# Adaptive rewiring based on diffusion balances stability and plasticity in weighted networks while evolving 'brain-like' structure

Ilias Rentzeperis[1*], Cees van Leeuwen[1,2]

[1]University of Leuven, Belgium

[2] University of Technology Kaiserslautern, Germany

[*]Corresponding author:

E-mail: ilias.rentzeperis@gmail.com

## Abstract

Activity dependent plasticity is the brain's mechanism for reshaping neural connections. Representing activity by graph diffusion, we model plasticity as adaptive rewiring. The rewiring involves adding shortcut connections where diffusion on the graph is intensive while pruning underused ones. This process robustly steers initially random networks to high-levels of structural complexity reflecting the global characteristics of brain anatomy: modular or centralized small world topologies, depending on overall diffusion rate. We extend this result, known from binary networks, to weighted ones in order to evaluate the flexibility of their evolved states. Both with normally- and lognormally-distributed weights, networks evolve modular or centralized topologies depending on a single control parameter, the diffusion rate, representing a global homeostatic or normalizing regulation mechanism. Once settled, normally weighted networks lock into their topologies, whereas lognormal ones allow flexible switching between them, tuned by the diffusion rate. For a small range of diffusion rates networks evolve the largest variety of topologies: modular, centralized or intermediate. Weighted networks in the transition range show topological but not weighted rich-club structure matching empirical data in the human brain. The simulation results allow us to propose adaptive rewiring based on diffusion as a parsimonious model for activity-dependent reshaping of the brain's connections.

Abstract: 202 words

Keywords: evolving networks, rewiring, structural connectivity, functional connectivity, flexibility, stability

## Author Summary

The brain is adapting continuously to a changing environment by strengthening or adding new connections and weakening or pruning existing ones. This forms the basis of flexible and adaptable behaviors. On the other hand, uncontrolled changes to the wiring can compromise the stability of the brain as an adaptive system. We used an abstract model to investigate how this basic problem could be addressed from a graph-theoretical perspective. The model adaptively rewires an initially randomly connected network into a more structured one with properties akin to the human brain, such as small worldness and rich club structure. The adaptive changes made to the network follow the heat diffusion, an abstract representation of brain functional connectivity. Moreover, depending on a parameter of the model, the heat diffusion rate, either modular or centralized connectivity patterns emerge, both found across different regions of the brain. For a narrow range of intermediate heat diffusion rates, networks develop a full range from modular to centralized connectivity patterns. Once settled into a connectivity pattern networks with normally distributed weights lock into that state, whereas networks with lognormally distributed weights show greater flexibility to adjust, while maintaining their small-world and rich club properties. Networks with lognormally distributed weights, therefore, show the combination of stability and flexibility needed to address the fundamental requirements of adaptive networks.

## Introduction

From neuronal synapses to white matter tracts (1), brain anatomical networks are characterized by structural properties like small-worldness (2), modularity (3,4) and rich club organization (5). The pervasiveness of these properties raises the question whether they result from a common principle (4,6).

We proposed that these properties are a product of adaptive rewiring (7–11, for a review 12). Adaptive rewiring captures a crucial property of how the brain anatomical network is shaped over time. The mechanisms that shape the network show great variety, as they encompass brain growth, development *and* learning (13 for a review). Yet they are alike in their common dependency on the network's *functional connectivity*, i.e. the statistical dependencies between the nodes' activities (10,14). Adaptive rewiring formalizes this dependency in terms of graph theory, as it promotes adding shortcut links to regions of the network with intense functional connectivity, while pruning underused ones. These dynamical rewirings could be regarded as adaptive network optimization to function (15).

Adaptive rewiring, it turns out, robustly drives random networks to complex architectures matching the signature characteristics of brain anatomy; small world structure with a mixture of modular and centralized topologies. Adaptive rewiring models reach such "brain-like" structure, regardless of the specific choices made in representing the functional connectivity: rewiring works equally for oscillatory (7) or spiking activity (16), or for more abstract representations in terms of graph diffusion (9). Hence, adaptive rewiring may capture the common driving principle leading to such structure in the brain.

The robustness of 'brain-like' network evolution in adaptive rewiring models brings about a version of the well-known stability-plasticity dilemma (17). Adaptive rewiring involves a balancing act between two seemingly contrasting attributes: stability and flexibility. Whereas the robustness of adaptive rewiring warrants the stability of brain-like architectures, brains need also to be flexible enough to adapt to volatile circumstances. What are the conditions that enable flexibility in models that robustly develop 'brain-like' structure? Here we address this issue, which will involve a development of Jarman's and colleagues (9) most abstract, and hence most general, adaptive rewiring model.

## Model

Recently, computational studies have shown that a simple, linear model can reliably predict functional from anatomical connectivity (18). In these studies, random walks represent the traffic through a fixed anatomical network. This allows the amount of traffic to be stochastically approximated in terms of graph diffusion. Jarman and colleagues (9) showed that adaptive rewiring shapes 'brain-like' networks, even when functional connectivity is essentially a matter of random walks on a graph.

An important shortcoming of Jarman's and colleagues (9) adaptive rewiring model is that it uses binary networks to represents brain anatomy. Notwithstanding the popularity of binary graphs in brain network analysis, networks with differentially weighted connections offer a more appropriate representation of the sophisticated data on brain connectivity from contemporary

tract-tracing and other imaging studies (2). Adaptive rewiring of differentially weighted networks does not automatically guarantee robust convergence to small world topologies; doing so may be ineffective since low-weight connections may not impact the network flow, whereas highly weighted ones may resist pruning (19).

Here, we first probe whether graph diffusion can drive the evolution of brain-like connectivity structures in weighted networks. Graph diffusion, we show, allows adaptive rewiring to establish 'brain-like' network structure, at least for two different weight distributions: normal and lognormal. The distribution of presynaptic weights on single neurons has traditionally been modeled as normal, but recent studies favor lognormal distributions, exhibiting a long tail of strong connections (20,21). Our results show that both normal and lognormal weight distributions provide robust evolution to small world structure. Weighted networks also reproduce a key feature observed by Jarman and colleagues (9), involving the crucial character of *diffusion rate,* a parameter that represents in the model a global homeostatic or normalizing regulation mechanism (for a review 22). Diffusion rate is an important control parameter for the model: low diffusion rates give rise to modular structure, and higher ones to centralized ones.

In the intermediate range of diffusion rates there is a transition zone, in which networks show the greatest variability and complexity. After rewiring, both normal and lognormal networks at the transition zone can have modular or centralized structures. For normally distributed weights the network is equally likely to be modular, centralized or anything in between; for lognormal weights, the network are mostly in an intermediate state with both centralized and modular characteristics. In addition, models with both lognormally and normally

distributed weights in the transition zone develop realistic rich-club structures, akin to backbone subnetwork structures in the human brain (23).

We find, moreover, that the transition zone endows adaptive rewiring with the requisite flexibility; i.e., for networks with lognormal but not with normal weight distributions. When the latter settle into either the modular or centralized state, they reach a 'point of no return'. Networks with lognormal weight distributions, by contrast, can cross over from modular to centralized under continued rewiring. Outside of the transition zone, normally weighted networks continue to be in a 'point of no return' mode, whereas lognormally weighted networks are steered to a modular or centralized state (depending on whether the diffusion rate is below or above the transition zone), irrespective of their initial connectivity state. This combination of plasticity and stability, along with tuning by a single parameter, offers a solution to the stability-plasticity dilemma in the context of adaptive rewiring.

## Results

## Small-worldness

Small-worldness (S) is observed for all rewired networks, albeit with non-identical profiles across the system parameters $\tau$ and $p_{random}$ (Fig 1A). S develops faster as a function of $\tau$ for binary networks than for normal and lognormal ones, the latter being the slowest to increase (Fig 1B). For $\tau = \tau_{plateau}$ (different for each of the three weighting regimes), S reaches a maximum value, $S_{max}$, henceforth sustaining it. $S_{max}$ decreases significantly for $p_{rand} > 0.4$. (Fig 1B).

An increase in a network's average clustering coefficient (C) or a decrease in its average path length (L) induces by definition an increase in its small worldness. We probed the dynamics of C and L that drove up the small worldness of the networks. We found that along τ, S shows a strong positive correlation with C ($\rho_{binary}$ = 0.91, $\rho_{normal}$ = 0.96, $\rho_{lognormal}$ = 0.97) and a moderate one with L ($\rho_{binary}$ = 0.42, $\rho_{normal}$ = 0.52, $\rho_{lognormal}$ = 0.31). In the Watts and Strogatz model (24), the authors start with a structured network and parametrically swap ordered connections with random ones. This induces a significant decrease in L but a more moderate one in C, resulting in an increase in the small worldness of the rewired network. Our model is in a way the reverse process, in which rewiring induces a random network configuration to morph into a more structured one, and thus becoming small world due to an increase in C (See S1 Appendix for a more detailed description).

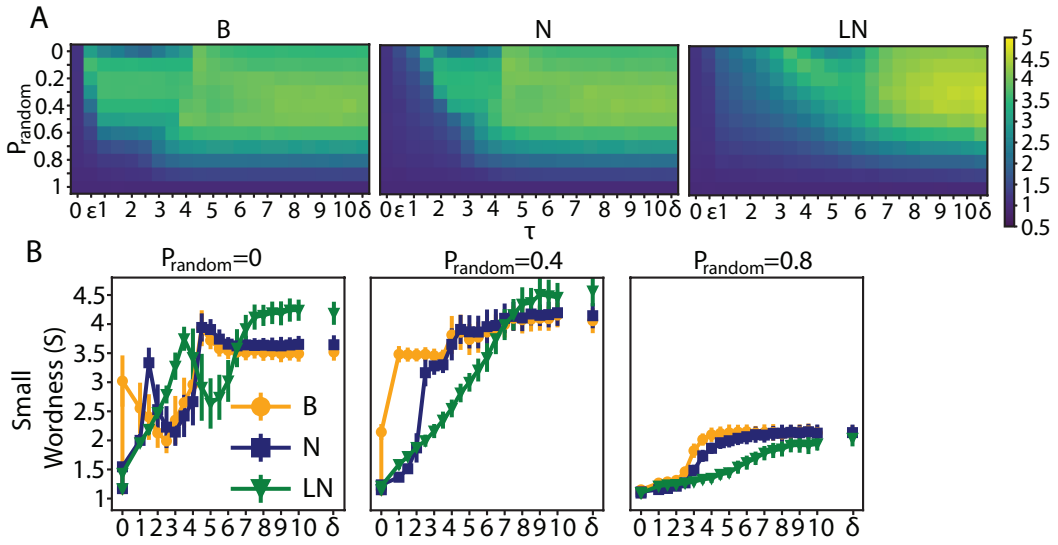


**Fig 1. Networks of binary, normal and lognormal weight distributions develop into small worlds.**

A. From left to right, small world values (S) for networks with binary, normal and lognormal weight distributions for different heat diffusion rates ($\tau$) and random rewiring probabilities ($p_{random}$). $\varepsilon = 10^{-15}$, $\delta = 10^{15}$. B. S as a function of $\tau$ for different $p_{random}$. Horizontal bars indicate standard deviation.

## Modularity

Small worldness characterizes networks with widely diverging topological qualities. For example, both of the adjacency matrices on the left and the right side of Fig 2 have the same S value. However, the ones on the left consist of clusters with dense intragroup and sparse intergroup connections. This characterizes modular connectivity patterns. On the other hand, the networks on the right comprise of a small number of nodes that are heavily connected, acting as hubs with the rest having very few connections. This is a feature of centralized topologies. Newman's (25) modularity index (Q) differentiates between the networks on the left and right side of Fig 2 (Q values for the left vs. right network: Figue 2A, $Q_{binary}$: 0.72 vs. 0.19; Fig 2B, $Q_{normal}$: 0.70 vs. 0.22; Fig 2C, $Q_{lognormal}$: 0.62 vs. 0.14).

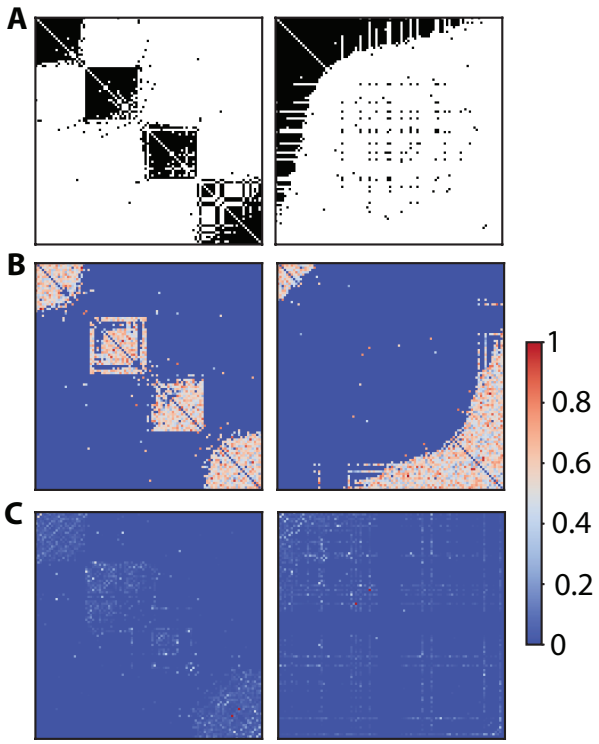


**Fig 2. Networks can have the same small world value but diverging connectivity pattern.** Adjacency matrices of small world networks with binary (A), normal (B), and lognormal (C) weight distributions. The adjacency matrices on the left are examples of modular networks where nodes within clusters are densely interconnected but sparsely interconnected between clusters. The ones on the right are examples of centralized networks where there is a degree imbalance between nodes, with the vast majority of the nodes having very few connections and the rest being heavily connected, acting as hubs. The networks on the left and right have the same small world value (A. S = 3.6; B. S=3.4; C. S = 2.3) despite their different topological characteristics. In all cases $p_{random}$ is 0.2. For A. The values of $\tau$ for the left and right networks are equal to 2 and 5 respectively, for B. 3 and 5, and for C. 4.5, and 7.

Binary, normally and lognormally weighted networks develop modularity as a function of $\tau$ and $p_{random}$ in a similar fashion (Fig 3A). Generally, Q initially increases as a function of $\tau$,

reaching a plateau, but eventually drops off to a value close to that of a random network ($\tau = 0$) (Fig 3B). Binary networks reach their maximum Q faster ($\tau=10^{-15}$) compared to normal ($\tau=2$), and lognormal networks ($\tau = 4$) (Fig 3B). Furthermore, the binary networks sustain the maximum Q value for a greater range of $\tau$ values compared to the two weighted regimes, with the lognormal networks having the smallest range. For all networks, random rewiring diminishes the maximum Q value, with the binary networks being the most resistant and the lognormal ones the most exposed to this effect (Fig 3B).

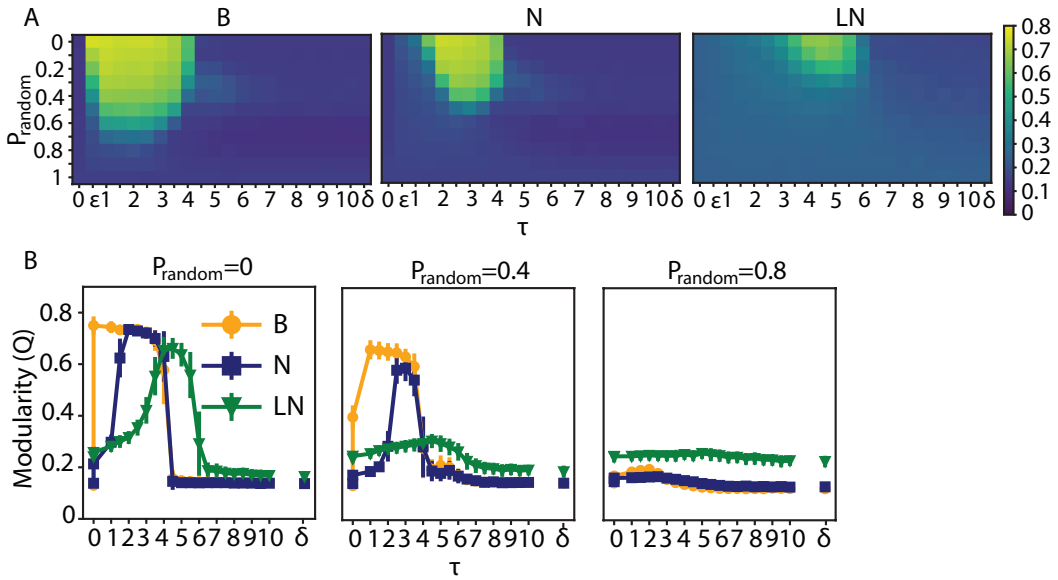


**Fig 3. Modularity profiles between networks with binary, normal and lognormal weight distributions have similar shape.**

A. From left to right, Modularity index Q for networks with binary, normal and lognormal weight distributions for different values of the system parameters diffusion rate $\tau$ and proportion random rewirings $p_{random}$. B. Q as a function of $\tau$ for different $p_{random}$.

## Degree and strength distributions of modular and centralized networks

For a fixed number of total connections, centralized networks can be distinguished from modular ones according to the number of nodes with degrees deviating significantly from the mean. Centralized networks have a small number of high-degree nodes acting as hubs and, correspondingly, a large number of nodes with low degree. Both of these groups constitute outliers in the connectivity distribution. Hence, the proportion of outliers characterizes the centralization in the network. Degrees in the vicinity of the mean were established as $<k> \pm 3\sigma_\kappa$, where $<k>$ is the mean degree of the network and $\sigma_\kappa$ its dispersion. We used the Poisson distribution with mean $<k>$ to calculate the value of the dispersion parameter ($\sigma_\kappa = <k>^{1/2}$). A Poisson distribution is a suitable baseline distribution since random networks by virtue of their construction have most node degrees close to the mean and are well approximated by it. We probed how the proportion of outliers changes with $\tau$, and which range of $\tau$ gives rise to centralized networks. We found that the proportion of outliers in networks as a function of $\tau$ follows a sigmoid function, with binary and normal networks having almost identical values but lognormal ones having a more gradual transition (Fig 4A).

Both the modularity and outlier analyses are in accord with each other: the range of $\tau$ values for which the networks have a small number of outliers includes the $\tau$ values for which the Q value is at its highest (it also includes $\tau$ values close to zero for which the Q value is small where the network is in a state starting to deviate from randomness but not quite reaching a definite structure). For $\tau$ values greater than this modular range, networks qualify as centralized according to our outliers criterion. Modular and centralized networks have different degree

distributions, the former's being concentrated around the mean (Fig 4B) and the latter's having a large spread with a heavy tail (Fig 4C). The $\tau$ and $p_{random}$ parameters giving rise to the distributions in Fig 4B are the same as in the example networks on the left side of Fig 2, for Fig 4C they are on the right side of Fig 2. Furthermore, weighted networks show the same degree and strength distributions (Fig 4B and 4C, inset plots).

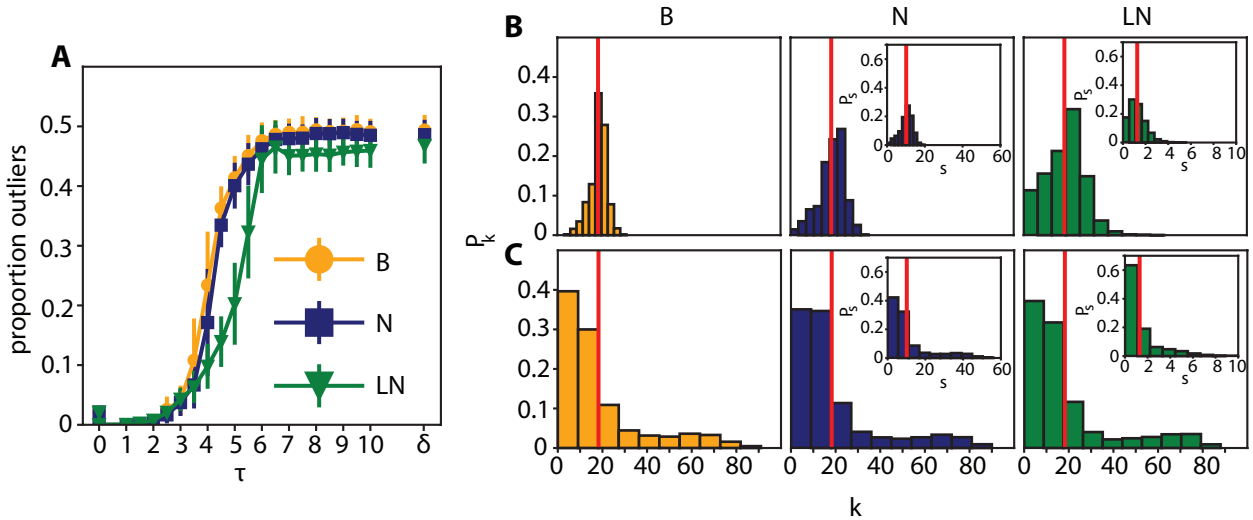


**Fig 4. Degree distribution for modular networks is close to the mean value; for centralized networks it is heavy tailed.**

A. Proportion of nodes with outlier degrees as a function of $\tau$ for binary, normal and lognormal networks. B. From left to right, the distribution of degrees for binary, normal and lognormal networks ($\tau_{binary} = 2$, $\tau_{normal} = 3$, $\tau_{lognormal} = 4.5$) which are modular. Inset plots show the strength distributions. C. Same as in B, but for centralized networks ($\tau_{binary} = 5$, $\tau_{normal} = 5$, $\tau_{lognormal} = 7$). In all cases $p_{random} = 0.2$

We probed the heavy tail degree and strength distributions of representative centralized networks ($\tau_{binary} = 5$, $\tau_{normal} = 5$, $\tau_{lognormal} = 7$, in all cases $p_{random} = 0.2$) The power-law and

lognormal distributions were a better fit than the exponential one (Fig S2). Power law distribution functions are of the form $P(k) \sim k^{-\alpha}$. For the degree distributions the exponent α that fitted best the data was 1.7, for the strength distributions it varied between 2 and 2.6 for the different types of networks (See S2 Appendix for a more detailed description of the analysis). We used publicly available code for this analysis (26).

## Network structure at the transition range

Our previous analyses showed that, for a given random rewiring probability, and depending on the value of the control parameter τ, the network could be rewired to be either in a modular or centralized state. At the boundary between those two states the structure of the network is ambiguous. Moving the control parameter τ across the boundary causes a phase transition from modular to centralized connectivity. Our aim in this section is to probe the properties of the network at this transition.

We estimated the τ value in the middle of the phase transition from modular to centralized networks ($\tau_{transition}$) by locating the point with the largest derivative value on each of the sigmoid curves in Fig 4A (inflection point). The inflection point is similar for both binary ($\tau_{transition} = 4.1$) and normal networks ($\tau_{transition} = 4.15$) but shifts to the right for lognormal ones ($\tau_{transition} = 5.5$). In all cases, $\tau_{transition}$ gives rise to networks with the highest variability of outlier values compared to all the other τ values. In a similar analysis on the modularity data we found the same $\tau_{transition}$ values (finding the highest absolute derivative value for the descending part of the modularity curves and the points with the highest variance).

Regarding the distribution of modularity values at $\tau_{transition}$, one hypothesis is that the distribution is bimodal with one peak centered at the modular region and the other at the distributed one. A contrasting hypothesis is that the distribution is unimodal with the peak in-between the modular and centralized regions. For both binary and normal networks, we found a compromise between those two hypotheses: the modularity distribution at $\tau_{transition}$ is uniform, with the modularity values on the left and right boundaries giving rise to modular and centralized networks respectively (Fig 5A and B, center plots). Furthermore, depending on the sign of small perturbations ($\delta\tau = 0.1$) to $\tau_{transition}$, the resulting modularity distributions show a distinct peak either at the modular region (for $\tau = \tau_{transition} - \delta\tau$) or the centralized one (for $\tau = \tau_{transition} + \delta\tau$) (Fig 5A and B, the leftmost and rightmost plots). On the other hand, the modularity distribution for lognormal networks is more in line with the second hypothesis: it is unimodal, and its peak is in-between the modular and centralized regions. Unlike the other two networks, its peak is not shifted for the small perturbations of its $\tau$ value (Fig 5C).

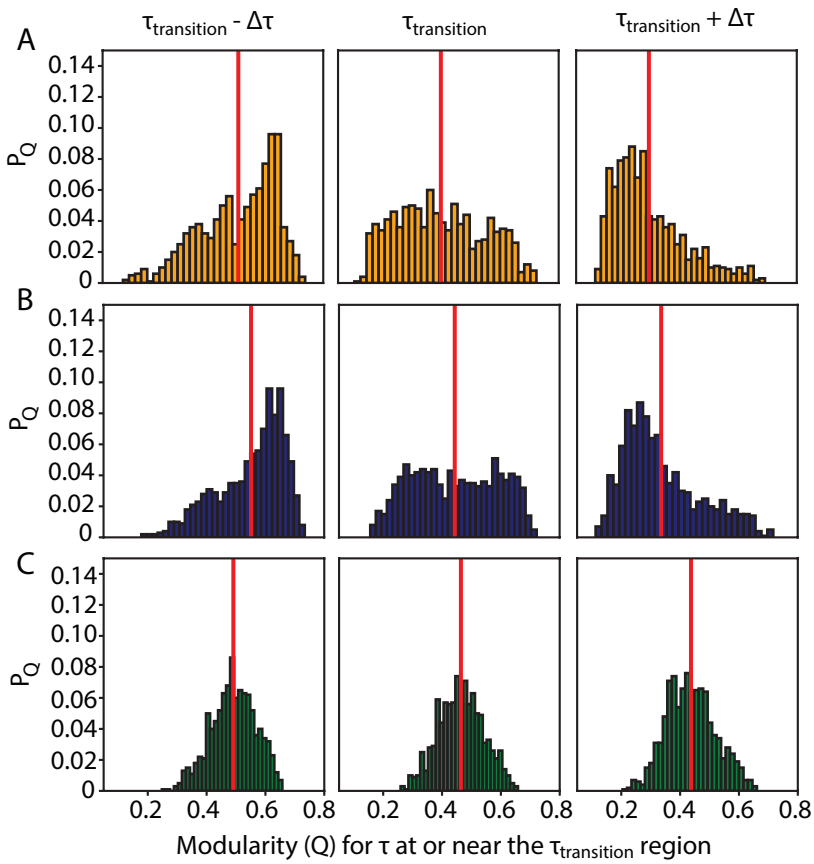


**Fig 5. Binary and normal networks have a uniform modularity distribution and are more prone to connectivity transitions after slight perturbations to $\tau_{transition}$ compared to lognormal ones which have a modularity distribution with a distinct peak.**

A. Binary network. From left to right, modularity distributions for $\tau$ = ($\tau_{transition}$ -$\delta\tau$, $\tau_{transition}$, $\tau_{transition}$ + $\delta\tau$) = (4.0, 4.1, 4.2). For each $\tau$ we obtained 1000 modularity values each corresponding to a different instantiation of the rewiring algorithm B. Same as A for the normal network $\tau$ = (4.05, 4.15, 4.25). C. Same as A for the lognormal network $\tau$ = (5.4, 5.5, 5.6).

Starting from a random configuration, the rewiring process at $\tau_{transition}$ can result in vastly different connectivity patterns ranging from highly modular to highly centralized networks (Fig 5). On the other hand, at the fringes of its variability range, the initial random network is biased

towards a more modular or centralized connectivity pattern. We asked whether small biases in the initial random network are amplified during rewiring, essentially predicting the connectivity pattern of the final network. We tested this hypothesis by comparing the modularity indices of 1000 networks before and after rewiring. We found a very weak correlation between the modulation values of the random and rewired networks ($\rho_{binary} = 0.10$, $\rho_{normal} = 0.14$, $\rho_{lognormal} = 0.08$; Fig 6A). This indicates that the connectivity pattern of the rewired network at $\tau_{transition}$ is nearly independent of the connectivity bias of the initial random network.

To observe whether established network connectivity patterns lock the rewiring at $\tau_{transition}$ in its configuration, we compared the modularity values of the final networks from the previous analysis (4000 rewirings) with the modularity values after 4000 additional rewirings (i.e. after 8000 rewirings in total). For both binary and normal networks we observed very strong correlations ($\rho_{binary} = 0.94$, $\rho_{normal} = 0.89$; Fig 6B) between the modularity values after 4000 and 8000 rewirings, indicating 'a point of no return': that is, once the network settles to a connectivity state, further rewiring no longer steers it away from that state. The correlation for lognormal networks is significantly less than the other two types ($\rho_{lognormal} = 0.59$; Fig 6B). This difference is also observed in the best (in the least squares sense) linear fits of the scatter plots in Fig 6B. The slope of the line for the binary (0.64) and normal networks is greater than the one for the lognormal (0.45) ones (Fig 6B). These results indicate that lognormal networks are more prone to change their connectivity state, most likely due to the small fraction of connections having significantly larger weights compared to the rest. Thus, in the sense that they do not lock to a specific state at $\tau_{transition}$, lognormally weighted networks show considerable flexibility.

Our next analyses probed the flexibility of network developed at $\tau_{transition}$, when $\tau$ is subsequently set to a different value. An initially random network was rewired 4000 times at $\tau_{transition}$; next another 4000 rewirings were subsequently performed at $\tau_{test}$. As long as $\tau_{test}$ deviates from $\tau_{transition}$ only slightly ($\tau_{test} = \tau_{transition} \pm 0.1$), the network's final connectivity pattern remains close to the one obtained after the first 4000 rewirings (more so for normal and binary networks than for lognormal ones; Fig S3). Fig 6C shows that with $\tau_{test} << \tau_{transition}$, all networks are driven to modular connectivity patterns irrespective of their initial topologies. On the other hand, for $\tau_{test} >> \tau_{transition}$, network evolution is strongly affected by the state originally reached at $\tau_{transition}$, at least for the binary and normal networks (Fig 6D). For lognormal networks, however, the correlation coefficient is low, indicating that the explanatory strength of the linear fit is weak.

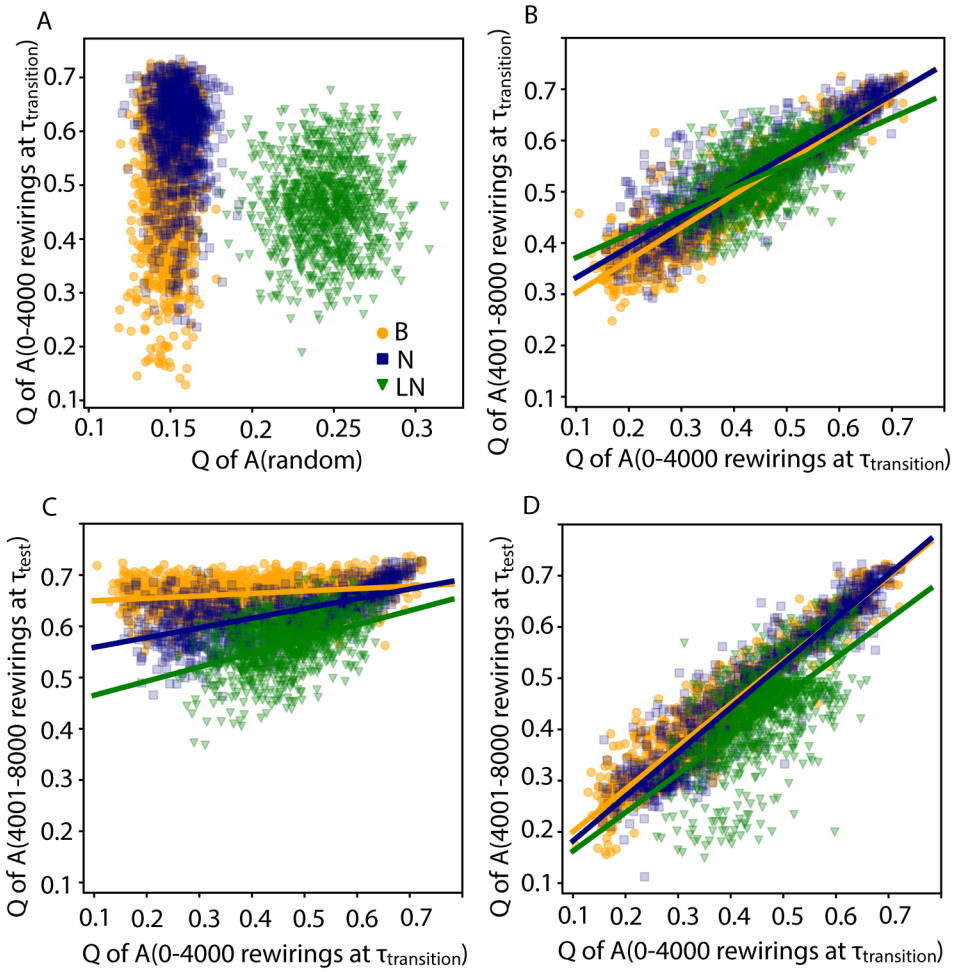

**Fig 6. Flexibility and locking-in of networks vary depending on the weight distribution of their edges.**

A. Scatter plot showing very weak correlations between the modularity values Q of networks after 4000 rewirings at $\tau_{transition}$ plotted against those of their initial random configuration. B Scatter plot showing a strong positive correlation of the modularity values between same networks after 4000 and 8000 rewirings at $\tau_{transition}$. C. Same as (B) but now between 4000 and 8000 rewirings the $\tau$ values equal ones producing modular topologies from random initial configurations ($\tau_{binary}$ =2, $\tau_{normal}$ = 3, $\tau_{lognormal}$ = 4.5) D. Same as (C), but with $\tau$ values equal to ones producing centralized topologies from random initial configurations ($\tau_{binary}$ =5, $\tau_{normal}$ = 5, $\tau_{lognormal}$ = 7). The lines in all cases indicate the best linear fits to the data.

For small $\tau_{test}$ values ($\tau_{test} < \tau_{transition}$) all types of networks show similar behavior; the greater the difference between $\tau_{test}$ and $\tau_{transition}$ the less dependent the rewired network's topology under $\tau_{test}$ from its initial topology (Fig 7A, correlation values for $\tau_{test} - \tau_{transition} < 0$). The same is true for $\tau_{test} - \tau_{transition} > 0$ but only for lognormal networks: as $\tau_{test}$ grows further from $\tau_{transition}$ the correlation in their topologies (modularity) decreases. Therefore, for lognormal weights, as the diffusion rate strays away from $\tau_{transition}$ in either direction, a network's topology after rewiring becomes less dependent on its initial connectivity state. However, this is not true for binary and normal networks for larger $\tau_{test}$ values ($\tau_{test} - \tau_{transition} > 0$): the correlation value does not diminish as the difference between $\tau_{test}$ and $\tau_{transition}$ increases (Fig 7A), indicating that for large diffusion values the topology of rewired normal and binary networks depends on their initial configuration ($Q_{transition}$).

We found the best linear fits for the modularity data produced for each pair of $\tau_{test}$ and $\tau_{transition}$ parameters (Fig 7B and 7C). More specifically, the linear fit takes the form $Q_{test} = \alpha Q_{transition} + \beta$, where $\alpha$ and $\beta$ are two scalar terms, the slope and the bias that we adjust to find the best fit of the data ($Q_{test}$, $Q_{transition}$) in the least squares error sense. The linear fits further cemented our findings from the correlation analysis. For small $\tau_{test}$ values ($\tau_{test} < \tau_{transition}$) the best fit parameters show the same trend for all types of networks: the further away $\tau_{test}$ is from $\tau_{transition}$ the smaller the slope and the greater the bias. This indicates that for $\tau_{test} << \tau_{transition}$ the rewired network will settle into a topological state that depends more on the bias term found from the fit (in the modular regime) than from its initial topological configuration (Fig 7B and 7C). For large $\tau_{test}$ values ($\tau_{test} > \tau_{transition}$), the binary and normal networks sustain both the slope and bias values that the network has when $\tau_{test} = \tau_{transition}$, but the lognormal networks do not, showing a tendency to decrease in slope and increase in bias as $\tau_{test}$ increases, an approximate mirror image of their behavior for small $\tau_{test}$ values (Fig 7B and 7C). Thus for $\tau_{test} > \tau_{transition}$, binary and normal networks lock in their initial connectivity pattern similarly to when $\tau_{test} = \tau_{transition}$, whereas the lognormal network shows a more flexible behavior with a bias nevertheless to switch into a more centralized state.

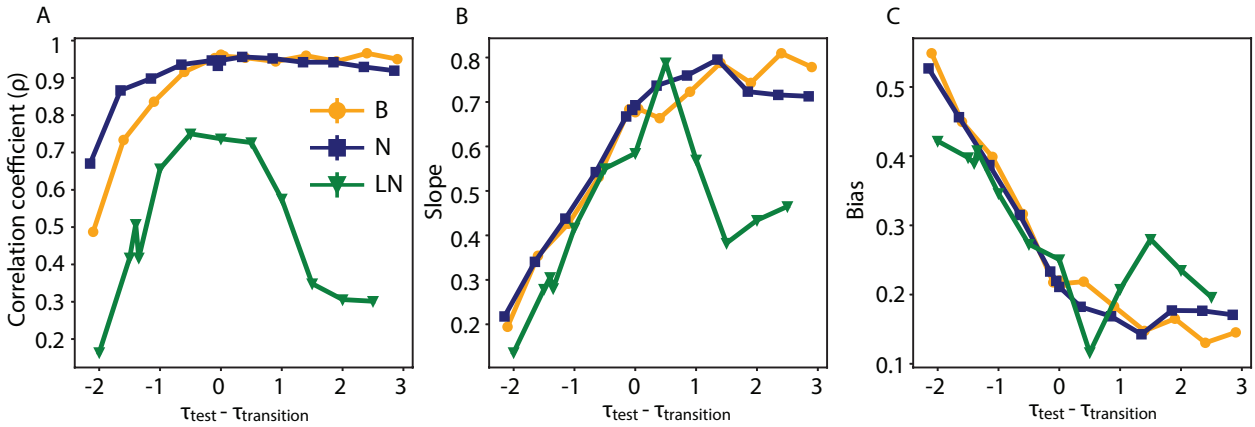


**Fig 7. Lognormal networks show different flexibility patterns compared to normal and binary networks, especially for $\tau_{test} > \tau_{transition}$**

A. Correlation coefficient of the modularity index between the same network that was initially rewired at $\tau_{transition}$ (4000 rewirings) and then re-rewired at different $\tau_{test}$ values (an additional 4000 times). Each correlation point is for 100 iterations. C. and D. The slope (C) and the bias (D) of the best-fitted line that explains the data obtained from A.

## Assortativity and rich club structure

We probed the possibility that the rewiring process favors homophily by measuring the topological assortativity coefficient, r. Networks with positive r are assortative, meaning that nodes with similar degrees tend to connect. Networks with negative r are dissassortative; in this case nodes with dissimilar degrees are more prone to connect. We found that for modular networks, r shows weak or zero assortativity; centralized networks dip into the dissassortative realm (Fig S4)

We further measured the rich club coefficient, $\Phi(k)$, of the networks. Topological rich club refers to the tendency of typically high degree nodes to connect with each other; when its normalized counterpart, $\Phi_{norm}(k)$, is above the baseline (greater than 1), then the subset of nodes with degree greater than k is more densely interconnected compared to a random control with the same degree distribution. The coefficient is not trivially connected to assortativity, since a dissasortative network could still be rich club and vice versa (27). We found that in the modular and transition range the weighted networks exhibited topological rich club behavior for a range of intermediate degrees (binary networks were not rich club for $\tau_{transition}$), whilst centralized networks were not rich club for any degree (Fig S5).

For normal and lognormal networks, we evaluated the normalized weighted rich club coefficient, $\Phi_{w,norm}(k)$. $\Phi_{w,norm}(k)$ above the baseline indicates that the edges of the rich club nodes have larger weights compared to a random control network (a network with the same topology but randomly reshuffled weights). We found that for the lognormal networks $\Phi_{w,norm}(k)$ did not deviate from 1; for the normal networks, $\Phi_{w,norm}(k)$ similarly hovered around 1 in the modular and centralized regimes. Hence for these states the rewiring process does not distribute the larger weights preferentially to the high degree nodes. In general the data show that there is a distinction between the topological and weighted measures; the former being above baseline for a range of degrees, but not the latter.

This distinction becomes more accentuated for normal networks in the transition zone, where $\Phi_{w,norm}(k)$ tilts below baseline for nodes with larger degrees ($k>20$) (Fig 8). Taken together, these topological and weighted coefficient profiles bode well with anatomical data on

the human brain. Specifically, van den Heuvel and colleagues (23) mapped human brain structural connectivity by estimating via diffusion tensor imaging (DTI) the number of connecting white matter streamlines between 1,170 subdivisions of the cortex. Based on topology the human network exhibits rich club behavior, but its weighted counterpart does not (from 28 Fig 3b and 3c), exhibiting the same qualitative properties as the normal –and the lognormal to a lesser extent- network at $\tau_{critical}$.

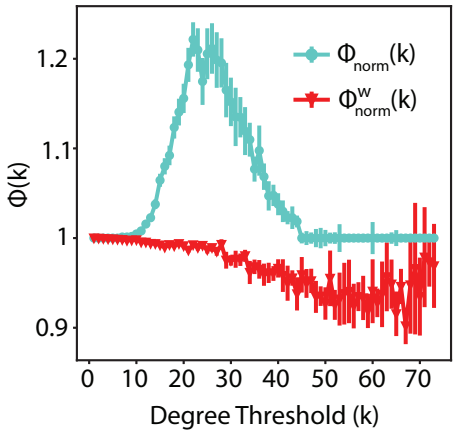


**Fig 8. Topological and weighted rich club coefficients diverge similarly to physiological data.** Topological and weighted rich club coefficient for the normal network at $\tau_{transition}$ = 4.15, $p_{random}$ = 0.2

We performed a fine-grained analysis at $\tau_{transition}$ to probe the rich club characteristics of networks with modular, centralized and in-between connectivity patterns. From the analysis data in Fig 5, we used the adjacency matrices corresponding to the 50 smallest modularity values to obtain the rich club measures for the centralized connectivity patterns, the 50 adjacency matrices corresponding to the largest modularity values for the modular patterns, and the 50 matrices corresponding to the modularity values between the mean one for the in-between patterns. The weighted networks showed rich club characteristics in all three cases (Fig 9A). Thus the

difference between the centralized networks at $\tau_{transition}$ and at $\tau > \tau_{transition}$ is that in the former case the network is rich club (for 20<k<40), in the latter it is not. This effect is more accentuated for the lognormal network. Visual inspection of the adjacency matrices of the different connectivity patterns at $\tau_{transition}$ indicate a smooth transition from modularity to centrality (Fig 9B and 9C).

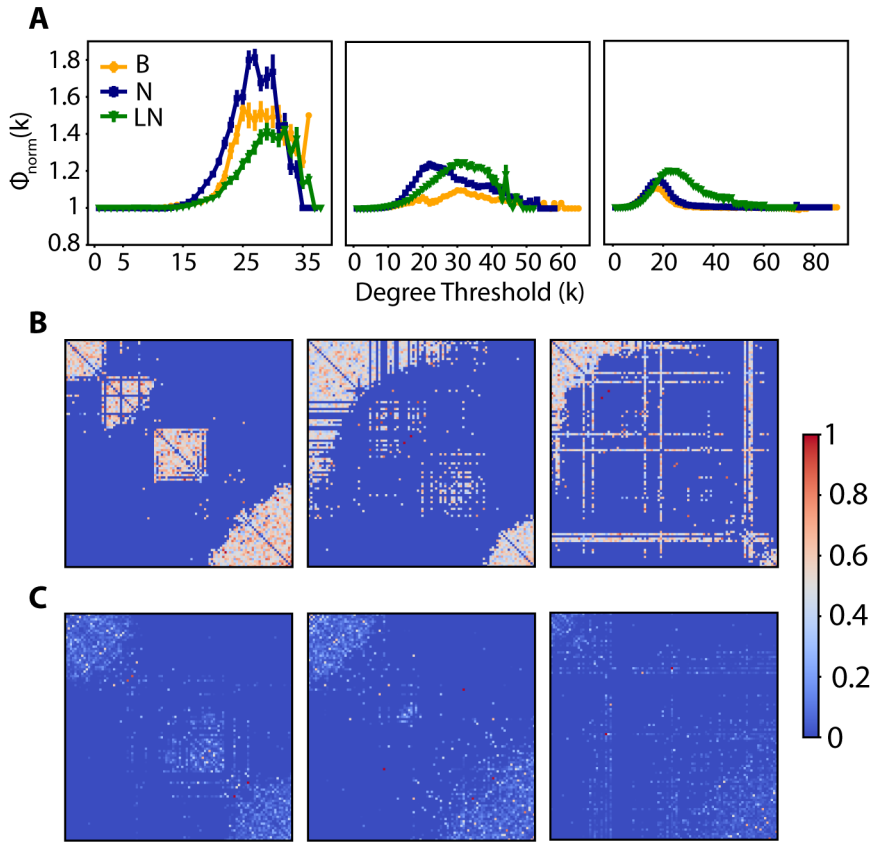


**Fig 9. Weighted networks are rich club for all the topological subdivision at $\tau_{transition}$.**

A. Topological normalized rich club, $\Phi_{norm}(k)$, for binary, normal and lognormal networks in the modular (left), centralized (right), and in between (center) parts of $\tau_{transition}$. B. Example adjacency matrices of normal networks in the modular (left), centralized (right), and in between (center) parts of $\tau_{transition}$. C. Same as in B but for lognormal networks.

## Discussion

## Graph diffusion and other rewiring models

We showed how in graph-theoretical terms, the stability-plasticity dilemma can be addressed in adaptive rewiring of weighted model networks. In these models, rewirings are made to bolster the functional connectivity. In the present model, functional connectivity was represented by heat diffusion. In previous adaptive rewiring models, functional connectivity was represented as synchronous activity (7,8,10,11,16,19,29,30) of linearly coupled nonlinear oscillators (31) or, somewhat more realistically, by spiking model neuron synchrony (16,32).

While these representations more closely resemble artificial neural networks than the current graph diffusion model, their shortcomings are twofold. They rely on arbitrary functional connectivity dynamics, whereas a representation in terms of diffusion is empirically adequate (18). Moreover, rewiring in the older models involved a global search for the best (i.e. most synchronized) rewiring target. Even though this problem was remedied in Jarman and colleagues (29) by taking rewiring distance into account, this introduced arbitrary assumptions about the spatial embedding of these networks. Jarman and colleagues (9) offers a graph-based solution to this problem, based on the rate of heat diffusion relative to that of rewiring. The lower the diffusion rate, the narrower the network search for the best rewiring candidate; higher diffusion rates allow the search to be incrementally broader. Diffusion rate offers a natural control parameter, to determine whether the resulting network structure tends to be modular or

centralized. The diffusion rate parameter may represent global homeostatic forms of plasticity regulating the network's dynamics (for a review 22).

## Graph diffusion and small worldness

For a wide range of species from C. elegans to humans, neuronal connections have been shown to form small world networks, with small path length and high clustering coefficient (2,24,33). Adaptive rewiring based on diffusion when driven by the network's own functional connectivity, leads to a similar small world structure. Networks are evolved to be small world for a wide range of diffusion rates and random rewiring probabilities (Fig 1). This is by and large because diffusion shapes the initial random network into a more structured connectivity pattern with high clustering coefficient while at the same time it maintains small path lengths (Fig S1)

Adaptive rewiring based on diffusion is an abstraction of the consequences of Hebbian learning, where connections with heavy traffic are strengthened whereas the ones with less activity are weakened or pruned. A similar process may lead to changes in the connectivity pattern of higher cognitive regions throughout childhood and adolescence. For example, using resting state functional connectivity MRI on children and adults Fair and colleagues showed that with age short-range functional connections between regions that are involved in top-down control regress whereas some long-range functional connections develop or increase in strength (34). This process leads the connectome of the regions involved mostly in higher cognitive functions to have a small world structure (35).

## Diffusion rate and network topology

Even though small-worldness is a key property of biological organs; many systems with similar small world value have diverging topologies perhaps because of differentially weighted trade-offs between multiple constraints (36). Diffusion-based models show this variety across the range of the diffusion rate parameter $\tau$. Across a wide range of diffusion rates, adaptive rewiring leads to approximately the same small world values, but different topological structures (Fig 2). For smaller $\tau$ values the emerging network is modular (Fig 3), with dense connectivity within clusters but sparse between them. For larger $\tau$ values, we obtain centralized topologies where a subset of nodes are acting as hubs (Fig 4C). Both of these patterns are present in the brain. Clustered neuronal modules facilitate local computations, an example being cortical minicolumns, the highly structured functional patterns in the sensory cortex. Neurons within a minicolumn and typically across the layers are densely connected, but tangential ones, across the layers to other minicolumns are sparsely connected (37). On the other hand, centralized connectivity patterns resemble brain modules that receive and integrate a distributed set of neural inputs, for example the different association areas in the cortex (38,39).

In the centralized regime, all types of networks show a similar degree distribution that is approximated by a power law distribution with an exponent of 1.7, for both binary and weighted networks (Fig S2). Diffusion tractography (40) and fMRI methods (41) have shown a degree distribution between cortical and subcortical regions that follow an exponentially truncated power law with exponents of 1.66 and 1.80 respectively. Brain functional connectivity inferred from fMRI in different tasks shows a scale free distribution with a power law exponent of 2 (42).

## Rich club structure in the transition zone

Brain networks at different levels and for different species have been shown to be topologically rich club (23,43,44), that is, regions or neurons acting as hubs are also preferentially connected among themselves. We found that the nodes of both the normal and lognormal networks show topological rich club behavior for small up to transition diffusion rates. Their corresponding weighted rich club index, which measures to what extent larger weights are attached to the connections within the rich club, was either at or below baseline (Fig 8). This interesting contrast between topological and weighted rich clubs is in line with physiological data on the human brain (23,28).

Furthermore, the centralized structures in the transition zone showed rich club behavior (Fig 9), but the centralized networks for larger $\tau$ values did not (Fig S5). The combination of centrality and rich club behavior is a signature of connectivity patterns in the brain (5). Rich club connections in the brain could constitute its backbone, communicating information from diverse functional modules. Empirical evidence has indicated that rich club connections extend over long distances forming a high capacity structure that facilitates global communication (23).

## Flexibility and stability of topology in and out of the transition zone

In the transition zone between modular and centralized structures, we observe the highest variability in network topology. For example, in contrast to the $\tau$ values producing robustly either modular or centralized topologies-when starting from a randomly connected one-, at the center of the boundary between them, $\tau_{transition}$, a network's connectivity pattern cannot be

predefined before rewiring since it is equally likely that it will have either topology and anything in-between in a continuum (Fig 5). Whereas this transition zone is similar in range for networks with binary and normally distributed connection weights, it covers a considerably broader parameter range for networks with lognormally distributed weights (Fig 4A).

Very small connectivity biases in the initial random network do not influence the final topology of the rewired network (Fig 6A), however, larger biases lock the network at the connectivity state it already is (Fig 6B). This 'point of no return' is not observed for networks with lognormal weights. Moreover, lognormal networks show a more controllable behavior as the diffusion rate moves away from the transition range in either direction: for diffusion values smaller than the transition ones networks will most likely evolve to become modular and for greater values they develop into a centralized topology irrespective of their initial connectivity pattern (Fig 7).

Throughout these modifications, the networks retain their small-world characteristics. The combination of maintaining the small-world topology and its flexible malleability, controlled by a single parameter, is a desirable characteristic for learning mechanisms associated with plastic modifications in the brain (45–47). In lognormally distributed networks the emphasis on modularity versus centrality around $\tau_{transition}$ varies smoothly with $\tau$. This smooth transition may also be considered advantageous for neuronal networks, since slight changes of a global regulating process will not have dramatic effects in their connectivity pattern. This combination of stability and flexibility adds to the advantages of lognormal networks mentioned elsewhere, such as sparsity in spontaneous firing rates (48) and efficient coding (49).

## Materials and Methods

## Graph preliminaries

A graph is defined as an ordered pair *G = (V, E, W)*, where *V* denotes the set of vertices (or nodes), *E* the edges (or connections) between them, and *W* the set of edge weights, $W = \{w^{ij} \in \mathbb{R}_{\geq 0} | (i,j) \in E\}$, i.e. only nonnegative weights are used. The cardinalities, $|V| = n$ and $|E| = m$, express the total number of nodes and connections respectively. The connectivity pattern of *G* can be conveniently represented by an n x n adjacency matrix, *A*, with its entries denoting the weights between edges ($A_{ij} = w^{ij}$). $w^{ij} = 0$ signals that edges *i* and *j* are not adjacent. If $w^{ij} > 0$ then vertices *i* and *j* are adjacent; the greater the value of the weight the stronger the connection between vertices. In the case of a binary network, the weights and the entries of the adjacency matrix can take only two values, 0 or 1; $A_{ij} = 1$ indicates that vertices *i* and *j* are adjacent ($(i,j) \in E$), and $A_{ij} = 0$ that they are not ($(i,j) \notin E$). In this paper both binary and weighted graphs are undirected and simple (there are no self loops), meaning *A* is symmetric ($A_{ij} = A_{ji}$), and zero in its diagonal entries ($A_{kk} = 0$) respectively. The strength of a vertex *j* is the sum of the weights from the edges incident to it. Practically the strength is found by summing the rows or the columns of the adjacency matrix: $s_j = \sum_{i=1}^{n} A_{ij}$. In the case of binary networks, the summation indicates the degree $d_j$ of node *j*.

## The graph Laplacian matrix

The graph Laplacian is defined as $L = D - A$, where D is a diagonal matrix having the strengths (or degrees for binary adjacency matrices) of the nodes in its diagonal entries ($D_{ii} = s_i$). It arises naturally in optimization problems such as graph partitioning (50) or nonlinear dimensionality reduction (51). It also emerges as the discrete analogue of the Laplace-Beltrami operator ($\nabla^2 f$) in the heat flow equation for graphs. The normalized graph Laplacian is defined as $\mathcal{L} = D^{-1/2} L D^{-1/2}$, with $D_{ii}^{-1/2} = 0$ for $s_i = 0$. Its entries take the values:

$$\mathcal{L}_{ij} = \begin{cases} 1 \; if \; i = j \\ \dfrac{-A_{ij}}{\sqrt{s_i s_j}} \; if \; (i,j) \in E \\ 0 \; otherwise \end{cases}$$

We use the normalized graph Laplacian, a more appropriate operator for graphs that are not regular, that is graphs with vertices that do not necessarily have the same strength (or degree). This is because the eigenvector of the normalized graph Laplacian corresponding to the zero eigenvalue captures the graph irregularity: $\lambda_0 = 0, v_0 = [\sqrt{s_1}, \dots, \sqrt{s_n}]^T$ (52). Thus, throughout the paper, any mention of the Laplacian refers to the normalized version.

The Laplacian is a symmetric positive semidefinite matrix with real values: $\mathcal{L}_{ij} = \mathcal{L}_{ji}$ and $z^T \mathcal{L} z \geq 0, \forall z \in \mathbb{R}^n$; consequently, its eigenvalues are nonnegative and real and its eigenvectors form an orthonormal set. The spectral decomposition of the Laplacian matrix is $\mathcal{L} = V \Lambda V^T$, where $\Lambda = diag\ (\lambda_0, \lambda_1 \dots \lambda_n)$ is a diagonal matrix, having in its diagonal entries the eigenvalues

of $\mathcal{L}$ ($0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_n \leq 2$) and $V = [v_0 v_1 \ldots v_n]$ is an nXn matrix having as its columns the corresponding orthonormal eigenvectors ($\mathcal{L}v_i = \lambda_i v_i$).

The eigenvalues and eigenvectors of the Laplacian yield valuable information about the graph they are derived from (53). The multiplicity of zero indicates the number of connected components in the graph; a single zero eigenvalue corresponds to a connected graph, two zero eigenvalues to a disconnected graph with two components, and so on. The second smallest eigenvalue, called the Fiedler value, and the largest one are related to several graph properties such as the graph's connectivity, diameter and convergence to a stationary probability distribution for random walks (54–57). The values of the eigenvector corresponding the smallest nonzero eigenvalue, the Fiedler vector, can be used to sort vertices so that the ones close to each other belong to the same community, and accordingly partition a graph (50,58). Nodes are sorted in the same order as the elements of the Fiedler vector in figures showing adjacency matrices.

## Heat kernel

The heat equation is a partial differential equation that describes how heat is spatially dissipated over time. It is defined as:

$$\frac{\partial h(t,\ x_{1,\ldots,}\ x_n)}{\partial t} = a\nabla^2 h(t,\ x_{1,\ldots,}\ x_n)$$

where $h$ is a function of n spatial variables $x_1, \ldots, x_n$ and time variable $t$, $a$ is a positive constant and $\nabla^2$ is the Laplacian operator, or else the divergence of the gradient.

The heat equation based on the graph Laplacian matrix is similarly describing the variation in the flow of heat (or energy or information) within a graph over time:

$$\frac{\partial h(t)}{\partial t} = -\mathcal{L}h(t)$$

The heat kernel, $h(t)$, and the graph Laplacian, $\mathcal{L}$, are both nXn matrices describing the flow of heat across edges and its rate of change respectively. $h(\tau)_{ij}$ is the amount of heat transferred from vertex i to vertex j after time τ. The heat equation has an explicit solution:

$$h(t) = e^{-t\mathcal{L}}$$

As it was shown in the previous section, $\mathcal{L}$ can be decomposed to its eigenspectrum. The heat kernel can then be written as:

$$h(t) = Ve^{-t\Lambda}V^T = \sum_{i=0}^{n-1} e^{-\lambda_i t} v_i v_i^T$$

The dynamics for which diffusion spreads in the graph are based on the eigenvalue/eigenvector pairs of the Laplacian matrix. In the summation the contribution of each eigenmode is not the same, but depends on the magnitude of the eigenvalue. Typically, as t increases, only the eigenmodes with the few smallest eigenvalues play a significant role in the summation. For a connected graph, in the extreme case of infinite time ($t = \infty$), only the first eigenmode will contribute to the summation, since $\lambda_0 = 0$.

## Adaptive rewiring algorithm

With the adaptive rewiring algorithm and its variations employed in this paper, we seek to probe the effects that a simple self-organizing rule will have on the properties of an initially randomly connected network. The core algorithm is essentially the same as the one in Jarman and colleagues (9), extended for application to both binary and weighted networks.

Starting with a random Erdös–Rényi type with $|V| = n$ vertices and $|E| = floor(2\frac{\log(n)}{n}n(n-1))$ connections- the latter guaranteeing the random network is fully connected (59)- the rewiring algorithm proceeds as follows:

Step 1: Select, with uniform random probability, a vertex, $k$ from the set of vertices in the graph that are of nonzero degree but also not connected to all other vertices ($k \in V|\, 0 < d_k < n-1$).

Step 2: With probability p, go to step 2.1 (random rewiring), otherwise go to step 2.2 (heat diffusion rewiring). In both cases, select vertex $j_1$ from the set of vertices that are not connected to $k$, ($j_1 \in \{j \in V|(j,k) \notin E\}$), and vertex $j_2$ from the set of vertices that are connected to $k$, ($j_2 \in \{j \in V|(j,k) \in E\}$), albeit in a differen. Delete the edge $(k, j_2)$, and add the edge $(k, j_1)$,. In the case of weighted networks, unless otherwise stated, use the weight of the previously connected edge $(k, j_2)$, for the edge $(k, j_1)$.

Step 2.1: For both $j_1$ and $j_2$ the selection is random and uniform among the elements of each set.

Step 2.2: Calculate the heat kernel, $h(\tau)$, of the adjacency matrix. $j_1$ is selected such that from all the vertices not connected to k, it is the one with the highest heat transfer with k. $j_2$ is selected such that from all the vertices connected to k, it is the one with the lowest heat transfer with k. In mathematical terms, we can express the statements above as follows:

$$j_1 = argmax_{(k,j)\notin E,\ k\neq j} h_{kj}(\tau)$$

$$j_2 = argmin_{(k,j)\in E,\ k\neq j}\ h_{kj}(\tau)$$

Step 3. Go back to Step 1 until k rewiring iterations are made.

## Graph measures

All the analyses were programmed in Python. The code package of the functions producing the rewired adjacency matrices, as well as the metrics described below characterizing them are publicly available (https://github.com/rentzi/netRewireAnalyze).

The *clustering coefficient*, C, measures the local cohesiveness of a network by quantifying the extent to which two nodes both adjacent to a common node are themselves connected to each other. The *weighted clustering coefficient*, $C_w$, includes the weight of the edges forming these triplets. It is defined as:

$$C_w = \sum_{i=1}^{n} c_i^w$$

where $c_i^w$ measures the clustering coefficient for node *i* and is expressed as:

$$c_i^w = \frac{1}{s_i(d_i - 1)} \sum_{j,h} \frac{w_{ij} + w_{ih}}{2} \alpha_{ij} \alpha_{ih} \alpha_{jh}$$

$s_i$ is the strength and $d_i$ the degree of node $i$, $w_{ij}$ the weight of the connection between nodes $i$ and $j$, and $\alpha_{ij}$ a binary index that is 1 if $w_{ij} > 0$, and 0 otherwise (60). The formula above also works for a binary network.

A network is *small world* if it is connected in such a way so that the nodes connected to a common node are likely to be connected as well (high $C$), but at the same time any node can reach any other node in a few steps. The latter property refers to the average path length, $L$, which has important ramifications in a network's behavior. For example, in a computational unit such as a visual module small $L$ suggests that the results of local scene computations can be relayed fast to other components of the network for the computation of global effects (for example normalization). The metric that measures the small worldness of a network is defined as:

$$S = \frac{\frac{C}{C_{random}}}{\frac{L}{L_{random}}}$$

where $C$ and $L$ are the average clustering coefficient and path length of the network, and $C_{random}$ and $L_{random}$ are the ones for a random network with the same number of nodes and connections. A network with S>1 is considered small world.

To obtain a measure of network *modularity*, we used the spectral algorithm for community detection introduced by Newman (25), which identifies the communities within the network and also provides a modularity index (Q). The index Q gives a measure of the density of

the interconnections within and between the found groups compared to a randomly connected graph with the same degree distribution: the greater Q is, the denser the connections within-and the sparser between- groups. For a division of a network into two clusters, we compute the eigenvector corresponding to the largest eigenvalue of the modularity matrix, *B*, and assign the nodes to one or the other group by the sign of the elements in the eigenvector. The entries of *B* are defined as:

$$B_{ij} = A_{ij} - \frac{s_i s_j}{2m}$$

where $s_i$ is the strength of node *i*, and *m* is the total strength of the network. Q for this division into two groups is then:

$$Q = \frac{1}{4m} g^T B g$$

where $g_i$ is 1 is node *i* belongs to group 1, and -1 if it belongs to group 2. We iteratively continue to subdivide the groups and add the additional contribution ΔQ. We stop until any additional split will give us a negative ΔQ. For a complete description of the iterative algorithm see Newman (25).

A network is *assortative* if nodes with similar characteristics are preferentially connected. In our study the characteristic is degree. The assortativity coefficient (61) formally expresses this property as:

$$r = \frac{\sum_{ij}(A_{ij} - s_i s_j/2m) d_i d_j}{\sum_{ij}(s_i \delta_{ij} - s_i s_j/2m) d_i d_j}$$

where $A_{ij}$ is the weight of the connection between the *i* and *j* nodes, $s_i$ is the strength of the *i*-th node, *m* is the total strength of the network and $d_i$ is the degree of node *i*. *r* can take values between 1 and -1, with 1 signifying a perfectly assortative network and -1 a perfectly

disassortative one. The equation is generalizable for both binary (where $s_i$ is then equal to $d_i$) and weighted network and is equivalent to the equations found in other studies for weighted networks (62,63).

*Rich club* (27,64) refers to high degree nodes preferentially connecting to each other. The topological rich club coefficient, $\Phi(k)$, disregards connection weights. It quantifies the tendency of the subgroup of nodes with degree greater than $k$ to connect with each other. It is expressed as the ratio of the existing connections between the nodes with degree greater than k over the maximum possible connections between them:

$$\Phi(k) = \frac{2M_{>k}}{N_{>k}(N_{>k} - 1)}$$

where $N_{>k}$ is the number of nodes with degree greater than k, and $M_{>k}$ is the number of connections between them. The rich club coefficient naturally increases with increasing k, even in randomly connected networks. We therefore normalize $\Phi(k)$, with the coefficient, $\Phi_{random}(k)$, obtained from a random network with the same number of nodes, connections, and degree distribution as the original one. We construct the random network by performing $100*M_{>k}$ double edge swaps from the original network. A double edge swap involves randomly choosing two connections a-b and c-d, and replacing them with the connections a-c and b-d. If the later connections already exist, then the algorithm does not count that iteration and chooses again two random connections (43). The normalized rich club coefficient is expressed as:

$$\Phi_{norm} = \frac{\Phi(k)}{\Phi_{random}(k)}$$

For the weighted rich club coefficient, $\Phi_w(k)$, the node selection is based, similarly to $\Phi(k)$, on the degree k. $\Phi_w(k)$ is expressed as:

$$\Phi_w(k) = \frac{W_{>k}}{\sum_{i=1}^{E_{>k}} w_i^{ranked}}$$

where $W_{>k}$ denotes the sum of the connection weights between the nodes with degree greater than k, $E_{>k}$ is the number of connections between them, and $w^{ranked}$ , the weights in the network in descending order. To normalize the weighted rich club coefficient, we use $\Phi_{W,random}(k)$, which has the same topology as the original network but reshuffled edge weights (28,65). The weighted normalized rich club is then:

$$\Phi_{W,norm} = \frac{\Phi_w(k)}{\Phi_{W,random}(k)}$$

## Simulations and analyses

Simulations and analyses were performed on 100-node networks (average degree was 18.24). Most simulations vary two parameters from the rewiring algorithm: the random rewiring probability ($p_{random}$) and the time variable in the heat equation ($t = \tau$). For each combination of parameter values, we run the rewiring algorithm 100 times. Unless otherwise stated, the figures show mean values of these 100 runs, with the vertical lines at the measured points indicating standard deviation. Each run involves 4000 rewirings, since by then networks typically have diverged from a random connectivity pattern.

## Acknowledgments

We thank Steeve Laquitaine for inspiring discussions.

## References


1. Sporns O. Networks of the brain. 2016.

2. Bassett DS, Bullmore ET. Small-world brain networks revisited. The Neuroscientist. 2017;23(5):499–516.

3. Hilgetag C-C, Burns GA, O'Neill MA, Scannell JW, Young MP. Anatomical connectivity defines the organization of clusters of cortical areas in the macaque and the cat. Philos Trans R Soc Lond B Biol Sci. 2000;355(1393):91–110.

4. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci. 2009;10(3):186.

5. van den Heuvel MP, Sporns O. Rich-Club Organization of the Human Connectome. J Neurosci. 2011 Nov 2;31(44):15775–86.

6. Kaiser M, Varier S. Evolution and development of brain networks: from Caenorhabditis elegans to Homo sapiens. Netw Comput Neural Syst. 2011;22(1–4):143–147.

7. Gong P, van Leeuwen C. Emergence of scale-free network with chaotic units. Phys Stat Mech Its Appl. 2003;321(3–4):679–688.

8. Gong P, van Leeuwen C. Evolution to a small-world network with chaotic units. EPL Europhys Lett. 2004;67(2):328.

9. Jarman N, Steur E, Trengove C, Tyukin IY, van Leeuwen C. Self-organisation of small-world networks by adaptive rewiring in response to graph diffusion. Sci Rep [Internet]. 2017 Dec [cited 2019 Jul 14];7(1). Available from: http://www.nature.com/articles/s41598-017-12589-9

10. Rubinov M, Sporns O, van Leeuwen C, Breakspear M. Symbiotic relationship between brain structure and dynamics. BMC Neurosci [Internet]. 2009 Dec [cited 2019 Jul 15];10(1). Available from: https://bmcneurosci.biomedcentral.com/articles/10.1186/1471-2202-10-55

11. Berg D van den, Leeuwen C van. Adaptive rewiring in chaotic networks renders small-world connectivity with consistent clusters. Europhys Lett EPL. 2004 Feb;65(4):459–64.

12. Papadopoulos L, Kim JZ, Kurths J, Bassett DS. Development of structural correlations and synchronization from adaptive rewiring in networks of Kuramoto oscillators. Chaos Interdiscip J Nonlinear Sci. 2017 Jul;27(7):073115.

13. Butz M, Woergoetter F, van Ooyen A. Activity-dependent structural plasticity. Brain Res Rev. 2009;60(2):287–305.

14. Avena-Koenigsberger A, Misic B, Sporns O. Communication dynamics in complex brain networks. Nat Rev Neurosci. 2018;19(1):17.

15. Latora V, Marchiori M. Efficient Behavior of Small-World Networks. Phys Rev Lett [Internet]. 2001 Oct 17 [cited 2019 Jul 14];87(19). Available from: https://link.aps.org/doi/10.1103/PhysRevLett.87.198701

16. Kwok HF, Jurica P, Raffone A, van Leeuwen C. Robust emergence of small-world structure in networks of spiking neurons. Cogn Neurodyn. 2007 Mar;1(1):39–51.

17. Grossberg S. How Does a Brain Build a Cognitive Code? In: Studies of Mind and Brain [Internet]. Dordrecht: Springer Netherlands; 1982 [cited 2019 Jul 14]. p. 1–52. Available from: http://www.springerlink.com/index/10.1007/978-94-009-7758-7_1

18. Abdelnour F, Voss HU, Raj A. Network diffusion accurately models the relationship between structural and functional brain connectivity networks. Neuroimage. 2014;90:335–347.

19. Hellrigel S, Jarman N, van Leeuwen C. Adaptive rewiring in weighted networks. Cogn Syst Res. 2019 Jun;55:205–18.

20. Teramae J, Fukai T. Computational Implications of Lognormally Distributed Synaptic Weights. Proc IEEE. 2014 Apr;102(4):500–12.

21. Buzsáki G, Mizuseki K. The log-dynamic brain: how skewed distributions affect network operations. Nat Rev Neurosci. 2014;15(4):264.

22. Turrigiano GG, Nelson SB. Homeostatic plasticity in the developing nervous system. Nat Rev Neurosci. 2004 Feb;5(2):97–107.

23. van den Heuvel MP, Kahn RS, Goni J, Sporns O. High-cost, high-capacity backbone for global brain communication. Proc Natl Acad Sci. 2012 Jul 10;109(28):11372–7.

24. Watts DJ, Strogatz SH. Collective dynamics of ‘small-world’ networks. Nature. 1998 Jun;393(6684):440–2.

25. Newman MEJ. Modularity and community structure in networks. Proc Natl Acad Sci. 2006 Jun 6;103(23):8577–82.

26. Alstott J, Bullmore E, Plenz D. powerlaw: a Python package for analysis of heavy-tailed distributions. PloS One. 2014;9(1):e85777.

27. Colizza V, Flammini A, Serrano MA, Vespignani A. Detecting rich-club ordering in complex networks. Nat Phys. 2006;2(2):110.

28. Alstott J, Panzarasa P, Rubinov M, Bullmore ET, Vértes PE. A unifying framework for measuring weighted rich clubs. Sci Rep. 2014;4:7258.

29. Jarman N, Trengove C, Steur E, Tyukin I, van Leeuwen C. Spatially constrained adaptive rewiring in cortical networks creates spatially modular small world architectures. Cogn Neurodyn. 2014 Dec;8(6):479–97.

30. van den Berg D, Gong P, Breakspear M, van Leeuwen C. Fragmentation: loss of global coherence or breakdown of modularity in functional brain architecture? Front Syst Neurosci [Internet]. 2012 [cited 2019 Jul 15];6. Available from: http://journal.frontiersin.org/article/10.3389/fnsys.2012.00020/abstract

31. Kaneko K. Overview of coupled map lattices. Chaos Interdiscip J Nonlinear Sci. 1992 Jul;2(3):279–82.

32. Damicelli F, Hilgetag CC, Hütt M-T, Messé A. Modular topology emerges from plasticity in a minimalistic excitable network model. Chaos Interdiscip J Nonlinear Sci. 2017;27(4):047406.

33. Stephan KE, Hilgetag C-C, Burns GA, O'Neill MA, Young MP, Kotter R. Computational analysis of functional connectivity between areas of primate cerebral cortex. Philos Trans R Soc Lond B Biol Sci. 2000;355(1393):111–126.

34. Fair DA, Dosenbach NU, Church JA, Cohen AL, Brahmbhatt S, Miezin FM, et al. Development of distinct control networks through segregation and integration. Proc Natl Acad Sci. 2007;104(33):13507–13512.

35. Dosenbach NU, Fair DA, Cohen AL, Schlaggar BL, Petersen SE. A dual-networks architecture of top-down control. Trends Cogn Sci. 2008;12(3):99–105.

36. Chen Y, Wang S, Hilgetag CC, Zhou C. Trade-off between multiple constraints enables simultaneous formation of modules and hubs in neural systems. PLoS Comput Biol. 2013;9(3):e1002937.

37. Kaas JH. Evolution of columns, modules, and domains in the neocortex of primates. Proc Natl Acad Sci. 2012 Jun 26;109(Supplement_1):10655–60.

38. van den Heuvel MP, Sporns O. Network hubs in the human brain. Trends Cogn Sci. 2013 Dec;17(12):683–96.

39. Zamora-López G, Zhou C, Kurths J. Cortical hubs form a module for multisensory integration on top of the hierarchy of cortical networks. Front Neuroinformatics. 2010;4:1.

40. Gong G, He Y, Concha L, Lebel C, Gross DW, Evans AC, et al. Mapping Anatomical Connectivity Patterns of Human Cerebral Cortex Using In Vivo Diffusion Tensor Imaging Tractography. Cereb Cortex. 2009 Mar;19(3):524–36.

41. Achard S, Salvador R, Whitcher B, Suckling J, Bullmore ED. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. J Neurosci. 2006;26(1):63–72.

42. Eguiluz VM, Chialvo DR, Cecchi GA, Baliki M, Apkarian AV. Scale-free brain functional networks. Phys Rev Lett. 2005;94(1):018102.

43. Towlson EK, Vertes PE, Ahnert SE, Schafer WR, Bullmore ET. The Rich Club of the C. elegans Neuronal Connectome. J Neurosci. 2013 Apr 10;33(15):6380–7.

44. Nigam S, Shimono M, Ito S, Yeh F-C, Timme N, Myroshnychenko M, et al. Rich-club organization in effective connectivity among cortical neurons. J Neurosci. 2016;36(3):670–684.

45. Trachtenberg JT, Chen BE, Knott GW, Feng G, Sanes JR, Welker E, et al. Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. Nature. 2002 Dec;420(6917):788–94.

46. Le Be J-V, Markram H. Spontaneous and evoked synaptic rewiring in the neonatal neocortex. Proc Natl Acad Sci. 2006 Aug 29;103(35):13214–9.

47. Hofer SB, Mrsic-Flogel TD, Bonhoeffer T, Hübener M. Experience leaves a lasting structural trace in cortical circuits. Nature. 2009 Jan;457(7227):313–7.

48. Ikegaya Y, Sasaki T, Ishikawa D, Honma N, Tao K, Takahashi N, et al. Interpyramid Spike Transmission Stabilizes the Sparseness of Recurrent Network Activity. Cereb Cortex. 2013 Feb 1;23(2):293–304.

49. Cossell L, Iacaruso MF, Muir DR, Houlton R, Sader EN, Ko H, et al. Functional organization of excitatory synaptic strength in primary visual cortex. Nature. 2015;518(7539):399.

50. Jianbo Shi, Malik J. Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell. 2000 Aug;22(8):888–905.

51. Belkin M, Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering. In: Advances in neural information processing systems. 2002. p. 585–591.

52. Chung F, Richardson RM. Weighted Laplacians and the sigma function of a graph. Contemp Math. 2006;415:93.

53. Chung FR. Lectures on spectral graph theory. CBMS Lect Fresno. 1996;6:17–21.

54. Fiedler M. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. Czechoslov Math J. 1975;25(4):619–633.

55. Fiedler M. Algebraic connectivity of graphs. Czechoslov Math J. 1973;23(2):298–305.

56. Mohar B, Alavi EY, Chartrand G, Oellermann OR, Schwenk AJ. THE LAPLACIAN SPECTRUM OF GRAPHS. :28.

57. Chung FR. Diameters and eigenvalues. J Am Math Soc. 1989;2(2):187–196.

58. Pothen A, Simon HD, Liou K-P. Partitioning Sparse Matrices with Eigenvectors of Graphs. SIAM J Matrix Anal Appl. 1990 Jul;11(3):430–52.

59. Erd\Hos P, Rényi A. On the evolution of random graphs. Publ Math Inst Hung Acad Sci. 1960;5(1):17–60.

60. Barrat A, Barthelemy M, Pastor-Satorras R, Vespignani A. The architecture of complex weighted networks. Proc Natl Acad Sci. 2004;101(11):3747–3752.

61. Newman MEJ. Mixing patterns in networks. Phys Rev E [Internet]. 2003 Feb 27 [cited 2019 Jul 15];67(2). Available from: https://link.aps.org/doi/10.1103/PhysRevE.67.026126

62. Leung CC, Chau HF. Weighted assortative and disassortative networks model. Phys Stat Mech Its Appl. 2007 May;378(2):591–602.

63. Farine DR. Measuring phenotypic assortment in animal social networks: weighted associations are more robust than binary edges. Anim Behav. 2014;89:141–153.

64. Zhou S, Mondragon RJ. The Rich-Club Phenomenon in the Internet Topology. IEEE Commun Lett. 2004 Mar;8(3):180–2.

65. Serrano MÁ. Rich-club vs rich-multipolarization phenomena in weighted networks. Phys Rev E [Internet]. 2008 Aug 4 [cited 2019 Jul 15];78(2). Available from: https://link.aps.org/doi/10.1103/PhysRevE.78.026101

## Supporting Information

## S1 Appendix

We performed pairwise correlations (Pearson coefficient; $\rho$) along the $\tau$ dimension, separately for each of the 100 iterations and $p_{random}$ bins (pairs of $p_{random}$ ,0.1 apart) to quantify the dependencies between path length (L), average clustering coefficient (C) and small-world (S) metrics. Qualitatively, C shows a similar pattern to S: steepest increase with respect to $\tau$ for binary networks, and increasingly moderate for the normal and lognormal distributions (Fig S1A). C is perfectly correlated with S (average across iterations and probability bins was $\rho_{binary} = 0.91$, $\rho_{normal} = 0.96$, $\rho_{lognormal} = 0.97$).

By contrast, L is not a monotonically increasing function of $\tau$ but reaches a peak and subsequently reverts back to values close to those of random connectivity networks; the peak is reduced to an almost flat line for $p_{rand} > 0.4$ (Fig S1B). Overall, across $\tau$, L was positively correlated with C, for all networks ($\rho_{binary} = 0.42$, $\rho_{normal} = 0.52$, $\rho_{lognormal} = 0.31$). The correlation results along with the S values indicate that the diffusion rewiring increases small worldness by increasing clustering at a greater rate compared to path length, something that is more accentuated for larger random rewiring probabilities (compare Figue S1A with S1B).

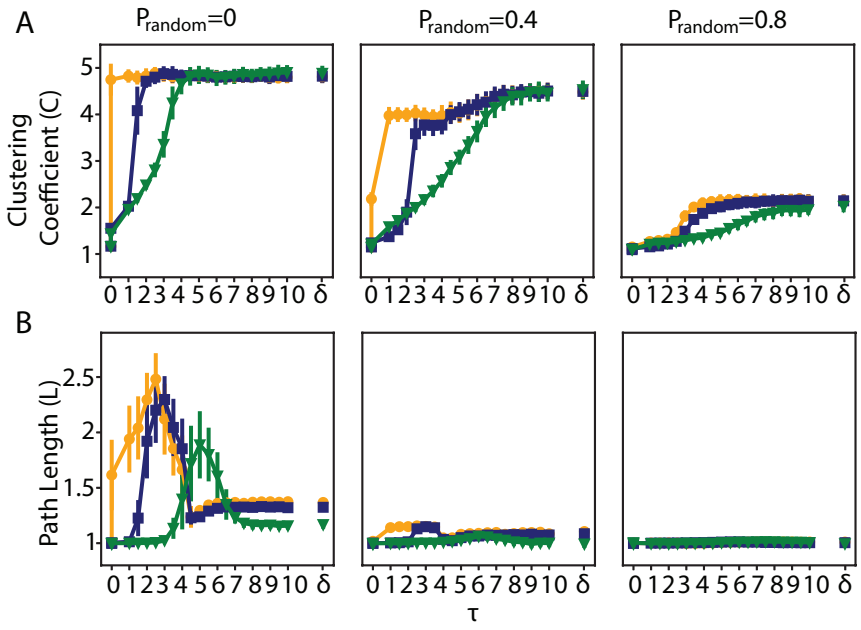


**Fig S1. An increase in C as a function of τ is accompanied by a more moderate increase in L.**

A. C as a function of τ for different $p_{random.}$ (0, 0.4, 0.8). B. Same as A for the L metric. Horizontal lines indicate standard deviation.

## S2 Appendix

To probe the heavy-tailed degree distributions of the centralized networks, we used $p_{random} = 0.2$, and $\tau_{binary} = 5$, $\tau_{normal} = 5$, $\tau_{lognormal} = 7$ as representative $\tau$ values for centralized networks. Using log-likelihood ratios, we determined whether they best fit a power-law, lognormal or exponential distribution. We first compared the lognormal with the power-law fit, and then the better fit of the two with the exponential distribution. We repeated this process for 100 different realizations of the network. In over 90% of the cases either power-law or lognormal were a better fit compared to exponential distributions, with no significant difference between lognormal and power-law. The fits were adjusted so that they optimally accommodate the tail of the distribution, i.e. if they did not improve the fit, bins with the smaller degrees were not included. For the optimal range of degrees at the tail end of the distributions, the lognormal and power-law fits were almost identical (Fig S2, fits on the left side of the plots). We performed the same analysis for networks of 1000 nodes (this time for one iteration due to the computational load) to probe whether the fit changes when the maximum allowable connections of a node increases by an order of magnitude with similar conclusions (but worse fits). Power law distribution functions are of the form $P(k) \sim k^{-\alpha}$. For all networks the exponent $\alpha$ was the same up to a decimal ($\alpha = 1.7$). The strength distributions of the normal and lognormal networks again showed a similar behavior to their degree counterparts with the only difference being that the distribution decayed faster (exponent $\alpha$ varied between 2 and 2.6).

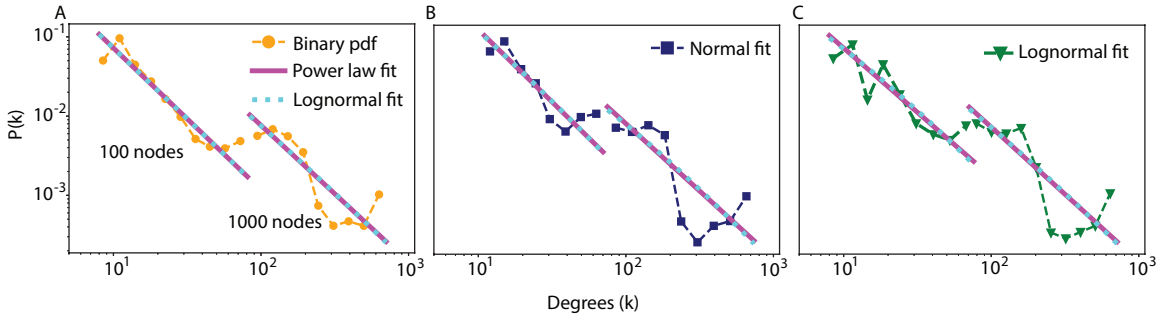


**Fig S2. Power law and lognormal fits are better fits of the degree distribution of centralized networks compared to the exponential fit.**

Log-log plot of the degree distribution of an example 100 node network and a 1000 node network along with their power-law and lognormal fits for (A) binary, (B) normal and (C) lognormal networks.

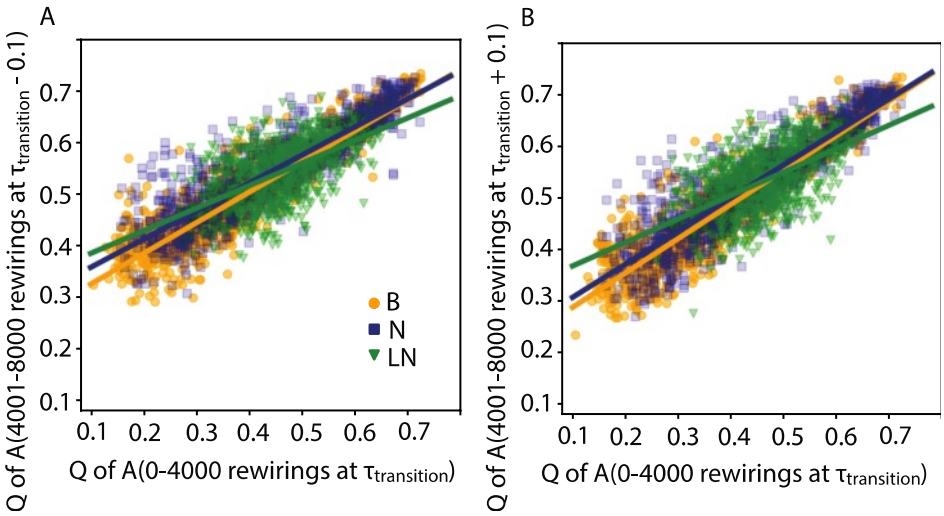


**Fig S3. τ values near $\tau_{transition}$ do not affect the modularity index correlations seen at $\tau_{transition}$.**

A. Scatter plot showing strong positive correlation of the modularity indices between the same network after 4000 rewirings at $\tau_{transition}$ and an additional 4000 rewirings at $\tau_{transition}$-0.1 (towards the more modular direction) B. Same as A but for $\tau_{transition}$+0.1 (towards the more centralized direction). The lines in all cases indicate the best linear fits of the data.

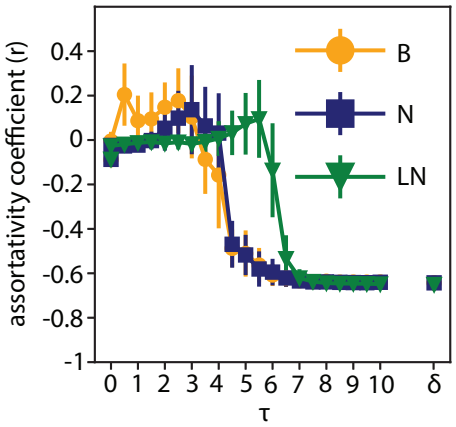


**Fig S4. Networks are in two assortative states: a low and a disassortative one.** The assortativity coefficient as a function of τ for binary, normal and lognormal networks. $p_{random} = 0.2$ for all networks.

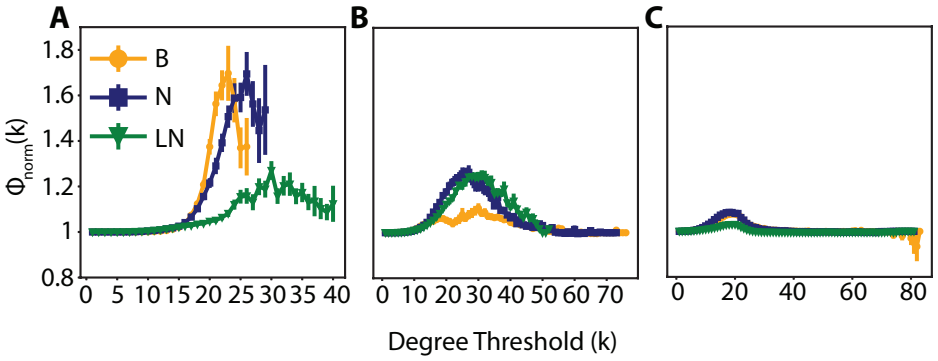


**Fig S5. The weighted networks are rich club at the modular and transition state, the binary only at the modular state.**

Topological normalized rich club, $\Phi_{norm}(k)$, for binary, normal and lognormal networks in the (A) modular ($\tau_{binary} = 2$, $\tau_{normal} = 3$, $\tau_{lognormal} = 4.5$) (B) critical ($\tau_{binary} = 4.1$, $\tau_{normal} = 4.15$, $\tau_{lognormal} = 5.5$) and (C) centralized ($\tau_{binary} = 5$, $\tau_{normal} = 5$, $\tau_{lognormal} = 7$) state. In all cases $p_{random} = 0$