Improving Network Inference of Oscillatory Systems: A Novel Framework To Reliably Identify the Correct Class Of Network

Gloria Cecchini, Björn Schelter, Marco Thiel and Linda Sommerlade

Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen, Scotland, UK

Abstract

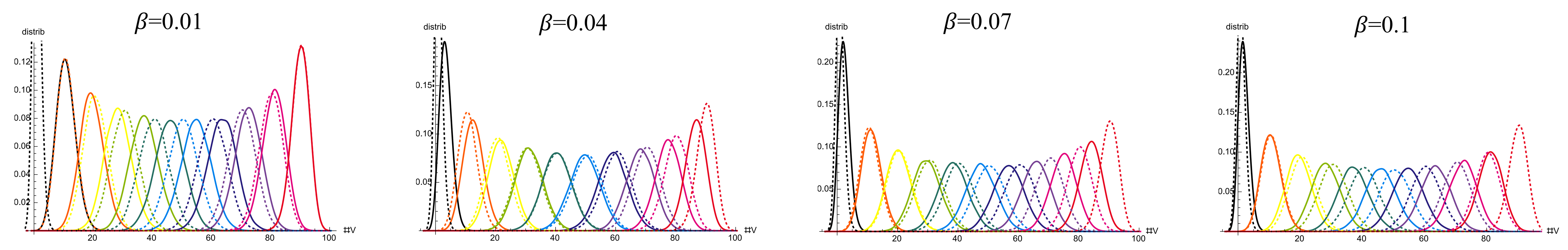
The reliable inference of network structure from data is of key interest in nowadays research. In our work we suggest a novel procedure to balance false positive conclusions about links in a network for a given analysis technique and the characteristics of the networks inferred. This enables us to obtain, e.g., better estimates for the average shortest path length or degree distribution for a given analysis technique. Our method suggests a procedure to infer the correct type of network with high probability.

Inference Reliability

We present a topological analysis of detected networks taking into account the role of two important parameters: false positive and false negative decisions about the presence and absence of links. Assume that α is the probability to detect a false positive connection and that β is the probability to detect a false negative connection. Statistical analysis revealed an inverse relationship between α and β, so it is not possible to have both α and β equal to zero. We find that classical statistical approaches may not always be the best choice. Therefore, we suggest a novel procedure how to optimally balance false positives connection to best estimate e.g. the degree distribution for a given analysis technique.

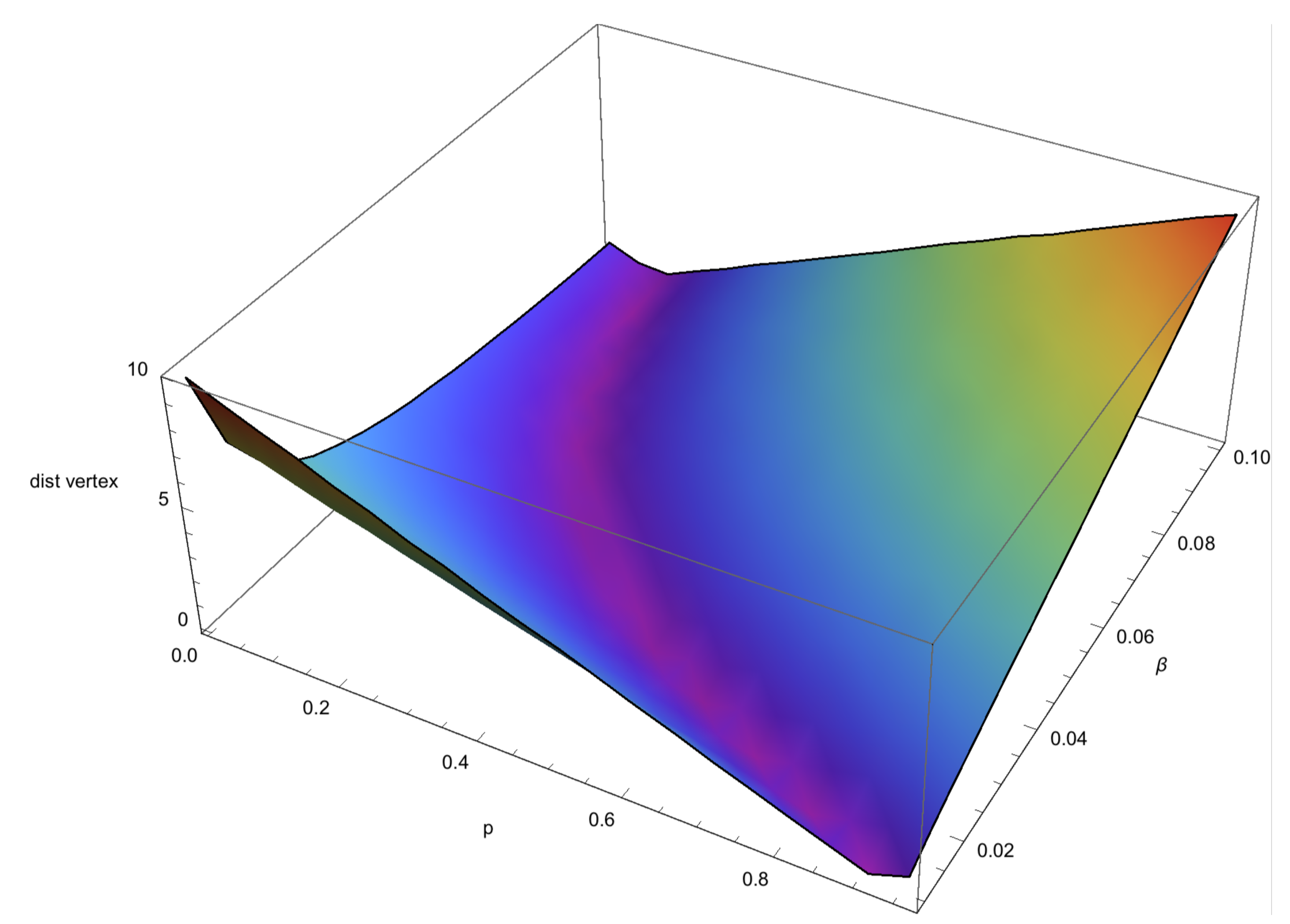
False Positive and False Negative Rate

Link	Present	Absent
Detected	$1-\alpha$	α
Not detected	β	$1-\beta$



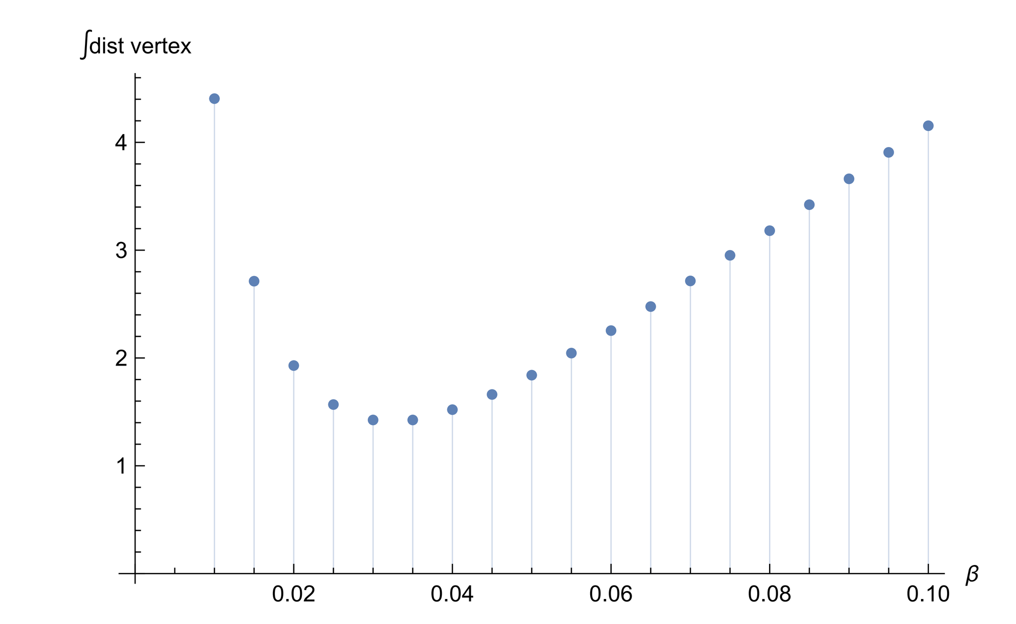
Vertex Degree Distributions of the original Erdos-Renyi networks (dotted lines) and the inferred ones (solid lines). Each color represents different p = 0.01, … ,0.9

For some values of p the distance between the distributions is negligible, but the optimal β to minimize this distance changes if p is modified. Marginalising out the influence of p, the integral with respect to p is taken for each β. We call this quantity the total distance.



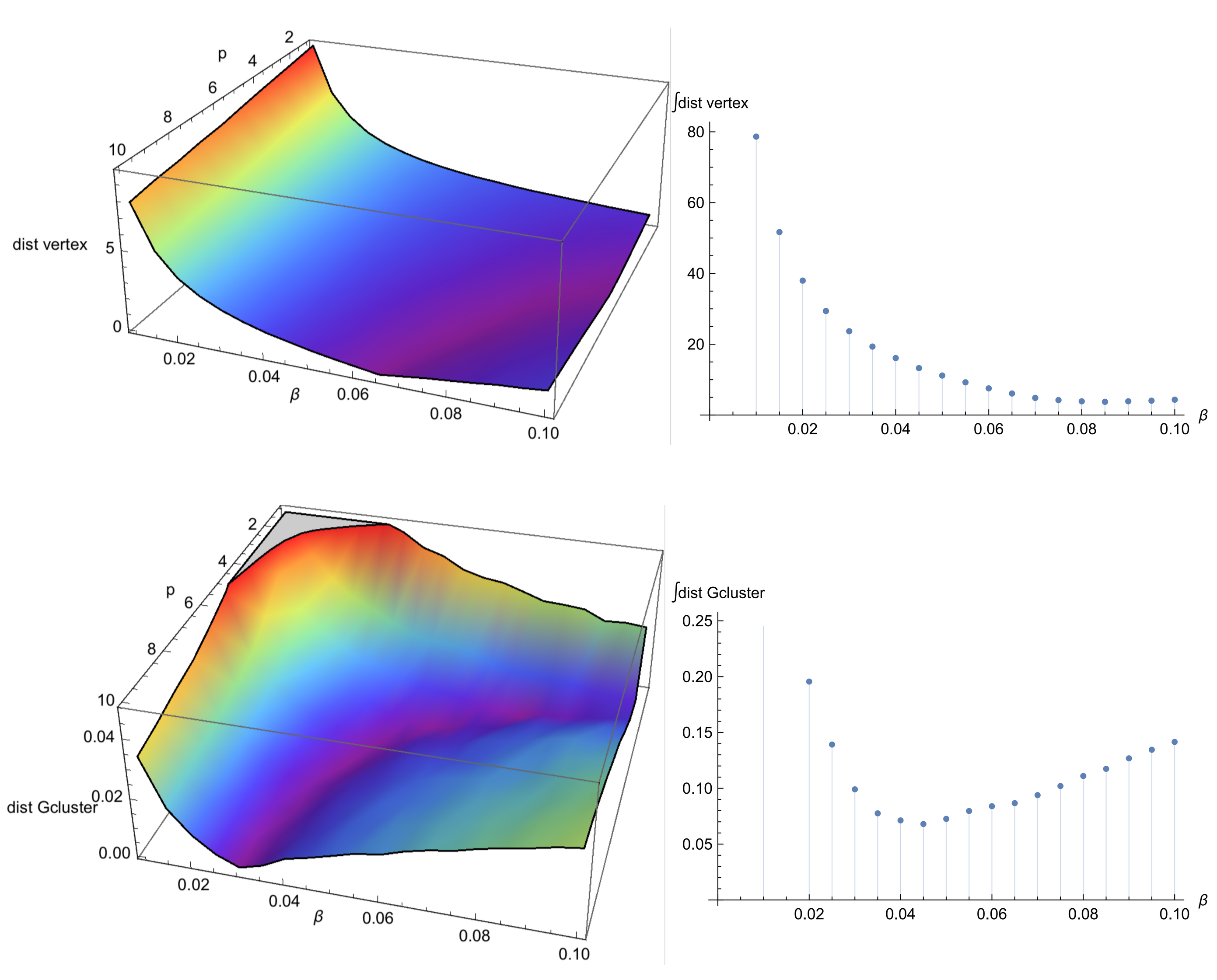
Distance between vertex degree distributions depending on β and p

$$\int_0^1 \left[\begin{array}{c}\text{distance}\\ \text{between}\\ \text{distrib.}\end{array}\right] dp$$



Total Distance between vertex degree distributions

Scale-Free Network



Network Reconstruction

Mis-estimated links in a network typically alter the key characteristics and therefore inferring the correct type of network becomes more complicated. Fixing α and having a strong assumption for β allow to reconstruct the type of the original network that was influenced by these parameters.

The number of edges and the vertex degree distribution are two important characteristics which can be used: compare the values collected from the detected network to the theoretical values typical of specific classes of networks, i.e., Erdos-Renyi, Small World or Scale Free Networks. We can use this procedure to assign certain probabilities to different kinds of network topologies. Also in the case explained in this section, we can suggest what is the best estimated network for a specific characteristic, e.g. vertex degree distribution, average shortest path length or global clustering coefficient.

Conclusions

In this work we analysed how α and β influence the network structure. Moreover, we developed a strategy to optimize these parameters to obtain the less biased result. We focused on the inverse problem, i.e. for a given analysis technique, infer the correct type of network with high probability.

Contact *Gloria Cecchini*, email address: gloria.cecchini@abdn.ac.uk

1495 UNIVERSITY OF ABERDEEN

COSMOS Complex Oscillatory Systems: Modeling and Analysis
Innovative Training Network
European Joint Doctorate

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

SUPA