Received: 1 April 2020 | Revised: 6 August 2020 | Accepted: 7 August 2020

DOI: 10.1111/ejn.14937

RESEARCH REPORT

EJN European Journal of Neuroscience FENS WILEY

On the effect of neuronal spatial subsampling in small-world networks

Mattia Bonzanni | **Kimberly M. Bockley** | **David L. Kaplan**

Department of Biomedical Engineering, Tufts University, Medford, MA, USA

Correspondence
Mattia Bonzanni and David L. Kaplan, Department of Biomedical Engineering, Tufts University, Medford, MA, USA.
Email: mattia.bonzanni@tufts.edu; david.kaplan@tufts.ed

Funding information
National Institutes of Health, Grant/Award Number: P41EB027062, R01NS092847 and U01EB014976

Abstract

The analysis of real-world networks of neurons is biased by the current ability to measure just a subsample of the entire network. It is thus relevant to understand if the information gained in the subsamples can be extended to the global network to improve functional interpretations. Here we showed how average clustering coefficient (CC), average path length (PL), and small-world propensity (SWP) scale when spatial sampling is applied to small-world networks. This extraction mimics the measurement of physical neighbors by means of electrical and optical techniques, both used to study neuronal networks. We applied this method to in silico and in vivo data and we found that the analyzed properties scale with the size of the sampled network and the global network topology. By means of mathematical manipulations, the topology dependence was reduced during scaling. We highlighted the behaviors of the descriptors that, qualitatively, are shared by all the analyzed networks and that allowed an approximated prediction of those descriptors in the global graph using the subgraph information. In contrast, below a spatial threshold, any extrapolation failed; the subgraphs no longer contain enough information to make predictions. In conclusion, the size of the chosen subgraphs is critical to extend the findings to the global network.

KEYWORDS

network subsampling, optical imaging, small-world networks, spatial neighbors

1 | INTRODUCTION

The development of new methodologies increased the capability to simultaneously record the electrical and optical activities of larger and larger fractions of neurons with an unprecedented resolution. These datasets are described and studied using graph theory, which identifies each neuron or region as a node and their anatomical and/or functional connections to neighbors as edges (Bassett, Zurn, & Gold, 2018). However, even with these advances, the fraction of analyzed neurons still represents a small portion with respect to the whole system. This spatial subsampling is, in most cases, unavoidable (Lee, Kim, & Jeong, 2006; Levina & Priesemann, 2017) and it poses two inquiries: (a) what is the likelihood that a given property of the global network is scale invariant? and (b) If the properties of a graph are not scale invariant, can we predict the behavior of the global network by means of the measures obtained from the fraction of the sampled networks? Growing evidence has indicated that several properties of global networks

Edited by Panayiota Poirazi

vary in sampled networks, suggesting a need to reevaluate relevance (Gerhard, Pipa, Lima, Neuenschwander, & Gerstner, 2011; Lee et al., 2006; Levina & Priesemann, 2017; She, Chen, & Chan, 2016; Stumpf, Wiuf, & May, 2005). One of the most notable properties to describe a neuronal network is its small-world propensity (SWP) (Bassett & Bullmore, 2017). A graph is defined as small-world if it has high average clustering coefficient (CC) and short average path lengths (PLs) in comparison with a lattice and a random null model (Bassett & Bullmore, 2017; Watts & Strogatz, 1998). This small-worldness has been identified in the whole brain as well as with in vitro neuronal cultures (Bassett & Bullmore, 2017; Sporns & Zwi, 2004), and it is associated with network robustness and efficiency of information transmission. Besides the topological nature of the global network (lattice, random, small-world, scale-free, etc.), a key feature of subsampling is the actual sampling method; for example, node sampling, link sampling, and snowball sampling (Lee et al., 2006). Of relevance in a neuronal context is the multi-electrode sampling method used by Gerhard and colleagues (Gerhard et al., 2011) and She and colleagues (She et al., 2016), which emphasizes the spatial component of the measurement, which is different from the aforementioned sampling methods that used topological information of the global network. Their data indicate that the SWP of the sampled network was always overestimated if compared to its value in the global network (Gerhard et al., 2011; She et al., 2016). Along with electrode-based measurements, calcium and voltage imaging are currently the gold-standard approaches to measure neuronal activity. With each imaging technique, the chosen field-of-view dictates the fraction of cells that are able to be measured during data acquisition as a function of the microscope magnification and the original surface area of the cell culture/tissue of study. The neurons retained in the field-of-view are merely spatial neighbors; no other topological considerations or knowledge are considered. Given the ubiquitous use of the imaging approach to evaluate functional/structural features of a neuronal culture, it is relevant to study the effect of subsampling on graph properties with the view of a pure spatial sampling approach. We thus addressed the following question: Given a sampled network originated from a single field-of-view measurement with defined spatial feature, is it possible to predict the topology of the global network and at which extent are the subnetwork data informative of the whole network?

Here we studied the scaling properties of the SWP, average CC, and PL in sampled networks using a spatial neighbor sampling (SNS) approach in three distinct in silico small-world networks (Watts-Strogatz unweighted; Watts-Strogatz weighted; and distance-dependent model) and in a previously published human functional connectivity dataset. The abovementioned properties were assessed by varying the number of nodes, node degree, and rewiring probability.

2 | MATERIALS AND METHODS

All the codes can be downloaded at https://github.com/mattiabonzanni/Subgraph-extraction-with-spatial-neighbors-sampling-SNS-approach-.

2.1 | Watts-Strogatz model construction

The Watts-Strogatz graph was constructed as previously described (Watts & Strogatz, 1998) using the MATLAB function "WattsStrogatz.m" (https://www.mathworks.com/help/matlab/math/build-watts-strogatz-small-world-graph-model.html). Briefly, a lattice network is constructed with N nodes and each node is attached to K/2 neighbors on each side, where K is the mean node degree. Each edge has a probability of rewiring equal to beta (β). For example, the edge connecting node i and node j can be replaced with an edge connecting node i and node k with a probability equal to β, where node k was not previously connected to node i. The Watts-Strogatz weighted model was constructed by calculating the edge weight between nodes (w_{ij}) and defined as follows:

$$w_{ij} = N - d_{ij}$$

where N is the total number of neurons and d_{ij} is the algebraical difference between the index of node i and node j. This guarantees that closer nodes (based on index) have greater weight values. The weight table is calculated before the rewiring step and, after the rewiring, used to construct the weighted graph. This implies that the edge weight is retained while rewiring, as previously described (Muldoon, Bridgeford, & Bassett, 2016).

2.2 | Distance-dependent model construction

The distance-dependent model is constructed using a custom MATLAB code. In detail:

- *Step 1 – Seeding process*

Choose N_G nodes and seed them with random x-y coordinates. The use of the *rand* function implies that for both the x-axis and y-axis the lower and upper limits are 0 and 1, respectively.

- *Step 2 – Weighted edges*

Each node is connected to N_G-1 nodes (self-looping is not permitted). Initially, the edge weight w_{ij} is directly proportional to the space distance d_{ij} between the two nodes (calculated using the Pythagoras' theorem). Considering that the edge weight is empirically greater between two spatially closer neurons (Muldoon et al., 2016), we assign the node edge weight accordingly to:

$$w_{ij} = D_{max} - d_{ij} \tag{1}$$

where $D_{max} = max\{d_{ij}\}$. It follows an inverse relation between w_{ij} and d_{ij}. All the pairwise w_{ij} are stored in a matrix (*distMatrix*). We then normalized the matrix by min (w_{ij}). This is practical to define all the edge weights, as well as the weight threshold w_t, in the range [0;1].

- *Step 3 – Thresholding*

From Step 2, all the possible pairwise connections between nodes are established. However, this does not mimic a real-world neuronal network. Based on previous reports of brain connectivity (She et al., 2016), a weight threshold w_t is thus defined in order to generate a network in which each node is connected in average to about 7%–10% of total nodes (8% is assumed as the default one but can be customized by the user). The thresholding process, moreover, showed a positive correlation between w_t and the emergence of small-world features, as detailed in Supplementary Figure 1a.

- *Step 4 – Random rewiring and characterization*

The distance between two neurons is one of several parameters that influence the connectivity between neurons. Without imposing additional connectivity considerations, we decided to apply a random rewiring step similar to the one in the Watts-Strogatz model (edge weight is retained). The rewiring process, moreover, showed the emergence of small-world behavior with increasing values of the rewiring probability β (Figure S1b); this behavior perfectly resembles the example described in Watts & Strogatz, (1998) starting with the ring lattice. On the other hand, additional connectivity rules can be easily implemented, if necessary. At the end of the rewiring process, the CC_G, PL_G, and SWP_G of the global graph are computed.

- *Step 5–6 – Subgraph extraction and characterization*

A spatial parameter rho (ρ) and an initial seed node *s* are chosen. Subsequently, the x and y coordinates of *s* (x_s and y_s, respectively) are extracted and used to identify all the nodes with the x coordinates in the [$x_s - \rho$; $x_s + \rho$] range and the y coordinates in the [$y_s - \rho$; $y_s + \rho$] range. These nodes and their existing edges are then extracted to create the subgraph. CC_S, PL_S, and SWP_S of the subgraphs are then computed.

2.3 | Parameters definition

The average CC was calculated as follows:

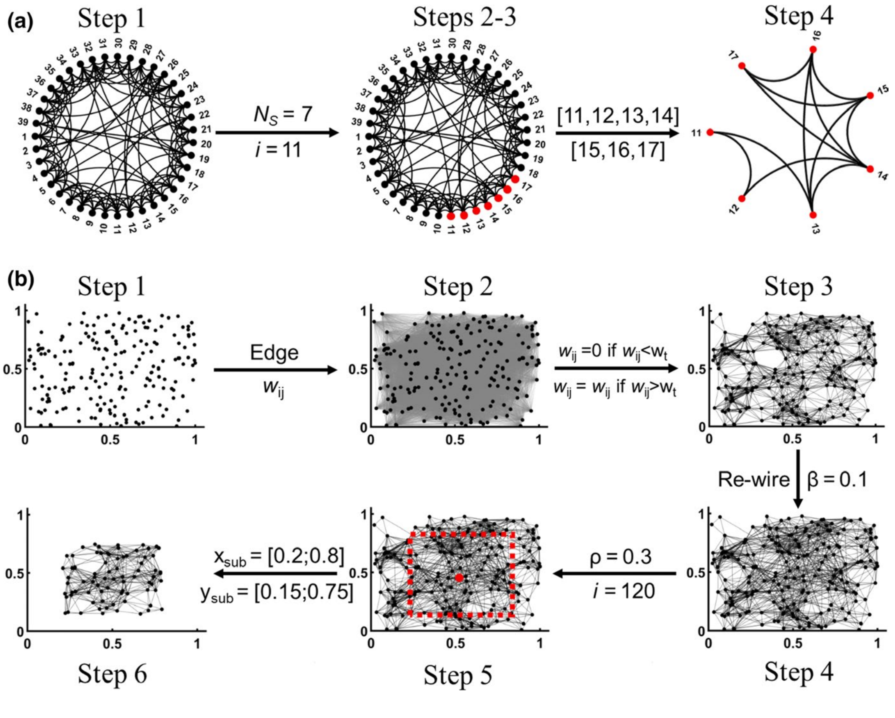

FIGURE 1 Representative subgraph extraction process. Representative graphs showing the SNS method for a given node applied to either an (a) unweighted Watts-Strogatz ($N_G = 39$, $k = 4$, $\beta = 0.1$) or (b) distance-dependent models ($N_G = 200$, $wt = 0.89$)

$$\mathrm{CC}=\frac{1}{N}\sum_{i=1}^{N}C_i \tag{2}$$

where N is the number of nodes and C_i is the local CC of the ith node. For a binary graph:

$$C_i=\frac{2\left|\left\{e_{kj}:v_k,v_j\in N_i,e_{kj}\in N_i\right\}\right|}{k_i\left(k_i-1\right)} \tag{3}$$

where e_{kj} is the edge connecting v_k and v_j vertices, N_i is the set of neighbor nodes of v_i, and k_i (degree of the ith node) is the number of vertices in the neighborhood N_i of v_i. It represents the fraction of neighbors of a node that are also connected to each other.

For a weighted graph:

$$C_i=\frac{1}{k_i\left(k_i-1\right)}\sum_{j,k}\sqrt[3]{\widehat{w}_{ij}\widehat{w}_{ik}\widehat{w}_{jk}} \tag{4}$$

where w_{ij} is the strength of a connection between v_i and v_j and $\widehat{w}_{ij}=w_{ij}/max(w)$ (Onnela, Saramaki, Kertesz, & Kaski, 2005). It follows that the average CC for a weighted graph represents the strength of connectivity of neighbors.

The characteristic PL was calculated as follows:

$$\mathrm{PL}=\frac{1}{N(N-1)}\sum_{i\neq j}d_{ij} \tag{5}$$

where d_{ij} is the shortest path between v_i and v_j and d_{ij} = $1/w_{ij}$ (Newman, 2001). For a binary graph, the shortest path represents the smallest number of edges that connect nodes i and j. For a weighted graph, the shortest path is calculated as $d_{ij}=1/w_{ij}$, where w_{ij} is the edge weight between the nodes i and j. Considering the direct proportionality between the edge weight and the activity similarities between two nodes, the shortest path between two nodes is the one which maximize the edge weight.

The SWP was calculated as follows (Muldoon et al., 2016):

$$\mathrm{SWP}=1-\sqrt{\frac{\Delta_C^2+\Delta_L^2}{2}} \tag{6}$$

where

$$\Delta_C=\frac{\mathrm{CC}_{\mathrm{Lattice}}-\mathrm{CC}_{\mathrm{Real}}}{\mathrm{CC}_{\mathrm{Lattice}}-\mathrm{CC}_{\mathrm{Random}}} \tag{7}$$

$$\Delta_L=\frac{\mathrm{PL}_{\mathrm{Real}}-\mathrm{PL}_{\mathrm{Random}}}{\mathrm{PL}_{\mathrm{Lattice}}-\mathrm{PL}_{\mathrm{Random}}} \tag{8}$$

Δ_C and Δ_L represents the fractional deviation of the measured $\mathrm{CC}_{\mathrm{Real}}$ and $\mathrm{PL}_{\mathrm{Real}}$ from its null models (lattice and random). The SWP represents the likelihood that a network possesses high CC and small PL if compared to both lattice and random null models.

The network degree was computed as:

$$E=\frac{1}{2}\sum_{ij}w_{ij} \tag{9}$$

where w_{ij} is the weight edge between nodes i and j.

The mean edge-based centrality was computed by averaging the values of centrality computed using the *centrality.m* MATLAB function, PageRank type. The measure scales the contribution that any neighbor of a node makes to its centrality by the degree of such neighbors.

The modularity is defined as:

$$Q=\frac{1}{2W}\sum_{ij}(w_{ij}-e_{ij}^{w})\delta\left(m_i,m_j\right) \tag{10}$$

where $W=\frac{1}{2}\sum_{ij}w_{ij}$ is the total weight of the unique edges of the network, w_{ij} is the weight of the edge linking nodes i and j, e_{ij}^{w} is the connectivity weight on edges linking nodes in the same community that is expected by chance, and $\delta\left(m_i,m_j\right)$ is the Kronecker delta function (equals to 1 if nodes i and j belong to the same module, and 0 otherwise).

2.4 | Null models

The generation of the lattice and null models was performed as previously described (Muldoon et al., 2016) (Maslov & Sneppen, 2002; Sporns & Zwi, 2004) while preserving degree distribution.

2.5 | Meta-analysis network of human whole-brain functional coactivation

The coactivation matrix used in Figure 5, with x, y, and z coordinates, was previously defined and characterized in Crossley et al., (2013). Briefly, Crossley and colleagues estimated the similarity (Jaccard index) of the activation patterns across experimental tasks between each pair of nodes (638) based on more than 1,600 studies of task-related activation acquired using fMRI or PET scan techniques. The subgraph extraction was performed for the distance-dependent model, with the difference of a volume rather than a surface extraction.

3 | RESULTS

3.1 | The spatial neighbors sampling (SNS) method applied to the in silico models

The SNS approach aims to retain spatial neighboring nodes in the subsample. The definition of spatial neighbors is either based on node indexing (nodes whose indices are consecutive integers in a given interval as seen in the Watts-Strogatz models) or Euclidian distance (nodes with coordinates within a 2D surface as seen in the distance-dependent model). The extracted graph is an induced subgraph as it contains all the edges connecting pairs of retained nodes. Apart from node, link, or snowball sampling methods, we did not use any topological information about the global network during the subgraph extraction. This approach was aimed to mimic the lack of information of the global network while optically selecting a single field-of-view.

Using a Watts-Strogatz network model (either unweighted or weighted) and the SNS method, a subgraph (S subscript) is extracted from the global graph (G subscript) as follows (Figure 1a):

1. Given N_G, k, and β (number nodes, average degree, and rewire probability, respectively), generate a Watts-Strogatz global network (either unweighted or weighted; see Material and Methods) and calculate SWP_G, CC_G, and PL_G. β were intentionally selected to achieve small-world networks;
2. Fix the percentage of nodes to extract, and calculate the resulting nodes N_S to be retained in each subgraph;
3. Choose a seed node, *s,* of index *i*;
4. Extract all the nodes with an index $\in \left[i;i+N_S-1\right]$. The edges are maintained if and only if both nodes are retained in the subgraph;
5. Calculate SWP_S, CC_S, and PL_S in the subgraph of dimension N_S created in step 4;
6. Repeat the process from step 3 to step 5 with the node of index $i + 1$ until the index of the seed node is equal to $i-1$ (to extract all the N_G subgraphs of dimension N_S).

Because the Watts-Strogatz model neglects any spatial information and considering the intrinsic spatial nature of the field-of-view selection, we developed a model which implemented spatial coordinates (distance-dependent model). We randomly seeded nodes and connected them to achieve a defined connectivity and a small-world topology. The attachment of two nodes was a function of the Euclidian distance between them; even if simplified, this approach has been previously shown to sufficiently mimic the neuronal architecture. To allow long-range connections, yet without imposing additional arbitrary spatial constrains, a beta-rewiring step was introduced. A single subgraph was then extracted selecting all the nodes contained in a surface of defined dimension. The subgraph extraction was repeated by considering each node as the center of the region, creating as many independent subgraphs as many nodes in the global graph. This mimicked the random and blind selection of a single field-of-view in real-world imaging experiments. Using a distance-dependent model and the SNS method, a subgraph was extracted as follows (Figure 1b):

1. Given N_G nodes, arrange them randomly in a 2D space. The 2D space has x and y coordinates ranging between 0 and 1;
2. Assuming an inverse relationship between physical distance and edge strength (Muldoon et al., 2016), assign edge weights w_{ij} according to the Euclidian distance d_{ij} between all the pairs of nodes as follows:

$$w_{ij} = D_{max} - d_{ij}$$

where $D_\text{max} = max\{d_{ij}\}$. Each node will have N_G-1 connections (except for the *ij* pair with distance d_{ij} equal to D_{max});

3. Eliminate connections with an edge weight below a weight threshold w_t. w_t is defined to achieve an average edge density between 7% and 10% in the network (She et al., 2016). This step led to the spontaneous emergence of small-world networks (Figure S1a);
4. Randomly rewire each edge with probability β (edge weight is retained). The inclusion of random rewiring guarantees that the network is not solely constructed as a function of physical distance, yet without imposing any additional rule. Moreover, as clear from Figure S1b, the introduction of the random rewiring led to a similar profile previously found in Watts & Strogatz, (1998);
5. Calculate SWP_G, CC_G, and PL_G;
6. Fix a spatial parameter rho (ρ) and a seed node, *s*, of index *i*;
7. Extract all the nodes with x coordinates $\in \left[x_s-\rho;x_s+\rho\right]$ and y coordinates $\in \left[y_s-\rho;y_s+\rho\right]$. An edge *ij* is preserved if and only if both nodes are retained in the subgraph;
8. Calculate SWP_S, CC_S, and PL_S of the subgraph created in step 7;
9. Repeat the process from step 6 to step 8 with a node of index $i + 1$ until the index of the seed node is equal to $i-1$ (to guarantee that all the N_G nodes of the global network are used as seed nodes).Repeat the process from step 6 to step 8 with a node of index $i + 1$ until the index of the seed node is equal to $i-1$ (to guarantee that all the N_G nodes of the global network are used as seed nodes).

3.2 | Impact of the SNS method on the Watts-Strogatz model

We tested the effect of the SNS method on the average CC, average *PL*, and SWP as a function of the size (number of nodes N_G: 250, 500, 1,000, 1,500, 2000, and 5,000) and β (0.05, 0.01, and 0.005) used to construct the global Watts-Strogatz network (k = 15). For each condition (N-β pair), we extracted the subgraphs varying the number of nodes N_S to retain. Figure 2 summarizes the variation of CC, PL, and SWP (expressed as a fold ratio between the parameter in the subgraph and in the global graph; Figure 2a–c, respectively) as a function of node's

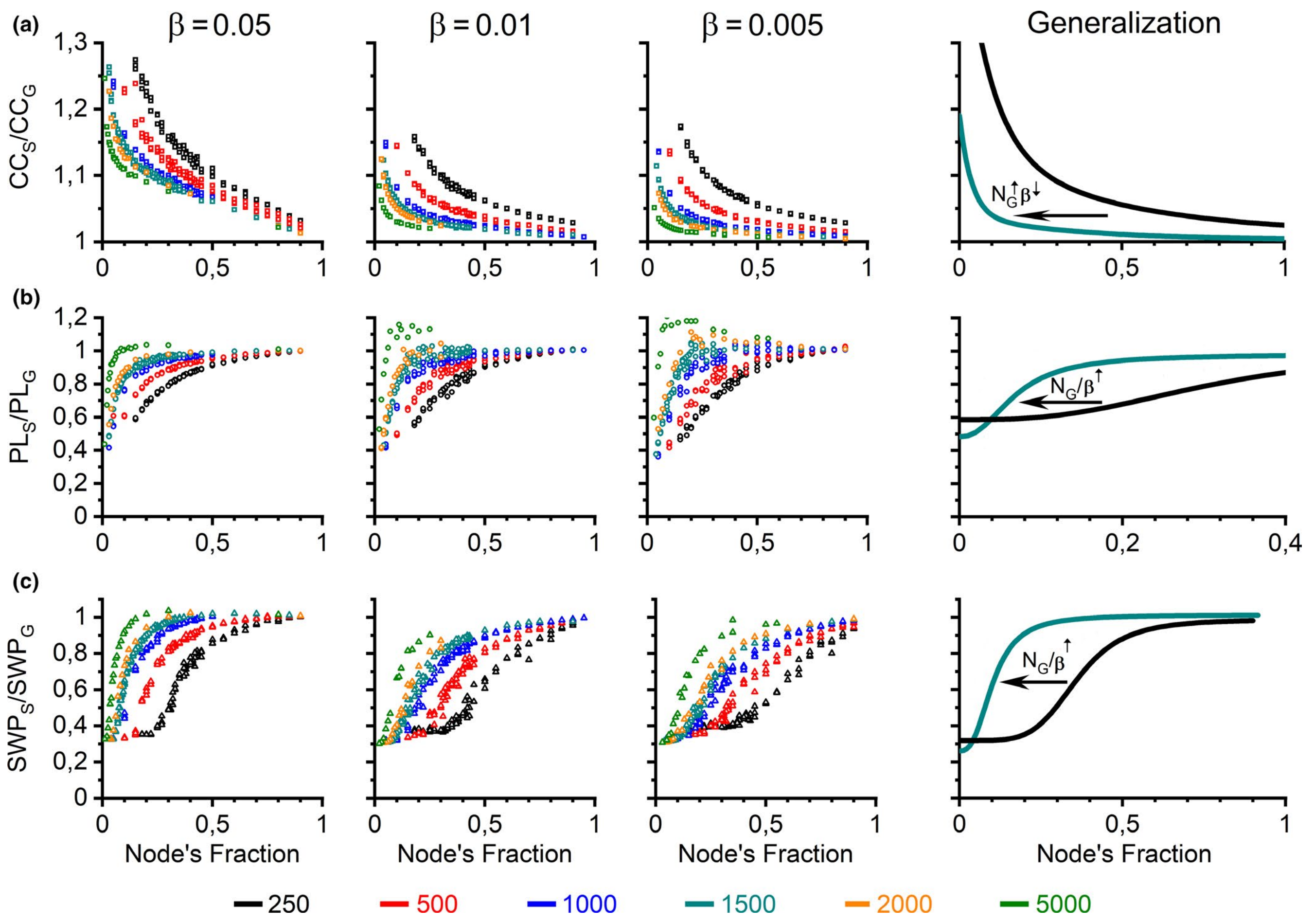

FIGURE 2 N_G and β dependence of subgraph parameters derived from the application of the SNS method to Watts-Strogatz graphs. Each point is the ratio between the (a) average clustering coefficients (CC), or (b) average path length (PL) or (c) small-world propensity (SWP) of the subgraph and the global graph from which is extracted (S and G, respectively) at different sizes and rewiring probabilities of the global network, as indicated. On the right, the fitting of the conditions $N_G = 250$ and $N_G = 1{,}500$ are used to graphically summarize the performance of each parameter as a function of N_G and β. Biexponential decay fitting: $(y=y_0+A_1e^{-\frac{x}{t_1}}+A_2e^{-\frac{x}{t_2}})$. Logistic fitting: $(y=\frac{A_1-A_2}{1+(\frac{x}{x_0})^p}1+A_2)$

fraction (calculated as the ratio N_S/N_G) for each condition, as indicated. The performance of each parameter was then graphically summarized in the right panel (generalization) as a function of both N_G and β using a fitting of the conditions $N_G = 250$ (black line) and $N_G = 1{,}500$ (green line).

The CC behavior (Figure 2a) can be approximated by a biexponential decay fitting $(y=y_0+A_1e^{-\frac{x}{t_1}}+A_2e^{-\frac{x}{t_2}})$, with a decay rate directly proportional to N_G and inversely proportional to its regularity (namely, inversely proportional to the β value). As the CC$_S$/CC$_G$ ratio is greater than 1 throughout the entire range of node fractions, it follows that CC$_S$ >CC$_G$, and thus CC$_S$ is overestimated for any subgraph. Moreover, for a given N_G, the deviation of CC$_S$ from CC$_G$ was more pronounced with an increased β. In summary, the larger ($N_G^{\uparrow}$) and more regular ($\beta^{\downarrow}$) the global network, the smaller the deviation of CC$_S$ from CC$_G$.

The PL behavior (Figure 2b) can be approximated with a logistic fitting $(y=\frac{A_1-A_2}{1+(\frac{x}{x_0})^p}1+A_2)$, where the curve steepness is directly proportional to both N_G and β. The PL$_S$/PL$_G$ ratio can be either greater or smaller than 1. Greater N_G and smaller β (larger and more regular global graphs) create a higher likelihood that PL$_S$/PL$_G$ >1, shifting the entire profile from a pure logistic relationship (see Figure 2b, $N_G = 5{,}000$, $\beta = 0.005$). In summary, the larger and more random the global network ($N_G^{\uparrow}$ and $\beta^{\uparrow}$), the smaller the deviation of PL$_S$ from PL$_G$ becomes.

The SWP behavior (Figure 2c) can be approximated by a logistic fitting $(y=\frac{A_1-A_2}{1+(\frac{x}{x_0})^p}1+A_2)$, with the steepness of the curve is directly proportional to both N_G and β. The SWP$_S$/SWP$_G$ is usually smaller than 1; however, there are cases in which the ratio is larger than 1. In summary, the larger and more random the global network ($N_G^{\uparrow}$ and $\beta^{\uparrow}$), the smaller the deviation of SWP$_S$ from SWP$_G$. The reduction of the SWP in the subgraphs as a function of the node fraction poses a question on their classification: while approaching the lower limit, the subgraph could either be more regular or more random when compared to the global graph. To solve this issue, we report in Figure S2 the trends of ΔC and ΔL (black square and red circle, respectively) as indicators of the subgraph behavior as a function of node fraction, N_G and β. The decreasing trend of ΔC (decreased variation from the lattice model) was always paired with an increasing trend of ΔL (increased variation from the random model) as a function of the node fraction. Therefore, all conditions classified the subgraphs as more regular than the global graph when the node fraction decreased. The slope of ΔC and ΔL is again a function of both N_G and β.

3.3 | Impact of the SNS method on the sum of the local clustering coefficients and of the shortest path length

The results summarized in Figure 2 clearly depict the conclusion that the information of N_G, paired with the information gained by the analysis of the subgraphs, was not sufficient to predict SWP_G, CC_G, and PL_G. Indeed, the knowledge of β in the global graph was necessary as well. We thus tested if, rather than the average CC and PL values, the calculation of the sum of local CC_S and of the shortest PL could be beneficial in the prediction of the global parameters. As the average CC is calculated as follows:

$$\mathrm{CC}=\frac{1}{N}\sum_{i=1}^{N}C_i \tag{11}$$

where N is the number of nodes and C_i is the local CC of the ith node, we can compute $\sum_{i=1}^{N}C_i$:

$$\sum_{i=1}^{N}C_i=CC\cdot N \tag{12}$$

Using Equation 12, we can calculate the ratio between the sum of the local $\mathrm{CC_S}$ of the subgraph and the sum of the global network. The ratio $\left(\sum_{i=1}^{N_S}C_i\right)_S/\left(\sum_{i=1}^{N_G}C_i\right)_G$ is then graphed with the respective node fraction in Figure 3a (top-Real Graph), in which data from global networks of different sizes (as indicated) and β are combined. The ratio of the local $\mathrm{CC_S}$ scales linearly with the node fraction, as confirmed by the line of best fit (blue line; $y = sx + q$).

As the average PL is calculated as follows:

$$\mathrm{PL}=\frac{1}{N(N-1)}ij\neq\widehat{\sum}d_{ij} \tag{13}$$

where N is the number of nodes and d_{ij} is the shortest path between nodes i and j, we can compute $\sum_{i\neq j}d_{ij}$:

$$ij\neq\widehat{\sum}d_{ij}=\mathrm{PL}\cdot N(N-1) \tag{14}$$

Using Equation 14, we can calculate the ratio between the sum of the local PL of the subgraph and the sum of the

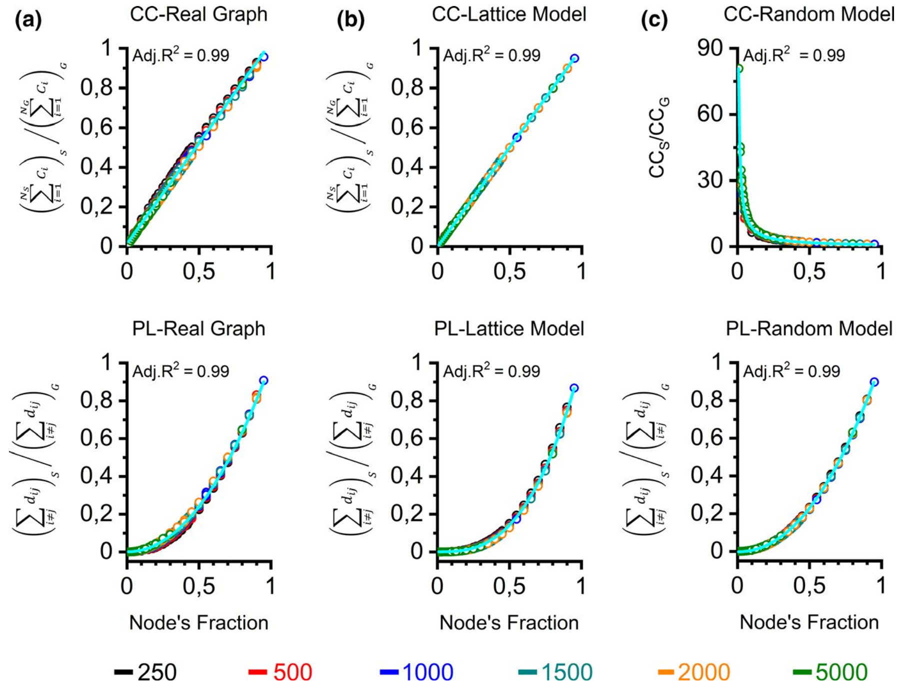

FIGURE 3 Relationship between either the sum of the local clustering coefficients or the sum of the shortest path length ratios and the node's fraction. The $\left(\sum_{i=1}^{N_S}C_i\right)_S/\left(\sum_{i=1}^{N_G}C_i\right)_G$—node fraction functions are shown for (a) the real graph (top) and (b) the lattice model (top). (c) The $\mathrm{CC_S/CC_G}$—node's fraction relation is graphed for the random model. The $\left(\sum_{i\neq j}d_{ij}\right)_S/\left(\sum_{i\neq j}d_{ij}\right)_G$—node fraction relationships are shown for the real, lattice, and random models (a–c, respectively; bottom). The graphs are made by combining different N_G (color-coded, as indicated) and β (0.005–0.01–0.05). The fitting is indicated with a light blue curve. Linear fitting: $y = sx + q$; Power fitting: $y=x^{\lambda}$. CC-Real: s = 1.015 ± 0.002, q = 0.0016 ± 0.0008; CC-Lattice: s = 0.999 ± 0.0001, q = 0.002 ± 0.00005; CC-Random: λ=−0.953 ± 0.0007; PL-Real: λ = 2.068 ± 0.002; PL-Lattice: λ = 2.828 ± 0.003; PL-Random: λ = 2.132 ± 0.00008

global network. The ratio $\left(\sum_{i\neq j} d_{ij}\right)_S / \left(\sum_{i\neq j} d_{ij}\right)_G$ is then graphed with the respective node fraction in Figure 3a (bottom-Real Graph), in which data from global networks of different sizes (as indicated) and rewiring probabilities are combined. The ratio of the local PLs that scaled with the node fraction follows a power law (blue line;$y = x^{\lambda}$). We then computed the CC_G and PL_G (Equations 6–10; See Supplementary Material) using the size of the global network (N_G), the size of the sampled network (N_S), and the calculated CC_S or PL_S. In Figure S3a,b, the distribution of error for the predicted CC_G and PL_G values is shown binning the node fraction (bin size = 0.1), revealing a larger error of prediction with lower node fractions.

To predict the SWP_G, we studied how the null models (lattice and random, which are used for the SWP computation (Muldoon et al., 2016)) scaled. In particular, the $\left(\sum_{i=1}^{N_S} C_i\right)_S / \left(\sum_{i=1}^{N_G} C_i\right)_G$ and $\left(\sum_{i\neq j} d_{ij}\right)_S / \left(\sum_{i\neq j} d_{ij}\right)_G$ ratios were calculated using the values from the null models of the subgraph and global networks (Figure 3b,c). We found that the CC_Lattice, PL_Lattice, and PL_Random models scaled similarly to the graphs shown in Figure 3a. Namely, the CC_Lattice model followed a linear function and the PL_Lattice and PL_Random models followed a power-law function. However, the $\left(\sum_{i=1}^{N_S} C_i\right)_S / \left(\sum_{i=1}^{N_G} C_i\right)_G$ ratio for the CC_Random models was not proportional to the node fraction; on the other hand, the $\text{CC}_\text{S}/\text{CC}_\text{G}$ ratio scaled with power profile (Figure 3c, top; $y = x^{\lambda}$). The authors do not have any explanation for this phenomenon. These relationships allowed us to compute the CC_G and PL_G for the lattice and random models. Using those values, we finally calculated the SWP_G (Muldoon et al., 2016); the distribution of error for the predicted SWP_G values is summarized in Figure S3c indicating an inverse relation between the accuracy of the SWP_G prediction and the subgraph's size. Also, we used the condition $N_G = 250$ to confirm that the abovementioned relations were not affected by either a binary versus a weighted global network comparison (Figure S4) or a different value of k ($k = 15$ vs $k = 30$; Figure S5).

3.4 | The SNS method applied to a model with spatial coordinates

One limitation of the Watts-Strogatz model is the lack of spatial coordinates associated with each node. In order to explore the impact of the SNS method considering the spatial proximity of the nodes, we generated the distance-dependent model (Figure 1b). We subsequently analyzed the impact of the SNS method on the CC, PL, and SWP values (Figure 4a) as a function of the size of the global distance-dependent models (number of nodes, N_G: 500, 1,000, 1,500). The same profiles found in Figure 2 emerged using the distance-dependent

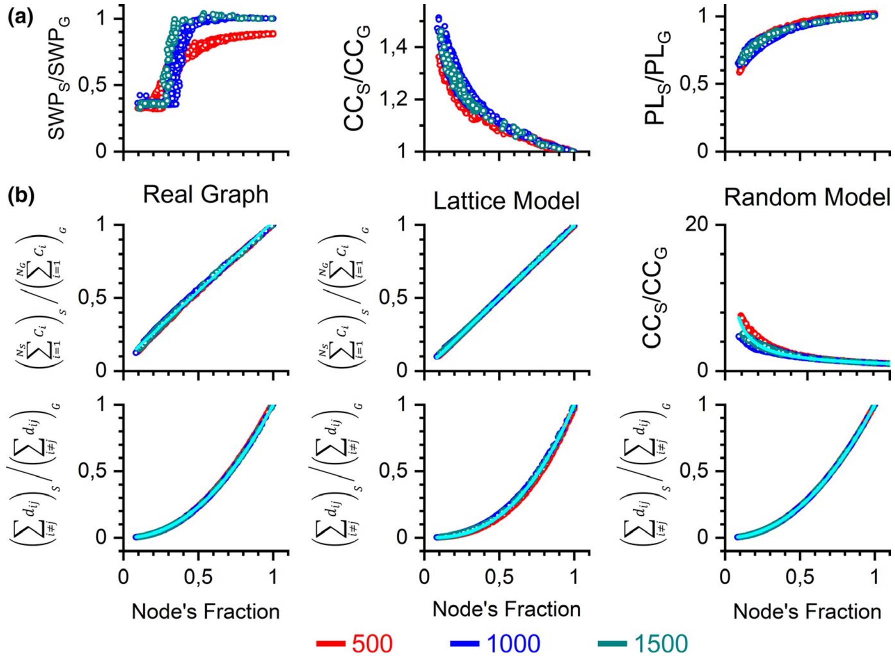

FIGURE 4 SNS method applied to a distance-dependent model. (a) Each point is the ratio between the SWP, CC, and PL values (left to right panels) of the subgraph and the global graph from which is extracted (S and G, respectively) at different sizes ($N_G = 500$, 1,000, and 1,500). (b) The $\left(\sum_{i=1}^{N_S} C_i\right)_S / \left(\sum_{i=1}^{N_G} C_i\right)_G$—node fraction functions are shown for the real graph and the lattice model. The $\text{CC}_\text{S}/\text{CC}_\text{G}$—node's fraction relation is graphed for the random model. The $\left(\sum_{i\neq j} d_{ij}\right)_S / \left(\sum_{i\neq j} d_{ij}\right)_G$—node fraction relationships are shown for the real, lattice, and random models. The graphs are made by combining different N_G (color-coded, as indicated) with $\beta = 0.05$. The fitting is indicated with a light blue curve. Linear fitting: $y = sx + q$; Power fitting: $y = x^{\lambda}$. CC-Real: s = 0.951 ± 0.001, q = 0.007 ± 0.0004; CC-Lattice: s = 0.985 ± 0.0002, q = 0.008 ± 0.0001; CC-Random: λ=−0.800 ± 0.002; PL-Real: λ = 2.101 ± 0.001; PL-Lattice: λ = 2.601 ± 0.003; PL-Random: λ = 2.100 ± 0.0004

models; namely a sigmoidal profile for both the SWP_S/SWP_G and PL_S/PL_G ratios and a biexponential profile for the CC_S/CC_G ratio. We then applied the relationships described in Equation 12 and Equation 14 (Figure 4b; empty circles) and found the same type of relationships summarized in Figure 3, indicating that similar rules dictate the scaling behavior of the parameters in the distance-dependent model graphs. We finally report in Figure S6 the error distributions of the predicted global graph descriptors by means of the fitting parameters binning the node fraction (bin size = 0.1); as previously found, the smaller the node's fraction, the less accurate the prediction becomes.

3.5 | The SNS method applied to a human coactivation matrix

We finally applied the SNS method to human real-world data; namely the coactivation matrix previously obtained by Crossley and colleagues acquired using fMRI or PET scan techniques was chosen (Crossley et al., 2013). We selected this graph as it is arranged in a small-world fashion ($SWP_G = 0.457$), and unlike the simulations, it does contain hub nodes, it is represented in a 3D volume, and the number of regions (638) is comparable with the data obtained from the presented simulations. Representative subgraphs (blue nodes) obtained upon application of the SNS method to the global graph (red nodes) are shown in Figure 5a for different views (inferior, superior, and lateral, respectively). Unlike Figures 2 and 4a, the profile of the ratio of values (Figure 5b) was not as defined, particularly for the SWP_S/SWP_G ratio. On the other hand, the application of the relationships between eq. 12 and eq. 14 was qualitatively invariant when compared to Figures 3 and 4 (the percentage of error distributions for the predicted global descriptors is shown in Figure S7). It is interesting to report a branching of the $\left(\sum_{i=1}^{N_S} C_i\right)_S / \left(\sum_{i=1}^{N_G} C_i\right)_G$ values for both the real data and lattice model when the node's fraction was greater than 0.5. This is a consequence of a broad distribution of CC_i (average CC value of the ith node) in the original graph, as shown in Figure S8, where the estimated density (Kernel smoothing) of CC_i values is shown for all the models. Moreover, we confirmed that, independently of the total area extracted from the human data (namely independently from the chosen ρ), the subgraphs represented in the high branch (black) were composed by a subpopulation of nodes with higher CC_i, as demonstrated by a rightward shift of the estimated density and a significantly higher mean and area above $CC_i = 0.5$ when compared to the subgraphs populating the low branch (red). On this dataset, we also tested the behavior of three more topological descriptors: network degree, mean degree-based centrality (PageRank), and modularity. Because in the in silico models we set the degree, the resulting narrow Gaussian degree distributions limit the potential diversity of connectivity, necessary for a proper partition of the graph in modules or a diversity in the centrality measures. Differently, the human dataset possesses the diversity necessary to test such parameters. The degree ratio between sampled and global graphs is shown in Figure 5d, indicating a power-like relation. A similar behavior was evident while plotting on the y-axis the ratio between the mean PageRank centrality measures (Figure 5e). While comparing the modularity ratios as a function of the node's fraction, we failed to detect a clear relationship (Figure 5f). As the modularity is defined as:

$$Q = \frac{1}{2W} ij \neq \widehat{\sum}(w_{ij} - e_{ij}^{w})\delta\left(m_i, m_j\right) \tag{15}$$

where $W = \frac{1}{2}\sum_{ij} w_{ij}$ is the total weight of the unique edges of the network, w_{ij}, is the weight of the edge linking nodes i and j, e_{ij}^{w} is the connectivity weight on edges linking nodes in the same community that is expected by chance, and $\delta\left(m_i, m_j\right)$ is the Kronecker delta function (equals to 1 if nodes i and j belong to the same module, and 0 otherwise), we decided to compute for the sampled modularity by the following parameter:

$$Q_S^{*} = Q_S \cdot 2W_S = ij \neq \widehat{\sum}(w_{ij} - e_{ij}^{w})\delta\left(m_i, m_j\right) \tag{16}$$

This formulation is thus normalized by the total weight of each unique sampled graphs. The Q_S^{*}/Q_G ratio, where Q_G is the modularity of the global graph, led to the emergence of a power-like relation with the node's fraction (Figure 5g). We did not apply Equation 16 to Q_G as the W_G value is unknown in real experiments. The robustness of those approximations can be inferred by the error distributions of the predicted global graph descriptors by means of the fitting parameters binning the node fraction (bin size = 0.1) in Figure S9.

3.6 | Data comparison

We finally compared the precision of the predictions of the graph descriptors using the results obtained in the aforementioned datasets. The average percentage of error for the predictions of the parameters is shown in Figure 6 for all the dataset (Watts-Strogatz, distant-dependent model and human data; black, red, and blue, respectively) as a function of binned node intervals (bin size = 0.05). As expected, the error was inversely proportional to node fraction for all the datasets. We finally overlapped the fitting generated for all the conditions (Figure 6b). Even if they shared qualitative similarities, the lines of best fit were dataset dependent and could be differentiated into simulations (black and red lines) and real data (blue line). In summary, the properties scale with the subgraph's size and, even if relationships can be established between a given descriptor in the subgraph versus

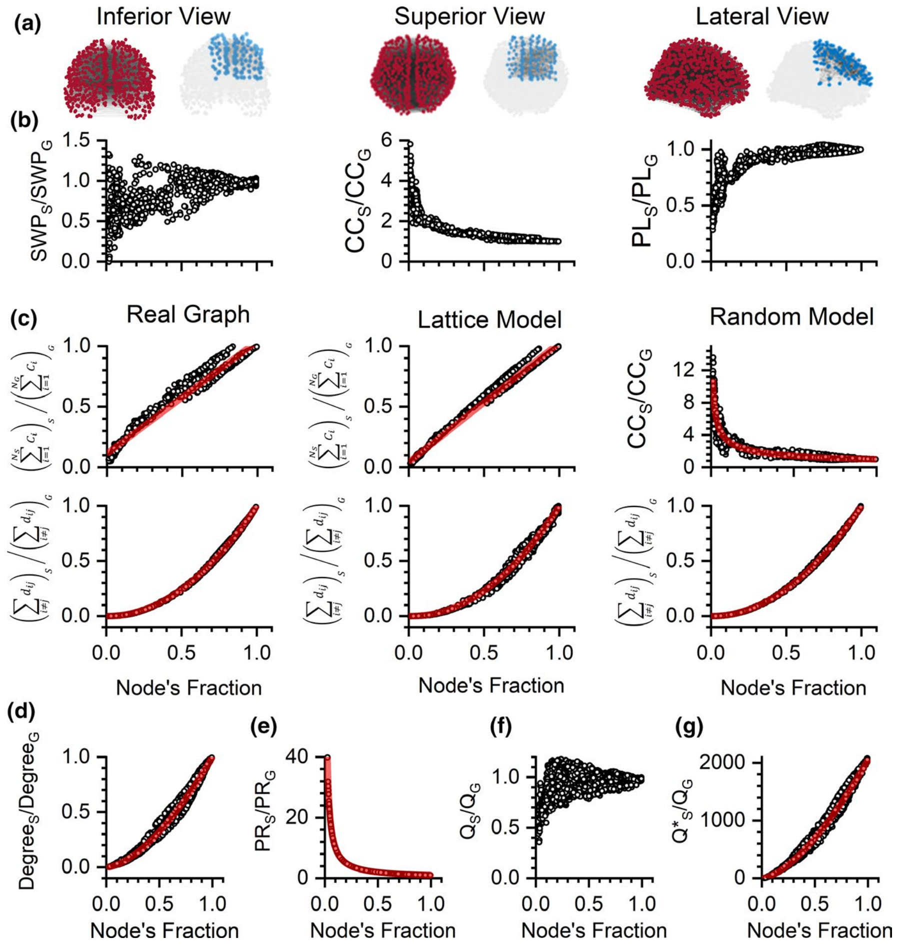

FIGURE 5 SNS method applied to a human coactivation matrix. (a) Representative subgraphs (blue nodes) isolated from the global graph (red nodes) in the 3D volume. (b) Ratio values of the SWP, CC, and PL (left to right, as indicated) in the subgraph versus the global graph. (c) The $\left(\sum_{i=1}^{N_S} C_i\right)_S / \left(\sum_{i=1}^{N_G} C_i\right)_G$—node fraction functions are shown for the real graph and the lattice model, as indicated. The CC_S/CC_G—node fraction relationship is graphed for the random model (top). The $\left(\sum_{i \neq j} d_{ij}\right)_S / \left(\sum_{i \neq j} d_{ij}\right)_G$—node fraction relationship is shown for the real, lattice, and random models (bottom). D-G) The ratio values of the network degree, mean PageRank (PR), modularity (Q), and transformed modularity (Q*) in the subgraph versus the global graph. The line of best fit for each condition is indicated by the red line. Linear fitting: $y = sx + q$; Power fitting: $y = x^{\lambda}$ Allometric: $y = ax^{\lambda}$. CC-Real: s = 0.928 ± 0.004, q = 0.114 ± 0.002; CC-Lattice: s = 0.985 ± 0.002, q = 0.043 ± 0.001; CC-Random: λ=−0.523 ± 0.002; PL-Real: λ = 2.032 ± 0.003; PL-Lattice: λ = 2.117 ± 0.008; PL-Random: λ = 2.020 ± 0.002; Degree: λ=−0.1.630 ± 0.004; PR: λ=−1 ± $3.9.10^{-6}$;Q*: a = 2046±3.42,: λ=−1.56 ± 0.005

global graph, the prediction accuracy was inversely correlated with the subgraph's size. Moreover, despite the fact that the relationships in Figure 6b were qualitatively similar, we cannot assume a general fitting irrespective of the analyzed dataset (simulations vs real data), as indicative of different rules shaping human brain connectivity.

3.7 | Limitations

As the study was conducted by means of numerical analysis, we cannot assume a generalization of the aforementioned relationships, as evident from Figure 6b. It follows that the quantitative descriptions in other scenarios, such as different generative models or a much higher N_G, must be calculated rather than be assumed. Moreover, it is important to underline that the data transformation in Figures 3–5 did not abolish the N_G-dependence of the parameters. As the percentage of error of those predictions increased while the node's fraction decreased, this does limit the usefulness of the identified relationships for the prediction of the parameters when considering low node fractions. In the present manuscript, we limited our study to just CC, PL, and SWP; however, the same pipeline can be easily adapted to study other descriptors as well. It is important to consider that the construction of the presented in silico models forces the user to set the mean degree, de facto limiting the analysis of parameters (i.e. centrality, hubs, and modularity) that rely on broad degree distribution. Finally, due to computational limits, the range of the nodes N_G explored was limited.

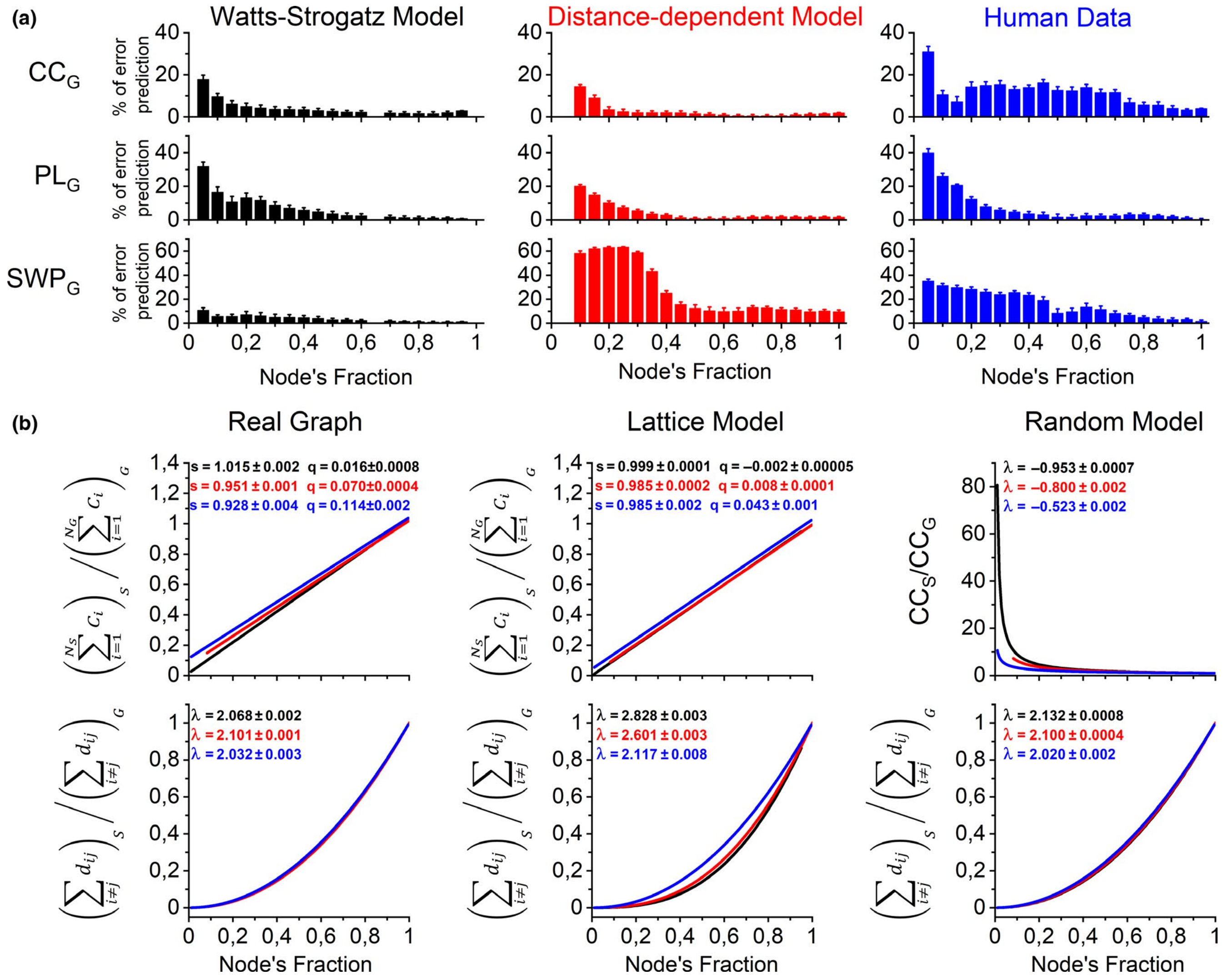

FIGURE 6 Accuracy of predictions and fitting comparison across datasets. (a) Bar graphs summarizing the percentage of error prediction (mean ± *SEM*) of CC_G, PL_G, and SWP_G (top to bottom, respectively) in the datasets using the fitting models. (b) Fitting comparisons between the datasets. The fitting parameters are reported (±*SD*) for all the datasets, as indicated. Watts-Strogatz: Black; Distance-dependent model: red; Human data: blue. Linear fitting: y = sx +q; Power fitting: $y=x^{\lambda}$

4 | DISCUSSION

Every time we observe a neuronal network under the microscope, we inevitably visualize just a fraction of the system, and the robustness of any statements must face the reality of this subsampling. Indeed, anytime a graph (which is a common representation of a neuronal network) is sampled, the graph's descriptors are not invariant (Stumpf et al., 2005), (Lee et al., 2006), and (Levina & Priesemann, 2017). In other words, a small fraction of the graph has different characteristics when compared to the global graph from which it was extracted. Moreover, we need to consider the methods by which the subgraphs were extracted; they usually use topological information from the global graph to operate the extraction. The field-of-view selection reduces the original samples to those cells that are just in physical proximity, independent of and blinded to the topology of the sample itself. We thus introduced the SNS method, in which the nodes of the subgraphs were selected ignoring any topological relations between them. Here we tested if the subsampled networks, created by a single selection of nodes of spatial neighbors, could carry enough information to predict global network descriptors. The key concept is the definition of spatial neighbors: two nodes are considered spatial neighbors if either (a) their indices are consecutive integers in a given range (graphs in which there are no spatial coordinates associated with each node—Watts-Strogatz) or (b) their node coordinates fall within a defined surface/volume (graphs in which the spatial position of the nodes is defined—distance-dependent model and human data). The absence of any topological bias between the nodes of the subgraphs correctly mimicked the random, blind selection of a single field-of-view while imaging. The global network was selected to be small-world, because data indicate that this topology is common while studying neuronal networks (Bassett & Bullmore, 2017). We took advantage of three different in silico models, namely the Watts-Strogatz (binary or weighted) and a distance-dependent approach (Figure 1). We focused our study on three parameters commonly used to describe and characterize a neuronal network: the average CC, the average PL, and the SWP, which combines both CC and PL in reference to null models (lattice and random) (Muldoon et al., 2016). Independently from the model used to generate the global graphs, all the parameters changed when the subgraphs became smaller (Figures 2–4). This finding indicated, as expected, that the small-world graphs were not self-similar across scales. In particular, it was clear that the $\mathrm{CC_S}$ was always overestimated in the sampled graphs, while the $\mathrm{PL_S}$ showed a N_G-dependence of its behavior (the greater the N_G is, the larger the fraction of $\mathrm{PL_s}$ values that are overestimated rather than underestimated in the sampled graphs) and the $\mathrm{SWP_S}$ was more often underestimated. Furthermore, a topology dependency was evident because the same parameters scaled differently based on the generative models (Figure 2 vs Figure 4a; Watts-Strogatz vs distance-dependent model, respectively) or the topological features for a given model (Figure 2, Watts-Strogatz model: $\beta = 0.05$ vs $\beta = 0.01$ vs $\beta = 0.005$), even if the overall behavior was largely maintained.

It follows that the actual topology of the global network, in addition to the results of the subgraph analysis, is necessary to study the scaling behaviors of $\mathrm{CC_G}$, $\mathrm{PL_G}$, and $\mathrm{SWP_G}$. Nevertheless, the knowledge of the topology is usually the desired output and not a priori knowledge while studying neuronal networks. Shifting the attention from the average values of CC and PL, the new parameters isolated by means of the Equations 12 and 14 minimized, yet did not eliminate the topological and node dependence previously found. Even if they are approximations, the relationships summarized in Figure 3 and 4b enlighten a clear scaling behavior and allowed us to make predictions of $\mathrm{CC_G}$, $\mathrm{PL_G}$, and $\mathrm{SWP_G}$ based on the subgraph values combined with the knowledge of the size of the global graph; such relations were empirically chosen and were not derived from first principles. In an in vitro experiment, the size of the global graph was represented by the number of seeded neurons, which is known by the experimenter. Even if the Watts-Strogatz and the distance-based models are constructed differently and their definition of neighbors differs, they both showed a strong bias toward local connectivity (they were, in fact, selected to be small-world) and it should not be a surprise that the SNS method led to similar results; we also adopted a similar conjecture for neurons. Indeed, we can assume that the probability that two neurons are connected rapidly decreases, but it is not null, when the distance between them increases, as previously outlined in Crossley et al., (2013). Namely, they showed that in a meta-analysis reconstruction of a human coactivation network, most connections or edges were short distance (median length of 57 mm; significantly shorter than random networks), and relatively few edges were long distance. Using this same dataset, we tested the effect of the SNS method on human data. We specifically chose this dataset as it was constructed from a large dataset; the selection could have been applied in a 3D volume with the number of nodes comparable to our simulations and, most importantly, it contained hubs. Indeed, the lack of a power-law distribution of the edge degree has to be acknowledged in both the Watts-Strogatz and distance-dependent models. We found the behavior of the ratio values for the human data (Figure 5b) less clear, yet the transformation of the ratios led to the same type of relationships (Figure 5c) previously found in the simulations (Figures 3 and 4), indicating a similarity in the scaling behavior across these datasets. For the same reasons, we also tested the impact of the SNS method on network degree, mean degree-based centrality, and modularity. We found that also for these descriptors, their global values can be inferred from their recorded values in the subgraphs.

Despite sharing similarities with the reports of Gerhard et al., (2011) and She et al., (2016), such as the use of random sampling regions with defined spatial resolutions and the use of distance-dependent models to study the sampling process, several differences emerged as well. The computation of the small-worldness metric differed; the SWP metric has the advantage to be robust in respect to network size, a crucial feature while comparing sampled versus global networks, and can be applied to weighted networks. The SNS method led to the identification of a single subgraph in which the nodes are spatial neighbors to each other while the multi-electrode approach led to the identification of a subgraph from multiple sites in which the nodes are not necessarily in physical proximity but are selected to be neighbors to the recording sites. These led to opposite conclusions; while She and colleagues and Gerhard and colleagues reported that the subgraph possessed a more pronounced small-worldness, we found that the subgraph is less small-world if compared to the global graph. It follows that, given a global network, the application of either the imaging or multi-electrode approaches led to different topological results and thus different biological conclusions about the sampled network.

Finally, it is evident from Figure 6a that the precision of the predictions decreased progressively when reducing the size of the subgraph in all the datasets. This implies that, below a certain size, even the above approximations failed to accurately foresee the parameter values in the global graphs. As the fittings of the Watts-Strogatz and the distance-dependent models were more similar to each other than the fittings obtained from the human dataset, we can assume that the human brain must follow different generative rules, which is expected. The authors suggest that the broader and leftward shifted probability distributions of the CC_i in the human versus simulated data, along with the presence of hub nodes in the human data, play a crucial role in guiding the different scaling performances. Yet, the similarities are qualitatively indisputable, and this most likely reflects the shared small-worldness.

Altogether, the data demonstrated that with the SNS method of subgraph extraction and independently from the type of small-world global graph: (a) the CC, PL, and SWP scale with the fraction of nodes retained in the subgraph, indicating a lack of self-similarity across the scales of the global graphs, and (b) an approximated prediction of the global values based on the subgraph analysis was possible, keeping in mind that its accuracy is directly proportional to the subgraph's size (the distributions of error can guide the experimenter in choosing a cut-off of accuracy). It follows that, below a certain subgraph size, we lose any predictive ability and, thus, the results obtained in the subgraph itself cannot be generalized to the entire network. In light of our findings, it is evident that the topologies resulting from brain recordings, such as fMRI, LFP, and EEG, or imaging of either brain structures or in vitro cultures must be carefully considered in view of the sampled network that is inevitably recorded. Any conceptual generalization to the entire sample, as well cross-comparison of network topologies reconstructed with different spatial resolutions, should be avoided. It thus appears a relevant try to close such gap in the near future.

In conclusion, given a graph and considering its topology, the magnitude of the subsampling (the ratio of the sampled and global graph sizes) is a critical parameter in deciding if the findings can or cannot be generalized to the whole global graph.

ACKNOWLEDGMENTS

We thank the NIH (P41EB027062, R01NS092847, and U01EB014976) for support of this work.

CONFLICT OF INTEREST

We declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Mattia Bonzanni: Conceptualization; Methodology; Investigation; Data curation; Formal analysis; Visualization; Project administration: Writing – Original draft; Writing - Review & Editing. Kimberly M. Bockley: Methodology; Investigation; Writing - Review & Editing. David L. Kaplan: Funding acquisition; Writing - Review & Editing.

DATA AVAILABILITY STATEMENT

The data and the codes for generating the models and apply the extraction are freely available on https://github.com/mattiabonzanni/Subgraph-extraction-with-spatial-neighbors-sampling-SNS-approach-.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.14937

ORCID

David L. Kaplan https://orcid.org/0000-0002-9245-7774

REFERENCES

Bassett, D. S., & Bullmore, E. T. (2017). Small-world brain networks revisited. *The Neuroscientist*, *23*, 499–516. https://doi.org/10.1177/1073858416667720

Bassett, D. S., Zurn, P., & Gold, J. I. (2018). On the nature and use of models in network neuroscience. *Nature Reviews Neuroscience*, *19*, 566–578. https://doi.org/10.1038/s41583-018-0038-8

Crossley, N. A., Mechelli, A., Vertes, P. E., Winton-Brown, T. T., Patel, A. X., Ginestet, C. E., … Bullmore, E. T. (2013). Cognitive relevance of the community structure of the human brain functional coactivation network. *Proceedings of the National Academy of Sciences*, *110*, 11583–11588. https://doi.org/10.1073/pnas.1220826110

Gerhard, F., Pipa, G., Lima, B., Neuenschwander, S., & Gerstner, W. (2011). Extraction of network topology from multi-electrode recordings: Is there a small-world effect? *Frontiers in Computational Neuroscience*, *5*, 4. https://doi.org/10.3389/fncom.2011.00004

Lee, S. H., Kim, P. J., & Jeong, H. (2006). Statistical properties of sampled networks. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *73*, 16102. https://doi.org/10.1103/PhysRevE.73.016102

Levina, A., & Priesemann, V. (2017). Subsampling scaling. *Nature Communications*, *8*, 15140. https://doi.org/10.1038/ncomms15140

Maslov, S., & Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Science*, *296*, 910–913. https://doi.org/10.1126/science.1065103

Muldoon, S. F., Bridgeford, E. W., & Bassett, D. S. (2016). Small-world propensity and weighted brain networks. *Scientific Reports*, *6*, 22057. https://doi.org/10.1038/srep22057

Newman, M. E. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *64*, 16132. https://doi.org/10.1103/PhysRevE.64.016132

Onnela, J. P., Saramaki, J., Kertesz, J., & Kaski, K. (2005). Intensity and coherence of motifs in weighted complex networks. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *71*, 65103. https://doi.org/10.1103/PhysRevE.71.065103

She, Q., Chen, G., & Chan, R. H. (2016). Evaluating the small-worldness of a sampled network: Functional connectivity of entorhinal-hippocampal circuitry. *Scientific Reports*, *6*, 21468. https://doi.org/10.1038/srep21468

Sporns, O., & Zwi, J. D. (2004). The small world of the cerebral cortex. *Neuroinformatics*, *2*, 145–162. https://doi.org/10.1385/NI:2:2:145

Stumpf, M. P., Wiuf, C., & May, R. M. (2005). Subnets of scale-free networks are not scale-free: Sampling properties of networks. *Proceedings of the National Academy of Sciences*, *102*, 4221–4224. https://doi.org/10.1073/pnas.0501179102

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*, 440–442. https://doi.org/10.1038/30918

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Bonzanni M, Bockley KM, Kaplan DL. On the effect of neuronal spatial subsampling in small-world networks. *Eur J Neurosci*. 2020;00:1–14. https://doi.org/10.1111/ejn.14937