Revealing Network Symmetries Using Time-Series Data

Ethan T.H.A. van Woerkom, Joseph D. Hart, Thomas E. Murphy and Rajarshi Roy

Abstract Complex dynamical networks may exhibit graph symmetries. These symmetries leave an imprint on network behaviour and statistics. This effect is first demonstrated in a small opto-electronic network. We then present the general conditions under which network statistics become invariant under the action of network symmetries. Statistical analyses can help reveal the symmetry group of a network graph without knowledge of the underlying network model. Finally, results from numerical experiments are also shown to align along network symmetries.

Ethan T.H.A. van Woerkom
University of Edinburgh, School of Physics and Astronomy, James Clerk Maxwell Building, Peter Guthrie Tait Road, Edinburgh, EH9 3FD , United Kingdom
e-mail: ethanvanwoerkom@yahoo.com

Joseph D. Hart
Institute for Research in Electronics and Applied Physics, University of Maryland, College Park, Maryland 20742, USA,
Department of Physics, University of Maryland, College Park, Maryland 20742, USA, e-mail: jhart12@umd.edu

Thomas E. Murphy
Institute for Research in Electronics and Applied Physics, University of Maryland, College Park, Maryland 20742, USA,
Department of Electrical and Computer Engineering, University of Maryland, College Park, Maryland 20742, USA, e-mail: tem@umd.edu

Rajarshi Roy
Institute for Research in Electronics and Applied Physics, University of Maryland, College Park, Maryland 20742, USA,
Department of Physics, University of Maryland, College Park, Maryland 20742, USA,
Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, USA, e-mail: rroy@umd.edu

1 Introduction

Network symmetries exist in many networks and give essential information about their structure. The difficult problem of reconstructing the graph of a complex network by studying its dynamics has been studied before [2, 3]. In this article we demonstrate a tool that instead of resolving individual network connections, reconstructs the symmetry group of a network using time-series statistics. This article details how network symmetries manifest themselves in network behaviour and time-series statistics and gives methods to infer network symmetries from statistics in the general case. In this article, unless otherwise stated, we study a dynamical system as in definition 1. In this article we study networks which contain symmetries. Definition 2 details the conditions for such a symmetry.

Definition 1 A dynamical network is defined by the dynamics of N nodes, each represented by state vectors, adhering to the following equation:

$$\dot{\boldsymbol{x}}_i = F_i(\boldsymbol{x}_1(t), ..., \boldsymbol{x}_N(t)) \quad \text{for} \quad i \in \{1, ..., N\}. \tag{1}$$

Here, $\boldsymbol{x}_i$ are the state vectors of the nodes in the dynamical system and $F_i(\boldsymbol{x}_1, ..., \boldsymbol{x}_N)$ represents the functions that give the derivative of each node as a function of all other nodes.

Definition 2 A dynamical system is defined to have symmetry g if $F_{g(i)}(\boldsymbol{x}_1, ..., \boldsymbol{x}_N) = F_i(\boldsymbol{x}_{g(1)}, ..., \boldsymbol{x}_{g(N)})$, where g is a permutation $g : \{1, ..., N\} \to \{1, ..., N\}$.

A consequence of the above definition of a symmetry of a dynamical system is that if there is a solution $\boldsymbol{s}_i(t)$, then $\boldsymbol{s}_{g(i)}(t)$ is also a solution.

This article is organised as follows. In section 2 we discuss an opto-electronic network experiment which demonstrates symmetries in time-averaged behaviour. In section 3 we present a theorem stating the conditions required for symmetries in network dynamics to appear in network statistics. We give a more elaborate example of the consequences of this theorem in section 4. In section 5 we conclude and discuss the possibility of using the presented methods to retrieve general network symmetries.

2 Statistical Symmetries in a Small Opto-Electronic Network

In this section we detail an experiment with a small opto-electronic network that demonstrates how network graph symmetries affect network behaviour and present themselves in network statistics.

2.1 Experimental Setup

A network of four coupled opto-electronic, time-delayed feedback systems is used [1]. Each of these systems, 'nodes', depicted in Fig.1, consists of a laser diode which passes a light signal through an integrated Mach-Zehnder modulator, altering the intensity of the signal with a $\cos^2(x + \phi_0)$ nonlinearity, where x is the normalised input voltage to the modulator. This signal is then passed on to other nodes through optical fibres and returned for self-feedback.

The two input signals, being the self-feedback signal, and the inputs from the other nodes respectively, are measured in separate photoreceivers. A Digital Signal Processing board is then used to apply a feedback and coupling delay, apply a digital filter and amplify the signal. The signal is then fed back into the Mach-Zehnder modulator. In this way, a coupled nonlinear chaotic oscillator with time delays is produced incorporating both coupling and self-feedback time delays. A two-pole digital Butterworth filter is used to filter the signal, with a high-pass frequency of $\omega_H/2\pi = 100Hz$ and a low-pass frequency of $\omega_L/2\pi = 2.5kHz$, operating at a sampling rate of 24 kSamples/s. Connections between nodes can be controlled by variable fibre-based attenuators.

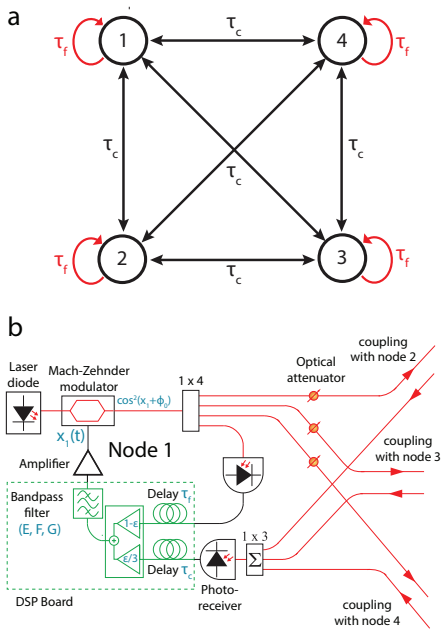


Fig. 1: (a)-Possible connections in network, with self-feedback time delays τ_f and coupling time delays τ_c included. (b)-Schematic of an opto-electronic node with connections to neighbours. Red connections are through optical fibres, whereas black connections are electronic. Figure from [1].

Equations 2-5 well describe the dynamics of the network. In reality, due to the digital sampling of the DSP board, the system dynamics is partitioned into discrete time steps. In these equations, $\boldsymbol{u}_i$ represents the state of the digital filter corresponding to each node, ε the coupling strength (ranging from 0-1), and β the round-trip gain. Furthermore, τ_f is the self-feedback time delay of each node, whereas τ_c is the coupling time delay, which controls the input delay from other nodes. ϕ_0 represents the DC offset of the Mach-Zehnder modulator, and ω_H and ω_L are the high-pass and low-pass filter constants respectively. Finally, A_{ij} is the network connectivity matrix and n_{in} is the number of input nodes per node.

$$\dot{\boldsymbol{u}}_i(t) = \boldsymbol{E}\boldsymbol{u}_i(t) - \boldsymbol{F}\beta\cos(x_i(t) + \phi_0), \tag{2}$$

$$x_i(t) = \boldsymbol{G}\Big(\boldsymbol{u}_i(t-\tau_f) + \frac{\varepsilon}{n_{in}}\sum_j A_{ij}(\boldsymbol{u}_j(t-\tau_c) - \boldsymbol{u}_i(t-\tau_f))\Big), \tag{3}$$

where,

$$\boldsymbol{E} = \begin{bmatrix} -(\omega_L+\omega_H) & -\omega_L \\ \omega_H & 0 \end{bmatrix}, \boldsymbol{F} = \begin{bmatrix} \omega_L \\ 0 \end{bmatrix}, \boldsymbol{G} = \begin{bmatrix} 1 & 0 \end{bmatrix},$$

$$\tau_f = \tau_c = 1.9ms, \quad \phi_0 = \pi/4, \quad \omega_H/2\pi = 100Hz, \quad \text{and} \quad \omega_L/2\pi = 2.5kHz.$$

2.2 Method

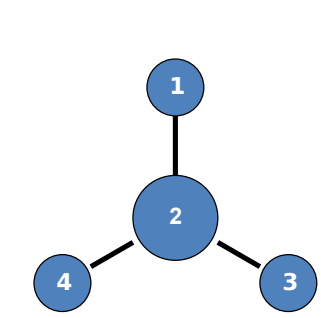
Trials on two different network configurations were conducted: the bidirectionally coupled star and chain networks in Fig. 2(a), 2(c). Consecutive runs were done on each network configuration, where ε, the coupling strength was increased in steps of 0.025, from 0 to 1. 2-second measurement runs were done using oscilloscopes. During the first 0.5 seconds, the nodes were allowed to oscillate with only self-feedback, and no coupling, in order to set them in a random, independent state. Coupling was then enabled and the next 0.1 seconds of data discarded. The remaining 1.4 seconds were used for data analysis.

2.3 Results

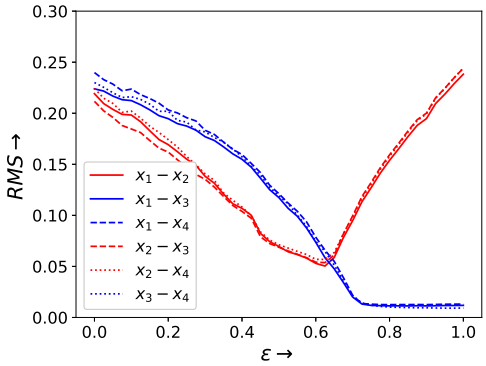
Root mean square differences between nodes were calculated as a function of ε for each possible combination of nodes in both networks ($\sqrt{\langle||\boldsymbol{x}_i - \boldsymbol{x}_j||^2\rangle}$ for $i < j$) and then plotted in Fig. 2(b), 2(d). As is visible, the star network achieved synchronisation between outer nodes for high values of coupling, whereas the chain network did not synchronise in any way, since the difference did not approach zero.

It can be seen in Fig. 2(b) that the RMS differences for node combinations 2-1, 2-3 and 2-4, and 1-3, 3-4 and 4-1 line up for all values of ε. This is due to the fact that there exist symmetries in the graph permuting these nodes to each other, since interchanging the 'arms' of the star graph does not alter the topology of the graph. Therefore one would expect that their general behaviour, and so any generic statistics, such as the one used here, would line up and give equal results. If these results were not equal, then that would suggest that symmetry had been broken in the experiment.

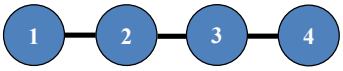
Similarly, it can be seen that in the chain network diagram 1-4 and 3-4 line up, as well as 1-3 and 2-4. This is due to the fact that the reflection symmetry in the chain graph permutes nodes 1 and 2 to 4 and 3, meaning that they have similar behaviour, and so one expects $RMS_{12} = RMS_{43} = RMS_{34}$. For the same reason, $RMS_{13} = RMS_{24}$. Combinations 1-4 and 2-3 can not be permuted to any other



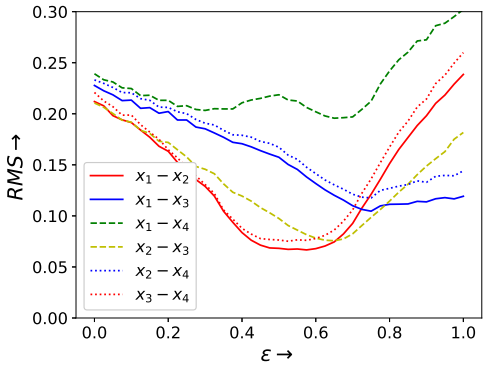
(a) Star network.



(b) Star network.



(c) Chain network.



(d) Chain network.

Fig. 2: (a, b)-Networks used in experimental setup. (c, d)-RMS difference calculated for each possible node combination as a function of coupling strength ε.

combination of nodes without changing the graph topology, and so they stand alone and do not cluster. We have now effectively identified the orbits of all two-node combinations under the action of the symmetry group in this graph.

These results may seem trivial. However, they are not. Many factors exist in this real-world setup that break the symmetry of the network. The round-trip gain β varies slightly in the different systems and can not be fixed exactly. The phase of the Mach-Zehnder modulators may shift slightly from the rest position and often needs to be recalibrated. Different lasers operate at slightly different intensities. These are all reasons for why the systems used in this setup are nominally homogeneous, but in reality only approximately the same. It can be concluded from this that network

graph symmetries can robustly present themselves as symmetries in the statistics of real-world experimental setups where different factors break exact symmetries.

3 Statistical Symmetries in the General Case

The previous section shows that network symmetries can cause symmetries in statistical data to arise. In this section we will detail which conditions are necessary for this to happen in the general case. This effort culminates in the 'Main Theorem' presented at the end of this section. We make use in the following of a so-called statistic of shape $S(\boldsymbol{a}_1, ..., \boldsymbol{a}_N) = H(\boldsymbol{s}_1(t), ..., \boldsymbol{s}_N(t))$. H can be imagined to be any calculation of the properties of a solution $s_i(t)$, such as the root mean square of the first node: $H(\boldsymbol{s}_1(t), ..., \boldsymbol{s}_N(t)) = \sqrt{\langle ||\boldsymbol{s}_1||^2 \rangle}$, or a cross-correlation between two nodes. First we define when a symmetry is present in a statistic using definition 3.

Definition 3 Define as a statistic a function $S(\boldsymbol{a}_1, ..., \boldsymbol{a}_N) = H(\boldsymbol{s}_1(t), ..., \boldsymbol{s}_N(t))$ that when applied to a particular initial condition $(\boldsymbol{a}_1, ..., \boldsymbol{a}_N)$, applies a function H to the corresponding solution of the initial conditions $(\boldsymbol{s}_1(t), ..., \boldsymbol{s}_N(t))$. Call a statistic S invariant under the action of symmetry g when it is averaged over some distribution of initial conditions p, if:

$$\int S(\boldsymbol{x}_1, ..., \boldsymbol{x}_N) p(\boldsymbol{x}_1, ..., \boldsymbol{x}_N) (dx)^{n \times N} = \int S(\boldsymbol{x}_{g(1)}, ..., \boldsymbol{x}_{g(N)}) p(\boldsymbol{x}_1, ..., \boldsymbol{x}_N) (dx)^{n \times N}. \tag{4}$$

In order to prove the main theorem we must first prove the following lemmas 1 and 2.

Lemma 1 (Symmetries in initial conditions continue down the line)
Let $\phi(\boldsymbol{x}_1, ..., \boldsymbol{x}_N)$ be a function of the variables of a dynamical system with symmetry g. Examine the initial conditions $\boldsymbol{c}_i$ and $\boldsymbol{c}_{g(i)}$, with $\boldsymbol{s}_i(t)$ and $\boldsymbol{z}_i(t)$ as respective solutions. Then:

- $\boldsymbol{z}_i(t) = \boldsymbol{s}_{g(i)}(t)$
- $\phi(\boldsymbol{z}_{g^{-1}(1)}(t), ..., \boldsymbol{z}_{g^{-1}(N)}(t)) = \phi(\boldsymbol{s}_1(t), ..., \boldsymbol{s}_N(t))$

Proof Due to the symmetry of the system, $\boldsymbol{s}_{g(i)}(t)$ is also a solution. Since $\boldsymbol{s}_{g(i)}(t) = \boldsymbol{c}_{g(i)} = \boldsymbol{z}_i(0)$, it follows from the uniqueness theorem that $\boldsymbol{s}_{g(i)}(t) = \boldsymbol{z}_i(t)$ and so, inserting $g^{-1}(i)$, we get $\phi(\boldsymbol{z}_{g^{-1}(1)}(t), ..., \boldsymbol{z}_{g^{-1}(N)}(t)) = \phi(\boldsymbol{s}_1(t), ..., \boldsymbol{s}_N(t))$. □

Lemma 2 *Let $\boldsymbol{a}_i$ be the initial conditions for a dynamical system with symmetry g and corresponding solution $\boldsymbol{s}_i(t)$. Let $\boldsymbol{b}_i = \boldsymbol{a}_{g(i)}$ and $\boldsymbol{z}_i(t)$ be the corresponding solution to the initial conditions $\boldsymbol{b}_i$. Define the statistic $S(\boldsymbol{a}_1, ..., \boldsymbol{a}_N) = H(\boldsymbol{s}_1(t), ..., \boldsymbol{s}_N(t))$ and $T(\boldsymbol{b}_1, ..., \boldsymbol{b}_N) = H(\boldsymbol{z}_{g^{-1}(1)}(t), ..., \boldsymbol{z}_{g^{-1}(N)}(t))$.*
Then $S(\boldsymbol{a}_1, ..., \boldsymbol{a}_N) = T(\boldsymbol{b}_1, ..., \boldsymbol{b}_N)$.

Proof

$$\begin{aligned} T(\boldsymbol{b}_1, ..., \boldsymbol{b}_N) &= H(\boldsymbol{z}_{g^{-1}(1)}(t), ..., \boldsymbol{z}_{g^{-1}(N)}(t)) \\ &= H(\boldsymbol{s}_1(t), ..., \boldsymbol{s}_N(t)) = S(\boldsymbol{a}_1, ..., \boldsymbol{a}_N). \end{aligned} \tag{5}$$

□

Main Theorem

Let $p(\boldsymbol{x}_1, ..., \boldsymbol{x}_N)$ be a distribution over the possible initial conditions. Let statistics $S(\boldsymbol{a}_1, ..., \boldsymbol{a}_N) = H(\boldsymbol{s}_1(t), ..., \boldsymbol{s}_N(t))$ and $T(\boldsymbol{b}_1, ..., \boldsymbol{b}_N) = H(\boldsymbol{z}_{g^{-1}(1)}(t), ..., \boldsymbol{z}_{g^{-1}(N)}(t))$. If p is invariant under the symmetry g, that is, $p(\boldsymbol{x}_1, ..., \boldsymbol{x}_N) = p(\boldsymbol{x}_{g(1)}, ..., \boldsymbol{x}_{g(N)})$, then:

$$\int_{\mathbb{R}^{N\times n}} p(\boldsymbol{x}_1, ..., \boldsymbol{x}_N) S(\boldsymbol{x}_1, ..., \boldsymbol{x}_N)(dx)^{N\times n} = \int_{\mathbb{R}^{N\times n}} p(\boldsymbol{x}_1, ..., \boldsymbol{x}_N) T(\boldsymbol{x}_1, ..., \boldsymbol{x}_N)(dx)^{N\times n}. \tag{6}$$

In other words, if a network has symmetry g and its initial conditions are invariant under g, then the statistic over the initial conditions will also be invariant under the action of g when averaged using distribution p.

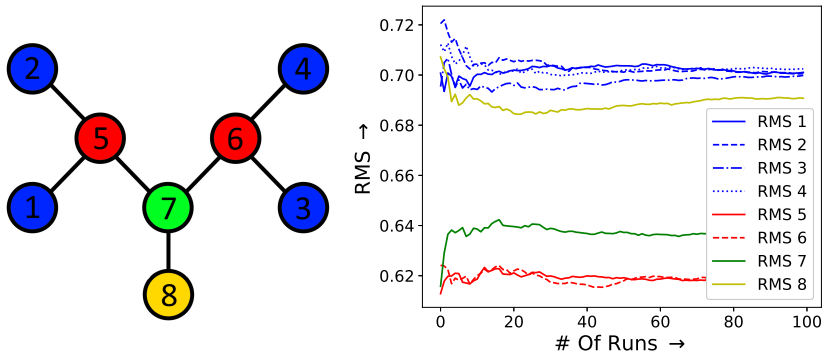
Proof

$$\begin{aligned} &\int_{\mathbb{R}^{N\times n}} p(\boldsymbol{x}_1, ..., \boldsymbol{x}_N) T(\boldsymbol{x}_1, ..., \boldsymbol{x}_N)(dx)^{N\times n} \\ &= \int_{\mathbb{R}^{N\times n}} p(\boldsymbol{y}_{g(1)}, ..., \boldsymbol{y}_{g(N)}) T(\boldsymbol{y}_{g(1)}, ..., \boldsymbol{y}_{g(N)})(dy)^{N\times n} * \\ &= \int_{\mathbb{R}^{N\times n}} p(\boldsymbol{y}_1, ..., \boldsymbol{y}_N) S(\boldsymbol{y}_1, ..., \boldsymbol{y}_N)(dy)^{N\times n} \\ &= \int_{\mathbb{R}^{N\times n}} p(\boldsymbol{x}_1, ..., \boldsymbol{x}_N) S(\boldsymbol{x}_1, ..., \boldsymbol{x}_N)(dx)^{N\times n}. \end{aligned} \tag{7}$$

* Here the change of variables $\boldsymbol{x}_i = \boldsymbol{y}_{g(i)}$ has been used, which has $|J| = 1$. □

This theorem is the main result of this article. It states that if an experiment on a network with a symmetry g is done in such a way that the choice of initial conditions for all trials p does not break the symmetry of the experiment, then all statistics which are averaged over the initial conditions will also be invariant under g. If the system is ergodic then the conditions can be much more lax: any time-averaged statistic will be invariant under g when averaged over sufficiently long time-series. In an ergodic system all symmetries will therefore be present in the data of a single run, instead of having to require that the statistics are averaged over the intial conditions.

4 Numerical Results



(a) 8-node opto-electronic network

(b) Average RMS after N runs

Fig. 3: (a)-Network used in simulation trials with nodes grouped into orbits by colour. (b)-Combined root mean square amplitudes from 100 trials of 2000 timesteps each

.

The results from the previous theoretical section have been demonstrated in two different, small, opto-electronic networks. The experimental setup is limited to experiments with at most 4 nodes. Simulations allow the results to be demonstrated in larger networks, with more control over the circumstances. The 8-node opto-electronic network from Fig. 3(a) was simulated. Simulation data from 100 trials with randomised initial conditions were combined. The root mean square amplitude ($\sqrt{\langle ||\boldsymbol{x}_i||^2 \rangle}$) of each node was calculated for every trial simulation of the opto-electronic network, as compared to the root mean square difference in section 2. This was done to identify the orbits of the nodes of the graph.

The nodes of any graph can be partitioned into distinct orbits O_i. An orbit O_i is a subset of the nodes of a graph with symmetry group G, where for each node $a, b \in O_i$, $\exists g \in G$, such that $a = g(b)$, and each orbit is closed under the action of the G. As is visible in Fig. 3(a), there are 4 orbits $\{1, 2, 3, 4\}, \{5, 6\}, \{7\}$, and $\{8\}$. Based on the main theorem, if network statistics are averaged over initial conditions in such a way that the conditions are invariant under the action of any symmetry, then one would expect these statistics to be invariant under the action of the network symmetries. Polling a network randomly over many different initial conditions is a close approximation to a continuous integral as stated in the main theorem, which is impossible to do in reality. Define a statistic RMS_i, which gives the root mean square of the signal from node i. If this network has symmetry g, then one expects that $RMS_i = RMS_{g(i)}$. This implies that any two nodes in the same orbit will have the same root mean square. We therefore expect in this numerical experiment to

see the root mean squares of the separate nodes to cluster along the lines of their respective orbital partitions as a consequence of the main theorem.

The data were found to converge into distinct groups. As is visible in Fig. 3(b), the four separate orbits of the graph in Fig. 3(a) can clearly be identified. We therefore confirm that the prediction of the main theorem holds in this case.

5 Conclusion

The main theorem in this work states that, when sampled under a distribution of initial conditions which is invariant under a symmetry g, any statistics calculated on these data must also be invariant under the action of the symmetry g. Symmetric networks therefore imply symmetric statistics. The converse was not shown to be true. Symmetries in time-series statistics are however still strong indicators of symmetries existing in a network. The work presented in both real and numeric experiments has shown that this result is indeed robust under ordinary real-world circumstances in real experiments, where symmetries are necessarily broken by small differences in the experiment. It has also been shown that the converse also reasonably holds for small networks. The question is whether the converse generally holds given that sufficient statistical testing and comparisons are done, and whether this can easily be extended to larger networks.

Assuming that the presence of network symmetries can be verified with a simple test, then in theory, the symmetry group of a network can be retrieved from experimental data. Since the symmetry group of a network with N nodes has up to $N!$ symmetries, it is unreasonable to check every symmetry. The question therefore arises how many tests need to be done to retrieve the full symmetry group, and what the algorithmic complexity is of this calculation. Regardless of the complexity of identifying the full group, this method can readily identify the orbits under the action of the entire network symmetry group, thereby already giving vital information about the network, and which clusters of synchrony may form [4].

Acknowledgements: This paper is a result of work performed as part of the TREND NSF REU program and the University of Maryland, College Park. JDH and RR are thankful for support from ONR Grant No. N000141612481. EvW is thankful for the aid of Olwen Enright and Timea Vitos for constructive input on the manuscript and Keshav Rakesh for his patient aid writing the proofs.

References

1. J.D. Hart, K. Bansal, T.E. Murphy and R. Roy, “Experimental observation of chimera and cluster states in a minimal globally coupled network”, Chaos **26**, 094801 (2016)
2. S. Pajevic, D. Plenz, “Efficient Network Reconstruction from Dynamical Cascades Identifies Small-World Topology of Neuronal Avalanches”, PLoS Comput Biol 5(1): e1000271 (2009)

3. A. Pikovsky, “Reconstruction of a neural network from a time series of firing rates”, Phys. Rev. E 93, 062313 (2016)
4. A.B. Siddique, L. Pecora, J.D. Hart, and F. Sorrentino “Symmetry- and input-cluster synchronization in networks”, Phys. Rev. E 97, 042217 (2018)