Functional brain networks reflect spatial and temporal autocorrelation

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

[1] Interdepartmental Neuroscience Program, Yale University
[2] Department of Psychiatry, Yale University
[3] Yale College, Yale University
[4] Department of Radiology and Biomedical Imaging, Yale University
[5] Department of Psychiatry, Psychotherapy and Psychosomatics, University Hospital for Psychiatry, Zurich
[6] University of Grenoble Alpes, CNRS, Inria, Grenoble INP, LJK, 38000 Grenoble, France
[7] Zanvyl Krieger Mind/Brain Institute, Johns Hopkins University
[8] Kavli Discovery Neuroscience Institute, Johns Hopkins University
[9] Department of Psychological and Brain Sciences, Johns Hopkins University
[10] Department of Neuroscience, Johns Hopkins University
[11] Department of Psychiatry, University of Cambridge
[12] Department of Physics, Yale University
[*] These authors contributed equally
[†] Correspondence: john.murray@yale.edu

February 22, 2023

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 examine resting-state fMRI (rs-fMRI) data using complex topology measures from network neuroscience. We show that spatial and temporal autocorrelation are reliable statistics which explain numerous measures of network topology. Surrogate timeseries 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.

Keywords: spatial autocorrelation, temporal autocorrelation, spatiotemporal, network neuroscience, resting-state fMRI, complex networks

Introduction

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 extensions to functional connectivity (FC)[1], the matrix of pairwise correlations across brain regions in resting-state fMRI (rs-fMRI). FC is a wildly successful tool for understanding brain function[2]. Nonlinear methods from network neuroscience build upon the success of FC by unlocking properties observable only at higher levels of abstraction[3]. 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[4].

Network neuroscience has profoundly influenced how we understand the organization of rs-fMRI activity across a variety of different domains[3]. One such example is aging, a complex process which impacts many regions of the brain differently. A rich literature shows network topology reflects numerous changes in functional organization as we age, including changes in local and global connectivity across the brain[5,6]. 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 sero-

tonergic receptor agonist lysergic acid diethylamide (LSD) causally impacts the modular organization and global integration of brain networks[7,8]. 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 timeseries 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 timeseries, may mediate the changes in network topology?

The basic statistical properties of rs-fMRI timeseries vary within and across individuals[9]. 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[10,11]. 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. While 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 timepoints in a timeseries, also detailed below. SA and TA are attractive features because they reflect a myriad of physical and biological mechanisms, including molecular[11,12], structural[13,14], activity state[15,16] and organizational[17,18] 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 timeseries statistics, spatial and temporal autocorrelation, 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 between SA, TA, and graph metrics, we developed a spatiotemporal surrogate timeseries model which is parameterized only by SA and TA. After fitting the model to timeseries from individual subjects, we found a large similarity in graph topology between the subject- 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 utilizing diverse methodologies implemented by different teams, focusing on the Human Connectome Project (HCP) dataset[19], with validation in the Yale Test-Retest (TRT)[20] and Cambridge Centre for Ageing and Neuroscience (Cam-CAN)[21] 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 (Figure 1b)[22,23]. 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 (Figure 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 (Figure ED1). 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 spatial and temporal autocorrelation are two such factors.

Spatial autocorrelation Spatial autocorrelation (SA) is the ubiquitous but often ill-defined phenomenon in neuroscience that nearby regions are more similar than distant regions[18,24]. In order to examine SA's test-retest reliability, i.e. 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 vs distance curve (Figure 1c), using Euclidean distance to accommodate both hemispheres and subcortex (Equation 1) (see 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[25,26], with values from 0.5-0.7 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 (Figure 1d) (SA-λ greater than all metrics, $p<.01$, SA-∞ greater than unweighted graph metrics, $p<.05$, bootstrap resampling). The effect cannot be explained by head motion within the scanner (accounting for motion, partial correlation $> .65$, $p < 10^{-10}$ for both SA-λ and SA-∞) or global signal regression (Figure ED1), and is strongest in the HCP dataset, which had the smallest motion and parcel size confounds (Figure ED2). Both SA-λ and SA-∞ are highly correlated with weighted and unweighted graph metrics (Figure 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 (Figure 1e). These results hold in all datasets (Figure ED1). Therefore, SA is reliable and topologically informative.

Temporal autocorrelation Temporal autocorrelation (TA) describes the smoothness, or memory, of the rs-fMRI timeseries over time, and is known to vary heterogeneously across brain regions[27,28], and influence both FC[16,29] and graph topology[30,31]. We quantify TA for each parcellated region in each subject as the Pearson correlation between adjacent time points of the region's timeseries, i.e. the lag-1 temporal autocorrelation (TA-Δ_1) (Figure 1c). TA-Δ_1 is a simple non-parametric measure of TA. While TA-Δ_1 in theory only measures correlation across a single timepoint in the timeseries, in practice, it was effective in measuring correlations at much longer timescales, including parametric estimates of long-memory dynamics (Figure ED3)[32]. Since preprocessing methodology and TR both influence TA-Δ_1, comparisons between TA-Δ_1 must be made within a single dataset. TA-Δ_1 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_1" and "global TA-Δ_1", respectively. We observed highest regional TA-Δ_1 in occipital and parietal regions, and lowest in limbic regions (Figure 1f).

TA is reliable across subjects, both at the whole brain and regional level. At the whole brain level, we computed each subject's global TA-Δ_1 by averaging regional TA-Δ_1 across all regions, and found that global TA-Δ_1 is highly reliable compared to graph metrics (Figure 1d). For each region, we also computed the reliability across subjects, and found that median regional reliability was higher for regional TA-Δ_1 than for other nodal graph metrics (Figure 1g). The reliability of regional TA-Δ_1 varies heterogeneously across the brain (Figure 1h)[33].

This reliability and heterogeneity suggests that regional TA-Δ_1 could be used to identify individual subjects across the population. To identify subjects, we utilized a "fingerprinting" analysis[34]. For a given measure, such as regional TA-Δ_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-Δ_1 identified the single matching session from the pool of 1765 sessions in over 62.4% of subjects, compared to 62.6% with a more traditional fingerprinting analysis using FC (Figure 1i). While 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[35]. We compared this to fingerprinting using several nodal graph metrics, which were unable to match the performance (Figure 1i). Similar results were obtained in other datasets (Figure ED1). We also tested whether this performance is due to differences in parcel boundaries or reliable inhomogeneity within the parcel by comparing regional TA-Δ_1 to the parcel's regional homogeneity. Regional homogeneity showed a moderate correlation with regional TA-Δ_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 (Figure ED4). This indicates that regional TA-Δ_1 is reliable enough to identify an individual subject from a population.

TA-Δ_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-Δ_1 and graph metrics (Figure 1e). To test its influence in conjunction with SA, we developed a linear model incorporating global TA-Δ_1, SA-λ, SA-∞, and a constant, training it on 50% of randomly-selected subjects to predict each graph metric. This model significantly predicted all graph metrics (Figure 1e), and almost all graph metrics in the other datasets (Figure ED1). At the regional level, for each subject, we computed the Spearman correlation between regional TA-Δ_1 and various graph metrics. We found that both weighted and unweighted nodal graph metrics were highly correlated with regional TA-Δ_1 (Figure 1j). Most notably, a node's degree, or the total number of connections a node makes to other regions, was predicted by regional TA-Δ_1 with a median correlation of 0.89 (Figure 1j), and this relationship was not driven by parcel size (median partial correlation 0.83 between degree and regional TA-Δ_1, accounting for parcel size). Remarkably, this implies that network hubs can be discovered without examining the topology of network[31,36].

Multiple sources of temporal autocorrelation Heterogeneity in TA across brain regions can be shaped by multiple underlying factors. One such factor is the region's intrinsic timescale, which measures the rate at which a signal in a region decays over time. Intrinsic timescales vary from region to region and have been previously described as a central factor of brain dynamics in both fMRI and electrophysiology[27,28]. If intrinsic timescale was the primary driver of TA heterogeneity, regions with longer timescales would lead to higher TA. However, brain maps of intrinsic timescale[27,28] do not correspond well to our brain map of regional TA-Δ_1 (Figure 1f). This suggests that other factors may be important in driving TA.

In addition to intrinsic timescale, the noise level of a brain region also influences its TA[37]. Decreasing the SNR or increasing the noise level will decrease TA. More generally, adding white noise to any timeseries will cause its TA to approach zero (see Supplement for proof). Thus, since 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 timescale, 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-Δ_1 show the lowest reliability in regional TA-Δ_1 (Figure 1k). By this logic, TA may influence FC by altering the fraction of shared variance between pairs of regions[38].

In what follows, we used this principle to build a spatiotemporal model for generating surrogate timeseries with regionally heterogeneous noise. We compare this with the "intrinsic timescale with SA" model whereby intrinsic timescale 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 timescale with SA" model, SA is independent of TA by construction. We also compare the spatiotemporal model with models which utilize 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 timeseries We designed a spatiotemporal surrogate timeseries model which, when fit to individual subjects, uses SA and TA to capture individual variation in network topology. The model operates at the level of the parcellated timeseries, meaning that comparisons may be drawn at multiple levels of the analysis pipeline (Figure 1a). Our model can be summarized in a small number of steps (Figure 2a). First, we generate long-memory ($1/f^{\alpha}$ spectrum) timeseries which have uniformly high TA. Concurrently, we spatially embed these timeseries 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 (see Methods). Then, we add uncorrelated white noise with a region-specific variance, thereby lowering TA heterogeneously. The resulting timeseries can then be analyzed the same way as rs-fMRI timeseries. 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 which are fit on the single-subject level, corresponding to noiseless SA-λ (SA-λ^{gen}) and noiseless SA-∞ (SA-∞^{gen}). We also utilize the subject's observed regional TA-Δ_1 to determine the variance of noise to add. We fit SA-λ^{gen} and SA-∞^{gen} through optimization to the distribution of eigenvalues

of the FC matrix, a property which is not a graph metric yet is well-captured by the model (Figure 2b). 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 timeseries. In principle, model fit could be assessed using several different criteria, such as the model's ability to capture as much variance as possible (e.g. using R^2), or alternatively, by its ability to capture individual variation (e.g. using Pearson correlation) (Table S1). Here, we assess model fit using Lin's concordance, a stricter criterion which is 1 only when the model captures both variance and individual variation, and zero or negative when either assessment is poor. Figure 2c demonstrates a schematic comparison between 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 which produces the complex, asymmetric patterns observed in the data (Figure 2d).

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 (Figure 2e). 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 (Figure 2e). The high Lin's concordance corresponds to a close fit to the diagonal of graph metrics plotted between the model and the data (Figure 3). As expected, we also confirmed the model matches the SA and TA of the original subjects (Figure ED5). Furthermore, it reproduces the degree distribution (Figure 2f), and also correlates with several nodal graph metrics on the subject level (Figure 2g).

To test alternative ways that graph metrics might be related to timeseries properties, we fit several additional models, depicted graphically in Figure ED6. First, we test whether both SA and TA are necessary within our model. To do this, we fit the spatiotemporal model with TA set to zero ("SA only" model) as well as the spatiotemporal model with an infinitely small SA-λ and a SA-∞ of zero ("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 timescale in conjunction with SA ("Intrinsic timescale + SA"). This model estimates a pink noise $1/f^{\alpha}$ exponent for each region, as well as SA-λ and SA-∞ for the full brain, and then generates surrogates using correlated spectral sampling (see 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 which matches these two statistics ("Zalesky matching"), previously shown to exhibit more brain-like network topology[30]. Additionally, we tested a model which 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.

None of these alternative models capture network topology (Figure 2e, 3, ED5) or the degree distribution (Figure 2f) as well as the spatiotemporal model. These results are not affected by the use of geodesic instead of Euclidean distance (Figure ED7). Across all other datasets, only the eigensurrogate model on the Yale-TRT dataset showed more explanatory power than the spatiotemporal model (Figure ED8, Table S1). 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 nonlinearity of constructing and analyzing a thresholded graph. Instead, we attempt 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-Δ_1 and degree (Figure 1j)[31,36], and shows why strong TA creates hubs. Degree is determined by the number of correlations the node makes which exceed the binarization threshold (Figure 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 moreso than a high mean. Thus, any process which increases a node's variance should also increase the node's degree. Two temporally autocorrelated timeseries are expected to have a higher variance in their pairwise correlation[31] (Supplement 1.5), and this relationship is reflected in our data (Figure 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[30]. This means that high–TA-Δ_1 nodes in the graph are more likely to share an edge[31,39], creating a clustered network topology.

We confirmed this reasoning using a two-parameter "economical clustering" model[23], which builds graphs directly by probabilistically connecting nodes based on their distance and clustering topology (Supplement 2). This model is known to reproduce several topological features of brain networks[23,40], 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 (Figure ED9). 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.

Spatial autocorrelation 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 timeseries used to produce them. Therefore, we ask 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[5,6], we looked for age-related changes in graph metrics, and show that perturbations in the SA-∞ parameter, but not the SA-λ or global TA-Δ_1 parameters, lead to a change in graph metrics which mirrors that of aging.

We analyzed the Cam-CAN dataset, containing cross-sectional rs-fMRI data from over 800 subjects ranging from age 18 to 90. The relationships between SA, TA, and graph metrics were unchanged despite the large variation in age (Figure ED1), and our surrogate model remained effective in this dataset (Figure ED8). Since motion was highly correlated with age in this dataset (Figure ED2), we performed all analyses using partial correlation controlling for motion. Several weighted and unweighted graph metrics are correlated with age (Figure 4), with global efficiency, var-FC, and kurt-FC showing significant correlations. Age was positively correlated with SA-λ, negatively correlated with SA-∞, and uncorrelated with global TA-Δ_1 (Figure 4a).

Our spatiotemporal model can be used to determine the extent to which SA-λ and SA-∞ govern the change in network topology. Since both SA-λ and SA-∞ 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. Since SA-∞ decreases with age, we perturbed the model by decreasing SA-∞^{gen} to obtain predictions for SA-∞-mediated aging. We found that all of the graph metrics which 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 (Figure 4b). We also perturbed the model by increasing SA-λ^{gen} to obtain predictions for SA-λ-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 (Figure 4b). This suggests that the impact of aging on network structure is mediated by changes in SA-∞ rather than SA-λ. 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.

Since 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-Δ_1 across the brain (Figure 4c). regional TA-Δ_1 decreased with age in the frontal subnetwork, but increased with age in cerebellar regions (Figure 4d), suggesting a difference in how these regions change over age. This difference may reflect age-related structural changes along an anterior-posterior gradient[41]. 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 While 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 criteria 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-λ, SA-∞, and global TA-Δ_1, partial on age and motion. Surprisingly, we found a significant negative partial correlation with SA-λ ($r = -0.184, p < 10^{-5}$), associating wider SA with reduced cognitive function (Figure 4e). There was also a weaker correlation with SA-∞ ($r = 0.118, p = 0.002$), but no significant relationship with global TA-Δ_1 ($r = -0.067, p = 0.09$) (Figure 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-λ and SA-∞ compared to graph metrics. The SA-λ-driven effect also contrasts with that of healthy aging, which was primarily driven by SA-∞. These findings hint at a relationship of clinical symptoms with SA and TA[42].

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 by 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[7] and psilocybin[43]. 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 timepoints after each administration. The same basic relationships between SA, TA, and graph metrics were preserved under the drugs (Figure ED10), despite a small sample size less than 3% that of HCP.

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 timepoints, in regional- and subject-averaged TA (Wilcoxon rank-sum test $p < 10^{-40}$ for all conditions) (Figure 5a). Subject-level drug effects on global TA-Δ_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 (Figure ED10).

Psychedelic effects of LSD and psilocybin are predominantly attributed to agonism of serotonin receptors, in particular the 5-HT_{2A} receptor[7,43]. A common serotonergic mechanism for LSD and psilocybin should produce similar cortex-wide topographies of the change in regional TA-Δ_1. Indeed, we found significant positive correlation among the cortical topographies for both drugs and both timepoints (Figure 5c).

Since both of these 5-HT_{2A} agonists produced a specific topographic pattern of reductions in regional TA-Δ_1, we predicted that a 5-HT_{2A} antagonist would produce the same pattern but as an increase in regional TA-Δ_1 instead of a decrease. In the LSD study, on a third visit, subjects were pre-administered ketanserin, a selective 5-HT_{2A} antagonist, before receiving LSD (LSD+Ket). We found that at the late timepoint, ketanserin strongly attenuated the effect of LSD on regional TA-Δ_1 (Figure 5a). At the early timepoint, there was a cortex-wide increase in regional TA-Δ_1 (Figure 5a). Furthermore, regions with the strongest decrease under LSD showed the strongest increase with pre-administration of ketanserin (Figure 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[44].

We used these experiments to construct a cortical map of overall regional TA-Δ_1 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 vs control contrasts from all participants. We found a single map (Figure 5d, left) explained approximately 50% of the variance of individual subjects for each of the three experiments at both timepoints (Figure 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 (Figure 5d, bottom right), consistent with their correlational structure (Figure 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 spatial and temporal autocorrelation—as parameterized by SA-λ, SA-∞, and TA-Δ_1—are highly reliable properties of rs-fMRI timeseries that correlate with, and are predictive of, network topology.

When we use a model to generate surrogate timeseries with subject-matched SA and TA, the timeseries produce networks that match the subject's graph metrics. The surrogate timeseries 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 pharmacological manipulations with multiple serotonergic drugs modulate TA in a distinct and reliable pattern with a large within-subject effect size. We anticipate comparable relationships with SA and TA will be present in other datasets stemming from a variety of conditions, diseases, and pharmacological states. The high reliability, interpretability, effect size, clinical relevance, and sensitivity to neurobiology make spatial and temporal autocorrelation serious candidates for fMRI-based biomarkers[33,42].

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[45], as well as in non-serotonin neuromodulators such as norepinephrine[12,46], dopamine[47], and acetylcholine[48]. They are also consistent with methodological studies, showing the statistical considerations of SA and TA[10,49] and the ability of null models to form complex networks[30,50]. SA is particularly important for the emerging literature on diverse cortical gradients[47]. Furthermore, SA and TA are influenced by local circuit connectivity and properties such as intrinsic timescale[27,51]. The influence of confounding physiological processes on SA and TA must be better understood[52,53], including the link between TA and signal variability[54], especially related to aging[45]. The hemodynamic response function imposes not only temporal but also spatial filtering to the BOLD signal[55], which could contribute to changes in SA and TA due to age[56] or serotonergic drugs[57]. Since anatomical structure serves as a scaffold for SA and TA[13,14], our test-retest reliability and fingerprinting performance could reflect individual differences in brain morphology[58,59]. Similarly, differences in functional network organization across participants could lead to reliable misalignment between parcel boundaries[60,61], which would contribute to the reliability we observed in regional TA-Δ_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-Δ_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-Δ_1. This suggests that the strongest underlying influences on TA-Δ_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 TR. On 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 preprocessing strategy, must be studied systematically in order 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 lower-dimensional subspaces capture network topology[23,40], and we showed SA and TA align with these dimensions through the economic clustering model. These dimensions may map onto other aspects of network topology, e.g., a high degree backbone and a lattice-like background[39]. 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. While it did not perform as well as our timeseries-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- or vertex-level analyses, but fitting the spatiotemporal model is computationally intractable for parcellations with many nodes. SA-∞ is conceptually similar to mean-FC, which has known links to aging[58,62]. It is also similar to global signal, but it remains reliable even after global signal regression, 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 less 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[32,42] and task-related changes[15,16]. The use of our spatiotemporal surrogate model—and of previously published spatially-[24,63] and temporally-informed[30] 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.

Acknowledgments

We thank Josh Burt for help with plotting; Amber Howell for assistance with HCP data; and Taku Ito and Jean C.C. Vila for helpful discussions. Funding was provided by the European Molecular Biology Organization grant ALTF 712-2021, the Winston Churchill Foundation of the United States, and the Gruber Foundation (MS); NIMH grant K00MH122372 (SN); Swiss National Science Foundation (P2ZHP1\161626) (KHP); Agence Nationale de la Recherche (ANR-20-NEUC-003-01) and MIAI@Grenoble Alpes (ANR-19-P3IA-0003) (SA); Usona Institute (2015-2056), Heffter Research Institute (1-190420), Swiss Neuromatrix Foundation (2015-103 and 2016-0111), and Swiss National Science Foundation under the framework of Neuron Cofund (01EW1908) (FXV); NIHR (Senior Investigator award) and the NIHR Cambridge Biomedical Research Centre (ETB); SFARI (Pilot Award) (AA and JDM); and the National Institute of Mental Health (NIMH) (R01MH112746) (JDM). Data collection and sharing for this project was provided in part by the Cambridge Centre for Ageing and Neuroscience (CamCAN). CamCAN funding was provided by the UK Biotechnology and Biological Sciences Research Council (grant number BB/H008217/1), together with support from the UK Medical Research Council and University of Cambridge, UK.

Author contributions

MS and ETB conceived the research. MS designed the experiments. MS and AH performed the experiments. MS and JDM analyzed and interpreted results. KHP, JLJ, FM, SN, DS, RTC, JHK, FXV, and AA contributed data, methodology, and resources. MS, LT, and SA performed the mathematical analysis. DL, ETB, and JDM provided supervision and funding. MS, AH, and LT wrote the first draft of the manuscript. All authors edited, revised, and approved the manuscript.

Ethics declarations

Competing interests KHP is currently an employee of Hoffmann-La Roche. JHK has consulting agreements (less than US$5,000 per year) with the following: Aptinyx, Inc.; Atai Life Sciences; AstraZeneca Pharmaceuticals; Biogen, Idec, MA; Biomedisyn Corporation; Bionomics, Limited (Australia); Boehringer Ingelheim International; Cadent Therapeutics, Inc.; Clexio Bioscience, Ltd.; COMPASS Pathways, Limited, United Kingdom; Concert Pharmaceuticals, Inc.; Epiodyne, Inc.; EpiVario, Inc.; Greenwich Biosciences, Inc.; Heptares Therapeutics, Limited (UK); Janssen Research & Development; Jazz Pharmaceuticals, Inc.; Otsuka America Pharmaceutical, Inc.; Perception Neuroscience Holdings, Inc.; Spring Care, Inc.; Sunovion Pharmaceuticals, Inc.; Takeda Industries; Taisho Pharmaceutical Co., Ltd. JHK serves on the scientific advisory boards of Biohaven Pharmaceuticals; BioXcel Therapeutics, Inc. (Clinical Advisory Board); Cadent Therapeutics, Inc. (Clinical Advisory Board); Cerevel Therapeutics, LLC; EpiVario, Inc.; Eisai, Inc.; Jazz Pharmaceuticals, Inc.; Lohocla Research Corporation; Novartis Pharmaceuticals Corporation; PsychoGenics, Inc.; Neumora Therapeutics, Inc.; Tempero Bio, Inc.; Terran Biosciences, Inc. JHK is on the board of directors of Freedom Biosciences, Inc. JHK has stock and/or stock options in Biohaven Pharmaceuticals; Sage Pharmaceuticals; Spring Care, Inc.; Biohaven Pharmaceuticals Medical Sciences; EpiVario, Inc.; Neumora Therapeutics, Inc.; Terran Biosciences, Inc.; Tempero Bio, Inc. JHK is editor of Biological Psychiatry with income greater than $10,000. DL is a co-founder of Neurogazer Inc. AA and JDM are co-founders of Manifest Technologies, serve on the technical advisory board of Neumora Therapeutics, and are consultants for Gilgamesh Pharmaceuticals. ETB serves on the scientific advisory board of Sosei Heptares and as a consultant for GlaxoSmithKline. The remaining authors declare no competing interests.

Tables

None

Figure legends/Captions (for main text figures)

Figure 1

Spatial and temporal autocorrelation are important features of rs-fMRI timeseries. (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_1 (bottom). To compute SA-λ and SA-∞, FC values are binned across distances D and the curve SA-∞ + (1 − SA-∞) $\exp(-D/\text{SA-}\lambda)$ is fit to the binned data. To compute global TA-Δ_1, the Pearson correlation is taken between the timeseries and the timeseries shifted by one timepoint. (d) 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-λ, the second SA-∞, and the third global TA-Δ_1, where # indicates p<.01, + indicates p<.05, and - indicates p>.05, by a two-sided bootstrap resampling procedure. Inset scatterplots show correlation across subjects for two example sessions. (e) Correlation across subjects between graph metrics and SA-λ, SA-∞, or global TA-Δ_1. "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_1, averaged across all subjects. (g) Distribution of reliability for each brain region. (h) Reliability of regional TA-Δ_1 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_1 with nodal graph metrics for each subject. (k) The regional TA-Δ_1 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 boxplots indicate the median, first/third quartiles, and range, with outliers hidden for visualization.

Figure 2

A spatiotemporal model captures connectome topology. (a) Schematic of our spatiotemporal modeling framework for surrogate timeseries 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) A schematic demonstrating Lin's concordance, our model fit statistic, on example data. Values close to one indicate both a correlation and a close match in value (gray). Values close to zero either indicate 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. Boxplots show median, first/third quartiles, and range, with outliers hidden for visualization, N=883.

Figure 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 scatterplots depict the relationships summarized by Lin's concordance in Figure 2e for the HCP dataset. Spearman correlation (r_s) and Lin's concordance (Lin) are inset. * indicates Spearman correlation two-tailed p<.05, and ** indicates p<.01.

Figure 4

Spatial autocorrelation 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<.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 10 runs of the model. The direction (positive or negative) of the change with age corresponds to

the data for SA-∞ but not SA-λ. (c) Partial correlation of regional TA-Δ_1 with age is shown for each region, partial on motion. (d) Mean partial correlation of regional TA-Δ_1 with age across brain regions, partial on motion, N=646 subjects. Boxplots indicate the median, first/third quartiles, and outlier-excluded range of partial correlation distribution. ** indicates p<.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-Δ_1 is not. Bottom: Var-FC was the only weighted or unweighted graph metric which was significant after Bonferroni FWER correction. Error bars show 95% CI. ** indicates p<.01, * indicates p<.05, two-sided Wilcoxon signed-rank test with Bonferroni FWER correction.

Figure 5

Pharmacological administration of serotonergic drugs causes changes in TA. (a) Cortical maps showing the mean difference in regional TA-Δ_1 between drug and placebo at early (left) and late (right) timepoints. Purple indicates lower regional TA-Δ_1 in the drug than placebo conditions (LSD N=24, Psilocybin N=23). (b) Change in the regional TA-Δ_1 mean across subjects. Boxplots show the median, first/third quartiles, and outlier-excluded range of the distribution of regional TA-Δ_1 mean across subjects for different regions. * indicates p<.05, ** indicates p<.01, two-sided Wilcoxon sign-rank test. (c) Similarity, measured by Spearman correlation, between "drug minus placebo" cortical maps for each drug and timepoint. * indicates p<.05, ** p<.01, two-tailed permutation test correcting for spatial autocorrelation[11] 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. Boxplots show the median, first/third quartiles, and outlier-excluded range of the distribution. * indicates p<.05, ** p<.01, two-sided Wilcoxon signed-rank test, N=646 subjects.

Main References

1. Buckner, R. L., Krienen, F. M. & Yeo, B. T. T. Opportunities and limitations of intrinsic functional connectivity MRI. *Nature Neuroscience* **16,** 832–837 (June 2013).
2. Smith, S. M. *et al.* Functional connectomics from resting-state fMRI. *Trends in Cognitive Sciences* **17,** 666–682 (Dec. 2013).
3. Fornito, A., Zalesky, A. & Bullmore, E. *Fundamentals of Brain Network Analysis* en. ISBN: 9780124081185 (Academic Press, Mar. 2016).
4. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage* **52,** 1059–1069 (Sept. 2010).
5. Betzel, R. F. *et al.* Changes in structural and functional connectivity among resting-state networks across the human lifespan. *NeuroImage* **102,** 345–357 (Nov. 2014).
6. Damoiseaux, J. S. Effects of aging on functional and structural brain connectivity. *NeuroImage* **160,** 32–40 (Oct. 2017).
7. 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.** doi:10.7554/elife.35082 (Oct. 2018).
8. Luppi, A. I. *et al.* LSD alters dynamic integration and segregation in the human brain. *NeuroImage* **227,** 117653 (Feb. 2021).
9. Murphy, K., Birn, R. M. & Bandettini, P. A. Resting-state fMRI confounds and cleanup. *NeuroImage* **80,** 349–359 (Oct. 2013).
10. Friston, K. J. *et al.* Analysis of fMRI Time-Series Revisited. en. *NeuroImage* **2,** 45–53. ISSN: 1053-8119 (Mar. 1995).
11. Burt, J. B., Helmer, M., Shinn, M., Anticevic, A. & Murray, J. D. Generative modeling of brain maps with spatial autocorrelation. en. *NeuroImage* **220,** 117038. ISSN: 1053-8119 (Oct. 2020).
12. 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. *The Journal of Neuroscience* **38,** 2304–2317 (Jan. 2018).

13. 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: An Interdisciplinary Journal of Nonlinear Science* **27,** 047405 (Apr. 2017).

14. Fallon, J. *et al.* Timescales of spontaneous fMRI fluctuations relate to structural connectivity in the brain. *Network Neuroscience* **4,** 788–806 (Jan. 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. *Cerebral Cortex* **29,** 230–241 (Dec. 2019).

16. Arbabshirani, M. R. *et al.* Autoconnectivity: A new perspective on human brain function. *Journal of Neuroscience Methods* **323,** 68–76 (July 2019).

17. Honey, C. *et al.* Slow Cortical Dynamics and the Accumulation of Information over Long Timescales. *Neuron* **76,** 423–434 (Oct. 2012).

18. Song, H. F., Kennedy, H. & Wang, X.-J. Spatial embedding of structural similarity in the cerebral cortex. *Proceedings of the National Academy of Sciences* **111,** 16580–16585 (Nov. 2014).

19. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536,** 171–178 (July 2016).

20. Noble, S. *et al.* Influences on the Test–Retest Reliability of Functional Connectivity MRI and its Relationship with Behavioral Utility. *Cerebral Cortex* **27,** 5415–5429 (Sept. 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 Neurology* **14.** doi:10.1186/s12883-014-0204-1 (Oct. 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. *Journal of Statistical Mechanics: Theory and Experiment* **2011,** P11018 (Nov. 2011).

23. Vértes, P. E. *et al.* Simple models of human brain functional networks. en. *Proceedings of the National Academy of Sciences* **109,** 5868–5873. ISSN: 0027-8424, 1091-6490 (Apr. 2012).

24. Roberts, J. A. *et al.* The contribution of geometry to the human connectome. *NeuroImage* **124,** 379–393 (Jan. 2016).

25. Cicchetti, D. V. & Sparrow, S. A. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American journal of mental deficiency* **86,** 127–137. ISSN: 0002-9351 (2 Sept. 1981). ppublish.

26. Koo, T. K. & Li, M. Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine* **15,** 155–163 (June 2016).

27. Murray, J. D. *et al.* A hierarchy of intrinsic timescales across primate cortex. *Nature Neuroscience* **17,** 1661–1663 (Nov. 2014).

28. Raut, R. V., Snyder, A. Z. & Raichle, M. E. Hierarchical dynamics as a macroscopic organizing principle of the human brain. *Proceedings of the National Academy of Sciences,* 202003383 (Aug. 2020).

29. Arbabshirani, M. R. *et al.* Impact of autocorrelation on functional connectivity. en. *NeuroImage* **102,** 294–308. ISSN: 1053-8119 (Nov. 2014).

30. Zalesky, A., Fornito, A. & Bullmore, E. On the use of correlation as a measure of network connectivity. en. *NeuroImage* **60,** 2096–2106. ISSN: 1053-8119 (May 2012).

31. Afyouni, S., Smith, S. M. & Nichols, T. E. Effective degrees of freedom of the Pearson's correlation coefficient under autocorrelation. en. *NeuroImage* **199,** 609–625. ISSN: 1053-8119 (Oct. 2019).

32. Shafiei, G. *et al.* Topographic gradients of intrinsic dynamics across neocortex. *eLife* **9.** doi:10.7554/elife.62116 (Dec. 2020).

33. Keitel, A. & Gross, J. Individual Human Brain Areas Can Be Identified from Their Characteristic Spectral Activation Fingerprints. *PLOS Biology* **14** (ed Engel, A. K.) e1002498 (June 2016).

34. Finn, E. S. *et al.* Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature Neuroscience* **18,** 1664–1671 (Oct. 2015).

35. Waller, L. *et al.* Evaluating the replicability, specificity, and generalizability of connectome fingerprints. *NeuroImage* **158,** 371–377 (Sept. 2017).

36. Baria, A. *et al.* Linking human brain local activity fluctuations to structural and functional network architectures. *NeuroImage* **73,** 144–155 (June 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 (Dec. 2019).

38. Cole, M. W., Yang, G. J., Murray, J. D., Repovš, G. & Anticevic, A. Functional connectivity change as shared signal dynamics. *Journal of Neuroscience Methods* **259,** 22–39 (Feb. 2016).

39. Achard, S. A Resilient, Low-Frequency, Small-World Human Brain Functional Network with Highly Connected Association Cortical Hubs. *Journal of Neuroscience* **26,** 63–72 (Jan. 2006).

40. Betzel, R. F. *et al.* Generative models of the human connectome. *NeuroImage* **124,** 1054–1064 (Jan. 2016).

41. Sullivan, E. V. & Pfefferbaum, A. Diffusion tensor imaging and aging. *Neuroscience & Biobehavioral Reviews* **30,** 749–761 (Jan. 2006).

42. Watanabe, T., Rees, G. & Masuda, N. Atypical intrinsic neural timescale in autism. *eLife* **8.** doi:10.7554/elife.42256 (Feb. 2019).

43. Preller, K. H. *et al.* Psilocybin Induces Time-Dependent Changes in Global Functional Connectivity. *Biological Psychiatry* **88,** 197–207 (July 2020).

44. Persson, B., Heykants, J. & Hedner, T. Clinical Pharmacokinetics of Ketanserin. *Clinical Pharmacokinetics* **20,** 263–279 (Apr. 1991).

45. Grady, C. L. & Garrett, D. D. Understanding variability in the BOLD signal and why it matters for aging. *Brain Imaging and Behavior* **8,** 274–283 (Sept. 2013).

46. Zerbi, V. *et al.* Rapid Reconfiguration of the Functional Connectome after Chemogenetic Locus Coeruleus Activation. en. *Neuron* **103,** 702–718.e5. ISSN: 0896-6273 (Aug. 2019).

47. Shafiei, G. *et al.* Dopamine Signaling Modulates the Stability and Integration of Intrinsic Brain Networks. *Cerebral Cortex* **29,** 397–409 (Oct. 2018).

48. Turchi, J. *et al.* The Basal Forebrain Regulates Global Resting-State fMRI Fluctuations. *Neuron* **97,** 940–952.e4 (Feb. 2018).

49. Davey, C. E., Grayden, D. B., Egan, G. F. & Johnston, L. A. Filtering induces correlation in fMRI resting state data. en. *NeuroImage* **64,** 728–740. ISSN: 1053-8119 (Jan. 2013).

50. Cantwell, G. T. *et al.* Thresholding normally distributed data creates complex networks. *Physical Review E* **101.** doi:10.1103/physreve.101.062302 (June 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 (Nov. 2020).

52. Chen, J. E. *et al.* Resting-state "physiological networks". *NeuroImage* **213,** 116707 (June 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 (Sept. 2020).

54. Baracchini, G. *et al.* Inter-regional BOLD signal variability is an organizational feature of functional brain networks. *NeuroImage,* 118149 (May 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 (Feb. 2010).

56. West, K. L. *et al.* BOLD hemodynamic response function changes significantly with healthy aging. *NeuroImage* **188,** 198–207 (Mar. 2019).

57. Cohen, Z., Bonvento, G., Lacombe, P. & Hamel, E. Serotonin in the regulation of brain microcirculation. *Progress in Neurobiology* **50,** 335–362 (Nov. 1996).

58. Geerligs, L., Tsvetanov, K. A., Cam-CAN & Henson, R. N. Challenges in measuring individual differences in functional connectivity using fMRI: The case of healthy aging. *Human Brain Mapping* **38,** 4125–4156 (May 2017).

59. Honey, C. J. *et al.* Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences* **106,** 2035–2040 (Feb. 2009).

60. Braga, R. M. & Buckner, R. L. Parallel Interdigitated Distributed Networks within the Individual Estimated by Intrinsic Functional Connectivity. *Neuron* **95,** 457–471.e5 (July 2017).

61. Kong, R. *et al.* Individual-Specific Areal-Level Parcellations Improve Functional Connectivity Prediction of Behavior. *Cerebral Cortex* **31,** 4477–4500 (May 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 Research: Neuroimaging* **221,** 37–42 (Jan. 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. *Physical Review E* **93,** 042308 (Apr. 2016).

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, Bern, Switzerland.

Datasets We analyzed the following datasets, which comprise a diversity of preprocessing steps and experimental methodologies, including different parcellations, sampling rates, spatial and temporal smoothings, covariate regressions, and noise removal strategies. Preprocessing 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 preprocessing strategies. Since 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[7,20,43,64,65].

Human Connectome Project (HCP) A total of 883 subjects aged 22-37 (459 female) from the Human Connectome Project 1200 subject data release underwent four resting-state scanning sessions spread across two days. Subjects provided informed consent and were compensated for participation. Resting-state scans lasted for 14.4 minutes with a TR of 0.72 s (sampling rate of 1.39 hz). Data were preprocessed with the Human Connectome Project minimal preprocessing 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)[19] with multimodal surface matching based on MSMAll[66], and with 2 mm FWHM surface spatial smoothing constrained to the parcel. The first 100 timepoints were removed to ensure steady state, consistent with prior work[64]. 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).

Human Connectome Project with Global Signal Regression (HCP-GSR) Subjects and scan parameters are identical to the HCP dataset, but global signal regression was added as a preprocessing step before parcellation. Additionally, we did not truncate the first 100 TRs, for a total of 1200 timepoints. 33 additional subjects were excluded due to nonconvergence in the global signal regression pipeline for a total of 850 subjects.

Yale Test-Retest A total of 12 subjects (6 female) aged 27-56 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, two days on each scanner, on days spaced approximately one 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 1.0 hz). Motion correction was applied, and images were spatially smoothed to achieve uniform spatial smoothness of a 2.5mm Gaussian kernel[67,68]. Images were coregistered to a common subject-specific space across days, and subsequently into MNI space, and parcellated using the Shen parcellation[69].

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.

Cambridge Centre for Ageing and Neuroscience (Cam-CAN) A total of 652 subjects aged 18-88 (326 female) were scanned using a TR of 1.97 s (sampling rate 0.508 hz). Subjects provided informed consent and were compensated for participation. We utilized the standard preprocessing pipeline and parcellation provided by the Cam-CAN project[21,70]. In summary, scans underwent motion correction and slice time correction before coregistration to T1 images and normalization to MNI space with DARTEL. We utilized the default AAL parcellation provided by the Cam-CAN project[71]. 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 utilized a double-blind randomized design[7]. Subjects provided informed consent and were compensated for participation. On each of the three days, N=24 participants aged 20-34 (5 female) received one of

the following treatments: (a) placebo (179 mg Mannitol and 1 mg Aerosil, orally) pretreatment followed by placebo (179 mg Mannitol and 1 mg Aerosil, orally) treatment; (b) placebo (179 mg Mannitol and 1 mg Aerosil, orally) pretreatment followed by LSD (100 μg, orally); and Ketanserin (40 mg, orally) pretreatment followed by LSD (100 μg, orally) treatment. Pretreatment was given 60 minutes before the treatment. The "early" resting-state scan occurred 75 min after treatment, and the "late" scan 300 min after treatment. Data were acquired with TR=2500 ms (sampling rate 0.4 hz). 25 subjects were enrolled in the study. One subject was excluded for technical faults in the data for a total of 24 subjects.

We utilized the same preprocessing pipeline described in Ref.[7], which is summarized below. First, the data were subject to the HCP minimal preprocessing 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 (>.008 hz). Several nuisance variables were regressed out: average signal in the ventricles, average signal in deep white matter, motion parameters, and the mean timeseries across gray matter (i.e. 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 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[19]. To ensure scrubbed frames did not impact TA-Δ_1 calculations, TA-Δ_1 was computed only on consecutive segments involving no dropped frames, such that TA-Δ_1 was a weighted average of the TA-Δ_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 utilized a double-blind randomized design[43], where 23 participants aged 20-40 (11 female) were scanned two separate days. Subjects provided informed consent and were compensated for participation. On each day, participants received either placebo (179 mg Mannitol and 1 mg Aerosil) or psilocybin (0.2 mg/kg), administered orally. rs-fMRI was performed at three timepoints following psilocybin administration: "immediate", performed 20 minutes after administration, "early" after 40 minutes, and "late" after 70 minutes. Data were acquired with TR=2430 ms. 24 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[7].

Data preprocessing was identical to that of the LSD dataset described above.

Estimating spatial autocorrelation We quantified spatial autocorrelation by decomposing it into two components: the rate at which correlations decrease exponentially with distance (SA-λ), and the spatially-invariant level of correlation to which it decays (SA-∞). This can be described quantitatively as

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

where x_i is the timeseries for region i, $D_{i,j}$ is the Euclidean distance between regions i and j, and corr is the Pearson correlation. The quantities SA-λ and SA-∞ are constants and do not depend on i or j.

We estimated the parameters SA-λ and SA-∞ 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 (Figure ED7). 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 vs distance curve (Figure 1a). We then found the least square fit of Equation 1 to this curve, optimizing with gradient descent, bounding SA-λ between 0 and 100 and SA-∞ between -1 and 1. While 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-λ and SA-∞.

To account for heteroskedasticity in SA-λ for the Cam-CAN dataset, we analyzed the logarithm of this parameter for this dataset.

Estimating temporal autocorrelation Our primary measure of temporal autocorrelation, TA-Δ_1, is a non-parametric measurement computed by taking the Pearson correlation of neighboring timepoints in the parcellated timeseries, i.e., for a timeseries $x[t]$, we have corr$(x[t], x[t+1])$. This measure is computationally efficient, and can be implemented with only a few lines of code. However, TA-Δ_1 is not comparable across datasets due to differing TR.

For datasets which included motion scrubbing, we computed TA-Δ_1 only on consecutive regions without scrubbed frames. More specifically, we split each timeseries into fragments at the location of the scrubbed frames. Then, we computed TA-Δ_1 on each fragment and took the weighted average. More formally, for each timeseries fragment x_i of length ℓ_i, TA-Δ_1 was computed as

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

Estimating long memory dynamics fMRI timeseries are thought to show characteristics of a long memory process. In other words, time points which are separated by a large lag show a high correlation[72–74]. These processes are often characterized by their autocorrelation function (ACF). For a timeseries x, the autocorrelation function is defined as

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

where corr is the Pearson correlation and k is the lag at which to evaluate the ACF. Long memory processes have an ACF which decays slowly across lags, meaning that relatively high correlations exist even at large lags. In this work, we use TA-Δ_1 as our primary measure of temporal autocorrelation, 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 10 randomly selected subsets for the HCP and Cam-CAN datasets, and 12 subsets for the TRT data. We fit the regression model

$$ACF_x(k) = \beta_0 + \beta_1 \text{TA-}\Delta_1$$

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[75] given by $H = d + 0.5$. Following previous work, we estimated d through a univariate wavelet-based Whittle estimator[76,77]. We restrict our analyses of d to the HCP, TRT, and Cam-CAN datasets, since they did not use motion scrubbing in their preprocessing pipeline. Under specific assumptions, the relationship between d and TA-Δ_1 can be described analytically[78,79]. Consistent with prior work, TA-Δ_1 is representative of the first and higher lag terms of the ACF[**Hassai2012**, 74,78,80].

FC and graph construction We defined the FC matrix $\rho_{i,j}$ as the matrix of Pearson correlation coefficients $\rho_{i,j} = \text{corr}(x_i, x_j)$ between each pairwise combination of regional timeseries x_i and x_j.

We constructed unweighted, undirected graphs from model or data FC matrices using standard techniques[3,81]. We first constructed a spanning tree backbone to ensure connectedness by transforming each element of the FC matrix $\rho_{i,j}$ by $\sqrt{2(1-\rho_{i,j})}$—an operation which turns the correlation similarity measure into a distance[82]—and then applying Kruskal's algorithm to find the minimum spanning tree[83]. 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 intraclass correlation coefficient (ICC). ICC measures the reliability of a particular scalar measure across subjects[37,84]. Let N be the total number of subjects, M be the number of sessions per subject, and γ_n^m be some scalar measure of interest for session m of subject n. Similar to an ANOVA, ICC decomposes the variance across subjects into variance from a common source, σ_γ^2,

and noise, σ_ϵ, by assuming that γ can be decomposed into a subject-specific term γ_n and a noise term $\epsilon_{n,m}$, i.e., $\gamma_n^m = \gamma_n + \epsilon_{n,m}$. Then, the ICC is defined as the fraction of variance explained by the subject-specific term, i.e.,

$$\text{ICC} = \frac{\sigma_\gamma}{\sigma_\gamma + \sigma_\epsilon}.$$

In theory, there are generalizations of ICC which can accommodate additional sources of variance. Here, we use the simplest form, ICC(1,1)[37]. Significance compared to zero and confidence intervals were computed using the formulas from Ref.[85]. Significant differences between ICC values was 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 Ref.[20,34,86]. Let γ_n^m be a vector describing some measure (e.g. regional TA-Δ_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

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

where 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 γ_n^m was equal to the number of regions in the parcellation R. For FC, the length of γ_n^m was $R(R-1)/2$, i.e. the number of distinct elements in the FC matrix.

In the HCP and HCP-GSR datasets, subjects were scanned four times over two days. For computing ICC, we used $M = 4$. For fingerprinting, in order 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 each session had only one correct match of the 1765 sessions in the pool, compared to 1 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 (i.e. $M = 4$ performed six times instead of $M = 24$ performed once), so subjects had 3 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, since only one rs-fMRI session was acquired per subject.

Regional homogeneity Regional homogeneity estimates how similar a voxel's timeseries is to its neighbors[87]. Here, since 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 timeseries within a parcel. More concretely,

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

where $r_{i,j}$ is the rank of the i'th timepoint in the j'th voxel, k is the number of voxels in the parcel, n is the number of timepoints, and T_i is the i'th triangular number.

Lin's concordance Lin's concordance[88] measures how well one variable reproduces another variable. In general, there are multiple ways for variables to be related to each other. On one hand, they can be correlated with each other, meaning that 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 scatterplot. Lin's concordance is defined for two samples x_1 and x_2 as

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

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

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

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 zero 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 $\mathcal{M} = \{m_{i,j} : i < j\}$.

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 only depends 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 which 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 (i.e. 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[4]:

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[89].

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

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

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, i.e.

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

where $s_{i,j}$ indicates the shortest path between nodes i and j and N is the number of nodes[90].

Local efficiency Local efficiency, similar to clustering, is the mean global efficiency on the subgraph of each node's nearest neighbors[90].

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 δ 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[91]

Transitivity The transitivity is the total number of number of 3-way reciprocal connections compared to the total possible number of such connections, i.e.,

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

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 which measures the total number of edges connected to a node, i.e.,

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

Centrality Betweenness centrality is a nodal metric which measures the faction of shortest paths between all pairs of nodes in a network which pass through the given node. It does not have a closed form equation and must be computed using an algorithm[92].

Graph metrics linear model To quantify the impact of SA and TA jointly, we utilized 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 Figure 1e and Figure ED1 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 SA-λ, SA-∞, and global TA-Δ_1 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. Since many HCP subjects are related, we used only 334 unrelated subjects to avoid the relatedness confound in our analysis, meaning there were 167 subjects in the training and test sets.

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

Diverse methods of generating timeseries which exhibit both TA and SA already exist, such as vector autoregressive (VAR) and Ornstein-Uhlenbeck processes. These models directly simulate timeseries in the time domain, introducing TA and SA locally. However, prior research demonstrating powerlaw dynamics in fMRI data[72–74] 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[76].

Instead, we seek to generate power spectra which represent the tradeoff between long-memory dynamics, i.e. filtered $1/f^{\alpha}$ 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 timeseries with a given covariance matrix C. In our case, C is given in Equation 2. The usual method of spatially embedding timeseries 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 timeseries, 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 timeseries 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 timeseries, and let T be the even length of the desired timeseries. Given an $N \times N$ correlation matrix $C_{i,j}$ and a set of power spectra $|X_i[k]|^2$, we seek to generate a set of timeseries $\{x_1[t], \ldots, x_N[t]\}$ such that each timeseries $x_i[t]$ has a specified estimate of the power spectrum $|X_i[k]|^2$ and the correlation matrix of the resulting timeseries has expected value $C_{i,j}$. 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 timeseries.*

1. *Pick the desired power spectra estimates $\{|X_1[k]|^2, \ldots, |X_N[k]|^2\}$, 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*

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

   is positive semi-definite, where sim denotes cosine similarity, i.e.,

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

2. *For each region $n \in \{1, \ldots, N\}$, 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 $a_n^I[T/2] = 0$ for all n. Because T is even, by way of symmetry the frequency-domain coefficients of the Nyquist frequency must be real. Set $a_n^I[0] = 0$ so that the resulting timeseries 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 timeseries $x_n[t]$.*

Each of the resulting timeseries $x_i[t]$ 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 Supplement 1.1.

Relationship between TA-Δ_1 and power spectra In addition to correlated spectral sampling, defining the spatiotemporal model requires us to determine the precise relationship between TA-Δ_1 and the power spectrum, with a specific focus on the case when the original timeseries is mixed with white noise.

Suppose we have a finite timeseries $x[t]$ with discrete Fourier transform $X[k]$. Then, its TA-Δ_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 Supplement 1.2. Importantly, for a fixed length timeseries with a $1/f^{\alpha}$ power spectrum, this formula implies a bijective relationship between α and TA-Δ_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-Δ_1. For a finite timeseries $y[t] = x[t] + w[t]$ 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 the Supplement 1.3. Note that this formula implies that, as the amount (or variance) of white noise is increased, the TA-Δ_1 approaches zero.

Spatiotemporal model The spatiotemporal model generates surrogate timeseries which can be analyzed like rs-fMRI timeseries. It takes two parameters—the noiseless SA-λ (SA-λ^{gen}), and the noiseless SA-∞ (SA-∞^{gen})—and also uses two pieces of information from the data—the TA-Δ_1 ϕ_i from each region i, and the Euclidean distance $D_{i,j}$ between the centroids of each pair of regions i and j.

The model operates in two basic steps. First, we generate random timeseries which are both spatially and temporally autocorrelated. For this first step, all timeseries have uniformly high TA, and SA is determined by the parameters SA-λ and SA-∞. Using correlated spectral sampling, we generate N high-pass filtered (cutoff frequency 0.01 hz, 4th order Butterworth filter) Brownian noise timeseries (frequency spectrum $1/f^2$) of length T such that, for timeseries $x_i[t]$ and $x_j[t]$, we have

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

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 $C_{i,j}$ required by correlated spectral sampling, such that

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

For non-negative SA-∞^{gen}, $C_{i,j}$ will be positive semidefinite.

Second, we reduce the temporal autocorrelation of each timeseries to match that of the original data. Brownian motion timeseries have high TA and adding white noise to a timeseries reduces TA, so we add white noise to each timeseries $x[i]$ until the TA-Δ_1 of the timeseries is equal to the empirical TA-Δ_1 ϕ_i. Thus, we have to choose a distinct amount of noise to add to each timeseries 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

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

where N is the length of the timeseries. The derivation of this formula is provided in Supplement 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 ϕ_i is truncated such that $\phi_i \geq 0.0001$ and $\sigma^2(\phi_i) \geq 0$. Note that SA-λ^{gen} and SA-∞^{gen} are parameters of the underlying process, and differ from the SA-λ and SA-∞ of the generated timeseries due to the addition of noise. The resulting timeseries exhibit a spatial embedding given by the parcel centroid distances $D_{i,j}$ and parameters SA-λ^{gen} and SA-∞^{gen}, and have regional TA-Δ_1 equal to the original timeseries.

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 SA-λ^{gen} and SA-∞^{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

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

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

Intrinsic timescale with SA In our spatiotemporal model, we make sure timeseries have the desired TA-Δ_1 by generating timeseries with uniformly high TA-Δ_1, and then adding different magnitudes of white noise to each timeseries to match TA-Δ_1 to the original timeseries. In this model, we do not add white noise to the timeseries. To match TA-Δ_1 to the original timeseries, we generate timeseries directly with matched TA-Δ_1. This is possible because correlated spectral sampling does not require the timeseries' power spectra to be identical. Thus, we achieve this diversity in TA-Δ_1 by assuming each timeseries has a filtered pink noise ($1/f^{\alpha}$) temporal dynamics. Then, for each timeseries i, we find an exponent α_i such that the high pass filtered $1/f^{\alpha_i}$ spectrum has expected TA-Δ_1 ϕ_i.

The success of this procedure requires us to choose α_i such that the high pass filtered $1/f^{\alpha_i}$ spectrum has TA-Δ_1 equal to ϕ_i. The mapping $\phi_i \rightarrow \alpha_i$ can be determined by numerically inverting the $\alpha_i \rightarrow \phi_i$ mapping implied by Supplement 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 SA-λ^{gen} and SA-∞^{gen} parameters directly from the data, without the need to fit parameters, using correlated spectral sampling. As in the "SA only" model, $\mathbb{E}$(SA-λ) = SA-λ^{gen} and $\mathbb{E}$(SA-∞) = SA-∞^{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 α_i for each region i are chosen such that the resulting TA-Δ_1 is equal to each region's TA-Δ_1 value; and (2) no noise is added to the powerlaw timeseries.

Homogeneous TA-Δ_1 To allow SA and TA to be independently manipulated, we developed a homogeneous variant of the spatiotemporal model which treats TA as a parameter. Rather than use TA-Δ_1 values computed from the original timeseries, this model uses a single fixed value of TA-Δ_1, TA-Δ_1^{gen}, for all regions. For simplicity, we fixed SA-∞^{gen} to be the mean SA-∞ across all networks. Thus, the model takes two parameters: TA-Δ_1^{gen} and SA-λ^{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 (Figure ED6). This tests the null hypothesis that an effect can be explained by its linear dimensionality. Unlike most of the other models we considered, this produces surrogate FC matrices instead of surrogate timeseries.

We first performed an eigendecomposition of the correlation matrix (FC matrix), and then applied the procedure of Ref.[94]. Briefly, 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.

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

Null models We also test several popular null models.

Phase randomization The power spectrum amplitude of individual timeseries is preserved, but the phases are randomized (Figure ED6). This procedure is described in detail in Ref.[95]. A Fourier transform was performed on each region's timeseries. 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 timeseries, contrary to many neuroimaging studies which use the same phase for a given frequency across all timeseries. This is because using the same phase for each timeseries preserves all cross-correlations between timeseries. In practice, this means that the surrogate FC matrix is identical to the original FC matrix, and hence, it produces a identical graph.

Zalesky matching This model matches the mean and the variance of the correlation matrix (Figure ED6), and is described in full in Ref.[30]. Briefly, it matches the first two moments by iteratively computing correlation matrices from timeseries of different durations with different ground-truth correlations. The process continues until the timeseries duration is found which maximally reproduces the mean and variance of the correlation matrix.

Edge reshuffle We preserve the degree of each node while scrambling the edges (Figure ED6). 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 5 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 TA-Δ_1 we do not report on timeseries 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 timeseries. Second, when fitting data using these models, the TA-Δ_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 comparable 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 SA-λ and SA-∞ 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 which operate at the level of the parcellated timeseries. This excluded approaches such as scrambling using a 3D Fourier transform.

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

Instead, the spatiotemporal model's SA-λ^{gen} and SA-∞^{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[97], 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-semidefiniteness 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 timeseries, 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 timeseries, we excluded these subjects from the analysis. Results were qualitatively similar when parameters were fit to the eigenvalue distribution as described above which forced parameters into valid regimes.

Economical clustering model As described previously, the spatiotemporal model produces timeseries for each brain region—graphs can then be constructed by processing these timeseries the same way as subjects' rs-fMRI timeseries. While 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 which construct graphs directly through the iterative addition of nodes or edges[23,40,98]. The economical clustering (EC) model is one such model which is popular for studying brain networks[23,40]. 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 10 networks per combination of parameters, fitting the homogeneous spatiotemporal model to each of these 10 networks. Since 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 Supplement 2.

Spatial correction for brain map similarity To assess the similarity between brain maps, the presence of spatial autocorrelation can induce a high false positive rate. To correct for this, we perform a permutation test using an SA-preserving surrogate method[11]. To compare a target brain map to a reference map, we generate 10000 surrogate brain maps which match the spatial autocorrelation of the reference map. Then, we compute the Spearman correlation between the target map to 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 which 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 which as at least as extreme in absolute value as the Spearman correlation between the two maps.

Singular value decomposition We computed a cortical map of serotonergic modulation using singular value decomposition (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 Σ is a diagonal matrix, and the rows of U and V form orthogonal bases. The diagonal elements of Σ 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 $M^T M = V\Sigma^2 V^T$, so Σ^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 Σ^2 gives the variance explained of each component. But in our case, since M is not centered, variance explained of the first k components is computed as

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

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, Bern, Switzerland, authorized the use of LSD and Psilocybin in humans.

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`. 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.

Code availability We prepared a software package which 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`. Code was implemented using the standard Python stack[99,100] and other libraries[101,102]. Source code was checked for correctness using software verification techniques[103].

Methods References

64. Demirtaş, M. *et al.* Hierarchical Heterogeneity across Human Cortex Shapes Large-Scale Neural Dynamics. *Neuron* **101,** 1181–1194.e13 (Mar. 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. *Network Neuroscience* **2,** 285–302 (Jan. 2018).

66. Robinson, E. C. *et al.* MSM: A new flexible framework for Multimodal Surface Matching. *NeuroImage* **100,** 414–426 (Oct. 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 (Oct. 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 (July 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 (Nov. 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 (Jan. 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 (Jan. 2002).

72. Bullmore, E. *et al.* Colored noise and computational inference in neurophysiological (fMRI) time series analysis: Resampling methods in time and wavelet domains. en. *Human Brain Mapping* **12,** 61–78. ISSN: 1097-0193 (2001).

73. Wagenmakers, E.-J., Farrell, S. & Ratcliff, R. Estimation and interpretation of 1/fα noise in human cognition. *Psychonomic bulletin & review* **11,** 579–615. ISSN: 1069-9384 (Aug. 2004).

74. Maxim, V. *et al.* Fractional Gaussian noise, functional MRI and Alzheimer's disease. en. *NeuroImage* **25,** 141–158. ISSN: 1053-8119 (Mar. 2005).

75. Achard, S., Bassett, D. S., Meyer-Lindenberg, A. & Bullmore, E. Fractal connectivity of long-memory networks. *Physical Review E* **77,** 036104 (Mar. 2008).

76. Sela, R. J. & Hurvich, C. M. Computationally efficient methods for two multivariate fractionally integrated models. *Journal of Time Series Analysis* **30,** 631–651 (Nov. 2009).

77. Achard, S. & Gannaz, I. Multivariate Wavelet Whittle Estimation in Long-range Dependence. *Journal of Time Series Analysis* **37,** 476–512 (Nov. 2015).

78. Lu, Z. *Analysis of stationary and non-stationary long memory processes : estimation, applications and forecast* Theses (École normale supérieure de Cachan - ENS Cachan, June 2009).

79. Hassani, H., Leonenko, N. & Patterson, K. The sample autocorrelation function and the detection of long-memory processes. *Physica A: Statistical Mechanics and its Applications* **391,** 6367–6379 (Dec. 2012).

80. Scheffer, M. *et al.* Early-warning signals for critical transitions. *Nature* **461,** 53–59 (Sept. 2009).

81. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. en. *Nature Reviews Neuroscience* **10,** 186–198. ISSN: 1471-0048 (Mar. 2009).

82. Mantegna, R. Hierarchical structure in financial markets. *The European Physical Journal B* **11,** 193–197 (Sept. 1999).

83. Kruskal, J. B. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society* **7,** 48–48 (Jan. 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. en. *NeuroImage* **142,** 172–187. ISSN: 1053-8119 (Nov. 2016).

85. McGraw, K. O. & Wong, S. P. Forming inferences about some intraclass correlation coefficients. *Psychological 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 (Apr. 2018).

87. Zang, Y., Jiang, T., Lu, Y., He, Y. & Tian, L. Regional homogeneity approach to fMRI data analysis. *NeuroImage* **22,** 394–400 (May 2004).

88. Lin, L. I.-K. A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics* **45,** 255 (Mar. 1989).

89. Newman, M. E. J. Assortative Mixing in Networks. *Physical Review Letters* **89.** doi:10.1103/physrevlett.89.208701 (Oct. 2002).

90. Latora, V. & Marchiori, M. Efficient Behavior of Small-World Networks. *Physical Review Letters* **87.** doi:10.1103/physrevlett.87.198701 (Oct. 2001).

91. Newman, M. E. J. & Girvan, M. Finding and evaluating community structure in networks. *Physical Review E* **69.** doi:10.1103/physreve.69.026113 (Feb. 2004).

92. Brandes, U. A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology* **25,** 163–177 (June 2001).

93. Timmer, J. & König, M. On Generating Power Law Noise. *A&a* **300,** 707–710 (1995).

94. Davies, P. I. & Higham, N. J. Numerically Stable Generation of Correlation Matrices and Their Factors. en. *BIT Numerical Mathematics* **40,** 640–651. ISSN: 1572-9125 (Dec. 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: Nonlinear Phenomena* **58,** 77–94 (Sept. 1992).

96. Maslov, S. & Sneppen, K. Specificity and Stability in Topology of Protein Networks. en. *Science* **296,** 910–913. ISSN: 0036-8075, 1095-9203 (May 2002).

97. Storn, R. & Price, K. Differential Evolution - A Simple and Efficient Heuristic for global Optimization over Continuous Spaces. *Journal of Global Optimization* **11,** 341–359 (1997).

98. Bassett, D. S. & Sporns, O. Network neuroscience. *Nature Neuroscience* **20,** 353–364 (Feb. 2017).

99. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585,** 357–362 (Sept. 2020).

100. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* **17,** 261–272 (Feb. 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** (ed Marinazzo, D.) e113838 (Dec. 2014).

102. Vallat, R. Pingouin: statistics in Python. *Journal of Open Source Software* **3,** 1026 (Nov. 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.) (Springer International Publishing, Cham, 2020), 143–160. ISBN: 978-3-030-41600-3.