Supporting Material and Methods
for

Adaptive reconfiguration of fractal small-world human brain functional networks

Danielle S. Bassett*†‡, Andreas Meyer-Lindenberg†§,
Sophie Achard*, Thomas Duke‡, Edward Bullmore*§

* Brain Mapping Unit, University of Cambridge, Department of Psychiatry, Addenbrooke's Hospital, Cambridge, CB2 2QQ, United Kingdom

† Unit for Systems Neuroscience in Psychiatry, Genes, Cognition and Psychosis Program, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, 20892, United States

‡ Biological and Soft Systems, University of Cambridge, Department of Physics, Cavendish Laboratory, Cambridge, CB3 0HE, United Kingdom

§ To whom correspondence should be addressed. Andreas Meyer-Lindenberg MD, PhD: Tel 301 4969671, Fax 301 4800169, Email andreasm@mail.nih.gov. Ed Bullmore MB, PhD: Tel +44(0)1223 336583/2, Fax +44(0)1223 336581, Email etb23@cam.ac.uk.

1 Materials and Methods

1.1 Sample and MEG Data Acquisition. Twenty-two healthy right-handed volunteers (12 male, 10 female) with mean age $= 31.0 \pm 6.5$ (SD) years were enrolled in the study. Eleven subjects (6 male, 5 female) performed a finger-tapping task, whereas 11 other subjects matched for age ($P = 0.93$, t test) and gender were used to procure resting data. A history of neurological or psychiatric disorders was excluded by clinical interview and examination. All subjects gave informed consent in writing. The study was ethically approved by the National Institute of Mental Health (NIMH) Institutional Review Board and conducted in accordance with National Institutes of Health (NIH) guidelines for research involving participation of human subjects.

Magnetoencephalographic (MEG) data were acquired at the NIMH using a 275 channel CTF MEG system (VSM MedTech Ltd., Coquitlam, BC, Canada), excluding one faulty sensor in the right frontal area (MRF43). Readings were taken at 600 Hz. Low frequency background noise was filtered using the third gradient. All MEG time series were mean subtracted and corrected for line noise using a 0.3-Hz-wide notch filter.

For the finger-tapping task, visual stimuli were presented at 1.2 Hz for 10.24 s by using a custom-built mechanical sensor. The stimulus was a visual flicker (a pulsating dot) that was tracked by a photodiode. Motor responses (taps of the right index finger) were registered using in-house software. This procedure was repeated four times for each subject.

In the resting state, data were acquired using the same MEG system in a single session while subjects remained quietly immobile with eyes open for 30 minutes. For the purposes of comparison to the shorter motor task-related data, four data segments of 10.24 s duration were sampled from the resting time series at equally spaced intervals excluding data acquired in the first two or last two minutes.

1.2 Wavelet Analysis and Correlations. Wavelet analysis, unlike Fourier analysis, allows both time and frequency localization of a signal; see Percival & Walden (1) for a comprehensive text on time series analysis using wavelets and Bullmore *et al.* (2) for a review of applications of wavelets to analysis of human functional MRI (fMRI) data. Wavelet correlation analysis is a linear calculation (1) and therefore does not allow quantization of the similarity be-

tween information at different frequencies and time lags. In order to include nonlinear effects, one has several options including a time delay embedding as described in Section 3.1 (3), a reconstruction of the wavelet transform based on the Gibbs function (4), or a combination of techniques (5).

All time series were decomposed using the maximum overlap discrete wavelet transform (MODWT), with the Daubechies 4 wavelet, because of its ability to decompose arbitrary signal lengths (6, 7) (Cornish, C. The WMTSA wavelet toolkit for MATLAB). Because MODWT is a circularly shifted redundant transform, the number of detail coefficients at all scales was constant (= 6144). Scales 1–6 collectively represented physiological activity in the frequency range 1-75 Hz: scale 1 (corresponding approximately to the classical EEG γ band = 37.5–75 Hz); scale 2 ($\sim \beta$) = 18.7–37.5 Hz; scale 3 ($\sim \alpha$) = 9.4–18.7 Hz; scale 4 ($\sim \theta$) = 4.9–9.6 Hz; scale 5 ($\sim$ high δ) = 2.4–4.8 Hz; scale 6 ($\sim$ low δ) = 1.1–2.2 Hz.

To quantify the strength of association at specific frequencies between MEG signals in different brain regions, we calculated the absolute value of the correlation between wavelet coefficients for each pair of sensors at each scale of the transform (see Appendix in Achard *et al.* (8) for mathematical details). This resulted in a set of 48 symmetric (275×275) wavelet correlation matrices: one for each scale (6), for each experimental run or segment of resting data (4), and in each behavioral state (2), for each subject. The four correlation matrices at a given scale and state were then averaged to give 6 scale-specific matrices for each of the two behavioral states in each subject.

To convert these continuous wavelet correlation matrices to an undirected graph G, we set to zero any correlations with value less than a threshold τ and set to one any correlations greater than τ. This operation transforms each wavelet correlation matrix to a binary adjacency matrix A which can be graphically represented as a network comprising nodes (brain regions) connected by an edge or line if the wavelet correlation between them was greater than τ. Further details on the choice of threshold are provided below.

1.3 Estimation of Small-World Parameters. The key parameters for a small-world analysis of an undirected graph are the degree, the clustering coefficient and the minimum path length.

The average degree $\langle k \rangle$ of each graph was found by summing the edges between nodes throughout the whole brain network and dividing by the total

number of nodes N, i.e., $\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i$ where k_i is the number of other nodes that node v_i is correlated with at magnitude greater than τ. The probability distribution for degree can be empirically estimated from the observed degrees of each node and various possible forms for the degree distribution can be comparatively evaluated in terms of their adequacy to account for the data. Small-world networks may have degree distributions best described by exponential, power law or exponentially-truncated power law degree distributions.

The average minimum path length L between any two nodes in the system was determined using Dijkstra's algorithm (9, 10) which finds the smallest number of edges that must be traversed to define a path between any two nodes.

The clustering coefficient was first introduced by Watts and Strogatz (11) as a metric to determine the inter-connectedness of a node's neighbors. It can be constructed by letting $m(v_i)$ be the number of opposite edges of a node v_i, $t(v_i)$ be the number of potential opposite edges of v_i, defined as $t(v_i) = k_i(k_i - 1)/2$ while the degree k_i of the ith node is greater than 2 (12). The clustering coefficient of node v_i is then defined as $c(v_i) = \frac{m(v_i)}{t(v_i)}$ while the clustering coefficient C of the graph is the average of the clustering coefficients of all nodes v_i, $C = \frac{1}{N} \sum_{i=1}^{N} c(v_i)$.

In order to determine whether the experimental networks have small-world topology, a comparison must be made to random graphs with the same number of nodes and average degree. Random graphs with a Gaussian degree distribution will have clustering coefficients given by $C_{rand} = \frac{\langle k \rangle}{N}$ (13). The path lengths of a random graph are given by $L_{rand} = \frac{\ln N}{\ln (\langle k \rangle)}$ (13). A small-world network will be characterized by an average clustering greater than a random network, and an average path length approximately equivalent to a comparable random graph, i.e. it will have a σ value greater than 1 where σ is defined as $\sigma = \frac{C}{C_{rand}} / \frac{L}{L_{rand}}$ (14). This comparison between experimental and random graphs can only occur in the range where these equations hold true for random graphs, i.e . where $\langle k \rangle \gg \ln N$ (8, 11, 15). One limitation of this comparison is that we have assumed a Gaussian degree distribution in the random graphs while our experimental graphs have a truncated power-law degree distribution. However, a proper method for the generation of a random network with a given degree distribution has not yet been shown (16).

Therefore, we chose our threshold τ using three constraints: (i) the false discovery rate (which controls the expected proportion of false positives among suprathreshold correlations) must be less than 5%; (ii) the average degree must be no smaller than $2 * \ln(N)$ to allow use of classic graph theory to estimate the small-world scalar σ; and (iii) at least 99% of the nodes of the brain must be connected since we were interested in global brain dynamics. Within these constraints, we chose the highest threshold possible to optimize the strength and thus biological plausibility of connections (17). A high threshold reduces the number of false positive edges in the graph and is consistent with the known relative sparsity of anatomical connections in the brain (18–25).

The chosen value of τ varied somewhat from one network to another but was typically $\tau > 0.4$. To exclude the possibility that results were unacceptably affected by the precise choice of τ we also estimated σ in networks thresholded with several values of τ and found evidence at all scales of small-world topology $\sigma > 1$ in networks thresholded by approximately $0.4 \leq \tau \leq 0.8$. Thus, the thresholded networks retain much important information about the original system. However, future work will further analyze weighted networks which have not been thresholded, especially for cases in which the correlation distributions or mean are distinct between the two states being studied.

1.4 Characterization of Network Hubs. Creating graphical depictions of the brain using these threshold values, we can determine which nodes are connected to the largest number of other nodes, i.e. which nodes are "hubs." We define a hub as a node that has a degree larger than the average degree of the network (26, 27). The topology and average spatial location of these hubs in several frequency bands can be seen in Fig. 2, which shows the degree of all nodes as a color distribution in several frequency bands. For hub distributions in both states and all frequency bands, see Fig. 6.

Hubs may be further classified according to the length and number of their connections as provincial, connector, or kinless hubs (26) (see Fig. 7 for a schematic). A provincial hub (P) will have up to 1/6 of its connections outside of a localized neighborhood of radius r around itself. A connector hub (C) has 1/6-1/2 of its connections outside r, whereas kinless hubs (K) have more than 1/2 of their connections outside r.

We note that if r is large enough ($r_{max} \sim 50$ cm, i.e., the diameter of the brain), provincial hubs dominate, whereas as r becomes small, connector and kinless hubs form the majority (see Fig. 4). The radius at which the number of provincial hubs ($\sharp P$) in a network is equal to the number of connector hubs ($\sharp C$), gives a characteristic length scale ζ of the network. For distributions of provincial and connector hubs in both states in all frequency bands, see Fig. 8.

It is also interesting to identify topologically pivotal nodes which might represent bottlenecks in information flow between modular subsystems (28). If a pivotal node is eliminated the subsystems will no longer be able to communicate. Pivotal nodes have high betweenness scores (29–31) and betweenness is defined more formally as $B_v = \sum \frac{1}{(N-1)(N-2)} \frac{g_{ivj}}{g_{ij}}$, where g_{ij} is the number of geodesic paths (i.e., on the network's surface) between nodes i and j, g_{ivj} is the number of paths between i and j passing through node v, and N is the number of nodes (32–36) (see Fig. 9 for betweenness spatial distributions across all frequencies in both states and Fig. 13 for the numerical distributions). The algorithm used to calculate betweenness centrality first computes the matrix $g_{i,j}$ as $g_{i,j} = A^{\alpha_{i,j}}$, where A is the adjacency matrix and α is the matrix of shortest paths found using the Dijkstra algorithm, giving the number of paths between i and j that have a length $\alpha_{i,j}$, i.e. the shortest possible length [the use of the power of the adjacency matrix is discussed in (36)]. The shortest possible distance between i and j, which also goes through node v is given by $d_v(i,j) = \alpha_{i,v} + \alpha_{v,j}$, where $\alpha_{i,v}$ gives the shortest distance between i and v, and $\alpha_{v,j}$ gives the shortest distance between v and j. If $d_v(i,j)$ is equal to $\alpha_{i,j}$, then a path can go through v on its way from i to j while remaining a shortest path (i.e., the Bellman criterion). In this case, g_{ivj} is given by $g_{i,v} * g_{v,j}$ (36) and describes the number of shortest paths between i and j that pass through v. Other algorithms for estimating betweenness centrality are possible (e.g., ref. 36) including those not based on shortest paths (37). Other types of centrality including information centrality, straightness centrality and closeness centrality have been shown to be useful measures of network structure (e.g., ref. 38).

1.5 Synchronizability. We can consider the dynamical consequences of an arbitrary network topology by first defining the diagonal matrix D whose elements are equal to the degree of the respective nodes: $D(i,i) = k_i$, where k_i is the degree of the ith node. We can define the Laplacian matrix of the

6

graph as $L = D - A$ where A is the adjacency matrix, i.e., the thresholded correlation matrix where 1s represent a connection, 0s represent no connection, and the rows and columns represent nodes. Although neural assemblies in the brain are likely to be coupled non-uniformly, results from uniformly coupled systems provide generalized dynamical tendencies of the system.

We suppose each individual node can be coupled to each other node through some equation of motion: $\dot{x}_i = F(x_i) + \phi \sum_{j=1}^{N} L_{ij} H(x_j), \ i = 1, ..., N$, where $\dot{x}_i$ is the activity of the ith node, ϕ is a uniform coupling strength, $H(x)$ is an output function, L_{ij} is the Laplacian matrix, and $\dot{x} = F(x)$ is the motion of an individual node independent of the system. A synchronized state of the system will then be characterized by $\{x_i(t) = x^*(t), \forall \ i\}$. Because the rows of L have zero sum, the smallest eigenvalue λ_1 is zero and, if the network is connected, $\lambda_2 > 0$. In fact, the magnitude of λ_2 is a measure of the connectivity of the graph. Further, because the coupling between these nodes must be large enough to synchronize the least coupled oscillators and small enough to just synchronize the highly coupled oscillators, the synchronization threshold will depend upon λ_2 and λ_N (39). Thus, in addition to connectivity, we can determine the synchronizability of a specific network, defined as $S = \frac{\lambda_2}{\lambda_N}$ where λ_N is the largest eigenvalue of the Laplacian matrix, and λ_2 is the second smallest eigenvalue of the Laplacian. Fully synchronized systems of various types of oscillators have $0.01 < S < 0.2$; systems with $S \sim 0.01$ are close to the transition from global order to disorder (40).

1.6 Intuitive Correlates of Graph Theory Parameters Used. Graph theory parameters can seem confusing and obtuse. However, social networks hold many analogies to graph theory parameters. For example, the average degree $\langle k \rangle$, i.e., number of connections emanating from or entering into a node, can be likened to the number of friends a person has. Person$_1$ is connected to Person$_i$, Person$_j$, etc by the "edges" of friendship, and therefore the "degree" of Person$_1$ is equal to the number of friends they have. Further, the clustering coefficient, C, represents in graph theory the probability that if node$_1$ is connected to node$_i$ and node$_j$, then node$_i$ and node$_j$ are also connected to each other. In the social construct, C would therefore represent the probability that if Person$_1$ has two friends, Person$_i$ and Person$_j$, then Person$_i$ and Person$_j$ are also friends with each other. Similarly, L describes how many nodes must be passed through to get between a randomly assigned start, n_i, and finish, n_f. This is similar to the Six Degrees of Kevin Bacon

in which any actor (i.e. the start node n_i) can be linked through their film roles to Kevin Bacon (i.e., the end node n_f) through at most six links (in graph theory terms, $L = 6$). This effect of being able to condense the entire social network of the film industry into just a few links is known as the "small-world" effect, and can be quantitatively compared to random networks via the parameter σ. The synchronizability, S, represents how likely it is that all nodes will produce the same wave pattern based on the topology of connections. This is similar to asking if your local group of friends is connected closely enough that if you initially all talked about different things at a party, by the end of the evening, you all were only talking about one thing in a single large group rather than in small cliques.

2 Supporting Results

2.1 A Note on the Synchronizability Measure. The spectrum of the eigenvalues of the Laplacian matrix has been used to illuminate dynamical properties and motifs of networks (41–43), including the synchronizability (40, 44, 45). To determine the maximum synchronizability possible in a network (i.e. the threshold above which synchronization can occur) we have used the methodology of the Master Stability Function as proposed by Pecora and Carroll (46, 47) which determines the threshold synchronizability at which a system of coupled oscillators will globally synchronize. This method has been used to show the relative synchronizability of small-world networks with respect to their regular lattice counterparts (32, 39, 48). This methodology has four assumptions: (i) The coupled oscillators (nodes) are all identical, (ii) The same function of the components from each oscillator is used to couple to other oscillators, (iii) The synchronization manifold is an invariant manifold, and (iv) The nodes are coupled in an arbitrary fashion which is well approximated near the synchronous state by a linear operator.

Synchronizability of these systems is found by computing the Laplacian spectrum (32, 40, 49), which is defined by the topology of the network in question (i.e., by using the adjacency matrix). Under these assumptions, it was proven that the ratio of the second Laplacian eigenvalue (θ_2) to the largest Laplacian eigenvalue (θ_{max}), i.e., ($S = \theta_2/\theta_{max}$), was greater than the synchronizability at which the system globally synchronizes ($S > S_{glob} = 1/\beta$) (where β is determined by the Master Stability Function) (40). With these assumptions, several oscillator systems have been studied (Lorenz, Rossler, double scroll, etc.) (40), and a range of S_{glob} has been identified as 0.01–0.20 depending

on the oscillator chosen. Therefore, to be most stringent, a network will be synchronizable if $S > 0.20$; a network is not synchronizable if $S < 0.01$.

The synchronization that we have found for our networks is at or below the lowest S_{glob} known, i.e. $S < 0.01$. To suggest what this may mean, we look again at the four assumptions above.

(i) First, our observables are MEG time series which reflect summary activity of relatively large neural assemblies. Clearly, therefore, a representation by identical oscillators is a simplification. However, because our primary interest is in the scale invariance of synchronizability, an identical oscillator system may be a good first-order approximation, especially since we use wavelet decomposition to restrict the topology to a given band, thereby constraining the main order parameter of an oscillator, and experimental data in our paradigm that could guide a selection of different oscillator parameters were not available.

(ii) This is a simplification in the interest of theoretical tractability that does not hold in the brain in general. However, our networks are very sparse, and we only allow highly correlated sensors to be "connected," i.e., to be present in our adjacency (and therefore Laplacian) matrix, which implies that this approximation may be reasonable, since we study a binary adjacency matrix after thresholding, and it is again not clear how data that would guide a more realistic estimate of coupling strength could be derived in humans in vivo.

(iii) This guarantees the existence of a synchronization hyperplane in the phase space.

(iv) Because we are studying linear correlations between brain areas as functionally significant, this assumption for linear coupling is true to our data.

Therefore, we suggest that these networks cannot globally synchronize no matter how high the coupling (40). This is the first analysis of synchronizability of the MEG sensor system, and, although we have used generalizations, this work can be followed up in the future by studying varied weightings (50), determining possible hierarchies in the topology (51), and determining which nodes are most important to S via edge removal (52).

2.2 Distance Distributions. The categorization of hubs into provincial, connector, and kinless hubs gave strong evidence for a characteristic spatial

length scale of connections dependent on task in several frequency bands. To support this result, we further computed the distribution of spatial distances comprised by all edges in all networks. More specifically, we created a list of the spatial edge lengths for the networks in all four runs for each of the 11 subjects at a particular wavelet level, and concatenated them together. We show the histogram of these concatenated lists in Fig. 10. It is evident that the larger characteristic length scale found in the γ, β, and α bands is robustly present in these distance distributions as well.

2.3 Degree Distribution Fits. We claim in the main article that the degree distributions found in our data in both tasks and all wavelet levels are best fit by an exponentially truncated power law ($P(k) \sim Ak^{\lambda-1}e^{k/k_c}$) with the values of the three parameters given in Fig. 1 and Table 2. Two other fits were tested, including the power law fit ($P(k) \sim Ak^{\alpha}$) and the exponential fit ($P(k) \sim Ae^{\beta}$). To quantify the strength of each fit, Akaike's Information Criterion was calculated and corrected for the small number of data points (a.k.a. AIC) (53). Fig. 5 shows the value of AIC at each wavelet level for all three fits where the AIC of the exponentially truncated power law is lower than the AIC for the power law and exponential fits for all levels, indicating that the exponentially truncated power law provides a better fit to the empirical degree distribution.

2.4 Effect of Spatial Proximity of Sensors on Observed Correlations. The correlation between wave patterns is less problematic in MEG sensors than it is with EEG electrodes. However, it is an issue which can not be perfectly described without fully reconstructing all statistically significant MEG sources and their distance and orientation from the surface. However, as a diagnostic test, we have calculated the correlation between each pair of sensors (all correlation values were placed in a vector "Corr") as well as the Euclidean distance between each pair of sensors (all Euclidean distances were placed in a vector "EucD" in the same order as the "Corr" vector). We then calculated the mutual information between the vector of correlations and the vector of distances, describing the percentage of the correlation matrix which can be predicted by the Euclidean distances between sensors. This was performed for both the resting and tapping data in each of the six wavelet bands, presented in Fig. 11. At the most (in the γ band of the resting data), the spatial proximity of sensors is never likely to accounted for no more than 18% percent of observed correlations.

2.5 Scale Invariance of Network Parameters. It was clear by preliminary inspection of network parameters estimated at different scales of the MODWT that most global topological and dynamical parameters were conserved within narrow bounds over all scales. To define more precisely the scaling regimes for each parameter, we fitted a simple linear regression model for the effect of scale on each of the parameters of interest and tested if it was significantly > 0. If so, we iteratively identified the scale corresponding to largest residuals of the fitted model, removed this scale from the model, and reestimated the model until the effect of scale was not statistically significant. This procedure was used to define the extent of scale invariance, or the scaling regime, for each topological and dynamical parameter in each behavioral state.

We have claimed that the topological parameters which we have studied are conserved within narrow bands over all scales as demonstrated in Fig. 1. The error bars in these plots give confidence intervals of standard errors of the means and are not derived from the subsequent linear regression procedure. Furthermore, the parameters are statistically similar to each other across frequency bands and yet their range is significantly smaller than their possible range. In order to determine the possible range for graphical parameters, we have calculated the clustering coefficient, average path length for a regular ring network as described in (11). From ref. 11, we know that C/C_{reg} and L/L_{reg} may range from 0 to 1, yet in our data C/C_{reg} ranges from approximately 0.65 to 0.75 and L/L_{reg} ranges from approximately 0.31 to 0.38. We have then performed the rewiring technique described in ref. 11 by randomly rewiring a percentage P of nodes and calculating the C and L. For the range of values of C and L, we determined the small-world parameter σ as $\frac{C/C_{rand}}{L/L_{rand}}$. The maximum σ for this system is ~ 38, whereas the minimum is ~ 1. The experimental range between 1.7 and 2.0 is significantly smaller than the possible range between 1.0 and 38.0. The possible range for k was 0 to 274, and the experimental range was 12.2 to 16.9. The synchronizability of a network is defined as λ_2/λ_{max}. The only requirements for these two variables is that $0 < \lambda_2 < \lambda_{max}$. Therefore, S may lie anywhere between 0 and 1, while the experimental values found for S lie between approximately 0.0052 and 0.0102 (see ref. 54 for the effect of topology on synchronizability). The percentage of the possible range taken up by the experimental range for these parameters is 10 percent for C, 7 percent for L, 0.8 percent for σ, 1.7 percent for k, and 0.5 percent for S. Therefore, by describing our bands as

narrow, we mean that the experimental range is 10 percent or less of the possible range of each of these parameters.

2.6 Correlation of Graph Theory Parameters. It has been suggested some graph theory parameters are highly correlated, and therefore if one parameter shows a certain behavior (such as scale invariance), other parameters may as well. For example, the average clustering coefficient C and the average degree k may be correlated (55), and C and L may as well (56). We have looked for inherent correlations between graph theory parameters in our work. We create a list of all the values of C for all four runs and all eleven subjects at a particular frequency and state (44 numbers in all) and calculated the correlation between this list and a similar list for, say, L. The parameters which we studied were τ, k, L, C, σ, ζ, and S. Correlation matrices were calculated for each state and each frequency, making 12 matrices in all. To facilitate visualization, we have depicted in Fig. 12 the parameters as nodes and the thickness and color of connections between them give the value and sign of the requisite correlations. It was found that state and frequency band differentially effect the correlations between these parameters in this experimental context.

3 Historical Context of Brain Complexity

Interest in the complexity of the brain's time-dependent activity such as is measured using EEG and MEG grew alongside the development of nonlinear dynamics and chaos theory. Classic measures used to characterize the complexity of EEG and MEG include the correlation dimension, Lyapunov exponents, and entropy (see ref. 57 for a detailed review). The correlation dimension is a measure of the dimensionality of the space occupied by a set of random points and is calculated by reconstruction of a time series using time-delay embedding (58, 59). The largest Lyapunov exponent as well as the Lyapunov spectrum describe in general the stability of the system's dynamics. The rate of information loss of a system's dynamics is its entropy which is equal to the sum of all positive Lyapunov exponents (60, 61). For use in neuroscience applications, the approximate entropy ($ApEn$) as first developed by Pincus is often applied, assigning to each time series a positive number whose magnitude is proportional to the amount of randomness in the series.

3.1 Nonlinear Complexity in EEG and MEG. Nonlinear structure, such

as is measured by the parameters above, has been shown to be present in EEG/MEG and in the interactions between separate EEG times series (62–64). Further, the effect of aging has been shown to increase the correlation dimension (65). It has been suggested that deeper sleep stages are associated with lower complexity (for example, refs. 66–68). Furthermore, anesthetic depth can be estimated by the magnitude of the correlation dimension (for example, refs. 69–71). Work has been done to show that seizure activity is highly nonlinear (72, 73) and low-dimensional (74, 75), whereas interictal EEG is only slightly nonlinear and high-dimensional (76) and furthermore, that nonlinear measures may be used to detect and predict seizures (for several reviews, see refs. 77–79). Several emotional and psychological states differentially effect the nonlinearity of EEG/MEG signals and connections (see, for example, refs. 80–82). Many forms of dementia have been associated with lower complexity (for two reviews, see refs. 83 and 84), whereas normal cognition, especially when applied to a task, has been associated with higher complexity (for example, see refs. 85–87).

3.2 Neural Complexity vs. Nonlinear Complexity. In all of this previous work, complexity is directly related to how random the system or process of the system is. However, recent definitions of "neural complexity" focus more on the intermediate state between randomness and order (first defined in ref. 88, a review given in ref. 89, and specific applications given in refs. 90 and 91). Neural complexity in this sense expresses the portion of the entropy of the system that can be accounted for by the interactions among its elements (92) and reflects a neural system's capacity to integrate distributed information (93). Beyond time series analysis, graphical networks which show the greatest neural complexity are small-world networks, i.e., between randomness and order. Therefore, the current interest in the small-world topology and the topological synchronizability on the border between order and disorder are graph theoretical correlates to "neural complexity" but not complexity in the sense of the nonlinear measures given above. Instead, the correct graphical correlate to the nonlinear time series measures given above would be a small average path length and small clustering, i.e., a random graph.

3.3 Graph Theory vs. Times Series Analysis. Besides the difference in the inherent meaning of the two definitions of complexity, there exists a very basic difference in the level of analysis in the two cases. Nonlinear

13

characteristics of the brain can be studied in a particular time series or in the interaction between two different time series. However, in order to apply graph theory to the brain, we first take the time series and define some continuous measure of association between any pair of time series (in our case, we have used wavelet correlation). From this continuous measure, we use a threshold to derive a binarized network, or graph to which graph theory methods can be applied. Thus, the dynamics of a system's time series can be significantly different than the dynamics a system is capable of due to its connection topology. Despite this inherent difference in the analysis methods and the information that can be extracted from them, it is possible to use a nonlinear measures such mutual information and the synchronization likelihood as the continuous measure of association from which the graph is finally derived.

1. Percival DB, Walden AT (2000) *Wavelet Methods for Time Series Analysis*, (Cambridge University Press, Cambridge, UK).

2. Bullmore ET, Fadili J, Maxim V, Sendur L, Whitcher B, Suckling J, Brammer MJ, Breakspear M (2004) *NeuroImage* 23:S234-S249.

3. Kantz H, Schreiber T (1997) *Nonlinear Time Series Analysis*, (Cambridge University Press, Cambridge, UK).

4. Scheller B, Zwissler B, Daunderer M, Schneider G, Schwender D, Rentschler I (2006) *Biol Cybern* 95:193-203.

5. Keylock, CJ (2006) *Phys Rev E* 73:036707.

6. Whitcher B, Guttop P, Percival DB (2000) *J Geophys Res* 105:941-962.

7. Whitcher B, Byers SD, Guttorp P, Percival DB (2002) *Water Resour Res* 38:10.1029.

8. Achard S, Salvador R, Whitcher B, Suckling J, Bullmore E (2006) *J Neurosci* 26:63-72.

9. Dijkstra EW (1959) *Numerische Mathematik* 1:269-271.

10. Tenenbaum JB, de Silva V, Langford JC (2000) *Science* 290:2319-2323.

11. Watts DJ, Strogatz SH (1998) *Nature* 393:440-442.

12. Shank T, Wagner D (2004) Technical Report 2004-9, Universitat Karlsruhe Fakultat fur Informatik (Karlsruhe, Germany).

13. Albert R, Barabási AL (2002) *Rev Mod Phys* 74:47-98.

14. Humphries MD, Gurney K, Prescott TJ (2006) *Proc Biol Sci* 273:503-511.

15. Bollabas B (1985) *Random Graphs* (Academic Press, London, UK).

16. Artzy-Randrup Y, Stone L (2005) *Phys Rev E* 72:056708.

17. Schneidman E, Berry MJ, Jr, Segev R, Bialek W (2006) *Nature* 440:1007-1012.

18. Thomson AM, Bannister AP (2003) *Cereb Cortex* 13:5-14.

19. Sjostrom PJ, Turrigiano GG, Nelson SB (2001) *Neuron* 32:1149-1164.

20. Markram H, Lubke J, Frotscher M, Roth A, Sakmann B (1997) *J Physiol* 500:409-440.

21. Markram H, Tsodyks M (1996) *Nature* 382:807-810.

22. Markram H, Lubke J, Frotscher M, Sakmann B (1997) *Science* 275:213-215.

23. Markram H (1997) *Cereb Cortex* 7:523-533.

24. Thomson AM, West DC, Wang Y, Bannister AP (2002) *Cereb Cortex* 12:936-953.

25. Holmgren C, Harkany T, Svennenfors B, Zilberter Y (2003) *J Physiol* 551:139-153.

26. Guimera R, Amaral LAN (2005) *Nature* 433:895-900.

27. Palla G, Derenyi I, Farkas I, Vicsek T (2005) *Nature* 435:814-818.

28. Bressler SL, Tognoli E (2006) *Int J Psychophysiol* 60:139-148.

29. Bavelas A (1948) *Human Organization* 7:16-30.

30. Freeman LC (1977) *Sociometry* 40:35-41.

31. Anthonisse JM (1971) *Technical Report BN 9/71* (Stichting Matehmatisch Centrum, Amsterdam, The Netherlands).

32. Nishikawa T, Motter AE, Lai Y, Hoppensteadt FC (2003) *Phys Rev Lett* 91:014101.

33. Costa LF, Rodrigues FA, Travieso G, Villas Boas PR (2005) *Arxiv (Cond-mat)* 0505185.

34. Newman ME (2001) *Phys Rev E* 64:016132.

35. Hong H, Kim BJ, Choi MY, Park H (2004) *Phys Rev E* 69:067105.

36. Brandes U (2001) *J Math Sociol* 25:163-177.

37. Friedkin NE (1991) *Am J Sociol* 96:1478-1504.

38. Crucitti P, Latora V, Porta S (2006) *Phys Rev E* 73:036125.

39. Motter AE, Zhou CS, Kurths J (2005) *Europhys Lett* 69:334-340.

40. Barahona M, Pecora LM (2002) *Phys Rev Lett* 89:054101.

41. Farkas IJ, Derenyi I, Barabási AL, Vicsek T (2001) *Phys Rev E* 64:026704.

42. Biggs NL (1974) *Algebraic Graph Theory* (Cambridge University Press, Cambridge, UK).

43. Donetti L, Munoz MA (2004) *J Stat Mech* P10012.

44. Chavez M, Hwang DU, Amann A, Hentschel HG, Boccaletti S (2005) *Phys Rev Lett* 94:218701.

45. Donetti L, Hurtado PI, Munoz MA (2005) *Phys Rev Lett* 95:188701.

46. Pecora LM, Carroll TL (1998) *Phys Rev Lett* 80:2109.

47. Fink KS, Johnson G, Carroll T, Mar D, Pecora L (2000) *Phys Rev E* 61:5080.

48. Hasegawa H (2003) *Phys Rev E* 72:056139.

49. Atay F-M, Bigikoglu T (2005) *Phys Rev E* 72:016217.

50. Zhou C, Kurths J (2006) *Phys Rev Lett* 96:164102.

51. Arenas A, Diaz-Guilera A, Perez-Vincente CJ (2006) *Phys Rev Lett* 96:114102.

52. Lu J, Yu K, Chen G, Cheng D (2004) *IEEE Transact Circuits Syst* 51:787-796.

53. Burnham KP, Anderson DR (2004) *Sociolog Methods Res* 33:261-304.

54. Park K, Lai YC, Gupte, S (2006) *Chaos* 16:015105.

55. Eguíluz VM, Chialvo DR, Cecchi GA, Baliki M, Apkarian AV (2005) *Phys Rev Lett* 94:018102.

56. Sporns O, Zwi JD (2004) *Neuroinformatics* 2:145-62.

57. Stam CJ (2005) *Clin Neurophysiol* 116:2266-2301.

58. Grassberger P, Procaccia I (1983) *Physica D* 9:189-208.

59. Grassberger P, Procaccia I (1983) *Phys Rev A* 28:2591-2593.

60. Jin SH, Na SH, Kim SY, Ham BJ, Lee DH, Lee JH, Lee H (2003) *Int J Psychophysiol* 491:1-15.

61. Pincus S (1995) *Chaos* 5:110-117.

62. Breakspear M (2002) *Hum Brain Mapp* 15:175-198.

63. Pritchard WS, Krieble KK, Duke DW (1995) *Psychophysiology* 32:486-491.

64. Stam CJ, Breakspear M, van Cappellen van Walsum AM, van Dijk BW (2003) *Hum Brain Mapp* 19:63-78.

65. Meyer-Lindenberg A (1996) *Electroenceph Clin Neurophysiol* 99:405-411.

66. Roschke J (1992) *Neuropsychobiology* 25:172-176.

67. Ferri R, Parrino L, Smerieri A, Terzano MG, Elia M, Musumeci SA, Pettinato S, Stam CJ (2002) *Int J Psychophysiol* 43:273-286.

68. Shen Y, Olbrich E, Achermann P, Meier PF (2003) *Clin Neurophysiol* 114:199-209.

69. Watt RC, Hameroff SR (1998) *Int J Clin Monitor Comput* 5:3-13.

70. Myles PS, Leslie K, McNeil J, Forbes A, Chan MTV (2004) *Lancet* 363:1757-1763.

71. Bruhn J, Ropkcke H, Rehberg B, Bouillon Th, Hoeft Al (2000) *Anesthesiology* 93:981-985.

72. Andrzejak RG, Widman G, Lehnertz K, Rieke C, David P, Elger CE (2001) *Epilepsy Res* 44:129-140.

73. Ferri R, Elia M, Musumeci SA, Stam CJ (2001) *Clin Neurophysiol* 112:2274-2280.

74. Iasemidis LD, Sackellares JCh, Zaveri HP, Williams WJ (1990) *Brain Topogr* 2:187-201.

75. Frank GW, Lookman T, Nerenberg MAH, Essex C, Lemieux J, Blume W *Physica D* 46:427-438.

76. Pijn JPM, Velis DN, van der Heyden M, DeGoede J, van Veelen WM, Lopes da Silva FH (1997) *Brain Topogr* 9:1-22.

77. Iasemidis LD (2003) *IEEE Trans Biomed Eng* 50:549-558.

78. Le van Quyen M (2005) *C R Biol* 328:187-198.

79. Lehnertz K, Mormann F, Kreuz Th, Andrzejak RG, Rieke Ch, Elger CE (2003) *IEEE Eng Med Biol Mag* 22:57-63.

80. Aftanas LI, Lotova NV, Koshkarov VI, Popov SA (1998) *Biol Psychol* 48:121-138.

81. Lee YJ, Zhu Y, Xu Y, Shen M, Tong S, Thakor NV (2001) *J Med Eng Technol* 25:79-83.

82. Breakspear M, Terry JR, Friston KJ, Harris AWF, Williams LM, Brown K, Brennan J, Gordon E (2003) *Neuroimage* 20:466-478.

83. Jeong J (2002) *Drug Develop Res* 56:57-66.

84. Jeong J (2004) *Clin Neurophysiol* 115:1490-1505.

85. Bizas E, Simos PG, Stam CJ, Arvanitis S, Terzakis D, Micheloyannis S (1999) *Brain Topogr* 33:85-95.

86. Meyer-Lindenberg A (1998) *Brain Topogr* 10:291-299.

87. Anokhin AP, Lutzenberger W, Birbaumer N (1999) *Int J Psychophysiol* 33:259-273.

88. Tononi G, Sporns O, Edelman GM (1994) *Proc Natl Acad Sci USA* 91:5033-5037.

89. Tononi G, Edelman GM, Sporns O (1998) *Trends Cogn Sci* 2:474-484.

90. van Cappellen van Walsum AM, Pijnenburg YAL, Berendse HW, van Dijk BW, Knol DL, Scheltens Ph, Stam CJ (2003) *Clin Neurophysiol* 114:1034-1040.

91. De Lucia M, Bottaccio M, Montuori M, Pietronero L (2005) *Phys Rev E* 71:016114.

92. Sporns O, Tononi G, Edelman GM (2000) *Cereb Cortex* 10:127-141.

93. Tononi G, Sporns O, Edelman GM (1996) *Proc Natl Acad Sci USA* 93:3422-3427.