Improving Network Inference: The Impact of False Positive and False Negative Conclusions about the Presence or Absence of Links

Gloria Cecchini[a,b,*], Marco Thiel[a], Björn Schelter[a], Linda Sommerlade[a]

[a] *Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Meston Building, Meston Walk, Aberdeen, AB24 3UE, United Kingdom*
[b] *Institute of Physics and Astronomy, University of Potsdam, Campus Golm, Karl-Liebknecht-Straße 24/25, 14476, Potsdam-Golm, Germany*

Abstract

Background

A reliable inference of networks from data is of key interest in the Neurosciences. Several methods have been suggested in the literature to reliably determine links in a network. To decide about the presence of links, these techniques rely on statistical inference, typically controlling the number of false positives, paying little attention to false negatives.

New Method

In this paper, by means of a comprehensive simulation study, we analyse the influence of false positive and false negative conclusions about the presence or absence of links in a network on the network topology. We show that different values to balance false positive and false negative conclusions about links should be used in order to reliably estimate network characteristics. We propose to run careful simulation studies prior to making potentially erroneous conclusion about the network topology.

Results

Our analysis shows that optimal values to balance false positive and false negative conclusions about links depend on the network topology and characteristic of interest.

*Corresponding author
Address: room 343, Meston Building, Meston Walk, Aberdeen, AB24 3UE, United Kingdom
Telephone Number: +441224272317
Email addresses: gloria.cecchini@abdn.ac.uk (Gloria Cecchini), *m.thiel@abdn.ac.uk* (Marco Thiel), *b.schelter@abdn.ac.uk* (Björn Schelter), *l.sommerlade@abdn.ac.uk* (Linda Sommerlade)

Comparison with Existing Methods

Existing methods rely on a choice of the rate for false positive conclusions. They aim to be sure about individual links rather than the entire network. The rate of false negative conclusions is typically not investigated.

Conclusions

Our investigation shows that the balance of false positive and false negative conclusions about links in a network has to be tuned for any network topology that is to be estimated. Moreover, within the same network topology, the results are qualitatively the same for each network characteristic, but the actual values leading to reliable estimates of the characteristics are different.

Keywords:
network inference, node degree distribution, false positive, false negative, statistical inference

---

1. Introduction

Recently, many research groups have focused on the inference of networks from data such as brain networks from observed electroencephalography or functional magnetic resonance imaging data (Bullmore and Sporns, 2009; Pessoa, 2014; Petersen and Sporns, 2015; Sporns et al., 2004). Particular emphasis is paid to the understanding of the normal functioning, e.g. healthy brain, as well as malfunctioning, e.g. diseased brain, of these networks. In the example of the brain, this promises to disclose information about how the brain processes signals and how alterations thereof cause specific diseases. A key hypothesis is that important characteristics are not specific to individual subjects but rather common in a given population. This is reflected by the fact that brain networks, but also other networks, are typically classified into few main prototypic networks (Newman, 2010, 2002), e.g., Erdős-Rényi (Erdős and Rényi, 1959, 1960), Watts-Strogatz (Watts, 1999; Watts and Strogatz, 1998), Barabási-Albert (Barabási and Albert, 1999; Barabási and Pósfai, 2016) networks. In our work we consider binary undirected networks of these three topologies.

These prototypical models for networks are in turn characterised by few parameters; procedures have been described to generate these networks with their well-established characteristics (Newman, 2010, 2002). Some of the key characteristics are the node degree distribution, the number of links, the global clustering coefficient, and the efficiency. We considered these characteristics in our study since they are meaningful in random networks and give a global description in large networks (Newman, 2010).

In the *Inverse Problem*, the challenge is to infer the network topology from data. Two challenges are particularly relevant: (i) the reliable inference of links in the network once the nodes have been fixed (Mader et al., 2015; Zerenner et al., 2014) and (ii) the successful usage of the characteristics above to uniquely determine the topology of network (Bialonski et al., 2010, 2011).

The correct reconstruction of networks is hampered not only by false conclusions about links due to statistical uncertainties, but also by unobserved processes (Elsegai et al., 2015; Guo et al., 2008; Ramb et al., 2013) and noise contamination (Nalatore et al., 2007; Newbold, 1978; Sommerlade et al., 2015) to name just a few challenges of network reconstruction. Classical statistical methods to estimate links in a network aim to identify present links with high certainty, e.g. (Jalili and Knyazeva, 2011; Quinn and Keough, 2002; Devore, 2011; Schinkel et al., 2011; De Vico Fallani et al., 2014; Chavez et al., 2010; Honey et al., 2007). Therefore, typically the rate of false positive conclusions about links is chosen and consequences for the rate of false negative conclusions about links are accepted. We investigate if these common rules of false positive conclusions and false negative conclusions should be modified to achieve a more reliable inference of the correct topology of network. To this aim, we analyse their influence on the network topology and characteristic.

The manuscript is structured as follows. An introduction to network topologies and their characteristics is given in Section 2.1. Section 2.2 explains statistical errors and their influence on the network topology. A simulation study in the case of Erdős-Rényi, Watts-Strogatz and Barabási-Albert networks is presented in Section 3.

2. Materials and Methods

In this section, network topologies and their characteristics are described (Section 2.1). We summarise statistical errors and suggest a distance measure to quantify their influence on the estimation of network characteristics (Section 2.2).

2.1. Network Characteristics

A network G is defined as a set of nodes with links between them. To quantify the topology of networks, different network characteristics have been described (Olbrich et al., 2010). Here, we consider four network characteristics: node degree, number of links, global clustering coefficient and efficiency. The node degree describes the number of links of a node. For example, if the node v has k links attached, its node degree is $d_v = k$. Typically the node degree distribution is used to characterise the entire network. The number of links refers to half of the sum over the node degrees.

The global clustering coefficient describes how well the neighbours of a node are connected. More precisely it measures the conditional probability that given one node connected to other two nodes, these are also connected to each other (Olbrich et al., 2010).

For two randomly selected nodes i, j in a network of n nodes, the shortest path length ℓ_{ij} measures the number of steps separating them if the shortest path is taken. The average path length $\gamma = \frac{1}{n(n-1)} \sum_{i \neq j} \ell_{ij}$ gives a measure of the sparsity of the network. The efficiency $\epsilon = \frac{1}{n(n-1)} \sum_{i \neq j} \frac{1}{\ell_{ij}}$ is defined as the sum of the inverse of the shortest paths lengths. Since the shortest

path is infinitely long for unconnected nodes, taking the average of the shortest path length in a network with unconnected nodes is not meaningful. Efficiency for unconnected nodes will be zero, therefore a meaningful network average of efficiency can be obtained.

Different network topologies have been described (Newman, 2003). Here, we investigate Erdős-Rényi (Erdős and Rényi, 1959, 1960), Watts-Strogatz (Watts, 1999; Watts and Strogatz, 1998) and Barabási-Albert (Barabási and Albert, 1999) networks, as key examples of networks.

Erdős-Rényi networks are random networks in which each pair of nodes is connected with independent probability p_c. The probability mass function of the node degree distribution of a Erdős-Rényi network

$$\mathbb{P}(d_v = k) = \binom{n-1}{k} p_c^k (1-p_c)^{n-1-k} \tag{1}$$

is a binomial distribution (Newman, 2010).

Watts-Strogatz networks are also referred to as small-world networks. They are characterised by a high local connectivity with some long-range "short-cuts". Watts-Strogatz networks are built from a regular network with node degree $2c$. Nodes are arranged on a circle; therefore, each node has c nearest clockwise as well as c nearest counterclockwise neighbours. With probability p_r each link connecting a node to one of its nearest neighbours is reconnected to another node randomly chosen. The node degree distribution has probability mass function

$$\mathbb{P}(d_v = k) = \sum_{i=\max(2c-k,0)}^{\min(n-1-k,2c)} \binom{2c}{i} \left(\frac{p_r}{2}\right)^i \left(1-\frac{p_r}{2}\right)^{2c-i} e^{-cp_r} \frac{(cp_r)^{k-2c+i}}{(k-2c+i)!}, \tag{2}$$

in the assumption that $n \gg c$, (Menezes et al., 2017). Here, we study Watts-Strogatz networks with $c = 2$.

Barabási-Albert networks are so-called scale free networks. They are constructed by adding nodes to an existing network. The degree of the existing nodes influences the probability for a new link. Each new node is connected to the network with a certain number b of links. The probability for one of these b links to be formed with any existing node is proportional the degree of that node. The node degree distribution has probability mass function (Barabási and Pósfai, 2016)

$$\mathbb{P}(d_v = k) = \frac{2b(b+1)}{k(k+1)(k+2)}. \tag{3}$$

Note that Eq. (3) is of type $\mathbb{P}(d_v = k) = c_1 k^{-c_2}$, where c_1 and c_2 are constants, i.e., it follows a power law.

2.2. Inference Reliability

In the Neurosciences different methods to identify nodes of a brain network exist. For our purposes the method of identifying nodes is not relevant. Therefore, we assume a fixed set of nodes. Once nodes have been fixed, several methods have been suggested in the literature to address the challenge of

reliable inference of links in the network. To determine the presence of links, these techniques usually rely on statistical inference, (Jalili and Knyazeva, 2011; Quinn and Keough, 2002; Devore, 2011; Schinkel et al., 2011; De Vico Fallani et al., 2014; Chavez et al., 2010; Honey et al., 2007).

Two types of errors exist when making these statistical inferences: (i) an absent link ($\overline{C}$) may be erroneously assumed to be present by the method (C^D), this is a false positive conclusion and referred to as a *type I error*; (ii) a present link (C) may remain undetected ($\overline{C^D}$) by the method, this is a false negative conclusion and referred to as a *type II error*. We call $\alpha = \mathbb{P}(C^D|\overline{C})$ the probability of a false positive conclusion and $\beta = \mathbb{P}(\overline{C^D}|C)$ the probability of a false negative conclusion. These two probabilities (α and β) are related and cannot be fixed independently. A standard choice is to set $\alpha = 0.05$ and neglect investigation of β, focussing on reliably detecting individual links of the network.

Let G denote the true network. As a consequence of the choice of α and thereby β, leading to a non-zero probability of detecting false positive and false negative links, the network we detect G^D will be a "mixture" of true links, false positive links, absent links and false negative links. Therefore, the number of detected links is generally different to the number of links of G. Also the node degree distribution, the global clustering coefficient and the efficiency are in general biased. We quantify the bias for each characteristic using a distance between distributions. Several distance measures are conceivable and have been investigated; for sake of simplicity and to make the arguments clearer, we only consider the distance

$$\delta = |\mu_1 - \mu_2| \tag{4}$$

between two distributions, as the modulus of the difference of the distribution's mean values. For example, the distance between the node degree distribution of G, which has mean μ_G, and the node degree distribution of G^D, which has mean μ_{G^D}, is $\delta = |\mu_G - \mu_{G^D}|$.

To investigate the relation between α and β numerically we simulate $N = 100$ data points taken from a bivariate normal distribution. To inspect in particular links with medium strength, we vary the simulated correlation between 0.36 and 0.46 in steps of 0.02. For each value, we repeated the simulation 10,000 times and tested for correlation using Pearson's correlation test. From this test we inferred the probability that the 100 simulated data points are not correlated, the so-called $p-$value of the test. Based on the 10,000 $p-$values of the correlation tests we inferred the relation between $1/\beta$ and α (see Fig. 1). Visual inspection of Fig. 1 shows that a linear relationship is a good approximation. Fitting linear functions to the curves shows that their respective slopes vary between $0.1 \cdot 10^{-3}$ and $1.1 \cdot 10^{-3}$. These slopes will differ if different simulation parameters, such as the number of data points N, are chosen. The more data points are considered the more accurate the analysis. Note that the inverse proportionality of α and β implies that an infinite number of data points N is needed to have both α and β equal to zero.

As an example of how the choice of α and consequently β affects the estimated network characteristics, we consider Erdős-Renyi networks G_{p_c}. The

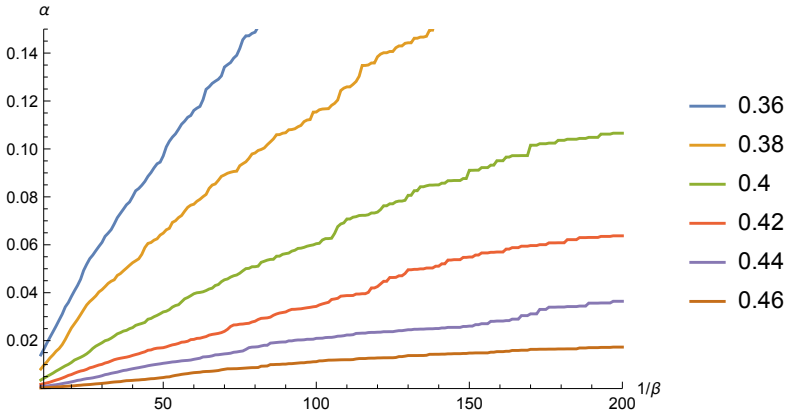


Figure 1: Relation between $1/\beta$ and α for correlation of 100 simulated data points. Colours refer to different correlation coefficients used for the simulation, as indicated in the legend. The slopes of linear functions fitted to these curves vary between $0.1 \cdot 10^{-3}$ and $1.1 \cdot 10^{-3}$.

parameter p_c was varied between 0.01 and 0.99 in steps of 0.01. The detected networks $G^D_{p_c}$ were generated by artificially introducing false positive links with probability α and false negative links with probability β. We varied α between 0.005 and 0.1 in steps of 0.001, the relation between α and β was fixed by

$$\beta = \frac{10^{-3}}{\alpha}, \tag{5}$$

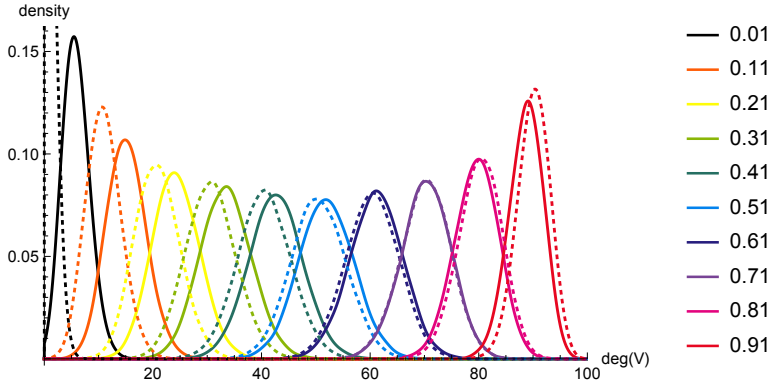
which represents a choice motivated by our simulations (Fig. 1). Moreover, this choice corresponds to a method, which has high sensitivity and specificity, i.e., $0.005 < \alpha, \beta < 0.2$. For each value of p_c and α, 200 networks with $n = 100$ nodes were generated. Figure 2 shows the true densities of the node degree derived from G_{p_c} (dashed lines) together with the average densities derived from the detected networks $G^D_{p_c}$ (solid lines). Results for two different values of α are shown. Different colours represent different Erdős-Renyi networks defined by the parameter p_c, for clarity, densities are plotted for p_c in steps of 0.1 only.

The distances (Eq. 4) between the true density and the detected density for each pair of p_c and α are shown in Fig. 3. For some values of p_c the distance is negligible, which means the detected node degree is almost identical to the true node degree. The optimal α, i.e. the one with the smallest distance between true density and detected density, depends on p_c.

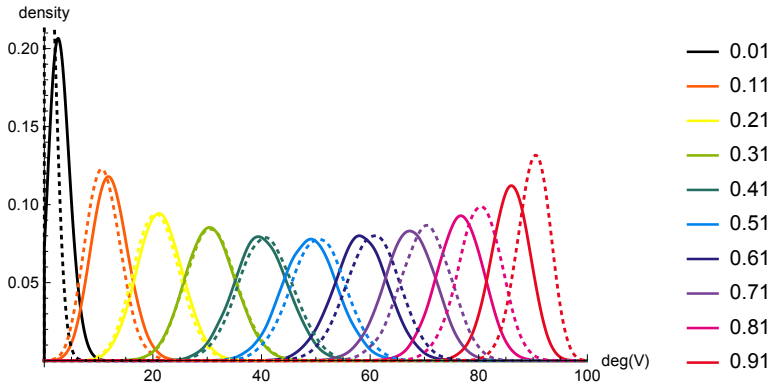
To have a general result for the optimal choice of α when estimating a network characteristic of a given network topology, we sum over p_c to marginalise out the influence of p_c for each α. We call this integrated quantity the total distance δ_{tot}, i.e.

$$\delta_{tot} = \sum_{p_c} \delta(p_c). \tag{6}$$

To identify the optimal choice of α, we are interested in finding where the minimum of the total distance δ_{tot} is located. Figure 4 shows δ_{tot} for the example



(a) Solid lines: $G^D_{p_c}$ with $\alpha = 0.05, \beta = 0.02$. Dotted lines: G_{p_c}



(b) Solid lines: $G^D_{p_c}$ with $\alpha = 0.02, \beta = 0.05$. Dotted lines: G_{p_c}

Figure 2: Densities of the node degree distributions for Erdős-Rényi networks of $n = 100$ nodes and different parameters $p_c = 0.01, ..., 0.91$ in steps of 0.1 represented by colour. The densities of the node degree distributions for the respective original networks G_{p_c} (dotted lines) and detected networks $G^D_{p_c}$ (solid lines) are shown.

of the node degree of Erdős-Rényi networks. In this example, the minimum of δ_{tot} is located at $\alpha = 0.030$. This suggests that in order to optimally reconstruct the node degree of an Erdős-Rényi network $\alpha = 0.03$ should be chosen, which is close to the standard choice of $\alpha = 0.05$ but distinctively smaller.

3. Results

We applied our analysis to Erdős-Rényi, Watts-Strogatz and Barabási-Albert networks. For each network topology, we investigated four different network characteristics: node degree, number of links, global clustering coefficient and efficiency. The distance δ, which depends on both α and the parameter of the network topology (p_c, b or p_r) is presented as density plot for all the investigated characteristics and network topologies in Fig. 5. All 12 investigated scenarios

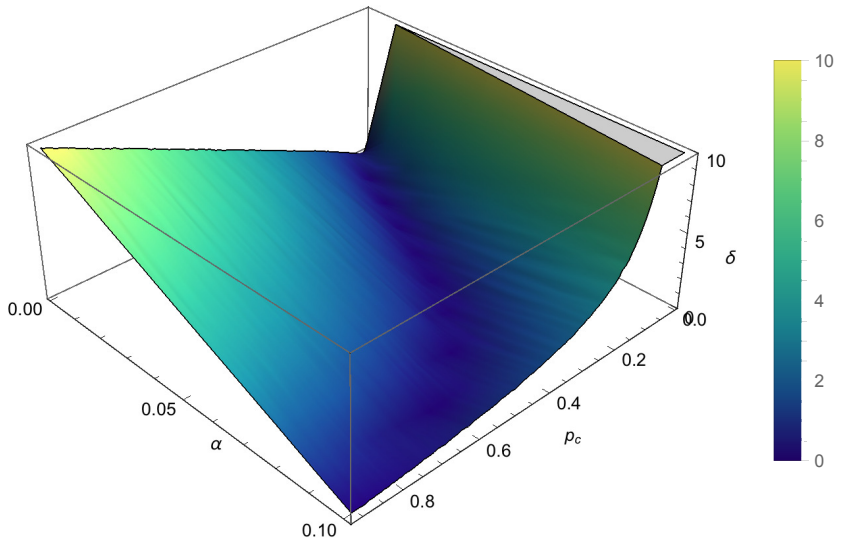


Figure 3: Distance δ between node degree distributions of Erdős-Rényi networks with 100 nodes depending on α and p_c. Distance δ is measured by calculating the difference between the mean of two corresponding distributions, Eq. (4). Colour code expresses distance values.

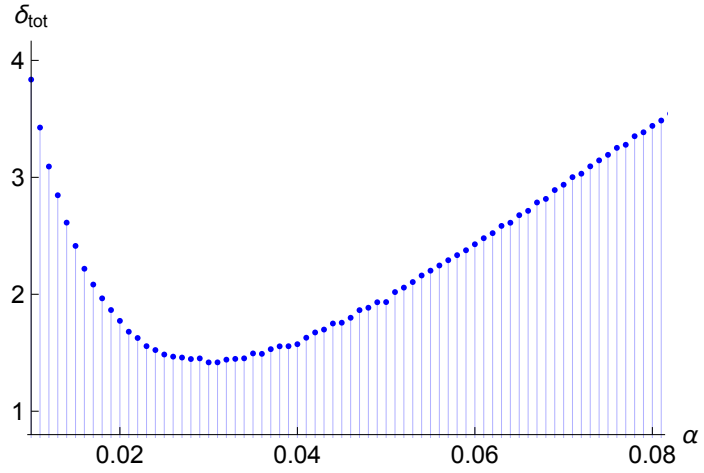


Figure 4: Total distances δ_{tot} between node degree distributions of Erdős-Rényi networks depending on α. The minimum is located at $\alpha = 0.030$.

show a dependence of the distance on the choice of α, suggesting that an optimum exists. For some scenarios, in particular node degree and number of links for Watts-Strogatz and Barabási-Albert networks, dependence of the distance on the parameter (p_r or b) is negligible. For other scenarios such as the global clustering coefficient in Watts-Strogatz networks the question arises if marginalising out the influence of p_r is distorting the results. Detailed results for each

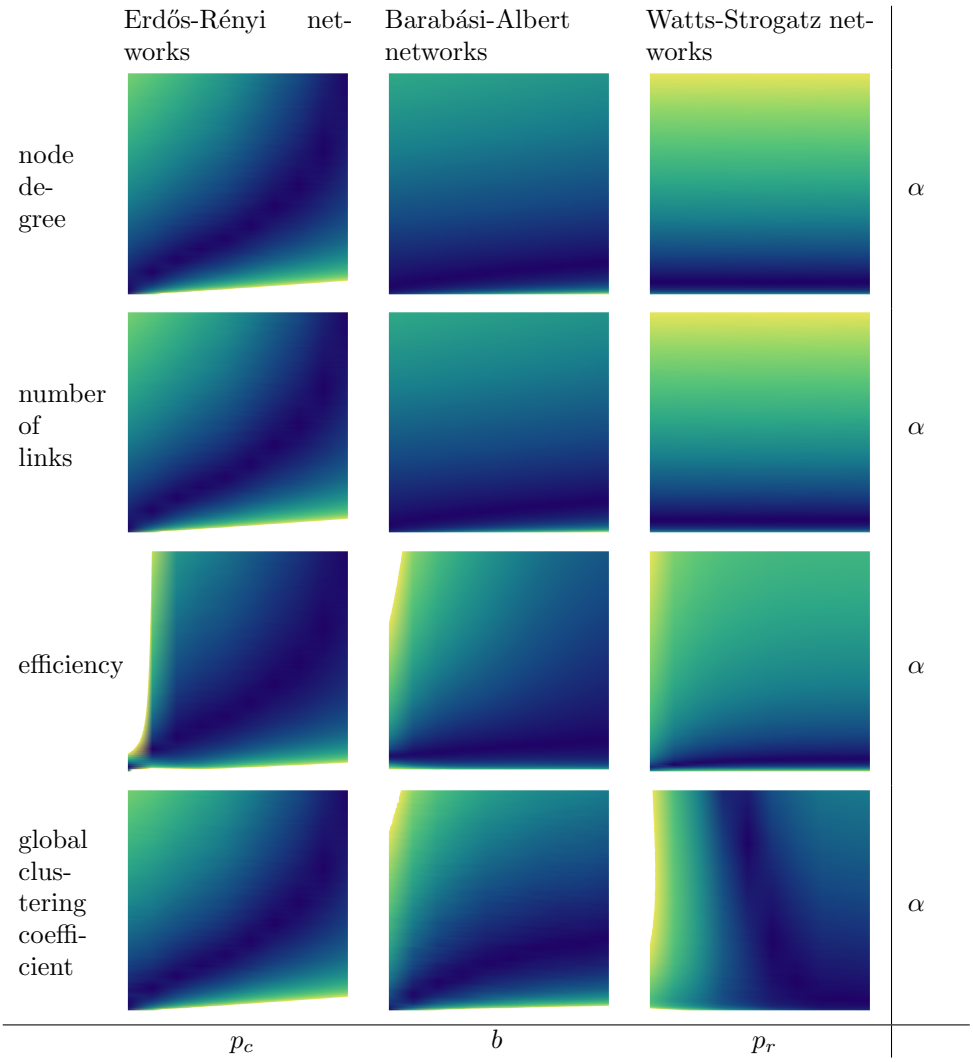


Figure 5: Distance δ for Erdős-Rényi, Barabási-Albert, and Watts-Strogatz networks with 100 nodes. Distance δ is measured by calculating the difference between the mean of two corresponding distributions, Eq. (4), and it is expressed by colour code, from blue to yellow. For each network topology we vary the control parameter p_c from 0.01 to 0.99 in steps of 0.01, b from 1 to 10 in steps of 1, or p_r from 0.01 to 0.99 in steps of 0.01 on the $x-$axis, and the probability of false positive α from 0.005 to 0.1 in steps of 0.001 on the $y-$axis.

network topology are presented below.

For Erdős-Rényi networks of $n = 50$, $n = 100$, and $n = 250$ nodes we varied p_c from 0.01 to 0.99 in steps of 0.01. Figure 2 shows an example of some of these values in steps of 0.1. The results of the total distance δ_{tot} for the node degree of Erdős-Rényi networks are shown in Fig. 4. The minimum of δ_{tot} is located at

$\alpha = 0.030$ ($\beta = 0.033$). For the remaining network characteristics, the Erdős-Rényi networks also show a clear minimum of the total distance in dependence on α. The specific values of α for the respective minimal total distances however vary; they are summarised in Table 1. The optimal α for efficiency is noticeably smaller than for the other network characteristics. Moreover, we chose a broad range for p_c to cover the broad spectrum of Erdős-Rényi networks. Marginalising out the dependence of the distance δ on p_c may therefore be distorting our results (see also dependence on p_c in Fig. 5). For a specific application we thus recommend narrowing the range of p_c to values relevant for the application.

Network Topology	$n = 50$		$n = 100$		$n = 250$	
and Characteristic	α	β	α	β	α	β
Erdős-Rényi:						
node degree	0.031	0.032	0.030	0.033	0.031	0.032
number of links	0.031	0.032	0.030	0.033	0.031	0.032
global clustering coeff	0.035	0.029	0.031	0.032	0.031	0.032
efficiency	0.016	0.063	0.012	0.083	0.020	0.050
Barabási-Albert:						
node degree	0.018	0.056	0.012	0.083	0.007	0.143
number of links	0.018	0.056	0.012	0.083	0.007	0.143
global clustering coeff	0.024	0.042	0.021	0.048	0.019	0.053
efficiency	0.015	0.067	0.010	0.100	0.007	0.143
Watts-Strogatz:						
node degree	0.009	0.111	0.007	0.143	0.004	0.250
number of links	0.009	0.111	0.007	0.143	0.004	0.250
global clustering coeff	0.008	0.125	0.006	0.167	0.004	0.250
efficiency	0.009	0.111	0.006	0.167	0.004	0.250

Table 1: Table of α and β values for minimal total distances δ_{tot} of each network topology and characteristic.

The set of Barabási-Albert networks of $n = 50$, $n = 100$, and $n = 250$ nodes was chosen with average degree typical for networks in neuroscience (Papo et al., 2014; Stanley et al., 2013) by using parameters $b = 1$ to $b = 10$ varying in steps of 1. The total distances δ_{tot} for the node degree of networks with $n = 100$ nodes are shown in Fig. 6. The minimum is found for $\alpha = 0.012$ ($\beta = 0.083$), it is more pronounced than that for the Erdős-Rényi networks. Again, the other network characteristics and number of nodes all show a single minimum. The values for optimal α and β are summarised in Table 1. For this network topology a noticeably different optimal value for α was found for the clustering coefficient.

Finally, we considered a set of Watts-Strogatz networks of $n = 50$, $n = 100$, and $n = 250$ nodes with parameter p_r varied between 0.01 and 0.99 in steps of 0.01. We analysed distances between the distributions of the node degree, the number of links, the global clustering coefficient and the efficiency. The minimum of the total distance for the node degree of networks with $n = 100$ nodes is found for $\alpha = 0.007$ ($\beta = 0.143$). The total distances between

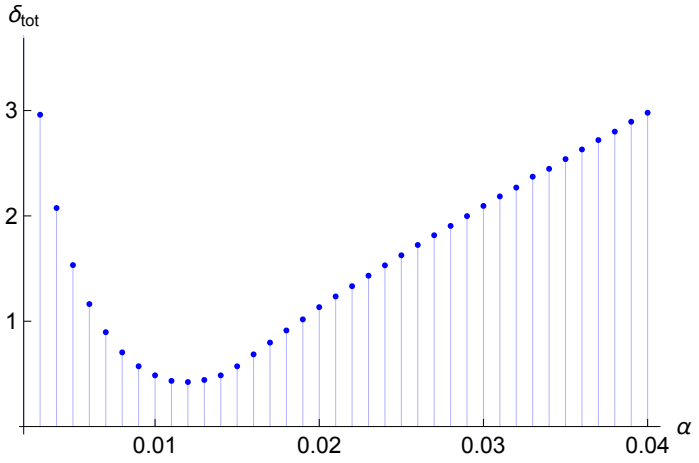


Figure 6: Total distance δ_{tot} between node degree distributions of Barabási-Albert networks of $n = 100$ nodes depending on α. The minimum is located at $\alpha = 0.012$ ($\beta = 0.083$).

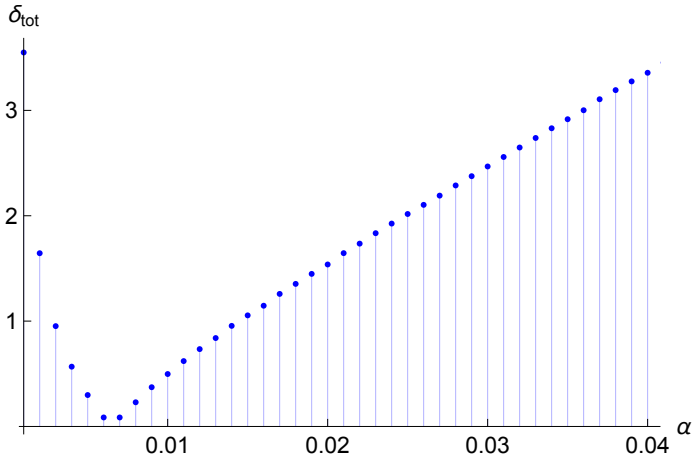


Figure 7: Total distance δ_{tot} between node degree distributions for Watts-Strogatz networks of $n = 100$ nodes depending on α. The minimum is located at $\alpha = 0.007$ ($\beta = 0.143$).

node degree distributions for these networks are shown in Fig. 7. For all four characteristics and different values of n clear minima can be identified and the values for optimal α are similar (Table 1). The results for the efficiency however have to be interpreted with case as the distance showed a clear dependence on the parameter p_r (see Fig. 5).

4. Discussion

We consider three topologies of networks Erdős-Rényi, Watts-Strogatz, and Barabási-Albert. For each topology, and for a specific characteristic, e.g. the

node degree distribution, the number of links, the efficiency or global clustering coefficient, the rate of false positive and false negative conclusions about links can be optimally chosen in order to have less biased reconstruction.

For Erdős-Rényi networks, the values for α identified with our method are close to standard choice of α of 0.05. Standard alpha values are suboptimal when the topology of network is different. For the set of Barabási-Albert networks we found that the value for $\alpha = 0.012$ and consequently $\beta = 0.083$ yields the most reliable results of the node degree. In this case, standard alpha values lead to a bigger distance between distributions of the node degree. The Watts-Strogatz networks yield the most reliable results for an even smaller value for $\alpha = 0.007$ and consequently $\beta = 0.154$. Moreover, for the optimal choice of α the corresponding β is rather high. This shows that the reliability of detecting individual false negative links in a network is less important than failing to recognise false positive links when network characteristics are estimated. Accepting a high rate of false negative links may thus be required when the aim is to infer a specific network characteristic.

This work shows that the standard choice of α of 0.05 is not optimal when the aim is to reconstruct the entire network topology. Moreover, α needs to be adjusted depending on specific network topologies and characteristics. For example, consider Erdős-Rényi networks with $p_c = 0.11$ and assume the relationship between α and β is Eq. 5. As result of 200 simulations, the mean of the node degree distribution of the original network G_{p_c} is 11 and the mean for estimation using $\alpha = 0.05$ is 15. Choosing $\alpha = 0.03$ results in a mean of the node degree distribution of 13. The choice of $\alpha = 0.03$ is motivated by the assumption that the original network is known to be an Erdős-Rényi network with unknown parameter p_c, see Table 1. For the same study, when the aim is to infer the efficiency ϵ, we calculate $\epsilon = 0.51$ for the true network, $\epsilon = 0.56$ for the one with $\alpha = 0.05$, and $\epsilon = 0.51$ when $\alpha = 0.012$. The more we know about the network we want to infer the more accurate the reconstruction is since the simulation study can be tuned accordingly.

As mentioned in Section 2.2, the relationship between α and β depends on the number of data points N, therefore the values of α and β leading to the minimal distance will change for different values of N. Nevertheless, the results will remain qualitatively the same.

The size of the network, i.e. the number of nodes, also influences the result. The number of false positive and false negative conclusions about the presence of links depends on the number of total links in the network. Keeping the same rate of α and β and increasing, for example, the size of the network, leads to larger number of false positive and false negative detections of links. As shown in Table 1, for Erdős-Rényi networks the values of α and β leading to the minimal distance almost do not change. The reason is that the number of links increases proportionally with the number of nodes for each p_c. This does not happen for Barabási-Albert and Watts-Strogatz networks; the values of α leading to the minimal distance present a decreasing trend because of their constructions.

We considered the node degree distribution, the number of links, the efficiency, and the global clustering coefficient as example characteristics to show

that the results depend on the characteristic under investigation. Nevertheless, our approach can be readily applied to other characteristics, as well as other network topologies.

5. Conclusion

False conclusions about the presence of links in a network typically alter network characteristics, such as the node degree distribution, the number of links, the global clustering coefficient and the efficiency. Identification of the underlying network topology relies on these characteristics and is thus hindered by false conclusions about links as well. For these reasons, the analysis of false positive and false negative conclusions about links is of key importance.

In this manuscript, assuming to know the underlying network topology, we investigate the influence of false positive and false negative conclusions about links in a network. We show that the values of α and β leading to minimal distance (difference in mean values) between the true network and the biased one change depending not only on the network topology, but also on the network characteristic of interest. Therefore, in the *Inverse Problem*, when the challenge is to infer the network topology from data, different values for α and β might be favourable when estimating different characteristics. We speculate that our simulation study can be used as an iterative procedure to achieve a better network reconstruction. Namely, when the network topology is not known a priori, various values for α can be chosen to perform the first iteration step of the network reconstruction. The result of this first step gives an idea of the network topology we want to infer. For the second iteration step the value for α can be adjusted according to the findings of the first step. This procedure can be iterated using the simulation study that we suggest in this paper in each iteration step, ultimately leading to a reconstruction of the network tailored to its previously unknown network topology.

This result suggests that in the Neurosciences, as well as in other scientific fields, various values for statistical inference could be considered within a simulation study to determine the optimal α for the network characteristic of interest. If several network characteristics are of interest, it may be useful to adjust the value of α for each characteristic.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 642563.

The authors declare no competing financial interests.

References

Barabási, A.-L., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512.

Barabási, A.-L., Pósfai, M., 2016. Network science. Cambridge University Press, Cambridge.

Bialonski, S., Horstmann, M.-T., Lehnertz, K., 2010. From brain to earth and climate systems: Small-world interaction networks or not? Chaos 20, 013134.

Bialonski, S., Wendler, M., Lehnertz, K., 2011. Unraveling spurious properties of interaction networks with tailored random networks. PLoS ONE 6, 1–13.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198.

Chavez, M., Valencia, M., Latora, V., Martinerie, J., 2010. Complex networks: new trends for the analysis of brain connectivity. Int. J. Bifurcation Chaos 20 (06), 1677–1686.

De Vico Fallani, F., Richiardi, J., Chavez, M., Achard, S., 2014. Graph analysis of functional brain networks: practical issues in translational neuroscience. Philos. T. R. Soc. B 369 (1653).

Devore, J., 2011. Probability and Statistics for Engineering and the Sciences. Cengage Learning.

Elsegai, H., Shiells, H., Thiel, M., Schelter, B., 2015. Network inference in the presence of latent confounders: The role of instantaneous causalities. J. Neurosci. Meth. 245, 91–106.

Erdős, P., Rényi, A., 1959. On random graphs. Publ. Math-Debrecen 6, 290–297.

Erdős, P., Rényi, A., 1960. On the evolution of random graphs. In: Publication of the Mathematical Institute of the Hungarian Academy of Sciences. pp. 17–61.

Guo, S., Seth, A. K., Kendrick, K. M., Zhou, C., Feng, J., 2008. Partial Granger causality-Eliminating exogenous inputs and latent variables. J. Neurosci. Meth. 172, 79–93.

Honey, C. J., Kötter, R., Breakspear, M., Sporns, O., 2007. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. Proc. Natl. Acad. Sci. 104 (24), 10240–10245.

Jalili, M., Knyazeva, M. G., 2011. Constructing brain functional networks from eeg: partial and unpartial correlations. J. Integr. Neurosci. 10 (02), 213–232.

Mader, W., Mader, M., Timmer, J., Thiel, M., Schelter, B., 2015. Networks: On the relation of bi- and multivariate measures. Sci. Rep. 5, 10805.

Menezes, M. B. C., Kim, S., Huang, R., 06 2017. Constructing a watts-strogatz network from a small-world network with symmetric degree distribution. PLoS One 12 (6), 1–9.

Nalatore, H., Ding, M., Rangarajan, G., 2007. Mitigating the effects of measurement noise on Granger causality. Phys. Rev. E 75, 1–10.

Newbold, P., 1978. Feedback induced by measurement errors. Int. Econ. Rev. 19, 787–791.

Newman, M., 2010. Networks: An Introduction. Oxford University Press, Inc., New York, NY, USA.

Newman, M. E. J., 2002. Random graphs as models of networks. In: Bornholdt, S., Schuster, H. G. (Eds.), Handbook of Graphs and Networks: From the Genome to the Internet. No. 1. pp. 34–68.

Newman, M. E. J., 2003. The structure and function of complex networks. SIAM Rev 45, 167–256.

Olbrich, E., Kahle, T., Bertschinger, N., Ay, N., Jost, J., 2010. Quantifying structure in networks. Eur. Phys. J. B 77, 239–247.

Papo, D., Buldu, J., Boccaletti, S., 2014. Network theory in neuroscience. Encyclopedia of Computational Neuroscience, 1852–1868.

Pessoa, L., 2014. Understanding brain networks and brain organization. Phys. Life Rev. 11, 400–435.

Petersen, S. E., Sporns, O., 2015. Brain Networks and Cognitive Architectures. Neuron 88, 207–219.

Quinn, G., Keough, M., 2002. Experimental Design and Data Analysis for Biologists. Cambridge University Press.

Ramb, R., Eichler, M., Ing, A., Thiel, M., Weiller, C., Grebogi, C., Schwarzbauer, C., Timmer, J., Schelter, B., 2013. The impact of latent confounders in directed network analysis in neuroscience. Philos. T. Roy. Soc. A 371, 20110612.

Schinkel, S., Zamora-López, G., Dimigen, O., Sommer, W., Kurths, J., 2011. Functional network analysis reveals differences in the semantic priming task. J. Neurosci. Meth. 197 (2), 333 – 339.

Sommerlade, L., Thiel, M., Mader, M., Mader, W., Timmer, J., Platt, B., Schelter, B., 2015. Assessing the strength of directed influences among neural signals: An approach to noisy data. J. Neurosci. Meth. 239, 47–64.

Sporns, O., Chialvo, D. R., Kaiser, M., Hilgetag, C. C., 2004. Organization, development and function of complex brain networks. Trends Cogn. Sci. 8, 418–425.

Stanley, M., Moussa, M., Paolini, B., Lyday, R., Burdette, J., Laurienti, P., 2013. Defining nodes in complex brain networks. Frontiers in Computational Neuroscience 7, 169.

Watts, D., 1999. Small Worlds: The Dynamics of Networks Between Order and Randomness. Princeton studies in complexity. Princeton University Press.

Watts, D. J., Strogatz, S. H., 1998. Collective dynamics of 'small-world' networks. Nature 393, 440–442.

Zerenner, T., Friederichs, P., Lehnertz, K., Hense, A., 2014. A gaussian graphical model approach to climate networks. Chaos 24, 023103.