On null models for temporal small-worldness in brain dynamics
Aurora Rossi, Samuel Deslauriers-Gauthier, Emanuele Natale

To cite this version:
Aurora Rossi, Samuel Deslauriers-Gauthier, Emanuele Natale. On null models for temporal small-worldness in brain dynamics. Network Neuroscience, 2024, pp.1-30. 10.1162/netn_a_00357. hal-04389639

HAL Id: hal-04389639
https://hal.science/hal-04389639
Submitted on 11 Jan 2024
On null models for temporal small-worldness in brain dynamics

Aurora Rossi¹, Samuel Deslauriers-Gauthier² and Emanuele Natale ¹

¹Université Côte d’Azur, COATI, INRIA, CNRS, I3S, France
²CRONOS, Inria Centre at Université Côte d’Azur, France

Keywords: Brain Networks, Temporal Networks, Small-worldness, Hyperbolic Graph, Null Model, fMRI

ABSTRACT

Brain dynamics can be modeled as a temporal brain network starting from the activity of different brain regions in functional magnetic resonance imaging (fMRI) signals. When validating hypotheses about temporal networks, it is important to use an appropriate statistical null model that shares some features with the treated empirical data. The purpose of this work is to contribute to the theory of temporal null models for brain networks by introducing the Random Temporal Hyperbolic Graph Model (RTH), an extension of the Random Hyperbolic Graph (RH), known in the study of complex networks for its ability to reproduce crucial properties of real-world networks. We focus on temporal small-worldness which, in the static case, has been extensively studied in real-world complex networks and has been linked to the ability of brain networks to efficiently exchange information. We compare the RTH Graph Model with standard null models for temporal networks and show it is the null model that best reproduces the small-worldness of resting brain activity. This ability to reproduce fundamental features of real brain networks, while adding only a single parameter compared to classical models, suggests that the RTH Graph Model is a promising tool for validating hypotheses about temporal brain networks.

AUTHOR SUMMARY

We show that the Random Temporal Hyperbolic Graph (RTH) is a suitable null model for testing hypotheses about brain dynamics, after comparing it with the current state of the art and two other geometric null models. The static version of this theoretical model captures properties of various real-world networks, and its temporal version exhibits the temporal small-world property, for which we
propose a new proper temporal definition. In particular, we show that the model best reproduces the temporal small-worldness measured in the empirical temporal network extracted from fMRI signals.
INTRODUCTION

The brain is always active, even when it is not performing a cognitive task. Therefore, it is interesting to investigate its fluctuations starting from this case: the resting state condition (Karahanoglu & Van De Ville, 2017; Maltbie, Yousefi, Zhang, Kashyap, & Keilholz, 2022; Preti, Bolton, & Van De Ville, 2017; Thompson, Brantefors, & Fransson, 2017). In this context, functional Magnetic Resonance Imaging (fMRI) techniques allow us to study the underlying functional architecture of the brain and its temporal evolution. We can represent cerebral activity by extracting and manipulating the Blood-Oxygenation-Level Dependent (BOLD) signal from the scans.

To test certain hypotheses about observed functional connectivity, after modeling brain activity as a complex temporal network (Bahrami, Laurienti, Shappell, Dagenbach, & Simpson, 2022; Sizemore & Bassett, 2018; Thompson et al., 2017), it is important to compare it with an appropriate null model (Lurie et al., 2020; Vasa & Misic, 2022). A null model is a random statistical object that has certain properties in common with the empirical data under consideration and is used to evaluate whether the latter has noteworthy features or properties that cannot be attributed to randomness or other constraints.

In the complex temporal networks of brain activity just mentioned, each node corresponds to a brain region, and edges represent the presence of interactions greater than a certain threshold between two connected regions. As the graph evolves over time, edges may appear and disappear, meaning that the interaction between two brain regions jumps over and below a certain value. Complex networks can be found across different fields such as biology, sociology, epidemiology and brain dynamics (Costa et al., 2011). They are graphs with non-trivial topological properties, two main properties that have come to define them are scale-freeness and small-worldness (Bassett & Bullmore, 2017; Chung & Lu, 2006; Tomasi, Shokri-Kojori, & Volkow, 2017).

In this paper, the common property of the null model and the real data we focus on is small-worldness, which expresses the efficiency of information exchange between nodes. It is defined as the ratio between two properties of the network: the clustering coefficient and the average path length. The clustering coefficient measures how many connections exist among a node’s neighbors, which reflects the local connectivity of the network. The average path length measures the average number of steps it takes to travel from one node to another, which reflects the global connectivity of the network. Given the fact that
we are considering temporal brain networks, the temporal version of small-worldness can be interpreted as a measure of efficient communication among brain regions. We propose a new definition of temporal small-worldness $S = C/L$ as the ratio of the temporal clustering coefficient $C$ and the temporal path length $L$. The temporal clustering coefficient $C$ is the time average of the clustering coefficients computed at each time step. The temporal path length $L$ is the average of the fastest temporal paths between all pairs of nodes. $S$ differs from $S_{SB}$, the temporal small-worldness definition of Sizemore and Bassett (2018), as it replaces the temporal version clustering coefficient with the temporal correlation coefficient $TC$, which measures the change between two consecutive time steps. In the Measures Section we will discuss the reasons behind the adoption of this new definition.

Until recently, neuroscientists mainly analyzed the static version of brain networks (Bassett & Bullmore, 2017; Fornito, Zalesky, & Bullmore, 2016; Liao, Vasilakos, & He, 2017; Váša & Mišić, 2022), as the temporal extension of graph models is a recent research topic even in computer science. Currently, the temporal null models, that have been considered in the literature, are the randomized edge model and the randomly permuted times model, which are based on empirical data (Sizemore & Bassett, 2018). For consistency with the other models used in this work, we will refer to them as the Random Temporal Edge Graph (RTE) Model and the Random Temporal Permuted Times Graph (RTPT) Model. In the RTE graph, each edge is rewired at each time step by changing one of its endpoints in the original graph. In the RTPT graph, the temporal structure is deteriorated by randomly permuting the time at which edges occur. However, as we show in this work, these random models do not reflect the typical structure of small-world networks, which exhibit short average path lengths and low clustering coefficients (Watts & Strogatz, 1998).

Lately, in the research community of real-world complex networks, the Random Hyperbolic Graph Model (RH) has become famous because it can exhibit both a high-tail degree distribution and small-worldness (Bode, Fountoulakis, & Muller, 2014; Fountoulakis, 2012; Fountoulakis & Müller, 2016; Krioukov, Papadopoulos, Kitsak, Vahdat, & Boguñá, 2010). It is a geometric model whose nodes are points that lie in the hyperbolic space. The latter cannot be embedded in Euclidean space, due to the exponential growth of its volume. It also happens in the case of functional brain networks, that are best represented in the hyperbolic disc as shown by Whi, Ha, Kang, and Lee (2022). A temporal version of
the RH Graph Model has recently been considered (Hartle, Papadopoulos, & Krioukov, 2021; von Looz & Meyerhenke, 2018).

We also consider two other temporal geometric graphs, the Random Temporal Square Graph (RTS) Model and the Random Temporal Torus Graph (RTT) Model as possible good candidates for null models. We compare them to the state of the art and to the networks extracted from the signals contained in the empirical data of 1047 subjects from the WU-Minn Human Connectome Project (WU-Minn HCP 1200 Subjects Data Release Reference Manual, 2018), which we collected in a dataset that we made publicly available (Rossi, Deslauriers-Gauthier, & Natale, 2023). We test whether our random temporal models have small-worldness values close to those of empirical data at different connectivity thresholds. In particular, we show that the RTH Graph Model is a suitable null model with respect to the small-worldness property, especially when compared to the previous RTPT and RTE graph null models considered by Sizemore and Bassett (2018).

METHODS

Extracting temporal brain networks

In this section, we describe the fMRI preprocessed data we started from and the process we implemented to extract the temporal brain networks from them (Rossi et al., 2023).

Data. The data we used were taken from the Human Connectome Project (WU-Minn HCP 1200 Subjects Data Release Reference Manual, 2018), specifically, the 3 Tesla Siemens fMRI dataset of 1047 selected resting state subjects (rfMRI) scanned at the University of Washington or the University of Minnesota. The data are preprocessed and minimally preprocessed according, respectively, to Smith et al. (2013) and Glasser et al. (2013), and had artifacts removed using ICA-FIX (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). During the acquisition of the brain images, the patients would lie quietly in a darkened room with their eyes open, looking at a fixed bright point on a dark background, without performing any tasks. This is arguably the simplest case one can imagine, and it is natural to start the investigation from this point because even when the subject does nothing, brain regions are activated and the signal fluctuates. Understanding this scenario is important to proceed with the comparison between
this state and more complicated situations, such as when the patient has some neural disease or is performing cognitive tasks involving memory and attention (Liao et al., 2017).

The rfMRI data were acquired in four runs of 14 minutes and 33 seconds each. The data were captured in both right-to-left (RL) and left-to-right (LR) phase encoding, with two scans per direction. We used only one of the two left-to-right scans. Regarding the temporal resolution of the data, the total number of time steps is 1200 and the interval between two time steps is 0.72 seconds. Regarding the spatial resolution, there are 91 voxels in the $x$ and $z$ axis and 109 in the $y$ axis, making the total number of voxels 902629. The dimension of a voxel is $2 \times 2 \times 2$ millimeters.

**Processing.** The pipeline we implemented takes the BOLD signals of the rfMRI scan for each voxel of the brain and transforms them into temporal graphs (Van Dijk et al., 2010). We release them in the publicly available dataset Rossi et al. (2023). The first step of the pipeline is to perform a linear regression on the movement parameters of the data, which reduces the contribution of the subject’s movements during the 14-minute procedure. The data also contains some noise caused by the patient’s involuntary respiratory and cardiac rates. This problem is mitigated by applying a bandpass filter that isolates the BOLD fluctuations within a range of $0.01 - 0.08 \, \text{Hz}$ (Van Dijk et al., 2010). To obtain the division of the brain into regions, that correspond to the nodes of our temporal networks, we use a brain atlas, such as the Gallardo (2018), Glasser et al. (2016) or Schaefer et al. (2018). Using an atlas we extract groups of voxels from gray matter areas that differ by anatomical and functional criteria. By averaging their signal, a time series is created for each brain region. Using the sliding window method, we individuate rectangular windows of length 60 seconds with an overlap of 30 seconds (Preti et al., 2017), each window being equivalent to a one-time step of the network. The Pearson correlation coefficient is calculated between each pair of brain region time series within each window; the resulting correlation value is the weight of the edge between the respective brain region nodes. In this way, we obtain an adjacency matrix for each window. The rows and columns of the matrix are ordered so that regions of the left hemisphere are grouped in the first block and regions of the right hemisphere are grouped in the second block (see the adjacency matrix in Figure 1). The two more correlated blocks (corresponding to the lighter pixels in the figure), upper left and lower right, represent strong intrahemispheric connectivity, while the two extra-diagonal blocks correspond to weaker
interhemispheric connections. Three diagonals can be distinguished, the middle one representing the 
correlation of a region with itself. The other two sub-diagonals correspond to the connectivity of each 
region with its counterpart in the opposite hemisphere (Fornito et al., 2016). The values range between -1 
and 1. If they are high (close to 1), the interpretation is clear: the two areas are synchronously activated, 
thus more likely to communicate. On the other hand, there is no consensus on the interpretation of 
negative indices. Given that many measures are defined for unweighted graphs, it is also common to 
threshold the data (Fornito et al., 2016). Currently, there is no consensus on the most appropriate 
threshold, so we analyze the data according to different positive thresholds applied to the Pearson 
correlation coefficient. The sequence of the obtained matrices defines the temporal brain networks.
Figure 1: Steps of the pipeline to extract temporal brain networks from an fMRI signal for each voxel of the brain. The first step is linear regression and bandpass filtering of the data. The second step is the division of the brain into regions according to an atlas. The third step is the sliding window method, which individuates rectangular windows within which, in the fourth step, the Pearson correlation coefficient is computed between each pair of brain region time series. The last step is the thresholding of the data.
Random graph models

In this paper, we consider five temporal random graph models. The first two models are the Random Temporal Permuted Times Graph Model (RTPT) and the Random Temporal Edges Graph Model (RTE) (Sizemore & Bassett, 2018), which are constructed by randomizing empirical data; in particular, the RTPT graph is obtained by randomizing the time at which each edge appears, while the RTE graph is obtained by randomly reassigning an edge endpoint to a different node within each time step. The other three models are the Random Temporal Square Graph Model (RTS), the Random Temporal Torus Graph Model (RTT), and the Random Temporal Hyperbolic Graph Model (RTH), which do not depend on empirical data. They are synthetic graphs obtained by randomly placing nodes on different spaces: the square, the torus, and the hyperboloid, respectively. At each time step, the temporal graph is constructed by connecting the nodes depending on their position. From one timestep to the next, the node positions are updated according to a movement model. Both the adjacency and their movement are computed according to the geometry of the underlying space. A summary of the five models and their parameters can be found in Table 1. Next, we provide their formal mathematical definitions. In the next section, we will compare them to the temporal brain networks extracted from data.

Temporal graph. A temporal graph, denoted as $G(V, E_T)$, is given by a set of nodes $V$ and a set of temporal edges $E_T$. Each temporal edge is a triple $(u, v, t)$ where $u, v \in V$ are the endpoints of the edge and $t \in [1, \ldots, T]$ is the time at which the edge appears thus $E_T = \{(u, v, t) : u, v \in V, t \in [1, \ldots, T]\}$.

Random Temporal Permuted Times Graph Model (RTPT($G$)). The Random Temporal Permuted Times Graph Model (Sizemore & Bassett, 2018) is obtained by taking the temporal graph $G(V, E_T)$ of a subject (see Figure 2a) and randomizing the time at which each edge appears (see Figure 2b). We construct a $G(V, E_T')$ from RTPT($G$) model by sampling a permutation $\pi$ uniformly at random and defining the set $E_T' = \{(u, v, \pi(t)) : (u, v, t) \in E_T\}$.

Random Temporal Edges Graph Model (RTE($G$)). The Random Temporal Edges Graph Model (Sizemore & Bassett, 2018) is obtained from the temporal graph $G(V, E_T)$ of a subject (see Figure 2a) by reassigning one endpoint of an edge to another node in each window time step (see Figure 2c). We
Figure 2: Example of 4-time steps illustrating how the RTPT and RTE graphs are obtained from a given temporal graph. The RTPT graph is obtained by randomizing the time each edge appears over different time steps. The RTE graph is obtained by reassigning one endpoint of an edge to a different node within each window time step.

Construct a $G(V, E_T')$ from RTE($G$) by sampling a permutation $\sigma$ uniformly at random of the set $V$ and by setting $(u, \sigma(v), t) \in E'$ for each contact $(u, v, t) \in E$.

**Random Square Graph Model (RS(r)).** A RS(r) graph $G(V, E)$ (Penrose, 2003) is obtained by distributing nodes independently and uniformly at random on a unit square $[0, 1]^2$ and connecting each pair of nodes whose Euclidean distance is less than radius $r$ (see nodes contained in orange balls in
Figure 3: Showing the main difference between the RS and the RT graphs. In both cases, the nodes are connected if they are at a distance less than a radius $r$. In the RS graph the nodes are distributed on a square, hence the node on the bottom right and the node on the top left are not connected; on the torus, on the contrary, they are connected.

Figure 3a). The set of edges is thus $E = \{(u, v) : (u, v) \in V \times V, ||p_u - p_v||_2 \leq r\}$ where $p_u$ and $p_v$ are the coordinates of the positions of the points $u$ and $v$.

**Random Torus Graph Model (RT($r$)).** A RT($r$) graph $G(V, E)$ is obtained by distributing nodes independently and uniformly at random on a torus $[0, 1]^2$ and connecting each pair of nodes that are at distance less than radius $r$ (see nodes contained in orange balls in Figure 3b). On the torus $[0, 1]^2$, the distance between two nodes $p_A, p_B$ with coordinates, respectively $(x_A, y_A), (x_B, y_B)$, is defined as

$$d(p_A, p_B) = \sqrt{\left(\frac{1}{2} - \frac{1}{2} - |x_A - x_B|\right)^2 + \left(\frac{1}{2} - \frac{1}{2} - |y_A - y_B|\right)^2}.$$

**Temporal version of Random Square and Torus Models.** To obtain the temporal version of both the RS and RT graphs models, which we will refer to them as RTS($r, v$) and RTT($r, v$) graph models, we can define a movement model according to which the positions of the points are updated and the distances are recomputed. The number of updates is set equal to the number of time steps of the real data. The new position of a point is determined by applying a displacement vector to it, whose direction is chosen uniformly in $[0, 2\pi)$ and whose length is chosen uniformly in $(0, v)$. $v$ is called the speed parameter. In the case of RT graph, there is no problem of updating the point position, since there are no boundaries.
the case of the unit square, boundaries need to be taken into account. If a point happens to move outside
the boundary of the unit square, its trajectory is reflected inwards, meaning that its position is updated as
if it had bounced off the boundary.

*Random Hyperbolic Graph Model (RH(ζ, α, R)).* Krioukov’s method for generating a Random
Hyperbolic graph is to distribute nodes quasi-uniformly within a disk of radius R centered on the upper
half of a hyperboloid of given negative curvature $K = -\zeta^2$ (Krioukov et al., 2010). Writing the node
positions in polar coordinates $(r, \theta)$, where $r \in [0, R]$ and $\theta \in [0, 2\pi]$, we have that the angular density is
$\rho(\theta) = \frac{1}{2\pi}$ and the radial coordinate density is $\rho(r) = \alpha \frac{\sinh(\alpha r)}{\cosh(\alpha R) - 1}$. The $\alpha$ parameter controls the spread
of the point positions. If $\alpha = \zeta$, the points follow a uniform distribution. If $\alpha > \zeta$, the points are more
likely to be near the border of the disk, otherwise, they are more likely to be near the center (see Figure
4a). Two nodes $u$ and $v$ are connected if they are at a hyperbolic distance $d_H$ less than $R$, where,
considering the polar coordinate of $u = (r_u, \theta_u)$ and $v = (r_v, \theta_v)$, $d_H$ is defined as

$$d_H((r_u, \theta_u), (r_v, \theta_v)) = \frac{1}{\zeta} \text{acosh}(\cosh(\zeta r_u \cosh(\zeta r_v - \sinh(\zeta r_u \sinh(\zeta r_v \cos(\pi - |\pi - |\theta_u - \theta_v|)))).$$

Thus, the parameter $R$ affects the average degree of the graph. The higher the value of $R$, the larger the
distances between the nodes and the lower the average degree (see Figure 4b). As for the curvature
parameter $\zeta$, we set it to $\zeta = 1$, since (Lemma 1.1 of Bode et al. (2014)) if two RH graphs have the same
ratio $\frac{\zeta}{\alpha}$ and the other parameters equal then they produce the same distribution on graphs.

To obtain a temporal version, which we will refer to it as the RTH($\zeta$, $\alpha$, $R$, $v$) Graph Model we update
the point position as many times as the number of time steps of the real data we want to compare with.
The point position update is chosen in such a way that guarantees that the marginal distribution of each
time step is the same as the initial one. Specifically, the polar coordinate is updated by adding to the
previous value a number chosen according to a uniform distribution over $(0, v)$, and then computing the
modulo $2\pi$. For the radial coordinate, we add to the previous value a number chosen according to a
uniform distribution over $(-v, v)$, and we reflect the result in the interval $[0, 1]$ to keep the distribution
uniform in $[0, 1]$ (Hartle et al., 2021). Thus, this parameter controls the movement of the points, and the
higher the value of $v$, the more the points move and the lower the percentage of edges that remain the
same between two consecutive time steps (see Figure 4c).
(a) Representations of the RTH graph on the Poincaré disk for different values of the parameter $\alpha$.

(b) Example showing how the parameter $R$ affects the average degree of the RTH graph when the other parameters are fixed $\alpha = 1.2, \zeta = 1$ and $\nu = 0.9$.

(c) Example showing how the parameter $\nu$ affects the percentage of edges that remain the same between two consecutive time steps when the other parameters are fixed $\alpha = 0.5, R = 4.5$ and $\zeta = 1$.

Figure 4: Behavior of the RTH Graph Model parameters.

Measures

In this section, we introduce some measures to define a temporal version of small-worldness. The first static qualitative definition of small-worldness was introduced by Watts and Strogatz (1998) as the combination of a high clustering coefficient, as in lattice graphs, and a short average path length, as in random graphs. Humphries and Gurney (2008) combined the previous measures into a quantifiable ratio:
<table>
<thead>
<tr>
<th>Name</th>
<th>Acronym</th>
<th>Data-based or Synthetic</th>
<th>Topological space</th>
<th>Parameters</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Temporal Square</td>
<td>RTS</td>
<td>Synthetic</td>
<td>Square $[0, 1]^2$</td>
<td>$r, v$</td>
<td>Penrose (2003)</td>
</tr>
<tr>
<td>Random Temporal Torus</td>
<td>RTT</td>
<td>Synthetic</td>
<td>Torus $[0, 1]^2$</td>
<td>$r, v$</td>
<td>Penrose (2003)</td>
</tr>
<tr>
<td>Random Temporal Hyperbolic</td>
<td>RTH</td>
<td>Synthetic</td>
<td>Hyperboloid with $\zeta &lt; 0$</td>
<td>$\zeta, \alpha, R, v$</td>
<td>Krioukov et al. (2010), von Looz and Meyerhenke (2018)</td>
</tr>
</tbody>
</table>

Table 1: Summary of the random temporal graph models described in the Random graph section.

The clustering coefficient divided by average path length, normalized by the corresponding measure calculated on a random graph, formally $\frac{C_g}{L_g} / \frac{C_{rand}}{L_{rand}}$. If the ratio is greater than 1, it indicates a significant small-world property of the network. The extension to temporal complex networks requires adapting the previous measures to capture the time evolution. The path length can naturally be extended to the average length of the fastest time-respecting path between each pair of nodes. Tang, Scellato, Musolesi, Mascolo, and Latora (2010) replaced the local clustering coefficient with the temporal correlation coefficient, which was also considered by Sizemore and Bassett (2018), which we will define in the next paragraph.

**Temporal clustering coefficient.** The clustering coefficient is defined as the ratio between the total number of closed triplets and the total number of triplets (open and closed). More specifically, closed triplets are sub-triangles in the graph that are counted three times, and open triplets are paths of length two, as illustrated in Figure 5.

![Figure 5: Open and closed triplets. The open triplet in this graph is the green path, instead, the orange triangle is a closed triplet.](image)

The clustering coefficient at a given time $t$ is denoted by $C(t)$ and is required to compute the temporal clustering coefficient.

$$C(t) = \frac{\sum_{i,j,k} A_{ij}(t)A_{jk}(t)A_{ki}(t)}{\sum_i k_i(k_i - 1)}$$

where $k_i = \sum_j A_{ij}(t)$ and $A_{ij}(t)$ is the value of the binary adjacency tensor of the network.
at time $t$ corresponding to nodes $i$ and $j$. The formula (1) is equivalent to the one introduced by Watt and Strogatz: the clustering coefficient is the average over all nodes of the local clustering coefficient

$$C_i = \frac{2|\{e_{jk} : v_j, v_k \in N_i, e_{jk} \in E\}|}{k_i(k_i - 1)},$$

$C_i$ measures the tendency of neighbors of a node to form a clique. Some examples are shown in Figure 6.

The temporal clustering coefficient $C$ is the average of the clustering coefficient computed at each time step:

$$C = \frac{1}{T} \sum_t C(t).$$

Figure 6: Local clustering of the orange node. In (a) the clustering coefficient is equal to zero because there are zero connections among the neighbors of the orange node, over three possible ones. In (b) there is one connection over three. In (c) there are all possible connections, hence the clustering coefficient has value 1, which is the maximum.

**Temporal path length.** The temporal path length is defined as the average of the fastest paths between all pairs of nodes. Paths are taken across time. Only one edge can be used at each time step, and pauses are allowed at nodes (i.e., an edge can be not selected to move forward in a time step). The temporal path length $L$ is thus defined as follows

$$L = \frac{1}{N(N - 1)} \sum_{i \neq j} l_{ij},$$
where \( l_{ij} \) is the minimum time needed to get from \( i \) to \( j \) starting at time 0. If there is no path between two nodes, the distance is set to the maximum distance, which is the number of time steps. Fornito et al. (2016) suggest other ways to address the latter case, such as considering only paths that exist or choosing a threshold that yields only connected graphs.

**Temporal correlation coefficient.** The temporal correlation coefficient quantifies how much the temporal structure changes over time. In particular, it shows the tendency of a node’s neighbors to remain connected to it over successive time steps. The higher the coefficient, the smaller the changes over time (Sizemore & Bassett, 2018). The temporal correlation coefficient of a node \( i \) is defined as

\[
TC_i = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{\sum_j A_{ij}(t)A_{ij}(t+1)}{\sqrt{\sum_j A_{ij}(t)\sum_j A_{ij}(t+1)}}.
\]

We can thus write the temporal correlation coefficient for the whole temporal network as

\[
TC = \frac{1}{N} \sum_i TC_i.
\]

**Temporal small-worldness.** In the static case, which corresponds to a graph that remains unchanged at each time step, the average path length is retrieved from the temporal path length. On the contrary, in the latter case, the temporal correlation coefficient does not coincide with the clustering coefficient. For this reason, the temporal correlation coefficient does not seem to properly generalize the clustering coefficient by Watts and Strogatz to the temporal setting, while our definition appears to be consistent with the latter. To illustrate the differences between the two measures, consider the following simple examples of temporal graphs with 4 time steps. In the first, shown in Figure 7a, we have a 5-node star graph that does not change over time in which the temporal correlation coefficient is 1, while the temporal clustering coefficient is 0 because there are no closed triplets. In the second case, we have a 6-node temporal graph in which a 3-node triangle subgraph alternates between two configurations, as shown in Figure 7b. In this case, the temporal correlation coefficient is 0, while the temporal clustering coefficient is 0.5. As a consequence, in our experiments, we consider and compare two versions of small-worldness:
The temporal small-worldness is the ratio between the clustering coefficient and the temporal path length, without normalization; formally

\[ S = \frac{C}{L}. \]

The Sizemore and Bassett temporal small-worldness is the ratio between the temporal correlation coefficient and the temporal path length, without normalization; in formulas

\[ S_{SB} = \frac{TC}{L}. \]

**EXPERIMENT SETUP**

In this section, we provide some technical information, such as the details of the machine on which we run the experiments, the inputs of the pipeline, and we describe the optimization methods we used to choose the model parameters.

*Computing Platform*
We run the pipeline on the Nef Cluster Computing Platform of Inria Sophia Antipolis. By setting the parameters described below, the average execution time of our pipeline for a subject is 250 seconds, on a 64-core node with the following technical characteristics:

- AMD dual-EPYC 7542 processors,
- 1024 GB of RAM.

**Pipeline**

The pipeline is implemented in Python. As explained in the Method Section, it transforms the BOLD signal into a temporal graph. The code, which is available at the following link

https://github.com/aurorarossi/TemporalBrainNetworksCode, takes various parameters as input, which are described in the next sections.

**Atlas.** Regarding the type of atlas, it is possible to choose among Gallardo (2018), Schaefer et al. (2018) or Glasser et al. (2016). For the first two, it is possible to select the number of regions into which the brain is divided, starting from 102 nodes up to a maximum of 1002 nodes. The Glasser atlas instead has 362 regions. In our experiment, we chose the Schaefer atlas with 302 nodes. Using an atlas with a higher number of regions could potentially introduce errors due to the small size of the regions, which could result in regions not matching across different subjects.

**Windows.** Several parameters describe the windows, such as the seconds corresponding to the window length and the window overlap, which we chose to be 60 and 30 seconds, respectively. Among the different types of windows, we chose to work with rectangular windows.

**Correlation.** As correlation measure we chose the classical Pearson’s correlation.

**Thresholds.** Since there was no consensus on the most appropriate threshold, we consider different values. Specifically, we consider all the thresholds from 0.2 to 0.9 with a step of 0.05 and from 0.9 to 0.98
with a step of 0.02. The output is a binary temporal network. The latter is sparse for high threshold values and dense for low values.

**Generators and measures**

All temporal graph generators and measures are written in the Julia programming language. In addition to optimizing the execution of serial code, we applied parallel programming techniques such as multithreading, which were exploited to make the scripts efficient. These scripts take as input 532 gigabytes of data generated by the pipeline starting from 975 gigabytes of raw data.

**Optimization: fitting parameters**

After creating the temporal graphs with the pipeline, our goal is to compare them with the random temporal models, which all depend on some parameters. The latter should be chosen to share as much as possible some common properties with the real data. In particular, we aim to have the same number of nodes, average degree, and similar temporal small-worldness values. We thus choose the set of parameters for which the random temporal models best match the aforementioned properties in the empirical data. In the RTPT and RTE Graph Models, the number of nodes and the average degree are inherited from the real data by construction and they do not depend on other parameters. For the RTS and RTT Graph Models, the number of nodes can be set directly and, by changing the connection parameter $r$, it is possible to obtain any average degree, instead, the velocity parameter $v$ has to be optimized. For the RTH Graph Model, the number of nodes has to be set, the average degree is obtained by changing $R$ and the parameters to be optimized are $\alpha$ and $v$, while $\zeta$ is set to 1, because, as explained in the Methods Section, we can fix $\zeta$ and vary just $\alpha$. We used two different optimization methods to find the best values of the parameters.

The first finds a set of parameters for the three random temporal models that would fit the small-worldness values of the real data for some average degree values in the interval $0 - 170$ (see Figure 8). To obtain different average degrees for the real data networks we use different threshold values, the higher the threshold value, the lower the average degree. For the other models, we mention how to set it in the previous paragraph.
The second optimization method is to select a set of parameters for each random model that minimizes the area between the temporal small-worldness curves of the empirical data and the random temporal model over all different average degrees. The curves are obtained by linear interpolation of the temporal small-worldness values at different average degrees (see Figure 9). For the RTH Graph Model, two different sets of parameters are shown, one optimizing each definition of temporal small-worldness, while for the RTS and RTT Graph Models, only one set of parameters is shown, corresponding to the result of optimization with respect to our definition of temporal small-worldness.

RESULTS

This section discusses the result of all the temporal measures described in this work, applied to the processed data from the Human Connectome Project (WU-Minn HCP 1200 Subjects Data Release Reference Manual, 2018) and to the random temporal models. We have calculated the values over different average degrees, namely in the range \(0 - 170\). Those of the empirical data vary by changing the threshold values. Consequently, this happens also for RTPT and RTE graph models. As for the RTS and RTT graph models, this property depends on the radius of the connection: the higher the radius, the higher the average degree. In the hyperbolic case, as shown in the section above, it depends on \(R\), which is the connection radius.

The plots in Figure 8 show the results of the first optimization method. The real data of all 1047 patients are shown in blue, the RTH Graph Model in red, and the RTT and RTS graph models in orange and green, respectively. Figure 8a shows that for any average degree value we can find a set of parameters for the RTH Graph Model that can approximate the real data value of our definition of temporal small-worldness, while this is not the case for RTT and RTS graph models. Figure 8b shows that for any average degree value we can find a set of parameters for all the random models that can approximate the real data value of the temporal small-worldness value of Sizemore and Bassett (2018).

In the plots of Figure 9, as above, the real data are shown in blue. RTT and RTS graph models are shown in orange and green, respectively. RTPT and RTE graph models are shown in grey and pink. There are also two RTH graph models with different parameters. The one with \(\alpha = 0.65, v = 0.6\) and \(\zeta = 1\), illustrated in violet, better matches the small-worldness \(S\) value of the real data, while the one with \(\alpha = 0.8, v = 0.9\) and \(\zeta = 1\), illustrated in red, better fits the definition of
small-worldness $S_{SB}$, which is based on the temporal correlation coefficient $TC$. For all models, the lines are the medians and the shadows correspond to the values between the first and third quartiles. Looking at the random models, the median of the RTPT Graph Model always has higher values with respect to those of the RTE Graph Model for all measures except the temporal path length (see Figure 9c). Note that for RTT, RTS and RTH graph models the shadows are barely visible due to the fact that the variance is very small. The quantitative results describing the plots 9a and 9b are given in Table 2 and Table 3, respectively. These tables illustrate the comparison between the median of the empirical data and that of the models. This comparison is made in two ways: by measuring the area between the curves, as explained in the Optimization subsection, and by using different types of norms, namely minimum, maximum, mean absolute, and mean Euclidean distance norms. The latter are computed across the differences between the linear interpolation of the empirical data and the model data values of $S$ and $S_{SB}$. These differences are computed on 16861 equispaced points in the range $[0, 170]$. The minimum captures the smallest difference between the curves, while the maximum captures the largest. The mean absolute value norm considers the absolute value of all differences at the selected points and then calculates their average. Analogously, the mean Euclidean distance norm calculates the Euclidean distance between the curves at the selected points and then calculates their average. The best results are
highlighted in bold and the second best is underlined. In Table 2 we can see that the RTH Graph Model with \( \alpha = 0.65, v = 0.6 \) is the one that best approximates the small-worldness \( S \) value of the real data, and the second best one is the other RTH Graph Model. In Table 3 we can see that the RTH Graph Model with \( \alpha = 0.8, v = 0.9 \) is the one that best approximates the small-worldness \( S_{SB} \) value of the real data, while the second best one is the RTPT Graph Model.

In all plots, the situation is more noisy in the sparse regime. For the sake of completeness, we did not remove this part even though these regimes are rarely considered in the literature in neuroscience.

<table>
<thead>
<tr>
<th>Model</th>
<th>Area</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean Absolute value</th>
<th>Mean Euclidean distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTH ( \alpha = 0.8, v = 0.9 )</td>
<td>2.84</td>
<td>0.0000014</td>
<td>0.067</td>
<td>0.017</td>
<td>0.00018</td>
</tr>
<tr>
<td>RTH ( \alpha = 0.65, v = 0.6 )</td>
<td>2.73</td>
<td>0.0000011</td>
<td>0.040</td>
<td>0.016</td>
<td>0.00015</td>
</tr>
<tr>
<td>RTS ( v = 0.1 )</td>
<td>8.02</td>
<td>0.0052702</td>
<td>0.166</td>
<td>0.048</td>
<td>0.00053</td>
</tr>
<tr>
<td>RTT ( v = 0.2 )</td>
<td>9.88</td>
<td>0.0000023</td>
<td>0.169</td>
<td>0.059</td>
<td>0.00057</td>
</tr>
<tr>
<td>RTE</td>
<td>33.92</td>
<td>0.0115314</td>
<td>0.248</td>
<td>0.201</td>
<td>0.00160</td>
</tr>
<tr>
<td>RTPT</td>
<td>21.42</td>
<td>0.0128656</td>
<td>0.166</td>
<td>0.127</td>
<td>0.00102</td>
</tr>
</tbody>
</table>

Table 2: Distances between the median of the empirical data and the one of the model for the temporal small-worldness \( S \), as discussed in Result Section. The best results are highlighted in bold and the second best is underlined.

DISCUSSION AND CONCLUSIONS

The main focus of this work is to investigate a novel appropriate null model for testing hypotheses about brain dynamics. A good null model must have similar properties to the empirical data under consideration. In this context, the two most important properties present in complex networks such as those modeling brain activity are small-worldness and scale-freeness. In this work, we focus on the small-worldness property.

To transform the data into temporal networks, the empirical data is collected in time and processed using the window sliding method. The main challenge is to adapt all definitions of measures and models
Figure 9: Results of the second optimization method: comparison between empirical and random networks with 302 nodes for different measures.
<table>
<thead>
<tr>
<th>Model</th>
<th>Area</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean Absolute value</th>
<th>Mean Euclidean distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTH $\alpha = 0.8$, $v = 0.9$</td>
<td>4.46</td>
<td>0.0000006</td>
<td>0.067</td>
<td>0.026</td>
<td>0.00024</td>
</tr>
<tr>
<td>RTH $\alpha = 0.65$, $v = 0.6$</td>
<td>9.60</td>
<td>0.0093862</td>
<td>0.096</td>
<td>0.057</td>
<td>0.00050</td>
</tr>
<tr>
<td>RTS $v = 0.1$</td>
<td>37.23</td>
<td>0.0000018</td>
<td>0.266</td>
<td>0.221</td>
<td>0.00174</td>
</tr>
<tr>
<td>RTT $v = 0.2$</td>
<td>17.91</td>
<td>0.0136169</td>
<td>0.130</td>
<td>0.106</td>
<td>0.00084</td>
</tr>
<tr>
<td>RTE</td>
<td>20.56</td>
<td>0.0396405</td>
<td>0.142</td>
<td>0.122</td>
<td>0.00095</td>
</tr>
<tr>
<td>RTPT</td>
<td>7.25</td>
<td>0.0000087</td>
<td>0.060</td>
<td>0.043</td>
<td>0.00035</td>
</tr>
</tbody>
</table>

Table 3: Distances between the median of the empirical data and the one of the model for the temporal small-worldness $S_{SB}$, as discussed in Result Section. The best results are highlighted in bold and the second best is underlined.

to their temporal versions. In particular, we provide a new definition of temporal small-worldness that is consistent with its static version: if the time steps of the temporal network are all identical, the temporal small-worldness is equal to the static small-worldness.

We compare five different temporal models by looking at how close their temporal small-worldness values are to the ones of the data at different thresholds. In Figure 8 we compare the model at each threshold individually, while in Figure 9 we compare them across different thresholds. The two models introduced by Sizemore and Bassett (2018) perform worse than the other three models for the $S$ measure. This may be not surprising given the fact that they do not have any parameters to optimize but depend only on the data. The RTS and RTT graph models perform better than the other two models, and the measures computed on them are consistently very close to each other. This similarity may be explained by the fact that the two graphs look locally the same for all points that are not close to the border of the square or the torus, since the connectivity only differs in the latter area (see Figure 3). The RTH Graph Model is the one that not only achieves the closest values of the temporal small-worldness, but its curve of values follows the same trend as that of empirical data across different average degrees. This result is consistent with the fact that RTH is a popular complex network model that exhibits both small-worldness and scale-freeness.
This analysis could be similarly extended to the more challenging investigation of a temporal version of the scale-freeness property. It would be interesting to see if the RTH Graph Model is still preferable. In this work, we focus on the resting state fMRI in healthy patients. A further development could be the comparison with patients with diseases that affect the functionality of the whole brain. For example, we could consider diseases such as Alzheimer’s and Parkinson’s but also ADHD, depression and schizophrenia. Again, the main question is whether the best model to fit the data remains the same across different parameter ranges. Another condition that could be investigated is the case of non-resting state, where patients are engaged in cognitive processes generated by tasks involving memory, learning, and attention. On the null models side, as future work, one could consider the case in which one starts from phase randomization and autoregressive randomization null data of brain dynamics, transform them in temporal networks and then apply the same analysis as in this work. This would allow to compare the results of the two different approaches (Liégeois, Laumann, Snyder, Zhou, & Yeo, 2017; Liégeois, Yeo, & Van De Ville, 2021).

ACKNOWLEDGMENTS

This work has been supported by the French government, through the UCA DS4H Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-17-EURE-0004. The authors are grateful to the OPAL infrastructure from Université Côte d’Azur for providing resources and support.

REFERENCES


95. Retrieved 2023-10-30, from
https://linkinghub.elsevier.com/retrieve/pii/S1053811914001815 doi: 10.1016/j.neuroimage.2014.03.034


TECHNICAL TERMS

- *Temporal graph*: a graph in which edges and nodes change over time. It can be represented as a sequence of graphs, each corresponding to a time step of the network at a given time.

- *Lattice graph*: a graph which contains a regular pattern.

- *Null model*: a theoretical model used to test hypotheses about empirical data, sharing certain properties with them.

- *Node degree*: number of edges incident to a node.

- *High tail distribution*: a distribution that presents a significant number of extreme events.

- *Brain atlas*: subdivision of the gray matter of the brain into smaller areas that share a common property.

- *Adjacency matrix*: a square matrix representing the graph structure, where the indices of rows and columns represent the nodes and the entries correspond to the edge weight between the respective nodes.

- *Time-respecting path*: a finite or infinite sequence of distinct edges, ordered in time, connecting a set of distinct nodes.

- *Threshold*: a criterion value that determines whether values above it are taken into account and values below it are not.

- *Binary temporal graph*: a temporal graph whose edges have weights of 1 or 0.