

Automatic reconstruction of complex dynamical networks

Marc Grau Leguia

TESI DOCTORAL UPF / 2019

DIRECTORS DE LA TESI:

Zoran Levnajić¹,

Ralph Gregor Andrzejak²

¹Faculty of Information Studies

²Universitat Pompeu Fabra, Department of Information and Communication Technologies



Agraïments — Agradecimientos

First and foremost, I would like to thank my two supervisors Zoran Levnajić and Ralph Gregor Andrzejak for giving me this opportunity and for their support and guidance throughout these years. Thanks Zoran for your passion, for your scientific discussions and for making sure that my time in Slovenia was pleasant. Hvala lepa. Thanks Ralph for welcoming me into your group and for your supervision and patience. I felt it was more than a secondment. I also thank you all the support you gave me to find my post-doc position in Bern.

Thanks all the PIs of the COSMOS consortium for guiding us all during these three years and for all the nice lectures and workshops that help us grow as a scientists. I especially thank Arkady, Michael and Antonio to make COSMOS possible and Caroline for your constant support and help. I also thank Bernard and Ljupčo for the nice discussions and meetings we had in the cold Ljubljana.

Thanks to all the COSMOS students for all the wonderful moments we had together. From Florence to Novo Mesto, each meeting was enjoyable thanks to the great atmosphere we created.

Thanks to my officemates in the UPF. Thanks Ola for the nice discussions and for not being mad at us when we switched all the time to Spanish. Thanks Petroula for your kindness and patience. Grazie Mario per stare sempre contento e prendermi in giro quando lo meritavo. Gracias Cristina por todo el "hateo" sano que hicieron más llevadero los problemas durante la tesi. Grazie Irene per le discussioni interessanti sul divano e per la passione che trasmetti sempre. Grazie Federico, per tutti i momenti divertenti facendo il "burro", siamo el "can culleretes" dell'ufficio.

Gràcies a tot el grup de companys d'universitat per tots els bons moments que hem passat. Gràcies Gemma per acollir-nos a Amsterdam i a la molt bonica ciutat de Rotterdam. Gràcies Toni per la loca Octoberfest que vem passar. Gràcies a la Elis per ajudar-me a sobreviure els primers mesos a Novo Mesto.

Gràcies Beltrol per les loques històries que ens expliques des de Hawai. Gràcies Bori, Hèctor, Reig i Toni pel gran viatge que fem fer per terres balcanes i per totes les barbacoes, jocs de taula etc que han fet possible que no acabés boig per la tesis. I Hèctor gràcies també per acollirme a terres Taiwaneses.

Grazie Giulia per tanti momenti belli che abbiamo passato e passeremo insieme. Ti voglio bene.

Finalment, dono les gràcies a la meva família pel seu suport i paciència que han tingut durant la tesi. Us estimo.

Resum

Un problema principal de la ciència de xarxes és com reconstruir (inferir) la topologia d'una xarxa real a partir de senyals mesurades de les seves unitats internes. Entendre la arquitectura d'una xarxa complexa és clau, no només per comprendre el seu funcionament, sinó també per predir i controlar el seu comportament. Els mètodes actualment disponibles es centren principalment en la detecció d'enllaços de xarxes no direccionals i sovint requereixen suposicions fortes sobre el sistema. Tanmateix, molts d'aquests mètodes no es poden aplicar a xarxes amb connexions direccionals. Per abordar aquest problema, en aquesta tesi ens centrarem en la inferència de xarxes direccionals. Concretament, desenvolupem un mètode de reconstrucció de xarxes basat en models que combina estadístiques de correlacions de derivades amb recuit simulat. A més, desenvolupem un mètode de reconstrucció basat en dades cimentat en una mesura d'interpedendència no lineal. Aquest mètode permet inferir la topologia de xarxes direccionals d'oscil·ladors caòtics de Lorenz per un subordre de la força d'acoblament i la densitat de la xarxa. Finalment, apliquem el mètode basat en dades a gravacions electroencefalogràfiques d'un pacient amb epilèpsia. Les xarxes cerebrals funcionals obtingudes a partir d'aquest mètode són coherents amb la informació mèdica disponible.

Paraules claus: Xarxes complexes, reconstrucció de xarxes, recuit simulat, mesura d'interpedendència no lineal, dinàmiques de Lorenz, gravacions electroencefalogràfiques (EEG), connectivitat funcional.

Resumen

Un problema principal de la ciencia de redes es cómo reconstruir (inferir) la topología de una red real usando las señales medidas de sus unidades internas. Entender la arquitectura de redes complejas es clave, no solo para entender su funcionamiento pero también para predecir y controlar su comportamiento. Los métodos existentes se focalizan en la detección de redes no direccionales y normalmente requieren fuertes suposiciones sobre el sistema. Sin embargo, muchos de estos métodos no pueden ser aplicados en redes con conexiones direccionales. Para abordar este problema, en esta tesis estudiamos la reconstrucción de redes direccionales. En concreto, desarrollamos un método de reconstrucción basado en modelos que combina estadísticas de correlaciones de derivadas con recocido simulado. Además, desarrollamos un método basado en datos cimentado en una medida d'interdependencia no lineal. Este método permite inferir la topología de redes direccionales de osciladores caóticos de Lorenz para un subrango de la fuerza de acoplamiento y densidad de la red. Finalmente, aplicamos el método basado en datos a grabaciones electroencefalográficas de un paciente con epilepsia. Las redes cerebrales funcionales obtenidas usando este método son consistentes con la información médica disponible.

Palabras claves: Redes complejas, reconstrucción de redes, recocido simulado, medida d'interdependencia no lineal, dinámicas de Lorenz, grabaciones electroencefalográficas (EEG), conectividad funcional.

Abstract

A foremost problem in network science is how to reconstruct (infer) the topology of a real network from signals measured from its internal units. Grasping the architecture of complex networks is key, not only to understand their functioning, but also to predict and control their behaviour. Currently available methods largely focus on the detection of links of undirected networks and often require strong assumptions about the system. However, many of these methods cannot be applied to networks with directional connections. To address this problem, in this doctoral work we focus at the inference of directed networks. Specifically, we develop a model-based network reconstruction method that combines statistics of derivative-variable correlations with simulated annealing. We furthermore develop a data-driven reconstruction method based on a nonlinear interdependence measure. This method allows one to infer the topology of directed networks of chaotic Lorenz oscillators for a subrange of the coupling strength and link density. Finally, we apply the data-driven method to multichannel electroencephalographic recordings from an epilepsy patient. The functional brain networks obtained from this approach are consistent with the available medical information.

Keywords: Complex networks, network reconstruction, simulated annealing, nonlinear interdependence measure, Lorenz dynamics, electroencephalographic recordings (EEG), functional connectivity.

Contents

Abstract	vii
Contents	x
List of figures	xi
List of tables	xv
Chapter 1 INTRODUCTION	1
1.1. Networks	1
1.1.1. Networks in nature	3
1.1.2. Network models	5
1.2. Dynamics on networks	6
1.3. Network inference	8
1.4. Contribution of this thesis	10
Chapter 2 EVOLUTIONARY OPTIMIZATION	13
2.1. Introduction	13
2.2. Methods	14
2.2.1. The derivative-variable correlation method	14
2.2.2. Generalization of the derivative-variable correlation method via simulated annealing	16
2.3. Results	19

2.3.1.	Comparison of two reconstruction schemes	19
2.3.2.	Network reconstruction from different dynamical regimes	21
2.3.3.	Comparison of reconstruction schemes for varying data quality	23
2.3.4.	Influence of noise	26
2.4.	Discussion	28
Chapter 3	INFERRING DIRECTED NETWORKS USING A RANK BASED CONNECTIVITY MEASURE	31
3.1.	Introduction	31
3.2.	Methods	32
3.2.1.	Rank based connectivity measure L	32
3.2.2.	Model	33
3.2.3.	Accuracy of the reconstruction	34
3.3.	Results	37
3.4.	Discussion	47
Chapter 4	INFERRING FUNCTIONAL CONNECTIVITY FROM EEG DATA	51
4.1.	Introduction	51
4.2.	Functional brain connectivity	51
4.3.	Dataset and Preprocessing	52
4.4.	Functional connectivity analysis using L	53
4.5.	Discussion	60
Chapter 5	DISCUSSION	63
5.1.	Future work	65
Appendix A	SPATIAL STRUCTURE OF ALL SEIZURES	67
References		71
SUBJECT INDEX		81

List of Figures

1.1. Hospital contact network generated from all the contacts between workers and patients in a hospital for an hour of recordings (9am)	5
1.2. Random directed graphs with $N = 16$ at three different link probabilities $\rho = 0.01, 0.1, 0.5$ (a, b, c, respectively).	5
2.1. Schematic representation on how we compute the reconstructed time series $y_p(t_l)$	16
2.2. Both annealing schemes outperform the random guessing.	20
2.3. We improve the reconstruction by minimizing the trajectory error Δ_T via simulated annealing.	21
2.4. The dynamical regime of the system plays a key role in the correlation between Δ_T and Δ_A	22
2.5. An interplay between quality of derivatives and phase space coverage is found for the adjacency matrix error.	24
2.6. We find good reconstruction for as few as 14 data points.	26
2.7. Minimizing directly the error of the noisy trajectories gets you closer to the real trajectories than to the noisy ones.	27
3.1. Illustration of the ROC space including different ROC curves with three distinct behaviours.	36
3.2. Different coupling strengths and link densities lead to different qualities of the network reconstruction.	37

3.3.	ROC curves allow us to quantify the accuracy of the reconstruction for a variable threshold.	38
3.4.	Performance of the network reconstruction for the noise-free heterogeneous dynamics is highest for sparse networks with intermediate coupling between nodes.	40
3.5.	Compared to the results for the heterogeneous noise-free system, the performance of the network reconstruction is degraded for the homogeneous noise-free system.	41
3.6.	For intermediate and high coupling strengths, heterogeneity helps the reconstruction.	42
3.7.	Depending of the choice of ε and ρ , there exists an ideal level of noise which improves the reconstruction as compared to the homogeneous noise-free case.	44
3.8.	Dynamical noise increases the performance for a wide range of parameters.	45
3.9.	Performance of the network reconstruction is decreased for increasing network size M when the link density ρ is not adjusted.	46
3.10.	Performance of the network reconstruction for the noise-free heterogeneous dynamics does not change substantially if the system size grows as long as the link density is properly adjusted. Same as described in the caption of Fig. 3.9 but using ρ_M	48
4.1.	Temporal evolution of the EEG recordings from the seizure number three during 16 seconds. Labels in the ordinate axis illustrate the regions of the brain from which the signals were recorded. Signals plotted in red depict the channels that were discarded because they were located in the white matter of the brain. The vertical red line marks the start of the seizure. The total number of initial channels is 56.	53
4.2.	Same as Fig. 4.1 but discarding the red signals of Fig. 4.1 and making a bipolar montage. The number of channels is now 35.	54

4.3.	Connectivity matrices computed from seizure three from windows corresponding to the beginning of the seizure. Labels in the connectivity matrices depict the regions of the brain from which the signals were recorded. The matrices are computed from 20 consecutive windows. Time ordering in the connectivity matrices goes from left to right and from top to bottom.	54
4.4.	Same as Fig. 4.3 but for the end of seizure three.	55
4.5.	High variability of connectivity structure is found across the different periods included in the EEG recording from seizure 3.	56
4.6.	High variability of the topological structure is found across the different periods included in the EEG recording from seizure 3.	56
4.7.	A high overall connectivity is found during the seizure with a predominant direction from the SOZ to the remaining brain areas at the beginning of the seizure.	60
A.1.	Same as Fig. 4.5 but for seizure 1.	68
A.2.	Same as Fig. 4.5 but for seizure 2.	68
A.3.	Same as Fig. 4.5 but for seizure 4.	68
A.4.	Same as Fig. 4.5 but for seizure 5.	69
A.5.	Same as Fig. 4.5 but for seizure 6.	69
A.6.	Same as Fig. 4.5 but for seizure 7.	69

List of Tables

3.1.	Confusion matrix of the comparison between the true adjacency matrix A and the reconstructed one R^n	35
4.1.	The average outdegree is always higher during the ictal period.	58
4.2.	The average indegree is always higher during the ictal period. Same as described in the caption of Tab. 4.1 but for the indegree.	58

CHAPTER 1

Introduction

A rich variety of natural and man-made systems are composed of separate units that interact in a complex way. In network science, these systems are modeled using networks where each component is represented by a node and their interaction is conveyed by links between the nodes. This rapidly growing field has helped to advance our understanding of real-world systems ranging from social to natural sciences (Newman, 2007; Barabási and Pósfai, 2016). To understand the collective dynamics of these complex systems, it is essential to characterize not only the behaviour of individual nodes, but also the network structure that makes them interact via links. This interaction creates a collective effect that does not exist at the level of individual nodes. However, unravelling the network structure from the available data is far from easy. This is the core topic of this doctoral dissertation.

1.1 Networks

The first time a graph was used to represent a real-world problem can be traced back to the XVIII century. Citizens in the town of Königsberg did not know if one could cross all the bridges in the city only once. Euler, the famous Swiss mathematician, solved this problem mathematically using a very simple graph representation. He showed that it was not

possible to cross all bridges only once (Alexanderson, 2006). However only two decades ago, networks started to get attention from other areas of science.

A network is a simple representation of the interaction of a given system. The nodes are the entities of the system that interact among themselves via links. We say that a node p is connected to node q if there is a direct link between them. Using graph theory, we mathematically represent a network via an adjacency matrix:

$$A_{pq} = \begin{cases} 1 & \text{if there is a link from } p \text{ to } q; \\ 0 & \text{otherwise;} \end{cases} \quad (1.1)$$

where the indices of the matrix go from $p, q \in (1, 2, \dots, M)$, and M is the number of nodes or entities in the system. We additionally define E as the number of links of the system. For an undirected network, the connections are bidirectional so the adjacency matrix A is symmetrical $A_{pq} = A_{qp}$. When the connections have a direction, $A_{pq} = A_{qp}$ does not hold in general, and we say that we have a directed network. In this kind of network, the direction of the links determines the flow of the information in the network. The maximum number of links in a network depends only on the number of nodes M . For a undirected network, the maximum number of links without counting self-loops is $E_{max} = M(M - 1)/2$. Alternatively, for a directed network this quantity is doubled $E_{max} = M(M - 1)$. Moreover, networks can be used to represent the strength of some inter-unit interaction. In this case, we talk about weighted networks where the adjacency matrix is no longer a binary matrix and each link is represented with a real value $A_{pq} = w_{pq}$ with $w_{pq} \in \mathbb{R}$. We furthermore define the number of links pointing to the node and the number of links going from that node to other nodes as the indegree (d_p^{in}) and outdegree (d_p^{out}), respectively:

$$d_p^{in} = \sum_{q=1}^M A_{qp}, \quad d_p^{out} = \sum_{q=1}^M A_{pq}. \quad (1.2)$$

For undirected networks the indegree is equal to the outdegree: $d_p^{in} = d_p^{out}$. In weighted networks, the degree does not longer correspond to the number of links pointing to or going from a node, and it is more a measure of participation of the node in the network.

This way one can represent networks (real-life systems with empirically determined properties) via graphs (mathematical objects with formal properties expressible through equations). Of course, networks and graphs can have multiple structural characteristics, and have a myriad of other properties. For the purposes of this thesis, we will not define further structural characteristics since the core topic of the thesis is to study methods to reconstruct a network.

1.1.1 Networks in nature

Networks are used as models of a wide variety of systems in nature, society and technology. These includes:

1. Social networks, as the most traditional category, have individuals, companies, societies, groups or even countries as nodes, and examine diverse types of relationships (links) between them (Killworth, Johnsen, Bernard, Ann Shelley, and McCarty, 1990; Russell Bernard, Johnsen, Killworth, and Robinson, 1991; Wasserman and Faust, 1994; Amaral, Scala, Barthélemy, and Stanley, 2000). These include friendship, romance, trade/business or import/export. These networks serve not only to understand better the relationships in society, but also to predict and tackle events such as business cycles and economic downturns.
2. Networks in biology and neuroscience assign biological entities to nodes, which include genes (Mittler, Vanderauwera, Gollery, and Breusegem, 2004), proteins (Jeong, Mason, Barabási, and Oltvai, 2001), neurons (or regions of the brain) (White, Southgate, Thomson, and Brenner, 1986) or animal species (Bascompte, Jordano, Melián, and Olesen, 2003). Research then focuses on the relationships between these entities, drawing conclusions that range from efficiency of gene or protein interaction networks to the stability of an ecosystem investigated via food networks among animal species.
3. Networks in infrastructure and technology assign diverse transport units such as airports and bus stops to nodes, and then look at the diversification and efficiency of, say, an infrastructure/transportation system or network (Kansky, 1963; Banavar, Maritan, and Rinaldo, 1999; Amaral et al., 2000; Gastner and Newman, 2004). Such

analysis can for instance yield insights into the stability of power grids, crucial for the functioning of economy.

4. Information networks nowadays are sometimes confused with social networks, since the bulk of social interactions today occur online, which blurs the border between true friendship networks and, say, Facebook network (Huberman, 2001; Lewis, Kaufman, Gonzalez, Wimmer, and Christakis, 2008). Along the same lines research studies knowledge networks such as Wikipedia and alike (Korfiatis, Poulos, and Bokos, 2006).

Networks as a framework are also used to model and analyze many other complex systems in nature, society and technology (Newman, 2007; Barabási and Pósfai, 2016), but for the purposes of this work we will not dive further into them. One of the main characteristics of these real-world networks is that most of them have considerably fewer links than the possible maximum number ($E \ll E_{max}$) (Barabási and Pósfai, 2016). This characteristic which we call *sparsity* is a very common feature among real networks (Newman, 2007). We say that a network is sparse if as the network size grows, the number of links per node remains constant. Similarly, the density of the network (ratio between links and total possible links of the network) tends to zero as the network size goes to infinity. For the cases where the network can grow (like most social networks), we can use this definition to better identify its sparseness.

In Fig. 1.1 we illustrate a social network of workers and patients in a hospital during one hour in the morning shift (Vanhems et al., 2013). This network depicts all the contacts made in a hospital during one hour (9–10 am). We observe that as other real-world cases, the network is sparse. Moreover, many of the nodes are isolated since they did not make any contact during this hour.

Other studies in network science over the last two decades revealed a whole range of other details on networks and their topologies (Watts and Strogatz, 1998; Barabási and Albert, 1999; Faloutsos, Faloutsos, and Faloutsos, 1999; Amaral et al., 2000; Adamic, Lukose, Puniyani, and Huberman, 2001; Liljeros, Edling, Amaral, Stanley, and Aaberg, 2001; Bassett and Bullmore, 2006; Boccaletti, Latora, Moreno, Chavez, and Hwang, 2006).

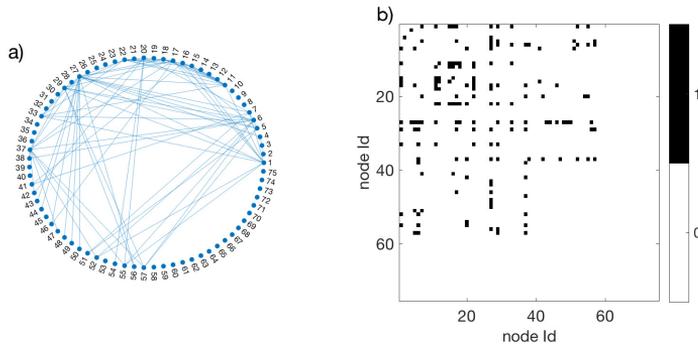


Figure 1.1: Hospital contact network generated from all the contacts between workers and patients in a hospital for an hour of recordings at 9–10 am (a) ($M=75$, $E=82$). Matrix representation of the hospital network (b). The figure is made using the recordings from Ref. (Vanhems et al., 2013).

1.1.2 Network models

Researchers often design mathematical models of networks, designed to replicate and capture the empirical properties observed in nature and society. One such example is to assume that all connections between nodes are generated randomly. A random or an Erdős-Rényi network (Erdős and Rényi, 1959) is defined by the link probability ρ . In this model, each possible link of the network has a probability ρ to be generated. The parameter ρ controls how dense the network is. Accordingly, we will use ρ as the network density throughout the thesis. For $\rho = 0$, we obtain a completely disconnected network whereas for $\rho = 1$ we generate a complete graph. For values between these two extremes, the network

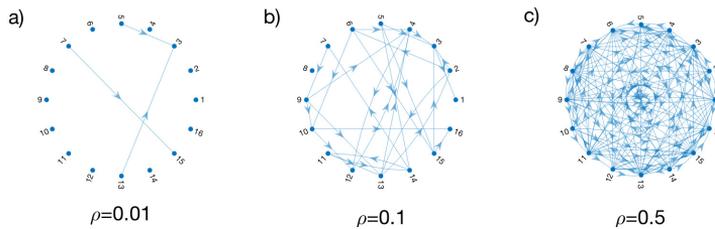


Figure 1.2: Random directed graphs with $N = 16$ at three different link probabilities $\rho = 0.01, 0.1, 0.5$ (a, b, c, respectively).

gets progressively dense (Fig. 1.2).

Many similar models of synthetic networks are also in use to grasp and

predict the properties of real networks. Models such as Watts-Strogatz (Watts and Strogatz, 1998) and Barabási-Albert (Barabási and Albert, 1999) are very well known for replicating the small-world and the scale-free propriety, respectively. However, for the purposes of this thesis, we will limit our attention only to the Erdős-Rényi model. One of the reasons why we consider this artificial network model is because we can adjust the sparsity to the levels of observed networks, so that this crucial property is adequately captured.

1.2 Dynamics on networks

In the previous section we introduced concepts from graph theory related with the structure and the mathematical form of a network. A lot of work is being done regarding the topological proprieties of a network. Understanding the topological features of a network is key to interpret the behavior when some dynamics run on the nodes of the network. A very clear example can be found in epidemic models which show how a disease is spreading on different topologies (Pastor-Satorras, Castellano, Van Mieghem, and Vespignani, 2015). As opposed to mere *social networks* which includes persons (individuals) as nodes, here we attach a state to each node, that can be either *healthy* or *sick*. These states then change over time, as a model of disease spread, generating non-trivial dynamics on this social network. Similarly as the spreading of viruses, we can assign ideas or opinions to the nodes (Clifford and Sudbury, 1973; Castellano, Fortunato, and Loreto, 2009, 2; Jackson, 2013) and study how the topology affects their spreading (Weng, Menczer, and Ahn, 2013; Lerman, Yan, and Wu, 2016).

In addition, a lot of effort is also put at studying very well-known dynamical systems coupled in a network. In these cases, similar to classical statistical physics, the emergent behaviour of the system cannot be understood completely by the sum of its components and the topology is key. In this area, the emergence of synchrony (Pikovsky, Rosenblum, and Kurths, 2001) in complex topologies is very well studied (Arenas, Díaz-Guilera, Kurths, Moreno, and Zhou, 2008). In general, a complex dynamical system whose elements are coupled in a network can be

described as follows:

$$\dot{x}_p = f(x_p) + \sum_{q=1}^M A_{qp} h(x_p, x_q), \quad (1.3)$$

where x is a time-dependent variable which describes the state of the node. The function f accounts for the intrinsic behaviour of the individual nodes and h for the inter-node interaction. The connectivity is encoded in the adjacency matrix A and therefore nodes p and q interact only if there is a link between them. As we said, a complex topology in Eq. 1.3 can generate non-trivial phenomena. It is worth to mention the case where the variables in Eq. 1.3 are phases and the function f is a constant phase velocity (ω_p) :

$$\dot{\theta}_p = \omega_p + \sum_{q=1}^M A_{qp} h(\theta_q - \theta_p), \quad (1.4)$$

which leads to a Kuramoto-like model when $h = \sin$ (Kuramoto, 1975; Acebrón, Bonilla, Pérez Vicente, Ritort, and Spigler, 2005). Studies on models similar to Eq. 1.4 have proved to be very prolific and have revealed a range of interesting collective dynamical phenomena (Kuramoto, 1975; Strogatz, 2001; Gómez-Gardeñes, Gómez, Arenas, and Moreno, 2011; Porter and Gleeson, 2016; Rodrigues, Peron, Ji, and Kurths, 2016).

Additionally, plenty of dynamical models were developed that capture other natural systems of physical interest. These include various models of neuronal interactions, gene interactions or interaction among bacteria (Hansel and Sompolinsky, 1993; Widder, Schicho, and Schuster, 2007; Ermentrout and Terman, 2010; Bressloff, 2011; Hoppensteadt and Izhikevich, 2012). An additional approach is to use standard systems from chaotic dynamics such as Lorenz dynamics (Duan and Chen, 2009), Rössler dynamics (Rössler, 1976; C. Li and Chen, 2004; Zhou and Kurths, 2006) or Logistic maps (Grassberger and Procaccia, 1983; Atay, Jost, and Wende, 2004), couple them in a complex way and study global properties such as their synchronization.

We here also note that dynamics on networks is not to be confused with dynamics of networks, where one studies a topology that changes over time. To clarify, we are here interested only in networks in which the nodes states are dynamically changing, not the network's topology.

1.3 Network inference

As mentioned before, the topology of a network has a key role in the resulting dynamics of the system. Most of the research in network science focuses on studying the dynamics resulting from a given network topology. However, our knowledge about the structure of real networks is often unreliable, mainly due to incomplete accessibility and observability of the system, which can lead to infer networks with missing or spurious links (Newman, 2007; Guimerà and Sales-Pardo, 2009; Timme and Casadiego, 2014). Alternatively, data from dynamics can often be measured with good precision. Nonetheless, inferring the structure of a network from dynamical data is far from being simple. This calls for the development of methods for reconstructing the network structure from the available dynamical data. Knowledge about the network structure is not just useful for understanding its functioning, but can help designing methods for its control or long-term prediction of its behaviour. Also, grasping the architecture principles of networks, especially those that underwent evolution (e.g. biological systems) gives inspiration for engineering artificial networks with prescribed functions. We refer to network inference or reconstruction as the area in network science that aims to infer the connections of a network using the information of the signals of the nodes. Studies that aim at this inference can be found in multiple scientific areas including gene regulatory networks (Margolin et al., 2006; Z. Li, Li, Krishnan, and Liu, 2011; Trejo Banos, Millar, and Sanguinetti, 2015), neurobiology (Chicharro, Andrzejak, and Ledberg, 2011; Oates et al., 2014), ecology (Sugihara et al., 2012), epidemiology (Tomovski and Kocarev, 2015) and climate science (Deza, Barreiro, and Masoller, 2015; Tirabassi, Sommerlade, and Masoller, 2017), among others.

In general, we can address the inference problem in two different ways depending on the type of assumptions we make:

1. *Model-based*: One approach is to use certain models for the nodes and possibly also for the type of interaction. In combination with the temporal evolution of the dynamics of each node, this information allows one to reconstruct the network (Timme, 2007; Levnajić and Pikovsky, 2011; Shandilya and Timme, 2011; Levnajić and Pikovsky, 2014; Ching, Lai, and Leung, 2015; Han, Shen, Wang, and Di, 2015; Pikovsky, 2016; Wang, Lai, and Grebogi, 2016; Cestnik and Rosenblum, 2017). Some reconstruction methods are

focused on Kuramoto-like systems (Kuramoto, 1975), where the system can synchronize (Arenas et al., 2008). In that case the reconstruction is not trivial because the synchronization overshadows most of the information of the system. These methods hence often employ external driving or phase resets to infer the topology (Timme, 2007; Levnajić and Pikovsky, 2011; Prigamo and Díaz-Guilera, 2012). Contrary to the methods developed for the study of networks of phase oscillators, there is a range of methods applicable to general dynamical systems (Levnajić and Pikovsky, 2014; Ching et al., 2015; Pikovsky, 2016). The study of reconstructing networks that vary in time also has been a subject of interest (Mei, Wu, Chen, and Lu, 2016). Some reconstruction methods are based on power expansions (Wang, Yang, Lai, Kovanis, and Harrison, 2011) or Laplacian spectra (Ipsen and Mikhailov, 2002). Several methods are invasive, i.e., rely on perturbing the system (Timme, 2007; Levnajić and Pikovsky, 2011), which in practice is often difficult or even impossible. Moreover, some methods make use of large dynamical data sets, which is rarely experimentally feasible. Indeed, this field calls for more research, particularly by exploring other known methods from other areas that could be of use for the reconstruction problems. Along those lines is the non-invasive method proposed in Ref. (Levnajić and Pikovsky, 2014), which is applicable to short data sets.

2. *Data-driven*: The main problem of the model-based methods is that they can only work when the assumptions about the model are correct. To avoid this problem, one can use a data-driven approach using exclusively the knowledge of the temporal evolution of the nodes without selecting any model. Although data-driven approaches make implicit assumptions about the measured signals, for example whether the dynamics are linear or nonlinear, it remains more applicable to real-world systems. Using this approach, measures like cross correlation, mutual information, or mutual information rate can perform well (Rubido et al., 2014; Tirabassi, Sevilla-Escoboza, Buldú, and Masoller, 2015; Bianco-Martinez, Rubido, Antonopoulos, and Baptista, 2016). However, these measures are not able to determine the direction of interaction between nodes. Partial mutual information from mixed embedding is able to infer networks which have almost the same topological structures

like the original ones (Kugiumtzis, 2013a; Koutlis and Kugiumtzis, 2016). Furthermore, measures like partial directed coherence and partial transfer entropy were used for small systems with particular topologies (Winterhalder, Schelter, and Timmer, 2007; Jachan et al., 2009; Schelter, Timmer, and Eichler, 2009; Sommerlade et al., 2009; Kugiumtzis, 2013b).

Despite the progress summarized above, inferring the exact connectivity matrix for a directed network without a specific choice of the network configuration remains a difficult task.

In this context we also mention the problem of network observability. A complex system (network) is called observable if we can reconstruct its complete internal state from measurable outputs (Liu, Slotine, and Barabási, 2013; Yan et al., 2015; Whalen, Brennan, Sauer, and Schiff, 2015). However, making measurements on complex networks can present several practical and technical difficulties. In fact, any description of real networks is ultimately limited by our ability to estimate (extract) the system's internal state (and its changes over time) from experimentally accessible outputs. Consider electroencephalographic (EEG) measurements, which are aimed at observing the internal dynamics of the brain. Those measurements are done using a limited number of electrodes, and are at best indirectly measuring the actual brain's state and dynamics. However, even such incomplete information on a complicated system (brain) is useful for estimating brain networks with certain precision. While we recognize the problem of network observability as an important one, in this thesis we will assume that we have complete observability of the system we study. Accordingly, we do not consider that any hidden node or agent is affecting the dynamics of our systems.

1.4 Contribution of this thesis

This thesis is part of the COSMOS project which involved working in two different universities. In particular, I carried out my work in the Faculty of Information Studies, in Novo mesto, Slovenia and in the Universitat Pompeu Fabra, in Barcelona, Spain. This collaboration led to study the inference problem from two different perspectives. First, in Chap. 2 we studied the inference problem from a model-based perspective and we introduced an extension of a derivative-variable correlation method

described in Ref. (Levnajić and Pikovsky, 2014). Our approach tries to improve the results in Ref. (Levnajić and Pikovsky, 2014) by exploiting simulated annealing. In the second part of the study, we address the reconstruction problem using a data-driven approach (Chap. 3). For this purpose, we use the state space measure L defined in (Chicharro and Andrzejak, 2009). We furthermore apply the data-driven method in Chap. 3 to multichannel EEG recordings of an epilepsy patient (Chap. 4). Finally, in Chap. 5 we summarize and discuss our results. Moreover, we discuss possible future extensions of our work.

Hence, our contribution is the development of two new network reconstruction methods, designed on two different setups and applicable in different contexts. We furthermore apply our later reconstruction method to a real-world data example.

CHAPTER 2

Evolutionary optimization of network reconstruction from derivative-variable correlations

2.1 Introduction

In Chap. 1 we mentioned that there were two ways to face the reconstruction problem; One model-based and the other data-driven. In this chapter we will approach the inference problem with a model-based perspective making assumptions on the equations controlling the behavior of each of the nodes in the system. We will follow closely the text of our study in Ref. (Leguia, Andrzejak, and Levnajić, 2017) published in the peer-reviewed journal of physics A: mathematical and theoretical. The author of this thesis is also the first author of this paper.

We present an extension of a derivative-variable correlations method described in Ref. (Levnajić and Pikovsky, 2014) (Sec. 2.2.1). Our approach tries to improve the results in Ref. (Levnajić and Pikovsky, 2014) by exploiting the simulated annealing (evolutionary optimization algorithm described in Sec. 2.2.2). We study the performance of the method for different dynamical regimes (Sec. 2.3.2). Moreover, we investigate the effect of the data quality and observational noise in the reconstruction (Sec. 2.3.3). Finally, we discuss our results in Sec. 2.4.

2.2 Methods

2.2.1 The derivative-variable correlation method

Here, we summarize the derivative-variable correlation method developed in Ref. (Levnajić and Pikovsky, 2014). We consider a general complex dynamical system composed of M nodes, whose dynamical states are represented by the time-dependent variables x_p . We consider a particular case of Eq. 1.3 where the dynamics can be described by:

$$\dot{x}_p = f_p(x_p) + \sum_{q=1}^M A_{qp} h_q(x_q), \quad (2.1)$$

where \dot{x}_p is the time-derivative of x_p , the function f_p represents the local dynamics, and h_q is the inter-node interaction, which expresses how node q acts on the other nodes. The adjacency matrix A_{pq} represents the structure of the network by capturing the connections between node p and q . It is this adjacency matrix that we seek to find. This dynamical set-up is widely used in computational neuroscience (Ermentrout and Terman, 2010; Bressloff, 2011; Hoppensteadt and Izhikevich, 2012).

We assume that:

1. The system behaves as Eq. 2.1,
2. The functions f_p and h_q are known.

After setting these assumptions, we define the following correlation matrices:

$$\begin{aligned} B &= \langle g(x_p) \dot{x}_q \rangle, \\ C &= \langle g(x_p) f_q(x_q) \rangle, \\ E &= \langle g(x_p) h_q(x_q) \rangle. \end{aligned} \quad (2.2)$$

Where $\langle \cdot \rangle$ represents time-averaging, the indices go from $p, q = (1, 2 \dots M)$, and $g(x)$ is a function that plays a key role in maximizing the precision of the reconstruction. Its explicit details will be explained later. If we multiply Eq. 2.1 by g we get:

$$g(x_k) \dot{x}_p = g(x_k) f_p(x_p) + \sum_{q=1}^M A_{qp} g(x_k) x_q. \quad (2.3)$$

If we make averages across time, we can identify the matrices defined in Eq. 2.2 and rewrite Eq. 2.3 as:

$$R = E^{-1}(B - C), \quad (2.4)$$

where R represents the reconstructed adjacency matrix. This equation is always applicable as long as the system behaves as Eq. 2.1. The matrices we defined in Eq. 2.2 can always be computed since our observable are L values ($x_p(t_1) \dots x_p(t_L)$) for each of the M nodes. These measurements are equally separated in time by τ . For infinitely long time series, the choice of g become less important. In realistic scenarios, however, where the time series are short and/or noisy, the selection of g plays the crucial role by allowing for tuning of the reconstruction equation above. In Ref. (Levnajić and Pikovsky, 2014) the function g was taken to be a periodic function on the interval covered by our observable data. For simplicity, g was defined as the first 10 Fourier components:

$$g(x) = \sum_{m=1}^{10} [a_m \sin(mx) + b_m \cos(mx)]. \quad (2.5)$$

Tuning of the reconstruction Eq. 2.4 was implemented by choosing the coefficients a_m and b_m at random (many times) and then use the g 's which lead to best precision. To quantify the precision for various realizations of g , we define the reconstruction error Δ_A as the deviation of the reconstructed adjacency matrix R from the original adjacency matrix A :

$$\Delta_A = \sqrt{\frac{\sum_{pq} (R_{pq} - A_{pq})^2}{\sum_{pq} A_{pq}^2}}. \quad (2.6)$$

In real-world applications the matrix A is of course unknown, so accordingly, Δ_A cannot be calculated directly. We need another way to quantify the precision of our measurements. To do so, we use R to check how well we reproduce our original data. For instance we start the dynamics from $x_p(t_1)$ and run Eq. 2.1 using R as an estimate for A until $t = t_2$ to obtain $y_p(t_2)$. Then we restart the run with $x_p(t_2)$ until $t = t_3$ obtaining $y_p(t_3)$ and so on (Fig. 2.1). The discrepancy between the reconstructed time series $y_p(t_i)$ and the original time series $x_p(t_i)$ is a measure of the reconstruction precision based only on the available data. This leads to the definition of the trajectory error Δ_T as:

$$\Delta_T = \frac{1}{M} \sum_{p=1}^M \sqrt{\frac{\langle (x_p - y_p)^2 \rangle}{\langle (x_p - \langle x_p \rangle)^2 \rangle}}. \quad (2.7)$$

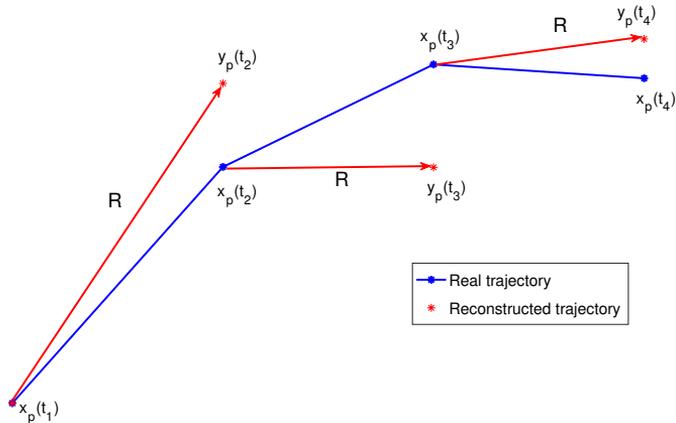


Figure 2.1: Schematic representation on how we compute the reconstructed time series $y_p(t_l)$

Where $\langle \cdot \rangle$ represents time-averaging. One of the key results from Ref. (Levnajić and Pikovsky, 2014) is that in general, Δ_A and Δ_T are correlated. Thus, low values of Δ_T generally imply low values of Δ_A . Each different choice of the coefficients a_m and b_m in Eq. 2.5 leads to a different reconstruction precision. In other terms, g is a tunable parameter that we can use to maximize the precision, by considering only the choices of g that lead to minimal Δ_T . In contrast to Ref. (Levnajić and Pikovsky, 2014), the Fourier coefficients will not be chosen at random, but we will instead apply a simulated annealing algorithm to mutate the Fourier coefficients. We explain the simulating annealing in the following section. As we show in what follows, this indeed in general leads to an enhanced performance.

2.2.2 Generalization of the derivative-variable correlation method via simulated annealing

It is easy to acknowledge that taking a big sample of random functions g and using each of them with Eq. 2.4 might not be the optimal approach. The search space we are dealing with is not known at all and is in general very rugged. So even a very large number of random guesses for g , does

not guarantee a sufficient exploration of the search space.

Fortunately, scenarios like this - because of their notoriety - are well studied, particularly in computer science, with a range of heuristics applicable depending on the specific case at hand. Among the most standard ones are evolutionary optimization algorithms, such as simulated annealing (Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller, 1953). Simulated annealing is an optimization technique that is used to explore rugged and large search spaces while avoiding the local shallow minima (Isakov, Zintchenko, Rønnow, and Troyer, 2015), used in a vast variety of scientific problems. It is related to the Metropolis-Hastings algorithm (Metropolis et al., 1953) often used for statistical mechanics problems. The approach was proven useful also in network reconstruction scenarios, where its employment leads to reconstruct the essential features of pore geometry and connectivity of chalk (Talukdar, Torsaeter, Ioannidis, and Howard, 2002).

Simulated annealing involves a function (energy) that needs to be minimized and proceeds by systematically modifying (mutating) the proposed solution. Mutations leading to a decrease of energy are accepted and those leading to its increase are either rejected or accepted. One defines a control parameter (temperature) which specifies (typically in a probabilistic way) which worsening mutations are accepted. Accepting none of them can quickly make the system get stuck in a local minimum, while accepting too many of them leads to excessive wandering without ever settling into any minimum. The level of worsening mutations which it is best to accept strongly depends on the problem in question, with several different strategies to choose from. In general, one wants to start with a high temperature to allow for a big region of the search space to be initially explored, and then gradually decrease the temperature so that the solution eventually settles into a minimum. This is in fact the strategy we will use.

Before starting the annealing, we need a proper definition of the cost function (energy) that is to be minimized. In our case this is the trajectory error Δ_T , since it is the quantity that can be measured for any reconstructed matrix R_{pq} . Hence, the annealing procedure will minimize Δ_T , which we expect to be in general correlated with minimization of the adjacency matrix error Δ_A . We start the annealing by setting randomly each Fourier coefficient for the function g from a log-uniform probability between $[0, 100]$. Then we carry out the following steps:

1. We start by picking randomly one of the Fourier coefficients as defined in Eq. 2.5 (for example, one of the a_m -s), and modify it as follows: $a'_m = a_m + \alpha a_m$, where $\alpha = U/\log(k)$ with k being the current step of the annealing and U random numbers taken from a uniform distribution between $[-1:1]$.
2. Compute the modified trajectory error Δ'_T using the mutated coefficient a'_m and compare it to the old trajectory error Δ_T , obtaining the difference $\Delta H = \Delta'_T - \Delta_T$,
3. Accept the mutation (keep a'_m) according to the probability p defined as (Metropolis et al., 1953):

$$p = \begin{cases} 1 & \text{if } \Delta H \leq 0; \\ e^{-\Delta H/\beta} & \text{if } \Delta H > 0; \end{cases}$$

where $\beta = \beta(k)$ is the temperature, which decreases with annealing steps as specified below,

4. Continue the process at point (1) by randomly mutating another Fourier coefficient in the function g .

Next we aim to establish the most suitable annealing scheme for our reconstruction problem. To that end, we study two annealing configurations as follows.

- (a) The logarithmic temperature decrease

$$\beta_{\log} = \frac{\beta_0}{\ln(k)}$$

with the initial temperature $\beta_0 = 10^{-5}$ (Ingber et al., 1996).

- (b) The exponential temperature decrease

$$\beta_{\exp} = \beta_0 \lambda^{k/10}$$

with $\lambda = 0.99$ and initial temperature $\beta_0 = 10^{-3}$.

In general, scheme (a) gives a very slow temperature decrease which protects against freezing too quickly, but it requires a lot of steps if we want to cover a wider range of temperature. In contrast, scheme (b) gives a much faster convergence to the minimum, but it might miss a better minimum due to fast freezing. In Sec. 2.3.1 we test both schemes, but start from a lower temperature for logarithmic decreasing so that both schemes would eventually freeze after a comparable number of steps.

2.3 Results

To test the performance of our new method and compare it with the one presented in Ref. (Levnajić and Pikovsky, 2014), we employ a simple toy-model dynamical system containing $M = 6$ nodes with the Hansel-Sompolinsky model, a simple model of neuronal dynamics, running on it (Hansel and Sompolinsky, 1993):

$$\dot{x}_p = -x_p + \sum_{q=1}^M A_{qp} \tanh(x_q). \quad (2.8)$$

The adjacency matrix A_{pq} that we wish to reconstruct is weighted, with weights randomly chosen between -10 and 10. Starting from random initial conditions we integrate the system between $t = 0$ to $t = 4$ and extract 20 values of x_p for each node p , equally spaced in time by $\tau = 0.2$. The obtained time series for all nodes are shown in Fig. 2.4, top left panel.

2.3.1 Comparison of two reconstruction schemes

We start by examining the actual difference between the two reconstructing schemes and between the two annealing schemes. To that end we consider the problem formulated above with time series from Fig. 2.4 top left panel and run three reconstruction procedures: (i) the random guessing one [from Ref. (Levnajić and Pikovsky, 2014)], (ii) simulated annealing with logarithmic temperature decrease, and (iii) simulated annealing with exponential temperature decrease. We record the decrease of the trajectory error Δ_T for each reconstruction scheme, and plot it in Fig. 2.2 left panel as function of the number of steps (iterations of the reconstruction scheme). For the case of the random guessing, we are always taking the overall minimum found so far. We have applied an adaptive stopping criteria. This means, that the annealing does not stop at a given number of steps but when the annealing is "frozen" i.e. when the annealing stops accepting any mutation. In our case, this is when during 2000 steps there is no change or the average change is less than an arbitrary but small threshold. This leads the annealing to stop at different numbers of steps. In the right panel in Fig. 2.2 we show the histogram of the final Δ_T values for the three reconstruction schemes, obtained from several runs of each scheme. It is clear from both panels of this figure

that both annealing schemes on average outperform the random guessing scheme. In fact, the maximum observed Δ_T for β_{exp} is smaller than the minimum Δ_T found for random guessing. As for choosing between the annealing schemes, we from now on rely on β_{exp} , since we found systematically better result for that scheme (as illustrated in the left panel in Fig. 2.2).

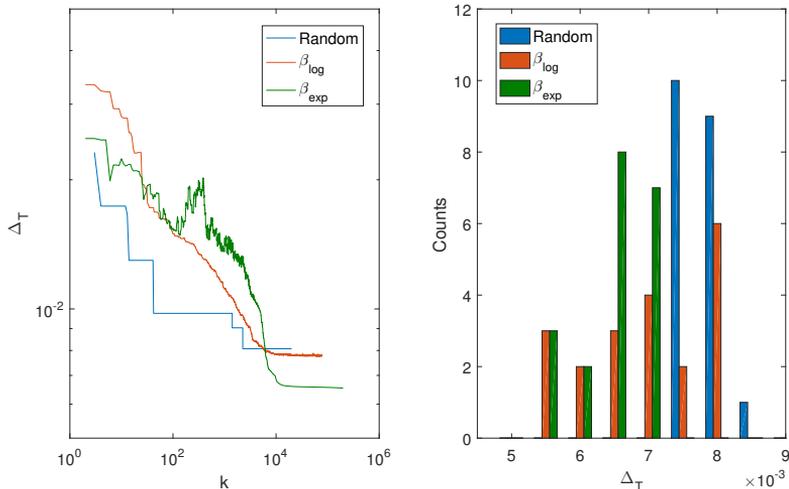


Figure 2.2: Both annealing schemes outperform the random guessing. Left: Trajectory error as function of the annealing steps for two different annealing schemes β_{log} and β_{exp} and random guessing scheme using one realization of each run. Right: Histogram of the final trajectory errors Δ_T for the same three schemes using 20 independent realizations of each scheme.

To further analyse the benefits of annealing, we compare in Fig. 2.3 the scatter plot for random guessing against several realizations of the annealing scheme. Scatter plots are obtained by identifying each reconstruction step via its Δ_A and Δ_T and representing it with a point in 2D, where the x-coordinate is the trajectory error and the y-coordinate is the adjacency matrix error. We can see how decreasing the trajectory error indeed decreases the adjacency matrix error for all schemes. Practically all realizations of the annealing achieve better results for both Δ_T and Δ_A than the random guessing. This is further clarified in the inset in Fig. 2.3. Despite the fact that some realizations converge to similar final precision, we found that each reconstruction comes with its own function g . We also found that there is no clear pattern according to which coefficients

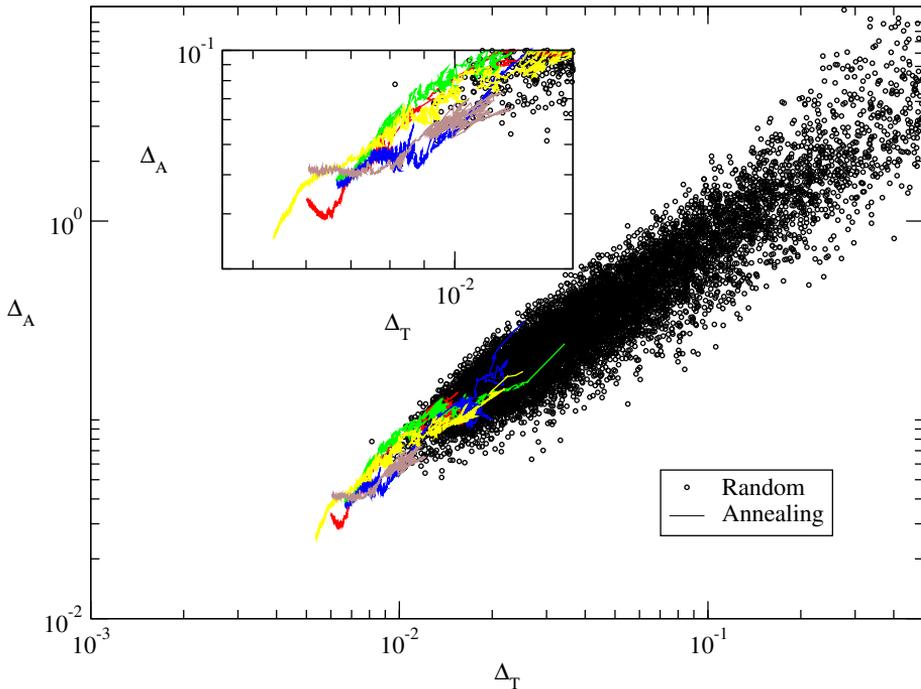


Figure 2.3: We improve the reconstruction by minimizing the trajectory error Δ_T via simulated annealing. Comparison of the original method using random choices of g (black dots) with several realizations of the simulated annealing approach (colored lines). Minimizing Δ_T in general leads to minimizing Δ_A . Inset: Zoom of the final result of the simulated annealing and random guessing.

contribute more to the final precision. This nevertheless confirms that, at least in this example, the annealing reconstruction scheme outperforms the random guessing scheme.

2.3.2 Network reconstruction from different dynamical regimes

To better appreciate the extent of applicability of random guessing versus annealing schemes, we next look at how the two reconstruction schemes perform for data from different dynamical regimes. This illustrates the fact that in real experiments the variability of fields (Pikovsky, 2016) of the available data can vary.

To investigate this further, we generate another set of time series from the same system Eq. (2.8) and show it in top right panel of the

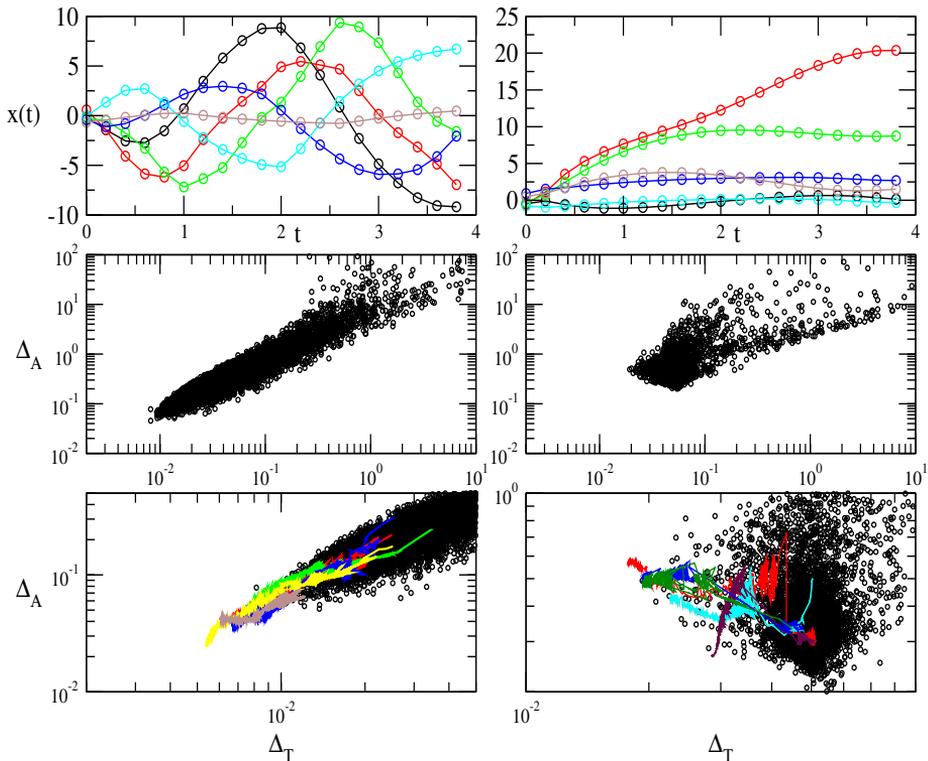


Figure 2.4: The dynamical regime of the system plays a key role in the correlation between Δ_T and Δ_A . Top panels: two sets of time series generated by the system Eq. 2.8 (see text). Middle panels: two scatter plots respectively corresponding to two scatter plots obtained via random guessing (old) method. Bottom panels: evolution plots of reconstruction via simulated annealing (several runs in different colors) compared with the scatter plots from middle panels (black dots).

Fig. 2.4. Clearly, these time series do not display as much variability as the one on the top left panel. Two scatter plots obtained using the original random guessing method are shown in the middle two panels of Fig. 2.4. As expected, the scatter plot in the left gives significantly better precision as it reaches to far lower values of both errors Δ_T and Δ_A . Moreover, this scatter plot exhibits a more visible correlation between the two errors. This indeed confirms that the left panel time series contain more usable information than the right one, thus allowing for more information to be extracted, which consequently leads to better reconstruction precision. Next we test the performance of the simulated

annealing approach (exponentially temperature decrease) and show the respective evolution plots in the bottom two panels of the Fig. 2.4 in comparison with the random guessing scatter plots from middle panels. As already observed, in case of time series on the left panel, simulated annealing indeed improves the reconstruction precision compared to random guessing method. All evolution plots eventually reach smaller final error values. They follow similar paths during the annealing, but ultimately finish at slightly different precisions.

In contrast, in the case of the time series on the right panel, we first observe that the correlation between two errors is far weaker, if present at all. This means that minimizing Δ_T does not in general guarantee minimization of Δ_A . And indeed, annealing does not improve the random guessing result. In fact, evolution plots do not follow similar paths and do not reach final points close to each other. Instead, since the quantity to minimize (energy) was the trajectory error Δ_T , most evolution plots do improve the trajectory error result, but that does not lead to the improvement of the adjacency matrix error result, due to a missing correlation between them. We can conclude that neither the original random guessing method nor the simulated annealing really works with the right panel time series. This suggests that complexity (information content) of the time series plays a major role in the reconstruction precision, which merits further research.

2.3.3 Comparison of reconstruction schemes for varying data quality

At the outset of the discussion so far we assumed that the empirical data (time series) come with a given sampling frequency and number of data points (total length of the time series). Yet for a reconstruction method to be of practical use, it should be implementable in realistic scenarios where due to experimental limitations these parameters might not be adjustable. In the remainder of this section, we investigate the robustness of both reconstruction schemes to these two limitations. For simplicity, we limit ourselves only to the case of time series on the left panel in Fig. 2.4.

First we examine the impact of the sampling frequency τ . In particular, since measurement frequency is becoming less of a problem for modern experimental equipment, we look at reducing τ below the original

value of $\tau = 0.2$. We keep the total number of recorded points fixed to 20. Lowering τ means that the total recording time no longer goes up to $t = 4$, but it is less than that, which means that time series are shorter (in the sense of physical time). This leads to data points covering a smaller portion of the dynamical phase space, which is expected to worsen the reconstruction. On the other hand, shrinking τ makes the derivative estimates better, which is expected to improve the reconstruction. To test this we consider 15 values of τ between 0 and 0.2, and run 20 realizations of each reconstruction scheme for each value of τ .

In Fig. 2.5 we plot the final average values of Δ_T in the left panel and Δ_A in the right panel as a function of τ . The errorbars represent the standard deviation over different realizations.

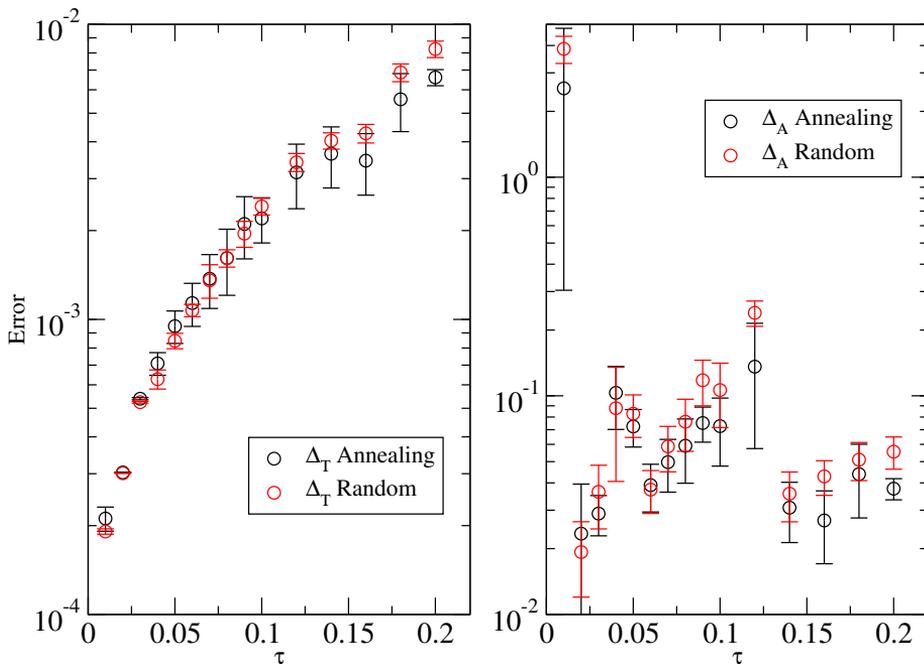


Figure 2.5: An interplay between quality of derivatives and phase space coverage is found for the adjacency matrix error. Average obtained value of the trajectory error (left) and adjacency matrix error (right) as a function of the sampling frequency for the random guessing scheme (red) and the annealing scheme (black). We use the time series of the top-left panel in Fig. 2.4. The errorbars represent the standard deviation of the different realizations obtained over 20 realizations for each scheme and each value of τ .

The trajectory error displays a well pronounced trend to increase

with increasing τ for both schemes. At larger τ annealing shows lower trajectory errors. This is largely expected, since adjusting to trajectory is easier with smaller τ , so there annealing makes no significant improvement, and both schemes in fact perform very well. As for the adjacency matrix error in the right panel, there is no clear trend mirroring the pattern found for the trajectory error. It seems that actually the precision is worse at smallest considered $\tau = 0.01$, then good for $\tau = 0.02$ and slightly worsening until $\tau = 0.12$, then again good for $\tau = 0.14$ and slightly worsening until $\tau = 0.2$. This indicates an interplay between phase space coverage (total length of time series) and quality of derivative estimates (resolution), which makes the reconstruction precision change as observed. For instance, it appears that near $\tau = 0.12$ exploring a wider region of the phase space gradually becomes more important than the derivative estimates. Interestingly, this interplay is found in both reconstruction schemes, although for increasing τ annealing seems to gradually perform better than random guessing.

We also studied the performance of reconstruction schemes by varying the total recording time while keeping the recording frequency fixed to $\tau = 0.2$. This is to say that the number of data points varies, and consequently the phase space coverage also varies, but the derivative estimates are of unchanged quality. In Fig. 2.6 we plot the trajectory and the adjacency matrix errors as function of the number of points, again using the time series in the top left panel of Fig. 2.4. Errorbars again represent the standard deviation over 20 different realizations for both schemes. Both Δ_A and Δ_T for both reconstruction schemes improve with the number of points and arrive at a plateau, where adding more points does not improve the precision significantly. This gives an insight into how the method performance reaches the best precision with a certain phase space coverage. This also has to do with intrinsic 'reconstructability' of the system. When we do not have enough points (until 12), both the annealing and the random evolution do not have enough information to reconstruct the network and both Δ_A and Δ_T are high. At about 12 data points, the method finally gets enough information to reconstruct the network and precision sharply improves, but then quickly saturates. Again, annealing performs slightly better with increasing number of data points. Note that in this case, the pattern observed for Δ_A and Δ_T is very similar.

We conclude by noting that both schemes are in principle applicable in realistic scenarios with scarce data, but the annealing scheme, at least

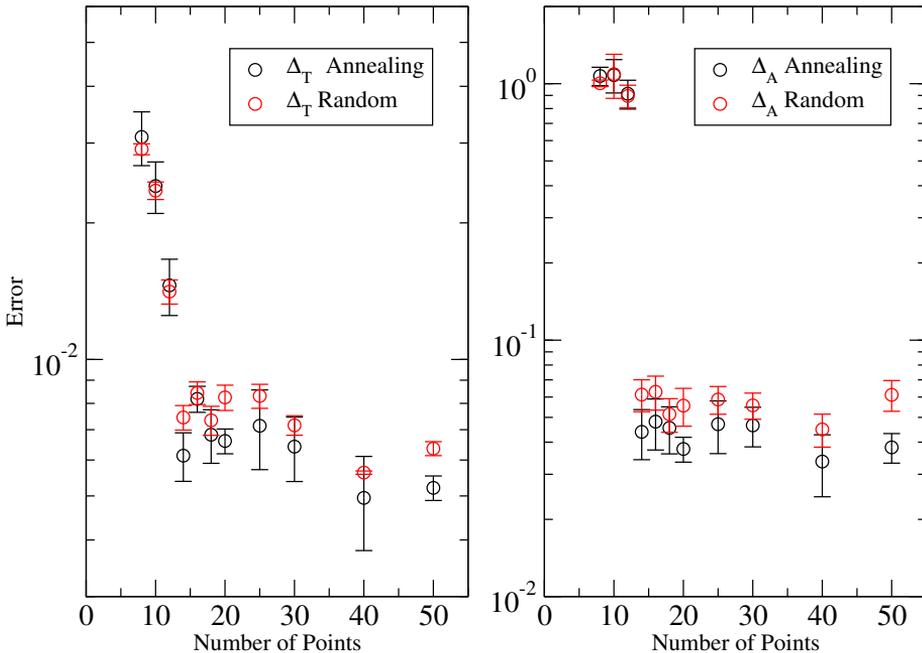


Figure 2.6: We find good reconstruction for as few as 14 data points. Average obtained value of the trajectory error (left) and adjacency matrix error (right) as a function of the number of points keeping $\tau = 0.2$ for the random guessing scheme (red) and the annealing scheme (black). We use the time series of the top left panel in Fig. 2.4. The errorbars represent the standard deviation of the different realizations obtained over 20 realizations for each scheme and each number of points.

with time series considered here, does indeed perform somewhat better.

2.3.4 Influence of noise

Finally we also tested the impact of noise on our method. In Ref. (Levnajić and Pikovsky, 2014), the authors applied a smoothing technique to reduce the impact of noise on the method. Here, to complement the approach of Ref. (Levnajić and Pikovsky, 2014), we study what happens if we minimize the noisy trajectory error. To do so, we use the time series of the left top panel in Fig. 2.3 and add white Gaussian noise with $\sigma^2 = 0.1$. These time series are displayed in the inset of Fig. 2.7. In Fig. 2.7 we also see the scatter plot of the error of the noisy trajectories with the error of the real trajectories. Using the noisy time series as starting point,

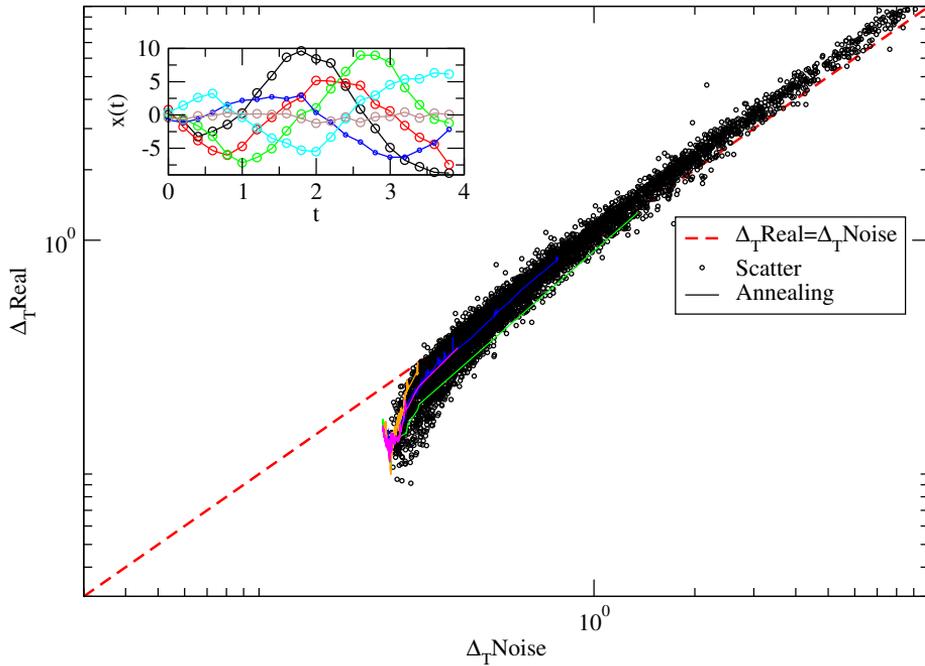


Figure 2.7: Minimizing directly the error of the noisy trajectories gets you closer to the real trajectories than to the noisy ones. Scatter plot of the error of the noisy trajectories with the error of the real trajectories. The lines correspond to several realizations of the annealing. The dotted line correspond to $\Delta_T \text{Real} = \Delta_T \text{Noise}$ relation. **Inset:** Time series from top left panel in Fig. 2.4 with the addition of white Gaussian noise with $\sigma^2 = 0.1$.

the real trajectory error is computed by comparing how the time series generated by the method resembles the noise-free one. When the error is big, there is high correlation (red dotted line). However, once the error of the noisy time series gets smaller, the error of the real trajectories gets smaller at a faster pace. This implies that minimizing directly the error of the noisy trajectories gets you closer to the real trajectories than to the noisy ones. We see this behaviour also for the solutions obtained from simulated annealing (lines in Fig. 2.7). This can be explained by recalling that the method uses smooth basis functions (Fourier series of order 10) to approximate the noisy signal. Therefore, it cannot approximate well the noisy edges in the curves which are caused by the superposition of the signal with white noise. However, the method still is successful to get the overall shape. And this overall shape is exactly the original signal without noise. Therefore, the trajectory error with regard to the original signal becomes smaller than the one with regard to the noisy time series. We are aware that by using Fourier series of higher order, the method could approximate noisy signals better. Therefore, minimizing the trajectory error of the noisy time series will no longer minimize at a faster pace the error of the real time series. That is why some prior knowledge about the scale in which the noise occurs is needed to select the right amount of harmonics.

2.4 Discussion

In this chapter we showed how the evolutionary algorithm of simulated annealing can help to improve network reconstruction methods. Simulated annealing is indeed capable of improving the existing network reconstruction methods. Additionally, we emphasize that the logic of evolutionary algorithms - in our case simulated annealing - can be fruitful for enhancing the existing, or designing new network reconstruction methods, related to the problems in physics and other sciences.

On the other hand, limitations coming from different dynamical regimes seem to pose a considerable problem that requires additional research. As expected, for enough variability of fields in the dynamics the method performs better, as clear from all realizations. This opens the question of how to establish whether a given good reconstruction is also the best theoretically obtainable reconstruction under any method. In other words, how can we know if our method was able to extract all

the information contained in the time series and use it for reconstruction. This amounts to computing the information content of time series, and examining its relation with the final precision. Our investigation of the method performance in more realistic scenarios seems to confirm such a relation, which presents an interesting line of further work.

We also found that different time series attributes, such as measurement frequency or number of points, interplay in yielding the final precision, as they contribute to variation in the phase space coverage and derivative estimates. Of course, further adjustments of our method, such as a temperature reduction scheme, can render it more applicable for a given specific scenario. Additionally, we showed that the method can be applied when noise is added to the time series. However, a more detail characterization of the effect of noise is a pending task.

The application of the method to bigger systems is still possible, however since the computation time grows with the system size, some optimization of the computing algorithm has to be made to go to bigger networks. Nevertheless, the application to realistic data sets requires the knowledge of the interaction functions f and h . This has just been recently generalized in Ref. (Pikovsky, 2016) where the specific knowledge of the functions details are not needed. Still, some a priori knowledge of the structure of the system is needed. So a method which uses small data sets and has as minimal assumptions as possible is an open problem.

Inferring directed networks using a rank based connectivity measure

3.1 Introduction

In Chap. 2 we studied a model-based reconstruction method. There, using the knowledge of the equation of motion along with the signals of each node, we were able to infer the adjacency matrix of the system. In this chapter, instead, we address the inference problem using a data-driven approach. For this purpose, we use the state space measure L defined in Ref. (Chicharro and Andrzejak, 2009). This measure belongs to a family of measures which aim to detect the direction of couplings between dynamics from pairs of signals by using the asymmetric state similarity criterion (Schiff, So, Chang, Burke, and Sauer, 1996; Arnhold, Grassberger, Lehnertz, and Elger, 1999; Quiroga, Arnhold, and Grassberger, 2000; Quiroga, Kraskov, Kreuz, and Grassberger, 2002; Serra, Serra, and Andrzejak, 2009; Andrzejak and Kreuz, 2011; Sugihara et al., 2012). Applications of L to real-world data include electroencephalographic recordings from epilepsy patients (Andrzejak, Chicharro, Lehnertz, and Mormann, 2011; Andrzejak, Schindler, and Rummel, 2012) and neurophysiological signals (Pereda, Quiroga, and Bhattacharya, 2005). In Ref. (Laiou and Andrzejak, 2017) the efficacy of L in the detection of bidirectional couplings was studied, and in Refs. (Andrzejak and Kreuz,

2011; Malvestio, Kreuz, and Andrzejak, 2017) L was used to detect coupling from spike train dynamics. Here, we apply L to multivariate signals extracted from networks by analyzing all pairwise combinations of individual signals, resulting in estimates of the connectivity matrix.

We study the performance of our method using random networks with Lorenz oscillators as the dynamics for each node. We use a diffusive coupling to convey the interaction between nodes (Sec. 3.2.2). In Sec. 3.3, we study the influence of parameters like the coupling strength, and the link density on the network inference. In particular, we show that for a range of these parameters, we are able to infer perfectly the underlying network of the system. We furthermore show that the addition of dynamical noise can in some cases improve on the performance of the method. Finally, we study the effect of the system size in the performance of the method. A paper containing results of this chapter and Chap. 4 is published in the peer-review journal *Physical Review E* (Leguia et al., 2019). The author of this thesis is also the first author of that paper.

3.2 Methods

3.2.1 Rank based connectivity measure L

In the following we describe the measure L (Chicharro and Andrzejak, 2009) which is based on the asymmetric state similarity criterion. Assume that we have two dynamical systems X and Y from which we derive scalar signals x_i and y_i for $i = 1, \dots, N^*$. From these signals, we reconstruct the state space of the dynamics using delay vectors $\mathbf{x}_i = (x_i, x_{i-\tau}, \dots, x_{i-(m-1)\tau})$, $\mathbf{y}_i = (y_i, y_{i-\tau}, \dots, y_{i-(m-1)\tau})$, where m is the embedding dimension and τ is the time delay. The index i now goes from $i = i_o = 1, \dots, N = N^* - (m-1)\tau$, so the delay vector with index i has the scalar value with that same index i as a leading element. For $j = 1, \dots, k$, we denote by $v_{i,j}$ and $w_{i,j}$ the time indices of the k spatially nearest neighbors of \mathbf{x}_i and \mathbf{y}_i , respectively.. Here the Euclidean distance is used as metric, and temporally nearest neighbors are discarded by applying a Theiler window of length T (Theiler, 1990). Accordingly, the time indices must satisfy that $|v_{i,j} - i| > T$ and $|w_{i,j} - i| > T$. We sort the distances between \mathbf{x}_i and $\mathbf{x}_{i_o \neq i}$, and we introduce g_{i,i_o} , which stands for the rank of the distance between \mathbf{x}_i and \mathbf{x}_{i_o} in an ascending list. To quantify the interdependence from $X \rightarrow Y$ we use $L(X|Y)$ defined in

the following way. We take the X dynamics as reference, and define the Y -conditioned mean rank as: $G_i^k(X|Y) = \frac{1}{k} \sum_{j=1}^k g_{i,w_{i,j}}$. We then introduce:

$$L(X|Y) = \frac{1}{N} \sum_{i=1}^N \frac{G_i(X) - G_i^k(X|Y)}{G_i(X) - G_i^k(X)}, \quad (3.1)$$

where $G_i(X) = \frac{N}{2}$ and $G_i^k(X) = \frac{k+1}{2}$ correspond to the mean and minimal mean rank of the distances in the X signal, respectively. To compute $L(Y|X)$, which quantifies the interdependence from $Y \rightarrow X$, we just exchange the roles of X and Y . For identically synchronized dynamics ($X = Y$), we get $L(X|Y) = 1$. In contrast, for independent dynamics the expected value of $L(Y|X)$ is distributed around zero. The number of nearest neighbors is set to $k = 5$, and we use $m = 5$ and $\tau = 5$ for the embedding dimension and the time delay, respectively. Finally, for the Theiler window we use $T = 15$. These values are taken from Ref. (Laiou and Andrzejak, 2017) to avoid any in-sample parameter optimization.

3.2.2 Model

We apply the measure L defined in Sec. 3.2.1 to directed networks with a dynamical system at each of the M network nodes. We use random networks, which are defined by the link probability ρ (Sec. 1.1.2). By construction, we exclude self-loops, i.e. nodes are never connected with themselves. Links between nodes are expressed by the binary adjacency matrix A as defined in Sec. 1.1. As we commented above, our aim in this chapter is to infer the matrix A using only the information of the scalar signals derived from each of the M nodes.

We consider directed networks of size $M = 16$ with Lorenz dynamics (Lorenz, 1963) at each node. The nodes are connected via a diffusive coupling with strength ε through the x components:

$$\begin{aligned} \dot{x}_p(t) &= 10(y_p(t) - x_p(t)) + \varepsilon \sum_{q=1}^M A_{qp} [x_q(t) - x_p(t)] + \xi_p^x(t), \\ \dot{y}_p(t) &= x_p(t)[b_p - z_p(t)] - y_p(t) + \xi_p^y(t), \\ \dot{z}_p(t) &= x_p(t)y_p(t) - \frac{8}{3}z_p(t) + \xi_p^z(t). \end{aligned} \quad (3.2)$$

We study two different settings. A heterogeneous system by taking b_p from a uniform distribution in the interval $[28, 48]$ or a homogeneous

system by setting b_p to a constant value of 28. In this setting, the model is chaotic at zero coupling strength. The quantity $\xi(t)$ stands for independent Gaussian noise with zero mean and with the following correlation function $\langle \xi_p^l(t)\xi_q^{l'}(t') \rangle = 2D\delta_{ll'}\delta_{pq}\delta(t-t')$ with l and l' being the three components of the Lorenz dynamics. For a given link probability ρ and coupling strength ε , we draw $S = 10$ independent random realizations of the adjacency matrix A . For each of these realizations of the network structure, we generate $R = 200$ realizations of the dynamics. For each realization of the dynamics, we start the system at random initial conditions and use a fourth-order Runge-Kutta scheme to integrate the dynamics. We use an integration step of 0.05 time units and downsample by a factor of 6. We discard initial transients, and extract the x component of the Lorenz dynamics for each of the network nodes. We use 4096 data points, which corresponds to approximately 200 oscillation periods.

For a given adjacency matrix A and each of the $R = 200$ realizations of the dynamics, we compute the matrix elements $W_{pq} = L(X_p|X_q)$. This matrix contains the L values for all the possible pairs of signals in the network. In general, the values $L(X_p|X_q)$ and $L(X_q|X_p)$ are not equal, and we therefore get an asymmetric matrix. We skip the diagonal elements which correspond to $L(X_p|X_p) = 1$. We average W across the $R = 200$ dynamical realizations to obtain \bar{W} . Our objective is to reconstruct the true underlying topology A by separating the continuous valued entries of the matrix \bar{W} into two sets. To achieve this separation, we apply a threshold η to obtain a reconstructed binary adjacency matrix R^η :

$$R_{pq}^\eta = \begin{cases} 0 & \text{if } \bar{W}_{pq} \leq \eta \\ 1 & \text{if } \bar{W}_{pq} > \eta \end{cases} \quad (3.3)$$

with $p, q \in (1, \dots, M)$ without considering diagonal elements $p = q$.

3.2.3 Accuracy of the reconstruction

In Sec. 3.2.2, we showed that by thresholding the connectivity matrix \bar{W} , we can obtain a reconstructed adjacency matrix R^η . Assuming that we know the true adjacency matrix A , the possible types of errors that can occur in the reconstruction are shown in the confusion matrix (Table. 3.1).

The confusion matrix not only shows which are the links and non-links that are correctly predicted (a and d respectively), but also the

		A	
		Link	No link
R^η	Pred. link	True positive (a)	False positive (b)
	Pred. no link	False negative (c)	True negative (d)
Total		$a + c$	$b + d$

Table 3.1: Confusion matrix of the comparison between the true adjacency matrix A and the reconstructed one R^η .

number of non-links predicted as links (false positives b) and the number of links predicted as non-links (false negatives c). These two errors are known as type I and type II errors, respectively. The total number of positive classes (links) in the true set A is given by the quantity $(a + c)$. Similarly, the total number of negative classes (non-links) is given by the quantity $(b + d)$. Using the information of the confusion matrix, we compute:

$$\lambda^\eta = \frac{a}{a + c}, \quad \kappa^\eta = \frac{b}{b + d}. \quad (3.4)$$

Where the true positive rate λ^η is the number of correctly detected links divided by the total number of links present in the original adjacency matrix (A). Similarly, the false positive rate κ^η is defined by the number of incorrectly detected links divided by the total number of absent links in A .

We then compute the reconstructed matrices R^η in dependence on the threshold η and determine the resulting values of λ^η and κ^η . By plotting λ^η versus κ^η for all values of η we get the so-called receiver operating characteristic (ROC) curve (Fig. 3.1). The main diagonal in the ROC space (dashed black line in Fig. 3.1), depicts a random classification. Points along this main diagonal represent the expected value of both λ^η and κ^η for a random reconstruction. Curves over (under) the main diagonal shows a better (worse) than random inference of the adjacency matrix A . By using the information of the ROC curve, an optimal threshold can be chosen depending if one prefers to avoid false positives (threshold for low κ^η) or to avoid false negative (threshold for high λ^η). For a perfect classification exists a threshold [point at $(\kappa^\eta = 0, \lambda^\eta = 1)$] that makes $R^\eta = A$. These characteristics are always true for any proportion of links and non-links in the adjacency matrix A .

Moreover, the area under the ROC curve quantifies the accuracy of the classification. This area, which we denote by Ω , attains its maximal

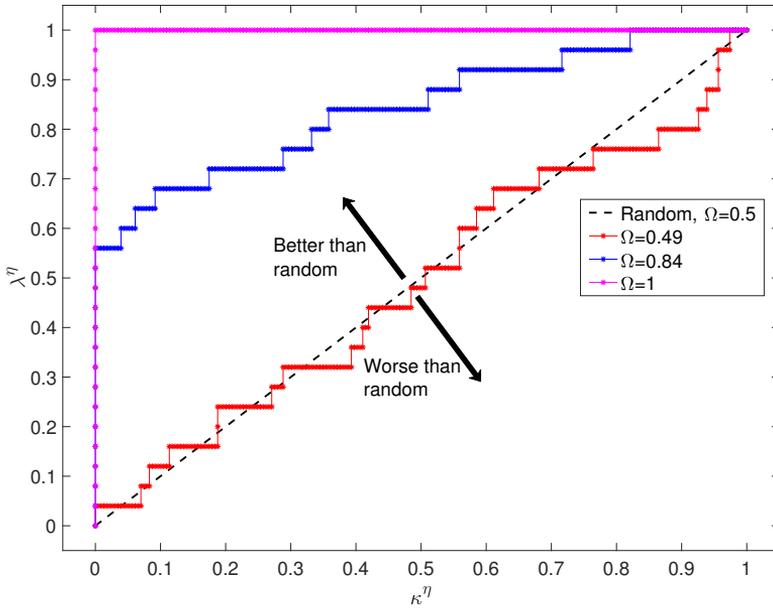


Figure 3.1: Illustration of the ROC space including different ROC curves with three distinct behaviours. We show the area under the ROC curve (Ω) in the legend. The dashed line depicts the expected performance for a random classification.

value of one for a perfect classification. In this case, a threshold η exists for which we get $R^\eta = A$. In contrast, the expected value for a random classification of this area is $\Omega = 0.5$. The measure Ω quantifies how precise the reconstruction is without the need of choosing a threshold η . Hence, we will use it as the measure of accuracy of the reconstruction throughout the chapter.

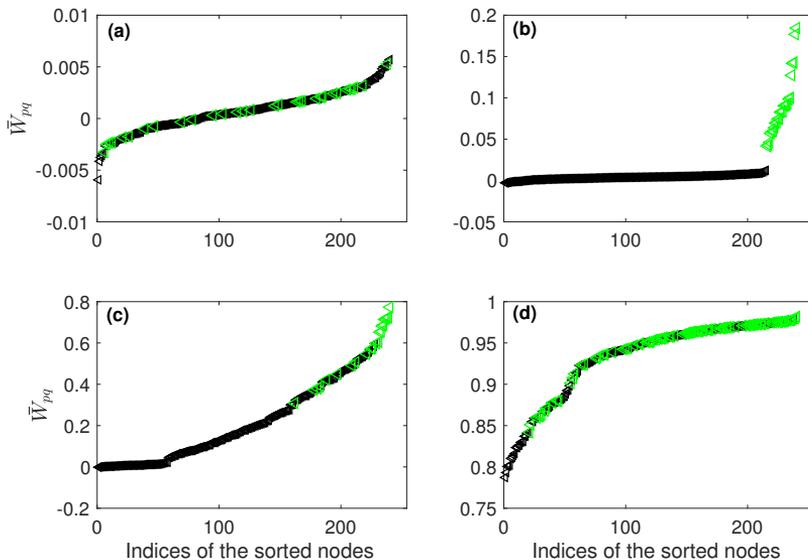


Figure 3.2: Different coupling strengths and link densities lead to different qualities of the network reconstruction. Sorted \bar{W} values for four different exemplary cases (a–d) computed for the heterogeneous noise-free case. For the cases (a–c), we fix $\rho = 0.1$ with $\varepsilon = 0, 0.665$, and 4.92 , respectively. For the case (d), we use $\rho = 0.4$ and $\varepsilon = 4.92$. Values corresponding to links (no links) are marked in green (black).

3.3 Results

We start by analyzing the elements of the matrix \bar{W} for four exemplary cases [Figs. 3.2(a)–3.2(d)]. All examples are computed from the heterogeneous noise-free system of Lorenz oscillators. In Figs. 3.2(a)–3.2(c), we use the coupling strengths $\varepsilon = 0, 0.665$, and 4.92 , respectively, while the link density is set to $\rho = 0.1$. In Fig. 3.2(d), we use $\varepsilon = 4.92$ but with higher link density $\rho = 0.4$. To evaluate \bar{W} , we plot its elements \bar{W}_{pq} in an ascending order (Rubido et al., 2014), in green if there is a link ($A_{pq} = 1$) and black otherwise ($A_{pq} = 0$). For zero coupling [Fig. 3.2(a)], all \bar{W}_{pq} values are around zero and there is no difference between the values where $A_{pq} = 1$ and $A_{pq} = 0$. In contrast, for intermediate coupling [Fig. 3.2(b)], there is a clear gap between results obtained from pairs of signals from connected nodes on the one hand, and results obtained from unconnected nodes on the other hand. Any threshold η taken from within

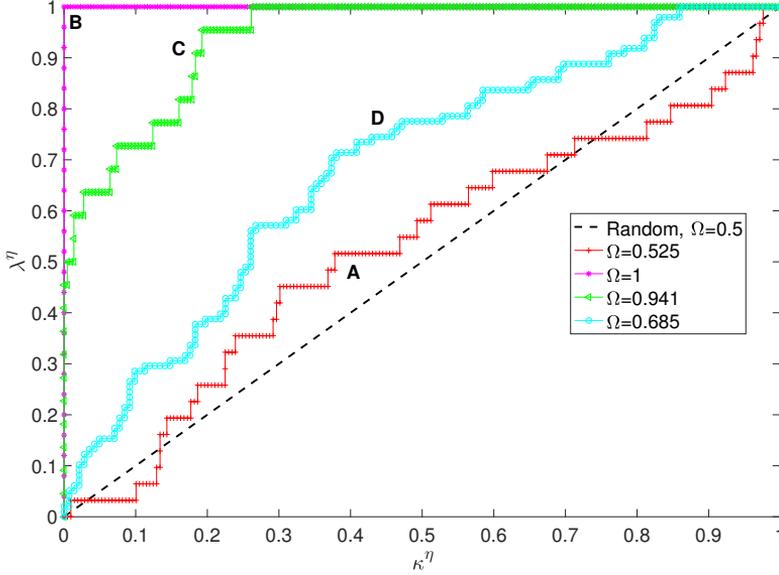


Figure 3.3: ROC curves allow us to quantify the accuracy of the reconstruction for a variable threshold. ROC curves for a heterogeneous noise-free system using different values of ρ and ε . The curves A–D correspond to the cases (a–d) in Fig. 3.2. The dashed black line corresponds to the ROC curve of a random classification. In the legend we show the area under the ROC curve Ω .

this gap will lead to a perfect reconstruction of the adjacency matrix: $R^\eta = A$. For high coupling [Fig. 3.2(c)], the \bar{W}_{pq} generally increase. Due to this increase, values obtained from pairs of signals from connected nodes and values obtained from unconnected nodes mix. Therefore, the gap we had in Fig. 3.2(b) closes. Nonetheless, results for connected nodes still tend to be higher than results for unconnected nodes. Finally, in Fig. 3.2(d), the increased link density paired with the high coupling leads to an almost synchronous temporal evolution for all nodes. As a result, the elements of \bar{W}_{pq} are close to the upper bound of the underlying measure L which is 1. From this display of the results in Fig. 3.2(d), we still see a tendency of connected nodes to be at higher values, but a clear distinction between the values from connected and unconnected nodes is no longer possible. For cases where A is not known, an analysis of the ordered set of \bar{W} could also be performed to help with the threshold decision.

Figure 3.2(b) shows that at intermediate coupling and low link density,

a gap between the sorted \bar{W} values from connected and unconnected nodes appears. Any threshold η placed within this gap, allows us to reconstruct the underlying network perfectly. In contrast, in Figs. 3.2(a), 3.2(c), and 3.2(d), no threshold leads to a perfect reconstruction, and there is no single optimal threshold. To further differentiate these cases, we show in Fig. 3.3 the ROC statistics to quantify the classification performance for these different cases [Figs. 3.2(a)–3.2(d)]. The resulting ROC curves (A–D) and corresponding areas under ROC curves Ω , complement the information shown in Fig. 3.2. At zero coupling (curve A), the ROC curve has a similar shape and area ($\Omega = 0.525$) as expected for a random classification ($\Omega = 0.5$). This correctly reflects that for zero coupling there is no difference between connected and unconnected nodes. For the curve B, we have a perfect ROC curve with $\Omega = 1$. This means that upon increasing the threshold we first find all the links $\lambda^\eta = 1$. Only when increasing the threshold further we would find false positives, i.e., classify pairs of unconnected nodes as links. At high coupling strength and link density $\rho = 0.1$ (curve C), we find some false positives before finding all the links. This reflects the results displayed in Fig. 3.2(c) where the values of connected and unconnected nodes mix. Still, the classification is very high as shown by $\Omega = 0.941$. Finally, for the curve D, the area drops substantially to $\Omega = 0.685$ but remains higher than the random case. The fact that the classification performance remains better than random indicates that for the curve D, although all the points had very similar values in Fig. 3.2(d), the results for connected nodes have a tendency to be higher than the ones for unconnected nodes.

Above we showed detailed results for individual realizations of the heterogeneous noise-free system at exemplary values of the link density ρ and the coupling strength ε . We now study the performance of our method for ranges of these parameters, and averaged over independent realizations of the adjacency matrix A (Fig. 3.4). For zero and small coupling strengths, the reconstruction is similar to a random classification ($\Omega \approx 0.5$), for all ρ . At intermediate couplings ($\varepsilon \approx 0.07$), the performance quickly improves, and for low ρ we even have a region with perfect reconstructions $\Omega = 1$ for all S realizations (green area in Fig. 3.4). This region is surrounded by an area of high performance $\Omega \approx 0.9$ (dark blue area in Fig. 3.4). For small ρ this region of high performance reaches high coupling values. However, as ρ gets bigger, the broadness of this area reduces. For higher link densities and higher couplings the dynamics of the nodes becomes almost synchronous. It is known that

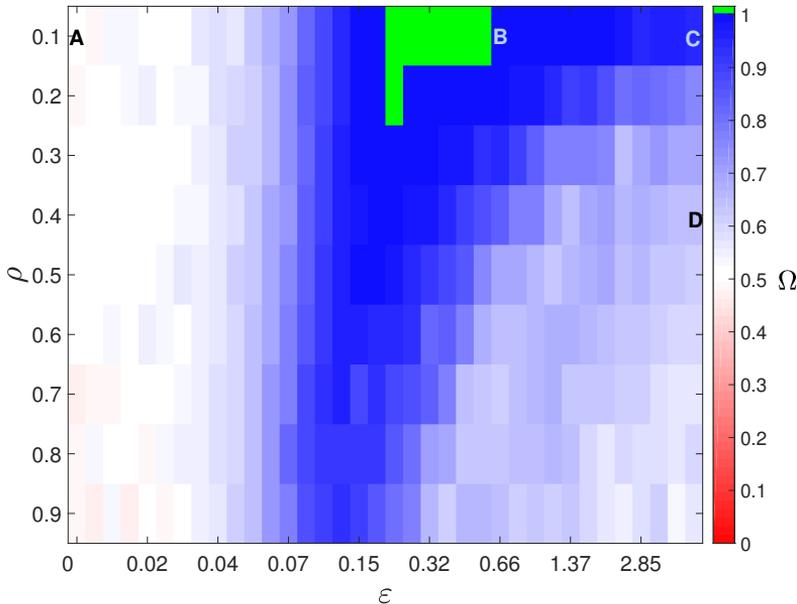


Figure 3.4: Performance of the network reconstruction for the noise-free heterogeneous dynamics is highest for sparse networks with intermediate coupling between nodes. Area under the ROC curve Ω averaged across $S = 10$ independent realizations of the adjacency matrix A in dependence on the link probability ρ and the coupling strength ε . Values of ε are placed equidistantly on a logarithmic scale. Only when $\Omega = 1$ across all the S realizations we plot the corresponding field in green. Labels from A to D mark the set of parameters (ρ, ε) used in both Figs. 3.2 and 3.3.

a reliable extraction of directional couplings from almost synchronous dynamics is not possible (Pikovsky et al., 2001; Smirnov and Andrzejak, 2005; Chicharro and Andrzejak, 2009). For this reason the performance is approaching the chance level when both the link density and coupling strengths are high.

The system that we have analyzed so far is heterogeneous in the sense that the Lorenz oscillators across network nodes are nonidentical in their parameter b_q . We continue to study a noise-free but now homogeneous network by setting $b_q = 28$ for all nodes. The area under the ROC curve Ω averaged across all the S resets of this homogeneous system is displayed in Fig. 3.5. The green region with perfect reconstruction is smaller than the corresponding region in Fig. 3.4. Moreover, the dark

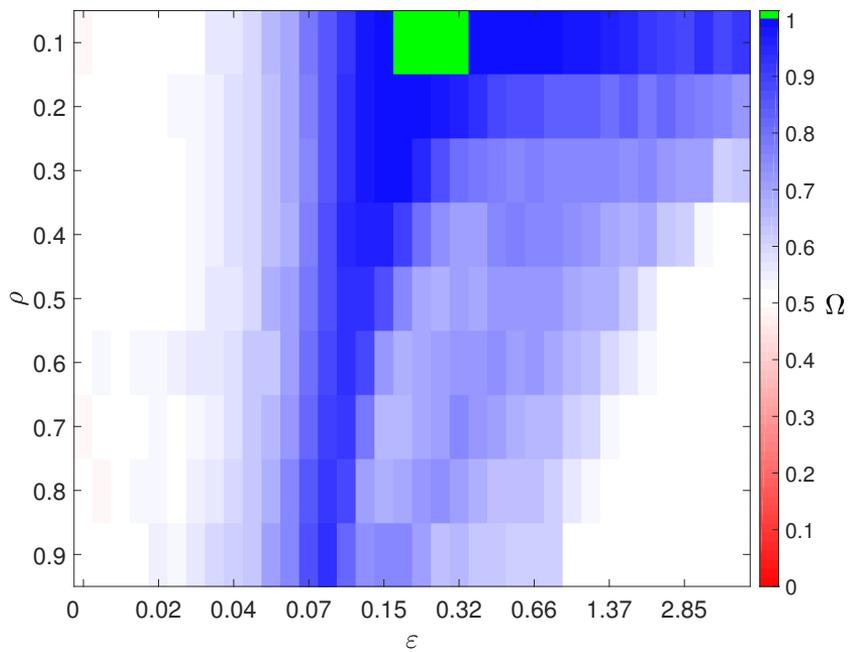


Figure 3.5: Compared to the results for the heterogeneous noise-free system, the performance of the network reconstruction is degraded for the homogeneous noise-free system. Same as Fig. 3.4 but here for the homogeneous noise-free case.

blue area of high performance ($\Omega \approx 0.9$) is also narrower than the one in the heterogeneous system. For high ε and ρ we find an area with exactly $\Omega = 0.5$. In this region, the nodes are identically synchronized and therefore we find that $L(X_p|X_q) = 1$ for all the p, q . While for thresholds $\eta > 1$, we obtain $\lambda^\eta, \kappa^\eta = 0$, we obtain $\lambda^\eta, \kappa^\eta = 1$ for $\eta \leq 1$. Hence, the ROC curve for this configuration coincides with the diagonal, resulting in $\Omega = 0.5$, which reflects that it is impossible to reconstruct a network when all nodes behave the same.

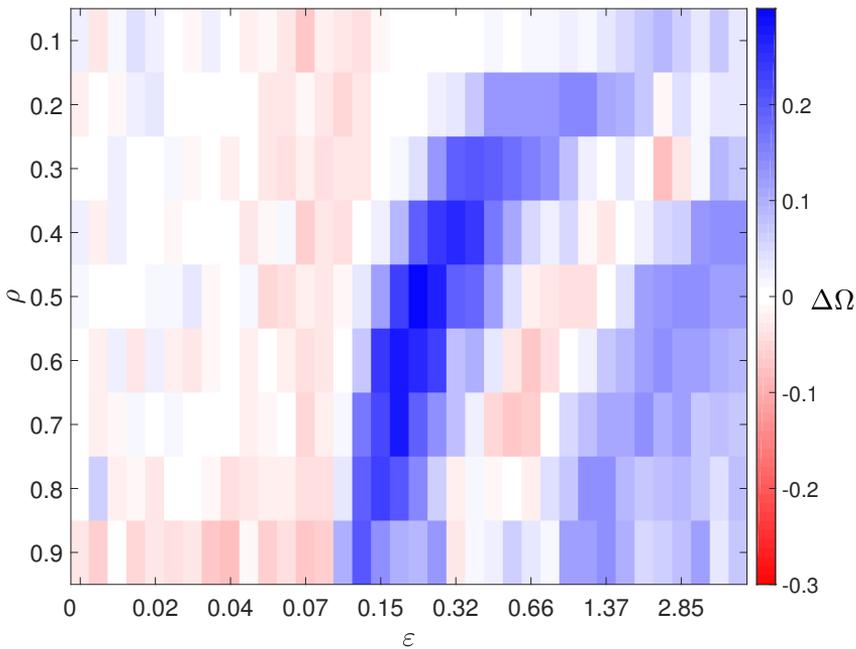


Figure 3.6: For intermediate and high coupling strengths, heterogeneity helps the reconstruction. Difference in performance between the heterogeneous and the homogeneous systems $\Delta\Omega$ shown in Figs. 3.4 and 3.5. Blue and red correspond to higher performance obtained for the heterogeneous and homogeneous system, respectively.

To further compare the results of our approach obtained for heterogeneous and homogeneous systems we inspect the differences between their reconstruction performances $\Delta\Omega$ (Fig. 3.6). For low couplings $\varepsilon \leq 0.07$, the method performs better for the homogeneous system. This shows that small couplings have stronger effects between identical dynamics as

compared to nonidentical dynamics. As a result, our method can detect them better. In contrast, when the coupling strength is increased, heterogeneity facilitates the reconstruction. This results in an area of positive $\Delta\Omega$ (dark blue) in the center of the parameter space of ε and ρ depicted in Fig. 3.6. Here, the coupling strength is high enough for the homogeneous system to substantially increase the overall interdependence between signals, making it more difficult to distinguish real links from indirect connections. In contrast, for the heterogeneous case, we need higher coupling to find this effect. This results in a dark blue area, indicating that the method performs much better for heterogeneous system than for the homogeneous one. For parameters located right from the central dark blue area, some substantial changes of our dynamics have to be mentioned for the interpretation of our results. For some realizations of the adjacency matrix and some initial conditions, the dynamics of a subset of the nodes degenerates from the original complex two-wing Lorenz attractor to much simpler structures and even fixed point solutions. The presence of such degenerated solutions has a negative impact on the average performance. This effect takes place at different regions of the parameter space for the homogeneous versus the heterogeneous system. This disparity is the main reason why right from the central dark blue region we find a red region for which the homogeneous system results in a better performance. Finally, for high couplings and high link density, no degeneration takes place. The performance is again better in the heterogeneous system leading to a triangular shaped blue area in the lower-right of Fig. 3.6. This again is because homogeneous systems are easier to synchronize than heterogeneous ones, and in consequence, it is more difficult to infer connectivity in the former type of systems.

So far we analyzed the system in Eq. 3.2 keeping the noise level at zero. We now study the homogeneous system with a variable degree of dynamical noise. The averaged performance Ω as a function of the noise for three different characteristic ε keeping $\rho = 0.5$ is plotted in Fig. 3.7. At low coupling strength ($\varepsilon = 0.017$), the addition of noise does not change the performance, and we have a random classification for all values of the noise. For $\varepsilon = 0.107$ the performance is high for the noise-free case and decreases with increasing noise level. If we further increase the coupling strength ($\varepsilon = 0.266$), the increased coupling makes the dynamics of each node more alike and the overall interdependence in the dynamics is increased. Here, moderate noise reduces this dynamical interdependence of the signals and therefore, it is easier to infer the

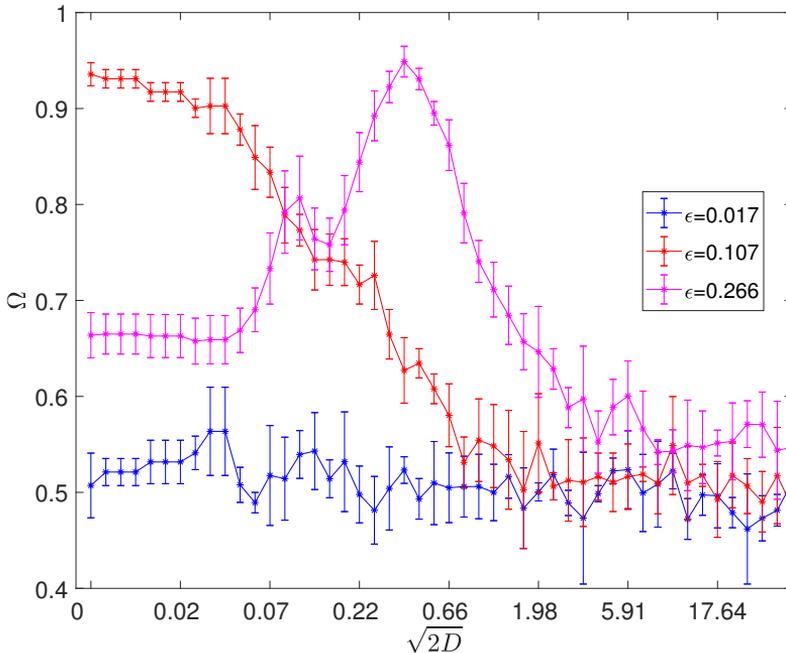


Figure 3.7: Depending of the choice of ε and ρ , there exists an ideal level of noise which improves the reconstruction as compared to the homogeneous noise-free case. Average performance of the reconstruction for $S = 10$ adjacency matrix realizations as a function of the noise level for some characteristic values of ε keeping $\rho = 0.5$. The error bars represent the standard deviation across the five different realizations.

connectivity of the system. This effect is strongest for an optimal value of the noise $\sqrt{2D} \approx 0.38$ for which the performance Ω achieves its maximum value. When the noise is increased further, the performance approaches the one of a random classification.

As shown in Fig. 3.7, the noise level leading to the highest performance depends on the parameter ε . In Fig. 3.8, we show the averaged performance of a fixed noise level of $\sqrt{2D} = 0.5$ in dependence on ε and ρ . For small coupling strengths, we still perform like a random classification. As the coupling strength increases ($\varepsilon \approx 0.07$), the noise hinders our approach to detect the connectivity and we continue to perform almost at random while in the noise-free case we have good performance at these parameter values. In contrast, once the coupling is high enough ($\varepsilon \approx 0.15$), dynamical noise results in increased Ω values and we find

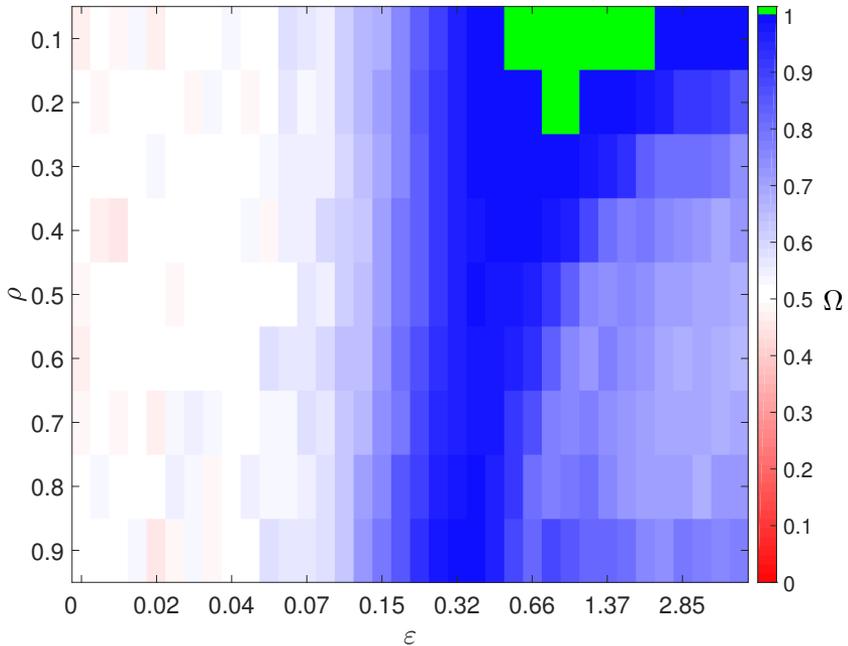


Figure 3.8: Dynamical noise increases the performance for a wide range of parameters. Same as Fig. 3.4 but here for the homogeneous case with noise level of $\sqrt{2D} = 0.5$.

larger regions with perfect reconstruction (green) and high performance (dark blue) than for the noise-free case. Moreover, at very high coupling we perform better than chance level in a parameter region where the classification for the noise-free case is completely random.

Finally, we study the performance of our method for increased values of the network size M keeping the length of the signal and the number of dynamical realizations R fixed. We study the noise-free heterogeneous case keeping $\rho = 0.1$ and for four different values of the coupling strength ε . The averaged performance for different network sizes M , is shown in Fig. 3.9. Due to the computational cost at increased network sizes, we limit the adjacency matrix realizations to $S = 5$. For sizes $M \leq 32$, the performance barely changes with the increased size. We notice that at $M = 32$ we do not perform perfectly for all the realizations but we still have a high Ω across the different coupling strengths. At $M = 64$, we still have very good performance for almost all the coupling strengths

considered. However, the performance decreases and approaches chance level upon further increase of M . This is specially true for the higher coupling strengths. To interpret this deterioration of the performance, one has to keep in mind that the number of possible links grows quadratically with the size M . This way, although we keep $\rho = 0.1$, the average number of links at any given node is much larger at $M = 512$ than at $M = 16$, which in turn makes the system more synchronizable. However, as we said in Chap. 1, most real-world networks tend to be sparse (Newman, 2007). So as the system size grows, the quantity that is conserved should be the link per node and not the density. To further illustrate this point,

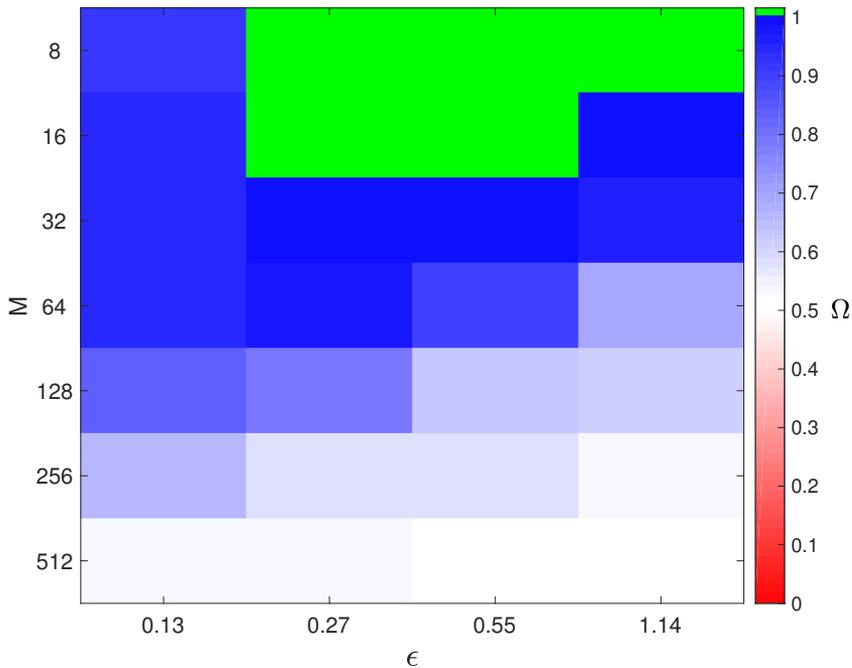


Figure 3.9: Performance of the network reconstruction is decreased for increasing network size M when the link density ρ is not adjusted. Area under the ROC curve Ω averaged across $S = 5$ independent realizations of the adjacency matrix A in dependence on the network size M and the coupling strength ε . Values of ε are placed equidistantly on a logarithmic scale. We use a density $\rho = 0.1$ for all the cases. Only when $\Omega = 1$ across all the S realizations, we plot the corresponding field in green.

we take as a reference quantities ($M_0 = 16$, $\rho_0 = 0.1$) and define:

$$\rho_M = \rho_0 \frac{M_0 - 1}{M - 1}, \quad (3.5)$$

where ρ_M is link density at size M . The measure ρ_M preserves the average links per node that we had at the one of the reference point ($M_0 = 16$, $\rho_0 = 0.1$). Similar to Fig. 3.9, in Fig. 3.10 we show the average performance for increased system sizes for some characteristic ε 's but using ρ_M for the link density. For all the ε , we observe that the performance does not change substantially with increased system size. At $\varepsilon = 0.27, 0.55$, the only difference as the system size increases is that we do not perform perfectly for all the realizations computed. Still, the average performance in these areas is almost perfect ($\Omega > 0.99$). These results confirm that the decrease in performance we observe in Fig. 3.9 is due to this increase of links which also make the system easier to synchronize.

3.4 Discussion

The aim of this study was to infer directed networks using solely the information of the signals of the nodes. To do so, we have presented a reconstruction method based on the state space measure L . We showed that for a subrange of the coupling strength and the link density, we were able to perfectly infer directed networks. This finding provided us with evidence that L can be an effective measure to evaluate directed connections from multivariate signals and could potentially be used in a wide variety of problems (Sugihara et al., 2012; Oates et al., 2014; Trejo Banos et al., 2015; Tirabassi et al., 2017). Other studies that addressed the same problem, typically dealt with small systems and particular topologies (Winterhalder et al., 2007; Schelter et al., 2009; Jachan et al., 2009; Sommerlade et al., 2009; Kugiumtzis, 2013b) or analyzed the resulting structural network properties (Koutlis and Kugiumtzis, 2016). Here, instead, we were able to infer directed networks without choosing specific topology configurations. We also showed that dynamical noise can be beneficial for the reconstruction. This was true in particular for strongly coupled systems where noise prevents the system to enter into a synchronous state. We furthermore studied the performance of the method for increased system sizes. We showed that the performance remains constant if we properly adjust the link density.

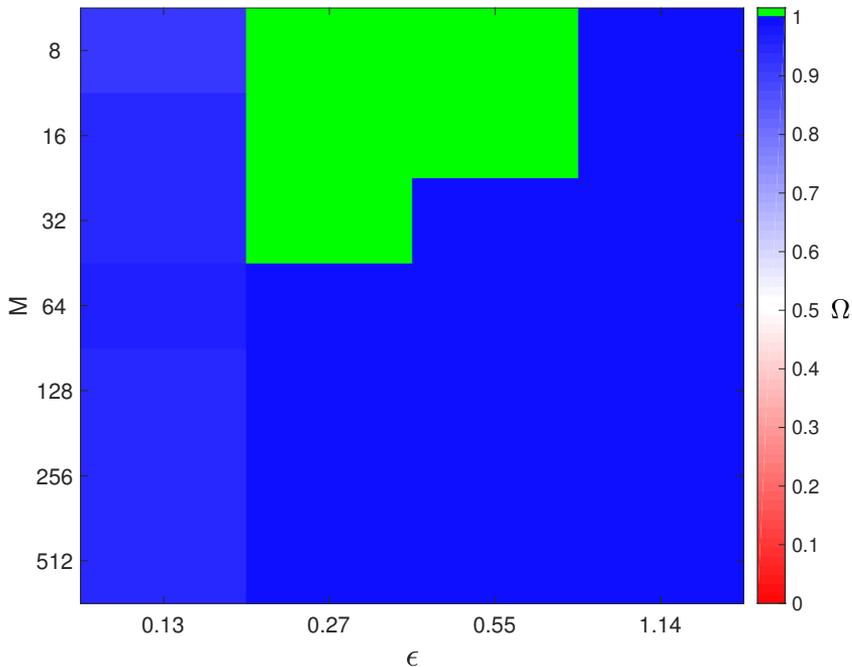


Figure 3.10: Performance of the network reconstruction for the noise-free heterogeneous dynamics does not change substantially if the system size grows as long as the link density is properly adjusted. Same as described in the caption of Fig. 3.9 but using ρ_M .

Although pairwise approaches are prone to mistakenly detect indirect spurious connections, L is performing well for a wide range of parameters. This good performance is specially noteworthy for high link density and high couplings at which the overall network dynamics become highly interdependent. This is because for such highly interdependent dynamics, the distinction between real links and indirect spurious links is difficult for any data-driven approach. To further improve the performance for highly interdependent dynamics, future work should aim at the development of multivariate extensions of the measure L , which are conditioned on subsets of the multivariate signals.

Finally, in the first part of our study, we have used the area under the ROC curve to assess the performance of our reconstruction. This is a good way for testing the performance of our method in different settings when we know the ground truth. However, this procedure cannot be

used when the exact topology is not known, and a decision on how to threshold has to be made. Setting the threshold at a low (high) value, we will likely find false positives (negatives). A possible way to determine an optimal threshold is by plotting the ordered set as in Fig. 3.2. In this case, without the ground truth knowledge of the matrix A , we can of course not add color labels. Nonetheless, in case a clear gap exists in this ordered set [see Fig. 3.2(b)], a threshold can be readily drawn from within this gap. The setting of the threshold will have impact on the resulting network topology (Cecchini, Thiel, Schelter, and Sommerlade, 2018), and the choice of an optimal threshold without previous information of the system is still an open question.

Inferring functional connectivity from EEG data

4.1 Introduction

In Chap. 3 we studied a reconstruction method based on the nonlinear interdependence measure L (Chicharro and Andrzejak, 2009). There, we tested its performance using simulated data from a network of Lorenz oscillators (Leguia et al., 2019). In this chapter, we instead apply the method we described in Chap. 3 to multichannel EEG recordings of seven seizures from an epilepsy patient. In Sec. 4.4, we illustrate the spatial connectivity prior, during, and after the seizure. We furthermore look at the temporal evolution of the connectivity throughout the duration of the recordings.

4.2 Functional brain connectivity

In Chap. 2 and 3 we aimed at the reconstruction of the adjacency matrix of the system we were studying. That is, the structure composed by links between the nodes of the studied complex system. Here, since we are applying our method to brain dynamics, we need to distinguish between functional and physical connectivity. We consider functional

connectivity, so we are not directly aiming at the inference of the physical connections (physical connectivity) between neurons but instead to reconstruct temporal interrelations of EEG activity between signals of different brain areas (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010; Friston, Moran, and Seth, 2013; Hassan, Dufor, Merlet, Berrou, and Wendling, 2014). Accordingly, the connections between signals taken from different electrodes represent correlations related to the activity of the brain and not real physical anatomical connections. In this chapter, we study the reconstruction of functional brain networks using the method in Chap. 3. Specifically, our approach aims at inferring directed functional brain networks that reflect causal relations between brain regions (Van Mierlo et al., 2013; Lehnertz and Dickten, 2015).

4.3 Dataset and Preprocessing

The EEG dataset that we analyze is composed of recordings of seven seizures from an epilepsy patient [patient 1 in Refs. (Vila-Vidal et al., 2017; Tauste Campo, Principe, Ley, Rocamora, and Deco, 2018)]. These recordings were performed prior to and independently from our study in the context of this patient’s presurgical diagnostics in Hospital del Mar, Barcelona, Spain. Five electrodes with a total of 56 recording channels were implanted intracranially in the following brain regions: frontal (FR), amygdala (AM), anterior hippocampus (AH), posterior hippocampus (PH), and temporobasal cortex (TB). The patient was diagnosed with a right-sided nonlesional mesial temporal lobe epilepsy and underwent epilepsy surgery resulting in complete seizure freedom with a follow-up of more than 5 years.

Our dataset contained EEG recordings from seven seizures. The EEG recording for each seizure started 60 s before the seizure onset and stopped 60 s after the seizure ended. The median duration of the recordings was 295 s (range 238 – 465 s). We refer to the period before, during and after the seizure as preictal, ictal and postictal, respectively. The recording channels from which the first signs of ictal activity and early propagation were recorded, the so-called seizure onset zone (SOZ), as well as the exact time of the beginning and end of the seizure were determined by a board-certified neurophysiologist (Rodrigo Rocamora). The data was recorded at 500 Hz, and we applied a high-pass filter with a cutoff frequency of 0.5 Hz and a low-pass filter at 40 Hz using

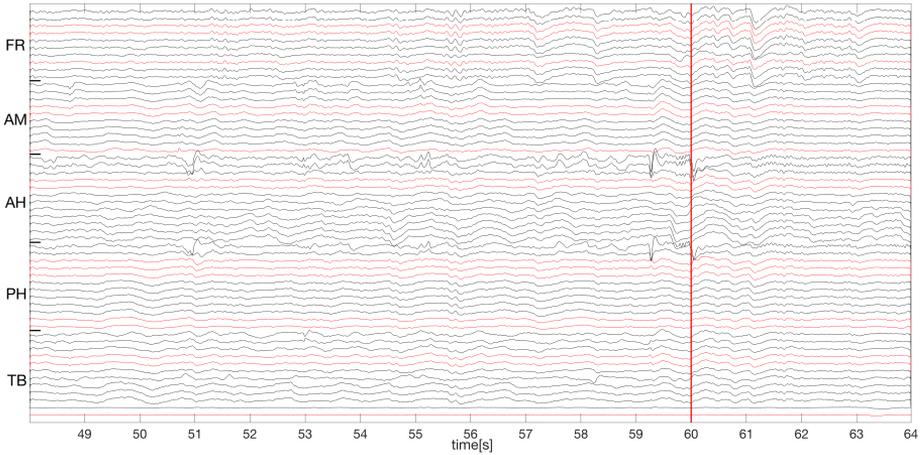


Figure 4.1: Temporal evolution of the EEG recordings from the seizure number three during 16 seconds. Labels in the ordinate axis illustrate the regions of the brain from which the signals were recorded. Signals plotted in red depict the channels that were discarded because they were located in the white matter of the brain. The vertical red line marks the start of the seizure. The total number of initial channels is 56.

fourth-order Butterworth filters. We discarded a total of 16 channels because they were located in the white matter of the brain (red lines in Fig. 4.1). Using all 40 remaining channels we made a bipolar reference between contacts that were neighbouring at the same electrodes, resulting in $M = 35$ signals (Fig. 4.2). We down-sampled the signals by a factor of two to a frequency of 250 Hz.

4.4 Functional connectivity analysis using L

The signals were analysed using a moving window technique with a window length of 16 s and an increment of one second between subsequent windows (93.75 % overlap). The total number of windows Z depended on the length of the EEG recording, which in turn depended on the duration of the corresponding seizure.

For each seizure, we computed L for all the pairs of signals and all windows with index $z = 1, \dots, Z$ leading to a connectivity matrix W^z for each window. As stated above, the windows belong either to the preictal, ictal or postictal period. We depict some of these connectivity matrices from seizure three in Fig. 4.3 and 4.4. Both Fig. 4.3 and 4.4

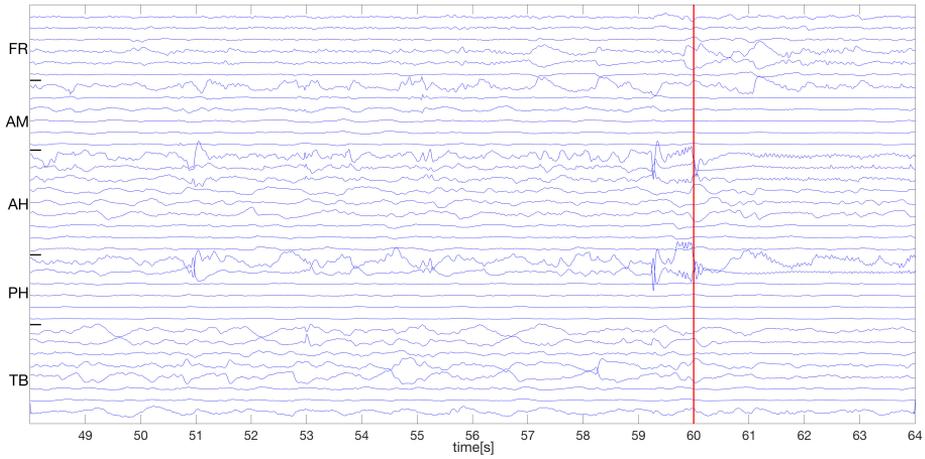


Figure 4.2: Same as Fig. 4.1 but discarding the red signals of Fig. 4.1 and making a bipolar montage. The number of channels is now 35.

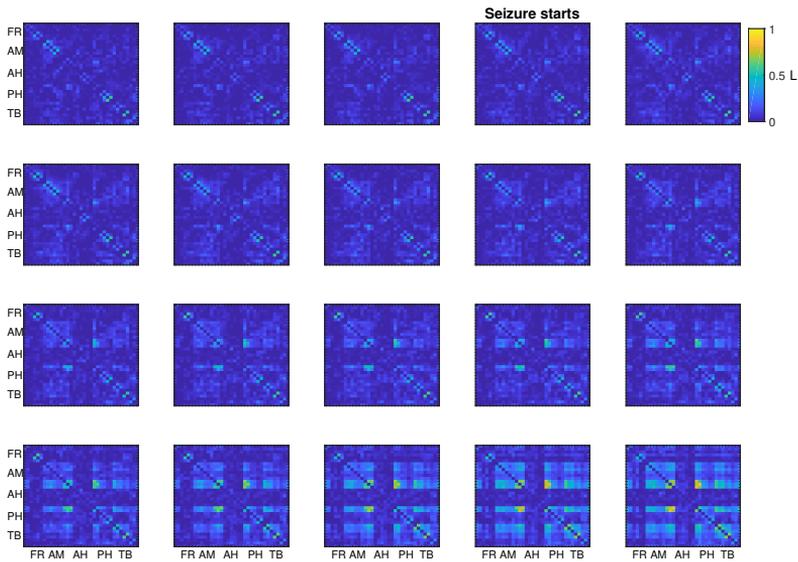


Figure 4.3: Connectivity matrices computed from seizure three from windows corresponding to the beginning of the seizure. Labels in the connectivity matrices depict the regions of the brain from which the signals were recorded. The matrices are computed from 20 consecutive windows. Time ordering in the connectivity matrices goes from left to right and from top to bottom.

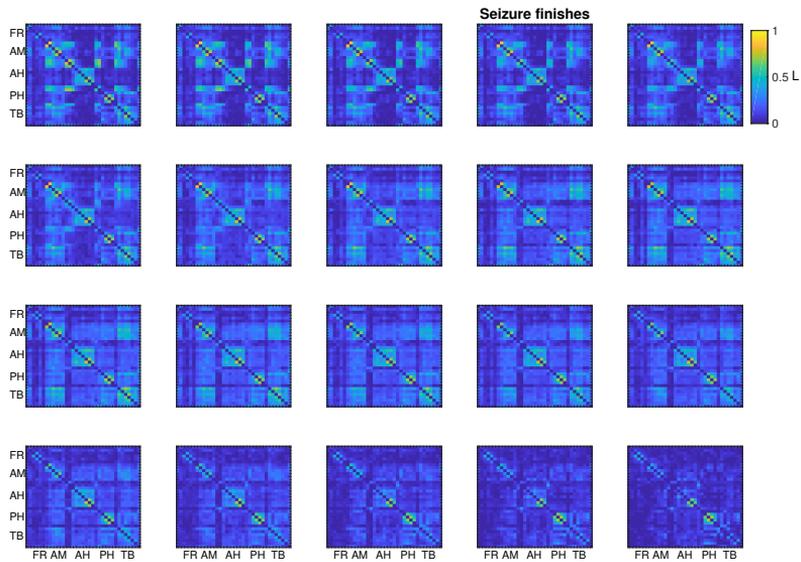


Figure 4.4: Same as Fig. 4.3 but for the end of seizure three.

show the continuous valued connectivity matrices W^z during the start and the end of the seizure, respectively. We observe that the change in connectivity from subsequent windows is difficult to appreciate since these matrices are computed from windows with high overlap. As the seizure starts (ends) the overall connectivity starts to grow (decrease). The first 15 matrices since the seizure starts (ends), show the results from two distinct ictal periods.

As the second step of evaluation, we averaged the connectivity matrices separately for each period resulting in $\langle W \rangle_{\text{period}}$. In these averages, we discarded all the windows that included either the beginning or the end of the seizure and therefore contained activity from two distinct periods.

Recall that in the analysis of the model systems presented in Chap. 3, we used a threshold η to convert our continuous-valued result matrix W to a binary adjacency matrix R^η . In this analysis of real-world data, we aim not only at detecting the existence, but also to estimate the strength of the directed connection between nodes. Accordingly, in Fig. 4.5 we display continuous-valued matrices $\langle W \rangle_{\text{period}}$ of the temporal means obtained separately for each period of each of an exemplary seizure. Values below the mean plus one standard deviation taken separately

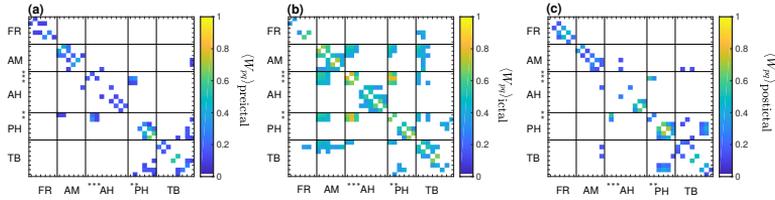


Figure 4.5: High variability of connectivity structure is found across the different periods included in the EEG recording from seizure 3. Temporal averages $\langle W \rangle_{\text{period}}$ for the third seizure for the periods preictal (a), ictal (b) and postictal (c). Values below the mean plus one standard deviation taken separately across the elements of the matrix for three periods are displayed in white. Black lines delimit the different regions of the brain where the electrodes were placed. Contact names placed in the SOZ are marked with an asterisk.

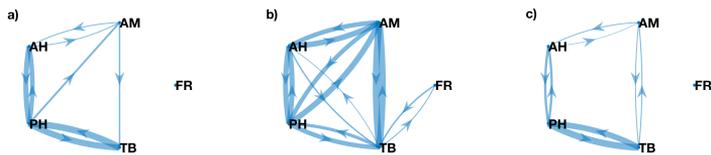


Figure 4.6: High variability of the topological structure is found across the different periods included in the EEG recordings from seizure 3. Graph representation of the average connectivity per brain region from the matrices depicted in Fig. 4.5 for the preictal (a), ictal (b) and postictal (c) periods. The thickness of the links are proportional to their strength. For simplicity we do not depict connections between the same brain region.

across the elements of the matrix for the three periods are thresholded and set to zero. The reconstructed matrix $\langle W \rangle_{\text{preictal}}$ shows connections mainly for contacts that are spatially close. We also infer connections between signals recorded from TB, PH and other regions. Moreover, we also observe connections within the SOZ sampled by contacts in AH and PH. In the ictal period, the reconstructed matrix $\langle W \rangle_{\text{ictal}}$ shows an overall stronger interconnectivity between brain regions as compared to $\langle W \rangle_{\text{preictal}}$. Furthermore, some of the strongest connections are found for contacts placed in the SOZ. In contrast, the frontal region remains less connected as assessed by our inference results. This is in agreement with the clinical information that the seizure did not spread to this brain area. Finally, for $\langle W \rangle_{\text{postictal}}$, as in the preictal period, most of the connections are found for spatially close contacts. However, in the postictal period, we infer stronger overall connectivity and more connections between brain regions. These results are not a special case for the particular seizure used as an example, and we found similar patterns regarding the structure of the reconstruction for all the other seizures (See Fig.A.1-A.6 of the Appendix). This consistency of connectivity patterns across seizures is in agreement with the consistency of relative activation patterns that was found for the same patient in Ref. (Vila-Vidal et al., 2017). Additionally, in Fig. 4.6 we show the graph representation of the connectivity matrix of Fig. 4.5. For simplicity, we average all connections within the same brain region and we do not show self-loops. In this graph representation, we observe that the connectivity between brain regions at the beginning and at the end of the seizure is quite low [Fig.4.6(a) and Fig.4.6(c), respectively] whereas during the ictal period [Fig.4.6(b)], the regions where the seizure actually spreads (AM, AH, PH and TB) form a complete graph with strong connectivity between them.

In Fig. 4.5 we saw that the contacts placed in the SOZ had higher connectivity in comparison with the ones placed in the rest of the brain. We can quantitatively compute this effect by averaging the outdegree and the indegree for all the contacts placed in the SOZ ($\langle d^{\text{out}} \rangle_{\text{SOZ}}$, $\langle d^{\text{in}} \rangle_{\text{SOZ}}$) and for the rest of the brain ($\langle d^{\text{out}} \rangle_R$, $\langle d^{\text{in}} \rangle_R$) separately.

The results for $\langle d^{\text{out}} \rangle_{\text{SOZ}}$ and $\langle d^{\text{out}} \rangle_R$ computed for all recordings and ictal periods are displayed in Tab. 4.1. We notice that the degrees $\langle d^{\text{out}} \rangle_{\text{SOZ}}$ and $\langle d^{\text{out}} \rangle_R$ are always higher in the ictal period than in the other two periods. This confirms what we saw in Fig. 4.5 where there was an increase of the overall connectivity during the ictal period. Moreover, we observe that both $\langle d^{\text{out}} \rangle_{\text{SOZ}}$ and $\langle d^{\text{out}} \rangle_R$ tend to be higher in the pos-

Seizure	$\langle d^{out} \rangle_{SOZ}$			$\langle d^{out} \rangle_R$		
	Preictal	Ictal	Postictal	Preictal	Ictal	Postictal
1	0.68	6.64	0.56	0.51	1.60	0.53
2	0.59	3.49	0.59	0.42	1.35	0.73
3	0.48	4.77	0.47	0.44	1.28	0.84
4	0.40	2.51	0.42	0.45	1.13	1.09
5	0.45	4.69	0.74	0.46	1.30	0.90
6	0.53	4.82	0.77	0.52	1.05	0.73
7	0.65	5.04	0.65	0.45	1.46	0.96

Table 4.1: The average outdegree is always higher during the ictal period. Average outdegree computed from all the contacts placed in SOZ ($\langle d^{out} \rangle_{SOZ}$), and average outdegree computed from all the other contacts ($\langle d^{out} \rangle_R$). We compute the outdegree for the three periods of each seizure recording separately.

Seizure	$\langle d^{in} \rangle_{SOZ}$			$\langle d^{in} \rangle_R$		
	Preictal	Ictal	Postictal	Preictal	Ictal	Postictal
1	0.67	6.37	0.55	0.50	1.58	0.52
2	0.59	3.55	0.65	0.43	1.30	0.73
3	0.43	4.17	0.44	0.47	1.33	0.83
4	0.39	2.30	0.42	0.45	1.15	1.08
5	0.47	4.07	0.74	0.45	1.41	0.88
6	0.53	4.21	0.77	0.53	1.19	0.74
7	0.65	4.83	0.62	0.43	1.40	0.96

Table 4.2: The average indegree is always higher during the ictal period. Same as described in the caption of Tab. 4.1 but for the indegree.

tictal period compared to the preictal period. During the preictal period, we do not see many differences between the average outdegree computed from the contacts placed in the SOZ ($\langle d^{out} \rangle_{SOZ}$) compared to the average outdegree computed from the contacts placed in the rest of the brain ($\langle d^{out} \rangle_R$). This is not true during the ictal period, in which we find that $\langle d^{out} \rangle_{SOZ} > \langle d^{out} \rangle_R$ for all the seizure recordings. This is concordance with the observation in Fig. 4.5 of strong directed connections during the ictal period from contacts placed in the SOZ. Alternatively, during the postictal period, $\langle d^{out} \rangle_{SOZ} < \langle d^{out} \rangle_R$ for most seizure recordings.

We now do the same calculations as in Tab. 4.1 but now for the average indegree (Tab. 4.2). Results for $\langle d^{in} \rangle_{SOZ}$ and $\langle d^{in} \rangle_R$ have the

same patterns that we found when we computed $\langle d^{out} \rangle_{SOZ}$ and $\langle d^{out} \rangle_R$. However, during the ictal period, $\langle d^{out} \rangle_{SOZ} > \langle d^{in} \rangle_{SOZ}$ for most of the seizure recordings. This asymmetry which is not observed for the contacts placed in the rest of the brain, shows that the SOZ has a tendency of sending more information than it receives.

In Fig. 4.5-4.6 and Tab. 4.1-4.2 we looked at the spatial structure of the connections by making a temporal average across the windows of the same period. At the top row in Fig. 4.7, instead, we look at the temporal evolution of the connectivity by making an average across all contacts within the same window:

$$\bar{W}^z = \frac{1}{N(N-1)} \sum_{p,q:p \neq q}^M W_{p,q}^z \quad (4.1)$$

For all the seizures, we see that during the preictal period, the overall interdependence of the system is low. When the seizure starts, the quantity \bar{W}^z quickly grows. In particular, for seizures two through six, we see two phases during the ictal period. One with higher interdependence at the beginning of the seizure, and one with smaller interdependence afterwards. In the postictal period, the quantity \bar{W}^z goes down but remains higher than in the preictal period.

We continue by looking at the temporal evolution of the connectivity but now using the directional information provided by the measure L . We separately average across the matrix elements corresponding to the SOZ driving the rest of the contacts ($\bar{W}_{S \rightarrow R}$) on the one hand, and the matrix elements corresponding to the opposite direction on the other hand ($\bar{W}_{R \rightarrow S}$). Contacts placed in the resected area but not in the SOZ are not included in any of the averages. The temporal profile of $\bar{W}_{S \rightarrow R} - \bar{W}_{R \rightarrow S}$ is shown in the bottom row of Fig. 4.7. For all seizures, during both the pre- and postictal periods the connectivity is almost symmetric, and therefore the differences $\bar{W}_{S \rightarrow R} - \bar{W}_{R \rightarrow S}$ remain small. For the ictal period, apart from some variability across seizures, one common pattern is consistent across all seizures. Shortly after the seizure onset, we find a prominent asymmetry in connectivity in the sense that the SOZ is driving the remaining areas.

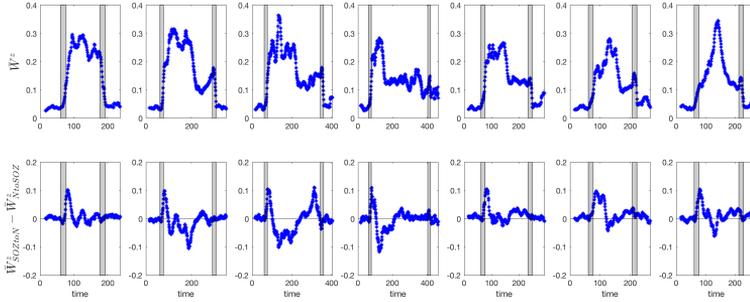


Figure 4.7: A high overall connectivity is found during the seizure with a predominant direction from the SOZ to the remaining brain areas at the beginning of the seizure. **Top row:** Spatially averaged connectivity matrix \bar{W}^z in dependence on time for the seven seizures. **Bottom row:** Spatially mean connectivity of SOZ driving the rest of the contacts ($\bar{W}_{S \rightarrow R}$) minus SOZ being driven ($\bar{W}_{R \rightarrow S}$) by the rest of the contacts for the same seven seizures. The gray frames depict windows that lay between two distinct periods. Left from the left frame is the preictal period, between the frames is the ictal period, and right from the right frame is the postictal period.

4.5 Discussion

Applying our method to analyze EEG recordings from an epilepsy patient, we estimated the functional connectivity for the three seizure periods. This resulted in three distinct structures corresponding to the three periods. We also observed that the strongest connections were the ones between contacts placed in the SOZ. We furthermore computed the temporal mean connectivity and noticed that it reflects the start and the end of the seizures. Finally, we also observed an asymmetrical pattern where the contacts situated in the SOZ were driving the rest of the contacts at the beginning of the seizure. This result is in agreement with other studies that reported driving of the SOZ (Flamm, Graef, Pirker, Baumgartner, and Deistler, 2013; Van Mierlo et al., 2013; Epstein, Adhikari, Gross, Willie, and Dhamala, 2014; Korzeniewska et al., 2014; Lehnertz and Dickten, 2015; Dickten, Porz, Elger, and Lehnertz, 2016). In interpreting these results, we have to keep in mind that the available amount of data was limited, and a study of more patients, seizures and different anatomical locations of the SOZ should be carried out to make any strong claims. Nevertheless, these results show the potential of the method for this type

of real-world data applications (Andrzejak et al., 2011; Andrzejak et al., 2012).

During our application of the reconstruction method of Chap. 3 to EEG recordings we did not know the exact functional connectivity to validate our results. Nevertheless, our findings were coherent with the medical information indicating that our method is extracting meaningful information about the system, and therefore it is a promising approach for medical applications.

CHAPTER 5

Discussion

In this thesis we developed methods that allow us to infer the network structure from dynamical data. We investigated the problem with two different perspectives. Using a model-based approach, we tested if evolutionary algorithms such as simulated annealing can be applied to the inference problem. Moreover, we studied the minimal amount of data needed to reconstruct a given system. Alternatively, using a data-driven approach, we did not make any assumptions about the model of the system and we did not limit ourselves with the amount of data. For measures like cross-correlation and mutual information, it has been shown that it was possible to perfectly infer bidirectional networks (Rubido et al., 2014). However, we studied if it was possible to reliably infer directed networks using only the time series from the dynamics of its nodes.

First (Chap. 2), we studied an extension of a model-based reconstruction method by using simulated annealing (Metropolis et al., 1953). There, we showed that the new method outperformed the previous results of Ref.(Levnajić and Pikovsky, 2014). Moreover, we investigated the robustness of the method for parameters such as data quality or noise level.

In the second part of the thesis (Chap. 3), we used a data-driven approach to infer directed networks. There, by using the state-space measure L (Chicharro and Andrzejak, 2009), we were able to reliably infer directed networks given the right choice of parameters such as

coupling strength and link density. We furthermore saw that under certain conditions, dynamical noise helps the reconstruction.

Finally, after studying network inference using L under controlled conditions, we applied the method to real-world data (Chap. 4). We used multichannel EEG recordings from an epilepsy patient. Our results using EEG data agree with other studies on the same topic and demonstrate the potential of the method for inferring real-world topologies. Full details of the discussion of the results in the studies can be found in Sec. 2.4, 3.4 and 4.5.

In both reconstruction methods (Chap. 2, 3), complete recordings of the signals of all nodes were required. In the first method (Chap. 2), by making strong assumptions about the underlying dynamics, we were able to infer networks with very small amount of data (almost of the order of the system size). However, these strong assumptions made it difficult to apply the method to real-world examples. Alternatively, in the second part of the thesis (Chap. 3), by using a data-driven approach, we did not make any strong assumption regarding the model of the system. In this case, we needed much more data than the method in Chap. 2 to reconstruct the network. Accordingly, a trade-off exists between the minimum amount of data required to infer the network and the assumptions made about such complex system. Using L , we could relax the assumptions on the system and develop a method that can be applied to real data such as EEG recordings from an epilepsy patient (Chap. 4).

The results of this thesis show that evolutionary algorithms can be used in the reconstruction problem. In our work we demonstrated that by using simulated annealing we could improve the results of an existing method (Levnajić and Pikovsky, 2014). Moreover, we showed that we were able to consistently infer directed networks of Lorenz dynamics by using the measure L (Chicharro and Andrzejak, 2009). To the best of our knowledge, it is the first time such networks were inferred without any assumptions on the network configuration. We furthermore studied the applicability of the data-driven method to real-world data by using a multichannel EEG recording from an epilepsy patient. The results of the analysis of real world data were in accordance with the medical information provided by the board-certified neurophysiologist (Rodrigo Rocamora). The validity of our method is supported by the fact that other studies show our same results about the driving action of the SOZ on the rest of the brain (Flamm et al., 2013; Van Mierlo et al., 2013; Epstein et al., 2014; Korzeniewska et al., 2014; Lehnertz and Dickten, 2015;

Dickten et al., 2016).

Contribution to Science

1. We developed a model-based reconstruction method which uses the simulated annealing and studied its performance for different dynamical regimes and data qualities.
2. We developed a data-driven reconstruction method which is able to perfectly infer networks of Lorenz chaotic oscillators.
3. We applied the data-driven method to infer directed functional brain networks from EEG data and show that the results were in coherence with the medical information about the epilepsy patient.

5.1 Future work

In the first part of the study (Chap. 2), we saw that an interplay between the phase space exploration and the quality of the derivatives computed from the data exists. Future work may involve to further study this interplay. Moreover, in the second part of the study, we could consider using surrogates to see how this affects to the reconstruction (Schreiber and Schmitz, 2000; Rummel et al., 2011; Andrzejak et al., 2012). Additionally, we could also evaluate an extension of the measure L . For our data-driven method in Chap. 3, we used a pair-wise approach. However, an extension of our method to a multivariate approach could potentially improve our results. Moreover, in both methods we assumed complete knowledge of the signals of all nodes. Nonetheless, this assumption is not always true and one can think of systems where information about some entities is hidden. Future work may involve studying those systems.

Finally, in this thesis we presented two distinct reconstruction methods. In the first we made strong assumptions but we needed very few data points. In contrast, in the second method we relaxed the assumptions but we demanded long data sets. That is why we started looking whether it was possible to construct a reconstruction method which did not make strong assumptions about the underlying dynamics but at the same time

did not need long data sets. The current work we briefly introduce here is not in a state of publication, and we just introduce its basic ideas. We are using machine learning algorithms such as Random Forest (Breiman, 2001) or RreliefF (Kononenko, Šimec, and Robnik-Šikonja, 1997) to reconstruct networks. Machine learning algorithms had been used in a huge variety of areas in science including the inference of networks (Simidjievski et al., 2018). In our setup, we assume a general complex system:

$$\dot{x}_p = f_p(x_1, x_2 \dots x_M), \quad (5.1)$$

where the f_p are unknown functions. The main idea of the method is to consider that the variables that are more important for describing the dynamics of node p are the ones that are physically connected to node p . To further illustrate the idea let us consider the following function:

$$y = x_1 + \epsilon x_2 + \epsilon^2 x_3, \quad (5.2)$$

where y is the target variable, x_1 , x_2 and x_3 are features and $\epsilon \ll 1$ is a parameter. Based on some learning data examples with combinations of values for y , x_1 , x_2 and x_3 , a hypothetical feature ranking method may assign a ranking score to each feature with the following ascending order: $x_1 > x_2 > x_3$. Similarly, by assigning ranking score to a system like Eq. 5.1 for all nodes we obtain a connectivity matrix W which we can compare with the true adjacency matrix A . Our preliminary results using feature ranking methods to infer networks are promising and we are able to reconstruct networks using a low amount of data. We think that the results will be able to provide further evidence that machine learning algorithms can be used in the framework of the network inference and show good performance.

APPENDIX A

Spatial structure of all seizures

In Sec. 4.4 of Chap. 4 we displayed the spatial structure for seizure three at the preictal, ictal, and postictal periods. In Figs.A.1-A.6 of the appendix, for completeness, we show the mentioned spatial structure for all the seizures our dataset. As mentioned in Chap. 4, we find an overall stronger connectivity during the ictal period. Furthermore, we observe high connectivity for contacts situated in the SOZ.

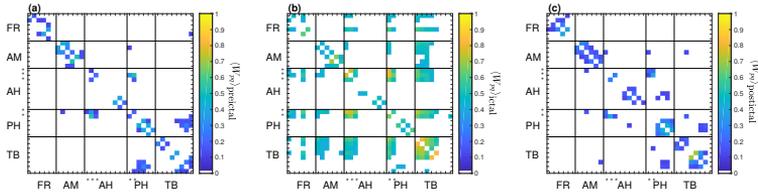


Figure A.1: Same as Fig. 4.5 but for seizure 1.

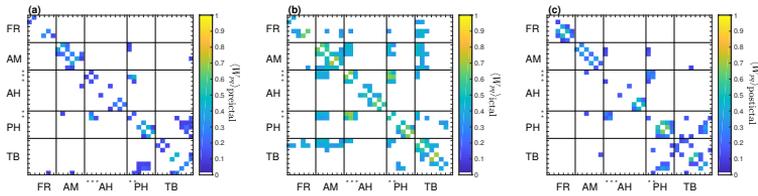


Figure A.2: Same as Fig. 4.5 but for seizure 2.

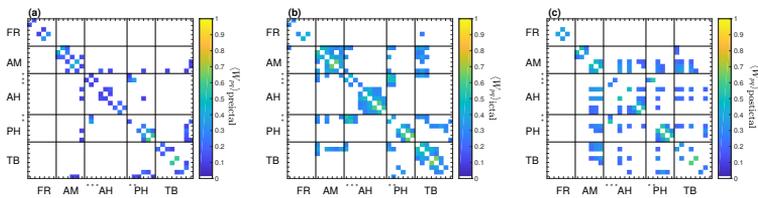


Figure A.3: Same as Fig. 4.5 but for seizure 4.

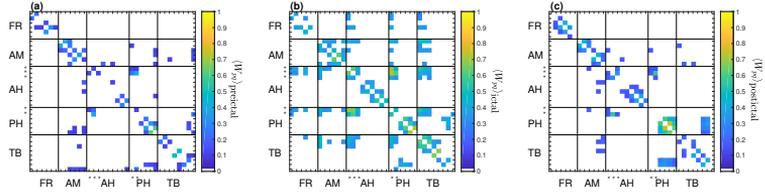


Figure A.4: Same as Fig. 4.5 but for seizure 5.

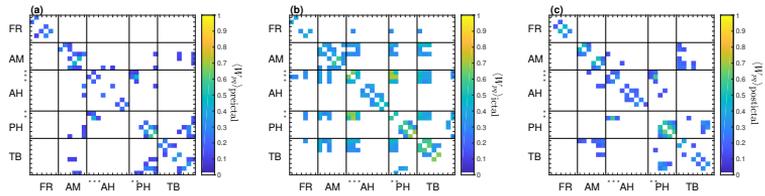


Figure A.5: Same as Fig. 4.5 but for seizure 6.

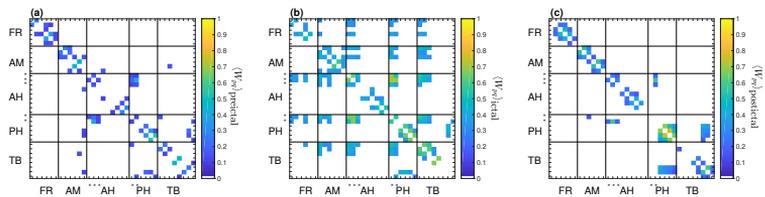


Figure A.6: Same as Fig. 4.5 but for seizure 7.

References

- Acebrón, J. A., Bonilla, L. L., Pérez Vicente, C. J., Ritort, F., & Spigler, R. (2005). The kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of Modern Physics*, 77(1), 137–185.
- Adamic, L. A., Lukose, R. M., Puniyani, A. R., & Huberman, B. A. (2001). Search in power-law networks. *Physical Review E*, 64(4), 046135.
- Alexanderson, G. L. (2006). Euler and königsberg’s bridges: A historical view. *Bulletin of the American Mathematical Society*, 43, 567.
- Amaral, L. A. N., Scala, A., Barthélémy, M., & Stanley, H. E. (2000). Classes of small-world networks. *Proceedings of the National Academy of Sciences*, 97(21), 11149–11152.
- Andrzejak, R. G., Chicharro, D., Lehnertz, K., & Mormann, F. (2011). Using bivariate signal analysis to characterize the epileptic focus: The benefit of surrogates. *Physical Review E*, 83(4), 046203.
- Andrzejak, R. G. & Kreuz, T. (2011). Characterizing unidirectional couplings between point processes and flows. *Europhysics Letters*, 96(5), 50012.
- Andrzejak, R. G., Schindler, K., & Rummel, C. (2012). Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients. *Physical Review E*, 86(4), 046206.
- Arenas, A., Díaz-Guilera, A., Kurths, J., Moreno, Y., & Zhou, C. (2008). Synchronization in complex networks. *Physics reports*, 469(3), 93–153.

- Arnhold, J., Grassberger, P., Lehnertz, K., & Elger, C. E. (1999). A robust method for detecting interdependences: Application to intracranially recorded EEG. *Physica D: Nonlinear Phenomena*, *134*(4), 419–430.
- Atay, F. M., Jost, J., & Wende, A. (2004). Delays, connection topology, and synchronization of coupled chaotic maps. *Physical Review Letters*, *92*(14), 144101.
- Banavar, J. R., Maritan, A., & Rinaldo, A. (1999). Size and form in efficient transportation networks. *Nature*, *399*(6732), 130.
- Barabási, A.-L. & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(6804), 509–512.
- Barabási, A.-L. & Pósfai, M. (2016). *Network science*. Cambridge university press.
- Bascompte, J., Jordano, P., Melián, C. J., & Olesen, J. M. (2003). The nested assembly of plant–animal mutualistic networks. *Proceedings of the National Academy of Sciences*, *100*(16), 9383.
- Bassett, D. S. & Bullmore, E. (2006). Small-world brain networks. *The Neuroscientist*, *12*(6), 512–523.
- Bianco-Martinez, E., Rubido, N., Antonopoulos, C. G., & Baptista, M. S. (2016). Successful network inference from time-series data using mutual information rate. *Chaos*, *26*(4), 043102.
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, *424*, 175–308.
- Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32.
- Bressloff, P. C. (2011). Spatiotemporal dynamics of continuum neural fields. *Journal of Physics A: Mathematical and Theoretical*, *45*(3), 033001.
- Bullmore, E. & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*(3), 186.
- Castellano, C., Fortunato, S., & Loreto, V. (2009). Statistical physics of social dynamics. *Reviews of Modern Physics*, *81*, 591–646.
- Cecchini, G., Thiel, M., Schelter, B., & Sommerlade, L. (2018). Improving network inference: The impact of false positive and false negative conclusions about the presence or absence of links. *Journal of Neuroscience Methods*, *307*, 31–36.

- Cestnik, R. & Rosenblum, M. (2017). Reconstructing networks of pulse-coupled oscillators from spike trains. *Physical Review E*, 96(1), 012209.
- Chicharro, D. & Andrzejak, R. G. (2009). Reliable detection of directional couplings using rank statistics. *Physical Review E*, 80(2), 026217.
- Chicharro, D., Andrzejak, R. G., & Ledberg, A. (2011). Inferring and quantifying causality in neuronal networks. *BMC Neuroscience*, 12(Suppl 1), P192.
- Ching, E. S., Lai, P. Y., & Leung, C. Y. (2015). Reconstructing weighted networks from dynamics. *Physical Review E*, 91(3), 030801(R).
- Clifford, P. & Sudbury, A. (1973). A model for spatial conflict. *Biometrika*, 60(3), 581–588.
- Deza, J. I., Barreiro, M., & Masoller, C. (2015). Assessing the direction of climate interactions by means of complex networks and information theoretic tools. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(3), 033105.
- Dickten, H., Porz, S., Elger, C. E., & Lehnertz, K. (2016). Weighted and directed interactions in evolving large-scale epileptic brain networks. *Scientific Reports*, 6(1), 34824.
- Duan, Z. & Chen, G. (2009). Global robust stability and synchronization of networks with lorenz-type nodes. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 56(8), 679–683.
- Epstein, C. M., Adhikari, B. M., Gross, R., Willie, J., & Dhamala, M. (2014). Application of high-frequency Granger causality to analysis of epileptic seizures and surgical decision making. *Epilepsia*, 55(12), 2038–2047.
- Erdős, P. & Rényi, A. (1959). On random graphs. *I. Publicationes Mathematicae*, 6, 290–297.
- Ermentrout, G. B. & Terman, D. H. (2010). *Mathematical foundations of neuroscience*. Springer Science & Business Media.
- Faloutsos, M., Faloutsos, P., & Faloutsos, C. (1999). On power-law relationships of the internet topology. *SIGCOMM Computer Communication Review*, 29(4), 251–262.
- Flamm, C., Graef, A., Pirker, S., Baumgartner, C., & Deistler, M. (2013). Influence analysis for high-dimensional time series with an application to epileptic seizure onset zone detection. *Journal of Neuroscience Methods*, 214(1), 80–90.

- Friston, K., Moran, R., & Seth, A. K. (2013). Analysing connectivity with granger causality and dynamic causal modelling. *Current opinion in neurobiology*, 23(2), 172–178.
- Gastner, M. T. & Newman, M. E. J. (2004). Diffusion-based method for producing density-equalizing maps. *Proceedings of the National Academy of Sciences*, 101(20), 7499–7504.
- Gómez-Gardeñes, J., Gómez, S., Arenas, A., & Moreno, Y. (2011). Explosive synchronization transitions in scale-free networks. *Physical Review Letters*, 106(12), 128701.
- Grassberger, P. & Procaccia, I. (1983). Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, 9(1-2), 189–208.
- Guimerà, R. & Sales-Pardo, M. (2009). Missing and spurious interactions and the reconstruction of complex networks. *Proceedings of the National Academy of Sciences*, 106(52), 22073–22078.
- Han, X., Shen, Z., Wang, W.-X., & Di, Z. (2015). Robust Reconstruction of Complex Networks from Sparse Data. *Physical Review Letters*, 114(6684), 028701.
- Hansel, D. & Sompolinsky, H. (1993). Solvable Model of Spatiotemporal Chaos. *Physical Review Letters*, 71(17), 2710–2714.
- Hassan, M., Dufor, O., Merlet, I., Berrou, C., & Wendling, F. (2014). Eeg source connectivity analysis: From dense array recordings to brain networks. *PloS one*, 9(8), e105041.
- Hoppensteadt, F. C. & Izhikevich, E. M. (2012). *Weakly connected neural networks*. Springer Science & Business Media.
- Huberman, B. A. (2001). *The laws of the web*. MIT Press.
- Ingber, L. et al. (1996). Adaptive simulated annealing (ASA): Lessons learned. *Control and cybernetics*, 25, 33–54.
- Ipsen, M. & Mikhailov, A. S. (2002). Evolutionary reconstruction of networks. *Physical Review E*, 66, 46109.
- Isakov, S. V., Zintchenko, I. N., Rønnow, T. F., & Troyer, M. (2015). Optimized simulated annealing for Ising spin glasses. *Computer Physics Communications*, 192, 265–271.
- Jachan, M., Henschel, K., Nawrath, J., Schad, A., Timmer, J., & Schelter, B. (2009). Inferring direct directed-information flow from multivariate nonlinear time series. *Physical Review E*, 80(1), 011138.
- Jackson, D., Matthew O. and López-Pintado. (2013). Diffusion and contagion in networks with heterogeneous agents and homophily. *Network Science*, 1(1), 49–67.

- Jeong, H., Mason, S. P., Barabási, A.-L., & Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *Nature*, *411*(6833), 41–42.
- Kansky, K. J. (1963). Structure of transportation networks : Relationships between network geometry and regional characteristics. Thesis. Chicago : Department of Geography, The University of Chicago.
- Killworth, P. D., Johnsen, E. C., Bernard, H., Ann Shelley, G., & McCarty, C. (1990). Estimating the size of personal networks. *Social Networks*, *12*(4), 289–312.
- Kononenko, I., Šimec, E., & Robnik-Šikonja, M. (1997). Overcoming the myopia of inductive learning algorithms with relieff. *Applied Intelligence*, *7*(1), 39–55.
- Korfatis, N. T., Poulos, M., & Bokos, G. (2006). Evaluating authoritative sources using social networks: An insight from wikipedia. *Online Information Review*, *30*(3), 252–262.
- Korzeniewska, A., Cervenka, M., Jouny, C., Perilla, J., Harezlak, J., Bergey, G., ... Crone, N. (2014). Ictal propagation of high frequency activity is recapitulated in interictal recordings: Effective connectivity of epileptogenic networks recorded with intracranial EEG. *NeuroImage*, *101*, 96–113.
- Koutlis, C. & Kugiumtzis, D. (2016). Discrimination of coupling structures using causality networks from multivariate time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *26*(9), 093120.
- Kugiumtzis, D. (2013a). Direct-coupling information measure from nonuniform embedding. *87*(6), 062918.
- Kugiumtzis, D. (2013b). Partial transfer entropy on rank vectors. *The European Physical Journal Special Topics*, *222*(2), 401–420.
- Kuramoto, Y. (1975). Self-entrainment of a population of coupled nonlinear oscillators. In *International symposium on mathematical problems in theoretical physics* (pp. 420–422). Springer.
- Laiou, P. & Andrzejak, R. G. (2017). Coupling strength versus coupling impact in nonidentical bidirectionally coupled dynamics. *Physical Review E*, *95*(1), 012210.
- Leguia, M. G., Andrzejak, R. G., & Levnajić, Z. (2017). Evolutionary optimization of network reconstruction from derivative-variable correlations. *Journal of Physics A: Mathematical and Theoretical*, *50*(33), 334001.
- Leguia, M. G., Martínez, C. G. B., Malvestio, I., Campo, A. T., Rocamora, R., Levnajić, Z., & Andrzejak, R. G. (2019). Inferring directed net-

- works using a rank-based connectivity measure. *Physical Review E*, 99(1), 012319.
- Lehnertz, K. & Dickten, H. (2015). Assessing directionality and strength of coupling through symbolic analysis: an application to epilepsy patients. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 373(2034), 20140094.
- Lerman, K., Yan, X., & Wu, X.-Z. (2016). The "majority illusion" in social networks. *PLOS ONE*, 11(2), 1–13.
- Levnajić, Z. & Pikovsky, A. (2011). Network reconstruction from random phase resetting. *Physical Review Letters*, 107(3), 34101.
- Levnajić, Z. & Pikovsky, A. (2014). Untangling complex dynamical systems via derivative-variable correlations. *Scientific Reports*, 4, 5030.
- Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., & Christakis, N. (2008). Tastes, ties, and time: A new social network dataset using facebook.com. *Social Networks*, 30(4), 330–342.
- Li, C. & Chen, G. (2004). Phase synchronization in small-world networks of chaotic oscillators. *Physica A: Statistical Mechanics and its Applications*, 341, 73–79.
- Li, Z., Li, P., Krishnan, A., & Liu, J. (2011). Large-scale dynamic gene regulatory network inference combining differential equation models with local dynamic Bayesian network analysis. *Bioinformatics*, 27(19), 2686–2691.
- Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E., & Aaberg, Y. (2001). The web of human sexual contacts. *Nature*, 411(6840), 907.
- Liu, Y.-Y., Slotine, J.-J., & Barabási, A.-L. (2013). Observability of complex systems. *Proceedings of the National Academy of Sciences*, 110(7), 2460–2465.
- Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2), 130–141.
- Malvestio, I., Kreuz, T., & Andrzejak, R. G. (2017). Robustness and versatility of a nonlinear interdependence method for directional coupling detection from spike trains. *Physical Review E*, 96(2), 022203.
- Margolin, A. A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Favera, R. D., & Califano, A. (2006). Aracne: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics*, 7(1), S7.

- Mei, G., Wu, X., Chen, G., & Lu, J.-A. (2016). Identifying structures of continuously-varying weighted networks. *Scientific Reports*, *6*, 26649.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. *Journal Chemical Physics*, *21*(6), 1087–1092.
- Mittler, R., Vanderauwera, S., Gollery, M., & Breusegem, F. V. (2004). Reactive oxygen gene network of plants. *Trends in Plant Science*, *9*(10), 490–498.
- Newman, M. E. J. (2007). *Networks, An Introduction*. Oxford UP.
- Oates, C. J., Dondelinger, F., Bayani, N., Korkola, J., Gray, J. W., & Mukherjee, S. (2014). Causal network inference using biochemical kinetics. *Bioinformatics*, *30*(17), i468–i474.
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P., & Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of Modern Physics*, *87*(3), 925–979.
- Pereda, E., Quiroga, R. Q., & Bhattacharya, J. (2005). Nonlinear multivariate analysis of neurophysiological signals.
- Pikovsky, A. (2016). Reconstruction of a neural network from a time series of firing rates. *Physical Review E*, *93*(6), 062313.
- Pikovsky, A., Rosenblum, M., & Kurths, J. (2001). *Synchronization: A universal concept in nonlinear science* (C. U. Press, Ed.).
- Porter, M. A. & Gleeson, J. P. (2016). *Dynamical systems on networks*. Frontiers in Applied Dynamical Systems: Reviews and Tutorials.
- Prigamo, L. & Díaz-Guilera, A. (2012). Extracting topological features from dynamical measures in networks of Kuramoto oscillators. *Physical Review E*, *85*(3), 36112.
- Quiroga, R. Q., Arnhold, J., & Grassberger, P. (2000). Learning driver-response relationships from synchronization patterns. *Physical Review E*, *61*(5), 5142.
- Quiroga, R. Q., Kraskov, A., Kreuz, T., & Grassberger, P. (2002). performance of different synchronization measures in real data: A case study on electroencephalographic signals. *Physical Review E*, *65*(4), 041903.
- Rodrigues, F. A., Peron, T. K. D., Ji, P., & Kurths, J. (2016). The kuramoto model in complex networks. *Physics Reports*, *610*, 1–98.
- Rssler, O. (1976). An equation for continuous chaos. *Physics Letters A*, *57*(5), 397–398.

- Rubido, N., Martí, A. C., Bianco-Martínez, E., Grebogi, C., Baptista, M. S., & Masoller, C. (2014). Exact detection of direct links in networks of interacting dynamical units. *New Journal of Physics*, *16*(9), 093010.
- Rubinov, M. & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage*, *52*(3), 1059–1069.
- Rummel, C., Abela, E., Müller, M., Hauf, M., Scheidegger, O., Wiest, R., & Schindler, K. (2011). Uniform approach to linear and nonlinear interrelation patterns in multivariate time series. *Physical Review E*, *83*(6), 066215.
- Russell Bernard, H., Johnsen, E. C., Killworth, P. D., & Robinson, S. (1991). Estimating the size of an average personal network and of an event subpopulation: Some empirical results. *Social Science Research*, *20*(2), 109–121.
- Schelter, B., Timmer, J., & Eichler, M. (2009). Assessing the strength of directed influences among neural signals using renormalized partial directed coherence. *Journal of Neuroscience Methods*, *179*(1), 121–130.
- Schiff, S. J., So, P., Chang, T., Burke, R. E., & Sauer, T. (1996). Detecting dynamical interdependence and generalized synchrony through mutual prediction in a neural ensemble. *Physical Review E*, *54*(6), 6708.
- Schreiber, T. & Schmitz, A. (2000). Surrogate time series. *Physica D: Nonlinear Phenomena*, *142*(3-4), 346–382.
- Serra, J., Serra, X., & Andrzejak, R. G. (2009). Cross recurrence quantification for cover song identification. *New Journal of Physics*, *11*(9), 093017.
- Shandilya, S. G. & Timme, M. (2011). Inferring network topology from complex dynamics. *New Journal of Physics*, *13*(1), 013004.
- Simidjievski, N., Tanevski, J., Zenko, B., Levnajić, Z., Todorovski, L., & Dzeroski, S. (2018). Decoupling approximation robustly reconstructs directed dynamical networks. *New Journal of Physics*, *20*(11), 113003.
- Smirnov, D. A. & Andrzejak, R. G. (2005). Detection of weak directional coupling: Phase-dynamics approach versus state-space approach. *Physical Review E*, *71*(3), 036207.
- Sommerlade, L., Eichler, M., Jachan, M., Henschel, K., Timmer, J., & Schelter, B. (2009). Estimating causal dependencies in networks of

- nonlinear stochastic dynamical systems. *Physical Review E*, 80(5), 051128.
- Strogatz, S. H. (2001). Exploring complex networks. *Nature*, 410(6825), 268.
- Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, 1227079.
- Talukdar, M. S., Torsaeter, O., Ioannidis, M. A., & Howard, J. J. (2002). Stochastic reconstruction, 3D characterization and network modeling of chalk. *Journal of Petroleum Science and Engineering*, 35(1–2), 1–21.
- Tauste Campo, A., Principe, A., Ley, M., Rocamora, R., & Deco, G. (2018). Degenerate time-dependent network dynamics anticipate seizures in human epileptic brain. *PLoS Biology*, 16(4), e2002580.
- Theiler, J. (1990). Estimating fractal dimension. *Journal of the Optical Society of America A*, 7(6), 1055.
- Timme, M. (2007). Revealing network connectivity from response dynamics. *Physical review letters*, 98(22), 224101.
- Timme, M. & Casadiego, J. (2014). Revealing networks from dynamics: An introduction. *Journal of Physics A: Mathematical and Theoretical*, 47(34), 343001.
- Tirabassi, G., Sevilla-Escoboza, R., Buldú, J. M., & Masoller, C. (2015). Inferring the connectivity of coupled oscillators from time-series statistical similarity analysis. *Scientific Reports*, 5, 10829.
- Tirabassi, G., Sommerlade, L., & Masoller, C. (2017). Inferring directed climatic interactions with renormalized partial directed coherence and directed partial correlation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(3), 035815.
- Tomovski, I. & Kocarev, L. (2015). Network topology inference from infection statistics. *Physica A: Statistical Mechanics and its Applications*, 436, 272–285.
- Trejo Banos, D., Millar, A. J., & Sanguinetti, G. (2015). A Bayesian approach for structure learning in oscillating regulatory networks. *Bioinformatics*, 31(22), 3617.
- Van Mierlo, P., Carrette, E., Hallez, H., Raedt, R., Meurs, A., Vandenberghe, S., . . . Vonck, K. (2013). Ictal-onset localization through connectivity analysis of intracranial EEG signals in patients with refractory epilepsy. *Epilepsia*, 54(8), 1409–1418.

- Vanhems, P., Barrat, A., Cattuto, C., Pinton, J.-F., Khanafer, N., Regis, C., ... Voirin, N. (2013). Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. *PLoS ONE*, 8(9), e73970.
- Vila-Vidal, M., Principe, A., Ley, M., Deco, G., Campo, A. T., & Rocamora, R. (2017). Detection of recurrent activation patterns across focal seizures: Application to seizure onset zone identification. *Clinical Neurophysiology*, 128(6), 977–985.
- Wang, W.-X., Lai, Y.-C., & Grebogi, C. (2016). Data based identification and prediction of nonlinear and complex dynamical systems. *Physics Reports*, 644, 1–76.
- Wang, W.-X., Yang, R., Lai, Y.-C., Kovanis, V., & Harrison, M. A. F. (2011). Time-series-based prediction of complex oscillator networks via compressive sensing. *Europhysics Letters*, 94(4), 48006.
- Wasserman, S. & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.
- Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 409–410.
- Weng, L., Menczer, F., & Ahn, Y.-Y. (2013). Virality prediction and community structure in social networks. *Scientific reports*, 3, 2522.
- Whalen, A. J., Brennan, S. N., Sauer, T. D., & Schiff, S. J. (2015). Observability and controllability of nonlinear networks: The role of symmetry. *Physical Review X*, 5(1), 011005.
- White, J. G., Southgate, E., Thomson, J. N., & Brenner, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 314(1165), 1–340.
- Widder, S., Schicho, J., & Schuster, P. (2007). Dynamic patterns of gene regulation i: Simple two-gene systems. *Journal of theoretical biology*, 246(3), 395–419.
- Winterhalder, M., Schelter, B., & Timmer, J. (2007). Detecting Coupling Directions in Multivariate Oscillatory Systems. *International Journal of Bifurcation and Chaos*, 17(10), 3735–3739.
- Yan, G., Tsekenis, G., Barzel, B., Slotine, J.-J., Liu, Y.-Y., & Barabási, A.-L. (2015). Spectrum of controlling and observing complex networks. *Nature Physics*, 11(9), 779.
- Zhou, C. & Kurths, J. (2006). Dynamical weights and enhanced synchronization in adaptive complex networks. *Physical review letters*, 96(16), 164102.

SUBJECT INDEX

- Biological networks, 3
- Complex Systems, 1
- Confusion matrix, 34
- Data-driven, 10, 32
- Derivative-variable correlation,
14
- Directed networks, 2, 33
- EEG, 10, 51
- Epidemics, 6
- Epilepsy, 51
- Functional connectivity, 52
- Hansel-Sompolinsky model, 19
- Information networks, 4
- Kuramoto model, 7
- Lorenz dynamics, 7, 33
- Model-based, 9, 13
- Network inference, 8, 19, 34
- Networks, 1, 4, 8
- Neuronal dynamics, 19
- Noise, 28, 44
- Random networks, 5
- ROC curve, 35, 38
- Scale-free, 6
- Simulated annealing, 17
- Small-world, 6
- Social networks, 3, 6
- Sparsity, 4, 47
- Synchronization, 7, 40
- Transport networks, 3
- Weighted networks, 2, 19

