coefficient. In this study, we used Euclidean distance rather than geodesic distance as our primary distance measure because Euclidean distance is defined across cortical hemispheres and within subcortex. However, geodesic and other distance measurements can be used instead and produce similar results in practice (Extended Data Fig. 7). We binned each pair of brain regions according to their Euclidean distance, using 1-mm bins for the HCP, HCP-GSR, LSD and psilocybin datasets, and 5-mm bins for the TRT and Cam-CAN datasets due to the fewer number of parcels. We then computed the mean Pearson correlation of pairs in each bin to generate correlation versus distance curve (Fig. 1a). We then found the least square fit of Equation (1) to this curve, optimizing with gradient descent, bounding  $SA\cdot\lambda$  between 0 and 100 and  $SA\cdot\infty$  between -1 and 1. Although it is possible to fit Equation (1) directly to the distance and correlation of each pair without binning, our approach puts more weight on nearby and distant correlations, which are less represented in the data but most critical for determining  $SA\cdot\lambda$  and  $SA\cdot\infty$ .

To account for heteroskedasticity in  $SA\cdot\lambda$  for the Cam-CAN dataset, we analyzed the logarithm of this parameter for this dataset.

### Estimating TA

Our primary measure of TA,  $TA\cdot\Delta_1$ , is a non-parametric measurement computed by taking the Pearson correlation of neighboring time points in the parcellated time series—that is, for a time series  $x[t]$ , we have  $\text{corr}(x[t], x[t+1])$ . This measure is computationally efficient and can be implemented with only a few lines of code. However,  $TA\cdot\Delta_1$  is not comparable across datasets due to differing TR.

For datasets that included motion scrubbing, we computed  $TA\cdot\Delta_1$  only on consecutive regions without scrubbed frames. More specifically, we split each time series into fragments at the location of the scrubbed frames. Then, we computed  $TA\cdot\Delta_1$  on each fragment and took the weighted average. More formally, for each time series, fragment  $x_i$  of length  $\ell_i$ ,  $TA\cdot\Delta_1$  was computed as

$$TA\cdot\Delta_1 = \sum_i \text{corr}(x_i[t], x_i[t+1]) \frac{\ell_i - 1}{\sum_j (\ell_j - 1)}.$$

### Estimating long-memory dynamics

fMRI time series are thought to show characteristics of a long-memory process. In other words, time points that are separated by a large lag show a high correlation<sup>72–74</sup>. These processes are often characterized by their autocorrelation function (ACF). For a time series  $x$ , the ACF is defined as

$$ACF_x(k) = \text{corr}(x[t], x[t+k])$$

where  $\text{corr}$  is the Pearson correlation, and  $k$  is the lag at which to evaluate the ACF. Long-memory processes have an ACF that decays slowly

across lags, meaning that relatively high correlations exist even at large lags. In this work, we use  $TA\cdot\Delta_1$  as our primary measure of TA, which is identical to the lag-1 term of the ACF,  $ACF_x(1)$ .

To evaluate the ability of the first term of the ACF to predict higher terms of the ACF, we divided the data into ten randomly selected subsets for the HCP and Cam-CAN datasets and 12 subsets for the TRT data. We fit the regression model

$$ACF_x(k) = \beta_0 + \beta_1 TA\cdot\Delta_1$$

to one of these subsets. Then, we evaluated the  $R^2$  of the model on each of the other subsets. Note that this is a more ‘difficult’ prediction than traditional  $k$ -fold cross validation, as it uses less training data to make the predictions. We evaluated this fit on the remaining subsets.

As a more direct measure of long-memory dynamics, we analyzed the fractional integration constant  $d$  from an ARFIMA model, frequently used to capture long-memory dynamics. One attractive property of  $d$  is its close relationship to other measures of long-memory dynamics, such as the Hurst exponent  $H$ —under specific conditions, there is a one-to-one mapping<sup>75</sup> given by  $H = d + 0.5$ . Following previous work, we estimated  $d$  through a univariate wavelet-based Whittle estimator<sup>76,77</sup>. We restricted our analyses of  $d$  to the HCP, TRT and Cam-CAN datasets, because they did not use motion scrubbing in their pre-processing pipeline. Under specific assumptions, the relationship between  $d$  and  $TA\cdot\Delta_1$  can be described analytically<sup>78,79</sup>. Consistent with prior work,  $TA\cdot\Delta_1$  is representative of the first and higher lag terms of the ACF<sup>74,78–80</sup>.

### FC and graph construction

We defined the FC matrix  $\rho_{ij}$  as the matrix of Pearson correlation coefficients  $\rho_{ij} = \text{corr}(x_i, x_j)$  between each pairwise combination of regional time series  $x_i$  and  $x_j$ .

We constructed unweighted, undirected graphs from model or data FC matrices using standard techniques<sup>3,81</sup>. We first constructed a spanning tree backbone to ensure connectedness by transforming each element of the FC matrix  $\rho_{ij}$  by  $\sqrt{2(1 - \rho_{ij})}$ —an operation that turns the correlation similarity measure into a distance<sup>82</sup>—and then applying Kruskal’s algorithm to find the minimum spanning tree<sup>83</sup>. We iteratively added the strongest edges in the FC matrix to the spanning tree until the graph contained 10% of all possible edges (proportional thresholding). This produced an unweighted, undirected graph with a fixed number of edges.

### Quantifying reliability

We quantify univariate reliability using the ICC. ICC measures the reliability of a particular scalar measure across subjects<sup>37,84</sup>. Let  $N$  be the total number of subjects,  $M$  be the number of sessions per subject and  $y_n^m$  be some scalar measure of interest for session  $m$  of subject  $n$ . Similarly to an ANOVA, ICC decomposes the variance across subjects into variance from a common source,  $\sigma_y^2$ , and noise,  $\sigma_e$ , by assuming that  $y$  can be decomposed into a subject-specific term  $\gamma_n$  and a noise term  $\epsilon_{n,m}$ —that is,  $y_n^m = \gamma_n + \epsilon_{n,m}$ . Then, the ICC is defined as the fraction of variance explained by the subject-specific term—that is,

$$\text{ICC} = \frac{\sigma_y}{\sigma_y + \sigma_e}.$$

In theory, there are generalizations of ICC that can accommodate additional sources of variance. Here, we use the simplest form,  $\text{ICC}(1,1)$ <sup>37</sup>. Significance compared to 0 and confidence intervals (CIs) were computed using the formulas from ref. 85. Significant differences between ICC values were computed using bootstrapping. Data distribution was assumed to be normal, but this was not formally tested.

We quantify multivariate reliability using fingerprinting, similar to that performed in refs. 20,34,86. Let  $y_n^m$  be a vector describing some measure (for example, regional  $TA\cdot\Delta_1$ ) from a single session  $m \in \{1, \dots, M\}$

for subject  $n \in \{1, \dots, N\}$ . For each session  $m$  of each subject  $n$ , we compute

$$\Gamma(n, m) = \arg \max_{n'} \left( \max_{m', (n, m) \neq (n', m')} \text{corr}(\gamma_n^m, \gamma_{n'}^{m'}) \right),$$

where  $\text{corr}$  is Pearson correlation. The fingerprinting performance is given by

$$\frac{1}{NM} \sum_{m,n} I_{\Gamma(n,m)=n},$$

where  $I$  is the indicator function. Under this measure, chance performance is  $(M - 1) / (NM - 1)$ . For regional measures, the length of vector  $\gamma_n^m$  was equal to the number of regions in the parcellation  $R$ . For FC, the length of  $\gamma_n^m$  was  $R(R - 1) / 2$ —that is, the number of distinct elements in the FC matrix.

In the HCP and HCP-GSR datasets, subjects were scanned four times over 2 days. For computing ICC, we used  $M = 4$ . For fingerprinting, to obtain error estimates and increase the difficulty of fingerprinting, we performed fingerprinting on each possible pair of sessions 1–4 ( $M = 2$  performed six times instead of  $M = 4$  performed once). This means that each session had only one correct match of the 1,765 sessions in the pool, compared to one correct match for 126 in the pool in ref. 34. In the TRT dataset, subjects were scanned in six sessions across four different days for a total of 24 sessions per subject. Likewise, we used  $M = 24$  to compute ICC. By contrast, we performed fingerprinting for each session independently (that is,  $M = 4$  performed six times instead of  $M = 24$  performed once), so subjects had three correct matching sessions out of 47 in the pool. Without such increases in difficulty on these datasets, fingerprinting performance was near 100%. We could not compute ICC or perform fingerprinting on the Cam-CAN dataset, because only one rs-fMRI session was acquired per subject.

### Regional homogeneity

Regional homogeneity estimates how similar a voxel's time series is to its neighbors<sup>87</sup>. Here, because we used a parcellated analysis, we tested the homogeneity of the voxels within a parcel. Following standard practice, we defined regional homogeneity as the Kendall's  $W$  of time series within a parcel. More concretely,

$$\text{ReHo} = \frac{12}{k^2(n^3 - n)} \sum_{i=1}^n \left( \sum_{j=1}^k r_{i,j} - kT_i/n \right)^2$$

where  $r_{i,j}$  is the rank of the  $i$ -th time point in the  $j$ -th voxel;  $k$  is the number of voxels in the parcel;  $n$  is the number of time points; and  $T_i$  is the  $i$ -th triangular number.

### Lin's concordance

Lin's concordance<sup>88</sup> measures how well one variable reproduces another variable. In general, there are multiple ways for variables to be related to each other. On the one hand, they can be correlated with each other, meaning that variation in one corresponds to variation in the other. This is generally quantified using the Pearson correlation coefficient. On the other hand, it is possible for two variables to have very different values while still being highly correlated. In this case, there would be a high Pearson correlation coefficient while also having a high mean squared error. Lin's concordance addresses this issue by modifying the definition of Pearson's correlation to account for differences in the values of the variables, returning its maximum only if the data have both a high Pearson correlation and a low mean squared error. In other words, it measures the scaled distance of the sample from the unity line in the scatter plot. Lin's concordance is defined for two samples  $x_1$  and  $x_2$  as

$$\frac{2\text{cov}(x_1, x_2)}{\text{var}(x_1) + \text{var}(x_2) + (\bar{x}_1 - \bar{x}_2)^2}$$

where  $\bar{x}$  is the mean of  $x$ . Alternatively, it can be formulated as a reweighted version of the Pearson correlation, using the equivalent definition

$$\text{corr}(x_1, x_2) \frac{2\sqrt{\text{var}(x_1)}\sqrt{\text{var}(x_2)}}{\text{var}(x_1) + \text{var}(x_2) + (\bar{x}_1 - \bar{x}_2)^2}.$$

Like Pearson correlation, Lin's concordance can take values ranging from -1 to 1. It is 1 if and only if the variables are identical. It is close to 0 if there is no correlation or if the means or variances of the two populations differ. This can be seen in the second definition based on Pearson correlation, where the weight on the Pearson correlation becomes 1 as the means and variances of  $x_1$  and  $x_2$  approach each other. Likewise, Lin's concordance is negative if the means and variances are similar but the Pearson correlation is negative.

### Weighted graph metrics

We consider three weighted graph metrics: the mean (mean-FC), variance (var-FC) and kurtosis (kurt-FC) of the FC matrix. Each is calculated by finding the corresponding statistic (mean, variance or kurtosis) of the upper triangular portion of the FC matrix, excluding the diagonal. In other words, for the FC matrix  $M$ , this is the mean, variance and kurtosis of the set  $\mathcal{M} = \{m_{ij} : i < j\}$ .

Note that the thresholding procedure destroys all explicit information about mean-FC, var-FC and kurt-FC. To understand why, recall that the thresholding procedure keeps a fixed fraction of edges. Thus, if we threshold an FC matrix, the number of edges depends only on the fixed fraction of edges we keep. For an FC matrix of size  $N$ , if we keep  $q$  fraction of edges, then the moments are identical to those of the binomial distribution: mean-FC of the thresholded matrix is  $q$ ; the variance is  $q(1 - q)$ ; and the kurtosis is  $1/pq - 6$ . Therefore, no explicit information remains about mean-FC, var-FC or kurt-FC after thresholding.

Because TA can be measured for individual nodes, we also consider how it impacts local topology. To do this, we also define the corresponding nodal graph metrics for each that operates on rows of the FC matrix instead of the upper triangle: nodal mean-FC, the mean of the row; nodal var-FC, the variance of the row; and nodal kurt-FC, the kurtosis of the row. We exclude self-connectivity (that is, the diagonal of ones in the FC matrix) from these calculations. Thresholding does not destroy nodal mean-FC, nodal var-FC or nodal kurt-FC.

### Unweighted graph metrics

We consider six popular graph metrics to quantify the topology of the connectome<sup>4</sup>:

**Assortativity.** Assortativity is a preference of high-degree nodes, or hubs, to connect to each other. Mathematically, this is the Pearson correlation between the degree of nodes connected by edges<sup>89</sup>.

**Clustering coefficient.** The clustering coefficient for a single node is the average relative number of triangles around a node. For adjacency matrix  $A$ , the nodal clustering coefficient for node  $i$  is

$$\frac{1}{(\sum_j A_{i,j})((\sum_j A_{i,j}) - 1)} \sum_{j,k, j \neq k} A_{i,j} A_{j,k} A_{i,k}.$$

The clustering coefficient for the network is the average of the nodal clustering coefficients.

**Global efficiency.** The global efficiency is related to the average topological distance between nodes. Mathematically, it is the mean of the inverse of shortest path lengths between each pair of nodes—that is,

$$\frac{1}{N(N-1)} \sum_{i,j,i \neq j} \frac{1}{s_{i,j}},$$

where  $s_{i,j}$  indicates the shortest path between nodes  $i$  and  $j$ , and  $N$  is the number of nodes<sup>90</sup>.

**Local efficiency.** Local efficiency, similar to clustering, is the mean global efficiency on the subgraph of each node's nearest neighbors<sup>90</sup>.

**Modularity.** Modularity quantifies the ability to break a network into 'communities' such that the number of edges within the community is maximized and outside the community is minimized. Let  $C_i$  be some community assignment of node  $i$ , and let  $\delta$  be the Kronecker delta function. We can compute the quality of the community assignment  $C_i$  with the equation

$$\sum_{i,j,i \neq j} (2A_{i,j} - 1)(2\delta(C_i, C_j) - 1).$$

The first term in the sum is 1 if the neurons are connected and -1 if they are not, and the second term is 1 if they are in the same community and -1 if they are not. Thus, this can be maximized if connected nodes fall into the same community and unconnected nodes do not. The modularity is defined as the maximum value of this function across all potential community assignments  $\{C_i\}_i$ , rescaled to fall between -0.5 and 1 (ref. 91).

**Transitivity.** The transitivity is the total number of three-way reciprocal connections compared to the total possible number of such connections—that is,

$$\frac{1}{\sum_i (\sum_j A_{i,j}) ((\sum_j A_{i,j}) - 1)} \sum_{i,j,k,i \neq j \neq k} A_{i,j} A_{j,k} A_{i,k}$$

Note that this is distinct from local efficiency because it considers the network as a whole rather than considering each node  $i$  individually and then averaging.

In addition to considering these graph metrics, we also consider two nodal graph metrics:

**Degree.** The degree is a nodal metric that measures the total number of edges connected to a node—that is,

$$\text{degree}(i) = \sum_j A_{i,j}$$

**Centrality.** Betweenness centrality is a nodal metric that measures the fraction of shortest paths between all pairs of nodes in a network that pass through the given node. It does not have a closed-form equation and must be computed using an algorithm<sup>92</sup>.

### Graph metrics linear model

To quantify the impact of SA and TA jointly, we used a linear model. We fit the 'SA-λ + SA-∞' model

$$\text{metric} = \beta_1 \text{SA-}\lambda + \beta_2 \text{SA-}\infty$$

or the 'All' model

$$\text{metric} = \beta_1 \text{SA-}\lambda + \beta_2 \text{SA-}\infty + \beta_3 \text{TA-}\Delta_1$$

on 50% of the data, randomly chosen from all sessions. Shown in Fig. 1e and Extended Data Fig. 1 is the Spearman correlation of the 50% held-out data with the predictions of this linear model. We used Spearman correlation instead of  $R^2$  for fair comparison of SA-λ, SA-∞ and

global TA-Δ<sub>1</sub> individually. Because the fit uses held-out data, the correlation with the linear model does not necessarily need to be higher than the correlation with any of the individual factors in the model. Because many HCP subjects are related, we used only 334 unrelated subjects to avoid the relatedness confound in our analysis, meaning that there were 167 subjects in the training and test sets.

### Correlated spectral sampling

The spatiotemporal model relies on generating spatially autocorrelated time series, each with a given power spectrum. We introduce the correlated spectral sampling algorithm to produce a set of time series related by a given covariance matrix, where each time series has a given power spectrum.

Diverse methods of generating time series that exhibit both TA and SA already exist, such as vector autoregressive (VAR) and Ornstein–Uhlenbeck processes. These models directly simulate time series in the time domain, introducing TA and SA locally. However, prior research demonstrating power law dynamics in fMRI data<sup>72–74</sup> suggests that VAR and similar models are inadequate for our purposes, as they are unable to model complex power spectral properties. More complex variants of these models, such as VARFI and FIVAR, are able to capture long-memory dynamics, but they are unable to simulate from arbitrary power spectra<sup>76</sup>.

Instead, we seek to generate power spectra that represent the tradeoff between long-memory dynamics—that is, filtered 1/f<sup>a</sup> noise for frequencies above 0.01 Hz—and optional white noise, with a flat power spectrum. To introduce SA in our method, we spatially embed these time series with a given covariance matrix  $C$ . In our case,  $C$  is given in Equation (2). The usual method of spatially embedding time series with a given covariance matrix  $C$  is to multiply them by the matrix square root of  $C$ , equivalent to sampling from a multivariate normal distribution. However, this process involves the linear combination of distinct time series, which changes their power spectra, destroying the desired power spectrum. To avoid this confound, we introduce the correlated spectral sampling algorithm, a generalization of ref. 93, which is able to produce correlated time series with arbitrary power spectra. Correlated spectral sampling can alternatively be seen as a generalization of phase randomization to obtain specified correlational structure.

### Algorithm: Correlated spectral sampling.

Let  $N$  be the number of desired time series, and let  $T$  be the even length of the desired time series. Given an  $N \times N$  correlation matrix  $C_{ij}$  and a set of power spectra  $|X_i[k]|^2$ , we seek to generate a set of time series  $\{x_1[t], \dots, x_N[t]\}$  such that each time series  $x_i[t]$  has a specified estimate of the power spectrum  $|X_i[k]|^2$ , and the correlation matrix of the resulting time series has expected value  $C_{ij}$ . Correlated spectral sampling operates by generating a set of complex frequency-domain coefficients in the domain  $[-T/2 + 1, T/2]$ , introducing SA and TA independently, and then performing an inverse discrete Fourier transform to create the desired set of time series.

- Pick the desired power spectra estimates  $\{|X_1[k]|^2, \dots, |X_N[k]|^2\}$ , representing the desired temporal structure, and  $N \times N$  correlation matrix  $C$ , representing the desired spatial structure. These choices must satisfy the requirement that the matrix

$$\Sigma_{i,j} = \frac{C_{ij}}{\text{sim}(|X_i[k]|, |X_j[k]|)}$$

is positive semi-definite, where  $\text{sim}$  denotes cosine similarity—that is,

$$\text{sim}(|X_i[k]|, |X_j[k]|) = \frac{\sum_k |X_i[k]| |X_j[k]|}{\sqrt{\sum_k |X_i[k]|^2} \sqrt{\sum_k |X_j[k]|^2}}.$$

- For each region  $n \in \{1, \dots, N\}$ , and for each frequency  $k$  in the positive sub-Nyquist Fourier domain,  $1 \leq k \leq T/2$ , sample two vectors from a multivariate normal distribution  $a_n^R[k], a_n^I[k] \sim N(0, \Sigma)$ .

3. Set  $a'_n[T/2] = 0$  for all  $n$ . Because  $T$  is even, by way of symmetry the frequency-domain coefficients of the Nyquist frequency must be real. Set  $a'_n[0] = 0$  so that the resulting time series will be real.
4. Form the complex frequency-domain coefficients by letting  $c_n[k] = |X_i[k]|(a_n^R[k] + ia_n^I[k])$  for  $0 \leq k \leq T/2$ , and  $c_n[-k] = \overline{c_n[k]}$  for  $0 < k < T/2$ .
5. Perform an inverse discrete Fourier transform for each  $c_n[k]$  to obtain the time series  $x_n[t]$ .

Each of the resulting time series  $x_i[t]$  will have a frequency spectrum approximately equal to  $|X_i[k](a_n^R[k] + ia_n^I[k])|^2 = |X_i|^2$ , the desired temporal structure. The proof of this algorithm is given in Supplementary Information 1.1.

### Relationship between TA- $\Delta_1$ and power spectra

In addition to correlated spectral sampling, defining the spatiotemporal model requires us to determine the precise relationship between TA- $\Delta_1$  and the power spectrum, with a specific focus on the case when the original time series is mixed with white noise.

Suppose we have a finite time series  $x[t]$  with discrete Fourier transform  $X[k]$ . Then, its TA- $\Delta_1$  is

$$\text{corr}(x[t], x[t+1]) = \frac{1}{\sum_{k=1}^{N-1} |X[k]|^2} \sum_{k=1}^{N-1} |X[k]|^2 \cos(2\pi k/N).$$

A proof is provided in Supplementary Information 1.2. Notably, for a fixed-length time series with a  $1/f^\alpha$  power spectrum, this formula implies a bijective relationship between  $\alpha$  and TA- $\Delta_1$  in the domain  $\alpha \in [0, 2]$ .

Additionally, adding white noise will cause a change in the power spectrum and, hence, will change TA- $\Delta_1$ . For a finite time series  $y[t] = x[t] + w[t]$  where  $w[t] \sim N(0, \sigma)$  of length  $N$ ,

$$\mathbb{E}[\text{corr}(y[t], y[t+1])] = \frac{1}{N\sigma^2 + \sum_{k=1}^{N-1} |X[k]|^2} \sum_{k=1}^{N-1} |X[k]|^2 \cos(2\pi k/N).$$

A proof is provided in Supplementary Information 1.3. Note that this formula implies that, as the amount (or variance) of white noise is increased, the TA- $\Delta_1$  approaches 0.

### Spatiotemporal model

The spatiotemporal model generates surrogate time series that can be analyzed like rs-fMRI time series. It takes two parameters—the noiseless SA- $\lambda$  ( $\text{SA-}\lambda^{\text{gen}}$ ) and the noiseless SA- $\infty$  ( $\text{SA-}\infty^{\text{gen}}$ )—and also uses two pieces of information from the data—the TA- $\Delta_1(\phi_i)$  from each region  $i$  and the Euclidean distance  $D_{ij}$  between the centroids of each pair of regions  $i$  and  $j$ .

The model operates in two basic steps. First, we generate random time series that are both spatially and temporally autocorrelated. For this first step, all time series have uniformly high TA, and SA is determined by the parameters SA- $\lambda$  and SA- $\infty$ . Using correlated spectral sampling, we generate  $N$  high-pass-filtered (cutoff frequency of 0.01 Hz, 4th order Butterworth filter) Brownian noise time series (frequency spectrum  $1/f^\alpha$ ) of length  $T$  such that, for time series  $x_i[t]$  and  $x_j[t]$ , we have

$$\mathbb{E}(\text{corr}(x_i[t], x_j[t])) = \text{SA-}\infty^{\text{gen}} + (1 - \text{SA-}\infty^{\text{gen}}) e^{-D_{ij}/\text{SA-}\lambda^{\text{gen}}} \quad (2)$$

where corr is the Pearson correlation, similar to Equation (1). The matrix consisting of all such expected correlations between regions  $i$  and  $j$  forms the matrix  $C_{ij}$  required by correlated spectral sampling, such that

$$C_{i,j} = \mathbb{E}(\text{corr}(x_i[t], x_j[t])).$$

For non-negative  $\text{SA-}\infty^{\text{gen}}$ ,  $C_{ij}$  will be positive semi-definite.

Second, we reduce the TA of each time series to match that of the original data. Brownian motion time series have high TA, and adding white noise to a time series reduces TA, so we add white noise to each time series  $x_i[t]$  until the TA- $\Delta_1$  of the time series is equal to the empirical TA- $\Delta_1(\phi_i)$ . Thus, we have to choose a distinct amount of noise to add to each time series such that  $\text{corr}(x_i[t], x_i[t+1]) = \phi_i$ . We update  $x_i[t] \leftarrow x_i[t] + N(0, \sigma^2(\phi_i))$  for some function  $\sigma^2(\phi_i)$  given by

$$\sigma^2(\phi) = \frac{1}{N^2 \phi} \sum_{k=1}^{N-1} |X[k]|^2 (\cos(2\pi k/N) - \phi). \quad (3)$$

where  $N$  is the length of the time series. The derivation of this formula is provided in Supplementary Information 1.3. The function  $\sigma^2(\phi_i)$  is defined for  $\phi_i > 0$ , and the variance of a random variable must be non-negative, so  $\phi_i$  is truncated such that  $\phi_i \geq 0.0001$  and  $\sigma^2(\phi_i) \geq 0$ . Note that  $\text{SA-}\lambda^{\text{gen}}$  and  $\text{SA-}\infty^{\text{gen}}$  are parameters of the underlying process and differ from the  $\text{SA-}\lambda$  and  $\text{SA-}\infty$  of the generated time series due to the addition of noise. The resulting time series exhibit a spatial embedding given by the parcel centroid distances  $D_{ij}$  and parameters  $\text{SA-}\lambda^{\text{gen}}$  and  $\text{SA-}\infty^{\text{gen}}$  and have regional TA- $\Delta_1$  equal to the original time series.

### Model variants

We test several variants of the model to determine the importance of three key components of the model: added uncorrelated noise, TA and SA.

**TA only.** This model modifies the spatiotemporal model to remove the spatial embedding. Specifically, it fixes the parameters  $\text{SA-}\lambda^{\text{gen}}$  and  $\text{SA-}\infty^{\text{gen}}$  to 0, such that  $\text{corr}(x_i[t], x_j[t]) = I_{i,j}$  where  $I$  is the indicator function. It takes no parameters.

**SA only.** We generate random multivariate Gaussian noise with mean zero and covariance matrix

$$C_{i,j} = \text{SA-}\infty^{\text{gen}} + (1 - \text{SA-}\infty^{\text{gen}}) \exp(-D_{i,j}/\text{SA-}\lambda^{\text{gen}})$$

similar to Equation (1). It takes two parameters,  $\text{SA-}\lambda^{\text{gen}}$  and  $\text{SA-}\infty^{\text{gen}}$ . Unlike the spatiotemporal model,  $\mathbb{E}(\text{SA-}\lambda) = \text{SA-}\lambda^{\text{gen}}$  and  $\mathbb{E}(\text{SA-}\infty) = \text{SA-}\infty^{\text{gen}}$ .

**Intrinsic time scale with SA.** In our spatiotemporal model, we make sure time series have the desired TA- $\Delta_1$  by generating time series with uniformly high TA- $\Delta_1$  and then adding different magnitudes of white noise to each time series to match TA- $\Delta_1$  to the original time series. In this model, we do not add white noise to the time series. To match TA- $\Delta_1$  to the original time series, we generate time series directly with matched TA- $\Delta_1$ . This is possible because correlated spectral sampling does not require the time series' power spectra to be identical. Thus, we achieve this diversity in TA- $\Delta_1$  by assuming that each time series has a filtered pink noise ( $1/f^\alpha$ ) temporal dynamics. Then, for each time series  $i$ , we find an exponent  $\alpha_i$  such that the high-pass-filtered  $1/f^{\alpha_i}$  spectrum has expected TA- $\Delta_1$  equal to  $\phi_i$ .

The success of this procedure requires us to choose  $\alpha_i$  such that the high-pass-filtered  $1/f^{\alpha_i}$  spectrum has TA- $\Delta_1$  equal to  $\phi_i$ . The mapping  $\phi_i \rightarrow \alpha_i$  can be determined by numerically inverting the  $\alpha_i \rightarrow \phi_i$  mapping implied by Supplementary Information 1.2. High-pass filtering is performed at the level of the power spectrum by multiplying the square of the amplitude response of the filter by the power spectrum. Additionally, in this model, it is possible to determine the  $\text{SA-}\lambda^{\text{gen}}$  and  $\text{SA-}\infty^{\text{gen}}$  parameters directly from the data, without the need to fit parameters, using correlated spectral sampling. As in the 'SA only' model,  $\mathbb{E}(\text{SA-}\lambda) = \text{SA-}\lambda^{\text{gen}}$  and  $\mathbb{E}(\text{SA-}\infty) = \text{SA-}\infty^{\text{gen}}$ .

Thus, in summary, this model makes the following modifications to the spatiotemporal model: (1) rather than simulating random walks ( $1/f^\alpha$  where  $\alpha = 2$ ), the spectral exponents  $\alpha_i$  for each region  $i$  are chosen

such that the resulting  $\text{TA}-\Delta_1$  is equal to each region's  $\text{TA}-\Delta_1$  value; and (2) no noise is added to the power law time series.

**Homogeneous  $\text{TA}-\Delta_1$ .** To allow SA and TA to be independently manipulated, we developed a homogeneous variant of the spatiotemporal model that treats TA as a parameter. Rather than use  $\text{TA}-\Delta_1$  values computed from the original time series, this model uses a single fixed value of  $\text{TA}-\Delta_1$ ,  $\text{TA}-\Delta_1^{\text{gen}}$ , for all regions. For simplicity, we fixed  $\text{SA}-\infty^{\text{gen}}$  to be the mean  $\text{SA}-\infty$  across all networks. Thus, the model takes two parameters:  $\text{TA}-\Delta_1^{\text{gen}}$  and  $\text{SA}-\lambda^{\text{gen}}$ .

### Eigensurrogate model

Due to our use of eigenvalues for fitting, we developed the eigensurrogate model to test whether eigenvalues alone are capable of reproducing a phenomenon (Extended Data Fig. 6). This tests the null hypothesis that an effect can be explained by its linear dimensionality. Unlike most of the other models that we considered, this produces surrogate FC matrices instead of surrogate time series.

We first performed an eigendecomposition of the correlation matrix (FC matrix) and then applied the procedure of ref. 94. In brief, we sampled a random set of eigenvectors and then applied a series of rotations to set ones on the diagonal. We used the method as implemented by Scipy in the numerical Python stack.

Because this model creates surrogate FC matrices, we produced time series by sampling the maximum entropy time series that would produce such a correlation matrix. To do this, we numerically computed the matrix square root and multiplied it by an  $N \times T$  matrix of standard normal iid random variables. This is equivalent to sampling each time point independently from a multivariate Gaussian distribution. In principle, it is possible to create temporally autocorrelated time series from this model (multiplying by temporally autocorrelated time series instead of iid standard normal random variables), but, because it operates at the level of the FC matrix, all methods of generating time series from the eigensurrogate method will produce an identical FC matrix.

### Null models

We also test several popular null models.

**Phase randomization.** The power spectrum amplitude of individual time series is preserved, but the phases are randomized (Extended Data Fig. 6). This procedure is described in detail in ref. 95. A Fourier transform was performed on each region's time series. Each element of the complex-valued Fourier transform was randomly rotated on the unit circle, and the inverse transform was performed on the phase-randomized spectra.

Note that we used distinct random phases for each time series, contrary to many neuroimaging studies that use the same phase for a given frequency across all time series. This is because using the same phase for each time series preserves all cross-correlations between time series. In practice, this means that the surrogate FC matrix is identical to the original FC matrix, and, hence, it produces an identical graph.

**Zalesky matching.** This model matches the mean and the variance of the correlation matrix (Extended Data Fig. 6) and is described in full in ref. 30. In brief, it matches the first two moments by iteratively computing correlation matrices from time series of different durations with different ground truth correlations. The process continues until the time series duration is found that maximally reproduces the mean and variance of the correlation matrix.

**Edge reshuffle.** We preserve the degree of each node while scrambling the edges (Extended Data Fig. 6). This is accomplished by an iterative algorithm described in detail in ref. 96. Two edges from a graph were selected at random, and the connections were swapped. This swap was

iterated  $k$  times, where  $k$  was chosen here to be five times the number of edges in the network. The result is that each node has the same number of connections, but those connections are randomized.

### Models not considered

Despite our parameterization based on the  $\text{TA}-\Delta_1$  we do not report on time series generated using autoregressive (AR) or vector autoregressive (VAR) models. These models have two limitations within this context. First, they do not reproduce the observed long-memory processes observed within rs-fMRI time series. Second, when fitting data using these models, the  $\text{TA}-\Delta_1$  parameter fits to values very close to 1, resulting in parameter degeneracy.

Likewise, we did not directly fit the economical clustering model to data. Due to the variability in individual instantiations of this model and the lack of smoothness of the parameter space, this model is incompatible with our numerical fitting algorithm, thus preventing the use of similar methodology to perform the fitting. These issues, combined with the long execution time of the model, made such individual-level fitting infeasible.

We did not consider models in which only one of  $\text{SA}-\lambda$  and  $\text{SA}-\infty$  is fit and the other is fixed. These two parameters were fit by optimizing a function to the subject's measured SA. However, we were unable to get consistent and interpretable fits to this function when only one of these two parameters was fit. In other words, fixing one of these parameters precludes reliable estimation of the other.

Lastly, we only considered approaches that operate at the level of the parcellated time series. This excluded approaches such as scrambling using a three-dimensional Fourier transform.

### Model fitting procedure

Due to the fact that uncorrelated random noise was added after the spatial embedding in our model, the  $\text{SA}-\lambda^{\text{gen}}$  and  $\text{SA}-\infty^{\text{gen}}$  parameters in the spatiotemporal model were not identical to the observed  $\text{SA}-\lambda$  and  $\text{SA}-\infty$ . We derived a mathematical method for directly matching the spatiotemporal model's  $\text{SA}-\lambda^{\text{gen}}$  and  $\text{SA}-\infty^{\text{gen}}$  parameters to the data without the use of fitting, but technical constraints prevented an implementation of the procedure (Supplementary Information 1.4). Therefore, we could not directly estimate the model based on observed  $\text{SA}-\lambda$  and  $\text{SA}-\infty$ .

Instead, the spatiotemporal model's  $\text{SA}-\lambda^{\text{gen}}$  and  $\text{SA}-\infty^{\text{gen}}$  parameters were fit to the eigenvalue distribution of each individual subject's FC matrix. A predicted FC matrix was generated by the model, and this procedure was iterated until a match was obtained in the eigenvalue distribution. The model was fit using differential evolution<sup>97</sup>, a gradient-free global heuristic search method, to the mean squared error between the sorted eigenvalue distributions of the subject and model FC matrices. Eigenvalues were non-negative due to the positive semi-definiteness of the correlation matrix. The model was implemented in such a way that they preserved smoothness with respect to the parameters, meaning that, for a given random seed, small perturbations of the parameters caused only small changes in the time series and, hence, in the structure of the graph. Optimization was performed on the mean objective function from two random seeds. All reported statistics and metrics about the models come from a single instantiation of the model using a different random seed than either seed used during fitting. Parameters for the 'SA only' model were also fit using the same procedure.

Most alternative models had no parameters ('TA only', 'Phase randomization', 'Zalesky matching' and 'Edge reshuffle'). For the 'Noiseless' model, parameters could be fit directly using correlated spectral sampling. This formalism is not guaranteed to converge for all subjects. When correlated spectral sampling was unable to produce valid time series, we excluded these subjects from the analysis. Results were qualitatively similar when parameters were fit to the eigenvalue distribution as described above that forced parameters into valid regimes.

## Economical clustering model

As described previously, the spatiotemporal model produces time series for each brain region–graphs can then be constructed by processing these time series the same way as subjects’ rs-fMRI time series. Although the spatiotemporal model can be called a ‘generative model’ by most definitions, within the graph theory literature, the term ‘generative model’ often refers to models that construct graphs directly through the iterative addition of nodes or edges<sup>23,40,98</sup>. The economical clustering (EC) model is one such model that is popular for studying brain networks<sup>23,40</sup>. In this model, connections between nodes are determined by one parameter governing the impact of distance and one for the impact of clustered topology. The probability of an edge forming between two brain regions is proportional to the product of the Euclidean distance between the regions raised to some power (the distance parameter) and the fraction of shared neighbors between them raised to some power (the clustering parameter). Full model details are provided in ref. 23.

To compare the two models, we simulated the EC model across a spectrum of distance and clustering parameters and then fit a spatiotemporal model to the simulated networks. Because our spatiotemporal model takes two SA-related parameters and obtains TA on a regional level directly from the data, we compared the EC model to the homogeneous variant of the spatiotemporal model, which includes one SA parameter and one TA parameter. We used the EC model to simulate ten networks per combination of parameters, fitting the homogeneous spatiotemporal model to each of these ten networks. Because the EC model produced graphs rather than FC matrices, we could not use our previous approach of fitting by eigenvalues, nor could we derive an analytic approach to parameter estimation. Thus, we fit using the objective function from ref. 40. Full details are provided in Supplementary Information 2.

## Spatial correction for brain map similarity

To assess the similarity between brain maps, the presence of SA can induce a high false-positive rate. To correct for this, we perform a permutation test using an SA-preserving surrogate method<sup>11</sup>. To compare a target brain map to a reference map, we generate 10,000 surrogate brain maps that match the SA of the reference map. Then, we compute the Spearman correlation between the target map and the reference map as well as between the target map and the surrogate maps. The *P* value of the two-tailed test is determined as the fraction of correlations with the surrogate maps that are at least as extreme in absolute value as with the reference map. For cases in which maps cannot be designated as a target or reference map, we perform this procedure twice, once with each map taking the role of the reference map and the other as the target map, and compute the two-tailed *P* value as the total number of target-to-surrogate Spearman correlations that are at least as extreme in absolute value as the Spearman correlation between the two maps.

## SVD

We computed a cortical map of serotonergic modulation using SVD, which bears many similarities to principal component analysis (PCA). For a data matrix  $M$ , we can rewrite  $M$  as

$$M = U\Sigma V^T$$

where  $\Sigma$  is a diagonal matrix, and the rows of  $U$  and  $V$  form orthogonal bases. The diagonal elements of  $\Sigma$  are called ‘singular values’; the rows of  $V$  are called ‘singular vectors’ (each element of which is a ‘loading’); and the projection of  $M$  onto  $V$  (or, equivalently,  $U\Sigma$ ) are the ‘scores’. Note also that the product  $M^T M = \Sigma^2 V^T$ , so  $\Sigma^2$  and  $V$  are the eigenvalues and eigenvectors, respectively, of  $M^T M$ . If  $M$  is centered, then  $M^T M$  is the covariance matrix, and SVD is equivalent to PCA, meaning that  $\Sigma^2$  gives the variance explained of each component. But, in our case,

because  $M$  is not centered, variance explained of the first  $k$  components is computed as

$$\text{var} \left( \sum_{i=1}^k \sigma_i u_i v_i^T \right).$$

To compute the variance explained by experimental condition, we find the variance explained by each experiment from each subject individually and then average across the experimental condition.

## Ethics declaration

All participants provided written informed consent statements before participation in the study. The HCP data were acquired using protocols approved by the Washington University institutional review board. The Yale-TRT data were collected with approval by the Yale University institutional review board. The Cam-CAN data were collected with approval by the Cambridgeshire 2 Research Ethics Committee. The LSD and psilocybin data were collected with approval by the Cantonal Ethics Committee of Zurich, and the Swiss Federal Office of Public Health authorized the use of LSD and psilocybin in humans.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

The HCP data are available at <https://www.humanconnectome.org/study/hcp-young-adult>. The Yale-TRT data are available at [http://fcon\\_1000.projects.nitrc.org/indi/retro/yale\\_trt.html](http://fcon_1000.projects.nitrc.org/indi/retro/yale_trt.html). The Cam-CAN data are available at <https://www.cam-can.org/index.php?content=dataset>. The LSD data and psilocybin data are available upon reasonable request.

## Code availability

We prepared a software package that allows the principal analyses in this paper to be performed quickly and easily. The ‘spatiotemporal’ Python package, which can be installed through pip or downloaded at <https://github.com/murraylab/spatiotemporal>, offers a more user-friendly way of applying the analyses described here. The raw source code used to perform the analyses in this paper can be downloaded at [https://github.com/murraylab/spatial\\_and\\_temporal\\_paper](https://github.com/murraylab/spatial_and_temporal_paper). Code was implemented using the standard Python stack<sup>99,100</sup> and other libraries<sup>101,102</sup>. Source code was checked for correctness using software verification techniques<sup>103</sup>.

## References

64. Demirtas, M. et al. Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics. *Neuron* **101**, 1181–1194 (2019).
65. Morgan, S. E., Achard, S., Termenon, M., Bullmore, E. T. & Vértes, P. E. Low-dimensional morphospace of topological motifs in human fMRI brain networks. *Netw. Neurosci.* **2**, 285–302 (2018).
66. Robinson, E. C. et al. MSM: a new flexible framework for multimodal surface matching. *Neuroimage* **100**, 414–426 (2014).
67. Friedman, L., Glover, G. H., Krenz, D. & Magnotta, V. Reducing inter-scanner variability of activation in a multicenter fMRI study: role of smoothness equalization. *Neuroimage* **32**, 1656–1668 (2006).
68. Scheinost, D., Papademetris, X. & Constable, R. T. The impact of image smoothness on intrinsic functional connectivity and head motion confounds. *Neuroimage* **95**, 13–21 (2014).
69. Shen, X., Tokoglu, F., Papademetris, X. & Constable, R. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage* **82**, 403–415 (2013).

70. Taylor, J. R. et al. The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. *Neuroimage* **144**, 262–269 (2017).
71. Tzourio-Mazoyer, N. et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273–289 (2002).
72. Bullmore, E. et al. Colored noise and computational inference in neurophysiological (fMRI) time series analysis: resampling methods in time and wavelet domains. *Hum. Brain Mapp.* **12**, 61–78 (2001).
73. Wagenmakers, E.-J., Farrell, S. & Ratcliff, R. Estimation and interpretation of  $1/f^{\alpha}$  noise in human cognition. *Psychon. Bull. Rev.* **11**, 579–615 (2004).
74. Maxim, V. et al. Fractional Gaussian noise, functional MRI and Alzheimer’s disease. *Neuroimage* **25**, 141–158 (2005).
75. Achard, S., Bassett, D. S., Meyer-Lindenberg, A. & Bullmore, E. Fractal connectivity of long-memory networks. *Phys. Rev. E* **77**, 036104 (2008).
76. Sela, R. J. & Hurvich, C. M. Computationally efficient methods for two multivariate fractionally integrated models. *J. Time Ser. Anal.* **30**, 631–651 (2009).
77. Achard, S. & Gannaz, I. Multivariate wavelet whittle estimation in long-range dependence. *J. Time Ser. Anal.* **37**, 476–512 (2015).
78. Lu, Z. Analysis of stationary and non-stationary long memory processes: estimation, applications and forecast. PhD thesis, École normale supérieure de Cachan - ENS Cachan (2019); <https://theses.hal.science/tel-00422376/document>
79. Hassani, H., Leonenko, N. & Patterson, K. The sample autocorrelation function and the detection of long-memory processes. *Phys. A* **391**, 6367–6379 (2012).
80. Scheffer, M. et al. Early-warning signals for critical transitions. *Nature* **461**, 53–59 (2009).
81. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**, 186–198 (2009).
82. Mantegna, R. Hierarchical structure in financial markets. *Eur. Phys. J. B* **11**, 193–197 (1999).
83. Kruskal, J. B. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc. Am. Math. Soc.* **7**, 48–48 (1956).
84. Termenon, M., Jaillard, A., Delon-Martin, C. & Achard, S. Reliability of graph analysis of resting state fMRI using test-retest dataset from the Human Connectome Project. *Neuroimage* **142**, 172–187 (2016).
85. McGraw, K. O. & Wong, S. P. Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* **1**, 30–46 (1996).
86. Horien, C. et al. Considering factors affecting the connectome-based identification process: comment on Waller et al. *Neuroimage* **169**, 172–175 (2018).
87. Zang, Y., Jiang, T., Lu, Y., He, Y. & Tian, L. Regional homogeneity approach to fMRI data analysis. *Neuroimage* **22**, 394–400 (2004).
88. Lin, L. I.-K. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **45**, 255 (1989).
89. Newman, M. E. J. Assortative mixing in networks. *Phys. Rev. Lett.* **89**, 208701 (2002).
90. Latora, V. & Marchiori, M. Efficient behavior of small-world networks. *Phys. Rev. Lett.* **87**, 198701 (2001).
91. Newman, M. E. J. & Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **9**, 026113 (2004).
92. Brandes, U. A faster algorithm for betweenness centrality. *J. Math. Sociol.* **25**, 163–177 (2001).
93. Timmer, J. & König, M. On generating power law noise. *Astron. Astrophys.* **300**, 707–710 (1995).
94. Davies, P. I. & Higham, N. J. Numerically stable generation of correlation matrices and their factors. *BIT Numer. Math.* **40**, 640–651 (2000).
95. Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Farmer, J. D. Testing for nonlinearity in time series: the method of surrogate data. *Physica D* **58**, 77–94 (1992).
96. Maslov, S. & Sneppen, K. Specificity and stability in topology of protein networks. *Science* **296**, 910–913 (2002).
97. Storn, R. & Price, K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **11**, 341–359 (1997).
98. Bassett, D. S. & Sporns, O. Network neuroscience. *Nat. Neurosci.* **20**, 353–364 (2017).
99. Harris, C. R. et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).
100. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in python. *Nat. Methods* **17**, 261–272 (2020).
101. LaPlante, R. A., Douw, L., Tang, W. & Stufflebeam, S. M. The connectome visualization utility: software for visualization of human brain networks. *PLoS ONE* **9**, e113838 (2014).
102. Vallat, R. Pingouin: statistics in Python. *J. Open Source Softw.* **3**, 1026 (2018).
103. Shinn, M. Refinement type contracts for verification of scientific investigative software. in *Verified Software. Theories, Tools, and Experiments* (eds Chakraborty, S. & Navas, J. A.) 143–160 (Springer, 2020).

## Acknowledgements

We thank J. Burt for help with plotting; A. Howell for assistance with HCP data; and T. Ito and J. C. C. Vila for helpful discussions. Funding was provided by European Molecular Biology Organization grant ALTF 712-2021, the Winston Churchill Foundation of the United States and the Gruber Foundation (M.S.); National Institute of Mental Health (NIMH) grant K00MH122372 (S.N.); the Swiss National Science Foundation (P2ZHP1|161626) (K.H.P.); Agence Nationale de la Recherche (ANR-20-NEUC-003-01) and MIAI@Grenoble Alpes (ANR-19-P3IA-0003) (S.A.); the Usona Institute (2015-2056), the Heffter Research Institute (1-190420), the Swiss Neuromatrix Foundation (2015-103 and 2016-0111) and the Swiss National Science Foundation under the framework of Neuron Cofund (01EW1908) (F.X.V.); the National Institute for Health and Care Research (NIHR) (Senior Investigator Award) and the NIHR Cambridge Biomedical Research Centre (E.T.B.); SFARI (Pilot Award) (A.A. and J.D.M.); and the NIMH (R01MH112746) (J.D.M.). Data collection and sharing for this project were provided, in part, by the Cambridge Centre for Ageing and Neuroscience (Cam-CAN). Cam-CAN funding was provided by the UK Biotechnology and Biological Sciences Research Council (grant BB/H008217/1), together with support from the UK Medical Research Council and the University of Cambridge.

## Author contributions

M.S. and E.T.B. conceived the research. M.S. designed the experiments. M.S. and A.H. performed the experiments. M.S. and J.D.M. analyzed and interpreted results. K.H.P., J.L.J., F.M., S.N., D.S., R.T.C., J.H.K., F.X.V. and A.A. contributed data, methodology and resources. M.S., L.T. and S.A. performed the mathematical analysis. D.L., E.T.B. and J.D.M. provided supervision and funding. M.S., A.H. and L.T. wrote the first draft of the manuscript. All authors edited, revised and approved the manuscript.

## Competing interests

K.H.P. is currently an employee of Hoffmann-La Roche. J.H.K. has consulting agreements (less than \$5,000 per year) with the following:

Aptinyx; Atai Life Sciences; AstraZeneca Pharmaceuticals; Biogen; Biomedisyn Corporation; Bionomics; Boehringer Ingelheim; Cadent Therapeutics; Clexio Bioscience; COMPASS Pathways; Concert Pharmaceuticals; Epiodyne; EpiVario; Greenwich Biosciences; Heptares Therapeutics; Janssen Research & Development; Jazz Pharmaceuticals; Otsuka America Pharmaceutical; Perception Neuroscience Holdings; Spring Care; Sunovion Pharmaceuticals; Takeda Industries; and Taisho Pharmaceutical Company. J.H.K. serves on the scientific advisory boards of Biohaven Pharmaceuticals; BioXcel Therapeutics (Clinical Advisory Board); Cadent Therapeutics (Clinical Advisory Board); Cerevel Therapeutics; EpiVario; Eisai; Jazz Pharmaceuticals; Lohocla Research Corporation; Novartis Pharmaceuticals Corporation; PsychoGenics; Neumora Therapeutics; Tempero Bio; and Terran Biosciences. J.H.K. is on the board of directors of Freedom Biosciences. J.H.K. has stock and/or stock options in Biohaven Pharmaceuticals; Sage Pharmaceuticals; Spring Care; Biohaven Pharmaceuticals Medical Sciences; EpiVario; Neumora Therapeutics; Terran Biosciences; and Tempero Bio. J.H.K. is editor of *Biological Psychiatry* with income greater than \$10,000. D.L. is a co-founder of Neurogazer. A.A. and J.D.M. are co-founders of Manifest Technologies and serve on the technical advisory board of Neumora

Therapeutics. E.T.B. serves on the scientific advisory board of Sosei Heptares and as a consultant for GlaxoSmithKline. The remaining authors declare no competing interests.

## Additional information

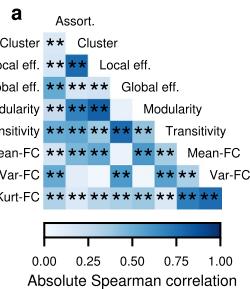
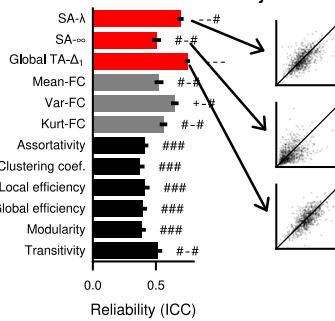
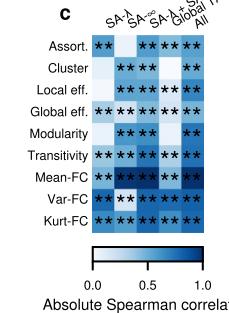
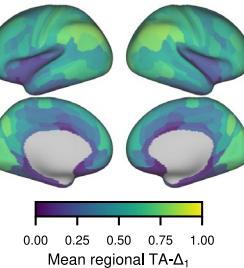
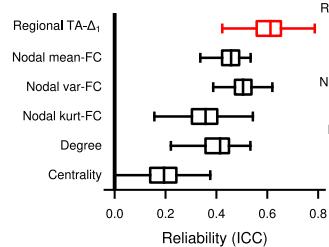
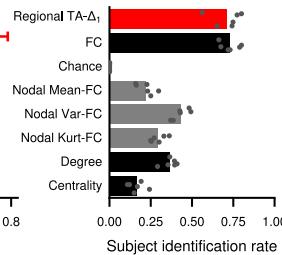
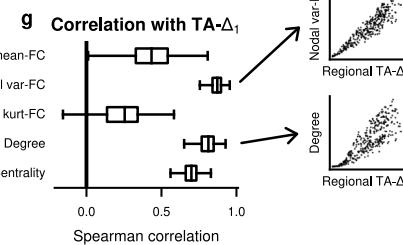
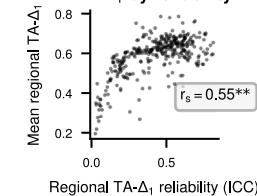
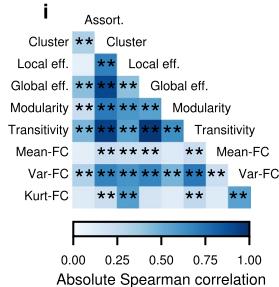
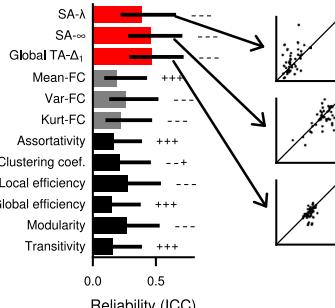
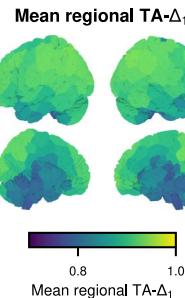
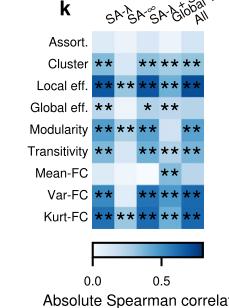
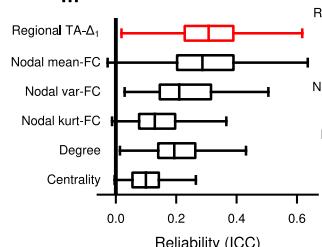
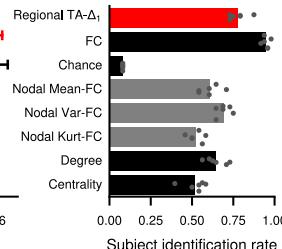
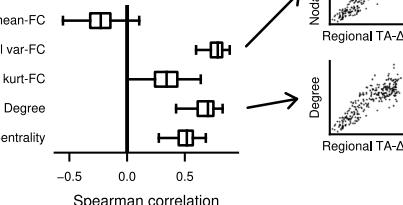
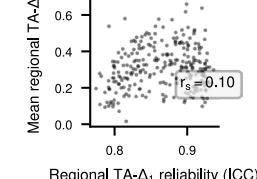
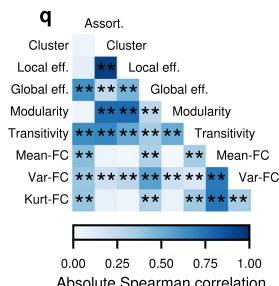
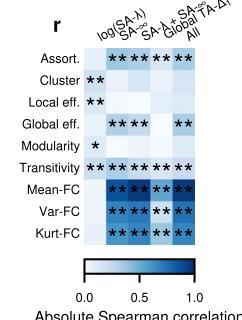
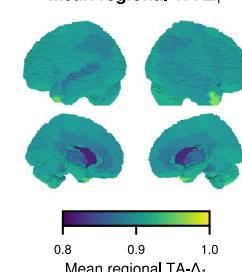
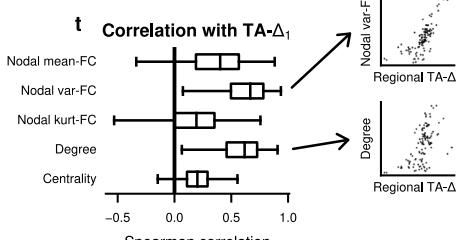
**Extended data** is available for this paper at <https://doi.org/10.1038/s41593-023-01299-3>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41593-023-01299-3>.

**Correspondence and requests for materials** should be addressed to John D. Murray.

**Peer review information** *Nature Neuroscience* thanks Janine Bijsterbosch, Javier Gonzalez-Castillo, Lucina Uddin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

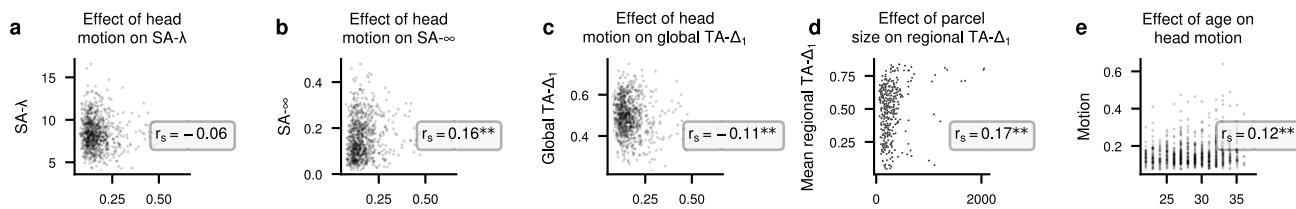
**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**HCP-GSR****Correlation among graph metrics****b Test-retest reliability****Correlation with graph metrics****d Mean regional TA-Δ<sub>1</sub>****e Reliability of nodal metrics****f Fingerprinting****g****h TA-Δ<sub>1</sub> by reliability****Yale Test-Retest****Correlation among graph metrics****j Test-retest reliability****Correlation with graph metrics****m Reliability of nodal metrics****n Fingerprinting****o Correlation with TA-Δ<sub>1</sub>****p TA-Δ<sub>1</sub> by reliability****Cam-CAN****Correlation among graph metrics****r Correlation with graph metrics****s Mean regional TA-Δ<sub>1</sub>****t****Extended Data Fig. 1 | See next page for caption.**

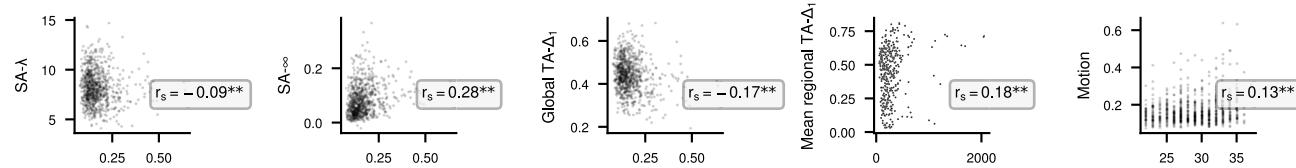
**Extended Data Fig. 1 | SA and TA are important features of rs-fMRI.** Data from the HCP-GSR (a-h, N=850 subjects), Yale-TRT (i-p, N=12 subjects, 4 sessions each), and Cam-CAN (q-t, N=646 subjects) datasets. For all subpanels, unless otherwise indicated, \* indicates  $p < 0.05$  and \*\* indicates  $P < 0.01$  on a two-sided test, and boxplots indicate the median, first/third quartiles, and range, with outliers hidden for visualization. (a,i,q) Correlation across all subjects in the HCP dataset, with Bonferroni FWER corrected two-sided  $P$ -values. (b,j) Test-retest reliability of graph metrics, quantified by intraclass correlation coefficient (ICC). Error bars indicate 95% CI. Following each bar is a string of three characters indicating significance: the first indicates significantly less than SA- $\lambda$ , the second SA- $\infty$ , and the third global TA- $\Delta_1$ , where # indicates  $P < .01$ , + indicates  $P < .05$ , and – indicates

$P > .05$ , by a one-sided bootstrap resampling procedure. Inset scatterplots show correlation across subjects for two example sessions. (c,k,r) Correlation across subjects between graph metrics and SA- $\lambda$ , SA- $\infty$ , or global TA- $\Delta_1$ . ‘SA- $\lambda$  + SA- $\infty$ ’ and ‘all’ indicate a cross-validated linear model with two or three terms, respectively, using Bonferroni FWER corrected two-sided  $P$ -values. (d,l,s) The brain map depicting regional TA- $\Delta_1$ , averaged across all subjects. (e,m) Distribution of reliability for each brain region. (f,n) Mean fraction of subjects correctly identified by a fingerprinting analysis. Points indicate identification performance on each of six possible test-retest pairs from the four sessions. (g,o,t) Correlation across regions of regional TA- $\Delta_1$  with nodal graph metrics for each subject. (k) The regional TA- $\Delta_1$  for each region plotted against its reliability.

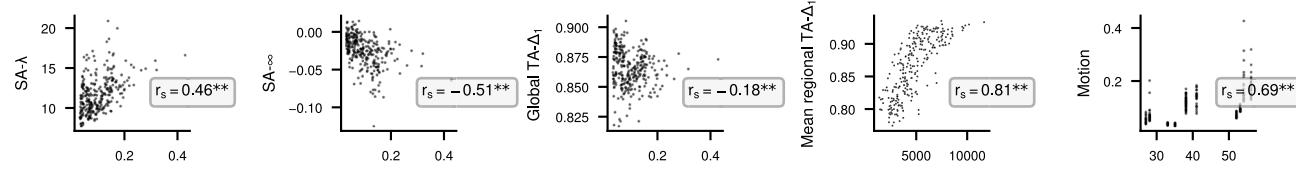
## HCP:



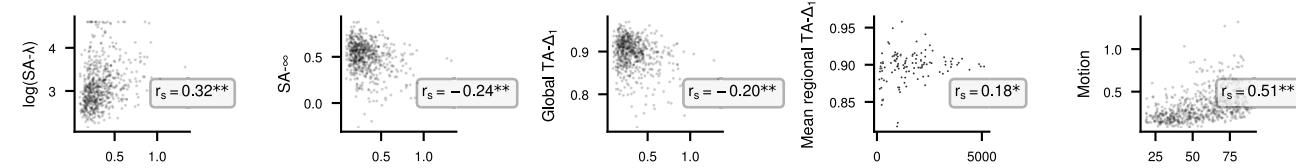
## HCP-GSR:



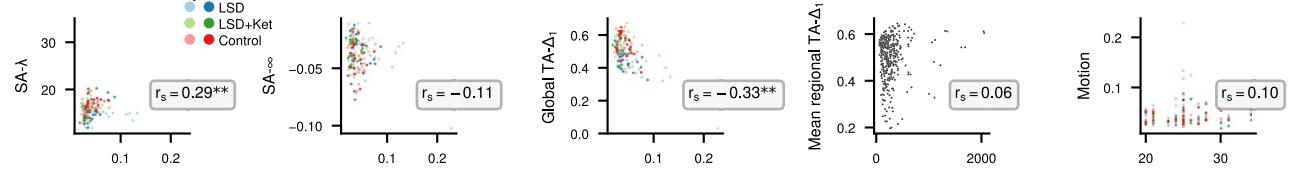
## Yale Test-Retest:



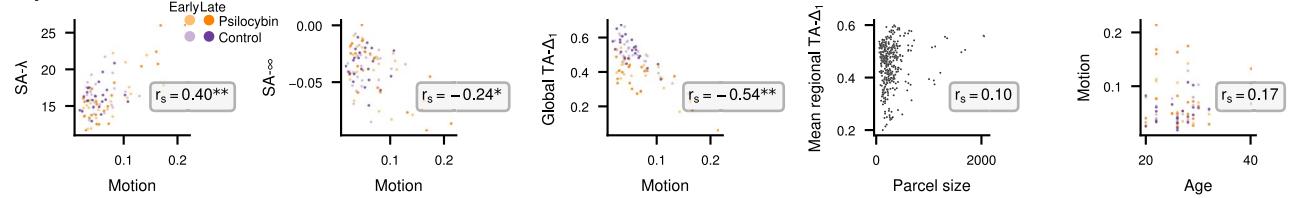
## Cam-CAN:



## LSD Dataset:



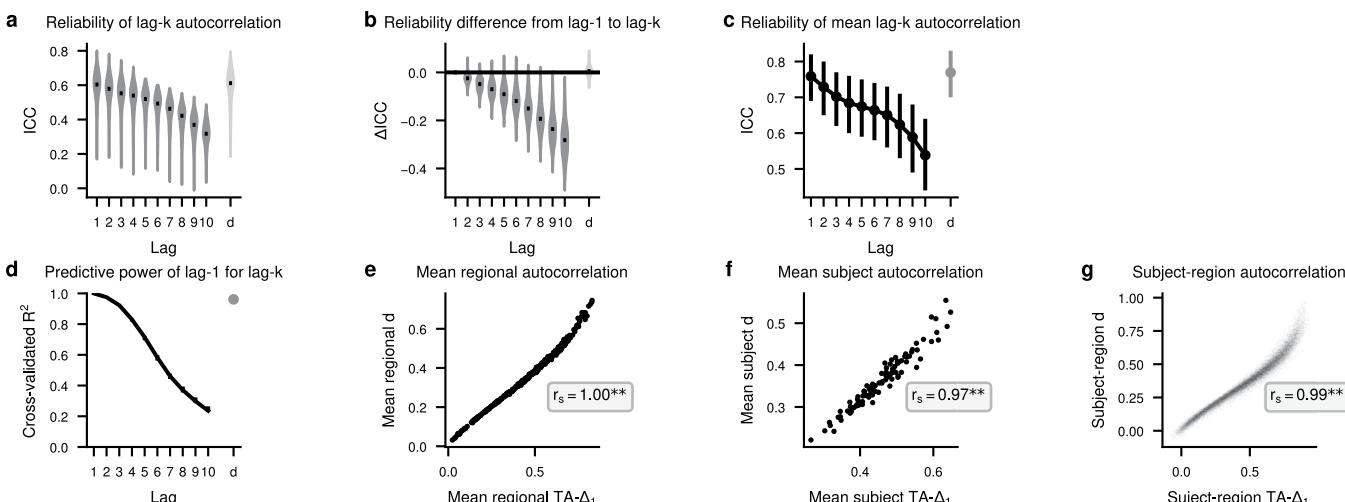
## Psilocybin Dataset:



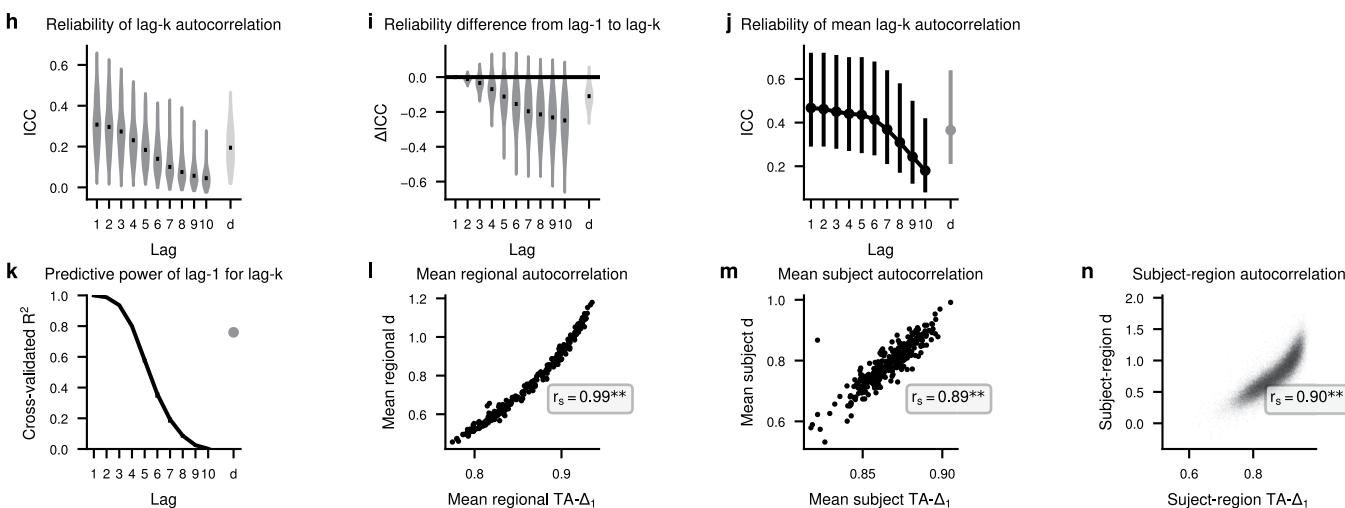
**Extended Data Fig. 2 | Correlation of motion and parcel size with spatial and temporal autocorrelation.** (a-c) For each dataset, the mean framewise displacement ('motion') is compared to (a)  $SA-\lambda$ , (b)  $SA-\infty$ , and (c)  $TA-\Delta_1$ , with inset Spearman correlation ( $r_s$ ). (d) The average regional  $TA-\Delta_1$  across all subjects is plotted against the region's parcel size, with inset Spearman correlation. Parcel size is measured in surface area for HCP and HCP-GSR, and number of voxels for

Yale-TRT and Cam-CAN. (e) Subject's mean framewise displacement ('motion') is plotted against the subject's age. \* indicates Spearman correlation two-sided  $P < .05$ , and \*\* indicates  $P < .01$ . HCP: N=883, HCP-GSR: N=850, Yale-TRT: N=12 subjects, 4 sessions each, Cam-CAN: N=646, LSD: N=24, Psilocybin: N=23 subjects.

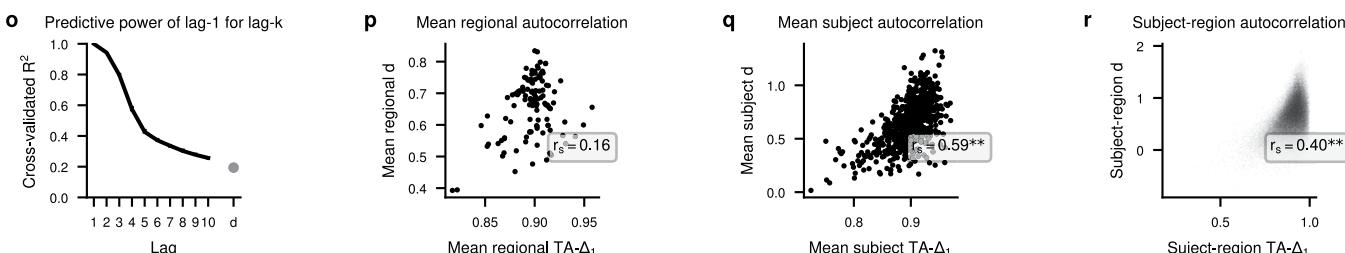
## HCP



## Yale-TRT

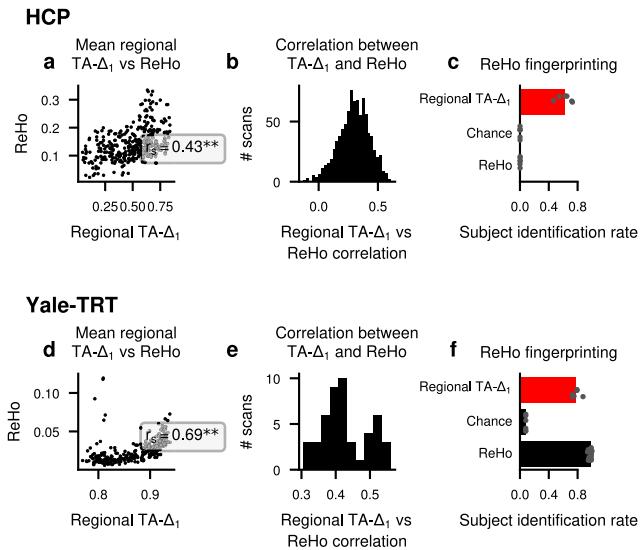


## Cam-CAN

Extended Data Fig. 3 | TA-Δ<sub>1</sub> captures individual variation in long memory dynamics.

(a,c,h-j) TA-Δ<sub>1</sub>, the first lag term in the ACF, is more reliable than higher lag terms. (a,h) For each lag  $k$ , we computed the reliability of the corresponding term in the ACF,  $\text{ACF}_x(k)$ , as measured by ICC. The median reliability across brain regions decreased as  $k$  increased, and regional TA-Δ<sub>1</sub> (that is,  $\text{ACF}_x(1)$ ) maximized median reliability. Additionally,  $d$  had similar reliability to regional TA-Δ<sub>1</sub>. (b,i) To confirm this occurs for each brain region individually, we found the difference in reliability between the  $\text{ACF}_x(l)$  and  $\text{ACF}_x(k)$  ( $\Delta\text{ICC}$ ), which also decreased across lags.  $d$  was similar to regional TA-Δ<sub>1</sub>. (c,j) Global TA-Δ<sub>1</sub> reliability also decreased with increasing lag, and  $d$  averaged across regions had similar reliability to global TA-Δ<sub>1</sub>, as measured by ICC. Error bars indicate 95% confidence interval. Thus, the reliability at short compared to long time lags, and the similarity to  $d$ , motivate a prioritization for TA-Δ<sub>1</sub> over higher lags. (d,k,o) TA-Δ<sub>1</sub> is predictive of higher terms

of the ACF. TA-Δ<sub>1</sub> was used to predict higher ACF terms in a regression model (see Methods). Mean cross-validated R<sup>2</sup> is shown, where error bars (sometimes hidden under the line) indicate maximum and minimum cross-validated R<sup>2</sup>. (e-g,l-n,p-r) Individual variation in  $d$ , the long-memory or fractional integration term from an ARFIMA(0,  $d$ , 0) model, can be captured by individual variation in TA-Δ<sub>1</sub>. (e,l,p) regional TA-Δ<sub>1</sub> averaged across subjects is highly correlated with regional  $d$  averaged across subjects. (e:  $P=0$ , l:  $P=10^{-260}$ , p:  $P=.10$ ) (f,m,q) Global TA-Δ<sub>1</sub> is highly correlated with  $d$  averaged across regions within a subject. (f:  $P=0$ , m:  $P=10^{-101}$  q:  $10^{-61}$ ) (g,n,r) Without averaging, across regions and subjects,  $d$  is correlated with TA-Δ<sub>1</sub>. (g:  $P=0$ , n:  $P=0$ , r:  $P=0$ ) For all figures, unless otherwise indicated,  $r_s$  indicates Spearman correlation, where \* indicates two-sided  $P < .05$ , and \*\* indicates  $P < .01$ ,  $N=883$  subjects.



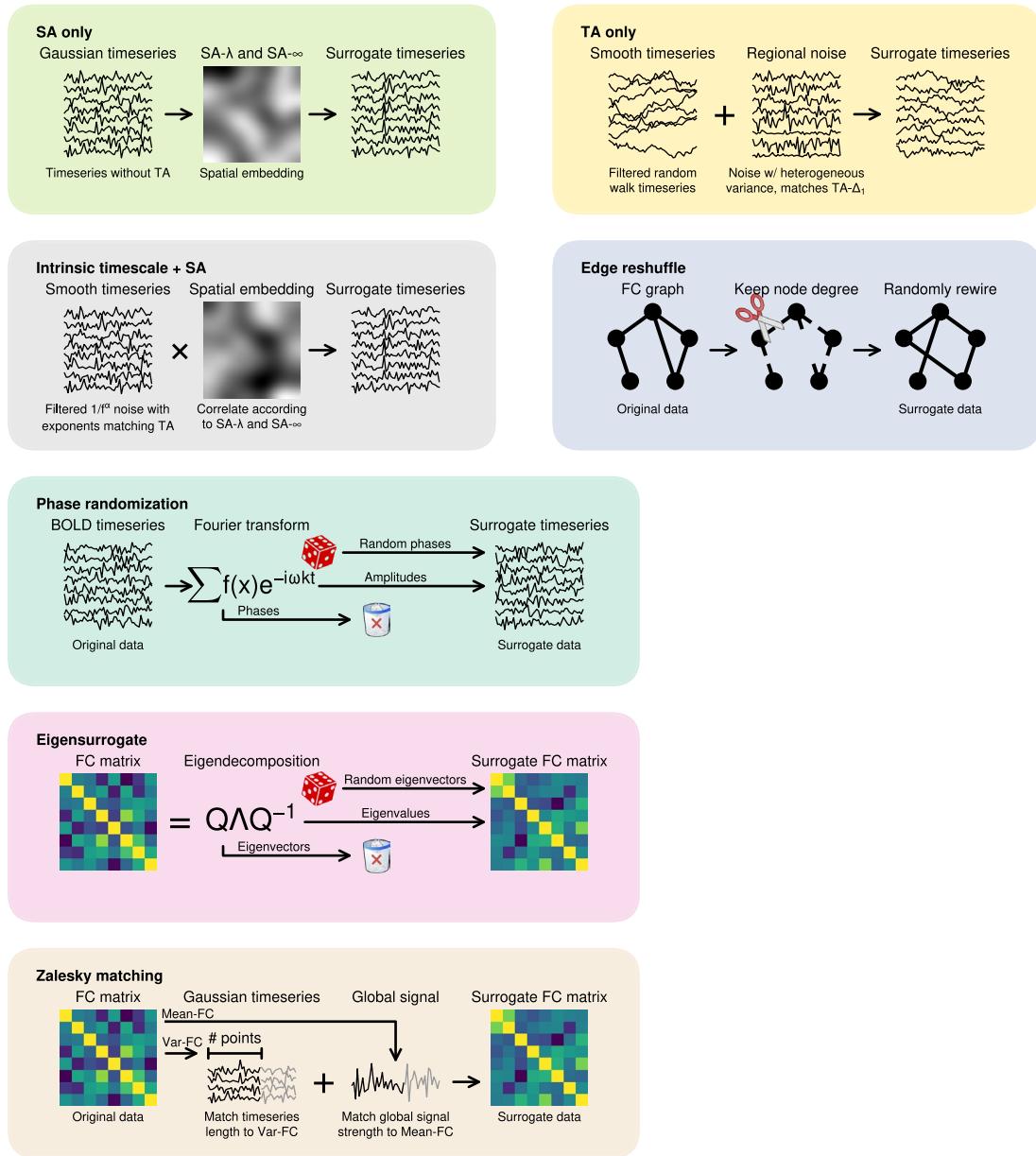
**Extended Data Fig. 4 | Relationship between regional TA- $\Delta_1$  and regional homogeneity.** Comparisons are shown for the HCP (a-c, N=883) and Yale-TRT (d-f, N=12, 4 sessions each) datasets. (a,d) The global TA- $\Delta_1$  averaged across subjects, compared to the regional homogeneity (ReHo) averaged across subjects. Each point represents a parcel. \*\* indicates two-sided Spearman

correlation  $P < .01$ . (a:  $P=10^{-18}$ , d:  $P=10^{-39}$ ) (b,e) Distribution of correlations of global TA- $\Delta_1$  and regional homogeneity. Compare to Fig. 1j. (c,f) Average fingerprinting performance of regional homogeneity, compared to global TA- $\Delta_1$  and chance. Points are overlaid for each pair of datasets compared. Compare to Fig. 1i.

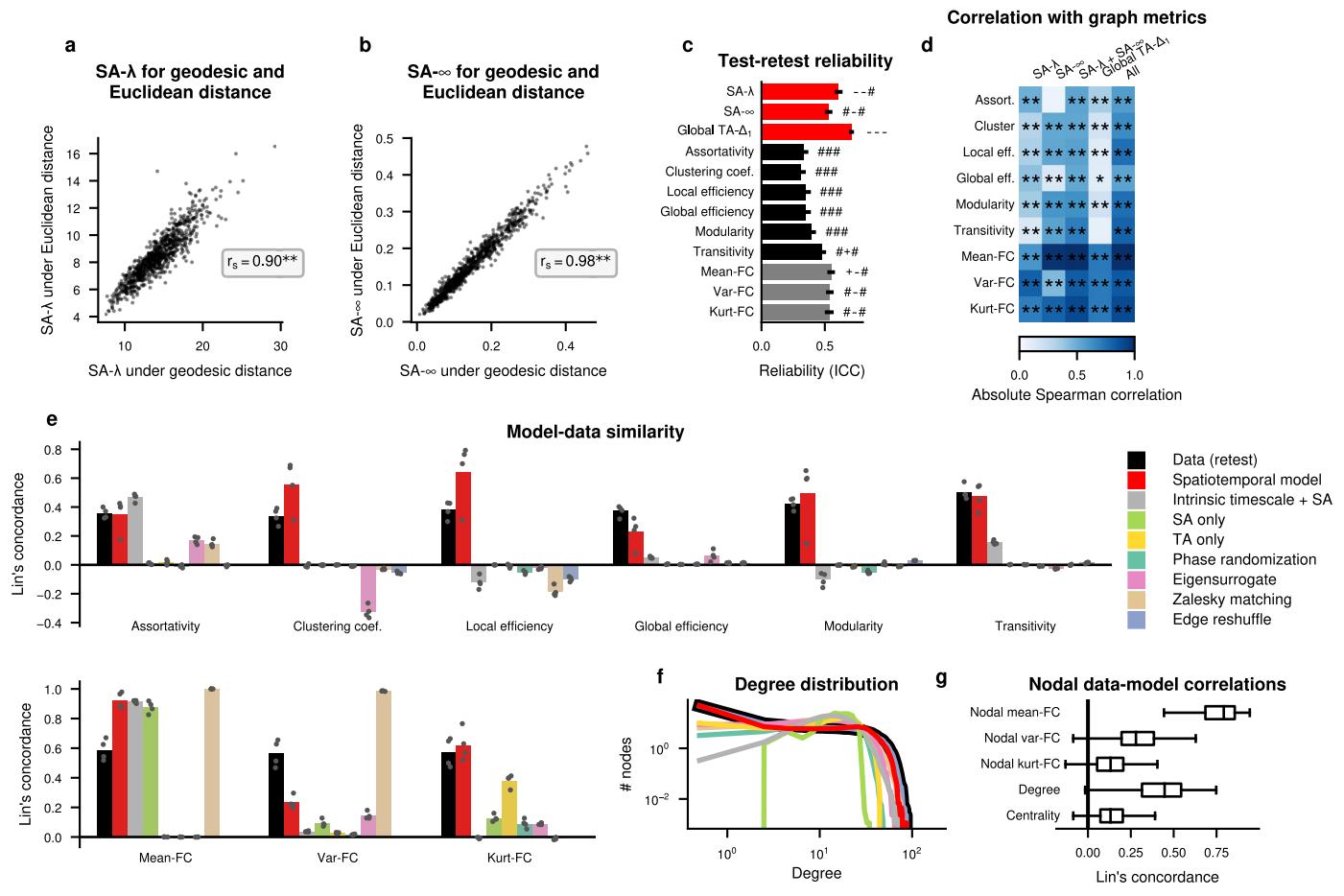


**Extended Data Fig. 5 | Correlation of model and data graph metrics for all models.** For each model, each subject's empirical graph metrics are plotted against model graph metrics for metrics from Fig. 2e. Spearman correlation ( $r_s$ ) and Lin's concordance (Lin) are inset. \* indicates Spearman correlation two-sided

$P < .05$ , and \*\* indicates  $P < .01$ , N=883 subjects. Zalesky matching operates at the level of the FC matrix, and thus, TA- $\Delta_1$  could not be computed. Edge reshuffle operates at the level of the graph, preventing any of these measures from being computed.

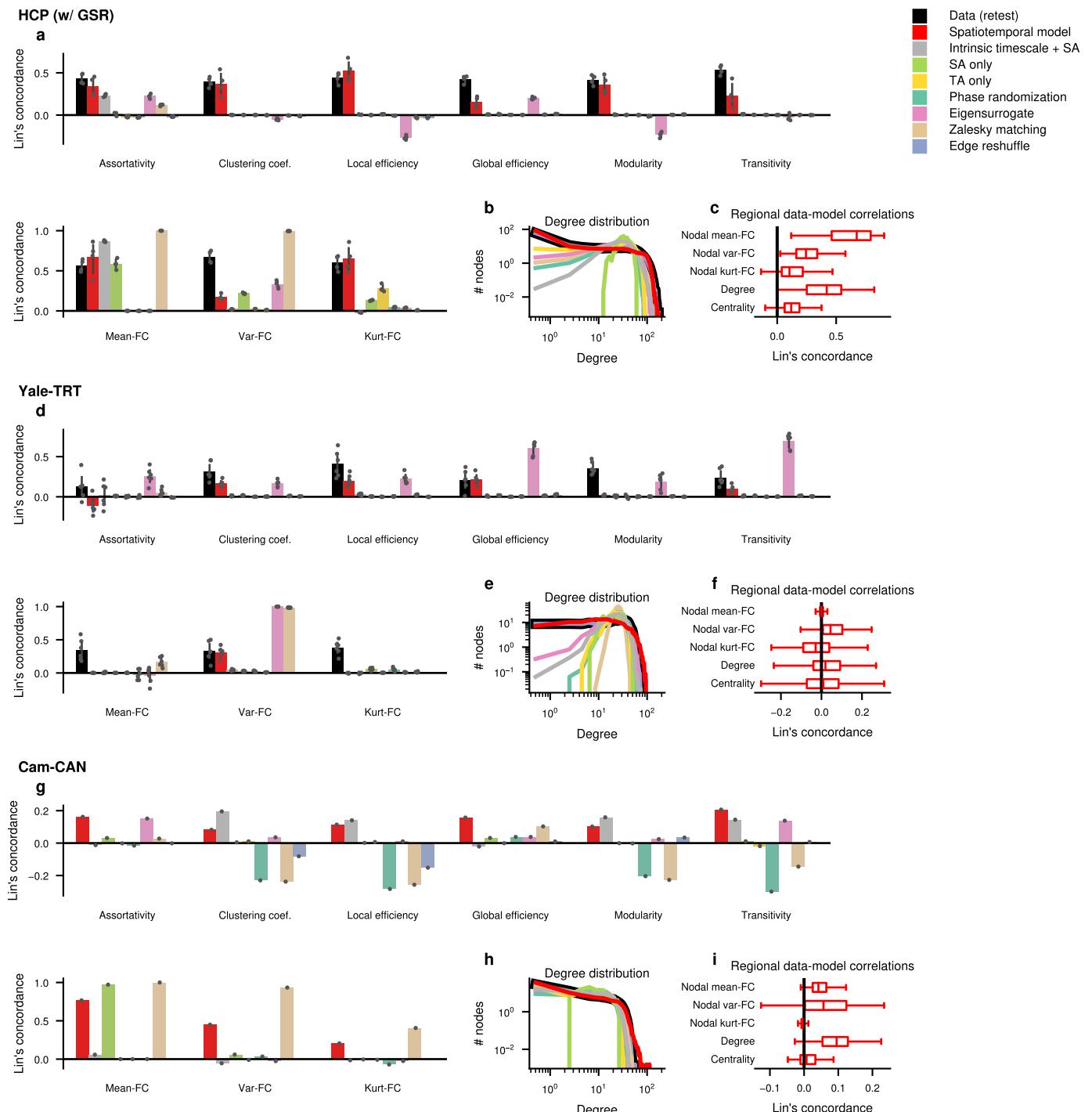


**Extended Data Fig. 6 | Schematics of all models.** All models considered are described in their corresponding schematics. The spatiotemporal model is shown in Fig. 2a.



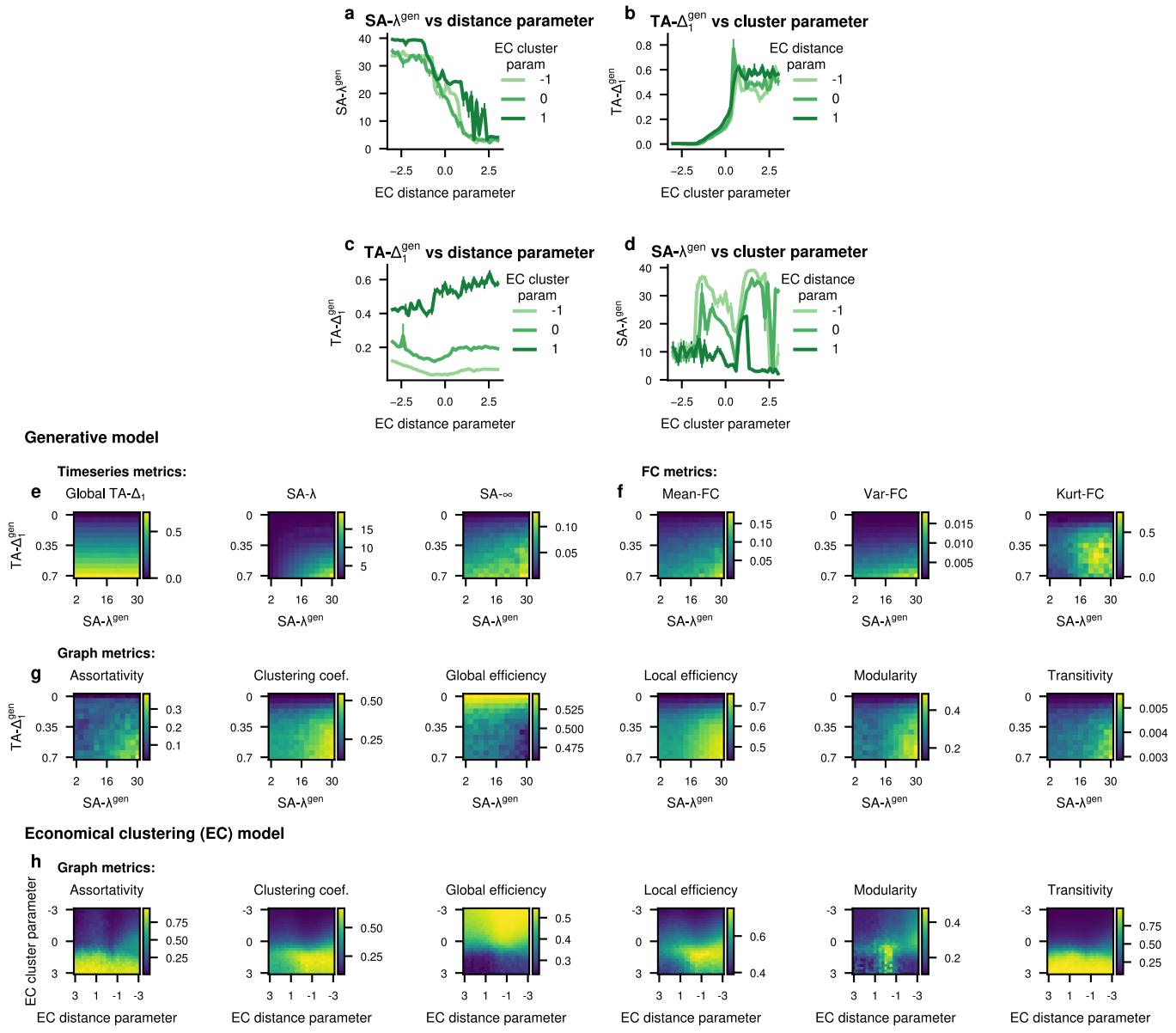
**Extended Data Fig. 7 | Model fitting under geodesic distance.** We analyzed the data and model using geodesic instead of Euclidean distance in the right hemisphere. (a-b) We computed  $SA-\lambda$  and  $SA-\infty$  under geodesic distance for the right hemisphere. The geodesic and Euclidean distances lead to highly correlated measurements for  $SA-\lambda$  (a) and  $SA-\infty$  (b). \* indicates Spearman correlation two-sided  $P < .05$ , and \*\* indicates  $P < .01$ . (a:  $P=0$ , b:  $P=0$ ) (c) Comparison of reliability of graph measures to  $SA-\lambda$  and  $SA-\infty$  when computed using geodesic distance, as measured using ICC. Following each bar is a string of three characters: the first indicates significantly less than  $SA-\lambda$ , the second  $SA-\infty$ , and the third global  $TA-\Delta_1$ , where # indicates  $P < .01$ , + indicates  $P < .05$ , and - indicates  $P > .05$  by a bootstrap resampling procedure on a one-sided test. Compare

to Fig. 1d. (d) Correlation of graph metrics with  $SA-\lambda$ ,  $SA-\infty$ , a linear model of both (' $SA-\lambda + SA-\infty$ '), and a linear model incorporating both of these plus global  $TA-\Delta_1$  ('all'), when computed using geodesic distance instead of Euclidean distance. Compare to Fig. 1e. \* indicates Spearman correlation two-sided  $P < .05$ , and \*\* indicates  $P < .01$ . (e-g) The spatiotemporal model and all comparison models were fit using geodesic instead of Euclidean distance. (e) The similarity of graph metrics in the model and data, as measured by Lin's concordance, under geodesic distance. Compare to Fig. 2e. (f) The similarity of degree distribution in the model and data under geodesic distance. Compare to Fig. 2f. (g) The similarity of nodal graph metrics in the model and data under geodesic distance. Compare to Fig. 2g. For all panels,  $N=883$  subjects.



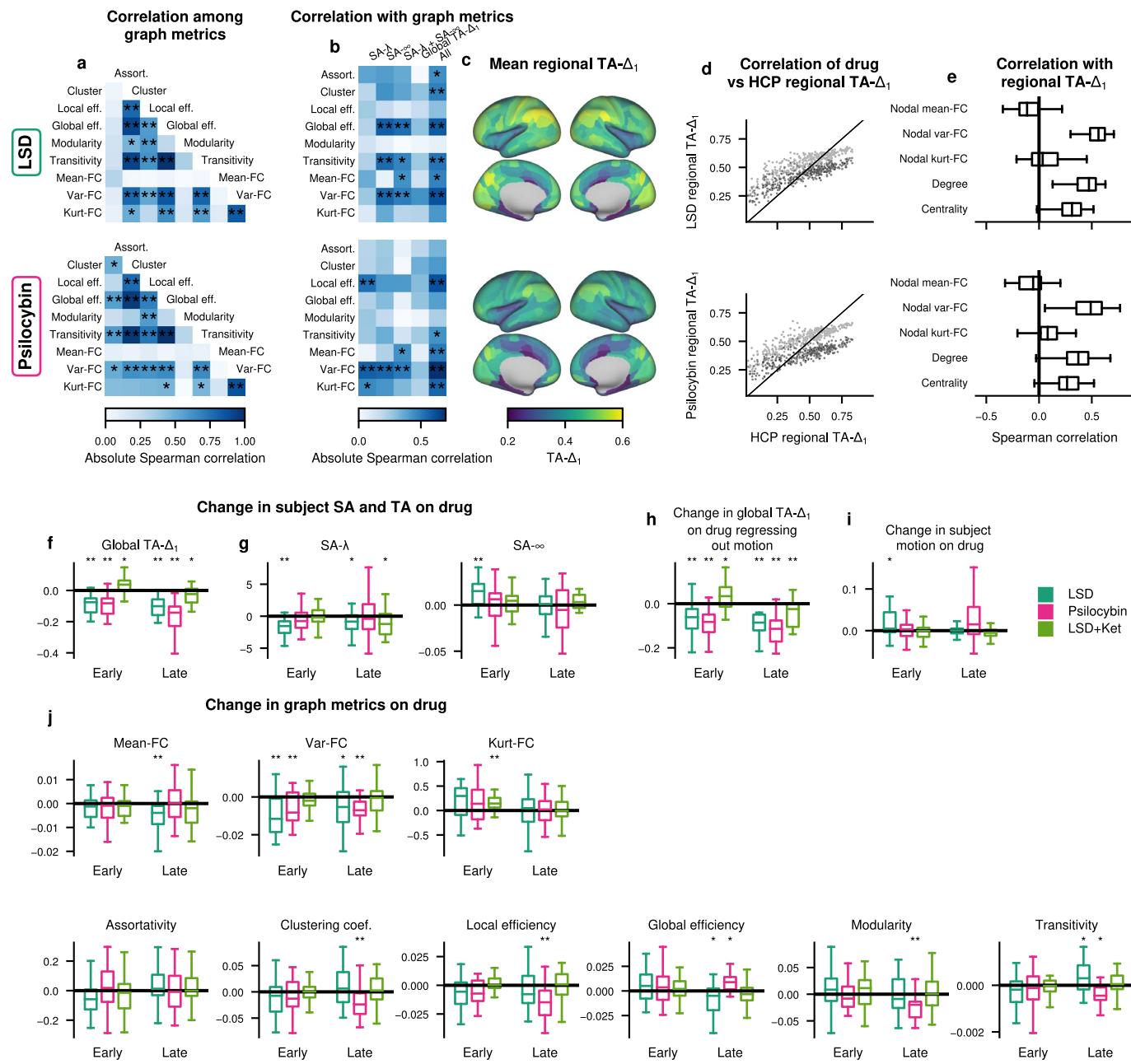
**Extended Data Fig. 8 | Comparison of model fits for all datasets.** (a,d,g) Lin's concordance between model and data for each model. Bars represent the mean across four (HCP-GSR), six (Yale-TRT), or one (Cam-CAN) scanning sessions, and points indicate the Lin's concordance between the model and data for each session. For comparison, black indicates Lin's concordance between separate sessions from the same subject in the HCP-GSR and Yale-TRT datasets, where

dots indicate pairs of sessions. (b,e,h) Log-log degree distribution for each model compared to the data (black). (c,f,i) Distribution of Lin's concordance of nodal metrics between model and data for each region. Boxplots show median, first/ third quartiles, and range, with outliers hidden for visualization, N=883 subjects. Compare to Fig. 2e-g. Statistics of these fits can be found in Table S1.



**Extended Data Fig. 9 | Relationship between the spatiotemporal model and economical clustering (EC) model.** (a) For three given values of the EC cluster parameter ( $y$ ), the distance parameter ( $\eta$ ) was varied and the  $TA-\Delta_1^{gen}$  parameter of the best fit spatiotemporal model is shown. (b) For three given values of the EC distance parameter, the EC cluster parameter was varied and the  $SA-\lambda_{gen}$  parameter of the best fit spatiotemporal model is shown. Error bars show standard error across 10 simulations of the EC model. (e-g) The Homogeneous

variant of the model was simulated for different values of parameters  $SA-\lambda_{gen}$  and  $TA-\Delta_1^{gen}$ . For each timeseries metric (e), weighted graph metric (f), and unweighted graph metric (g), the metric value is plotted as a heatmap. (h) The economical clustering (EC) model was simulated for different values of its distance and cluster parameters. The value of each of the graph metrics is plotted as a heatmap. For all panels,  $N=883$  subjects.



**Extended Data Fig. 10 | SA and TA under serotonergic modulation.** (a-c,e) Analysis of scans after administration of LSD (top) or psilocybin (bottom). (a) Correlation of graph metrics across all subjects, with Bonferroni FWER corrected two-sided  $P$ -values, \* indicates  $P < .05$ , \*\* indicates  $P < .01$ . Compare to Figure 1b. (b) Correlation across subjects between graph metrics and SA- $\lambda$ , SA- $\infty$ , or global TA- $\Delta_1$ . ‘SA- $\lambda$ +SA- $\infty$ ’ and ‘all’ columns indicate a linear model with leave-one-out cross-validation and Bonferroni FWER corrected two-sided  $P$ -values, \* indicates  $P < .05$ , \*\* indicates  $P < .01$ . Compare to Figure 1e. (c) The brain map depicting regional TA- $\Delta_1$ , averaged across all subjects. (d) Correlation of regional TA- $\Delta_1$  in HCP with those under LSD (top) or psilocybin (bottom). Spearman correlation  $> .85$ , two-sided significance  $P < .0001$  of the correlation of HCP and LSD or psilocybin regional TA- $\Delta_1$  determined with SA-preserving scrambles<sup>11</sup>.

Gray points indicate the placebo condition. (e) Correlation across regions of regional TA- $\Delta_1$  with nodal graph metrics for each subject. Boxplots show the median, first/third quartiles, and outlier-excluded minimum and maximum of the distribution. Compare to Figure 1j. (f-j) Difference between drug and control across subjects for LSD (blue), psilocybin (pink), and LSD with ketanserin (green), for both early and late scans. Metrics plotted are TA as quantified by TA- $\Delta_1$  (f), SA as quantified by SA- $\lambda$  and SA- $\infty$  (g), the residual of a regression model using motion to predict TA- $\Delta_1$ , subject motion as quantified by mean framewise displacement (i), and graph metrics (j). \* indicates  $P < .05$ , \*\* indicates  $P < .01$ , two-sided Wilcoxon sign-rank test. Boxplots show the median, first/third quartiles, and outlier-excluded minimum and maximum of the distribution. For all panels, LSD: N=24, Psilocybin: N=23 subjects.

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

|                 |   |
|-----------------|---|
| Data collection | No software was used for data collection.   |
| Data analysis   | We released a software package which allows the central analyses in this paper to be performed quickly and easily, the "spatiotemporal" Python package, which can be installed through pip or downloaded at <a href="https://github.com/murraylab/spatiotemporal">https://github.com/murraylab/spatiotemporal</a> . Analysis utilized Python 3.6 with the libraries numpy 1.19.5, scipy 1.5.4, bctpy 0.5.0, pingouin 0.3.8, and paranoid-scientist 0.2.2. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The HCP data are available at: <https://www.humanconnectome.org/study/hcp-young-adult>. The Human Connectome Project with Global Signal Regression (HCP-

GSR) dataset was derived from the same data as the HCP dataset. The Yale-TRT data are available at: [http://fcon\\_1000.projects.nitrc.org/indi/retro/yale\\_trt.html](http://fcon_1000.projects.nitrc.org/indi/retro/yale_trt.html). The Cam-CAN data are available at: <https://www.cam-can.org/index.php?content=dataset>. The LSD data and Psilocybin data are available upon request.

## Human research participants

Policy information about [studies involving human research participants](#) and [Sex and Gender in Research](#).

|                             |   |
|-----------------------------|---|
| Reporting on sex and gender | Our sample consisted of both males and females, but effects of sex were not analyzed in our current study.  |
| Population characteristics  | The HCP dataset consisted of subjects aged 22-37; the Yale-TRT dataset aged 27-56; the Cam-CAN dataset aged 18-88; the LSD dataset aged 20-34; and the Psilocybin dataset aged 20-40. No other population characteristics were analyzed.  |
| Recruitment                 | Participants were recruited using diverse strategies to achieve goals which were not relevant to the current study, such as related sibling pairs for the HCP data. Recruitment strategy is unlikely to impact our conclusions.   |
| Ethics oversight            | All participants provided written informed consent statements before participation in the study. The HCP data were acquired using protocols approved by the Washington University institutional review board. The Yale-TRT data were collected with approval by the Yale University institutional review board. The Cam-CAN data were collected with approval by the Cambridgeshire 2 Research Ethics Committee. The LSD and Psilocybin data were collected with approval by the Cantonal Ethics Committee of Zurich, and the Swiss Federal Office of Public Health, Bern, Switzerland, authorized the use of LSD and Psilocybin in humans. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

|                 |  |
|-----------------|--|
| Sample size     | We did not perform sample size calculations, and instead used all data available from the respective datasets.   |
| Data exclusions | HCP: 33 subjects were excluded from the GSR analysis due to non-convergence of the preprocessing pipeline. Cam-CAN: Six subjects and one cerebellar region were excluded due to missing data. Yale TRT: No subjects were excluded. LSD: One subject was excluded due to failed registration. Psilocybin: One subject was excluded due to missing data. |
| Replication     | We replicated our results on five different datasets, as described in the paper.   |
| Randomization   | In the LSD and Psilocybin data, the order of experimental sessions (drug and placebo) was randomised. The randomisation was balanced and completed by a study nurse who had no other role in the trial.  |
| Blinding        | The LSD and Psilocybin datasets were double-blinded. No other group allocation was performed in this study, so no other blinding was necessary.  |

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

|                                     |                               |
|-------------------------------------|-------------------------------|
| n/a                                 | Involved in the study         |
| <input checked="" type="checkbox"/> | Antibodies                    |
| <input checked="" type="checkbox"/> | Eukaryotic cell lines         |
| <input checked="" type="checkbox"/> | Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | Animals and other organisms   |
| <input checked="" type="checkbox"/> | Clinical data                 |
| <input checked="" type="checkbox"/> | Dual use research of concern  |

## Methods

|                                     |                        |
|-------------------------------------|------------------------|
| n/a                                 | Involved in the study  |
| <input checked="" type="checkbox"/> | ChIP-seq               |
| <input checked="" type="checkbox"/> | Flow cytometry         |
| <input type="checkbox"/>            | MRI-based neuroimaging |

## Magnetic resonance imaging

### Experimental design

Design type

Resting state

Design specifications

HCP dataset: four 14.5 minute scans per subject on two different days, TR=720. Yale-TRT dataset: 24 six-minute scans per subject spread across four days, TR=1000. Cam-CAN dataset: one 8.5 minute scan per subject, TR=1970. LSD dataset: Two 10 minute scans for each of the three treatment groups in each subject, TR=2500. Psilocybin dataset: three 10 minute scans for both treatment groups, TR=2430.

Behavioral performance measures

Behavioural performance measures are not applicable to this study.

### Acquisition

Imaging type(s)

Functional

Field strength

3T

Sequence &amp; imaging parameters

All scans were acquired using a Gradient-Echo Echo-Planar Imaging (EPI) sequence. Full details are provided in the respective publications. Parameters are: HCP: 72 slices, TR = 720 ms, TE = 33.10 ms, multiband factor = 8, flip angle = 52, voxel size = 2 x 2 x 2 mm, FOV = 208 x 180 x 144 mm; TRT: 75 slices, TR = 1000 ms, TE = 30 ms, flip angle = 55, voxel size = 2 x 2 x 2 mm; Cam-CAN: 32 slices, slice thickness of 3.7 mm with an interslice gap of 20%, TR = 1970 ms, TE = 30 ms, flip angle = 78, FOV = 192 x 192 mm; voxel size = 3 x 3 x 4.44 mm; LSD: 45 slices, TR = 2500 ms, TE = 27 ms, field of view = 240 x 240 mm, voxel size = 3 x 3 x 3 mm; sensitivity-encoding reduction factor = 2.0; Psi: 45 slices, TR = 2430 ms, TE = 27 ms, slice thickness = 3 mm, field of view = 240 x 240 mm; voxel size = 3 x 3 x 3 mm; sensitivity-encoding reduction factor = 2.0.

Area of acquisition

Whole brain

Diffusion MRI

 Used Not used

### Preprocessing

Preprocessing software

HCP data, LSD data, and Psilocybin data were preprocessed with the HCP minimal preprocessing pipeline and the Connectome Workbench suite. Yale TRT data were preprocessed with BiolImage Suite. Cam-CAN data were processed with Automatic Analysis 4.2.

Normalization

HCP data, LSD data, and Psilocybin data used surface-based analysis. Yale TRT: Data were normalized to MNI template through affine and non-linear transformation. Cam-CAN: Normalization to MNI template through affine transformation.

Normalization template

MNI

Noise and artifact removal

HCP: Motion correction was performed, and data were denoised using ICA-FIX, and high-pass filtered at 0.01 Hz. A 2 mm spatial smoothing was applied on the cortical surface constrained to the parcel. The first 100 timepoints were discarded.

Yale TRT: Motion correction was performed with SPM5, and data were spatially smoothed with a 2.5 mm gaussian filter. Nuisance regression was performed, including linear, quadratic, and cubic drift, a 24-parameter model of motion, mean cerebrospinal fluid signal, and mean white matter signal.

Cam-CAN: We also applied a second-order Butterworth low-pass filter at half the Nyquist frequency (0.127 Hz) to account for high-frequency motion artifacts.

LSD and Psilocybin: Nuisance regression was performed, including mean ventricle signal, white matter, and motion parameters. Motion scrubbing was applied to remove the frames with the highest movement. All measurements of temporal autocorrelation accounted for the scrubbing.

Volume censoring

Performed only for LSD and Psilocybin data. Frames were removed which satisfied one of the following criteria: (a) the sum of displacement across all six rigid body movement correction parameters exceeded 0.5 mm (assuming 50 mm cortical

sphere radius); (b) RMS of differences in intensity between current and preceding frame normalized by frame intensity which exceeded 1.6 times the median across scans.

## Statistical modeling & inference

### Model type and settings

We used the following models, as described in detail in the paper: Spatiotemporal, TA-only, SA-only, Intrinsic timescale with SA, Homogeneous TA, eigensurrogate, phase randomisation, Zalesky matching, edge reshuffle, economical clustering.

### Effect(s) tested

We tested the models' ability to reproduce graph metrics.

### Specify type of analysis:

Whole brain     ROI-based     Both

### Statistic type for inference (See [Eklund et al. 2016](#))

We used spatial and temporal autocorrelation statistics, defined in the paper.

### Correction

Where applicable, we used Bonferroni correction for multiple comparisons.

## Models & analysis

### n/a Involved in the study

- 
- Functional and/or effective connectivity
- 
- Graph analysis
- Multivariate modeling or predictive analysis

### Functional and/or effective connectivity

FC was computed using Pearson correlation

### Graph analysis

Weighted and binarized subject-level graphs were analyzed. Connectivity measures for weighted graphs: mean-FC, var-FC, kurt-FC, nodal mean-FC, nodal var-FC, nodal kurt-FC. Connectivity measures for binarized graphs: assortativity, clustering coefficient, local efficiency, global efficiency, modularity (Newman-Girvan Q), transitivity, nodal degree, nodal betweenness centrality.