New Journal of Physics
The open access journal at the forefront of physics
Deutsche Physikalische Gesellschaft DPG | IOP Institute of Physics

Brain networks: small-worlds, after all?

Lyle Muller, Alain Destexhe and Michelle Rudolph-Lilith
Unité de Neurosciences, Information et Complexité (UNIC), Centre National de la Recherche Scientifique (CNRS), 1 Avenue de la Terrasse, Gif-sur-Yvette, France
E-mail: rudolph@unic.cnrs-gif.fr

Received 21 April 2014, revised 26 June 2014
Accepted for publication 4 August 2014
Published 6 October 2014
New Journal of Physics **16** (2014) 105004
doi:10.1088/1367-2630/16/10/105004

Abstract
Since its introduction, the 'small-world' effect has played a central role in network science, particularly in the analysis of the complex networks of the nervous system. From the cellular level to that of interconnected cortical regions, many analyses have revealed small-world properties in the networks of the brain. In this work, we revisit the quantification of small-worldness in neural graphs. We find that neural graphs fall into the 'borderline' regime of small-worldness, residing close to that of a random graph, especially when the degree sequence of the network is taken into account. We then apply recently introducted analytical expressions for clustering and distance measures, to study this borderline small-worldness regime. We derive theoretical bounds for the minimal and maximal small-worldness index for a given graph, and by semi-analytical means, study the small-worldness index itself. With this approach, we find that graphs with small-worldness equivalent to that observed in experimental data are dominated by their random component. These results provide the first thorough analysis suggesting that neural graphs may reside far away from the maximally small-world regime.

S Online supplementary data available from stacks.iop.org/njp/16/105004/mmedia

Keywords: brain networks, small-worlds, graph theory, random graphs

1. Introduction

The small-world network model of Watts and Strogatz is quite possibly the pinnacle of network science. In their seminal work [1], the authors showed that there exists an intermediate state between regular and random graphs, where high clustering of network connections and low distance between any two nodes coexist, and that this intermediate 'small-world' state is reached by random rewiring of only a few connections in a regular ring graph. Because of these dual properties stemming from the structure of a regular graph combined with the efficiency of a random graph, the pertinence of the small-world network model to real-world systems, both natural and technological, was immediately clear.

Since that time, indications of the small-world property have been uncovered in a vast array of real-world datasets. From the world-wide web [2] to the electric power grid [3], social networks [4], and various neural systems [5], network connectivity has often been observed to exhibit clustering exceeding that of a random graph, while simultaneously maintaining a low average number of links between nodes. While in many networks the small-world property may be strikingly clear in comparison to a random graph, however, in some cases the difference may not be particularly large, especially in the case of small systems. For this reason, a ratio-based index that quantifies the 'small-worldness' of a graph was recently introduced [6], where values beyond unity indicate the presence of the small-world property.

Immediately upon the introduction of this measure, it was apparent that, while systems such as the electric power grid or the co-authorship network of mathematicians could exhibit a small-worldness index ranging from 100 even to 10 000 and above, certain other systems fall into a 'borderline' category, defined by small-worldness indices ranging between 1 and 3 [6]. Among these borderline cases is the neural connectivity diagram of the macaque cortex (MC), and further, since the introduction of this index, several neural graphs have been reported to fall within this regime [7, 8], including the area-level connectivity diagram of the human cerebral cortex [9].

Here, motivated by interest in the small-world property and the meaning of this borderline regime, we present a comprehensive analysis of small-worldness in neural graphs. The connectivity diagrams used for this analysis span in scale from the microscopic neural connectivity of *C. elegans* to the macroscopic connectivity of cortical regions in the monkey brain, and in size from a few tens to several hundred nodes. We quantify clustering and distance in the neural graphs using several available measures, in order to consider all relevant formulations for the small-worldness index in this analysis. Within this multifaceted approach, a consistent result nevertheless emerges: neural graphs fall in general into the borderline small-worldness regime, particularly when the degree sequence of the network is taken into account.

To investiate the implications of this borderline regime, we employ a recently introduced exact expression for the clustering coefficient of small-world and random graphs [10] and an excellent analytical approximation for the average geodesic graph distance [11]. With these expressions, we derive theoretical bounds for the maximal and minimal small-worldness index of graph models constrained by neural data. Furthermore, through a semi-analytical approach, we study the behavior of the small-worldness index directly. To reproduce the small-worldness observed in the data, we find that corresponding graph models require a dominant random rewiring, in all cases one order of magnitude larger than the small rewiring first noted to bring a regular graph into the maximally small-world regime [1]. Finally, we discuss these results and their implications for the study of network structure in the nervous system.

2. Graph theory preliminaries

Mathematically, a (relational) graph or network is a pair $G = (\mathcal{N}, \mathfrak{a})$, where $\mathcal{N}$ denotes a set of N_N nodes (vertices) and $\mathfrak{a}$: $\mathcal{N} \times \mathcal{N} \longrightarrow \{0, 1\}$ an adjacency relation which assigns to each pair of nodes (i, j) a number $a_{ij} \in \{0, 1\}$. If $a_{ij} = 1$, the nodes i and j are said to be adjacent (connected) to each other, and the adjacency relation defines an edge between both nodes. For $a_{ij} = 0$, no link (edge) exists between nodes i and j. If the relation $\mathfrak{a}$ is symmetric, i.e. $\forall\ (i, j)$: $a_{ij} = a_{ji}$, the graph is said to be undirected. A graph is directed, if at least one node pair (i,j) exists for which $a_{ij} \neq a_{ji}$. For such digraphs, edges are endowed with a direction pointing from a source node to a target node. If $a_{ii} = 1$, node i is said to be self-looped. A graph with at least one self-looped node is called self-looped graph.

The adjacency relations a_{ij} form a matrix, which provides a complete description of a relational graph's structure. This adjacency matrix defines a whole host of graph theoretical measures, with which a given graph can be further characterized (e.g, see [12, 13] for comprehensive reviews). The connectedness (or connectance) Co, a measure of relative graph connectivity, of a not self-looped graph is defined as

$$Co = \frac{1}{N_N\left(N_N - 1\right)} \sum_{i,j=1}^{N_N} a_{ij}, \tag{1}$$

with $0 \leqslant Co \leqslant 1$. The node in- and out-degrees for digraphs are defined as

$$a_i^{\{\text{in, out}\}} = \sum_{j=1}^{N_N} a_{\{ji,ij\}}, \tag{2}$$

and the average node degree as

$$\left\langle a_i \right\rangle = \frac{1}{N_N} \sum_{i=1}^{N_N} a_i^{\text{in}} = \frac{1}{N_N} \sum_{i=1}^{N_N} a_i^{\text{out}}. \tag{3}$$

A digraph is comprised of a number of strongly connected components, that is pairwise disjoint subgraphs in which each node can reach and is reachable by any other node in the subgraph. More specifically, a strongly connected component is defined as a subgraph consisting of a set of nodes from which all other nodes in the subgraph can be reached, and which can be reached from all other nodes, by following existing edges. The largest of these strongly connected components is called giant strongly connected component. The analysis in this paper will will restrict to the giant strongly connected components of neural digraphs.

3. Analysis methods

In this paper, we study structural aspects of a number of publicly available biological neural graphs used in the literature in the past two decades. These data include areal connectivity graphs of the cat cortex (CC1 and CC2), the macaque brain (MB1), MC1 and MC2, macaque visual cortex (MVC1 and MVC2) and macaque neocortex (MNC1), as well as the neuronal connectivity graph of the nematode *C. elegans* (CE1, CE2 and CE3). References for each data source are provided in supplementary table 1. Brief descriptions of the circuit reconstruction

Table 1. Geodesic distance analysis across the neural graphs. Shown are values for the number of connected node pairs N_C, the geodesic graph distance D, graph radius r and diameter d, (equations (7)–(9)), average geodesic graph and node distance ($\langle d \rangle$ and $\langle d_i \rangle$, respectively; equations (10) and (11)), as well as the characteristic path length L, equation (12), and geodesic connectivity length ℓ, equation (13).

Geodesic distance analysis								
	N_C	r	d	D	$\langle d \rangle$	$\langle d_i \rangle$	L	ℓ
CC1	9025	2	4	17359	1.9234	182.7263	1.8842	1.6894
CC2	2704	2	4	4899	1.8118	94.2115	1.8077	1.5814
CE1	57121	4	14	227938	3.9904	953.7155	3.1674	3.1299
CE3	75076	4	7	215363	2.8686	785.9964	2.8376	2.5700
MB1	123201	3	6	322355	2.6165	918.3903	2.5385	2.3513
MC1	4900	3	5	11296	2.3053	161.3714	2.2143	1.9756
MC2	7225	3	5	13453	1.8620	158.2706	1.8000	1.5794
MNC1	2209	3	4	4537	2.0539	96.5319	2.0851	1.7551
MVC1	900	2	4	1553	1.7256	51.7667	1.7333	1.5128
MVC2	900	2	4	1553	1.7256	51.7667	1.7333	1.5128

techniques, which range from electron microscopy of individual neurons and synaptic contacts in *C. elegans* to tract-tracing studies at the level of cortical regions in the monkey brain, are provided in the supplementary information (section 2, 'Historical' neural graphs). Note that some of these graphs have experienced slight modifications and refinement since their first investigation. Throughout this study, we analyze the giant strongly connected component of the network (see supplementary information, section 3, connected components).

3.1. Numerical tools

Numerical analysis was performed using the custom software *cygraph* and *Mathematica*. A *cygraph* binary (Mac OSX), all graph data, and analysis protocols are available for download [14].

3.2. Network models

The assessment of the small-world properties of the investigated biological neural graphs requires their comparison with equivalent random graphs [6, 15]. To that end, we constructed surrogate randomized networks with the same number of nodes (N_N), number of edges (N_E), and node in- and out-degrees (a_i^{in} and a_i^{out}, respectively) as in the neural data using the well-known edge swapping algorithm of Maslov and Sneppen (MS) [16]. All results presented here utilize this algorithm for generating degree-matched randomized networks. Because the sampling uniformity and convergence properties of the MS algorithm have been subject to recent critique [17, 18], we next utilized an algorithm for sampling from the ensemble of exact degree-matched Erdős–Rényi (ER) (EDM) random graphs introduced in [17, 19]. The results obtained with the EDM algorithm showed excellent quantitative agreement to those obtained with the MS (data not shown). Finally, to quantitatively understand the effect of the network degree distribution on the small-worldness index, we constructed ER random graphs with the same number of nodes and edges as in the neural data, without taking the degree-distribution

into account (data not shown). Also here, results showed qualitiative agreement with the presented findings, although quantiative differences occur when the specific degree distribution of the network is not taken into account.

In order to compare the obtained results with a classical model of small-worlds, we constructed Watts–Strogatz (WS) graph models [1] with the same number of nodes and initial node degree to match the number of edges in the network. To construct WS digraphs, we first constructed appropriately matched undirected WS graphs, after which the direction of each edge was randomized with a probability of 0.5. When considering the MS, EDM, ER, and WS network models, 1000 random realizations were used at each point in parameter space to ensure statistical stability.

4. Results

In the following sections, we first present results of the geodesic distance and clustering analysis of the considered neural graphs and degree-preserved randomized counterparts. These results are then used to quantitatively assess the small-worldness index (section 4.3). We conclude in section 4.4 by applying analytical expressions for these network measures to derive theoretical bounds for these quantities and conducting a semi-analytical calculation for the small-worldness of WS graphs matched to the neural data, expressed as a function of the rewiring parameter.

We note that the analyses presented here greatly expand upon the methodology introduced in [6]. First, we consider neural graphs in their original, directed form. Secondly, we compare the clustering and distance measures observed in the neural graphs to those in matched random graphs which take the node degree distribution into account. Thirdly, we employ recently introduced analytical expressions for the clustering in small-world and random graphs that are exact and valid non-asymptotically, and an analytical approximation for the distance in random graphs directly verified to hold for the parameter range considered here. Finally, the distance in small-world graphs, for which no valid analytic approximation exists outside a narrow regime of small rewiring, is evaluated numerically and with high precision, to complete a rigorous evaluation of the small-worldness property in neural graphs.

4.1. Geodesic distances

First, we performed a classical geodesic distance analysis. In the past, various measures of characteristic distance in relational graphs were proposed (for a review of various measures, see [12, 13]), all based on the geodesic distance d_{ij} between two nodes i and j. Here, d_{ij} is defined as the number of edges in the shortest path connecting two nodes, and the set of d_{ij} forms a matrix called the (geodesic) distance matrix. Typically, unconnected nodes are assigned a geodesic distance of infinity, but as we restrict our analysis to the (strongly) connected components of each graph, no unconnected pair of nodes can occur. The diagonal elements d_{ii} of the distance matrix are called geodesic loops, or closed geodesics. In looped graphs, the closed geodesics of a self-looped node i (i.e. a node for which $a_{ii} = 1$) is the shortest nontrivial path which connects the node with itself. This definition contrasts the commonly used notion, in which d_{ii} of self-looped nodes is set to 0. The analysis presented in this paper will include closed geodesics of all nodes.

The geodesic node distance matrix d_{ij} is a $N_N \times N_N$ matrix from which a number of measures can be conceived. First, the geodesic node in- and out-distance degree (also

sometimes called distance strength) $d_i^{\{\text{in, out}\}}$ is defined for directed graphs as

$$d_i^{\{\text{in, out}\}} = \sum_{j=1}^{N_N} d_{\{ji,ij\}} \tag{4}$$

and, due to the symmetry of the geodesic distance matrix, for undirected graphs by

$$d_i = d_i^{\text{in}} = d_i^{\text{out}} = \sum_{j=1}^{N_N} d_{ij}. \tag{5}$$

If closed geodesics are excluded, the sum in the above equations runs over $j \neq i$.

Secondly, the maximal geodesic distance between a given node i and any other node in the graph

$$\varepsilon_i = \max_j \left\{ d_{ij} \right\}, \tag{6}$$

defines the geodesic node eccentricities. In the case closed geodesics are excluded from the analysis, the maximum is evaluated over all $d_{ij}, j \neq i$ for a given i. With these measures, one has

$$D = \begin{cases} \sum_{i,j=1}^{N_N} d_{ij} & \text{closed geodesics included} \\ \sum_{\substack{i,j=1 \\ i \neq j}}^{N_N} d_{ij} & \text{closed geodesics excluded,} \end{cases} \tag{7}$$

$$r = \min_i \{ \varepsilon_i \}, \tag{8}$$

$$d = \begin{cases} \max_{i,j} \left\{ d_{ij} \right\} & \text{closed geodesics included} \\ \max_{\substack{i,j=1 \\ i \neq j}} \left\{ d_{ij} \right\} & \text{closed geodesics excluded,} \end{cases} \tag{9}$$

as the geodesic graph distance, geodesic graph radius and diameter, respectively. Furthermore

$$\langle d \rangle = \frac{D}{N_C}, \tag{10}$$

$$\left\langle d_i \right\rangle = \frac{D}{N_N}, \tag{11}$$

define the average geodesic graph and node distance, respectively, where N_C denotes the number of connected node pairs in a given graph. In the case of (strongly) connected components, $N_C = N_N^2$ in the case closed geodesics are included, and $N_C = N_N (N_N - 1)$ if closed geodesics are not considered.

Finally, the characteristic geodesic path length is given by the median of the means of the shortest geodesic distances connecting each node to all other nodes, i.e.

$$L = \underset{i}{\text{median}}\left\{\frac{1}{N_N} d_i\right\}, \tag{12}$$

and the geodesic connectivity length is defined by the harmonic mean of geodesic node distances

$$\ell = N_C \left\{\sum_{i,j=1}^{N_N} \frac{1}{d_{ij}}\right\}^{-1}. \tag{13}$$

If closed geodesics are excluded from the analysis, the sum in the last equation runs over all i, j with $i \neq j$. Results of the application of the above measures to the considered biological neural graphs are summarized in table 1.

In accordance with expectation, the various geodesic distance measures will decrease for increasing connectedness of a graph. This is illustrated in figure 1, which shows the characteristic geodesic path length, geodesic connectivity length and average geodesic graph distance as function of the graph connectedness Co (figures 1(a)–(c), respectively) for both the neural graphs (green dots) and their degree-preserved randomized counterparts (black dots). A rough theoretical estimate for the average graph distance in ER random graphs is known [20], and given by

$$\langle d \rangle_{\text{ER}}^{\text{theor}} = \frac{\ln\left(N_N\right)}{\ln\left(\langle a_i \rangle\right)}, \tag{14}$$

where $\langle a_i \rangle$ denotes the average node degree. Excluding for simplicity self-loops, the average node degree of a classical ER random graph is $\langle a_i \rangle = Co\,(N_N - 1)$, which yields

$$\langle d \rangle_{\text{ER}}^{\text{theor}} = \frac{\ln\left(N_N\right)}{\ln\left(Co\left(N_N - 1\right)\right)} \tag{15}$$

for the expected average geodesic graph distance. This theoretical prediction qualitatively matches with the results found for the biological neural graphs when considering $\langle d \rangle$ as a function of connectedness, but generally underestimate the observed values (figure 1(c), compare dots with upward triangles). Recently, a more accurate theoretical prediction of $\langle d \rangle$ for classical ER random graphs was provided, which takes the form [11]

$$\langle d \rangle_{\text{ER}}^{\text{theor}} = \frac{\ln\left(N_N\right) - \gamma}{\ln\left(Co \cdot N_N\right)} + \frac{1}{2}, \tag{16}$$

where γ denotes the Euler–Mascheroni constant. Comparing the numerical results of equation (16) with the average geodesic graph distance for the corresponding biological neural graphs shows an excellent agreement both qualitatively and quantitatively (figure 1(c), compare dots with downward triangles). Moreover, this agreement provides further support for the recent finding that the structural makeup of small neural graphs is dominated by a classical ER random component constrained only by a non-scale-free node-degree distribution [21, 22].

To further quantify the comparison between the observed and randomized graphs, we defined the following ratios [6]

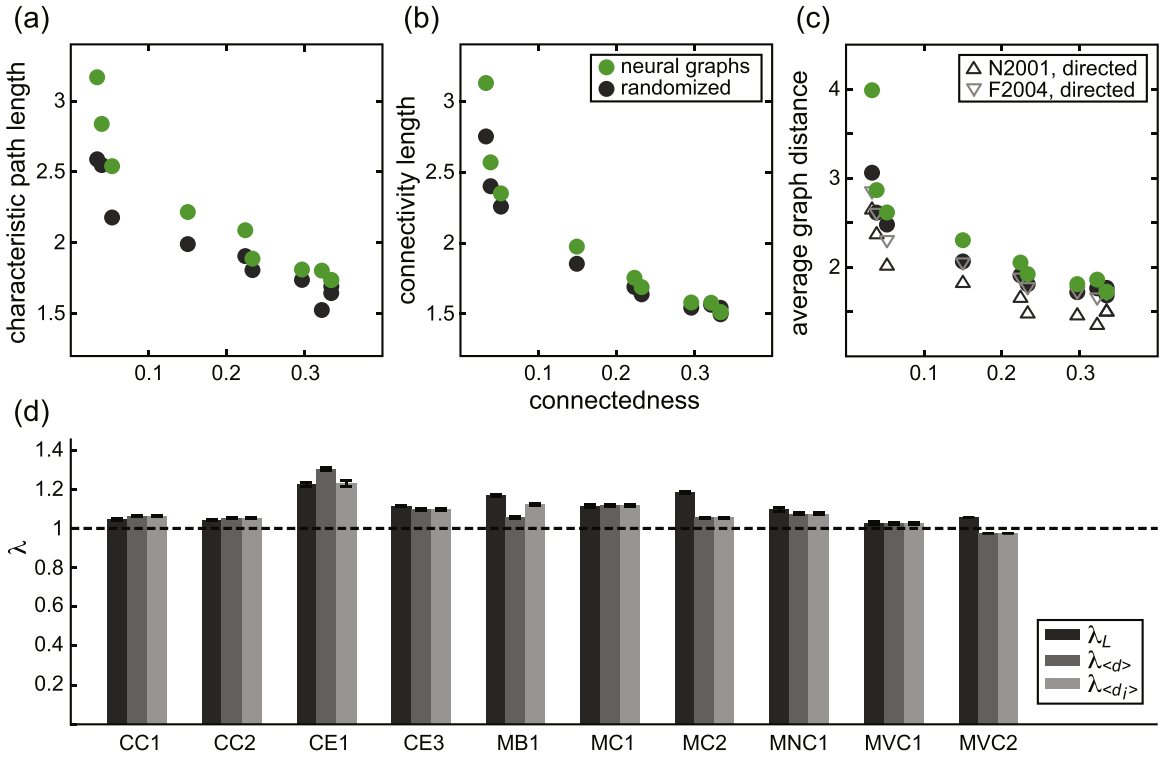

Figure 1. Geodesic distances in neural graphs. (a)–(c) Characteristic geodesic path length L (equation (12)), geodesic connectivity length ℓ (equation (13)) and average geodesic graph distance $\langle d \rangle$ (equation (10)) as function of graph connectedness Co (green) and their degree-preserved randomized counterparts (black). Theoretical predictions for the average geodesic graph distance are displayed in grey (upward triangles: equation (15) from [20]; downward triangles: equation (16) from [11]). (d) Ratio (mean ±SD) λ_L (equation (17); black), $\lambda_{\langle d \rangle}$ (equation (18); dark grey) and $\lambda_{\langle d_i \rangle}$ (equation (19); light grey) between the characteristic geodesic path length L, average geodesic graph distance $\langle d \rangle$ and average geodesic node distance $\langle d_i \rangle$ of the considered biological graphs and their corresponding degree-preserved randomized counterparts, respectively.

$$\lambda_L = \frac{L}{L_{\mathrm{MS}}}, \tag{17}$$

$$\lambda_{\langle d \rangle} = \frac{\langle d \rangle}{\langle d \rangle_{\mathrm{MS}}}, \tag{18}$$

$$\lambda_{\langle d_i \rangle} = \frac{\langle d_i \rangle}{\langle d_i \rangle_{\mathrm{MS}}} \tag{19}$$

between the characteristic geodesic path length L, average geodesic graph distance $\langle d \rangle$ and average geodesic node distance $\langle d_i \rangle$ of a given graph and its corresponding randomized surrogate (L_{MS}, $\langle d \rangle_{\mathrm{MS}}$ and $\langle d_i \rangle_{\mathrm{MS}}$, respectively). In all cases and for all considered biological neural graphs, we observe λ-values very close to 1 (figure 1(d); $1.03 \leqslant \lambda_L \leqslant 1.22$, $0.98 \leqslant \lambda_{\langle d \rangle} \leqslant 1.30$ and $0.98 \leqslant \lambda_{\langle d_i \rangle} \leqslant 1.23$). In section 4.3, the various λ ratios defined in equations (17)–(19) will be used to assess the small-worldness of the investigated graphs.

4.2. Clustering

Secondly, we investigated the clustering properties of the neural graphs. The notion of 'clustering' was first introduced as 'transitivity' in studies of social graphs to measure the extent to which nodes are linked together in triangular relationships [23]. Specifically, three nodes A, B, and C are said to be transitive if, for existing links between nodes A and B and nodes B and C, nodes A and C are also linked. Mathematically, the graph transitivity T can be defined as the ratio between the number of closed paths of length two and the total number of paths of length two [13], i.e.

$$\begin{aligned} T &= \frac{\#\text{ closed paths of length 2}}{\#\text{ paths of length 2}} \\ &\equiv \frac{3\times\ \#\text{ triangles}}{\#\text{ connected triples}}, \end{aligned} \tag{20}$$

with $0 \leqslant T \leqslant 1$. In a similar fashion, the node clustering coefficient is defined as [1, 12]

$$\begin{aligned} C_i &= \frac{2\times\#\text{ edges in subgraph of neighbors of node } i}{a_i\left(a_i-1\right)} \\ &= \frac{1}{a_i\left(a_i-1\right)}\sum_{\substack{j=1\\ j\neq i}}^{N_N}\sum_{\substack{h=1\\ h\neq i,j}}^{N_N} a_{ij}a_{ih}a_{jh} \end{aligned} \tag{21}$$

($0 \leqslant C_i \leqslant 1$), and measures how likely $a_{mn} = 1$ for two neighbors m, n of node i, that is the ratio between the number of existing edges in the subgraph spanned by the neighbors of a node i and the number of possible edges present within this subgraph. The average node clustering coefficient is then given by

$$\left\langle C_i\right\rangle = \frac{1}{N_N}\sum_{i=1}^{N_N} C_i. \tag{22}$$

The clustering measures in equations (20)–(22) are defined for undirected graphs only, but can be generalized to hold for digraphs as well. Here, however, the edge direction has to be taken into account. For the single triangular relationship which defines graph transitivity, a total of eight relationships can be drawn for digraphs (figure 2(a)). From these, four equivalent pairs exist, leaving four distinct definitions of global clustering coefficients for digraphs

$$C^{(000)} = \frac{T_{\Delta}^{(000)}}{T_{<}^{(01)}}, \tag{23}$$

$$C^{(001)} = \frac{T_{\Delta}^{(001)}}{T_{<}^{(00)}}, \tag{24}$$

$$C^{(010)} = \frac{T_{\Delta}^{(010)}}{T_{<}^{(01)}}, \tag{25}$$

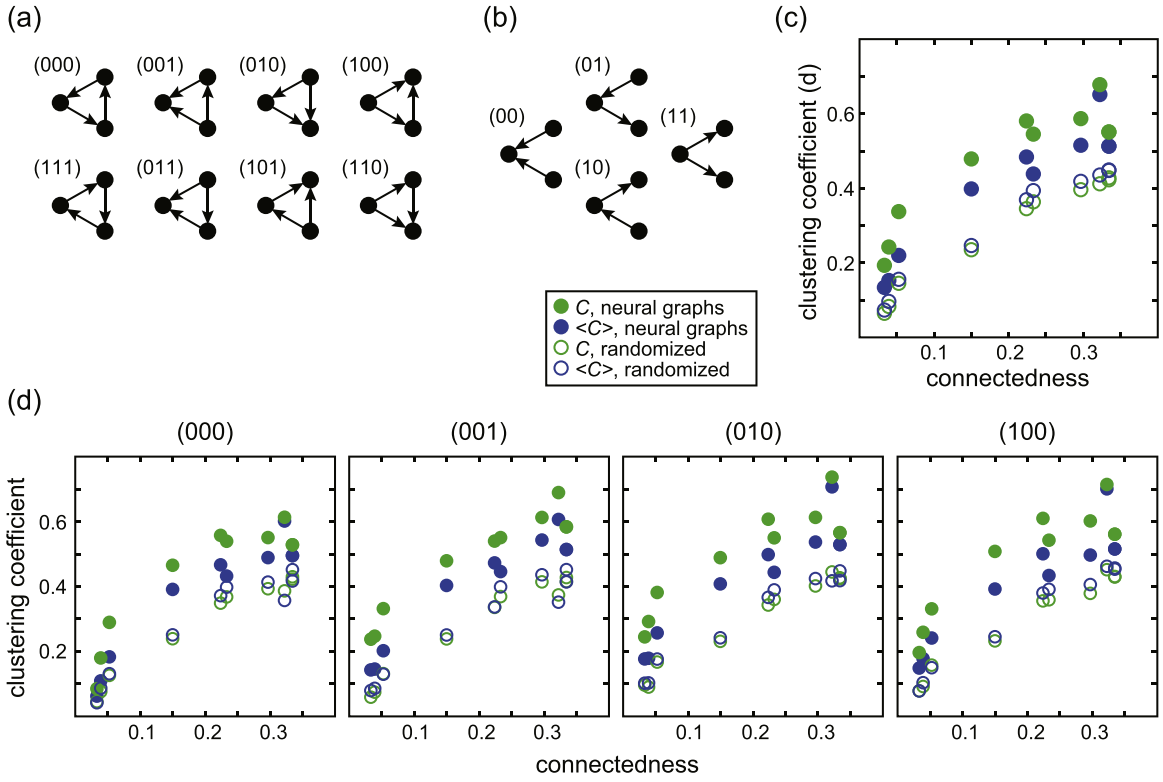

Figure 2. Clustering in neural graphs. (a)–(b) Adjacency relations contributing to the clustering in digraphs. There are eight possible triangular relationships when taking the edge direction into account, defining triangles of type (abc), and four relationships defining connected triples of type (ab). From these eight triangles, four equivalent pairs exist, leaving four distinct motifs defining clustering coefficients for digraphs (equations (23)–(27) and equations (28)–(32)). (c) Total graph clustering coefficient C^d (equation (27)) and average node clustering coefficient $\langle C_i^d \rangle$ (equation (34)) as a function of graph connectedness. Filled dots indicate the neural graphs, while open dots indicate their degree-preserved randomized counterparts. (d) Individual graph clustering coefficients $C^{(abc)}$ (equations (23)–(26)) and node clustering coefficients $\langle C_i^{(abc)} \rangle$ (equation (33)), plotted as in panel (c).

$$C^{(100)} = \frac{T_\Delta^{(100)}}{T_<^{(11)}}, \tag{26}$$

where $T_\Delta^{(abc)}$, $a, b, c \in \{0, 1\}$, denotes the number of triangles of type (abc) (figure 2(a)) and $T_<^{(ab)}$ the number of connected triples of type (ab) (figure 2(b)). Furthermore, the sum over all possible types of triangles and connected triples yields a total global clustering coefficient for digraphs [24]

$$C^d = \frac{1}{2\sum_{i=1}^{N_N}\left(a_i^{\text{tot}}\left(a_i^{\text{tot}} - 1\right) - 2a_i^{\,b}\right)} \sum_{i=1}^{N_N}\sum_{\substack{j=1 \\ j\neq i}}^{N_N}\sum_{\substack{h=1 \\ h\neq i,j}}^{N_N} \left(a_{ij} + a_{ji}\right)\left(a_{ih} + a_{hi}\right)\left(a_{jh} + a_{hj}\right), \tag{27}$$

where $a_i^{\text{tot}} = a_i^{\text{in}} + a_i^{\text{out}}$ and

$$a_i^{b} = \sum_{\substack{j=1 \\ j\neq i}}^{N_N} a_{ij} a_{ji}.$$

Note that $C^{(abc)}$ and C^{d} are set to 0 if the denominator in their respective defining equations vanishes, and that $0 \leqslant C^{(abc)}, C^{d} \leqslant 1$.

Similar to the undirected case and motivated by equations (23)–(27), one can furthermore define node clustering coefficients for digraphs according to [24]

$$C_i^{(000)} = \frac{T_{\Delta i}^{(000)}}{T_{<i}^{(01)}} = \frac{1}{2\left(a_i^{\text{in}} a_i^{\text{out}} - a_i^{b}\right)} \sum_{\substack{j=1 \\ j\neq i}}^{N_N} \sum_{\substack{h=1 \\ h\neq i,j}}^{N_N} \left(a_{ij} a_{jh} a_{hi} + a_{ih} a_{hj} a_{ji}\right), \tag{28}$$

$$C_i^{(001)} = \frac{T_{\Delta i}^{(001)}}{T_{<i}^{(00)}} = \frac{1}{2a_i^{in}\left(a_i^{\text{in}} - 1\right)} \sum_{\substack{j=1 \\ j\neq i}}^{N_N} \sum_{\substack{h=1 \\ h\neq i,j}}^{N_N} a_{ji} a_{hi} \left(a_{jh} + a_{hj}\right), \tag{29}$$

$$C_i^{(010)} = \frac{T_{\Delta i}^{(010)}}{T_{<i}^{(01)}} = \frac{1}{2\left(a_i^{\text{in}} a_i^{\text{out}} - a_i^{b}\right)} \sum_{\substack{j=1 \\ j\neq i}}^{N_N} \sum_{\substack{h=1 \\ h\neq i,j}}^{N_N} \left(a_{ij} a_{hj} a_{hi} + a_{ih} a_{jh} a_{ji}\right), \tag{30}$$

$$C_i^{(100)} = \frac{T_{\Delta i}^{(100)}}{T_{<i}^{(11)}} = \frac{1}{2a_i^{\text{out}}\left(a_i^{\text{out}} - 1\right)} \sum_{\substack{j=1 \\ j\neq i}}^{N_N} \sum_{\substack{h=1 \\ h\neq i,j}}^{N_N} a_{ij} a_{ih} \left(a_{jh} + a_{hj}\right), \tag{31}$$

$$C_i^{d} = \frac{1}{2\left(a_i^{\text{tot}}\left(a_i^{\text{tot}} - 1\right) - 2a_i^{b}\right)} \sum_{\substack{j=1 \\ j\neq i}}^{N_N} \sum_{\substack{h=1 \\ h\neq i,j}}^{N_N} \left(a_{ij} + a_{ji}\right)\left(a_{ih} + a_{hi}\right)\left(a_{jh} + a_{hj}\right), \tag{32}$$

where $T_{\Delta i}^{(abc)}$ denotes the number of triangles of type (abc) connecting two nodes in the neighborhood of node i and node i, and $T_{<i}^{(ab)}$ the number of connected triples of type (ab) with node i being the base node and two other nodes m, n in the neighborhood of node i. Also here, $0 \leqslant C_i^{(abc)}, C_i^{d} \leqslant 1$. Finally, the average over the node clustering coefficients yields the corresponding average node clustering coefficients

$$\left\langle C_i^{(abc)} \right\rangle = \frac{1}{N_N} \sum_{i=1}^{N_N} C_i^{(abc)}, \tag{33}$$

$$\left\langle C_i^{d} \right\rangle = \frac{1}{N_N} \sum_{i=1}^{N_N} C_i^{d}. \tag{34}$$

Numerical results of the application of the various clustering coefficients to the investigated biological neural graphs are summarized in table 2 and visualized in figures 2(c)–(d). As expected, the clustering of a graph increases with its connectedness, with the graph clustering coefficients taking larger values than the corresponding average node clustering coefficients (figures 2(c)–(d), compare blue and green filled dots). Interestingly, in all

Table 2. Clustering analysis across neural graphs. Shown are values for the global clustering coefficients $C^{(abc)}$ and C^{d}, equations (23)–(27), average node clustering coefficients $\langle C^{(abc)} \rangle$ and $\langle C^{d} \rangle$, equations (33)–(34).

Clustering analysis					
	C^{d}	$C^{(000)}$	$C^{(001)}$	$C^{(010)}$	$C^{(100)}$
CC1	0.4384	0.4315	0.4453	0.4430	0.4339
CC2	0.5154	0.4883	0.5423	0.5361	0.4967
CE1	0.1339	0.0624	0.1429	0.1766	0.1488
CE3	0.1525	0.1093	0.1455	0.1788	0.1778
MB1	0.2199	0.1831	0.2019	0.2568	0.2413
MC1	0.3982	0.3905	0.4027	0.4075	0.3922
MC2	0.6514	0.6008	0.6056	0.7058	0.7006
MNC1	0.4838	0.4660	0.4720	0.4974	0.5006
MVC1	0.5129	0.4945	0.5131	0.5283	0.5155
MVC2	0.5129	0.4945	0.5131	0.5283	0.5155
	$\langle C_i^{d} \rangle$	$\langle C_i^{(000)} \rangle$	$\langle C_i^{(001)} \rangle$	$\langle C_i^{(010)} \rangle$	$\langle C_i^{(100)} \rangle$
CC1	0.5451	0.5386	0.5498	0.5490	0.5426
CC2	0.5867	0.5499	0.6118	0.6120	0.6019
CE1	0.1933	0.0851	0.2373	0.2445	0.1965
CE3	0.2424	0.1798	0.2469	0.2918	0.2591
MB1	0.3373	0.2892	0.3313	0.3811	0.3311
MC1	0.4786	0.4647	0.4784	0.4880	0.5080
MC2	0.6780	0.6125	0.6882	0.7357	0.7142
MNC1	0.5805	0.5567	0.5391	0.6065	0.6098
MVC1	0.5510	0.5276	0.5829	0.5649	0.5609
MVC2	0.5510	0.5276	0.5829	0.5649	0.5609

cases, $C^{(000)}$ and $\langle C^{(000)} \rangle$, which correspond to the occurrence of closed geodesics of length three in digraphs, were smallest among their peers. The various clustering coefficients for the degree-preserved randomized surrogates are likewise visualized as open dots in figure 2. Also here, the average node clustering coefficients were found to be larger than the corresponding graph clustering coefficients (figure 2(d), compare open blue and green dots), but with all triangular motifs represented with equal frequency. Moreover, as for the neural graphs, a general increase of the clustering for increasing connectedness was observed in the randomized graphs. However, in comparison to the randomized graphs, the neural data exhibits a clustering which, in most cases, is less than two times larger. This finding is somewhat surprising, as previous claims of a prevalent small-world structure in biological neural graphs would demand a significantly larger clustering than in corresponding random graphs.

To quantify the dissimilarity of the clustering coefficients in the investigated neural graphs and degree-preserved randomized graphs, we defined the following ratios [6]

$$\gamma^{(abc)} = \frac{C^{(abc)}}{C^{(abc)}_{\mathrm{MS}}}, \tag{35}$$

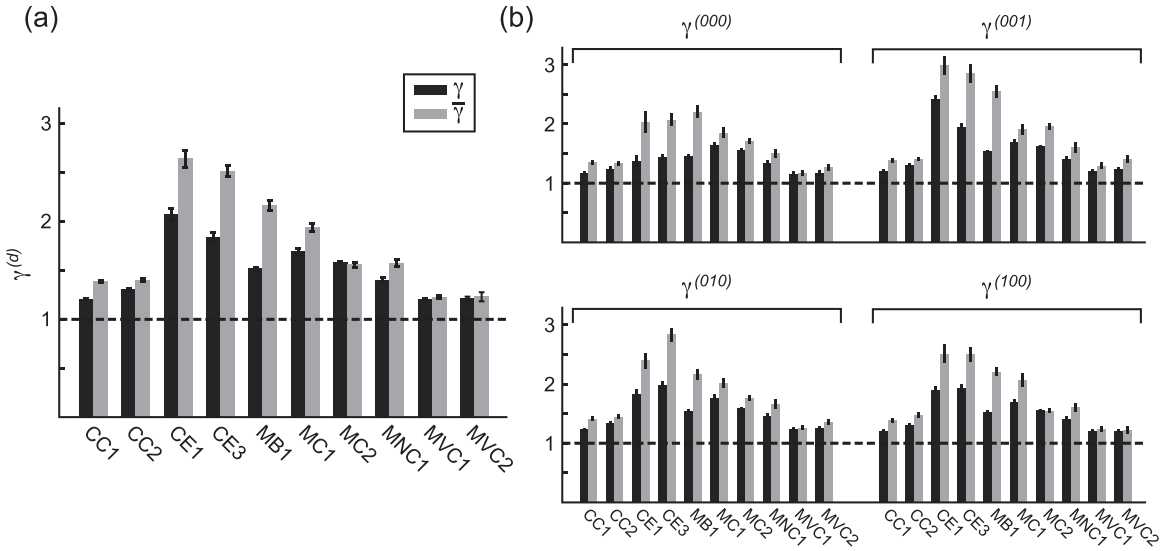

Figure 3. Clustering in neural graphs. (a) The ratio (mean ± SD) γ^d between the total clustering coefficient C^d in the observed and randomized graph (equations (35)–(36); dark gray). $\bar{\gamma}^d$ (equation (38); light grey) is plotted in the same way. (b) The ratios $\gamma^{(abc)}$ and $\bar{\gamma}^{(abc)}$ for the individual graph and node clustering coefficients.

$$\gamma^d = \frac{C^d}{C^d_{\mathrm{MS}}}, \tag{36}$$

$$\bar{\gamma}^{(abc)} = \frac{\left\langle C_i^{(abc)} \right\rangle}{\left\langle C_i^{(abc)} \right\rangle_{\mathrm{MS}}}, \tag{37}$$

$$\bar{\gamma}^d = \frac{\left\langle C_i^d \right\rangle}{\left\langle C_i^d \right\rangle_{\mathrm{MS}}} \tag{38}$$

between the graph clustering coefficients $C^{(abc)}$, C^d and average node clustering coefficients $\langle C_i^{(abc)} \rangle$, $\langle C_i^d \rangle$ of a given digraph and its corresponding random graph ($C^{(abc)}_{\mathrm{MS}}$, C^d_{MS}, $\langle C_i^{(abc)} \rangle_{\mathrm{MS}}$ and $\langle C_i^d \rangle_{\mathrm{MS}}$, respectively). In most cases, we observed γ-values smaller than 2 (figure 3). Only for the neural connectivity graphs of *C. elegans* (CE1 and CE3), the various γ take larger values up to 2.98. In accordance with the above-mentioned finding, γ-values for digraph clustering coefficients of type (000) are generally smaller, hinting at the underrepresentation of motifs describing closed geodesics of length three in biological graphs (see figure 2(d), $\gamma^{(000)}$). In the next section, γ defined in equations (35)–(38) will be used to quantify the small-worldness of the considered graphs.

4.3. Small-worldness

The notion of 'small-worlds' was introduced to describe a specific class of graphs in which the typical geodesic distance (shortest path length) between two nodes is comparable to that in ER random graphs with the same number of nodes and connectedness, and in which the geodesic

distance scales logarithmically with the number of nodes in the graph. Furthermore, small-worlds are endowed with an overabundance of local triangular relationships, i.e. display high clustering, when compared with their random counterparts [1]. Using the notation introduced above, a graph is labeled a 'small-world' if both $\lambda \geqslant 1$ and $\gamma \gg 1$ (e.g., see [6]).

In section 4.1, we showed that for biological neural graphs the various λ-values (equations (17)–(19)) range from 0.98 to 1.30 for the considered neural graphs (see figure 1(d)), thus suggesting a strong similarity between biological neural graphs and corresponding random graphs when considering the typical geodesic distance between nodes. Moreover, the obtained results for the average graph distance scale in accordance with the expectation for random graphs (figure 1(c), compare filled dots with solid triangles), i.e. logarithmically in the number of nodes (see equation (16) for analytical form). Thus, the first condition $\lambda \geqslant 1$, allowing for a characterization of biological neural graphs as small-worlds, is met.

The evidence for the small-world property in the context of the various clustering coefficients is, however, much less clear. Here, we found that for the majority of the considered biological digraphs, γ takes values in the range between 1.15 and 2. Only the neural connectivity digraphs of *C. elegans* (CE1 and CE3) yield values up to 2.98 for $\gamma^{(001)}$. This finding suggests that, at least for the biological neural graphs studied here, the condition $\gamma \gg 1$ characterizing small-world graphs must be met with greater care.

In order to better quantify the level of 'small-worldness', we employed a number of small-worldness indices [6, 21]. Global small-worldness indices for digraphs can be defined by

$$S^{(abc)} = \frac{\gamma^{(abc)}}{\lambda_{\langle d\rangle}} = \frac{C^{(abc)}}{C^{(abc)}_{\mathrm{MS}}} \frac{\langle d\rangle_{\mathrm{MS}}}{\langle d\rangle}, \tag{39}$$

$$S^{d} = \frac{\gamma^{d}}{\lambda_{\langle d\rangle}} = \frac{C^{d}}{C^{d}_{\mathrm{MS}}} \frac{\langle d\rangle_{\mathrm{MS}}}{\langle d\rangle}. \tag{40}$$

Similarly, two sets of local small-worldness indices are defined for digraphs as

$$\bar{S}^{(abc)} = \frac{\bar{\gamma}^{(abc)}}{\lambda_{\langle d_i\rangle}} = \frac{\left\langle C_i^{(abc)}\right\rangle}{\left\langle C_i^{(abc)}\right\rangle_{\mathrm{MS}}} \frac{\left\langle d_i\right\rangle_{\mathrm{MS}}}{\left\langle d_i\right\rangle}, \tag{41}$$

$$\bar{S}^{d} = \frac{\bar{\gamma}^{d}}{\lambda_{\langle d_i\rangle}} = \frac{\left\langle C_i^{d}\right\rangle}{\left\langle C_i^{d}\right\rangle_{\mathrm{MS}}} \frac{\left\langle d_i\right\rangle_{\mathrm{MS}}}{\left\langle d_i\right\rangle} \tag{42}$$

and

$$\tilde{S}^{(abc)} = \frac{\bar{\gamma}^{(abc)}}{\lambda_{L}} = \frac{\left\langle C_i^{(abc)}\right\rangle}{\left\langle C_i^{(abc)}\right\rangle_{\mathrm{MS}}} \frac{L_{\mathrm{MS}}}{L}, \tag{43}$$

$$\tilde{S}^{d} = \frac{\bar{\gamma}^{d}}{\lambda_{L}} = \frac{\left\langle C_i^{d}\right\rangle}{\left\langle C_i^{d}\right\rangle_{\mathrm{MS}}} \frac{L_{\mathrm{MS}}}{L}. \tag{44}$$

With these, the two conditions $\lambda \geqslant 1$ and $\gamma \gg 1$ defining small-world graphs are then equivalent to $\{S^{\bullet}, \bar{S}^{\bullet}, \tilde{S}^{\bullet}\} > 1$, where $\bullet \in \{(abc), d\}$ [6].

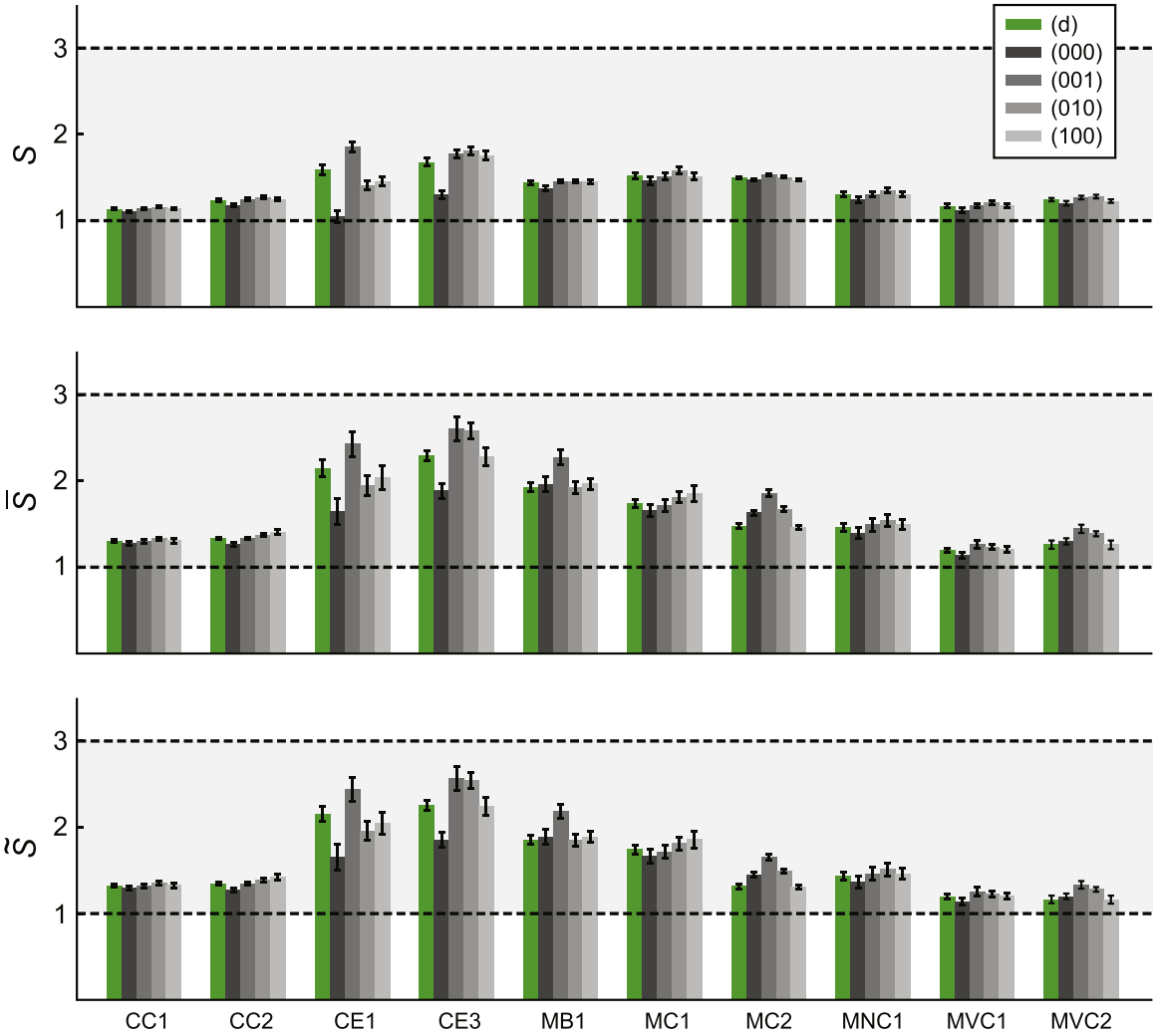

Figure 4. Small-worldness of biological neural graphs. Small-worldness indices (mean ±SD) S (equations (39)–(44); top), $\bar{S}$ (equations (41)–(42); middle) and $\tilde{S}$ (equations (43)–(44); bottom) for the total graph clustering coefficient (green) and the individual clustering coefficients (gray bars). Above $\{S^{\bullet}, \bar{S}^{\bullet}, \tilde{S}^{\bullet}\} = 1$ (black dashed), a graph is considered a small-world, with $1 < \{S^{\bullet}, \bar{S}^{\bullet}, \tilde{S}^{\bullet}\} < 3$ (light grey region) defining borderline small-worlds, according to [6].

For each graph studied here, we calculated the various small-worldness indices defined in equations (39)–(44). In all cases, $S^{\bullet}$, $\bar{S}^{\bullet}$ and $\tilde{S}^{\bullet}$ take values larger than 1 (figure 4), thus placing these graphs into the class of small-worlds. However, in most of the graphs the various small-worldness indices remained far below a value of 2, with only $\bar{S}^{(001)}$, $\bar{S}^{(010)}$ for the CE3 graph reaching values up to 2.61. This suggests that the investigated neural graphs, at best, fall into the class of 'borderline' small-world graphs [6].

4.4. Borderline small-worlds

In order to further explore and give meaning to the borderline small-world regime, we analytically assessed the small-worldness index of WS graphs. The latter is defined by

$$S^d = \frac{C^d_{\mathrm{SW}}}{C^d_{\mathrm{ER}}} \frac{\langle d \rangle_{\mathrm{ER}}}{\langle d \rangle_{\mathrm{SW}}}, \tag{45}$$

where C^d_{SW} and C^d_{ED} denote the total clustering coefficient for small-world and ER digraphs, $\langle d \rangle_{\mathrm{SW}}$ and $\langle d \rangle_{\mathrm{ER}}$ their average geodesic graph distance, respectively.

Recently, an analytically exact expression for the total global clustering coefficient C^d_{SW} of small-world digraphs with degree k and rewiring probability q, valid non-asymptotically for all graph sizes, was obtained [10]

$$C^d_{\mathrm{SW}} = \frac{64}{\mathcal{N}_{\mathrm{SW}}(N_N - 1)^3}\Big(k^3(N_N - 1)^3 + B^3_{\mathrm{SW}}A_{(3,0)} - 3kqB^2_{\mathrm{SW}}A_{(2,1)} + 3(kq)^2B_{\mathrm{SW}}A_{(1,2)} - (kq)^3A_{(0,3)}\Big), \tag{46}$$

where

$$B_{\mathrm{SW}} = (N_N - 1)\left(1 - q + \frac{2kq}{N_N - 1}\right),$$

$$\mathcal{N}_{\mathrm{SW}} = \frac{8kN_N}{N_N - 1}\Big(2k\left(q^2 - 2q + 2N_N - 2\right) - (N_N - 1)\left(q^2 - 2q + 2\right)\Big)$$

and

$$A_{(i,j)} = \sum_{m=1}^{N_N - 1} \frac{\left\{\sin\left(\frac{km}{N_N}\pi\right)\cos\left(\frac{(1+k)m}{N_N}\pi\right)\right\}^i \left\{\sin\left(\frac{(1+2k)m}{N_N}\pi\right)\right\}^j}{\left\{\sin\left(\frac{m}{N_N}\pi\right)\right\}^{-(i+j)}}. \tag{47}$$

For $q = 1$, equation (46) yields the corresponding expression for the total global clustering coefficient of ER graphs, namely

$$C^d_{\mathrm{ER}} = \frac{64}{\mathcal{N}_{\mathrm{ER}}(N_N - 1)^3}\Big(k^3(N_N - 1)^3 + B^3_{\mathrm{ER}}A_{(3,0)} - 3kB^2_{\mathrm{ER}}A_{(2,1)} + 3k^2B_{\mathrm{ER}}A_{(1,2)} - k^3A_{(0,3)}\Big) \tag{48}$$

with $B_{\mathrm{ER}} = 2k$ and

$$\mathcal{N}_{\mathrm{ER}} = \frac{8kN_N}{N_N - 1}\big(2k(2N_N - 3) - (N_N - 1)\big).$$

For the average geodesic graph distance, we utilize equation (16), i.e.

$$\langle d \rangle_{\mathrm{ER}} = \frac{\ln(N_N) - \gamma}{\ln(Co \cdot N_N)} + \frac{1}{2}, \tag{49}$$

which provides an excellent analytical approximation for the geodesic distance for ER random graphs [11] (see figure 1(c), compare black dots with downward triangles). Unfortunately, no valid analytical expression or approximation for the geodesic distance has yet been proposed for WS small-world graphs. The only known expression is [25]

$$\langle d\rangle_{\mathrm{SW}} = \frac{N_N}{2k\sqrt{\left(kqN_N\right)^2 + 2kqN_N}} \tanh^{-1}\sqrt{\frac{kqN_N}{kqN_N + 2}}, \tag{50}$$

which provides reasonable approximations only in a narrow parameter regime for $q \ll 1$ and small k, and therefore will not be considered here.

To assess the small-worldness index (45) analytically, taking the lack of an analytical expression for $\langle d\rangle_{\mathrm{SW}}$ into account, we focused on three different approaches. We first studied the maximal value S^d can take. As $\langle d\rangle_{\mathrm{SW}}$ is a monotonically decreasing function of q with $\langle d\rangle_{\mathrm{SW}} = \langle d\rangle_{\mathrm{ER}}$ for $q = 1$, we can define

$$S^d_{\max} = \frac{C^d_{\mathrm{SW}}}{C^d_{\mathrm{ER}}}, \tag{51}$$

which provides an upper bound for S^d, i.e. $S^d \leqslant S^d_{\max}$. We note that, due to the monotonic behavior of $\langle d\rangle_{SW}$ as function of the rewiring probability, S^d will always be smaller, especially for small q. Thus, equation (51) provides a valid upper bound of the small-worldness index for $q \ll 1$, i.e. close to symmetric ring-graph regime of the WS graph model.

Secondly, we investigated the theoretical lower bound of S^d. As $\langle d\rangle_{\mathrm{SW}}$ is maximal for $q = 0$, i.e. a symmetric ring-graph, we define

$$S^d_{\min} = \frac{C^d_{\mathrm{SW}}}{C^d_{\mathrm{ER}}}\frac{\langle d\rangle_{\mathrm{ER}}}{\langle d\rangle_{\mathrm{RG}}}, \tag{52}$$

where

$$\langle d\rangle_{\mathrm{RG}} = \frac{1}{N_N - 1}\left(\frac{N_N - 1}{2}(a + 1) - \frac{k}{2}a^2 - \frac{k}{2}a\right)$$

with

$$a = \left\lfloor \frac{N_N - 1}{2k} \right\rfloor.$$

As for a given graph $\langle d\rangle_{\mathrm{ER}}/\langle d\rangle_{\mathrm{RG}}$ is fixed, and C^d_{SW} is monotonically decreasing with rewiring probability q, reaching its minimum $C^d_{\mathrm{SW}} = C^d_{\mathrm{ER}}$ for $q = 1$, the small-worldness index S^d will always be larger, especially for larger values of q. Thus, equation (52) provides a valid lower bound of the small-worldness index for $q \gg 0$, i.e. close to random regime of the WS graph model.

Finally, we explored the behavior of the small-worldness index in a semi-analytical fashion by replacing $\langle d\rangle_{\mathrm{SW}}$ with numerical values obtained from analyzing corresponding semi-analytical small-world models constrained by choosing N_N and $k = \lfloor \frac{1}{2}Co\,(N_N - 1)\rfloor$ from their neural counterparts. This allows to define

$$S^d_{\mathrm{sa}} = \frac{C^d_{\mathrm{SW}}}{C^d_{\mathrm{ER}}}\frac{\langle d\rangle_{\mathrm{ER}}}{\langle d\rangle^{\mathrm{num}}_{\mathrm{SW}}}, \tag{53}$$

where $\langle d\rangle^{\mathrm{num}}_{\mathrm{SW}}$ denotes the mean of the average geodesic graph distance obtained from numerical analysis of constructed WS graphs for a given parameter set.

Figure 5(a) visualizes the borderline small-world regime (small-worldness between 1 and 3) obtained from analytical ($S^d_{\min}$, equation (52); $S^d_{\max}$, equation (51)) and semi-analytical (S^d_{sa},

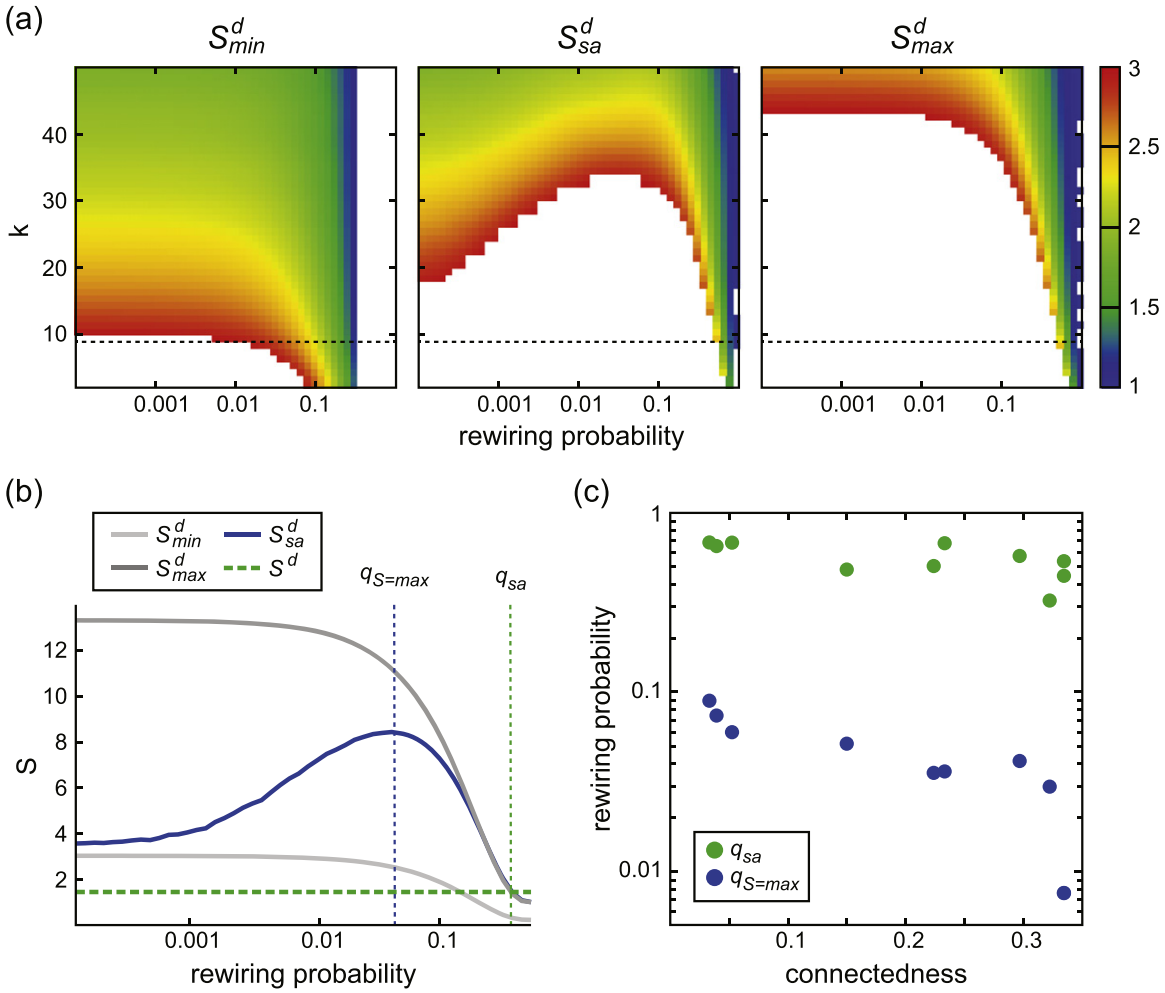

Figure 5. Analytical and semi-analytical investigation of borderline small-worlds. (a) Small-worldness index $S^d_{\min}$ (equation (52)), S^d_{sa} (equation (53)) and $S^d_{\max}$ (equation (51)) as function of rewiring probability q and degree k of Watts–Strogatz graphs for $N_N = 350$ (corresponding to the representative MB1 neural graph). Displayed is the borderline region, defined as $1 \leqslant S^d_{\min}, S^d_{\mathrm{sa}}, S^d_{\max} \leqslant 3$. The dashed line marks the degree $k = \lfloor \frac{1}{2} Co\,(N_N - 1) \rfloor = 9$ which corresponds to the investigated MB1 graph. (b) Minimal ($S^d_{\min}$, equation (52); light grey), maximal ($S^d_{\max}$, equation (51); dark grey) small-worldness indices, and small-worldness index in the semi-analytical model (S^d_{sa}, equation (53); blue) of a small-world graph with $N_N = 350$ and $k = 9$. The horizontal dashed line indicates the small-worldness index observed in the MB1 graph. The vertical green dotted line marks the rewiring probability q_{sa} at which the analytical curves cross the observed S^d value (green dashed), and the blue dotted line the rewiring probability $q_{S=\max}$ at which the semi-analytical model reaches maximal small-worldness. (c) Pooled result of q_{sa} and $q_{S=\max}$ values for the investigated neural graphs. Note that q_{sa} is about one order of magnitude larger than $q_{S=\max}$ signifying the maximally small-world regime.

equation (53)) considerations in a representative graph of size $N_N = 350$ for $1 \leqslant k \leqslant 50$ and $0.0001 \leqslant q \leqslant 1$. We observe that the borderline region resides, for sufficiently small value of k (i.e. sufficiently sparse graphs), only at larger values of the rewiring probability q (figure 5(a), dotted lines). This suggests, in accordance with the findings reported earlier, a dominant random component in the makeup of the corresponding graphs.

To quantify this finding, we then investigated the behavior of $S^d_{\min}$, $S^d_{\max}$ and S^d_{sa} as function of q in the semi-analytical graph models with N_N and k chosen from their corresponding neural graphs (figure 5(b)). Specifically, we numerically assessed in each semi-analytical model the rewiring probability q_{sa} necessary to reproduce the small-worldness index S^d observed in the corresponding neural graph (figure 5(b), intersection of green dashed and solid blue line), i.e. $S^d_{\mathrm{sa}}(q_{\mathrm{sa}}) = S^d$. Furthermore, we estimated the rewiring probability $q_{S=\max}$ at which S^d_{sa} reaches its maximum (figure 5(b), blue dotted line), in order to quantify the maximally small-world regime for a given graph model.

The pooled result of this analysis applied to all investigated neural graphs is presented in figure 5(c). We observe that the q_{sa} for all neural graphs is large ($q_{\mathrm{sa}} \geqslant 0.32$) and can reach values of 0.68 for graphs with small connectedness, in general agreement with the range reported in [6]. This shows that for the investigated neural graphs, under the premise that the latter are small-worlds, a significant rewiring of at least 30% and up to 70% of edges is required to construct the neural graphs with a given small-worldness index from the WS model. More importantly, the rewiring probability $q_{S=\max}$ at which the semi-analytical small-world model reaches its maximally small-world regime is in all investigated cases about one order of magnitude lower than q_{sa}. This shows that neural graphs, even under the premise of being small-worlds, reside far away from the maximally small-world regime of the WS model, and that the latter has to be taken into account when assessing the small-worldness properties of graphs.

5. Discussion

In this work, we have conducted a comprehensive analysis of the small-world property in structural neural graphs. We started with a classical numerical quantification of the clustering coefficient and average graph geodesic distance (sections 4.1 and 4.2, respectively). From this, we then calculate in each neural graph the small-worldness index, which precisely quantifies the coexistence of high clustering and low average path length in relation to that observed in random graphs (section 4.3). We observed that, in none of the neural graphs considered here does the small-worldness index exceed a value of 3, which suggests that neural graphs fall into a 'borderline' regime close to random graphs [6]. For densely connected graphs, high clustering and short path length are expected in their randomized counterparts. This will narrow the range of posssible small-worldness values, with small-worldness approaching 1 for increasing connectedness. Interestingly, however, we observe that all graphs considered, whether dense (MVC and MC2) or sparse (CE and MB1), fall into this borderline regime.

To corroborate this observation, we then employed known analytical expressions for the clustering coefficient in random and small-world graphs, and for the average geodesic distance in random graphs, and constructed semi-analytical small-world models corresponding to the investigated neural graphs (section 4.4). By studying the numerical range the global small-worldness index takes in these models as a function of the remaining free parameter, the

rewiring probability, we conceived of a novel method to quantify the small-worldness property of a given graph. Taken alone, the small-worldness index introduced in [6] is highly variable and strongly depends on the size and connectivity of the graph in question. It is thus difficult to interpret directly. Only in relation to the maximal small-worldness possible for a given graph can an objective quantification of the small-world property be made. Specifically, while the maximal small-worldness index in the semi-analytical small-world graphs is highly variable, with values ranging from 1.79 (for the MVC1) to 9.70 (for the *C. elegans* neural connectivity graph, CE3), the actual values observed in the corresponding neural graphs are, in all cases, much lower (1.17 for MVC1, 1.68 for CE3).

In order to quantify this difference, we then needed to ask the question: which rewiring is required to obtain, in these semi-analytical small-world models, the small-worldness index observed in the neural data? We found that the observed small-worldness can only be reproduced by a rewiring that is about one order of magnitude larger than that which produces the maximal small-worldness. This finding holds for all structural neural graphs investigated here, and it means that the rewiring required to reproduce the small-worldness observed in the neural data is far larger than the rewiring of only a few edges first proposed by Watts and Strogatz as a mechanism to bring a regular ring-graph into the small-world regime. This result suggests that neural graphs reside far closer to randomness than to maximal small-worldness.

In recent years, neural connectivity diagrams of the human cerebral cortex with a higher number of nodes than the graphs considered here have been introduced [9]. There, the small-worldness index was observed to increase from 1.54 in regional connectivity to 10.64 in high-resolution connectivity. Such an increase, however, is to be expected, given the known linear scaling of small-worldness with network size [6]. Because the linear increase of small-worldness with network size will also apply to the maximal small-worldness index, a quantification of the small-world properties must again take both the actual and maximal possible small-worldness index into account. Indeed, using the parameters of the high-resolution connectivity map from [9] in the framework of our study, we found that the maximal small-worldness index for the constructed semi-analytical model is 16.28, reached at a rewiring probability of 2% of edges, while the reported small-worldness index of 10.64 requires a rewiring of 11.5% of edges. This additional analysis thus demonstrates that the findings of this study are applicable even to the next generation of large-scale reconstructions of neural connectivity.

These findings, consistently observed across neural connectivity diagrams spanning multiple scales of organization in the nervous system, range from the microscopic neural connectivity of *C. elegans*, to connections between individual neocortical areas, up to the regional-level connectiviy in the primate brain. Here, we focus solely on the structural connectivity. However, in future work, it will be interesting to extend this analysis to time-dependent functional connectivity and its implications on the system-level. Similar to recent work assessing the interplay within networks of independent physiological systems [26], it is possible that the specific structural organization of neural connectivity that we begin to uncover here (see also [22]) allows individual neurons and brain regions to interact as a functional whole. In line with this point, this specific large-scale organization of neural connectivity at multiple scales may reflect a high-dimensional structure that is not strictly subject to the constraint of wiring minimization ([27]; see also [28]).

In conclusion, in this study we confirm the often reported finding, that the clustering of neural connectivity exceeds that of a random graph while the distance is approximately the

same. We continue, however, to observe that the quantification of these measures is not enough for a complete assessment of the small-world property in a given real-world graph. Here, we propose that a meaningful assessment of the small-worldness must take the maximal small-worldness into account, as well as relate the rewiring necessary to achieve this maximal and actual small-worldness. By doing so, we have observed in this study that neural graphs, though often discussed in the context of the small-world property, in fact reside far outside the classical small-world regime.

Acknowledgements

The authors wish to thank OD Little for inspiring comments. Work supported by the CNRS and the European Community (BrainScales project, FP7-269921). LM is a PhD fellow from École des Neurosciences de Paris (ENP), and also received support from ANR (Complex-V1).

References

[1] Watts D J and Strogatz S H 1998 *Nature* **393** 440–2
[2] Albert R, Hawoong J and Barabási A 1999 *Nature* **401** 130
[3] Amaral L A N, Scala A, Barthelemy M and Stanley H E 2000 *Proc. Natl Acad. Sci. USA* **97** 11149
[4] Conyon M J and Muldoon M R 2006 *J. Bus., Finance Account.* **33** 1321
[5] Sporns O and Zwi J 2004 *Neuroinformatics* **2** 145
[6] Humphries M D and Gurney K 2008 *PLoS One* **3** e0002051
[7] He Y, Chen Z J and Evans A C 2007 *Cereb. Cortex* **17** 2407
[8] Li M, Chen H, Wang J, Liu F, Long Z, Wang Y, Iturria-Medina Y, Zhang J, Chunshui Y and Chen H 2014 *Brain Connectivity* **4** 145
[9] Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey C J, Wedeen V J and Sporns O 2008 *PLoS Biol.* **6** e159
[10] Rudolph-Lilith M and Muller L E 2014 *Phys. Rev.* E **89** 012812
[11] Fronczak A, Fronczak P and Hołyst J A 2004 *Phys. Rev.* E **70** 056110
[12] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D-U 2006 *Phys. Rep.* **424** 175
[13] Newman M E J 2010 *Networks: An Introduction* (Oxford: Oxford University Press)
[14] http://www.cydyns.com; http://www.newscienceportal.com/MLR
[15] Varshney L R, Chen B L, Paniaqua E, Hall D H and Chklovskii D B 2011 *PLoS Comput. Biol.* **7** e1001066
[16] Maslov S and Sneppen K 2002 *Science* **296** 910
[17] del Genio C I, Kim H, Toroczkai Z and Bassler K E 2010 *PLoS One* **5** e10012
[18] Roberts E S and Coolen A C C 2012 *Phys. Rev.* E **85** 046103
[19] Kim H, del Genio C I, Bassler K E and Toroczkai Z 2012 *New J. Phys.* **14** 023012
[20] Newman M E J, Strogatz S H and Watts D J 2001 *Phys. Rev.* E **64** 026118
[21] Rudolph-Lilith M, Destexhe A and Muller L E 2012 arXiv:1208.3383v1
[22] Rudolph-Lilith M and Muller L E 2014 *Biol. Cybern.* **108** 381
[23] Holland P W and Leinhardt S 1971 *Comp. Group Stud.* **2** 107
[24] Fagiolo G 2007 *Phys. Rev.* E **76** 026107
[25] Newman M E J, Moore C and Watts D J 2000 *Phys. Rev. Lett.* **84** 3201
[26] Bashan A, Bartsch R P, Kantelhardt J W, Havlin S and Ivanov P C 2012 *Nat. Commun.* **3** 702
[27] Mathias N and Gopal V 2001 *Phys. Rev.* E **63** 021117
[28] Kaiser M and Hilgetag C C 2005 *PLoS Comput. Biol.* **2** e95